Package ‘BSDA’
March 23, 2012

Type Package

Title Basic Statistics and Data Analysis

Version 1.01

Date 2012-03-23

LazyLoad yes

LazyData yes

Author Alan T. Arnholt

Description Data sets for book ‘Basic Statistics and Data Analysis’ by Larry J. Kitchens

Maintainer Alan T. Arnholt <arnholt@math.appstate.edu>

Depends e1071, lattice

License GPL (>= 2)

R topics documented:

BSDA-package ........................................ 8
Abbey ............................................... 8
ABC ............................................... 9
Abilene ........................................... 10
Ability ............................................ 10
Abortion .......................................... 11
Absent ............................................ 12
Achieve .......................................... 13
Adsales ........................................... 13
Aggress .......................................... 14
Aid ............................................... 15
Aids ............................................. 15
Airdisasters .................................... 16
Airline .......................................... 17
Alcohol ......................................... 18
Allergy ......................................... 18
Anesthet ....................................... 19
Anxiety ......................................... 20
Apolipop ....................................... 20
R topics documented:

Append ................................. 21
Appendec ................................ 22
Aptitude ................................ 22
Archeo .................................. 23
Arthriti .................................. 24
Artifici .................................. 25
Asprin .................................... 25
Asthmati .................................. 26
Attorney ................................... 27
Autogear .................................. 28
Backtoback ............................... 28
Bbsalaries ................................. 29
Bigten ...................................... 30
Biology ..................................... 30
Birth ....................................... 31
Blackedu ................................... 32
Blood ....................................... 33
Board ....................................... 33
Bones ....................................... 34
Books ....................................... 35
Bookstor ................................... 35
Brain ....................................... 36
Bumpers ................................... 37
Bus .......................................... 37
Bypass ...................................... 38
Cabinets ................................... 39
Cancer ..................................... 40
Carbon ...................................... 40
Cat .......................................... 41
Censored ................................... 42
Challeng ................................... 42
Chemist ..................................... 43
Chesapea ................................... 44
Chevy ....................................... 45
Chicken ..................................... 45
Chipavg ..................................... 46
Chips ....................................... 47
Cigar ....................................... 48
Cigaret ..................................... 48
Clsim ....................................... 49
Citrus ....................................... 50
Clean ....................................... 51
Coaxial ..................................... 52
Coffee ....................................... 52
Coins ....................................... 53
Combinations ............................. 54
Commute ................................... 54
Concept ..................................... 55
Concrete ..................................... 56
Corn ......................................... 56
Correlat ..................................... 57
Counsel ..................................... 58
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cpi</td>
<td>58</td>
</tr>
<tr>
<td>Crime</td>
<td>59</td>
</tr>
<tr>
<td>Darwin</td>
<td>60</td>
</tr>
<tr>
<td>Dealers</td>
<td>60</td>
</tr>
<tr>
<td>Defectiv</td>
<td>61</td>
</tr>
<tr>
<td>Degree</td>
<td>62</td>
</tr>
<tr>
<td>Delay</td>
<td>62</td>
</tr>
<tr>
<td>Depend</td>
<td>63</td>
</tr>
<tr>
<td>Detroit</td>
<td>64</td>
</tr>
<tr>
<td>Develop</td>
<td>64</td>
</tr>
<tr>
<td>Devmath</td>
<td>65</td>
</tr>
<tr>
<td>Dice</td>
<td>65</td>
</tr>
<tr>
<td>Diesel</td>
<td>66</td>
</tr>
<tr>
<td>Diplomat</td>
<td>67</td>
</tr>
<tr>
<td>Disposal</td>
<td>68</td>
</tr>
<tr>
<td>Dogs</td>
<td>68</td>
</tr>
<tr>
<td>Domestic</td>
<td>69</td>
</tr>
<tr>
<td>Dopamine</td>
<td>70</td>
</tr>
<tr>
<td>Dowjones</td>
<td>70</td>
</tr>
<tr>
<td>Drink</td>
<td>71</td>
</tr>
<tr>
<td>Drug</td>
<td>72</td>
</tr>
<tr>
<td>Dyslexia</td>
<td>72</td>
</tr>
<tr>
<td>Earthqk</td>
<td>73</td>
</tr>
<tr>
<td>EDA</td>
<td>74</td>
</tr>
<tr>
<td>Educat</td>
<td>75</td>
</tr>
<tr>
<td>Eggs</td>
<td>75</td>
</tr>
<tr>
<td>Elderly</td>
<td>76</td>
</tr>
<tr>
<td>Energy</td>
<td>77</td>
</tr>
<tr>
<td>Engineer</td>
<td>78</td>
</tr>
<tr>
<td>Entrance</td>
<td>78</td>
</tr>
<tr>
<td>Epaminicompact</td>
<td>79</td>
</tr>
<tr>
<td>Epatwoseater</td>
<td>80</td>
</tr>
<tr>
<td>Executiv</td>
<td>81</td>
</tr>
<tr>
<td>Exercise</td>
<td>81</td>
</tr>
<tr>
<td>Fabric</td>
<td>82</td>
</tr>
<tr>
<td>Faithful</td>
<td>83</td>
</tr>
<tr>
<td>Family</td>
<td>83</td>
</tr>
<tr>
<td>Ferraro1</td>
<td>84</td>
</tr>
<tr>
<td>Ferraro2</td>
<td>85</td>
</tr>
<tr>
<td>Fertility</td>
<td>85</td>
</tr>
<tr>
<td>Firstchi</td>
<td>86</td>
</tr>
<tr>
<td>Fish</td>
<td>87</td>
</tr>
<tr>
<td>Fitness</td>
<td>87</td>
</tr>
<tr>
<td>Florida2000</td>
<td>88</td>
</tr>
<tr>
<td>Fluid</td>
<td>89</td>
</tr>
<tr>
<td>Food</td>
<td>90</td>
</tr>
<tr>
<td>Framingh</td>
<td>90</td>
</tr>
<tr>
<td>Freshman</td>
<td>91</td>
</tr>
<tr>
<td>Funeral</td>
<td>92</td>
</tr>
<tr>
<td>Galaxie</td>
<td>92</td>
</tr>
<tr>
<td>Gallup</td>
<td>93</td>
</tr>
<tr>
<td>Gasoline</td>
<td>94</td>
</tr>
<tr>
<td>Topics Documented</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------</td>
<td>------</td>
</tr>
<tr>
<td>German</td>
<td>95</td>
</tr>
<tr>
<td>Golf</td>
<td>95</td>
</tr>
<tr>
<td>Governor</td>
<td>96</td>
</tr>
<tr>
<td>Gpa</td>
<td>97</td>
</tr>
<tr>
<td>Grades</td>
<td>97</td>
</tr>
<tr>
<td>Graduate</td>
<td>98</td>
</tr>
<tr>
<td>Greenriv</td>
<td>99</td>
</tr>
<tr>
<td>Gmriv2</td>
<td>99</td>
</tr>
<tr>
<td>Groupabc</td>
<td>100</td>
</tr>
<tr>
<td>Groups</td>
<td>101</td>
</tr>
<tr>
<td>Gym</td>
<td>101</td>
</tr>
<tr>
<td>Habits</td>
<td>102</td>
</tr>
<tr>
<td>Haptologo</td>
<td>103</td>
</tr>
<tr>
<td>Hardware</td>
<td>103</td>
</tr>
<tr>
<td>Hardwood</td>
<td>104</td>
</tr>
<tr>
<td>Heat</td>
<td>104</td>
</tr>
<tr>
<td>Heating</td>
<td>105</td>
</tr>
<tr>
<td>Hodgkin</td>
<td>106</td>
</tr>
<tr>
<td>Homes</td>
<td>107</td>
</tr>
<tr>
<td>Homework</td>
<td>108</td>
</tr>
<tr>
<td>Honda</td>
<td>108</td>
</tr>
<tr>
<td>Hostile</td>
<td>109</td>
</tr>
<tr>
<td>Housing</td>
<td>110</td>
</tr>
<tr>
<td>Hurrican</td>
<td>111</td>
</tr>
<tr>
<td>Iceberg</td>
<td>112</td>
</tr>
<tr>
<td>Income</td>
<td>112</td>
</tr>
<tr>
<td>Independent</td>
<td>113</td>
</tr>
<tr>
<td>Indian</td>
<td>114</td>
</tr>
<tr>
<td>Indiapol</td>
<td>115</td>
</tr>
<tr>
<td>Indy500</td>
<td>115</td>
</tr>
<tr>
<td>Inflatio</td>
<td>116</td>
</tr>
<tr>
<td>Inletoil</td>
<td>117</td>
</tr>
<tr>
<td>Inmate</td>
<td>117</td>
</tr>
<tr>
<td>Inspect</td>
<td>118</td>
</tr>
<tr>
<td>Insulate</td>
<td>119</td>
</tr>
<tr>
<td>Iqgpa</td>
<td>119</td>
</tr>
<tr>
<td>Irises</td>
<td>120</td>
</tr>
<tr>
<td>Jdpower</td>
<td>121</td>
</tr>
<tr>
<td>Jobsat</td>
<td>121</td>
</tr>
<tr>
<td>Kidsmoke</td>
<td>122</td>
</tr>
<tr>
<td>Kilowatt</td>
<td>123</td>
</tr>
<tr>
<td>Kinder</td>
<td>123</td>
</tr>
<tr>
<td>Laminect</td>
<td>124</td>
</tr>
<tr>
<td>Lead</td>
<td>125</td>
</tr>
<tr>
<td>Leader</td>
<td>125</td>
</tr>
<tr>
<td>Lethal</td>
<td>126</td>
</tr>
<tr>
<td>Life</td>
<td>127</td>
</tr>
<tr>
<td>Lifespan</td>
<td>127</td>
</tr>
<tr>
<td>Lightmonth</td>
<td>128</td>
</tr>
<tr>
<td>Lodge</td>
<td>129</td>
</tr>
<tr>
<td>Longtail</td>
<td>130</td>
</tr>
<tr>
<td>Lowabil</td>
<td>130</td>
</tr>
</tbody>
</table>
R topics documented:

Magnesium ................................................. 131
Malpractice .................................................. 132
Manager ......................................................... 132
Marked .......................................................... 133
MATH ............................................................. 134
Mathcomp ....................................................... 134
Mathpro ........................................................ 135
Maze ............................................................. 136
Median .......................................................... 136
Mental ........................................................... 137
Mercury .......................................................... 138
Metreng .......................................................... 138
Miller ............................................................ 139
Miller1 ........................................................... 139
Moisture ......................................................... 140
Monoxide ......................................................... 141
Movie ........................................................... 141
Music ............................................................ 142
Name ............................................................. 143
Nascar ............................................................ 143
Nervous .......................................................... 144
Newsstand ....................................................... 145
Nfldraf2 .......................................................... 145
Nfldraft .......................................................... 146
Nicotine .......................................................... 147
normarea ........................................................ 148
nsize .............................................................. 148
ntester ............................................................ 149
Orange ........................................................... 150
Orioles ............................................................ 151
Oxytocin .......................................................... 151
Parented .......................................................... 152
Patrol ............................................................. 153
Pearson ........................................................... 153
Phone ............................................................. 154
Poison ............................................................ 155
Politic ............................................................. 155
Pollution .......................................................... 156
Porosity ........................................................... 156
Poverty ........................................................... 157
Precinct ........................................................... 158
Prejudice .......................................................... 159
President ........................................................ 159
Press .............................................................. 160
Prognost .......................................................... 161
Program .......................................................... 161
Psat ............................................................... 162
Psych ............................................................. 163
Puerto ............................................................. 163
Quail ............................................................. 164
Quality ............................................................ 164
Rainks ............................................................ 165
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randd</td>
<td>166</td>
</tr>
<tr>
<td>Rat</td>
<td>167</td>
</tr>
<tr>
<td>Ratings</td>
<td>167</td>
</tr>
<tr>
<td>Reaction</td>
<td>168</td>
</tr>
<tr>
<td>Reading</td>
<td>169</td>
</tr>
<tr>
<td>Readiq</td>
<td>169</td>
</tr>
<tr>
<td>Referend</td>
<td>170</td>
</tr>
<tr>
<td>Region</td>
<td>171</td>
</tr>
<tr>
<td>Register</td>
<td>171</td>
</tr>
<tr>
<td>Rehab</td>
<td>172</td>
</tr>
<tr>
<td>Remedial</td>
<td>173</td>
</tr>
<tr>
<td>Rentals</td>
<td>174</td>
</tr>
<tr>
<td>Repair</td>
<td>174</td>
</tr>
<tr>
<td>Retail</td>
<td>175</td>
</tr>
<tr>
<td>Ronbrown1</td>
<td>175</td>
</tr>
<tr>
<td>Ronbrown2</td>
<td>176</td>
</tr>
<tr>
<td>Rural</td>
<td>177</td>
</tr>
<tr>
<td>Salary</td>
<td>178</td>
</tr>
<tr>
<td>Salinity</td>
<td>178</td>
</tr>
<tr>
<td>Sat</td>
<td>179</td>
</tr>
<tr>
<td>Saving</td>
<td>180</td>
</tr>
<tr>
<td>Scales</td>
<td>180</td>
</tr>
<tr>
<td>Schizop2</td>
<td>181</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>182</td>
</tr>
<tr>
<td>Seatbelt</td>
<td>182</td>
</tr>
<tr>
<td>Selfdefe</td>
<td>183</td>
</tr>
<tr>
<td>Senior</td>
<td>184</td>
</tr>
<tr>
<td>Sentence</td>
<td>184</td>
</tr>
<tr>
<td>Shkdrug</td>
<td>185</td>
</tr>
<tr>
<td>Shock</td>
<td>186</td>
</tr>
<tr>
<td>Shoplift</td>
<td>186</td>
</tr>
<tr>
<td>Short</td>
<td>187</td>
</tr>
<tr>
<td>Shuttle</td>
<td>188</td>
</tr>
<tr>
<td>SIGN.test</td>
<td>188</td>
</tr>
<tr>
<td>Simpson</td>
<td>191</td>
</tr>
<tr>
<td>Situp</td>
<td>192</td>
</tr>
<tr>
<td>Skewed</td>
<td>192</td>
</tr>
<tr>
<td>Skin</td>
<td>193</td>
</tr>
<tr>
<td>Slc</td>
<td>194</td>
</tr>
<tr>
<td>Smokyph</td>
<td>194</td>
</tr>
<tr>
<td>Snore</td>
<td>195</td>
</tr>
<tr>
<td>Snow</td>
<td>196</td>
</tr>
<tr>
<td>Soccer</td>
<td>196</td>
</tr>
<tr>
<td>Social</td>
<td>197</td>
</tr>
<tr>
<td>Sophomor</td>
<td>198</td>
</tr>
<tr>
<td>South</td>
<td>198</td>
</tr>
<tr>
<td>Speed</td>
<td>199</td>
</tr>
<tr>
<td>Spellers</td>
<td>200</td>
</tr>
<tr>
<td>Spelling</td>
<td>200</td>
</tr>
<tr>
<td>Sports</td>
<td>201</td>
</tr>
<tr>
<td>Spouse</td>
<td>202</td>
</tr>
<tr>
<td>SRS</td>
<td>202</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>Stable</td>
<td>203</td>
</tr>
<tr>
<td>Stamp</td>
<td>204</td>
</tr>
<tr>
<td>Statclas</td>
<td>204</td>
</tr>
<tr>
<td>Statelaw</td>
<td>205</td>
</tr>
<tr>
<td>Statisti</td>
<td>206</td>
</tr>
<tr>
<td>Step</td>
<td>206</td>
</tr>
<tr>
<td>Stress</td>
<td>207</td>
</tr>
<tr>
<td>Study</td>
<td>208</td>
</tr>
<tr>
<td>Submarin</td>
<td>208</td>
</tr>
<tr>
<td>Subway</td>
<td>209</td>
</tr>
<tr>
<td>Sunspot</td>
<td>209</td>
</tr>
<tr>
<td>Superbowl</td>
<td>210</td>
</tr>
<tr>
<td>Supercar</td>
<td>211</td>
</tr>
<tr>
<td>Tablrock</td>
<td>212</td>
</tr>
<tr>
<td>Teacher</td>
<td>213</td>
</tr>
<tr>
<td>Tenness</td>
<td>213</td>
</tr>
<tr>
<td>Tensile</td>
<td>214</td>
</tr>
<tr>
<td>Test1</td>
<td>215</td>
</tr>
<tr>
<td>Thermal</td>
<td>215</td>
</tr>
<tr>
<td>Tiaa</td>
<td>216</td>
</tr>
<tr>
<td>Ticket</td>
<td>217</td>
</tr>
<tr>
<td>Toaster</td>
<td>218</td>
</tr>
<tr>
<td>Tonsils</td>
<td>218</td>
</tr>
<tr>
<td>Tort</td>
<td>219</td>
</tr>
<tr>
<td>Toxic</td>
<td>220</td>
</tr>
<tr>
<td>Track</td>
<td>221</td>
</tr>
<tr>
<td>Track15</td>
<td>222</td>
</tr>
<tr>
<td>Treatments</td>
<td>222</td>
</tr>
<tr>
<td>Trees</td>
<td>223</td>
</tr>
<tr>
<td>Trucks</td>
<td>224</td>
</tr>
<tr>
<td>tsum.test</td>
<td>224</td>
</tr>
<tr>
<td>Tv</td>
<td>227</td>
</tr>
<tr>
<td>Twin</td>
<td>228</td>
</tr>
<tr>
<td>Undergrad</td>
<td>229</td>
</tr>
<tr>
<td>Vacation</td>
<td>230</td>
</tr>
<tr>
<td>Vaccine</td>
<td>230</td>
</tr>
<tr>
<td>Vehicle</td>
<td>231</td>
</tr>
<tr>
<td>Verbal</td>
<td>232</td>
</tr>
<tr>
<td>Victoria</td>
<td>232</td>
</tr>
<tr>
<td>Viscosit</td>
<td>233</td>
</tr>
<tr>
<td>Visual</td>
<td>234</td>
</tr>
<tr>
<td>Vocab</td>
<td>234</td>
</tr>
<tr>
<td>Wastewat</td>
<td>235</td>
</tr>
<tr>
<td>Weather94</td>
<td>236</td>
</tr>
<tr>
<td>Wheat</td>
<td>236</td>
</tr>
<tr>
<td>Windmill</td>
<td>237</td>
</tr>
<tr>
<td>Window</td>
<td>238</td>
</tr>
<tr>
<td>Wins</td>
<td>238</td>
</tr>
<tr>
<td>Wool</td>
<td>239</td>
</tr>
<tr>
<td>Yearsunspot</td>
<td>240</td>
</tr>
<tr>
<td>z.test</td>
<td>241</td>
</tr>
<tr>
<td>zsum.test</td>
<td>243</td>
</tr>
</tbody>
</table>
The package BSDA provides data and functions for the book *Basic Statistics and Data Analysis*

**Author(s)**
Alan T. Arnholt  
Maintainer: <arnholtat@appstate.edu>

**References**

**Abbey**  
Daily price returns (in pence) of Abbey National shares between 7/31/91 and 10/8/91

**Description**
Data used in problem 6.39

**Usage**
Abbey

**Format**
A data frame with 50 observations on the following variable.

C1 a numeric vector

**Source**
Three samples to illustrate analysis of variance

Description

Data used in Exercise 10.1

Usage

Abc

Format

A data frame with 18 observations on the following 3 variables.

- **GroupA** a numeric vector
- **GroupB** a numeric vector
- **GroupC** a numeric vector

Source


Examples

```
attach(Abc)
STACKED <- stack(Abc)
STACKED[1:5,]
boxplot(values ~ ind, col = c("red", "blue", "green"), data = STACKED)
anova(lm(values ~ ind, data = STACKED))
remove(STACKED)
detach(Abc)
```
Abilene  

**Description**

Data used in Exercise 1.23 and 2.79

**Usage**

Abilene

**Format**

A data frame with 8 observations on the following 5 variables.

- **Crime**: a factor with levels Aggravated assault Arson Burglary Forcible rape Larceny theft Murder Robbery Vehicle theft
- **X1992**: a numeric vector
- **X92percent**: a numeric vector
- **X1999**: a numeric vector
- **X99percent**: a numeric vector

**Source**


**Examples**

```r
data(Abilene)
attach(Abilene)
par(mfrow=c(2,1))
barplot(X1992,names.arg=c("Murder","Rape","Robbery","Assault","Burglary","Larceny","V.Theft","Arson"),col="blue",main="Crime 1992")
par(mfrow=c(1,1))
detach(Abilene)
```

---

Ability  

**Description**

Data used in Exercise 8.57

**Usage**

Ability
Format
A data frame with 2 observations on the following 6 variables.

- gender: a factor with levels boys girls
- hopeless: a numeric vector
- belowavg: a numeric vector
- average: a numeric vector
- aboveavg: a numeric vector
- superior: a numeric vector

Source

Examples
X <- as.matrix(Ability[1:2,2:6])
chisq.test(X)

Abortion

Description
Data use in Exercise 8.51

Usage
Abortion

Format
A data frame with 51 observations on the following 9 variables.

- state: a factor with levels alabama alaska arizona arkansas california colorado connecticut delaware dist of columbia florida georgia hawaii idaho illinois indiana iowa kansas kentucky louisiana maine maryland massachusetts michigan minnesota mississippi missouri montana nebraska nevada new hampshire new jersey new mexico new york north carolina north dakota ohio oklahoma oregon pennsylvania rhode island south carolina south dakota tennessee texas utah vermont virginia washington west virginia wisconsin wyoming
- region: a factor with levels midwest northeast south west
- recode: a numeric vector
- X88rate: a numeric vector
- X92rate: a numeric vector
- X96rate: a numeric vector
- X88provid: a numeric vector
- X92provid: a numeric vector
- lowhigh: a numeric vector

Abortion rate by region of country
Absent

Source


Examples

```r
attach(Abortion)
AbortionRate <- cut(X96rate,breaks=c(0,20,10000) )
levels(AbortionRate) <- c("Low","High")
table(region,AbortionRate)
chisq.test(table(region,AbortionRate))
detach(Abortion)
```

---

<table>
<thead>
<tr>
<th>Absent</th>
<th>Number of absent days for 20 employees</th>
</tr>
</thead>
</table>

Description

Data used in Exercise 1.28

Usage

Absent

Format

A data frame with 20 observations on the following 4 variables.

- days  a numeric vector
- days_1  a numeric vector
- Count  a numeric vector
- Percent  a numeric vector

Source


Examples

```r
data(Absent)
attach(Absent)
table(days)
barplot(table(days),col="pink")
detach(Absent)
```
Achieve

*Math achievement test scores by gender for 25 high school students*

**Description**
Data used in Example 7.14 and Exercise 10.7

**Usage**
Achieve

**Format**
A data frame with 25 observations on the following 4 variables.
- Score  a numeric vector
- Gender  a numeric vector
- Female  a numeric vector
- Male  a numeric vector

**Source**

**Examples**
- str(Achieve)
- attach(Achieve)
- anova(lm(Score~Gender))
- t.test(Female,Male,var.equal=TRUE)
- detach(Achieve)

Adsales

*Number of ads versus number of sales for a retailer of satellite dishes*

**Description**
Data used in Exercise 9.15

**Usage**
Adsales

**Format**
A data frame with 6 observations on the following 2 variables.
- ads  a numeric vector
- sales  a numeric vector
Aggress

Source


Examples

attach(Adsales)
plot(ads, sales)
linmod <- lm(sales ~ ads)
abline(linmod)
summary(linmod)
detach(Adsales)

---

<table>
<thead>
<tr>
<th>Aggress</th>
<th>.Aggressive tendency scores for a group of teenage members of a street gang</th>
</tr>
</thead>
</table>

Description

Data used in Exercises 1.61 and 1.81

Usage

Aggress

Format

A data frame with 28 observations on the following variable.

aggress a numeric vector

Source


Examples

str(Aggress)
attach(Aggress)
EDA(aggres)
# OR
IQR(aggres)
diff(range(aggres))
detach(Aggress)
Aid

**Monthly payments per person for families in the AFDC federal program**

**Description**

Data used in Exercises 1.91 and 3.68

**Usage**

Aid

**Format**

A data frame with 51 observations on the following 2 variables.


- **payment** a numeric vector

**Source**


**Examples**

```r
str(Aid)
attach(Aid)
hist(payment)
boxplot(payment)
library(lattice)
dotplot(State~payment)
detach(Aid)
```

Aids

**Incubation times for 295 patients thought to be infected with HIV by a blood transfusion**

**Description**

Data used in Exercise 6.60

**Usage**

Aids
Airdisasters

Format
A data frame with 295 observations on the following 11 variables.
duration a numeric vector
age a numeric vector
group a numeric vector
duratio1 a numeric vector
children a numeric vector
duratio2 a numeric vector
adults a numeric vector
duratio3 a numeric vector
elderly a numeric vector
SRES1 a numeric vector
FITS1 a numeric vector

Source

Examples
str(Aids)
attach(Aids)
EDA(duration)
t.test(duration,mu=3/zero.noslash,alternative="greater")
SIGN.test(duration,md=24,alternative="greater")
detach(Aids)

Airdisasters  Aircraft disasters in five different decades

Description
Data used in Exercise 1.12

Usage
Airdisasters

Format
A data frame with 141 observations on the following 7 variables.
year  a numeric vector
deaths  a numeric vector
X1950  a numeric vector
X1960  a numeric vector
X1970  a numeric vector
X1980  a numeric vector
X1990  a numeric vector
Airline

Source


Examples

attach(Airdisasters)
STA <- stack(Airdisasters[,3:7])
library(lattice)
dotplot(ind~values,data=STA)
stripchart(x=list(X1950,X1960,X1970,X1980,X1990),method="stack",main="",pch=1,
 xlab="Number of Fatalities")
title(main="Aircraft Disasters 1950-1990")
detach(Airdisasters)

<table>
<thead>
<tr>
<th>Airline</th>
<th>Percentage of on-time arrivals and number of complaints for 11 airlines</th>
</tr>
</thead>
</table>

Description

Data for Exercise 2.9

Usage

Airline

Format

A data frame with 11 observations on the following 3 variables.

airline  a factor with levels Alaska Amer West American Continental Delta Northwest Pan Am Southwest TWA United USAir
ontime  a numeric vector
complnt  a numeric vector

Source


Examples

str(Airline)
attach(Airline)
barplot(complnt, names.arg=airline, col="lightblue")
plot(ontime, complnt)
detach(Airline)
Alcohol

Ages at which 14 female alcoholics began drinking

Description
Data used in Exercise 5.79

Usage
Alcohol

Format
A data frame with 14 observations on the following variable.

age  a numeric vector

Source

Examples
attach(Alcohol)
qqnorm(age)
qqline(age)
SIGN.test(age, md=2, conf.level=0.99)
detach(Alcohol)

Allergy

Allergy medicines by adverse events

Description
Data used in Exercise 8.22

Usage
Allergy

Format
A data frame with 3 observations on the following 4 variables.

C1.T  a factor with levels Drowsiness Headache Insomnia
Seldane  a numeric vector
Pseudoep  a numeric vector
Placebo  a numeric vector
**Source**


**Examples**

```r
attach(Allergy)
X <- as.matrix(Allergy[1:3,2:4])
chisq.test(X)
Xr <- as.matrix(Allergy[2:3,2:4])
chisq.test(Xr)
detach(Allergy)
```

---

<table>
<thead>
<tr>
<th>Anesthet</th>
<th>Recovery times for anesthetized patients</th>
</tr>
</thead>
</table>

**Description**

Data used in Exercise 5.58

**Usage**

`Anesthet`

**Format**

A data frame with 10 observations on the following variable.

- `recover` a numeric vector

**Source**


**Examples**

```r
attach(Anesthet)
str(Anesthet)
qqnorm(recover)
qqline(recover)
t.test(recover,conf.level=0.90)$conf
detach(Anesthet)
```
Apolipop

Math test scores versus anxiety scores before the test

Description
Data used in Exercise 2.96

Usage
Anxiety

Format
A data frame with 20 observations on the following 2 variables.
- anxiety a numeric vector
- math a numeric vector

Source

Examples
attach(Anxiety)
plot(anxiety,math)
cor(anxiety,math)
linmod <- lm(math~anxiety)
abline(linmod)
summary(linmod)
detach(Anxiety)

Apolipop

Level of apolipoprotein B and number of cups of coffee consumed per day for 15 adult males

Description
Data used in Examples 9.2 and 9.9

Usage
Apolipop

Format
A data frame with 15 observations on the following 4 variables.
- coffee a numeric vector
- apolipB a numeric vector
- SRES1 a numeric vector
- FITS1 a numeric vector
Append

Source

Examples
attach(Apolipop)
str(Apolipop)
plot(coffee, apolipB)
linmod <- lm(apolipB~coffee)
summary(linmod)
# plot(linmod)
detach(Apolipop)

---

Append

Median costs of an appendectomy at 20 hospitals in North Carolina

Description
Data for Exercise 1.119

Usage
Append

Format
A data frame with 20 observations on the following variable.

fee a numeric vector

Source

Examples
str(Append)
attach(Append)
ll <- mean(fee)-2*sd(fee)
ul <- mean(fee)+2*sd(fee)
limits <- c(ll, ul)
limits
fee[fee<ll | fee>ul]
detach(Append)
Appendec

*Median costs of appendectomies at three different types of North Carolina hospitals*

**Description**

Data for Exercise 10.60

**Usage**

Appendec

**Format**

A data frame with 59 observations on the following 5 variables.

- **Cost**: a numeric vector
- **Region**: a numeric vector
- **Rural**: a numeric vector
- **Regional**: a numeric vector
- **Metropol**: a numeric vector

**Source**


**Examples**

```r
str(Appendec)
attach(Appendec)
boxplot(Cost~Region)
anova(lm(Cost~as.factor(Region)))
detach(Appendec)
```

---

**Aptitude**

*Aptitude test scores versus productivity in a factory*

**Description**

Data for Exercises 2.1, 2.35 and 2.51

**Usage**

Aptitude
Archaeo

Format
A data frame with 8 observations on the following 4 variables.

- **aptitude**  a numeric vector
- **product**  a numeric vector
- **SRES1**  a numeric vector
- **FITS1**  a numeric vector

Source

Examples

```r
str(Aptitude)
attach(Aptitude)
plot(aptitude, product, main = "Exercise 2.1")
model1 <- lm(product ~ aptitude)
model1
abline(model1, col = "red", lwd = 3)
resid(model1)
fitted(model1)
cor(product, aptitude)
detach(Aptitude)
```

Archaeo  Radiocarbon ages of observations taken from an archaeological site

Description
Data for Exercises 5.120, 10.20 and Example 1.16

Usage

Archaeo

Format
A data frame with 60 observations on the following 6 variables.

- **phase1**  a numeric vector
- **phase2**  a numeric vector
- **phase3**  a numeric vector
- **phase4**  a numeric vector
- **age**  a numeric vector
- **phase**  a numeric vector

Source
Examples

```r
str(Arthriti)
attach(Arthriti)
boxplot(Time~Treatmnt)
anova(lm(Time~as.factor(Treatmnt)))
detach(Arthriti)
```

Description

Data for Exercise

Usage

Arthriti

Format

A data frame with 51 observations on the following 5 variables.

- TreatA: a numeric vector
- TreatB: a numeric vector
- TreatC: a numeric vector
- Time: a numeric vector
- Treatmnt: a numeric vector

Source


Examples

```r
str(Arthriti)
attach(Arthriti)
boxplot(Time~Treatmnt)
anova(lm(Time~as.factor(Treatmnt)))
detach(Arthriti)
```
Artifici

**Durations of operation for 15 artificial heart transplants**

**Description**
Data for Exercise 1.107

**Usage**
Artifici

**Format**
A data frame with 15 observations on the following variable.
duration  a numeric vector

**Source**

**Examples**
str(Artifici)
attach(Artifici)
stem(duration)
summary(duration)
values <- duration[duration<6.5]
values
summary(values)
detach(Artifici)
remove(values)

Asprin

**Dissolving time versus level of impurities in aspirin tablets**

**Description**
Data for Exercise 10.51

**Usage**
Asprin

**Format**
A data frame with 5 observations on the following 3 variables.X1.  a numeric vector
X5.  a numeric vector
X10.  a numeric vector
Asthmati

Source

Examples

str(Asprin)
attach(Asprin)
STACKED <- stack(Asprin)
STACKED[1:5,]
boxplot(values~ind,col=c("red","blue","green"),data=STACKED)
anova(lm(values~ind,data=STACKED))
remove(STACKED)
detach(Asprin)

Asthmati

Asthmatic relief index on 9 subjects given a drug and a placebo

Description
Data for Exercise 7.52

Usage
Asthmati

Format
A data frame with 9 observations on the following 3 variables.

Drug  a numeric vector
Placebo  a numeric vector
differ  a numeric vector

Source

Examples

str(Asthmati)
attach(Asthmati)
qqnorm(differ)
qline(differ)
shapiro.test(differ)
t.test(Placebo,Drug,paired=TRUE,mu=0,alternative="greater")
detach(Asthmati)
Description

Data for Exercises 2.2, 2.43 and 2.57

Usage

Attorney

Format

A data frame with 88 observations on the following 3 variables.

Staff   a numeric vector
Convict a numeric vector
District a factor with levels Albuquerque, Alexandria, Anchorage, Asheville, Atlanta, Baltimore, Baton Rouge, Billings, Birmingham, Boise, Boston, Buffalo, Burlington, Charleston, Cheyenne, Cincinnati, Cleveland, Columbia, Concord, Denver, Des Moines, East St. Louis, Fargo, Fort Smith, Fort Worth, Grand Rapids, Greensboro, Honolulu, Houston, Indianapolis, Jackson, Miss, Kansas City, Knoxville, Las Vegas, Lexington, Little Rock, Los Angeles, Louisville, Memphis, Miami, Milwaukee, Minneapolis, Mobile, Ala, Montgomery, Ala, Muskogee, Ok, Nashville, New Haven, Conn, New Orleans, New York (Brooklyn), New York (Manhattan), Newark, NJ, Oklahoma, City, Omaha, Oxford, Miss, Pensacola, Fl, Philadelphia, Phoenix, Pittsburgh, Portland, Maine, Portland, Ore, Providence, RI, Raleigh, NC, Roanoke, Va, Sacramento, Salt Lake City, San Antonio, San Diego, San Francisco, Savannah, Ga, Scranton, Pa, Seattle, Shreveport, La, Sioux Falls, SD, South Bend, Ind, Spokane, Wash, Springfield, Ill, St. Louis, Syracuse, NY, Tampa, Topeka, Kan, Tulsa, Tyler, Tex, Washington, Wheeling, WV, a factor with levels Albuquerque, Alexandria, Anchorage, Asheville, Atlanta, Baltimore, Baton Rouge, Billings, Birmingham, Boise, Boston, Buffalo, Burlington, Charleston, Cheyenne, Cincinnati, Cleveland, Columbia, Concord, Denver, Des Moines, East St. Louis, Fargo, Fort Smith, Fort Worth, Grand Rapids, Greensboro, Honolulu, Houston, Indianapolis, Jackson, Miss, Kansas City, Knoxville, Las Vegas, Lexington, Little Rock, Los Angeles, Louisville, Memphis, Miami, Milwaukee, Minneapolis, Mobile, Ala, Montgomery, Ala, Muskogee, Ok, Nashville, New Haven, Conn, New Orleans, New York (Brooklyn), New York (Manhattan), Newark, NJ, Oklahoma, City, Omaha, Oxford, Miss, Pensacola, Fl, Philadelphia, Phoenix, Pittsburgh, Portland, Maine, Portland, Ore, Providence, RI, Raleigh, NC, Roanoke, Va, Sacramento, Salt Lake City, San Antonio, San Diego, San Francisco, Savannah, Ga, Scranton, Pa, Seattle, Shreveport, La, Sioux Falls, SD, South Bend, Ind, Spokane, Wash, Springfield, Ill, St. Louis, Syracuse, NY, Tampa, Topeka, Kan, Tulsa, Tyler, Tex, Washington, Wheeling, WV.

Source


Examples

```r
str(Attorney)
attach(Attorney)
par(mfrow=c(1,2))
plot(Staff, Convict, main="With Washington, D.C.")
plot(Staff[-86], Convict[-86], main="Without Washington, D.C.")
par(mfrow=c(1,1))
detach(Attorney)
```
### Autogear

**Number of defective auto gears produced by two manufacturers**

**Description**

Data for Exercise 7.46

**Usage**

`Autogear`

**Format**

A data frame with 20 observations on the following 2 variables.

- `A` a numeric vector
- `B` a numeric vector

**Source**


**Examples**

```r
str(Autogear)
attach(Autogear)
t.test(A, B)
wilcox.test(A, B)
t.test(A, B, var.equal=TRUE)
detach(Autogear)
```

### Backtoback

**Illustrates inferences based on pooled t-test versus Wilcoxon rank sum test**

**Description**

Data for Exercise 7.40

**Usage**

`Backtoback`

**Format**

A data frame with 24 observations on the following 3 variables.

- `score` a numeric vector
- `group` a numeric vector
- `ranks` a numeric vector
**Bbsalaries**

Baseball salaries for members of five major league teams

**Description**

Data for Exercise 1.11

**Usage**

Bbsalaries

**Format**

A data frame with 33 observations on the following 5 variables.

- ANGLES a numeric vector
- ORIOLES a numeric vector
- REDSOXS a numeric vector
- WHITESOXS a numeric vector
- INDIANS a numeric vector

**Source**


**Examples**

```
str(Bbsalaries)
attach(Bbsalaries)
stripchart(x=list(INDIANS,WHITESOXS,REDSOXS,ORIOLES,ANGLES), xlab="Salary", method="stack", main="", pch=1, col="blue", group.names=c("Indians","White Sox", "Red Sox", "Orioles","Angels"))
title(main="Major League Salaries")
detach(Bbsalaries)
```
Graduation rates for student athletes and nonathletes in the Big Ten Conf.

Description
Data for Exercises 1.124 and 2.94

Usage
Bigten

Format
A data frame with 11 observations on the following 5 variables.

- School: a factor with levels Illinois Indiana Iowa Michigan Michigan State Minnesota Northwestern Ohio State Penn State Purdue Wisconsin
- X1984.85students: a numeric vector
- X1984.85athletes: a numeric vector
- X1993.94students: a numeric vector
- X1993.94athletes: a numeric vector

Source

Examples
str(Bigten)
attach(Bigten)
boxplot(X1993.94students,X1993.94athletes, names=c("Students","Athletes"),
ylab="1993-1994 Graduation Rates")
plot(X1993.94students,X1993.94athletes, xlab="1993-1994 students",
ylab="1993-1994 athletes")
detach(Bigten)

Test scores on first exam in biology class

Description
Data for Exercise 1.49

Usage
Biology
**Birth**

*Format*

A data frame with 30 observations on the following variable.

- **score**: a numeric vector

*Source*


*Examples*

```r
str(Biology)
attach(Biology)
hist(score,breaks="scott",col="brown",prob=TRUE,main="Problem 1.49")
lines(density(score),lwd=3)
detach(Biology)
```

**Usage**

- **Birth**

*Format*

A data frame with 51 observations on the following 3 variables.


- **X199/zero.noslashrate**: a numeric vector

- **X1998rate**: a numeric vector

*Source*

Examples

str(Birth)
attach(Birth)
stem(X1998rate)
hist(X1998rate,breaks=seq(10.9,21.9,1.0),xlab="1998 Birth Rate",
main="Figure 1.14 in BSDA",col="pink")
hist(X1998rate,breaks=seq(10.9,21.9,1.0),xlab="1998 Birth Rate",
main="Figure 1.14 in BSDA",col="pink",prob=TRUE)
lines(density(X1998rate),col="red",lwd=2)
detach(Birth)

Blackedu

Education level of blacks by gender

Description

Data for Exercise 8.55

Usage

Blackedu

Format

A data frame with 5 observations on the following 3 variables.

education a factor with levels bachelor deg graduate deg high sch dropout high sch graduate some college
female a numeric vector
male a numeric vector

Source


Examples

str(Blackedu)
attach(Blackedu)
Blackedu
chisq.test(Blackedu[,2:3])
detach(Blackedu)
Blood

Blood pressure of 15 adult males taken by machine and by an expert

Description
Data for Exercise 7.84

Usage
Blood

Format
A data frame with 15 observations on the following 2 variables.

Machine a numeric vector
Expert a numeric vector

Source

Examples
str(Blood)
attach(Blood)
DIF <- Machine - Expert
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(Machine,Expert,paired=TRUE)
detach(Blood)
remove(DIF)

Board
Incomes of board members from three different universities

Description
Data for Exercise 10.14

Usage
Board

Format
A data frame with 7 observations on the following 3 variables.

UnivA a numeric vector
UnivB a numeric vector
UnivC a numeric vector
Bones

Source

Examples

str(Board)
attach(Board)
STACKED <- stack(Board)
STACKED[1:5,]
boxplot(values~ind,col=c("red","blue","green"),data=STACKED)
remove(STACKED)
detach(Board)

Bones

Bone density measurements of 35 physically active and 35 non-active women

Description
Data for Exercise 7.22

Usage
Bones

Format
A data frame with 70 observations on the following 5 variables.

Active a numeric vector
Nonactive a numeric vector
Density a numeric vector
group a numeric vector
Ranks a numeric vector

Source

Examples

str(Bones)
attach(Bones)
t.test(Active,Nonactive,alternative="greater")
wilcox.test(Active,Nonactive,alternative="greater")
detach(Bones)
Books

Number of books read and final spelling scores for 17 third graders

Description
Data for Exercise 9.53

Usage
Books

Format
A data frame with 17 observations on the following 2 variables.

book a numeric vector
spelling a numeric vector

Source

Examples
data(Books)

Bookstor

Prices paid for used books at three different bookstores

Description
Data for Exercise 10.30 and 10.31

Usage
Bookstor

Format
A data frame with 72 observations on the following 6 variables.

StoreA a numeric vector
StoreB a numeric vector
StoreC a numeric vector
Dollars a numeric vector
Store a numeric vector
Ranks a numeric vector
Brain

Source

Examples
```r
str(Bookstor)
attach(Bookstor)
boxplot(Dollars~Store)
kruskal.test(Dollars~as.factor(Store))
detach(Bookstor)
```

Brain weight versus body weight of 28 animals

Description
Data for Exercises 2.15, 2.44, 2.58 and Examples 2.3 and 2.20

Usage
Brain

Format
A data frame with 28 observations on the following 5 variables.

Species  a factor with levels African elephant Asian Elephant Brachiosaurus Cat Chimpanzee Cow Diplodocus Donkey Giraffe Goat Gorilla Gray wolf Guinea Pig Hamster Horse Human Jaguar Kangaroo Mole Mouse Mt Beaver Pig Potar monkey Rabbit Rat Rhesus monkey Sheep Triceratops

body.wt  a numeric vector

brain.wt  a numeric vector

logbody  a numeric vector

logbrain  a numeric vector

Source

Examples
```r
str(Brain)
attach(Brain)
plot(logbody,logbrain,pch=19,col="blue",main="Example 2.3")
model <- lm(logbrain~logbody)
abline(model)
detach(Brain)
```
**Bumpers**

*Repair costs of vehicles crashed into a barrier at 5 miles per hour*

**Description**

Data for Exercise 1.73

**Usage**

Bumpers

**Format**

A data frame with 23 observations on the following 2 variables.

- **Car**: a factor with levels Buick Century Buick Skylark Chevrolet Cavalier Chevrolet Corsica Chevrolet Lumina Dodge Dynasty Dodge Monaco Ford Taurus Ford Tempo Honda Accord Hyundai Sonata Mazda 626 Mitsubishi Galant Nissan Stanza Oldsmobile Calais Oldsmobile Ciera Plymouth Acclaim Pontiac 6000 Pontiac Grand Am Pontiac Sunbird Saturn SL2 Subaru Legacy Toyota Camry
- **repair**: a numeric vector

**Source**


**Examples**

```r
str(Bumpers)
attach(Bumpers)
EDA(repair)
sum(repair>(mean(repair)-sd(repair)) & repair < (mean(repair)+sd(repair)))/length(repair)
stripchart(repair,method="stack")
library(lattice)
dotplot(Car~repair)
detach(Bumpers)
```

**Bus**

*Attendance of bus drivers versus attendance*

**Description**

Data for Exercise 8.25

**Usage**

Bus
Bypass

Format
A data frame with 2 observations on the following 6 variables.

Attend a factor with levels Absent Present
AM a numeric vector
Noon a numeric vector
PM a numeric vector
Swing a numeric vector
Split a numeric vector

Source

Examples
str(Bus)
attach(Bus)
Bus
chisq.test(Bus[,2:6])
detach(Bus)

Bypass

Median charges for coronary bypass at 17 hospitals in North Carolina

Description
Data for Exercises 5.104 and 6.43

Usage
Bypass

Format
A data frame with 17 observations on the following 2 variables.

hospital a factor with levels Carolinas Med Ct Duke Med Ct Durham Regional Forsyth Memorial Frye Regional High Point Regional Memorial Mission Mercy Moore Regional Moses Cone Memorial NC Baptist New Hanover Regional Pitt Co. Memorial Presbyterian Rex Univ of North Carolina Wake County
charge a numeric vector

Source
Examples

```
str(Bypass)
attach(Bypass)
EDA(charge)
t.test(charge,conf.level=.9)$conf
t.test(charge,mu=35000)
detach(Bypass)
```

---

Cabinets

*Estimates of costs of kitchen cabinets by two suppliers on 20 prospective homes*

Description

Data for Exercise 7.83

Usage

`Cabinets`

Format

A data frame with 20 observations on the following 3 variables.

- **Home**: a numeric vector
- **SupplA**: a numeric vector
- **SupplB**: a numeric vector

Source


Examples

```
str(Cabinets)
attach(Cabinets)
DIF <- SupplA - SupplB
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(SupplA,SupplB,paired=TRUE)
wilcox.test(SupplA,SupplB,paired=TRUE)
detach(Cabinets)
remove(DIF)
```
Cancer

Survival times of terminal cancer patients treated with vitamin C

**Description**

Data for Exercises 6.55 and 6.64

**Usage**

Cancer

**Format**

A data frame with 17 observations on the following 5 variables.

- stomach: a numeric vector
- bronchus: a numeric vector
- colon: a numeric vector
- ovary: a numeric vector
- breast: a numeric vector

**Source**


**Examples**

```r
str(Cancer)
attach(Cancer)
EDA(stomach)
SIGN.test(stomach, md=100, alternative="greater")
detach(Cancer)
```

---

Carbon

Carbon monoxide level measured at three industrial sites

**Description**

Data for Exercise 10.28 and 10.29

**Usage**

Carbon
**Cat**

**Format**

A data frame with 24 observations on the following 6 variables.

- SiteA: a numeric vector
- SiteB: a numeric vector
- SiteC: a numeric vector
- monoxide: a numeric vector
- Site: a numeric vector
- Ranks: a numeric vector

**Source**


**Examples**

```r
str(Carbon)
attach(Carbon)
boxplot(monoxide~Site)
kruskal.test(monoxide~as.factor(Site))
detach(Carbon)
```

---

**Cat**

Reading scores on the California achievement test for a group of 3rd graders

**Description**

Data for Exercise 1.116

**Usage**

Cat

**Format**

A data frame with 17 observations on the following variable.

- score: a numeric vector

**Source**


**Examples**

```r
str(Cat)
attach(Cat)
stem(score)
fivenum(score)
boxplot(score,main="Problem 1.116",col="green")
detach(Cat)
```
Censored

Entry age and survival time of patients with small cell lung cancer under two different treatments

**Description**

Data for Exercises 7.34 and 7.48

**Usage**

Censored

**Format**

A data frame with 121 observations on the following 8 variables.

- `surviva` a numeric vector
- `agea` a numeric vector
- `censora` a numeric vector
- `survivb` a numeric vector
- `ageb` a numeric vector
- `censorb` a numeric vector
- `survival` a numeric vector
- `group` a numeric vector

**Source**


**Examples**

```r
str(Censored)
attach(Censored)
boxplot(survival~group)
detach(Censored)
```

Challeng

Temperatures and O-ring failures for the launches of the space shuttle Challenger

**Description**

Data for Examples 1.11, 1.12, 1.13, 2.11 and 5.1

**Usage**

Challeng
Chemist

Format
A data frame with 25 observations on the following 4 variables.

flight  a factor with levels 1 2 3 4 41-b 41-c 41-d 41-g 5 51-a 51-b 51-c 51-d 51-f 51-g 51-i 51-j 6 61-a 61-b 61-c 61-i 7 8 9
temp   a numeric vector
Failures a numeric vector

Source

Examples
str(Challeng)
attach(Challeng)
stem(temp)
summary(temp)
IQR(temp)
quantile(temp)
fivenum(temp)
stem(sort(temp)[-1])
summary(sort(temp)[-1])
IQR(sort(temp)[-1])
quantile(sort(temp)[-1])
fivenum(sort(temp)[-1])
par(mfrow=c(1,2))
qqnorm(temp)
qqline(temp)
qqnorm(sort(temp)[-1])
qqline(sort(temp)[-1])
par(mfrow=c(1,1))
detach(Challeng)

Chemist

Starting salaries of 50 chemistry majors

Description
Data for Example 5.3

Usage
Chemist

Format
A data frame with 50 observations on the following variable.
salary  a numeric vector
Source


Examples

```r
str(Chemist)
attach(Chemist)
EDA(salary)
detach(Chemist)
```

---

**Chesapeake**

*Surface salinity measurements taken offshore from Annapolis, Maryland in 1927*

Description

Data for Exercise 6.46

Usage

```r
Chesapeake
```

Format

A data frame with 16 observations on the following variable.

```r
salinity  a numeric vector
```

Source


Examples

```r
str(Chesapeake)
attach(Chesapeake)
qqnorm(salinity)
qqline(salinity)
shapiro.test(salinity)
t.test(salinity,mu=7)
detach(Chesapeake)
```
Chevy

Insurance injury ratings of Chevrolet vehicles for 1990 and 1993 models

Description
Data for Exercise 8.35

Usage
data(Chevy)

Format
A data frame with 2 observations on the following 6 variables.

year  a factor with levels 88–90 91–93
A  a numeric vector
B  a numeric vector
C  a numeric vector
D  a numeric vector
F  a numeric vector

Source

Examples
str(Chevy)
attach(Chevy)
Chevy
chisq.test(Chevy[,2:6])
detach(Chevy)

Chicken

Weight gain of chickens fed three different rations

Description
Data for Exercise 10.15

Usage
data(Chicken)
**Chipavg**

**Description**

Data for Exercises 6.49 and 7.47

**Usage**

Chipavg

**Format**

A data frame with 30 observations on the following 3 variables.

- **wafer1** a numeric vector
- **wafer2** a numeric vector
- **thickness** a numeric vector

**Source**


---

**Chipavg**

Measurements of the thickness of the oxide layer of manufactured integrated circuits

**Format**

A data frame with 13 observations on the following 3 variables.

- **Ration1** a numeric vector
- **Ration2** a numeric vector
- **Ration3** a numeric vector

**Source**


**Examples**

```r
str(Chicken)
attach(Chicken)
STACKED <- stack(Chicken)
STACKED[1:5,]
boxplot(values~ind,col=c("red","blue","green"),data=STACKED)
anova(lm(values~ind,data=STACKED))
remove(STACKED)
detach(Chicken)
```
Chips

Examples

str(Chipavg)
attach(Chipavg)
EDA(thickness)
t.test(thickness,mu=1)
boxplot(wafer1,wafer2,name=c("Wafer 1","Wafer 2"))
shapiro.test(wafer1)
shapiro.test(wafer2)
t.test(wafer1,wafer2,var.equal=TRUE)
detach(Chipavg)

Chips

Four measurements on a first wafer and four measurements on a second wafer selected from 30 lots

Description

Data for Exercise 10.9

Usage

Chips

Format

A data frame with 30 observations on the following 8 variables.

wafer11 a numeric vector
wafer12 a numeric vector
wafer13 a numeric vector
wafer14 a numeric vector
wafer21 a numeric vector
wafer22 a numeric vector
wafer23 a numeric vector
wafer24 a numeric vector

Source


Examples

str(Chips)
attach(Chips)
boxplot(wafer11,wafer12,wafer13,wafer14,wafer21,wafer22,wafer23,wafer24)
detach(Chips)
Cigar

Milligrams of tar in 25 cigarettes selected randomly from 4 different brands

Description
Data for Example 10.4

Usage
Cigar

Format
A data frame with 100 observations on the following 6 variables.
- `brandA` a numeric vector
- `brandB` a numeric vector
- `brandC` a numeric vector
- `brandD` a numeric vector
- `tar` a numeric vector
- `brand` a numeric vector

Source

Examples
```
str(Cigar)
attach(Cigar)
boxplot(tar~brand)
anova(lm(tar~as.factor(brand)))
detach(Cigar)
```

Cigarette

Effect of mother's smoking on birth weight of newborn

Description
Data for Exercise 2.27

Usage
Cigarette
**CIsim**

**Format**

A data frame with 16 observations on the following 2 variables.

- `cigarette` a numeric vector
- `weight` a numeric vector

**Source**


**Examples**

```r
str(Cigarette)
attach(Cigarette)
plot(cigarette,weight)
model <- lm(weight~cigarette)
abline(model)
cor(weight,cigarette)
detach(Cigarette)
```

---

**CIsim**  

*Confidence Interval Simulation Program*

**Description**

This program simulates random samples from which it constructs confidence intervals for one of the parameters mean (Mu), variance (Sigma), or proportion of successes (Pi).

**Usage**

```r
CIsim(samples=100, n=30, mu=0, sigma=1, conf.level = 0.95, type = "Mean")
```

**Arguments**

- `samples` the number of samples desired.
- `n` the size of each sample.
- `mu` if constructing confidence intervals for the population mean or the population variance, mu is the population mean (i.e., type is one of either "Mean", or "Var"). If constructing confidence intervals for the population proportion of successes, the value entered for mu represents the population proportion of successes (Pi), and as such, must be a number between 0 and 1.
- `sigma` the population standard deviation. sigma is not required if confidence intervals are of type "Pi".
- `conf.level` confidence level for the graphed confidence intervals, restricted to lie between zero and one.
- `type` character string, one of "Mean", "Var" or "Pi", or just the initial letter of each, indicating the type of confidence interval simulation to perform.
Details

Default is to construct confidence intervals for the population mean. Simulated confidence intervals for the population variance or population proportion of successes are possible by selecting the appropriate value in the type argument.

Value

Graph depicts simulated confidence intervals. The number of confidence intervals that do not contain the parameter of interest are counted and reported in the commands window.

Author(s)

Alan T. Arnholt

Examples

```R
CIsim(100, 30, 100, 10)
# Simulates 100 samples of size 30 from
# a normal distribution with mean 100
# and standard deviation 10. From the
# 100 simulated samples, 95% confidence
# intervals for the Mean are constructed
# and depicted in the graph.

CIsim(100, 30, 100, 10, type="Var")
# Simulates 100 samples of size 30 from
# a normal distribution with mean 100
# and standard deviation 10. From the
# 100 simulated samples, 95% confidence
# intervals for the variance are constructed
# and depicted in the graph.

CIsim(100, 50, .5, type="Pi", conf.level=.9)
# Simulates 100 samples of size 50 from
# a binomial distribution where the population
# proportion of successes is 0.5. From the
# 100 simulated samples, 90% confidence
# intervals for Pi are constructed
# and depicted in the graph.
```

Citrus

Percent of peak bone density of different aged children

Description

Data for Exercise 9.7

Usage

Citrus
### Format
A data frame with 9 observations on the following 2 variables.

- **age**: a numeric vector
- **percent**: a numeric vector

### Source

### Examples
```r
str(Citrus)
attach(Citrus)
model <- lm(percent~age)
summary(model)
anova(model)
detach(Citrus)
remove(model)
```

---

### Clean

**Residual contaminant following the use of three different cleansing agents**

### Description
Data for Exercise 10.16

### Usage
`Clean`

### Format
A data frame with 45 observations on the following 5 variables.

- **A**: a numeric vector
- **B**: a numeric vector
- **C**: a numeric vector
- **clean**: a numeric vector
- **agent**: a numeric vector

### Source

### Examples
```r
str(Clean)
attach(Clean)
boxplot(clean~agent,col=c("red","blue","green"))
anova(lm(clean~as.factor(agent)))
detach(Clean)
```
### Coaxial

**Signal loss from three types of coaxial cable**

**Description**
Data for Exercise 10.24 and 10.25

**Usage**
Coaxial

**Format**
A data frame with 45 observations on the following 5 variables.

- **Type.A**: a numeric vector
- **Type.B**: a numeric vector
- **Type.C**: a numeric vector
- **Signal**: a numeric vector
- **Cable**: a numeric vector

**Source**

**Examples**

```r
str(Coaxial)
attach(Coaxial)
boxplot(Signal~Cable)
kruskal.test(Signal~as.factor(Cable))
detach(Coaxial)
```

### Coffee

**Productivity of workers with and without a coffee break**

**Description**
Data for Exercise 7.55

**Usage**
Coffee

**Format**
A data frame with 9 observations on the following 4 variables.

- **Without**: a numeric vector
- **With**: a numeric vector
- **differ**: a numeric vector
- **sgnrnks**: a numeric vector
**Coins**

**Source**


**Examples**

```r
str(Coffee)
attach(Coffee)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
t.test(With,Without,paired=TRUE,alternative="greater")
wilcox.test(With,Without,paired=TRUE,alternative="greater")
detach(Coffee)
```

---

<table>
<thead>
<tr>
<th>Coins</th>
<th>Yearly returns on 12 investments</th>
</tr>
</thead>
</table>

**Description**

Data for Exercise 5.68

**Usage**

Coins

**Format**

A data frame with 12 observations on the following variable.

coins  a numeric vector

**Source**


**Examples**

```r
str(Coins)
attach(Coins)
qqnorm(coins)
qqline(coins)
eda(coins)
detach(Coins)
```
Combinations

Description
Computes all possible combinations of n objects taken k at a time.

Usage
Combinations(n, k)

Arguments
n  a number.
k  a number less than or equal to n.

Value
Returns a matrix containing the possible combinations of n objects taken k at a time.

See Also
SRS

Examples
Combinations(5,2)
  # The columns in the matrix list the values of the 10 possible
  # combinations of 5 things taken 2 at a time.

Commuting times for selected cities in 1980 and 1990

Description
Data for Exercises 1.13, and 7.85

Usage
Commute

Format
A data frame with 39 observations on the following 3 variables.

City  a factor with levels Atlanta Baltimore Boston Buffalo Charlotte Chicago Cincinnati Cleveland Columbus Dallas Denver Detroit Hartford Houston Indianapolis Kansas City Los Angeles Miami Milwaukee Minneapolis New Orleans New York Norfolk Orlando Philadelphia Phoenix Pittsburgh Portland Providence Rochester Sacramento Salt Lake City San Antonio San Diego San Francisco Seattle St. Louis Tampa Washington
X1980  a numeric vector
X1990  a numeric vector
Concept

Source


Examples

```r
str(Commute)
attach(Commute)
stripchart(x=list(X198/zero.noslash,X199/zero.noslash),method="stack",pch=1,cex=2,col=c("red","blue"),
group.names=c("1980","1990"),main="",xlab="minutes")
title(main="Commute Time")
boxplot(X198/zero.noslash,X199/zero.noslash,col=c("red","blue"),names=c("1980","1990"),horizontal=TRUE,las=1)
library(lattice)
commute <- stack(Commute)
commute[1:5,]
attach(commute)
stripplot(ind~values,jitter=TRUE)
dotplot(ind~values)
bwplot(ind~values)
remove(commute)
detach(Commute)
```

```
<table>
<thead>
<tr>
<th>Concept</th>
<th>Tennessee self concept scale scores for a group of teenage boys</th>
</tr>
</thead>
</table>

Description

Data for Exercise 16.8 and 1.82

Usage

Concept

Format

A data frame with 28 observations on the following variable.

self  a numeric vector

Source


Examples

```r
str(Concept)
attach(Concept)
summary(self)
sd(self)
diff(range(self))
IQR(self)
summary(self/10)
IQR(self/10)
sd(self/10)
diff(range(self/10))
detach(Concept)
```
Concrete

Compressive strength of concrete blocks made by two different methods

Description

Data for Exercise 7.17

Usage

Concrete

Format

A data frame with 20 observations on the following 3 variables.

- Strength: a numeric vector
- Method: a numeric vector
- Ranks: a numeric vector

Source


Examples

```r
str(Concrete)
attach(Concrete)
wilcox.test(Strength~Method,alternative="greater")
detach(Concrete)
```

Corn

Comparison of the yields of a new variety and a standard variety of corn planted on 12 plots of land

Description

Data for Exercise 7.77

Usage

Corn

Format

A data frame with 12 observations on the following 3 variables.

- New: a numeric vector
- Standard: a numeric vector
- differ: a numeric vector
Correlat

Source


Examples

str(Corn)
attach(Corn)
boxplot(differ)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
t.test(New,Standard,paired=TRUE,alternative="greater")
detach(Corn)

---

Correlat  Exercise to illustrate correlation

Description

Data for Exercise 2.23

Usage

Correlat

Format

A data frame with 13 observations on the following 2 variables.

X  a numeric vector
Y  a numeric vector

Source


Examples

str(Correlat)
attach(Correlat)
plot(X,Y)
model <- lm(Y~X)
abline(model)
detach(Correlat)
Counsel

Scores of 18 volunteers who participated in a counseling process

**Description**

Data for Exercise 6.96

**Usage**

Counsel

**Format**

A data frame with 18 observations on the following variable.

- **score** a numeric vector

**Source**


**Examples**

```r
str(Counsel)
attach(Counsel)
EDA(score)
t.test(score,mu=70)
detach(Counsel)
```

Cpi

Consumer price index from 1979 to 1998

**Description**

Data for Exercise 1.34

**Usage**

Cpi

**Format**

A data frame with 20 observations on the following 2 variables.

- **Year** a numeric vector
- **CPI** a numeric vector

**Source**

Examples

```
str(Cpi)
attach(Cpi)
plot(Year,CPI,type="l",lty=2,lwd=2,col="red")
names(CPI) <- Year
barplot(CPI,col="pink",las=2,main="Problem 1.34")
detach(Cpi)
```

Description

Data for Exercises 1.90, 2.32, 3.64, and 5.113

Usage

```
Crime
```

Format

A data frame with 51 observations on the following 3 variables.

```
X1983 a numeric vector
X1993 a numeric vector
```

Source


Examples

```
str(Crime)
attach(Crime)
boxplot(X1983,X1993, names=c("1983","1993"), xlab="Year",
ylab="Crime Rate per 100,000 Inhabitants", main="Problem 1.90")
plot(X1983,X1993)
detach(Crime)
```
Darwin

*Charles Darwin's study of cross-fertilized and self-fertilized plants*

**Description**

Data for Exercise 7.62

**Usage**

Darwin

**Format**

A data frame with 30 observations on the following 7 variables.

- `pot` a numeric vector
- `cross` a numeric vector
- `self` a numeric vector
- `height` a numeric vector
- `method` a numeric vector
- `differ` a numeric vector
- `sgnrnks` a numeric vector

**Source**


**Examples**

```r
str(Darwin)
attach(Darwin)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
wilcox.test(cross, self, paired=TRUE)
detach(Darwin)
```

---

Dealers

*Automobile dealers classified according to type dealership and service rendered to customers*

**Description**

Data for Example 2.22

**Usage**

Dealers
**Defectiv**

**Format**

A data frame with 6 observations on the following 2 variables.

- Replace  a numeric vector
- Recommnd a numeric vector

**Source**


**Examples**

```r
str(Dealers)
attach(Dealers)
Deal <- as.matrix(Dealers)
rownames(Deal) <- c("Honda","Toyota","Mazda","Ford","Dodge","Saturn")
Dealers
barplot(t(Deal),beside=TRUE,legend=TRUE)
detach(Dealers)
remove(Deal)
```

---

**Defectiv**

*Number of defective items produced by 20 employees*

**Description**

Data for Exercise 1.27

**Usage**

Defectiv

**Format**

A data frame with 20 observations on the following 4 variables.

- C1 a numeric vector
- number a numeric vector
- Count a numeric vector
- Percent a numeric vector

**Source**


**Examples**

```r
str(Defectiv)
attach(Defectiv)
table(C1)
barplot(table(C1),col="pink",ylab="Frequency",
       xlab="Defective Items Produced by Employees",main="Problem 1.27")
detach(Defectiv)
```
Degree

Percent of bachelor’s degrees awarded women in 1970 versus 1990

Description

Data for Exercise 2.75

Usage

Degree

Format

A data frame with 11 observations on the following 3 variables.

Field a factor with levels All fields Business Education Engineering Fine Arts Foreign Lng Health Life Sciences Physical Sci Psychology Social Science

X1970 a numeric vector

X1990 a numeric vector

Source


Examples

str(Degree)
attach(Degree)
Dmat <- as.matrix(Degree[,2:3])
rownames(Dmat) <- Field
colnames(Dmat) <- c("1970","1990")
Dmat
barplot(t(Dmat),beside=TRUE,legend=TRUE,cex.names=.5)
detach(Degree)
remove(Dmat)

Delay

Delay times on 20 flights from four major air carriers

Description

Data for Exercise 10.55

Usage

Delay
**Depend**

**Format**

A data frame with 80 observations on the following 6 variables.

- CarrierA a numeric vector
- CarrierB a numeric vector
- CarrierC a numeric vector
- CarrierD a numeric vector
- delay a numeric vector
- Carrier a numeric vector

**Source**


**Examples**

```r
str(Delay)
attach(Delay)
boxplot(delay~Carrier)
kruskal.test(delay~as.factor(Carrier))
detach(Delay)
```

**Description**

Data for Exercise 1.26

**Usage**

Depend

**Format**

A data frame with 50 observations on the following 4 variables.

- C1 a numeric vector
- number a numeric vector
- Count a numeric vector
- Percent a numeric vector

**Source**

Examples

```r
str(Depend)
attach(Depend)
table(C1)
barplot(table(C1),col="lightblue",main="Problem 1.26",
  xlab="Number of Dependent Children",ylab="Frequency")
detach(Depend)
```

Detroit Educational levels of a sample of 40 auto workers in Detroit

Description

Data for Exercise 5.21

Usage

Detroit

Format

A data frame with 40 observations on the following variable.

educ a numeric vector

Source


Examples

```r
str(Detroit)
attach(Detroit)
EDA(educ)
detach(Detroit)
```

Develop Demographic characteristics of developmental students at 2-year colleges and 4-year colleges

Description

Data used for Exercise 8.50

Usage

Develop

Format

**Devmath**

*Source*

*Examples*
Develop
chisq.test(Develop)

---

<table>
<thead>
<tr>
<th>Devmath</th>
<th>Test scores for students who failed developmental mathematics in the fall semester 1995</th>
</tr>
</thead>
</table>

*Description*
Data for Exercise 6.47

*Usage*
Devmath

*Format*
A data frame with 40 observations on the following variable.

- **score** a numeric vector

*Source*

*Examples*
str(Devmath)
attach(Devmath)
EDA(score)
t.test(score,mu=80,alternative="less")
detach(Devmath)

---

<table>
<thead>
<tr>
<th>Dice</th>
<th>Outcomes and probabilities of the roll of a pair of fair dice</th>
</tr>
</thead>
</table>

*Description*
Data for Exercise 3.109

*Usage*
Dice
Diesel

Format

A data frame with 11 observations on the following 2 variables.

- x a numeric vector
- P.x. a numeric vector

Source


Examples

str(Dice)
attach(Dice)
roll1 <- sample(1:6,2/000,replace=TRUE)
roll2 <- sample(1:6,2/000,replace=TRUE)
outcome <- roll1+roll2
table(outcome)/length(outcome)
detach(Dice)
remove(roll1,roll2,outcome)

Diesel

Diesel fuel prices in 1999-2000 in nine regions of the country

Description

Data for Exercise 2.8

Usage

Diesel

Format

A data frame with 65 observations on the following 11 variables.

- NatAvg a numeric vector
- EstCst a numeric vector
- NE a numeric vector
- CltAtl a numeric vector
- LwrAtl a numeric vector
Diplomat

Gulf a numeric vector
Rocky a numeric vector
WstMt a numeric vector
Coast a numeric vector
Calif a numeric vector

Source


Examples

```r
str(Diesel)
attach(Diesel)
boxplot(NatAvg,EstCst,Gulf,Rocky,Calif,
names=c("National Average","East Coast","Gulf","Rocky","California"),col="pink")
```

---

**Description**

Data for Exercises 1.14 and 1.47

**Usage**

Diplomat

**Format**

A data frame with 10 observations on the following 4 variables.

Country a factor with levels Brazil Bulgaria Egypt Indonesia Israel Nigeria Russia S. Korea Ukraine Venezuela
Number a numeric vector
rate a numeric vector
Code a factor with levels Br Bu Eg In Is Ni Ru SK Uk Ve

Source


Examples

```r
str(Diplomat)
attach(Diplomat)
par(mfrow=c(1,2))
names(Number) <- Country
dotchart(Number,main="Number of Tickets",col="blue",pch=1)
names(rate) <- Country
dotchart(rate,main="Tickets/Vehicle/Month",col="red",pch=2)
barplot(rate,col="pink")
detach(Diplomat)
```
### Disposal

Toxic intensity for plants producing herbicidal preparations

#### Description
Data for Exercise 1.127

#### Usage
Disposal

#### Format
A data frame with 29 observations on the following variable.
- **pounds**: a numeric vector

#### Source

#### Examples
- `str(Disposal)`
- `attach(Disposal)`
- `stem(pounds)`
- `fivenum(pounds)`
- `EDA(pounds)`
- `detach(Disposal)`

### Dogs

Rankings of the favorite breeds of dogs

#### Description
Data for Exercise 2.88

#### Usage
Dogs

#### Format
A data frame with 20 observations on the following 5 variables.
- **Dog**: a factor with levels Beagle Boxer Chihuahua Chow Dachshund Dalmatian Doberman Huskie Labrador Pomeranian Poodle Retriever Rotweiler Schnauzer Shepherd Shetland ShihTzu Spaniel Springer Yorkshire
- **X1992**: a numeric vector
- **X1993**: a numeric vector
- **X1997**: a numeric vector
- **X1998**: a numeric vector
Domestic

Source


Examples

```
str(Dogs)
attach(Dogs)
cor(Dogs[,2:5])
detach(Dogs)
```

---

<table>
<thead>
<tr>
<th>Domestic</th>
<th>Rates of domestic violence per 1,000 women by age groups</th>
</tr>
</thead>
</table>

Description

Data for Exercise 1.20

Usage

Domestic

Format

A data frame with 5 observations on the following 2 variables.

- **Age** a factor with levels 12-19 20-24 25-34 35-49 50-64
- **Rate** a numeric vector

Source


Examples

```
str(Domestic)
attach(Domestic)
names(Rate) <- Age
barplot(Rate,col="gold")
pie(Rate)
detach(Domestic)
```
Dopamine

*Dopamine b-hydroxylase activity of schizophrenic patients treated with an antipsychotic drug*

**Description**

Data for Exercises 5.14 and 7.49

**Usage**

data(Dopamine)

**Format**

A data frame with 25 observations on the following 4 variables.

- nonpsych a numeric vector
- psychotic a numeric vector
- DBH a numeric vector
- group a numeric vector

**Source**


**Examples**

```r
str(Dopamine)
attach(Dopamine)
boxplot(DBH~group,names=c("Non Psychotic","Psychotic"))
t.test(DBH~group,var.equal=TRUE)
detach(Dopamine)
```

Dowjones

*Closing yearend Dow Jones Industrial averages from 1896 through 2000*

**Description**

Data for Exercise 1.35

**Usage**

Dowjones

**Format**

A data frame with 105 observations on the following 3 variables.

- year  a numeric vector
- close  a numeric vector
- X.change a numeric vector
Drink

Source

Examples

```r
str(Dowjones)
attach(Dowjones)
plot(year,close,type="l",lty=2,lwd=2,col="blue")
barplot(close,col="blue",las=2,main="Problem 1.35",names.arg=FALSE)
detach(Dowjones)
```

---

Drink

*Opinion on referendum by view on moral issue of selling alcoholic beverages*

Description
Data for Exercise 8.53

Usage

```
Drink
```

Format
A data frame with 3 observations on the following 4 variables.

- **drink**  a factor with levels immoral ok tolerated
- **For**  a numeric vector
- **Against**  a numeric vector
- **undecide**  a numeric vector

Source

Examples

```r
str(Drink)
attach(Drink)
Drink
chisq.test(Drink[,2:4])
detach(Drink)
```
Dyslexia

Data on a group of college students diagnosed with dyslexia

Description
Data for Exercise 2.90

Usage
data(Dyslexia)

---

Drug

Number of trials to master a task for a group of 28 subjects assigned to a control and an experimental group

Description
Data for Exercise 7.15

Usage
Drug

Format
A data frame with 28 observations on the following 2 variables.

trials a numeric vector
group a numeric vector

Source

Examples
str(Drug)
attach(Drug)
boxplot(trials~group)
wilcox.test(trials~group)
detach(Drug)
Earthqk

Format
A data frame with 8 observations on the following 7 variables.

- words  a numeric vector
- age    a numeric vector
- gender a factor with levels f m
- handed a factor with levels l r
- weight a numeric vector
- height a numeric vector
- children a numeric vector

Source

Examples
str(Dyslexia)
attach(Dyslexia)
plot(weight,height)
plot(handed,words)
detach(Dyslexia)

Earthqk
One hundred year record of worldwide seismic activity(1770-1869)

Description
Data for Exercise 6.97

Usage
Earthqk

Format
A data frame with 100 observations on the following 2 variables.

- year   a numeric vector
- severity a numeric vector

Source

Examples
str(Earthqk)
attach(Earthqk)
EDA(severity)
t.test(severity,mu=100,alternative="greater")
detach(Earthqk)
EDA

Exploratory Data Analysis

Description

Function that produces a histogram, density plot, boxplot, and Q-Q plot.

Usage

EDA(x, trim = 0.05)

Arguments

x numeric vector. NAs and Infs are allowed but will be removed.
trim fraction (between 0 and 0.5, inclusive) of values to be trimmed from each end of the ordered data. If trim = 0.5, the result is the median.

Details

Will not return command window information on data sets containing more than 5000 observations. It will however still produce graphical output for data sets containing more than 5000 observations.

Value

Function returns various measures of center and location. The values returned for the Quartiles are based on the definitions provided in BSDA. The boxplot is based on the Quartiles returned in the commands window.

Note

Requires package e1071.

Author(s)

Alan T. Arnholt

Examples

EDA(rnorm(100))
# Produces four graphs for the 100 randomly generated standard normal variates.
Educat

| Educat | Crime rates versus the percent of the population without a high school degree |

Description
Data for Exercise 2.41

Usage
Educat

Format
A data frame with 51 observations on the following 3 variables.

nodegree  a numeric vector
crime  a numeric vector

Source

Examples
str(Educat)
attach(Educat)
plot(nodegree,crime,xlab="No Crime",ylab="Violent Crime Rate per 100,000")
detach(Educat)

---

Eggs

| Eggs | Number of eggs versus amounts of feed supplement |

Description
Data for Exercise 9.22

Usage
Eggs
Elderly

Format
A data frame with 12 observations on the following 7 variables.

- feed  a numeric vector
- eggs  a numeric vector
- SRES1 a numeric vector
- FITS1 a numeric vector
- c1sq  a numeric vector
- SRES2 a numeric vector
- FITS2 a numeric vector

Source

Examples

```r
str(Eggs)
attach(Eggs)
plot(feed,eggs)
model <- lm(eggs~feed)
abline(model)
summary(model)
detach(Eggs)
remove(model)
```

---

Elderly  Percent of the population over the age of 65

Description
Data for Exercise 1.92 and 2.61

Usage

Elderly

Format
A data frame with 51 observations on the following 3 variables.

- X85percent a numeric vector
- X98percent a numeric vector
Energy

Amount of energy consumed by homes versus their sizes

Description

Data for Exercises 2.5, 2.24, and 2.55

Usage

Energy

Format

A data frame with 12 observations on the following 5 variables.

Size  a numeric vector
kilowatt  a numeric vector
SRES1  a numeric vector
FITS1  a numeric vector
Residuals  a numeric vector

Source


Examples

str(Elderly)
attach(Elderly)
stripchart(x=list(X98percent,X85percent),method="stack",pch=19,
col=c("red","blue"),group.names=c("1998","1985"))
cor(X98percent,X85percent)
detach(Elderly)

str(Energy)
attach(Energy)
plot(Size,kilowatt)
cor(Size,kilowatt)
model <- lm(kilowatt~Size)
plot(Size,resid(model))
detach(Energy)
### Engineer

**Salaries after 10 years for graduates of three different universities**

**Description**

Data for Example 10.7

**Usage**

Engineer

**Format**

A data frame with 51 observations on the following 6 variables.

- UnivA, a numeric vector
- UnivB, a numeric vector
- UnivC, a numeric vector
- salary, a numeric vector
- university, a numeric vector
- ranks, a numeric vector

**Source**


**Examples**

```r
str(Engineer)
attach(Engineer)
boxplot(salary~university)
kruskal.test(salary~as.factor(university))
detach(Engineer)
```

### Entrance

**College entrance exam scores for 24 high school seniors**

**Description**

Data for Example 1.8

**Usage**

Entrance

**Format**

A data frame with 24 observations on the following variable.

- score, a numeric vector
Epaminicompact

Source

Examples
str(Entrance)
attach(Entrance)
stem(score)
detach(Entrance)

---

**Epaminicompact**

*Fuel efficiency ratings for compact vehicles in 2001*

Description
Data for Exercise 1.65

Usage
Epaminicompact

Format
A data frame with 22 observations on the following 10 variables.

- **Class**: a factor with levels MINICOMPACT CARS
- **Manufacturer**: a factor with levels AUDI BMW JAGUAR MERCEDES-BENZ MITSUBISHI PORSCHE
- **carline.name**: a factor with levels 325CI CONVERTIBLE 330CI CONVERTIBLE 911 CARRERA 2/4 911 TURBO CLK320 (CABRIOLET) CLK430 (CABRIOLET) ECLIPSE SPYDER JAGUAR XK8 CONVERTIBLE JAGUAR XKR CONVERTIBLE M3 CONVERTIBLE TT COUPE TT COUPE QUATTRO
- **displ**: a numeric vector
- **cyl**: a numeric vector
- **trans**: a factor with levels Auto(L5) Auto(S4) Auto(S5) Manual(M5) Manual(M6)
- **drv**: a factor with levels 4 F R
- **cty**: a numeric vector
- **hwy**: a numeric vector
- **cmb**: a numeric vector

Source

Examples
str(Epaminicompact)
attach(Epaminicompact)
summary(cty)
detach(Epaminicompact)
Epatwoseater  

Fuel efficiency ratings for two-seater vehicles in 2001

**Description**

Data for Exercise 5.8

**Usage**

Epatwoseater

**Format**

A data frame with 36 observations on the following 10 variables.

- **Class**: a factor with levels `TWO SEATERS`
- **Manufacturer**: a factor with levels `ACURA AUDI BMW CHEVROLET DODGE FERRARI HONDA LAMBORGHINI MAZDA MERCEDES-BENZ PLYMOUTH PORSCHE TOYOTA`
- **carline.name**: a factor with levels `BOXSTER BOXSTER S CORVETTE DB132/144 DIABLO FERRARI 360 MODENA/SPIDER FERRARI 550 MARANELLO/BARCHETTA INSIGHT MR2 MX-5 MIATA NSX PROWLER S2000 SL500 SL600 SLK230 KOMPRESSOR SLK320 TT ROADSTER TT ROADSTER QUATTRO VIPER CONVERTIBLE VIPER COUPE Z3 COUPE Z3 ROADSTER Z8`
- **displ**: a numeric vector
- **cyl**: a numeric vector
- **trans**: a factor with levels `Auto(L4) Auto(L5) Auto(S4) Auto(S5) Auto(S6) Manual(M5) Manual(M6)`
- **drv**: a factor with levels `4 F R`
- **cty**: a numeric vector
- **hwy**: a numeric vector
- **cmb**: a numeric vector

**Source**


**Examples**

```
str(Epatwoseater)
attach(Epatwoseater)
boxplot(cty)
detach(Epatwoseater)
```
Executive

Description

Data for Exercise 1.104

Usage

Executive

Format

A data frame with 25 observations on the following variable.

Age  a numeric vector

Source


Examples

str(Executiv)
attach(Executiv)
EDA(Age)
detach(Executiv)

Exercise

Weight loss for 30 members of an exercise program

Description

Data for Exercise 1.44

Usage

Exercise

Format

A data frame with 30 observations on the following variable.

Loss  a numeric vector

Source

**Fabric**

Measures of softness of 10 different clothing garments washed with and without a softener

**Description**

Data for Exercise 7.21

**Usage**

Fabric

**Format**

A data frame with 10 observations on the following 3 variables.

- **Type** a numeric vector
- **With** a numeric vector
- **Without** a numeric vector

**Source**


**Examples**

```r
str(Fabric)
attach(Fabric)
DIF <- With - Without
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
wilcox.test(With,Without,paired=TRUE,alternative="greater")
detach(Fabric)
remove(DIF)
```
Faithful

Waiting times between successive eruptions of the Old Faithful geyser

Description

Data for Exercise 5.12 and 5.111

Usage

Faithful

Format

A data frame with 299 observations on the following 2 variables.

Time  a numeric vector
Eruption  a numeric vector

Source


Examples

str(Faithful)
attach(Faithful)
hist(Time,prob=TRUE,xlab="Waiting time between eruptions",col="tomato")
lines(density(Time),col="red",lwd=3)
t.test(Time)$conf
detach(Faithful)

Family

Size of family versus cost per person per week for groceries

Description

Data for Exercise 2.89

Usage

Family

Format

A data frame with 20 observations on the following 2 variables.

Number  a numeric vector
Cost  a numeric vector
Source


Examples

```r
str(Ferraro1)
attach(Ferraro1)
plot(Number,Cost)
cor(Number,Cost)
lm(Cost~Number)
detach(Ferraro1)
```

---

**Ferraro1**

*Choice of presidential ticket in 1984 by gender*

Description

Data for Exercise 8.23

Usage

```r
Ferraro1
```

Format

A data frame with 2 observations on the following 4 variables.

- `gender`: a factor with levels `Men` `Women`
- `Reag.Bs`: a numeric vector
- `Mond.Fer`: a numeric vector
- `undecide`: a numeric vector

Source


Examples

```r
str(Ferraro1)
attach(Ferraro1)
Ferraro1
chisq.test(Ferraro1[,2:4])
detach(Ferraro1)
```
Ferraro2

Choice of vice presidential candidate in 1984 by gender

Description

Data for Exercise 8.23

Usage

Ferraro2

Format

A data frame with 2 observations on the following 4 variables.

- gender: a factor with levels Men Women
- Bush: a numeric vector
- Ferraro: a numeric vector
- undecide: a numeric vector

Source


Examples

str(Ferraro2)
attach(Ferraro2)
Ferraro2
chisq.test(Ferraro2[,2:4])
detach(Ferraro2)

Fertility

Fertility rates of all 50 states and DC

Description

Data for Exercise 1.125

Usage

Fertility
Format

A data frame with 51 observations on the following 2 variables.


rate a numeric vector

Source


Examples

```r
str(Fertility)
attach(Fertility)
library(lattice)
dotplot(State~rate)
stem(rate)
fivenum(rate)
EDA(rate)
detach(Fertility)
```

---

Firstchi

**Ages of women at the birth of their first child**

Description

Data for Exercise 5.11

Usage

Firstchi

Format

A data frame with 87 observations on the following variable.

age a numeric vector

Source


Examples

```r
str(Firstchi)
attach(Firstchi)
EDA(age)
detach(Firstchi)
```
**Fish**

*Length and number of fish caught with small and large mesh codend*

**Description**

Data for Exercises 5.83, 5.119, and 7.29

**Usage**

Fish

**Format**

A data frame with 767 observations on the following 5 variables.

- length: a numeric vector
- smallmesh: a numeric vector
- largemesh: a numeric vector
- smallmesh: a numeric vector
- largemesh: a numeric vector

**Source**


**Examples**

```
str(Fish)
attach(Fish)
median(smallmesh, na.rm=TRUE)
median(largemesh)
IQR(smallmesh, na.rm=TRUE)
IQR(largemesh)
SIGN.test(smallmesh, conf.level=.99)
SIGN.test(largemesh, conf.level=.99)
t.test(smallmesh, largemesh)
detach(Fish)
```

---

**Fitness**

*Number of sit-ups before and after a physical fitness course*

**Description**

Data for Exercise 7.71

**Usage**

Fitness
Florida2000

Format
A data frame with 9 observations on the following 2 variables.

Before  a numeric vector
After   a numeric vector

Source

Examples

```r
str(Fitness)
attach(Fitness)
DIF <- After - Before
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(After, Before, paired=TRUE, alternative="greater")
detach(Fitness)
```

Florida2000  Florida voter results in the 2000 presidential election

Description
Data for Statistical Insight Chapter 2

Usage
Florida2000

Format
A data frame with 67 observations on the following 12 variables.

County  a factor with levels ALACHUA BAKER BAY BRADFORD BREVARD BROWARD CALHOUN CHARLOTTE CITRUS CLAY COLLIER COLUMBIA DADE DE SOTO DIXIE DUVAL ESCAMBIA FLAGLER FRANKLIN GADSDEN GILCHRIST GLADES GULF HAMILTON HARDEE HENDRY HERNANDO HIGHLANDS HILLSBOROUGH HOLMES INDIAN RIVER JACKSON JEFFERSON LAFAYETTE LAKE LEE LEON LEVY LIBERTY MADISON MANATEE MARION MARTIN MONROE NASSAU OKALOOSA OKEECHOBEE ORANGE OCEOLA PALM BEACH PASCO PINELLAS POLK PUTNAM SANTA ROSA SARASOTA SEMINOLE ST. JOHNS ST. LUCIE SUMTER SUWANNEE TAYLOR UNION VOLUSIA WAKULLA WALTON WASHINGTON

GORE  a numeric vector
BUSH  a numeric vector
BUCHANAN a numeric vector
NADER  a numeric vector
BROWNE a numeric vector
HAGELIN a numeric vector
HARRIS  a numeric vector
Fluid

MCREYNOLDS  a numeric vector
MOOREHEAD  a numeric vector
PHILLIPS  a numeric vector
Total  a numeric vector

Source


Examples

str(Florida2/zero.noslash/zero.noslash/zero.noslash)
attach(Florida2/zero.noslash/zero.noslash/zero.noslash)
plot(Total,BUCHANAN,xlab="Total votes cast (in thousands)",
ylab="Votes for Buchanan")
detach(Florida2/zero.noslash/zero.noslash/zero.noslash)

<table>
<thead>
<tr>
<th>Fluid</th>
<th>Breakdown times of an insulating fluid under various levels of voltage stress</th>
</tr>
</thead>
</table>

Description

Data for Exercise 5.76

Usage

Fluid

Format

A data frame with 76 observations on the following 10 variables.

X26kV  a numeric vector
X28kV  a numeric vector
X30kV  a numeric vector
X32kV  a numeric vector
X34kV  a numeric vector
X36kV  a numeric vector
X38kV  a numeric vector
response  a numeric vector
group  a numeric vector
ln.resp.  a numeric vector

Source

Food

Annual food expenditures for 40 single households in Ohio

Description
Data for Exercise 5.106

Usage
Food

Format
A data frame with 40 observations on the following variable.

food a numeric vector

Source

Examples
str(Food)
attach(Food)
EDA(food)
detach(Food)

Framingh

Cholesterol values of 62 subjects in the Framingham Heart Study

Description
Data for Exercises 1.55, 1.75, 3.69, and 5.60

Usage
Framingh

Format
A data frame with 62 observations on the following variable.

cholest a numeric vector
**Freshman**

Data for Exercise 6.53

**Ages of a random sample of 30 college freshmen**

**Description**

Data for Exercise 6.53

**Usage**

Freshman

**Format**

A data frame with 30 observations on the following variable.

- **age** a numeric vector

**Source**


**Examples**

```r
str(Freshman)
attach(Freshman)
SIGN.test(age, md=19)
detach(Freshman)
```
Funeral

Cost of funeral by region of country

Description

Data for Exercise 8.54

Usage

Funeral

Format

A data frame with 4 observations on the following 4 variables.

Region  a factor with levels Central East South West
Less  a numeric vector
Average  a numeric vector
More  a numeric vector

Source


Examples

str(Funeral)
attach(Funeral)
Funeral
chisq.test(Funeral[,2:4])
detach(Funeral)

Galaxie

Velocities of 82 galaxies in the Corona Borealis region

Description

Data for Example 5.2

Usage

Galaxie

Format

A data frame with 82 observations on the following variable.

velocity  a numeric vector
Gallup

Source

Examples
```
str(Galaxie)
attach(Galaxie)
EDA(velocity)
detach(Galaxie)
```

Gallup

Results of a Gallup poll on possession of marijuana as a criminal offense conducted in 1980

Description
Data for Exercise 2.76

Usage
Gallup

Format
A data frame with 4 observations on the following 16 variables.

Gender a factor with levels Female Male
Crime1 a numeric vector
No.Crime1 a numeric vector
No.Opinion1 a numeric vector
Education a factor with levels College Grade School High School
Crime2 a numeric vector
No.Crime2 a numeric vector
No.Opinion2 a numeric vector
Age a factor with levels 18-24 25-29 30-49 50-older
Crime3 a numeric vector
No.Crime3 a numeric vector
No.Opinion3 a numeric vector
Religion a factor with levels Catholic Protestant
Crime4 a numeric vector
No.Crime4 a numeric vector
No.Opinion4 a numeric vector

Source
Gasoline

Price of regular unleaded gasoline obtained from 25 service stations

Description

Data for Exercise 1.45

Usage

Gasoline

Format

A data frame with 25 observations on the following variable.

price  a numeric vector

Source


Examples

str(Gasoline)
attach(Gasoline)
stem(price)
detach(Gasoline)
### German

Number of errors in copying a German passage before and after an experimental course in German

#### Description
Data for Exercise 7.60

#### Usage
German

#### Format
A data frame with 10 observations on the following 4 variables.

- **Before** a numeric vector
- **After** a numeric vector
- **differ** a numeric vector
- **sgnrnks** a numeric vector

#### Source

#### Examples
```
str(German)
attach(German)
qqnorm(differ)
qline(differ)
shapiro.test(differ)
wilcox.test(Before,After,paired=TRUE)
detach(German)
```

### Golf

Distances a golf ball can be driven by 20 professional golfers

#### Description
Data for Exercise 5.24

#### Usage
Golf

#### Format
A data frame with 20 observations on the following variable.

- **yards** a numeric vector
Governor

Source

Examples

```
str(Golf)
attach(Golf)
stem(yards)
EDA(yards)
detach(Golf)
```

---

Governor

<table>
<thead>
<tr>
<th>State</th>
<th>Annual salaries for state governors in 1994</th>
</tr>
</thead>
</table>

Description
Data for Exercise 5.112

Usage
Governor

Format
A data frame with 50 observations on the following 3 variables.

- **X1994salary**: a numeric vector
- **X1999salary**: a numeric vector

Source

Examples

```
str(Governor)
attach(Governor)
EDA(X1994salary)
detach(Governor)
```
**Gpa**

*High school GPA versus college GPA*

**Description**

Data for Exercise 2.13

**Usage**

Gpa

**Format**

A data frame with 10 observations on the following 2 variables.

- HSGPA  a numeric vector
- CollGPA a numeric vector

**Source**


**Examples**

```r
str(Gpa)
attach(Gpa)
plot(HSGPA,CollGPA)
model <- lm(CollGPA~HSGPA)
abline(model)
model
r <- resid(model)
yhat <- fitted(model)
Table2.1 <- cbind(HSGPA,CollGPA,yhat,r)
Table2.1
remove(r,yhat,model,Table2.1)
detach(Gpa)
```

---

**Grades**

*Test grades in a beginning statistics class*

**Description**

Data for Exercise 1.120

**Usage**

Grades
Format
A data frame with 29 observations on the following variable.

grades  a numeric vector

Source

Examples
```r
str(Grades)
attach(Grades)
EDA(grades)
detach(Grades)
```

Graduate  
Graduation rates for student athletes in the Southeastern Conf.

Description
Data for Exercise 1.118

Usage
Graduate

Format
A data frame with 12 observations on the following 3 variables.

School  a factor with levels Alabama Arkansas Auburn Florida Georgia Kentucky Louisiana St Mississippi Mississippi St South Carolina Tennessee Vanderbilt

Code  a factor with levels Al Ar Au Fl Ge Ke LSt Mi MSt SC Te Va

Percent  a numeric vector

Source

Examples
```r
str(Graduate)
attach(Graduate)
names(Percent) <- School
barplot(Percent,las=2,cex.names=.65,col="tomato")
detach(Graduate)
```
Greenriv

Varve thickness from a sequence through an Eocene lake deposit in the Rocky Mountains

Description

Data for Exercise 6.57

Usage

Greenriv

Format

A data frame with 37 observations on the following variable.

thick  a numeric vector

Source


Examples

str(Greenriv)
attach(Greenriv)
EDA(thick)
SIGN.test(thick,md=7.3,alternative="greater")
detach(Greenriv)

Grnriv2

Thickness of a varved section of the Green river oil shale deposit near a major lake in the Rocky Mountains

Description

Data for Exercises 6.45 and 6.98

Usage

Grnriv2

Format

A data frame with 101 observations on the following variable.

thick  a numeric vector

Source

Examples

```r
str(Grnriv2)
attach(Grnriv2)
EDA(thick)
t.test(thick,mu=8,alternative="less")
SIGN.test(thick,md=8,alternative="less")
detach(Grnriv2)
```

Description

Data for Exercise 10.42

Usage

`Groupabc`

Format

A data frame with 15 observations on the following 3 variables.

- **GroupA**: a numeric vector
- **GroupB**: a numeric vector
- **GroupC**: a numeric vector

Source


Examples

```r
str(Groupabc)
attach(Groupabc)
STACKED <- stack(Groupabc)
STACKED[1:5,]
boxplot(values~ind,col=c("red","blue","green"),data=STACKED)
anova(lm(values~ind,data=STACKED))
remove(STACKED)
detach(Groupabc)
```
Groups

An illustration of analysis of variance

Description
Data for Exercise 10.4

Usage
Groups

Format
A data frame with 26 observations on the following 3 variables.

GroupA  a numeric vector
GroupB  a numeric vector
GroupC  a numeric vector

Source

Examples

```r
str(Groups)
attach(Groups)
STACKED <- stack(Groups)
STACKED[1:5,]
boxplot(values~ind,col=c("red","blue","green"),data=STACKED)
anova(lm(values~ind,data=STACKED))
remove(STACKED)
detach(Groups)
```

Gym
Children’s age versus number of completed gymnastic activities

Description
Data for Exercises 2.21, 9.14, and 9.32

Usage
Gym

Format
A data frame with 8 observations on the following 3 variables.

age  a numeric vector
number  a numeric vector
x.  a numeric vector
Source


Examples

```r
str(Gym)
attach(Gym)
plot(age,number)
model <- lm(number~age)
abline(model)
cor(age,number)
detach(Gym)
```

---

Habits

*Study habits of students in two matched school districts*

Description

Data for Exercise 7.57

Usage

Habits

Format

A data frame with 11 observations on the following 4 variables.

- **A**: a numeric vector
- **B**: a numeric vector
- **differ**: a numeric vector
- **signrks**: a numeric vector

Source


Examples

```r
str(Habits)
attach(Habits)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
t.test(B,A,paired=TRUE,alternative="less")
wilcox.test(B,A,paired=TRUE,alternative="less")
detach(Habits)
```
Haptologo

Haptoglobin concentration in blood serum of 8 healthy adults

Description
Data for Example 6.9

Usage
Haptologo

Format
A data frame with 8 observations on the following variable.

concent a numeric vector

Source

Examples
str(Haptologo)
attach(Haptologo)
qqnorm(concent,col="blue")
qqline(concent,col="red")
shapiro.test(concent)
t.test(concent,mu=2,alternative="less")
detach(Haptologo)

Hardware

Daily receipts for a small hardware store for 31 working days

Description
Data for Example 2.18

Usage
Hardware

Format
A data frame with 31 observations on the following variable.

receipt a numeric vector

Source
**Examples**

```r
data(Hardwood)
```

**Description**

Data for Exercise 9.33

**Usage**

Hardwood

**Format**

A data frame with 19 observations on the following variable.

```r
str(Hardwood)
```

**Source**


**Examples**

```r
str(Hardwood)
attach(Hardwood)
```

**Description**

Data for Exercise 1.29

**Usage**

Heat
Heating

Format
A data frame with 6 observations on the following 4 variables.

Fuel a factor with levels Electricity Fuel Oil LP bottled gas Other Utility gas Wood
Reserv a numeric vector
All.US a numeric vector
Not.Rese a numeric vector

Source

Examples
str(Heat)
attach(Heat)
MAT <- cbind(Reserv, All.US, Not.Rese)
row.names(MAT) <- c("Utility Gas","LP bottled Gas","Electricity","Fuel Oil","Wood","Other Fuel")
MAT
barplot(t(MAT),beside=TRUE,legend=TRUE,main="Heating of American Indian Homes")
sum(Reserv)
sum(All.US)
sum(Not.Rese)
detach(Heat)

Heating Fuel efficiency ratings for three types of oil heaters

Description
Data for Exercise 10.32

Usage
Heating

Format
A data frame with 30 observations on the following 6 variables.

TypeA a numeric vector
TypeB a numeric vector
TypeC a numeric vector
Rating a numeric vector
Type a numeric vector
Ranks a numeric vector

Source
Hodgkin

Results of treatments for Hodgkin's disease

Description

Data for Exercise 2.77

Usage

Hodgkin

Format

A data frame with 4 observations on the following 4 variables.

- **Histological** a factor with levels LD LP MC NS
- **Positive** a numeric vector
- **Partial** a numeric vector
- **None** a numeric vector

Source


Examples

```
str(Hodgkin)
nattach(Hodgkin)
HOD <- as.matrix(Hodgkin[,2:4])
rownames(HOD) <- Histological
HOD
barplot(t(HOD),legend=TRUE,beside=TRUE)
detach(Hodgkin)
remove(HOD)
```
Homes

Median prices of single-family homes in 65 metropolitan statistical areas

Description

Data for Statistical Insight Chapter 5

Usage

Homes

Format

A data frame with 65 observations on the following 5 variables.

City a factor with levels Akron OH Albuquerque NM Anaheim CA Atlanta GA Baltimore MD Chicago, IL Cincinnati, OH Cleveland, OH Columbus, OH Corpus Christi, TX Dallas, TX Daytona Beach, FL Denver, CO Des Moines, IA Detroit, MI El Paso, TX Grand Rapids, MI Hartford, CT Honolulu, HI Houston, TX Indianapolis, IN Jacksonville, FL Kansas City, MO Knoxville, TN Las Vegas, NV Los Angeles, CA Louisville, KY Madison, WI Memphis, TN Miami, FL Milwaukee, WI Minneapolis, MN Mobile, AL Nashville, TN New Haven, CT New Orleans, LA New York, NY Oklahoma City, OK Omaha, NE Orlando, FL Philadelphia, PA Phoenix, AZ Pittsburgh, PA Portland, OR Providence, RI Sacramento, CA Salt Lake City, UT San Antonio, TX San Diego, CA San Francisco, CA Seattle, WA Spokane, WA St Louis, MO Syracuse, NY Tampa, FL Toledo, OH Tulsa, OK Washington, DC

X1994 a numeric vector

Region a factor with levels Midwest Northeast South West

X2000 a numeric vector
difference a numeric vector

Source


Examples

str(Homes)
attach(Homes)
EDA(X2000)
boxplot(X1994,X2000, names=c("1994","2000"), col=c("red","blue"), ylab="Cost")
boxplot(X2000~Region)
detach(Homes)
Homework

Number of hours per week spent on homework for private and public high school students

Description

Data for Exercise 7.78

Usage

Homework

Format

A data frame with 15 observations on the following 2 variables.

Private a numeric vector
Public a numeric vector

Source


Examples

```r
str(Homework)
attach(Homework)
boxplot(Private,Public)
t.test(Private,Public,conf.level=.98)
detach(Homework)
```

Honda

Miles per gallon for a Honda Civic on 35 different occasions

Description

Data for Statistical Insight Chapter 6

Usage

Honda

Format

A data frame with 35 observations on the following variable.

mileage a numeric vector

Source

Hostile

Examples

```r
str(Honda)
attach(Honda)
t.test(mileage,mu=4/zero.noslash,alternative="less")
detach(Honda)
```

---

<table>
<thead>
<tr>
<th>Hostile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hostility levels of high school students from rural, suburban, and urban areas</td>
</tr>
</tbody>
</table>

Description

Data for Example 10.6

Usage

Hostile

Format

A data frame with 45 observations on the following 6 variables.

- **Rural** a numeric vector
- **Suburban** a numeric vector
- **Urban** a numeric vector
- **HLT** a numeric vector
- **Type** a numeric vector
- **Ranks** a numeric vector

Source


Examples

```r
str(Hostile)
attach(Hostile)
boxplot(HLT~Type)
kruskal.test(HLT~as.factor(Type))
detach(Hostile)
```
Housing

Median home prices for 1984 and 1993 in 37 markets across the U.S.

**Description**

Data for Exercise 5.82

**Usage**

Housing

**Format**

A data frame with 37 observations on the following 3 variables.

- **City** a factor with levels Albany Anaheim Atlanta Baltimore Birmingham Boston Chicago Cincinnati Cleveland Columbus Dallas Denver Detroit Ft Lauderdale Houston Indianapolis Kansas City Los Angeles Louisville Memphis Miami Milwaukee Minneapolis Nashville New York Oklahoma City Philadelphia Providence Rochester Salt Lake City San Antonio San Diego San Francisco San Jose St Louis Tampa Washington

- **X1984** a numeric vector
- **X1993** a numeric vector

**Source**


**Examples**

```r
str(Housing) attach(Housing) stem(X1993) stem(X1984) par(mfrow=c(2,2)) stripchart(x=list(X1984,X1993),method="stack",pch=1,cex=1.2,col=c("orange","pink"),group.names=c("1984","1993"),title(main="Problem 5.82 \n We have not talked about this kind of graph before...")) hist(X1993,breaks="Scott",col="pink") hist(X1984,breaks="Scott",col="orange") plot(density(X1993),col="red",ylab="",main="",ylim=c(0,.00003)) lines(density(X1984),col="orange") par(mfrow=c(1,1)) boxplot(X1993,X1984,col=c("pink","orange"),names=c("1993","1984"),main="Problem 5.82") SIGN.test(X1984,conf.level=.98) SIGN.test(X1993,conf.level=.98) # 98% CI -> 63591.1 79622.56 and 85591.69 109915.4 # Placing on a common number line...
my.axis <- function(side, at, labels,...) {for(i in seq(along=at)) axis(side=side, at=at[i], labels=labels[i],...)}
plot(1,type="n",xlim=c(62000,110000),ylim=c(0,1),xlab="Median House Price",ylab="",yaxt="n",main="") title(main="98 Percent Confidence Intervals") my.axis(2,at=c(.25,.75),labels=c("1984","1993"),cex.axis=1.2,las=2)
```
Number of storms, hurricanes and El Nino effects from 1950 through 1995

Description

Data for Exercises 1.38, 10.19, and Example 1.6

Usage

Hurrican

Format

A data frame with 46 observations on the following 5 variables.

- year   a numeric vector
- storms a numeric vector
- hurrican a numeric vector
- ElNino a factor with levels cold neutral warm
- code   a numeric vector

Source


Examples

```
str(Hurrican)
attach(Hurrican)
barplot(table(hurrican),col="blue",main="Problem 1.38",
       xlab="Number of Hurricanes",ylab="Number of Seasons")
boxplot(storms~ElNino)
anova(lm(storms~ElNino))
detach(Hurrican)
```
Iceberg

Number of icebergs sighted each month south of Newfoundland and south of the Grand Banks in 1920

Description

Data for Exercise 2.46 and 2.60

Usage

Iceberg

Format

A data frame with 12 observations on the following 3 variables.

Month a factor with levels Apr Aug Dec Feb Jan Jul Jun Mar May Nov Oct Sep
Newfound a numeric vector
GrandBk a numeric vector

Source


Examples

str(Iceberg)
attach(Iceberg)
plot(GrandBk,Newfound)
abline(lm(Newfound~GrandBk))
detach(Iceberg)

Income

Percent change in personal income from 1st to 2nd quarter in 2000

Description

Data for Exercise 1.33

Usage

Income
Format

A data frame with 51 observations on the following 6 variables.


income a numeric vector
C3 a numeric vector
Class a numeric vector
freq a numeric vector
percent a numeric vector

Source


Examples

str(Income)
attach(Income)
CATS <- factor(cut(income, breaks=c(0, 1, 1.5, 2, max(income))))
table(CATS)
table(CATS)/length(income)
barplot(table(CATS), col="lightblue", main="Problem 1.33")
detach(Income)
remove(CATS)

Independent

Illustrates a comparison problem for long-tailed distributions

Description

Data for Exercise 7.41

Usage

Independent

Format

A data frame with 46 observations on the following 3 variables.

score a numeric vector
group a numeric vector
ranks a numeric vector
Source


Examples

```r
str(Independent)
attach(Independent)
boxplot(score~group)
wilcox.test(score~group)
detach(Independent)
```

---

Indian

*Educational attainment versus per capita income and poverty rate for American Indians living on reservations*

Description

Data for Exercise 2.95

Usage

Indian

Format

A data frame with 10 observations on the following 4 variables.

- **Reserv**: a factor with levels Blackfeet Fort Apache Gila River Hopi Navajo Papago Pine Ridge Rosebud San Carlos Zuni Pueblo
- **highsch**: a numeric vector
- **income**: a numeric vector
- **poverty**: a numeric vector

Source


Examples

```r
str(Indian)
attach(Indian)
par(mfrow=c(1,2))
plot(highsch,income,xlab="Percent High School Graduates", ylab="Per capita income")
plot(highsch,poverty,xlab="Percent High School Graduates", ylab="Poverty rate")
par(mfrow=c(1,1))
cor(cbind(highsch,income,poverty))
detach(Indian)
```
Indiapol

Average miles per hour for the winners of the Indianapolis 500 race

Description
Data for Exercise 1.128

Usage
Indiapol

Format
A data frame with 39 observations on the following 3 variables.
year a numeric vector
speed a numeric vector
yr.1960 a numeric vector

Source

Examples
str(Indiapol)
attach(Indiapol)
plot(year,speed,type="l")
detach(Indiapol)

Indy500
Qualifying miles per hour and number of previous starts for drivers in 79th Indianapolis 500 race

Description
Data for Exercises 7.11 and 7.36

Usage
Indy500

Format
A data frame with 33 observations on the following 4 variables.
driver a factor with levels andretti bachelart boesel brayton c.guerrero cheever fabi
fernandez ferran fittipaldi fox goodyear gordon gugelmin herta james johansson
jones lazier luyendyk matsuda matsushita pruett r.guerrero rahal ribeiro salazar
sharp sullivan tracy vasser villeneuve zampedri
qualif a numeric vector
starts a numeric vector
group a numeric vector
Inflatio

Source


Examples

    str(Indy500)
    attach(Indy500)
    stripchart(qualif~group, method="stack", pch=19, col=c("red","blue"))
    boxplot(qualif~group)
    t.test(qualif~group)
    detach(Indy500)

Inflatio  Private pay increase of salaried employees versus inflation rate

Description

Data for Exercises 2.12 and 2.29

Usage

    data(Inflatio)

Format

A data frame with 24 observations on the following 5 variables.

    year  a numeric vector
    pay   a numeric vector
    increase a numeric vector
    inflation a numeric vector
    C6.T  a factor with levels alow bmiddle high

Source


Examples

    str(Inflatio)
    attach(Inflatio)
    plot(inflation,increase)
    cor(inflation,increase,use="complete.obs")
    detach(Inflatio)
Inletoil

Inlet oil temperature through a valve

Description
Data for Exercises 5.91 and 6.48

Usage
Inletoil

Format
A data frame with 12 observations on the following variable.

temp  a numeric vector

Source

Examples
str(Inletoil)
attach(Inletoil)
t.test(temp)$conf
t.test(temp, mu=98, alternative="less")
detach(Inletoil)

Inmate

Type of drug offense by race

Description
Data for Statistical Insight Chapter 8

Usage
Inmate

Format
A data frame with 3 observations on the following 5 variables.

Race  a factor with levels black hispanic white
heroin  a numeric vector
crack  a numeric vector
cocaine  a numeric vector
marijuan  a numeric vector
Source


Examples

```r
str(Inmate)
attach(Inmate)
Inmate
chisq.test(Inmate[,2:5])
detach(Inmate)
```

---

Inspect

<table>
<thead>
<tr>
<th>Percent of vehicles passing inspection by type inspection station</th>
</tr>
</thead>
</table>

Description

Data for Exercise 5.89

Usage

Inspect

Format

A data frame with 6 observations on the following 4 variables.

- **Type**: a factor with levels auto inspection auto repair car care center gas station new car dealer tire store
- **less70**: a numeric vector
- **X70.85**: a numeric vector
- **great85**: a numeric vector

Source


Examples

```r
str(Inspect)
attach(Inspect)
Inspect
chisq.test(Inspect[,2:4])
detach(Inspect)
```
Insulate

Heat loss through a new insulating medium

Description
Data for Exercise 9.50

Usage
Insulate

Format
A data frame with 10 observations on the following 2 variables.
temp a numeric vector
loss a numeric vector

Source

Examples
str(Insulate)
attach(Insulate)
summary(lm(loss~temp))
detach(Insulate)

Iqgpa

GPA versus IQ for 12 individuals

Description
Data for Exercises 9.51 and 9.52

Usage
Iqgpa

Format
A data frame with 12 observations on the following 2 variables.
IQ a numeric vector
GPA a numeric vector

Source
Examples

```r
str(Iqgpa)
attach(Iqgpa)
plot(IQ,GPA)
model <- lm(GPA~IQ)
abline(model)
summary(model)
detach(Iqgpa)
remove(model)
```

Irises

R.A. Fishers famous data on sepal length of a species of Iris Setosa

Description

Data for Exercises 1.15 and 5.19

Usage

Irises

Format

A data frame with 150 observations on the following 14 variables.

- `sepalL1` a numeric vector
- `sepalW1` a numeric vector
- `petalL1` a numeric vector
- `petalW1` a numeric vector
- `sepalL2` a numeric vector
- `sepalW2` a numeric vector
- `petalL2` a numeric vector
- `petalW2` a numeric vector
- `sepalL3` a numeric vector
- `sepalW3` a numeric vector
- `petalL3` a numeric vector
- `petalW3` a numeric vector
- `sepalL` a numeric vector
- `sample` a numeric vector

Source


Examples

```r
str(Irises)
attach(Irises)
EDA(sepalL1)
t.test(sepalL1,conf.level=.99)$conf
detach(Irises)
```
**Jdpower**

*Number of problems reported per 100 cars in 1994 versus 1995s*

**Description**

Data for Exercise 2.14, 2.17, 2.31, 2.33, and 2.40

**Usage**

Jdpower

**Format**

A data frame with 29 observations on the following 3 variables.

- **Car** a factor with levels Acura BMW Buick Cadillac Chevrolet Dodge Eagle Ford Geo Honda Hyundai Infiniti Jaguar Lexus Lincoln Mazda Mercedes-Benz Mercury Mitsubishi Nissan Oldsmobile Plymouth Pontiac Saab Saturn Subaru Toyota Volkswagen Volvo
- **X1994** a numeric vector
- **X1995** a numeric vector

**Source**


**Examples**

```r
str(Jdpower)
athead(Jdpower)
plot(X1994,X1995)
model <- lm(X1995~X1994)
abline(model)
model
cor(X1995,X1994)
detach(Jdpower)
```

---

**Jobsat**

*Job satisfaction and stress level for 9 school teachers*

**Description**

Data for Exercise 9.60

**Usage**

Jobsat
Kidsmoke

Format
A data frame with 9 observations on the following 2 variables.

\- WSPT a numeric vector
\- satisfac a numeric vector

Source

Examples
```
str(Jobsat)
attach(Jobsat)
plot(WSPT,satisfac)
model <- lm(satisfac~WSPT)
abline(model)
summary(model)
detach(Jobsat)
remove(model)
```

---

Kidsmoke

*Smoking habits of boys and girls ages 12 to 18*

Description
Data for Exercise 4.85

Usage
Kidsmoke

Format
A data frame with 1000 observations on the following 2 variables.

\- gender a numeric vector
\- smoke a numeric vector

Source

Examples
```
str(Kidsmoke)
attach(Kidsmoke)
table(gender,smoke)
addmargins(table(gender,smoke))
addmargins(table(gender,smoke)/1000)
detach(Kidsmoke)
```
Kilowatt

Rates per kilowatt-hour for each of the 50 states and DC

Description

Data for Example 5.9

Usage

Kilowatt

Format

A data frame with 51 observations on the following 2 variables.

<table>
<thead>
<tr>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
</tr>
<tr>
<td>Alaska</td>
</tr>
<tr>
<td>Arizona</td>
</tr>
<tr>
<td>Arkansas</td>
</tr>
<tr>
<td>California</td>
</tr>
<tr>
<td>Colorado</td>
</tr>
<tr>
<td>Connecticut</td>
</tr>
<tr>
<td>Delaware</td>
</tr>
<tr>
<td>District of</td>
</tr>
<tr>
<td>Columbia</td>
</tr>
<tr>
<td>Florida</td>
</tr>
<tr>
<td>Georgia</td>
</tr>
<tr>
<td>Hawaii</td>
</tr>
<tr>
<td>Idaho</td>
</tr>
<tr>
<td>Illinois</td>
</tr>
<tr>
<td>Indiana</td>
</tr>
<tr>
<td>Iowa</td>
</tr>
<tr>
<td>Kansas</td>
</tr>
<tr>
<td>Kentucky</td>
</tr>
<tr>
<td>Louisiana</td>
</tr>
<tr>
<td>Maine</td>
</tr>
<tr>
<td>Maryland</td>
</tr>
<tr>
<td>Massachusetts</td>
</tr>
<tr>
<td>Michigan</td>
</tr>
<tr>
<td>Minnesota</td>
</tr>
<tr>
<td>Mississippi</td>
</tr>
<tr>
<td>Missouri</td>
</tr>
<tr>
<td>Montana</td>
</tr>
<tr>
<td>Nebraska</td>
</tr>
<tr>
<td>Nevada</td>
</tr>
<tr>
<td>New Hampshire</td>
</tr>
<tr>
<td>New Jersey</td>
</tr>
<tr>
<td>New Mexico</td>
</tr>
<tr>
<td>New York</td>
</tr>
<tr>
<td>North Carolina</td>
</tr>
<tr>
<td>North Dakota</td>
</tr>
<tr>
<td>Ohio</td>
</tr>
<tr>
<td>Oklahoma</td>
</tr>
<tr>
<td>Oregon</td>
</tr>
<tr>
<td>Pennsylvania</td>
</tr>
<tr>
<td>Rhode Island</td>
</tr>
<tr>
<td>South Carolina</td>
</tr>
<tr>
<td>South Dakota</td>
</tr>
<tr>
<td>Tennessee</td>
</tr>
<tr>
<td>Texas</td>
</tr>
<tr>
<td>Utah</td>
</tr>
<tr>
<td>Vermont</td>
</tr>
<tr>
<td>Virginia</td>
</tr>
<tr>
<td>Washington West</td>
</tr>
<tr>
<td>Virginia</td>
</tr>
<tr>
<td>Wisconsin</td>
</tr>
<tr>
<td>Wyoming</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>rate</th>
</tr>
</thead>
</table>

Source


Examples

str(Kilowatt)
attach(Kilowatt)
EDA(rate)
detach(Kilowatt)

Kinder

Reading scores for first grade children who attended kindergarten versus those who did not

Description

Data for Exercise 7.68

Usage

Kinder
Laminect

Format

A data frame with 8 observations on the following 3 variables.

Pair  a numeric vector  
Kinder  a numeric vector  
NoKinder  a numeric vector

Source


Examples

```r
str(Kinder)
attach(Kinder)
DIF <- Kinder - NoKinder
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(Kinder, NoKinder, paired=TRUE, alternative="greater")
detach(Kinder)
remove(DIF)
```

---

Laminect  
Median costs of laminectomies at hospitals across North Carolina in 1992

Description

Data for Exercise 10.18

Usage

Laminect

Format

A data frame with 46 observations on the following 5 variables.

cost  a numeric vector  
class  a numeric vector  
Rural  a numeric vector  
Regional  a numeric vector  
Metropol  a numeric vector

Source

Lead

Examples

\begin{verbatim}
str(Laminect)
attach(Laminect)
boxplot(cost~class)
anova(lm(cost~as.factor(class)))
detach(Laminect)
\end{verbatim}

<table>
<thead>
<tr>
<th>Lead</th>
<th>Lead levels in children’s blood whose parents worked in a battery factory</th>
</tr>
</thead>
</table>

Description

Data for Example 1.17

Usage

Lead

Format

A data frame with 33 observations on the following 3 variables.

- Pair  a numeric vector
- exposed  a numeric vector
- control  a numeric vector

Source


Examples

\begin{verbatim}
str(Lead)
attach(Lead)
boxplot(exposed,control, names=c("Exposed","Control"),col=c("red","blue"))
detach(Lead)
\end{verbatim}

<table>
<thead>
<tr>
<th>Leader</th>
<th>Leadership exam scores by age for employees on an industrial plant</th>
</tr>
</thead>
</table>

Description

Data for Exercise 7.31

Usage

Leader
Lethal

Format

A data frame with 34 observations on the following 2 variables.

under35  a numeric vector
over35  a numeric vector

Source


Examples

str(Leader)
attach(Leader)
boxplot(under35,over35, names=c("Under 35","Over 35"),col=c("green","brown"))
t.test(under35,over35)
detach(Leader)

Lethal

Survival time of mice injected with an experimental lethal drug

Description

Data for Example 6.12

Usage

Lethal

Format

A data frame with 30 observations on the following variable.

survival  a numeric vector

Source


Examples

str(Lethal)
attach(Lethal)
SIGN.test(survival,md=45,alternative="less")
detach(Lethal)
Life expectancy of men and women in U.S.

Description
Data for Exercise 1.31

Usage
Life

Format
A data frame with 8 observations on the following 3 variables.

year  a numeric vector
Men   a numeric vector
Women a numeric vector

Source

Examples
str(Life)
attach(Life)
plot(year,Men,type="l",ylim=c(min(Men,Women),max(Men,Women)),col="blue",
main="Life Expectancy versus Year",ylab="Age",xlab="Year")
lines(year,Women,col="red")
text(1955,65,"Men",col="blue")
text(1955,7/zero.noslash,"Women",col="red")
detach(Life)

Life span of electronic components used in a spacecraft versus heat

Description
Data for Exercise 2.4, 2.37, and 2.49

Usage
Lifespan
Ligntmonth

Format

A data frame with 6 observations on the following 4 variables.

heat  a numeric vector
life  a numeric vector
RESI1 a numeric vector
FITS1 a numeric vector

Source


Examples

str(Lifespan)
attach(Lifespan)
plot(heat,life)
model <- lm(life~heat)
model
resid(model)
sum((resid(model))^2)
anova(model)

detach(Lifespan)

Ligntmonth

Relationship between damage reports and deaths caused by lightning

Description

Data for Exercise 2.6

Usage

Ligntmonth

Format

A data frame with 12 observations on the following 4 variables.

deads  a numeric vector
injuries a numeric vector
damage  a numeric vector

Source

**Lodge**

**Examples**

```r
str(Lightmonth)
attach(Lightmonth)
plot(damage, deaths)
detach(Lightmonth)
```

---

**Lodge**

*Measured traffic at three prospective locations for a motor lodge*

**Description**

Data for Exercise 10.33

**Usage**

Lodge

**Format**

A data frame with 45 observations on the following 6 variables.

- **SiteA** a numeric vector
- **SiteB** a numeric vector
- **SiteC** a numeric vector
- **Traffic** a numeric vector
- **Site** a numeric vector
- **Ranks** a numeric vector

**Source**


**Examples**

```r
str(Lodge)
attach(Lodge)
boxplot(Traffic~Site)
anova(lm(Traffic~as.factor(Site)))
detach(Lodge)
```
Longtail  

*Long-tailed distributions to illustrate Kruskal Wallis test*

**Description**

Data for Exercise 10.45

**Usage**

Longtail

**Format**

A data frame with 60 observations on the following 6 variables.

- GroupA: a numeric vector
- GroupB: a numeric vector
- GroupC: a numeric vector
- score: a numeric vector
- Group: a numeric vector
- Ranks: a numeric vector

**Source**


**Examples**

```r
str(Longtail)
attach(Longtail)
boxplot(score~Group)
kruskal.test(score~as.factor(Group))
anova(lm(score~as.factor(Group)))
detach(Longtail)
```

Lowabil  

*Reading skills of 24 matched low ability students*

**Description**

Data for Example 7.18

**Usage**

Lowabil
**Magnesium**

**Format**

A data frame with 12 observations on the following 3 variables.

- **Pair**  a numeric vector
- **Experimt**  a numeric vector
- **Control**  a numeric vector

**Source**


**Examples**

```r
str(Lowabil)
attach(Lowabil)
DIF <- Experimt - Control
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(Experimt,Control,paired=TRUE)
detach(Lowabil)
remove(DIF)
```

**Description**

Data for Exercise 9.9

**Usage**

`Magnesium`

**Format**

A data frame with 20 observations on the following 4 variables.

- **distance**  a numeric vector
- **magnesiu**  a numeric vector
- **SRES1**  a numeric vector
- **FITS1**  a numeric vector

**Source**

Examples

str(Magnesiu)
attach(Magnesiu)
model <- lm(magnesiu~distance)
plot(distance,magnesiu)
abline(model)
summary(model)
detach(Magnesiu)
remove(model)

Malpract

Amounts awarded in 17 malpractice cases

Description

Data for Exercise 5.73

Usage

Malpract

Format

A data frame with 17 observations on the following variable.

award a numeric vector

Source


Examples

str(Malpract)
attach(Malpract)
SIGN.test(award,conf.level=.90)
detach(Malpract)

Manager

Advertised salaries offered general managers of major corporations in 1995

Description

Data for Exercise 5.81

Usage

Manager
**Marked**

**Format**

A data frame with 26 observations on the following variable.

- **salary** a numeric vector

**Source**


**Examples**

```r
str(Manager)
attach(Manager)
stem(salary)
SIGN.test(salary)
detach(Manager)
```

---

**Marked**

Percent of marked cars in 65 police departments in Florida

**Description**

Data for Exercise 6.100

**Usage**

Marked

**Format**

A data frame with 65 observations on the following variable.

- **percent** a numeric vector

**Source**


**Examples**

```r
str(Marked)
attach(Marked)
EDA(percent)
t.test(percent, mu=60, alternative="greater")
SIGN.test(percent, md=60, alternative="greater")
detach(Marked)
```
Mathcomp

Mathcomp

Standardized math test scores for 30 students

Description
Data for Exercise 1.69

Usage
MATH

Format
A data frame with 30 observations on the following variable.

math  a numeric vector

Source

Examples
str(MATH)
attach(MATH)
hist(math,col="pink")
CharlieZ <- (62-mean(math))/sd(math)
CharlieZ
detach(MATH)
remove(CharlieZ)

Mathcomp

Standardized math competency for a group of entering freshmen at a small community college

Description
Data for Exercise 5.26

Usage
Mathcomp

Format
A data frame with 31 observations on the following variable.

score  a numeric vector

Source
Mathpro

Examples

str(Mathcomp)
attach(Mathcomp)
stem(score)
EDA(score)
detach(Mathcomp)

Mathpro

Math proficiency and SAT scores by states

Description

Data for Exercise 9.24, Example 9.1, and Example 9.6

Usage

Mathpro

Format

A data frame with 51 observations on the following 10 variables.

state1 a factor with levels Conn D.C. Del Ga Hawaii Ind Maine Mass Md N. C. N. H. N. J. N. Y.
     Ore Pa R. I. S. C. Va Vt
Sat.M1 a numeric vector
Profic1 a numeric vector

state2 a factor with levels Ala Alaska Ariz Ark Calif Colo Fla Idaho Ill Iowa Kan Ky La
     Mich Minn Miss Mo Mont N. D. N. M. Neb Nev Ohio Okla S. D. Tenn Texas Utah W. V. Wash
     Wis Wyo
Sat.M2 a numeric vector
Profic2 a numeric vector

state a factor with levels Ala Alaska Ariz Ark Calif Colo Conn D. C. Del Fla Ga Hawaii Idaho
     Ill Ind Iowa Kan Ky La Maine Mass Md Mich Minn Miss Mo Mont N. C. N. D. N. H. N. J. N. M.
     N. Y. Neb Nev Ohio Okla Ore Pa R. I. S. C. S. D. Tenn Texas Utah Va Vt W. V. Wash Wis Wyo
Sat.M a numeric vector
Profic a numeric vector
Group a numeric vector

Source


Examples

str(Mathpro)
attach(Mathpro)
model <- lm(Sat.M1~Profic1)
plot(Profic1,Sat.M1)
abline(model)
model
detach(Mathpro)
remove(model)
Median

Error scores for four groups of experimental animals running a maze

Description
Data for Exercise 10.13

Usage
Median

Illustrates test of equality of medians with the Kruskal Wallis test

Description
Data for Exercise 10.52

Usage
Median

Illustrates test of equality of medians with the Kruskal Wallis test
Mental

Format

A data frame with 15 observations on the following 3 variables.

Sample1 a numeric vector
Sample2 a numeric vector
Sample3 a numeric vector

Source


Examples

str(Median)
attach(Median)
STACKED <- stack(Median)
STACKED[1:5,]
boxplot(values~ind,col=c("red","blue","green"),data=STACKED)
anova(lm(values~ind,data=STACKED))
kruskal.test(values~ind,data=STACKED)
remove(STACKED)
detach(Median)

Mental

Median mental ages of 16 girls

Description

Data for Exercise 6.52

Usage

Mental

Format

A data frame with 16 observations on the following variable.

age a numeric vector

Source


Examples

str(Mental)
attach(Mental)
SIGN.test(age,md=100)
detach(Mental)
Mercury

Concentration of mercury in 25 lake trout

Description

Data for Example 1.9

Usage

Mercury

Format

A data frame with 25 observations on the following variable.

mercury a numeric vector

Source


Examples

str(Mercury)
attach(Mercury)
stem(mercury)
detach(Mercury)

Metrent

Monthly rental costs in metro areas with 1 million or more persons

Description

Data for Exercise 5.117

Usage

data(Metrent)

Format

A data frame with 46 observations on the following variable.

rent a numeric vector

Source

Miller personality test scores for a group of college students applying for graduate school

Description

Data for Example 5.7

Usage

Miller

Format

A data frame with 25 observations on the following variable.

miller a numeric vector

Source


Examples

str(Miller)
attach(Miller)
stem(miller)
fivenum(miller)
boxplot(miller)
qqnorm(miller,col="blue")
qqline(miller,col="red")
detach(Miller)

Twenty scores on the Miller personality test

Description

Data for Exercise 1.41

Usage

Miller1
Moisture

Format
A data frame with 20 observations on the following variable.
miller a numeric vector

Source

Examples
str(Miller1)
attach(Miller1)
stem(miller)
stem(miller,scale=2)
detach(Miller1)

Moisture content and depth of core sample for marine muds in eastern Louisiana

Description
Data for Exercise 9.37

Usage
Moisture

Format
A data frame with 16 observations on the following 4 variables.
depth a numeric vector
moisture a numeric vector
lnmoist a numeric vector
depthsq a numeric vector

Source

Examples
str(Moisture)
attach(Moisture)
model <- lm(moisture~depth)
plot(depth,resid(model))
detach(Moisture)
remove(model)
Monoxide

Carbon monoxide emitted by smoke stacks of a manufacturer and a competitor

Description
Data for Exercise 7.45

Usage
Monoxide

Format
A data frame with 10 observations on the following 2 variables.

manufac a numeric vector
compet a numeric vector

Source

Examples
str(Monoxide)
attach(Monoxide)
t.test(manufac,compet)
wilcox.test(manufac,compet)
detach(Monoxide)

Movie
Moral attitude scale on 15 subjects before and after viewing a movie

Description
Data for Exercise 7.53

Usage
Movie

Format
A data frame with 12 observations on the following 3 variables.

Before a numeric vector
After a numeric vector
differ a numeric vector
**Source**


**Examples**

```r
str(Movie)
attach(Movie)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
t.test(After, Before, paired=TRUE, conf.level=.99)
wilcox.test(After, Before, paired=TRUE)
detach(Movie)
```

---

**Music**

*Improvement scores for identical twins taught music recognition by two techniques*

**Description**

Data for Exercise 7.59

**Usage**

Music

**Format**

A data frame with 12 observations on the following 3 variables.

- **Method1** a numeric vector
- **Method2** a numeric vector
- **differ** a numeric vector

**Source**


**Examples**

```r
str(Music)
attach(Music)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
t.test(Method1, Method2, paired=TRUE)
detach(Music)
```
Name

Estimated value of a brand name product and the company’s revenue

Description

Data for Exercises 2.28, 9.19, and Example 2.8

Usage

Name

Format

A data frame with 42 observations on the following 3 variables.

Brand a factor with levels Band-Aid Barbie Birds Eye Budweiser Camel Campbell Carlsberg Coca-Cola Colgate Del Monte Fisher-Price Gordon’s Green Giant Guinness Haagen-Dazs Heineken Hertz HGTV Hydro-Quik Indian Wells Jif Johnson’s Johnson Wax Kool Aid Labatt’s Libbys Lipton Listerine Lucozade Macintosh Markal Maxwell House Mennen Mint Mark Motor Oil Nabisco Nabs Nature’s Promise Old El Paso Old Navy Old Spice Olay Ortho Palmar’s Pain Relief Palmolive Pampers Perrier Pepsi Qantas Quaker Retsina Rice-A-Roni Rosemary’s Ritz Rolaids Royal Crown Sandies Sara Lee Scott Sheepish Slow’s Smith & Wesson Smirnoff Spear’s Speedmaster Speedy’s Sunkist Surfer’s Tampax Tastee-Treats Taylor’s Tazo Tea Tourism Times Traveling’ Can Terra-di-Gold Tena Tess Travel Toastea Top Ramen Troy Tuna Turner’s Turner’s Old Bay Udder’s Unilever Valenti Veet Vicks Vodka Vidal Sassoon Vidal Grant Viennetta Walters Wawa Willy’s Windex Windex Wrigley’s Yager’s Yale-Yale Yen Yenzer Yeti’s Zippo

value a numeric vector

revenue a numeric vector

Source


Examples

str(Name)
attach(Name)
plot(revenue,value)
model <- lm(value~revenue)
abline(model)
cor(value,revenue)
summary(model)
detach(Name)
remove(model)

Nascar

Efficiency of pit crews for three major NASCAR teams

Description

Data for Example 10.53

Usage

Nascar
Nervous

Format
A data frame with 36 observations on the following 6 variables.

- TeamA: a numeric vector
- TeamB: a numeric vector
- TeamC: a numeric vector
- Time: a numeric vector
- Team: a numeric vector
- Ranks: a numeric vector

Source

Examples

```r
str(Nascar)
attach(Nascar)
boxplot(Time~Team)
anova(lm(Time~as.factor(Team)))
detach(Nascar)
```

---

Nervous

Reaction effects of 4 drugs on 25 subjects with a nervous disorder

Description
Data for Exercise 10.3

Usage
Nervous

Format
A data frame with 25 observations on the following 2 variables.

- react: a numeric vector
- drug: a numeric vector

Source

Examples

```r
str(Nervous)
attach(Nervous)
boxplot(react~drug)
anova(lm(react~as.factor(drug)))
detach(Nervous)
```
### Newsstand

**Daily profits for 20 newsstands**

**Description**

Data for Exercise 1.43

**Usage**

Newsstand

**Format**

A data frame with 20 observations on the following variable.

- `profit` a numeric vector

**Source**


**Examples**

```r
str(Newsstand)
attach(Newsstand)
stem(profit)
stem(profit,scale=3)
detach(Newsstand)
```

### Nfldraf2

**Rating, time in 40-yard dash, and weight of top defensive linemen in the 1994 NFL draft**

**Description**

Data for Exercise 9.63

**Usage**

Nfldraf2

**Format**

A data frame with 47 observations on the following variable.

- `Rating.forty.weight` a factor with levels

<table>
<thead>
<tr>
<th>Rating</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0000000000e+000</td>
<td>4.9000000000e+000</td>
</tr>
<tr>
<td>5.2000000000e+000</td>
<td>5.0000000000e+000</td>
</tr>
<tr>
<td>5.3000000000e+000</td>
<td>4.7800000000e+000</td>
</tr>
<tr>
<td>5.4000000000e+000</td>
<td>4.8900000000e+000</td>
</tr>
<tr>
<td>5.4000000000e+000</td>
<td>5.0000000000e+000</td>
</tr>
</tbody>
</table>

```r
```
Rating, time in 40-yard dash, and weight of top offensive linemen in the 1994 NFL draft

Data for Exercises 9.10 and 9.16

A data frame with 29 observations on the following variable.

<table>
<thead>
<tr>
<th>Rating</th>
<th>forty</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.200000000e+00</td>
<td>5.230000000e+00</td>
<td>2.740000000e+00</td>
</tr>
<tr>
<td>5.300000000e+00</td>
<td>5.180000000e+00</td>
<td>2.500000000e+00</td>
</tr>
<tr>
<td>5.500000000e+00</td>
<td>5.090000000e+00</td>
<td>5.200000000e+00</td>
</tr>
<tr>
<td>5.500000000e+00</td>
<td>5.560000000e+00</td>
<td>5.500000000e+00</td>
</tr>
<tr>
<td>5.700000000e+00</td>
<td>5.290000000e+00</td>
<td>2.500000000e+00</td>
</tr>
<tr>
<td>5.900000000e+00</td>
<td>5.250000000e+00</td>
<td>2.900000000e+00</td>
</tr>
<tr>
<td>6.000000000e+00</td>
<td>5.270000000e+00</td>
<td>2.900000000e+00</td>
</tr>
<tr>
<td>6.100000000e+00</td>
<td>5.270000000e+00</td>
<td>2.900000000e+00</td>
</tr>
<tr>
<td>6.200000000e+00</td>
<td>5.230000000e+00</td>
<td>2.900000000e+00</td>
</tr>
<tr>
<td>6.300000000e+00</td>
<td>5.360000000e+00</td>
<td>3.110000000e+00</td>
</tr>
<tr>
<td>6.400000000e+00</td>
<td>5.260000000e+00</td>
<td>3.020000000e+00</td>
</tr>
</tbody>
</table>

Source


Examples

data(Nfldraft2)

Description

Data for Exercises 9.10 and 9.16

Usage

Nfldraft

Format

A data frame with 29 observations on the following variable.

Rating.forty.weight  a factor with levels 5.000000000e+000 5.300000000e+000 3.100000000e+002
Nicotine content versus sales for 8 major brands of cigarettes

<table>
<thead>
<tr>
<th>Nicotine</th>
<th>6.500000000e+00</th>
<th>5.180000000e+00</th>
<th>3.250000000e+00</th>
<th>7.000000000e+02</th>
<th>0.0000000000e+00</th>
<th>5.200000000e+00</th>
</tr>
</thead>
<tbody>
<tr>
<td>sales</td>
<td>7.000000000e+00</td>
<td>5.360000000e+00</td>
<td>3.170000000e+00</td>
<td>7.100000000e+02</td>
<td>0.0000000000e+00</td>
<td>5.060000000e+00</td>
</tr>
<tr>
<td></td>
<td>7.200000000e+00</td>
<td>5.200000000e+00</td>
<td>3.150000000e+00</td>
<td>7.600000000e+02</td>
<td>0.0000000000e+00</td>
<td>5.150000000e+00</td>
</tr>
</tbody>
</table>

Source


Examples

```r
str(Nfldraft)
attach(Nfldraft)
detach(Nfldraft)
```

---

### Description

Data for Exercise 9.21

### Usage

Nicotine

### Format

A data frame with 8 observations on the following 2 variables.

- nicotine a numeric vector
- sales a numeric vector

### Source


### Examples

```r
str(Nicotine)
attach(Nicotine)
model <- lm(sales~nicotine)
summary(model)
detach(Nicotine)
remove(model)
```
normarea  

Normal Area

Description

Function that computes and draws the area between two user specified values in a user specified normal distribution with a given mean and standard deviation.

Usage

\[
\text{normarea}(\text{lower} = -\text{Inf}, \text{upper} = \text{Inf}, \text{m}, \text{sig})
\]

Arguments

- **lower**: the lower value
- **upper**: the upper value
- **m**: the mean for the population
- **sig**: the standard deviation of the population

Author(s)

Alan T. Arnholt

Examples

\[
\text{normarea}(7/\text{zero.noslash}, 13/\text{zero.noslash}, 1/\text{zero.noslash}/\text{zero.noslash}, 15)
\]

# Finds \(P(7 < X < 13)\) given \(X < N(1,15)\).

nsize  

Required Sample Size

Description

Function to determine required sample size to be within a given margin of error.

Usage

\[
\text{nsize}(b, \text{sigma} = \text{NULL}, p = 0.5, \text{conf.level} = 0.95, \text{type} = "\text{mu}")
\]

Arguments

- **b**: the desired bound.
- **sigma**: population standard deviation. Not required if using type "\text{pi}".
- **p**: estimate for the population proportion of successes. Not required if using type "\text{mu}".
- **conf.level**: confidence level for the problem, restricted to lie between zero and one.
- **type**: character string, one of "\text{mu}" or "\text{pi}", or just the initial letter of each, indicating the appropriate parameter. Default value is "\text{mu}".
ntester

Details

Answer is based on a normal approximation when using type "pi".

Value

Returns required sample size.

Author(s)

Alan T. Arnholt

Examples

nsize(b=.03, p=708/1200, conf.level=.90, type="pi")
   # Returns the required sample size (n) to estimate the population
   # proportion of successes with a 0.9 confidence interval
   # so that the margin of error is no more than 0.03 when the
   # estimate of the population proportion of successes is 708/1200.
   # This is problem 5.38 on page 257 of Kitchen's BSDA.

nsize(b=.15, sigma=.31, conf.level=.90, type="mu")
   # Returns the required sample size (n) to estimate the population
   # mean with a 0.9 confidence interval so that the margin
   # of error is no more than 0.15. This is Example 5.17 on page
   # 261 of Kitchen's BSDA.

Description

Q-Q plots of randomly generated normal data of the same size as the tested data are generated and plotted on the perimeter of the graph while a Q-Q plot of the actual data is depicted in the center of the graph.

Usage

ntester(actual.data)

Arguments

actual.data a numeric vector. Missing and infinite values are allowed, but are ignored in the calculation. The length of actual.data must be less than 5000 after dropping nonfinite values.

Details

Q-Q plots of randomly generated normal data of the same size as the tested data are generated and plotted on the perimeter of the graph sheet while a Q-Q plot of the actual data is depicted in the center of the graph. The p-values are calculated from the Shapiro-Wilk W-statistic. Function will only work on numeric vectors containing less than or equal to 5000 observations.
Author(s)
Alan T. Arnholt

References

Examples
ntester(rexp(50,1))
  # Q-Q plot of random exponential data in center plot
  # surrounded by 8 Q-Q plots of randomly generated
  # standard normal data of size 50.

Orange
Price of oranges versus size of the harvest

Description
Data for Exercise 9.61

Usage
Orange

Format
A data frame with 6 observations on the following 2 variables.

  harvest a numeric vector
  price a numeric vector

Source

Examples
str(Orange)
attach(Orange)
summary(lm(price~harvest))
detach(Orange)
**Orioles**

*Salaries of members of the Baltimore Orioles baseball team*

**Description**

Data for Example 1.3

**Usage**

Orioles

**Format**

A data frame with 27 observations on the following 3 variables.

- **first.name** a factor with levels Albert, Arthur, B.J., Brady, Cal, Charles, d1-Delino, d1-Scott, Doug, Harold, Heathcliff, Jeff, Jesse, Juan, Lenny, Mike, Rich, Ricky, Scott, Sidney, Will, Willis
- **last.name** a factor with levels Amaral, Anderson, Baines, Belle, Bones, Bordick, Clark, Conine, Deshields, Erickson, Fetters, Garcia, Guzman, Johns, Johnson, Kamieniecki, Mussina, Orosco, O.fadez, Ponson, Reboulet, Rhodes, Ripken Jr., Slocumb, Surhoff, Timlin, Webster
- **X1999salary** a numeric vector

**Source**


**Examples**

```r
data(Orioles)
```

---

**Oxytocin**

*Arterial blood pressure of 11 subjects before and after receiving oxytocin*

**Description**

Data for Exercise 7.86

**Usage**

Oxytocin

**Format**

A data frame with 11 observations on the following 3 variables.

- **Subject** a numeric vector
- **Before** a numeric vector
- **After** a numeric vector
Source


Examples

```r
str(Oxytocin)
attach(Oxytocin)
DIF <- Before - After
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(Before,After,paired=TRUE)
detach(Oxytocin)
```

<table>
<thead>
<tr>
<th>Parented</th>
<th>Education backgrounds of parents of entering freshmen at a state university</th>
</tr>
</thead>
</table>

Description

Data for Exercise 1.32

Usage

Parented

Format

A data frame with 6 observations on the following 3 variables.

- **Educat**: a factor with levels `4yr college degree`, `Doctoral degree`, `Grad degree`, `H.S grad or less`, `Some college`, `Some grad school`
- **Mother**: a numeric vector
- **Father**: a numeric vector

Source


Examples

```r
str(Parented)
attach(Parented)
MAT <- cbind(Mother, Father)
row.names(MAT) <- Educat
MAT
barplot(t(MAT), beside=TRUE, legend=TRUE, col=c("blue", "red"))
detach(Parented)
remove(MAT)
```
Patrol

Years of experience and number of tickets given by patrolpersons in New York City

Description
Data for Example 9.3

Usage
Patrol

Format
A data frame with 10 observations on the following 7 variables.
tickets a numeric vector
years a numeric vector
ln.tickets. a numeric vector
SRES1 a numeric vector
FITS1 a numeric vector
SRES2 a numeric vector
FITS2 a numeric vector

Source

Examples
str(Patrol)
attach(Patrol)
model <- lm(tickets~years)
summary(model)
detach(Patrol)
remove(model)

Pearson

Karl Pearson’s data on heights of brothers and sisters

Description
Data for Exercise 2.20

Usage
Pearson
**Phone**

**Format**

A data frame with 11 observations on the following 2 variables.

- **brother** a numeric vector
- **sister** a numeric vector

**Source**


**Examples**

```r
str(Pearson)
attach(Pearson)
plot(brother,sister)
cor(brother,sister)
detach(Pearson)
```

---

**Phone**

Length of long-distance phone calls for a small business firm

**Description**

Data for Exercise 6.95

**Usage**

Phone

**Format**

A data frame with 20 observations on the following variable.

- **time** a numeric vector

**Source**


**Examples**

```r
str(Phone)
attach(Phone)
qqnorm(time)
qqline(time)
shapiro.test(time)
SIGN.test(time,md=5,alternative="greater")
detach(Phone)
```
### Poison

**Number of poisonings reported to 16 poison control centers**

**Description**

Data for Exercise 1.113

**Usage**

Poison

**Format**

A data frame with 6 observations on the following 2 variables.

- **Type**: a factor with levels Alcohol, Cleaning Agent, Cosmetics, Drugs, Insecticides, Plants
- **number**: a numeric vector

**Source**


**Examples**

```r
str(Poison)  
attach(Poison)  
names(number) <- Type  
barplot(number,col="red")
```

### Politic

**Political party and gender in a voting district**

**Description**

Data for Example 8.3

**Usage**

Politic

**Format**

A data frame with 250 observations on the following 2 variables.

- **Party**: a numeric vector
- **Gender**: a numeric vector

**Source**

Examples

```r
str(Politic)
attach(Politic)
table(Party, Gender)
chisq.test(table(Party, Gender))
detach(Politic)
```

Pollutio

Air pollution index for 15 randomly selected days for a major western city

Description

Data for Exercise 5.59

Usage

Pollutio

Format

A data frame with 15 observations on the following variable.

- **inde**: numeric vector

Source


Examples

```r
str(Pollutio)
attach(Pollutio)
EDA(inde)
t.test(inde, conf.level = .98)$conf
detach(Pollutio)
```

Porosity

Porosity measurements on 20 samples of Tensleep Sandstone, Pennsylvanian from Bighorn Basin in Wyoming

Description

Data for Exercise 5.86

Usage

Porosity
**Poverty**

**Format**

A data frame with 20 observations on the following variable.

- porosity a numeric vector

**Source**


**Examples**

```
str(Porosity)
attach(Porosity)
stem(porosity)
fivenum(porosity)
boxplot(porosity)
detach(Porosity)
```

---

**Description**

Data for Exercise 9.11 and 9.17

**Usage**

Poverty

**Format**

A data frame with 20 observations on the following 6 variables.

- City a factor with levels Atlanta Buffalo Cincinnati Cleveland Dayton, O Detroit Flint, Mich Fresno, C Gary, Ind Hartford, C Laredo Macon, Ga Miami Milwaukee New Orleans Newark, NJ Rochester, NY Shreveport St. Louis Waco, Tx
- Poverty a numeric vector
- Crime a numeric vector
- cindex a numeric vector
- popu a numeric vector
- ratio a numeric vector

**Source**

Examples

\begin{verbatim}
str(Poverty)
attach(Poverty)
plot(Crime,Poverty)
model <- lm(Poverty~Crime)
abline(model)
summary(model)
detach(Poverty)
remove(model)
\end{verbatim}

<table>
<thead>
<tr>
<th>Precinct</th>
<th>Robbery rates versus percent low income in 8 precincts</th>
</tr>
</thead>
</table>

Description

Data for Exercise 2.2 and 2.38

Usage

Precinct

Format

A data frame with 8 observations on the following 2 variables.

rate  a numeric vector
income  a numeric vector

Source


Examples

\begin{verbatim}
str(Precinct)
attach(Precinct)
plot(rate,income,main="Exercise 2.2")
model <- lm(income~rate)
model
abline(model,col="green")
detach(Precinct)
\end{verbatim}
Prejudic

**Racial prejudice measured on a sample of 25 high school students**

**Description**
Data for Exercise 5.10 and 5.22

**Usage**
Prejudic

**Format**
A data frame with 25 observations on the following variable.
- **prejud** a numeric vector

**Source**

**Examples**
str(Prejudic)
attach(Prejudic)
EDA(prejud)
detach(Prejudic)

Presiden

**Ages at inauguration and death of U.S. presidents**

**Description**
Data for Exercise 1.126

**Usage**
Presiden

**Format**
A data frame with 43 observations on the following 5 variables.
- **Birth** a factor with levels ARK CAL CONN GA IA ILL KY MASS MO NC NEB NH NJ NY OH PA SC TEX VA VT
- **Inaugage** a numeric vector
- **Deathage** a numeric vector
Source


Examples

```
str(Presiden)
attach(Presiden)
table(Birt)
pie(table(Birt))
stripchart(x=list(Inaug_age,Deathage),method="stack",
group.names=c("Inaugural Age","Death Age"),col=c("green","brown"),pch=19)
detach(Presiden)
```

<table>
<thead>
<tr>
<th>Press</th>
<th>Degree of confidence in the press versus education level for 20 randomly selected persons</th>
</tr>
</thead>
</table>

Description

Data for Exercise 9.55

Usage

Press

Format

A data frame with 20 observations on the following 4 variables.

- educat: a numeric vector
- confid: a numeric vector
- SRES1: a numeric vector
- FITS1: a numeric vector

Source


Examples

```
str(Press)
attach(Press)
summary(lm(confid~educat))
detach(Press)
```
Prognost

Klopfers prognostic rating scale for subjects receiving behavior modification therapy

Description
Data for Exercise 6.61

Usage
Prognost

Format
A data frame with 15 observations on the following variable.

score a numeric vector

Source

Examples
str(Prognost)
attach(Prognost)
EDA(score)
t.test(score, mu=9)
detach(Prognost)

Program
Effects of four different methods of programmed learning for statistics students

Description
Data for Exercise 10.17

Usage
Program

Format
A data frame with 11 observations on the following 4 variables.

Method1 a numeric vector
Method2 a numeric vector
Method3 a numeric vector
Method4 a numeric vector
Source


Examples

```r
str(Program)
attach(Program)
STACKED <- stack(Program)
STACKED[1:5,]
boxplot(values~ind,col=c("red","blue","green","yellow"),data=STACKED)
anova(lm(values~ind,data=STACKED))
remove(STACKED)
detach(Program)
```

<table>
<thead>
<tr>
<th>Psat</th>
<th>PSAT scores versus SAT scores</th>
</tr>
</thead>
</table>

Description

Data for Exercise 2.50

Usage

Psat

Format

A data frame with 7 observations on the following 4 variables.

- psat a numeric vector
- sat a numeric vector
- SRES1 a numeric vector
- FITS1 a numeric vector

Source


Examples

```r
str(Psat)
attach(Psat)
model <- lm(sat~psat)
plot(psat,resid(model))
detach(Psat)
```
Psych

Correct responses for 24 students in a psychology experiment

Description
Data for Exercise 1.42

Usage
Psych

Format
A data frame with 23 observations on the following variable.

score  a numeric vector

Source

Examples
str(Psych)
attach(Psych)
stem(score)
EDA(score)
detach(Psych)

Puerto

Weekly incomes of a random sample of 50 Puerto Rican families in Miami

Description
Data for Exercise 5.22 and 5.65

Usage
Puerto

Format
A data frame with 50 observations on the following variable.

income  a numeric vector

Source
Quail

Plasma LDL levels in two groups of quail

Description
Data for Exercise 1.53, 1.77, 1.88, 5.66, and 7.50

Usage
Quail

Format
A data frame with 20 observations on the following 2 variables.

placebo  a numeric vector

treatmen  a numeric vector

Source

Examples

```r
str(Quail)
attach(Quail)
boxplot(placebo,treatmen,names=c("Placebo","Treatment"),
horizontal=TRUE,xlab="LDL level",col=c("lightblue","yellow"))
boxplot(placebo,treatmen,names=c("Placebo","Treatment"),
ylab="LDL level",col=c("lightblue","yellow"))
detach(Quail)
```

Quality

Quality control test scores on two manufacturing processes

Description
Data for Exercise 7.81

Usage
Quality
Rainks

Format
A data frame with 8 observations on the following 2 variables.

Process1 a numeric vector
Process2 a numeric vector

Source

Examples

```r
str(Quality)
attach(Quality)
qqnorm(Process1)
qqline(Process1)
shapiro.test(Process1)
qqnorm(Process2)
qqline(Process2)
shapiro.test(Process2)
t.test(Process1,Process2)
detach(Quality)
```

---

Rainfalls

Rainfall in an area of west central Kansas and four surrounding counties

Description
Data for Exercise 9.8

Usage
Rainfalls

Format
A data frame with 35 observations on the following 5 variables.

rain a numeric vector
x1 a numeric vector
x2 a numeric vector
x3 a numeric vector
x4 a numeric vector

Source
Examples

```r
str(Rainks)  
attach(Rainks)  
cor(Rainks)  
lm(rain~x2)  
detach(Rainks)
```

---

**Randd**  
Research and development expenditures and sales of a large company

**Description**

Data for Exercise 9.36 and Example 9.8

**Usage**

```r
Randd
```

**Format**

A data frame with 12 observations on the following 5 variables.

- `rd` a numeric vector
- `sales` a numeric vector
- `SRES1` a numeric vector
- `FITS1` a numeric vector
- `RESI1` a numeric vector

**Source**


**Examples**

```r
str(Randd)
attach(Randd)
plot(rd,sales)
model <- lm(sales~rd)
abline(model)
summary(model)
# plot(model)
detach(Randd)
remove(model)
```
Rat

Survival times of 20 rats exposed to high levels of radiation

Description
Data for Exercise 1.52, 1.76, 5.62, and 6.44

Usage
Rat

Format
A data frame with 20 observations on the following variable.

survival.time a numeric vector

Source

Examples
str(Rat)
attach(Rat)
hist(survival.time)
qqnorm(survival.time,col="blue")
qqline(survival.time,col="red")
t.test(survival.time)$conf
t.test(survival.time,mu=100,alternative="greater")
detach(Rat)

Ratings
Grade point averages versus teacher’s ratings

Description
Data for Example 2.6

Usage
Ratings
Reaction

Format

A data frame with 250 observations on the following 7 variables.

F  a numeric vector
D  a numeric vector
C  a numeric vector
B  a numeric vector
A  a numeric vector
Rating  a factor with levels A B C D F
GPA  a numeric vector

Source


Examples

```r
str(Ratings)
attach(Ratings)
boxplot(GPA~Rating,xlab="Teacher's Rating",ylab="GPA",main="Example 2.6",col="pink")
detach(Ratings)
```

---

Reaction

Threshold reaction time for persons subjected to emotional stress

Description

Data for Example 6.11

Usage

Reaction

Format

A data frame with 12 observations on the following variable.

time  a numeric vector

Source


Examples

```r
str(Reaction)
attach(Reaction)
SIGN.test(time,md=15,alternative="less")
detach(Reaction)
```
Reading

Standardized reading scores for 30 fifth graders

Description
Data for Exercise 1.72 and 2.10

Usage
Reading

Format
A data frame with 30 observations on the following 4 variables.

- reading a numeric vector
- sorted a numeric vector
- trimmed a numeric vector
- winsoriz a numeric vector

Source

Examples

```r
str(Reading)
attach(Reading)
EDA(reading)
detach(Reading)
```

Readiq

Reading scores versus IQ scores

Description
Data for Exercises 2.10 and 2.53

Usage
Readiq

Format
A data frame with 14 observations on the following 2 variables.

- reading a numeric vector
- IQ a numeric vector
Referend

Source


Examples

```r
str(Readiq)
attach(Readiq)
plot(IQ, reading)
model <- lm(reading ~ IQ)
abline(model)
detach(Readiq)
```

Referend

<table>
<thead>
<tr>
<th>Opinion on referendum by view on freedom of the press</th>
</tr>
</thead>
</table>

Description

Data for Exercise 8.20

Usage

Referend

Format

A data frame with 3 observations on the following 4 variables.

- Response  a factor with levels A B C
- For      a numeric vector
- Against  a numeric vector
- undecide  a numeric vector

Source


Examples

```r
str(Referend)
attach(Referend)
chisq.test(Referend[,2:4])
detach(Referend)
```
Region

Pollution index taken in three regions of the country

Description

Data for Exercise 10.26

Usage

Region

Format

A data frame with 48 observations on the following 6 variables.

West a numeric vector
Central a numeric vector
East a numeric vector
Index a numeric vector
Region a numeric vector
Ranks a numeric vector

Source


Examples

str(Region)
attach(Region)
boxplot(Index~Region)
anova(lm(Index~as.factor(Region)))
detach(Region)

Register

Maintenance cost versus age of cash registers in a department store

Description

Data for Exercise 2.3, 2.39, and 2.54

Usage

Register
Rehab

Format

A data frame with 9 observations on the following 4 variables.

- **age**: a numeric vector
- **cost**: a numeric vector
- **SRES1**: a numeric vector
- **FITS1**: a numeric vector

Source


Examples

```r
str(Register)
attach(Register)
plot(age, cost, main = "Exercise 2.3")
model <- lm(cost ~ age)
abline(model)
plot(age, resid(model))
detach(Register)
```

Rehab

Rehabilitative potential of 20 prison inmates as judged by two psychiatrists

Description

Data for Exercise 7.61

Usage

Rehab

Format

A data frame with 20 observations on the following 3 variables.

- **Psych1**: a numeric vector
- **Psych2**: a numeric vector
- **differ**: a numeric vector

Source

Examples

```r
str(Remedial)
attach(Remedial)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
boxplot(Psych1,Psych2,col=c("pink","lightblue"))
t.test(Psych1,Psych2,paired=TRUE)
detach(Remedial)
```

Description

Data for Exercise 7.43

Usage

Remedial

Format

A data frame with 42 observations on the following 2 variables.

- female a numeric vector
- male a numeric vector

Source


Examples

```r
str(Remedial)
attach(Remedial)
boxplot(female,male,col=c("blue","red"))
wilcox.test(female,male,conf.int=TRUE)
t.test(female,male)
detach(Remedial)
```
Rentals

**Weekly rentals for 45 apartments**

**Description**

Data for Exercise 1.122

**Usage**

Rentals

**Format**

A data frame with 45 observations on the following variable.

- **rent** a numeric vector

**Source**


**Examples**

```r
str(Rentals)
attach(Rentals)
EDA(rent)
detach(Rentals)
```

---

Repair

**Recorded times for repairing 22 automobiles involved in wrecks**

**Description**

Data for Exercise 5.77

**Usage**

Repair

**Format**

A data frame with 22 observations on the following variable.

- **time** a numeric vector

**Source**

Examples

str(Repair)
attach(Repair)
stem(time)
SIGN.test(time,conf.level=.98)
detach(Repair)

---

**Retail**

Length of employment versus gross sales for 10 employees of a large retail store

**Description**

Data for Exercise 9.59

**Usage**

Retail

**Format**

A data frame with 10 observations on the following 2 variables.

- months: a numeric vector
- sales: a numeric vector

**Source**


**Examples**

str(Retail)
attach(Retail)
summary(lm(sales~months))
detach(Retail)

---

**Ronbrown1**

Oceanography data obtained at site 1 by scientist aboard the ship Ron Brown

**Description**

Data for Exercise 2.9

**Usage**

Ronbrown1
Ronbrown2

Format
A data frame with 75 observations on the following 12 variables.
depth  a numeric vector
downtemp1 a numeric vector
downtemp2 a numeric vector
downsalinity1 a numeric vector
downsalinity2 a numeric vector
downdensity a numeric vector
C7 a numeric vector
uptemp1 a numeric vector
uptemp2 a numeric vector
upsalinity1 a numeric vector
upsalinity2 a numeric vector
updensity a numeric vector

Source

Examples
str(Ronbrown2)
attach(Ronbrown2)
plot(depth,downtemp1)
detach(Ronbrown2)

Ronbrown2 Oceanography data obtained at site 2 by scientist aboard the ship Ron Brown

Description
Data for Exercise 2.56 and Example 2.4

Usage
Ronbrown2

Format
A data frame with 150 observations on the following 6 variables.
depth  a numeric vector
primarytemp a numeric vector
secondarytemp a numeric vector
primarysalinity a numeric vector
secondarysalinity a numeric vector
density a numeric vector


Source


Examples

str(Ronbrown2)
attach(Ronbrown2)
plot(depth,primarysalinity,xlab="Depth",ylab="Salinity",
main="Example 2.4",col="tomato")
detach(Ronbrown2)

Rural

Social adjustment scores for a rural group and a city group of children

Description

Data for Exercise 7.16

Usage

Rural

Format

A data frame with 33 observations on the following 4 variables.

Rural  a numeric vector
City   a numeric vector
score  a numeric vector
code   a numeric vector

Source


Examples

str(Rural)
attach(Rural)
wilcox.test(score~code)
wilcox.test(Rural,City)
detach(Rural)
### Salary

**Description**

Data for Exercise 3.66

**Usage**

Salary

**Format**

A data frame with 25 observations on the following variable.

- **salary** a numeric vector

**Source**


**Examples**

```r
data(Salary)
```

---

### Salinity

**Description**

Surface-water salinity measurements from Whitewater Bay, Florida

**Usage**

Salinity

**Format**

A data frame with 48 observations on the following variable.

- **salinity** a numeric vector

**Source**


**Examples**

```r
str(Salinity)
attach(Salinity)
EDA(salinity)
t.test(salinity,conf.level=.99)$conf
detach(Salinity)
```
SAT scores, percent taking exam and state funding per student by state for 1994, 1995 and 1999

Description

Data for Statistical Insight Chapter 9

Usage

data(Sat)

Format

A data frame with 51 observations on the following 16 variables.

state a factor with levels alabama alaska arizona arkansas california colorado connecticut delaware dist of columbia florida georgia hawaii idaho illinois indiana iowa kansas kentucky louisiana maine maryland massachusetts michigan minnesota mississippi missouri montana nebraska nevada new hampshire new jersey new mexico new york north carolina north dakota ohio oklahoma oregon pennsylvania rhode island south carolina south dakota tennessee texas utah vermont virginia washington west virginia wisconsin wyoming

verbal94 a numeric vector
math94 a numeric vector
total94 a numeric vector
percent94 a numeric vector
code94 a numeric vector
expend94 a numeric vector
verbal95 a numeric vector
math95 a numeric vector
total95 a numeric vector
verbal99 a numeric vector
math99 a numeric vector
total99 a numeric vector
percent99 a numeric vector
code99 a numeric vector
expend99 a numeric vector

Source


Examples

str(Sat)
attach(Sat)
pairs(Sat)
detach(Sat)
Scales

Description
Problem asset ration for savings and loan companies in California, New York, and Texas

Usage
Data for Exercise 10.34 and 10.49

Format
A data frame with 75 observations on the following 6 variables.

- calif: a numeric vector
- newyork: a numeric vector
- texas: a numeric vector
- PAR: a numeric vector
- state: a numeric vector
- ranks: a numeric vector

Source

Examples
str(Saving)
attach(Saving)
boxplot(PAR~state)
kruskal.test(PAR~as.factor(state))
detach(Saving)

Scales
Readings obtained from a 100 pound weight placed on four brands of bathroom scales

Description
Data for Exercise 1.89

Usage
Scales
Schizop2

Format
A data frame with 20 observations on the following 2 variables.

Brand  a factor with levels A B C D
reading  a numeric vector

Source

Examples
str(Scales)
attach(Scales)
boxplot(reading~Brand,ylab="Reading",xlab="Brand",main="Problem 1.89")
detach(Scales)

Schizop2

Exam scores for 17 patients to assess the learning ability of schizophrenics after taking a specified does of a tranquilizer

Description
Data for Exercise 6.99

Usage
Schizop2

Format
A data frame with 17 observations on the following variable.

score  a numeric vector

Source

Examples
str(Schizop2)
attach(Schizop2)
EDA(score)
SIGN.test(score,md=22,alternative="greater")
detach(Schizop2)
**Schizoph**

*Standardized exam scores for 13 patients to investigate the learning ability of schizophrenics after a specified dose of a tranquilizer*

**Description**

Data for Example 6.10

**Usage**

Schizoph

**Format**

A data frame with 13 observations on the following variable.

- **score** a numeric vector

**Source**


**Examples**

```r
str(Schizoph)
attach(Schizoph)
EDA(score)
t.test(score,mu=2/zero.noslash)
detach(Schizoph)
```

---

**Seatbelt**

*Injury level versus seatbelt usage*

**Description**

Data for Exercise 8.24

**Usage**

Seatbelt

**Format**

A data frame with 2 observations on the following 5 variables.

- **seatbelt** a factor with levels no yes
- **None** a numeric vector
- **Minimal** a numeric vector
- **Minor** a numeric vector
- **Major** a numeric vector
Source

Examples

```
str(Seatbelt)
attach(Seatbelt)
Seatbelt
chisq.test(Seatbelt[,2:5])
detach(Seatbelt)
```

```
Selfdefe
```

Description
Data for Example 7.19

Usage

```
Selfdefe
```

Format
A data frame with 9 observations on the following 3 variables.

- **Woman**  a numeric vector
- **Before** a numeric vector
- **After**  a numeric vector

Source

Examples

```
str(Selfdefe)
attach(Selfdefe)
DIF <- After-Before
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(After,Before,paired=TRUE,alternative="greater")
detach(Selfdefe)
remove(DIF)
```
Senior

Reaction times of 30 senior citizens applying for drivers license renewals

Description
Data for Exercise 1.83 and 3.67

Usage
Senior

Format
A data frame with 31 observations on the following variable.

reaction a numeric vector

Source

Examples
str(Senior)
attach(Senior)
fivenum(reaction)
boxplot(reaction,horizontal=TRUE,main="Problem 1.83 Part d.",col="orange")
detach(Senior)

Sentence

Sentences of 41 prisoners convicted of a homicide offense

Description
Data for Exercise 1.123

Usage
Sentence

Format
A data frame with 41 observations on the following variable.

months a numeric vector

Source
Shkdrug

Examples

str(Sentence)
attach(Sentence)
stem(months)
EDA(months)
ll <- mean(months)-2*sd(months)
ul <- mean(months)+2*sd(months)
limits <- c(ll,ul)
limits
detach(Sentence)

Shkdrug

**Effects of a drug and electroshock therapy on the ability to solve simple tasks**

Description

Data for Exercises 10.11 and 10.12

Usage

Shkdrug

Format

A data frame with 64 observations on the following 6 variables.

Drug.Shk  a numeric vector
Drug.NoS  a numeric vector
NoDrug.S  a numeric vector
NoDg.NoS  a numeric vector
Treatment  a factor with levels Drug/NoS Drug/Shk NoDg/NoS NoDrug/S
Response  a numeric vector

Source


Examples

str(Shkdrug)
attach(Shkdrug)
boxplot(Response~Treatment)
anova(lm(Response~Treatment))
detach(Shkdrug)
**Shoptlift**

*Effect of experimental shock on time to complete difficult task*

### Description

Data for Exercise 10.50

### Usage

Shoptlift

### Format

A data frame with 9 observations on the following 3 variables.

- **Group1** a numeric vector
- **Group2** a numeric vector
- **Group3** a numeric vector

### Source


### Examples

```r
str(Shoptlift)
attach(Shoptlift)
STACKED <- stack(Shoptlift)
STACKED[1:5,]
boxplot(values~ind,col=c("red","blue","green"),data=STACKED)
anova(lm(values~ind,data=STACKED))
remove(STACKED)
detach(Shoptlift)
```

**Shoplift**

*Sales receipts versus shoplifting losses for a department store*

### Description

Data for Exercise 9.58

### Usage

Shoplift

### Format

A data frame with 8 observations on the following 2 variables.

- **sales** a numeric vector
- **loss** a numeric vector
**Source**


**Examples**

```r
str(Shoplift)
attach(Shoplift)
summary(lm(loss~sales))
detach(Shoplift)
```

**Short**

*James Short's measurements of the parallax of the sun*

**Description**

Data for Exercise 6.65

**Usage**

`Short`

**Format**

A data frame with 158 observations on the following 10 variables.

- `Sample.1` a numeric vector
- `Sample.2` a numeric vector
- `Sample.3` a numeric vector
- `Sample.4` a numeric vector
- `Sample.5` a numeric vector
- `Sample.6` a numeric vector
- `Sample.7` a numeric vector
- `Sample.8` a numeric vector
- `Parallax` a numeric vector
- `Sample` a numeric vector

**Source**


**Examples**

```r
str(Short)
attach(Short)
hist(Parallax)
EDA(Parallax)
SIGN.test(Parallax,md=8.798)
t.test(Parallax,mu=8.798)
detach(Short)
```
Shuttle

Number of people riding shuttle versus number of automobiles in the downtown area

Description

Data for Exercise 9.20

Usage

Shuttle

Format

A data frame with 15 observations on the following 2 variables.

shuttle  a numeric vector
autos    a numeric vector

Source


Examples

str(Shuttle)
attach(Shuttle)
model <- lm(autos~shuttle)
summary(model)
detach(Shuttle)
remove(model)

SIGN.test

Sign Test

Description

This function will test a hypothesis based on the sign test and reports linearly interpolated confidence intervals for one sample problems.

Usage

SIGN.test(x, y = NULL, md = 0, alternative = "two.sided", conf.level = 0.95)
SIGN.test

Arguments

x numeric vector; NAs and Infs are allowed but will be removed.
y optional numeric vector; NAs and Infs are allowed but will be removed.
md a single number representing the value of the population median specified by the null hypothesis
alternative is a character string, one of "greater", "less", or "two.sided", or the initial letter of each, indicating the specification of the alternative hypothesis. For one-sample tests, alternative refers to the true median of the parent population in relation to the hypothesized value of the median.
conf.level confidence level for the returned confidence interval, restricted to lie between zero and one

Details

Computes a “Dependent-samples Sign-Test” if both x and y are provided. If only x is provided, computes the “Sign-Test”.

Value

A list of class htest, containing the following components:

statistic the S-statistic (the number of positive differences between the data and the hypothesized median), with names attribute “S”.
p.value the p-value for the test
conf.int is a confidence interval (vector of length 2) for the true median based on linear interpolation. The confidence level is recorded in the attribute conf.level. When the alternative is not "two.sided", the confidence interval will be half-infinite, to reflect the interpretation of a confidence interval as the set of all values k for which one would not reject the null hypothesis that the true mean or difference in means is k. Here infinity will be represented by Inf.
estimate is a vector of length 1, giving the sample median; this estimates the corresponding population parameter. Component estimate has a names attribute describing its elements.
null.value is the value of the median specified by the null hypothesis. This equals the input argument md. Component null.value has a names attribute describing its elements.
alternative records the value of the input argument alternative: "greater", "less", or "two.sided"
data.name a character string (vector of length 1) containing the actual name of the input vector x

Null Hypothesis

For the one-sample sign-test, the null hypothesis is that the median of the population from which x is drawn is md. For the two-sample dependent case, the null hypothesis is that the median for the differences of the populations from which x and y are drawn is md. The alternative hypothesis indicates the direction of divergence of the population median for x from md (i.e., "greater", "less", "two.sided").
Assumptions
The median test assumes the parent population is continuous.

Confidence Interval
A linear interpolation is returned for the related confidence interval (returned component `conf.int`) which can be obtained by interpolating between the possible achieved confidence levels closest to the desired level. Note that, as explained under the description of `conf.int`, the confidence interval will be half-infinite when alternative is not "two.sided"; infinity will be represented by Inf.

Note
The reported confidence interval is based on linear interpolation. The lower and upper confidence levels are exact.

Author(s)
Alan T. Arnholt

References

See Also
`z.test`, `zsum.test`, `tsum.test`

Examples
```r
x <- c(7.8, 6.6, 6.5, 7.4, 7.3, 7., 6.4, 7.1, 6.7, 7.6, 6.8)
SIGN.test(x, md=6.5)
# Computes two-sided sign-test for the null hypothesis
# that the population median for 'x' is 6.5. The alternative
# hypothesis is that the median is not 6.5. An interpolated 95%
# confidence interval for the population median will be computed.

              14.4, 15.8, 11.3, 15.0)
SIGN.test(reaction, md=15, alternative="less")
# Data from Example 6.11 page 330 of Kitchens BSDA.
# Computes one-sided sign-test for the null hypothesis
# that the population median is 15. The alternative
# hypothesis is that the median is less than 15.
# An interpolated upper 95% upper bound for the population
# median will be computed.
```
Description
Data for Example 1.18

Usage
Simpson

Format
A data frame with 100 observations on the following 15 variables.
gpa a numeric vector
spor a numeric vector
gender a numeric vector
gpamale a numeric vector
sptmale a numeric vector
gpafemal a numeric vector
sptfemal a numeric vector
bbgpa a numeric vector
genderbb a numeric vector
sogpa a numeric vector
genderso a numeric vector
tkgpa a numeric vector
gendertk a numeric vector
gradept a numeric vector
gender2 a numeric vector

Source

Examples
str(Simpson)
attach(Simpson)
par(mfrow=c(1,2))
boxplot(gpa~gender,col=c("blue","pink"),names=c("Male","Female"),
main="GPA versus Gender",xlab="Gender",ylab="Grade Point Average")
boxplot(gradept~gender2,las=2,col=c("blue","pink"),
names=c("M-BBALL","F-BBALL","M-SOCC","F-SOCC","M-TRAC","F-TRAC"),
ylab="Grade Point Average",main="GPA vs Gender by Sports")
par(mfrow=c(1,1))
detach(Simpson)
### Situp

**Maximum number of situps by participants in an exercise class**

**Description**

Data for Exercise 1.47

**Usage**

Situp

**Format**

A data frame with 20 observations on the following variable.

- **number**: a numeric vector

**Source**


**Examples**

```r
str(Situp)
attach(Situp)
stem(number)
hist(number,breaks=seq(0,70,10))
hist(number,breaks=seq(0,70,10),right=FALSE,col="blue",prob=TRUE,
  main="Problems 1.46 & 1.47")
lines(density(number),col="red",lwd=3)
detach(Situp)
```

### Skewed

**Illustrates the Wilcoxon Rank Sum test**

**Description**

Data for Exercise 7.65

**Usage**

Skewed

**Format**

A data frame with 21 observations on the following 2 variables.

- **C1**: a numeric vector
- **C2**: a numeric vector
Skin

Source


Examples

```r
str(Skewed)
attach(Skewed)
boxplot(C1,C2)
wilcox.test(C1,C2)
detach(Skewed)
```

```
Skin
Survival times of closely and poorly matched skin grafts on burn patients
```

Description

Data for Exercise 5.20

Usage

Skin

Format

A data frame with 11 observations on the following 2 variables.

- close a numeric vector
- poor a numeric vector

Source


Examples

```r
str(Skin)
attach(Skin)
DIFF <- close-poor
stem(DIFF)
EDA(DIFF)
remove(DIFF)
detach(Skin)
```
S1c

*Sodium-lithium countertransport activity on 190 individuals from six large English kindred*

**Description**

Data for Exercise 5.116

**Usage**

S1c

**Format**

A data frame with 190 observations on the following variable.

SLC a numeric vector

**Source**


**Examples**

str(S1c)
attach(S1c)
EDA(SLC)
detach(S1c)

Smokyph

*Water pH levels of 75 water samples taken in the Great Smoky Mountains*

**Description**

Data for Exercises 6.40, 6.59, 7.10, and 7.35

**Usage**

Smokyph

**Format**

A data frame with 75 observations on the following 5 variables.

waterph a numeric vector
code a factor with levels high low
elev a numeric vector
SRES1 a numeric vector
FITS1 a numeric vector
Snoring versus heart disease

Description
Data for Exercise 8.21

Usage
Snore

Format
A data frame with 2 observations on the following 5 variables.

heart  a factor with levels no yes
Non   a numeric vector
occasion a numeric vector
nearly a numeric vector
every  a numeric vector

Source

Examples

\begin{verbatim}
str(Snore)
attach(Snore)
chisq.test(Snore[,2:5])
detach(Snore)
\end{verbatim}
### Snow

**Description**

Data for Exercise 7.87

**Usage**

Snow

**Format**

A data frame with 34 observations on the following 4 variables.

- `antarc` a numeric vector
- `greenld` a numeric vector
- `concent` a numeric vector
- `site` a numeric vector

**Source**


**Examples**

```r
str(Snow)
attach(Snow)
boxplot(concent~site)
detach(Snow)
```

### Soccer

**Description**

Weights of 25 soccer players

**Usage**

Soccer

**Format**

A data frame with 25 observations on the following variable.

- `weight` a numeric vector
Examples

\begin{verbatim}
str(Soccer)
attach(Soccer)
stem(weight, scale=2)
hist(weight, breaks=seq(110, 210, 10), col="orange",
main="Problem 1.46 \n Weights of Soccer Players", right=FALSE)
detach(Soccer)
\end{verbatim}

---

Social

Median income level for 25 social workers from North Carolina

Description

Data for Exercise 6.63

Usage

Social

Format

A data frame with 25 observations on the following variable.

income a numeric vector

Source


Examples

\begin{verbatim}
str(Social)
attach(Social)
SIGN.test(income, md=27500, alternative="less")
detach(Social)
\end{verbatim}
<table>
<thead>
<tr>
<th>Sophomor</th>
<th>Grade point averages, SAT scores and final grade in college algebra for 20 sophomores</th>
</tr>
</thead>
</table>

**Description**

Data for Exercise 2.42

**Usage**

Sophomor

**Format**

A data frame with 20 observations on the following 4 variables.

- Student  a numeric vector
- GPA       a numeric vector
- SAT       a numeric vector
- Exam      a numeric vector

**Source**


**Examples**

```r
str(Sophomor)
attach(Sophomor)
cor(Sophomor)
detach(Sophomor)
```

<table>
<thead>
<tr>
<th>South</th>
<th>Murder rates for 30 cities in the South</th>
</tr>
</thead>
</table>

**Description**

Data for Exercise 1.84

**Usage**

South

**Format**

A data frame with 31 observations on the following variable.

- rate  a numeric vector
Source


Examples

str(South)
attach(South)
boxplot(rate,col="yellow",main="Problem 1.83")
detach(South)

Speed reading scores before and after a course on speed reading

Description

Data for Exercise 7.58

Usage

Speed

Format

A data frame with 15 observations on the following 4 variables.

Before  a numeric vector
After   a numeric vector
differ  a numeric vector
signrnks a numeric vector

Source


Examples

str(Speed)
attach(Speed)
qqnorm(differ)
qline(differ)
shapiro.test(differ)
t.test(After,Before,paired=TRUE,alternative="greater")
wilcox.test(After,Before,paired=TRUE,alternative="greater")
detach(Speed)
Spellers

Standardized spelling test scores for two fourth grade classes

Description
Data for Exercise 7.82

Usage
Spellers

Format
A data frame with 10 observations on the following 2 variables.

Fourth  a numeric vector
Colleag  a numeric vector

Source

Examples

```r
str(Spellers)
attach(Spellers)
t.test(Fourth,Colleag,alternative="greater")
detach(Spellers)
```

Spelling
Spelling scores for 9 eighth graders before and after a 2-week course of instruction

Description
Data for Exercise 7.56

Usage
Spelling

Format
A data frame with 9 observations on the following 3 variables.

Before  a numeric vector
After   a numeric vector
differ  a numeric vector
Source


Examples

```r
str(Spelling)
attach(Spelling)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
t.test(After,Before,paired=TRUE,alternative="greater")
detach(Spelling)
```

---

### Sports

**Favorite sport by gender**

<table>
<thead>
<tr>
<th>Sports</th>
<th>Gender</th>
<th>Football</th>
<th>Basketball</th>
<th>Baseball</th>
<th>Tennis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>3.800000000e+01</td>
<td>2.100000000e+01</td>
<td>3.800000000e+01</td>
<td>2.400000000e+01</td>
<td>5.000000000e+01</td>
</tr>
<tr>
<td>Male</td>
<td>3.300000000e+01</td>
<td>3.800000000e+01</td>
<td>2.400000000e+01</td>
<td>5.000000000e+01</td>
<td></td>
</tr>
</tbody>
</table>

Source


Examples

```r
str(Sports)
attach(Sports)
Sports
detach(Sports)
```
Spouse

Convictions in spouse murder cases by gender

Description
Data for Exercise 8.33

Usage
Spouse

Format
A data frame with 4 observations on the following 3 variables.

result a factor with levels acquitted convicted not prosecuted pleaded guilty
husband a numeric vector
wife a numeric vector

Source

Examples
str(Spouse)
attach(Spouse)
Spouse
chisq.test(Spouse[,2:3])
detach(Spouse)

SRS

Simple Random Sampling

Description
Computes all possible samples from a given population using simple random sampling.

Usage
SRS(POPvalues, n)

Arguments
POPvalues vector containing the population values.

n the sample size.

Value
Returns a matrix containing the possible simple random samples of size n taken from a population POPvalues.
Stable

Author(s)
Alan T. Arnholt

See Also
Combinations

Examples

SRS(c(5,8,3),2)
# The rows in the matrix list the values for the 3 possible
# simple random samples of size 2 from the population of 5,8, and 3.

Description

Data for Exercise 6.93

Usage

Stable

Format

A data frame with 9 observations on the following variable.

time  a numeric vector

Source


Examples

str(Stable)
attach(Stable)
EDA(time)
SIGN.test(time,md=98.5,alternative="greater")
detach(Stable)
**Stamp**

*Thicknesses of 1872 Hidalgo stamps issued in Mexico*

**Description**

Data for Statistical Insight Chapter 1 and Exercise 5.110

**Usage**

`Stamp`

**Format**

A data frame with 485 observations on the following 3 variables.

- thickness: a numeric vector
- thick: a numeric vector
- freq: a numeric vector

**Source**


**Examples**

```r
str(Stamp)
attach(Stamp)
hist(thickness, prob=TRUE, col="lightblue")
lines(density(thickness), lwd=2, col="blue")
t.test(thickness, conf.level=.99)$conf
detach(Stamp)
```

---

**Statclas**

*Grades for two introductory statistics classes*

**Description**

Data for Exercise 7.30

**Usage**

`Statclas`

**Format**

A data frame with 36 observations on the following 2 variables.

- X9am: a numeric vector
- X2pm: a numeric vector
Statelaw

Source


Examples

str(Statclas)
attach(Statclas)
t.test(X9am,X2pm)
detach(Statclas)

Statelaw

Operating expenditures per resident for each of the state law enforcement agencies

Description

Data for Exercise 6.62

Usage

Statelaw

Format

A data frame with 50 observations on the following 2 variables.


cost a numeric vector

Source


Examples

str(Statelaw)
attach(Statelaw)
summary(cost)
EDA(cost)
SIGN.test(cost,md=8,alternative="less")
detach(Statelaw)
Statisti

Test scores for two beginning statistics classes

Description

Data for Exercises 1.70 and 1.87

Usage

Statisti

Format

A data frame with 31 observations on the following 2 variables.

Class1 a numeric vector
Class2 a numeric vector

Source


Examples

str(Statisti)
attach(Statisti)
boxplot(Class1,Class2, names=c("Class 1","Class 2"),col=c("red","blue"),
main="Problem 1.87")
detach(Statisti)

Step

STEP science test scores for a class of ability-grouped students

Description

Data for Exercise 6.79

Usage

Step

Format

A data frame with 12 observations on the following variable.

score a numeric vector

Source

**Examples**

```r
str(Step)
attach(Step)
EDA(score)
t.test(score,mu=80,alternative="less")
detach(Step)
```

---

**Stress**

*Short-term memory test scores on 12 subjects before and after a stressful situation*

**Description**

Data for Example 7.20

**Usage**

`Stress`

**Format**

A data frame with 12 observations on the following 2 variables.

- `Prestre` a numeric vector
- `Poststre` a numeric vector

**Source**


**Examples**

```r
str(Stress)
attach(Stress)
DIF <- Poststre - Prestre
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(Poststre,Prestre,paired=TRUE,alternative="less")
detach(Stress)
remove(DIF)
```
**Study**

*Number of hours studied per week by a sample of 50 freshmen*

**Description**

Data for Exercise 5.25

**Usage**

Study

**Format**

A data frame with 50 observations on the following variable.

- **hours**  a numeric vector

**Source**


**Examples**

```r
str(Study)
attach(Study)
stem(hours)
EDA(hours)
detach(Study)
```

**Submarin**

*Number of German submarines sunk by U.S. Navy in World War II*

**Description**

Data for Exercises 2.16, 2.45, and 2.59

**Usage**

Submarin

**Format**

A data frame with 16 observations on the following 3 variables.

- **Month**  a numeric vector
- **reported**  a numeric vector
- **actual**  a numeric vector

**Source**

Examples

\begin{verbatim}
str(Submarin)
attach(Submarin)
plot(reported, actual)
model <- lm(actual ~ reported)
abline(model)
anova(model)
summary(model)
detach(Submarin)
\end{verbatim}

Subway

Time it takes a subway to travel from the airport to downtown

Description

Data for Exercise 5.19

Usage

Subway

Format

A data frame with 30 observations on the following variable.

\begin{verbatim}
time  a numeric vector
\end{verbatim}

Source


Examples

\begin{verbatim}
str(Subway)
attach(Subway)
EDA(time)
detach(Subway)
\end{verbatim}

Sunspot

Wolfer sunspot numbers from 1700 through 2000

Description

Data for Example 1.7

Usage

Sunspot
Superbowl

Format

A data frame with 35 observations on the following 5 variables.


winner.score a numeric vector


loser.score a numeric vector

margin a numeric vector

Source

Supercar

Examples

str(Superbowl)
attach(Superbowl)
stem(margin)
detach(Superbowl)

Supercar

Top speeds attained by five makes of supercars

Description

Data for Statistical Insight Chapter 10

Usage

Supercar

Format

A data frame with 30 observations on the following 7 variables.

Acura a numeric vector
Ferrari a numeric vector
Lotus a numeric vector
Porsche a numeric vector
Viper a numeric vector
speed a numeric vector
car a numeric vector

Source


Examples

str(Supercar)
attach(Supercar)
boxplot(speed~car)
anova(lm(speed~as.factor(car)))
detach(Supercar)
Tablrock

Ozone concentrations at Mt. Mitchell, North Carolina

Description
Data for Exercise 5.63

Usage
Tablrock

Format
A data frame with 719 observations on the following 16 variables.

- **hour** a factor with levels 00:00 01:00 02:00 03:00 04:00 05:00 06:00 07:00 08:00 09:00 10:00 11:00 12:00 13:00 14:00 15:00 16:00 17:00 18:00 19:00 20:00 21:00 22:00 23:00
- **X03** a numeric vector
- **tmp** a numeric vector
- **vdc** a numeric vector
- **wd** a numeric vector
- **ws** a numeric vector
- **amb** a numeric vector
- **dew** a numeric vector
- **so2** a numeric vector
- **no** a numeric vector
- **no2** a numeric vector
- **nox** a numeric vector
- **co** a numeric vector
- **co2** a numeric vector
- **gas** a numeric vector
- **air** a numeric vector

Source

Examples
```r
str(Tablrock)
attach(Tablrock)
EDA(X03)
t.test(X03,conf.level=.99)$conf
detach(Tablrock)
```
**Teacher**

*Average teacher's salaries across the states in the 70s, 80s, and 90s*

**Description**

Data for Exercise 5.114

**Usage**

Teacher

**Format**

A data frame with 51 observations on the following 4 variables.

- **X1973.74**: a numeric vector
- **X1983.84**: a numeric vector
- **X1993.94**: a numeric vector

**Source**


**Examples**

```r
str(Teacher)
attach(Teacher)
boxplot(X1973.74,X1983.84,X1993.94,
detach(Teacher)
```

**Tenness**

*Tennessee self concept scores for 20 gifted high school students*

**Description**

Data for Exercise 6.56

**Usage**

Tenness
Tensile

**Format**

A data frame with 20 observations on the following variable.

- **score**: a numeric vector

**Source**


**Examples**

```r
str(Tenness)
attach(Tenness)
EDA(score)
t.test(score,mu=3/zero.noslash,alternative="less")
SIGN.test(score,md=3/zero.noslash,alternative="less")
detach(Tenness)
```

---

**Tensile**

*Tensile strength of plastic bags from two production runs*

**Description**

Data for Example 7.11

**Usage**

Tensile

**Format**

A data frame with 72 observations on the following 4 variables.

- **Run.1**: a numeric vector
- **Run.2**: a numeric vector
- **Tensile**: a numeric vector
- **Run**: a numeric vector

**Source**


**Examples**

```r
str(Tensile)
attach(Tensile)
boxplot(Run.1,Run.2, names=c("Run 1","Run 2"), col=c("red","Blue"))
boxplot(Tensile~Run, names=c("Run 1","Run 2"), col=c("red","Blue"))
t.test(Tensile~Run)
t.test(Run.1,Run.2)
detach(Tensile)
```
Test1

Grades on the first test in a statistics class

Description
Data for Exercise 5.80

Usage
Test1

Format
A data frame with 25 observations on the following variable.
test1 a numeric vector

Source

Examples
str(Test1)
attach(Test1)
EDA(test1)
detach(Test1)

Thermal
Heat loss of thermal pane windows versus outside temperature

Description
Data for Example 9.5

Usage
Thermal

Format
A data frame with 12 observations on the following 3 variables.
temp a numeric vector
loss a numeric vector
x a numeric vector

Source
Examples

\begin{verbatim}
str(Thermal)
attach(Thermal)
model <- lm(loss~temp)
summary(model)
detach(Thermal)
\end{verbatim}

Tiaa

1999-2000 closing prices for TIAA-CREF stocks

Description

Data for your enjoyment

Usage

Tiaa

Format

A data frame with 365 observations on the following 4 variables.

crestk a numeric vector
crefgwt a numeric vector
tiaa a numeric vector
date a factor with levels 1/01/2000 1/02/2000 1/03/2000 1/04/2000 1/05/2000 1/06/2000
1/31/2000 1/01/1999 1/02/1999 1/03/1999 1/04/1999 1/05/1999 1/06/1999 1/07/1999
1/08/1999 1/09/1999 1/10/1999 1/11/1999 1/12/1999 1/13/1999 1/14/1999
1/15/1999 1/16/1999 1/17/1999 1/18/1999 1/19/1999 1/20/1999 1/21/1999 1/22/1999
\end{verbatim}
Ticket


Source


Examples

data(Tiaa)

Ticket

<table>
<thead>
<tr>
<th>Time to complete an airline ticket reservation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ticket</td>
</tr>
</tbody>
</table>

Description

Data for Exercise 5.18

Usage

Ticket

Format

A data frame with 20 observations on the following variable.

  time  a numeric vector

Source

**Examples**

```r
str(Ticket)
attach(Ticket)
EDA(time)
detach(Ticket)
```

---

**Description**

Data for Exercise 9.35

**Usage**

`Toaster`

**Format**

A data frame with 17 observations on the following 3 variables.

- **toaster** a factor with levels `Black&D SO2500GBlack&D T660GBlack&D TRO200Black&D TRO400 Black&D TRO510 DeLonghi XU14 DeLonghi XU20L Hamilton Beach 336 Munsey M88 Panasonic NT855U Proctor-Silex 03008 Proctor-Silex 03010 Proctor-Silex 03030 Sears Kenmore 48216 Toastmaster 319V Toastmaster 336V Toastmaster 342`
- **score** a numeric vector
- **cost** a numeric vector

**Source**


**Examples**

```r
data(Toaster)
```

---

**Description**

Size of tonsils collected from 1,398 children

**Usage**

`Tonsils`
The number of torts, average number of months to process a tort, and county population from the court files of the nation’s largest counties

Data for Exercise 5.13

Usage

Tort

Format

A data frame with 45 observations on the following 5 variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>county</td>
<td>A factor with levels alameda, ca, allegheny, pa, bergen, nj, bexar, tx, contra costa, ca, cook, il, cuyahoga, oh, dade, fl, dallas, tx, dupage, il, essex, ma, essex, nj, fairfax, va, fairfield, ct, franklin, oh, fresno, ca, fulton, gaharris, tx, hartford, ct, hennepin, mn, honolulu, hi, jefferson, ky, king, wa, los angeles, camaricopa, az, marin, in, middlesex, ma, middlesex, nj, milwaukee, w, new york, ny, norfolk, ma, oakland, mi, orange, ca, orange, fl, palm beach, fl, philadelphia, pa, pima, az, san bernadino, ca, san francisco, ca, santa clara, ca, st, louis, mo, suffolk, ma, ventura, ca, wayne, mi, worchester, ma</td>
</tr>
<tr>
<td>months</td>
<td>A numeric vector</td>
</tr>
<tr>
<td>populat</td>
<td>A numeric vector</td>
</tr>
<tr>
<td>torts</td>
<td>A numeric vector</td>
</tr>
<tr>
<td>rate</td>
<td>A numeric vector</td>
</tr>
</tbody>
</table>
Toxic

Source

Examples

```
str(Toxic)
attach(Toxic)
EDA(months)
detach(Toxic)
```

---

Toxic

*Hazardous waste sites near minority communities*

Description

Data for Exercises 1.55, 5.08, 5.109, 8.58, and 10.35

Usage

Toxic

Format

A data frame with 51 observations on the following 5 variables.

- `state`: a factor with levels alabama alaska arizona arkansas california colorado connecticut delaware dist of columbia florida georgia hawaii idaho illinois indiana iowa kansas kentucky louisiana maine maryland massachusetts michigan minnesota mississippi missouri montana nebraska nevada new hampshire new jersey new mexico new york north carolina north dakota ohio oklahoma oregon pennsylvania rhode island south carolina south dakota tennessee texas utah vermont virginia washington west virginia wisconsin wyoming
- `region`: a factor with levels midwest northeast south west
- `sites`: a numeric vector
- `minority`: a numeric vector
- `percent`: a numeric vector

Source


Examples

```
str(Toxic)
attach(Toxic)
hist(minority,prob=TRUE)
lines(density(minority))
SIGN.test(sites,conf.level=.98)
boxplot(sites~region)
kruskal.test(sites~as.factor(region))
detach(Toxic)
```
Description

Data for Exercises 2.97, 5.115, and 9.62

Usage

Track

Format

A data frame with 55 observations on the following 8 variables.

country a factor with levels argentina australia austria belgium bermuda brazil burma canada chile china colombia cookis costa czech denmark domrep dprkorea finland france frg gbni gdr greece guatemala hungary india indonesia ireland israel italy japan kenya korea luxembourg malaysia mauritius mexico netherlands norway nz philippines png poland portugal rumania singapore spain sweden switzerland taipei thailand turkey usa ussr wsamoa

X100m a numeric vector
X200m a numeric vector
X400m a numeric vector
X800m a numeric vector
X1500m a numeric vector
X3000m a numeric vector
marathon a numeric vector

Source


Examples

str(Track)
attach(Track)
cor(Track[,2:8])
pairs(Track[,2:8])
detach(Track)
Track15

*Olympic winning times for the men’s 1500-meter run*

### Description

Data for Exercise 1.36

### Usage

*Track15*

### Format

A data frame with 26 observations on the following 2 variables.

- **Year**: a numeric vector
- **X1500m**: a numeric vector

### Source


### Examples

```r
str(Track15)
attach(Track15)
plot(Year,X1500m,type="l",lwd=2,lty=2,col="red",xlab="Year",
     ylab="1500m Time (seconds)",main="Problem 1.36")
detach(Track15)
```

---

Treatments

*Illustrates analysis of variance for three treatment groups*

### Description

Data for Exercise 10.44

### Usage

*Treatments*

### Format

A data frame with 24 observations on the following 5 variables.

- **treat1**: a numeric vector
- **treat2**: a numeric vector
- **treat3**: a numeric vector
- **Treatment**: a numeric vector
- **Group**: a numeric vector

---
Trees

Source


Examples

```r
str(Treatments)
attach(Treatments)
anova(lm(Treatment~as.factor(Group)))
detach(Treatments)
```

---

**trees**

Number of trees in 20 grids

---

Description

Data for Exercise 1.50

Usage

Trees

Format

A data frame with 20 observations on the following variable.

- **number**: a numeric vector

Source


Examples

```r
str(Trees)
attach(Trees)
stem(number)
hist(number,breaks=seq(60,110,10),right=FALSE,col="green",main="Problem 1.5.0")
detach(Trees)
```
Trucks

Miles per gallon for standard 4-wheel drive trucks manufactured by Chevrolet, Dodge and Ford

Description

Data for Example 10.2

Usage

Trucks

Format

A data frame with 15 observations on the following 5 variables.

chevy a numeric vector
dodge a numeric vector
ford a numeric vector
gas.mileage a numeric vector
truck a factor with levels chevy dodge ford

Source


Examples

str(Trucks)
attach(Trucks)
anova(lm(gas.mileage~truck))
detach(Trucks)

tsum.test

Summarized t-test

Description

Performs a one-sample, two-sample, or a Welch modified two-sample t-test based on user supplied summary information. Output is identical to that produced with t.test.

Usage

tsum.test(mean.x, s.x = NULL, n.x = NULL, mean.y = NULL, s.y = NULL, n.y = NULL, alternative = "two.sided", mu = 0, var.equal = FALSE, conf.level = 0.95)
Arguments

- **mean.x**: a single number representing the sample mean of x
- **s.x**: a single number representing the sample standard deviation for x
- **n.x**: a single number representing the sample size for x
- **mean.y**: a single number representing the sample mean of y
- **s.y**: a single number representing the sample standard deviation for y
- **n.y**: a single number representing the sample size for y
- **alternative**: is a character string, one of "greater", "less" or "two.sided", or just the initial letter of each, indicating the specification of the alternative hypothesis. For one-sample tests, alternative refers to the true mean of the parent population in relation to the hypothesized value mu. For the standard two-sample tests, alternative refers to the difference between the true population mean for x and that for y, in relation to mu. For the one-sample and paired t-tests, alternative refers to the true mean of the parent population in relation to the hypothesized value mu. For the standard and Welch modified two-sample t-tests, alternative refers to the difference between the true population mean for x and that for y, in relation to mu. For the one-sample t-tests, alternative refers to the true mean of the parent population in relation to the hypothesized value mu. For the standard and Welch modified two-sample t-tests, alternative refers to the difference between the true population mean for x and that for y, in relation to mu.
- **mu**: is a single number representing the value of the mean or difference in means specified by the null hypothesis.
- **var.equal**: logical flag: if TRUE, the variances of the parent populations of x and y are assumed equal. Argument var.equal should be supplied only for the two-sample tests.
- **conf.level**: is the confidence level for the returned confidence interval; it must lie between zero and one.

Details

If y is NULL, a one-sample t-test is carried out with x. If y is not NULL, either a standard or Welch modified two-sample t-test is performed, depending on whether var.equal is TRUE or FALSE.

Value

A list of class htest, containing the following components:

- **statistic**: the t-statistic, with names attribute "t"
- **parameters**: is the degrees of freedom of the t-distribution associated with statistic. Component parameters has names attribute "df".
- **p.value**: the p-value for the test.
- **conf.int**: is a confidence interval (vector of length 2) for the true mean or difference in means. The confidence level is recorded in the attribute conf.level. When alternative is not "two.sided", the confidence interval will be half-infinite, to reflect the interpretation of a confidence interval as the set of all values k for which one would not reject the null hypothesis that the true mean or difference in means is k. Here infinity will be represented by Inf.
estimate vector of length 1 or 2, giving the sample mean(s) or mean of differences; these estimate the corresponding population parameters. Component estimate has a names attribute describing its elements.

null.value the value of the mean or difference in means specified by the null hypothesis. This equals the input argument mu. Component null.value has a names attribute describing its elements.

alternative records the value of the input argument alternative: "greater", "less" or "two.sided".

data.name a character string (vector of length 1) containing the names x and y for the two summarized samples.

Null Hypothesis

For the one-sample t-test, the null hypothesis is that the mean of the population from which x is drawn is mu. For the standard and Welch modified two-sample t-tests, the null hypothesis is that the population mean for x less that for y is mu.

The alternative hypothesis in each case indicates the direction of divergence of the population mean for x (or difference of means for x and y) from mu (i.e., "greater", "less", or "two.sided").

Test Assumptions

The assumption of equal population variances is central to the standard two-sample t-test. This test can be misleading when population variances are not equal, as the null distribution of the test statistic is no longer a t-distribution. If the assumption of equal variances is doubtful with respect to a particular dataset, the Welch modification of the t-test should be used.

The t-test and the associated confidence interval are quite robust with respect to level toward heavy-tailed non-Gaussian distributions (e.g., data with outliers). However, the t-test is non-robust with respect to power, and the confidence interval is non-robust with respect to average length, toward these same types of distributions.

Confidence Intervals

For each of the above tests, an expression for the related confidence interval (returned component conf.int) can be obtained in the usual way by inverting the expression for the test statistic. Note that, as explained under the description of conf.int, the confidence interval will be half-infinite when alternative is not "two.sided"; infinity will be represented by Inf.

Author(s)

Alan T. Arnholt

References

See Also

`z.test, zsum.test`

Examples

```r
tsum.test(mean.x=5.6, s.x=2.1, n.x=16, mu=4.9, alternative="greater")
# Problem 6.31 on page 324 of BSDA states: The chamber of commerce
# of a particular city claims that the mean carbon dioxide
# level of air pollution is no greater than 4.9 ppm. A random
# sample of 16 readings resulted in a sample mean of 5.6 ppm,
# and s=2.1 ppm. One-sided one-sample t-test. The null
# hypothesis is that the population mean for 'x' is 4.9.
# The alternative hypothesis states that it is greater than 4.9.
```

```r
x <- rnorm(12)
tsum.test(mean(x), sd(x), n.x=12)
# Two-sided one-sample t-test. The null hypothesis is that
# the population mean for 'x' is zero. The alternative
# hypothesis states that it is either greater or less
# than zero. A confidence interval for the population mean
# will be computed. Note: above returns same answer as:
t.test(x)
```

```r
x <- c(7.8, 6.6, 6.5, 7.4, 7.3, 7.0, 6.4, 7.1, 6.7, 7.6, 6.8)
y <- c(4.5, 5.4, 6.1, 6.1, 5.4, 5.0, 4.1, 5.5)
tsum.test(mean(x), s.x=sd(x), n.x=11 ,mean(y), s.y=sd(y), n.y=8, mu=2)
# Two-sided standard two-sample t-test. The null hypothesis
# is that the population mean for 'x' less that for 'y' is 2.
# The alternative hypothesis is that this difference is not 2.
# A confidence interval for the true difference will be computed.
# Note: above returns same answer as:
t.test(x, y)
```

```r
tsum.test(mean(x), s.x=sd(x), n.x=11, mean(y), s.y=sd(y), n.y=8, conf.level=0.90)
# Two-sided standard two-sample t-test. The null hypothesis
# is that the population mean for 'x' less that for 'y' is zero.
# The alternative hypothesis is that this difference is not
# zero. A 90% confidence interval for the true difference will
# be computed. Note: above returns same answer as:
t.test(x, y, conf.level=0.90)
```

---

**Tv**

Percent of students that watch more than 6 hours of TV per day versus national math test scores

Description

Data for Examples 2.1 and 2.7

Usage

Tv
Twin

Format
A data frame with 53 observations on the following 3 variables.


percent a numeric vector
test a numeric vector

Source

Examples
str(Tv)
attach(Tv)
plot(percent,test,col="blue")
cor(percent,test,use="complete.obs")
detach(Tv)

Twin

Intelligence test scores for identical twins in which one twin is given a drug

Description
Data for Exercise 7.54

Usage
Twin

Format
A data frame with 9 observations on the following 3 variables.

TwinA a numeric vector
TwinB a numeric vector
differ a numeric vector

Source
Examples

```r
str(Twin)
attach(Twin)
qqnorm(differ)
qqline(differ)
shapiro.test(differ)
t.test(TwinA,TwinB,paired=TRUE)
detach(Twin)
```

Undergrad

Data set describing a sample of undergraduate students

Description

Data for Exercise 1.15

Usage

Undergrad

Format

A data frame with 100 observations on the following 6 variables.

- **Gender**: a factor with levels Female Male
- **Major**: a factor with levels Accounting Biology Chemistry English Geology History Math Music Physics Psychology Sociology
- **Class**: a factor with levels Freshman Junior Senior Sophomore
- **GPA**: a numeric vector
- **SAT**: a numeric vector
- **Drops**: a numeric vector

Source


Examples

```r
str(Undergrad)
attach(Undergrad)
stripchart(GPA~Class,method="stack",col=c("blue","red","green","lightblue"),
pch=19,main="GPA versus Class")
stripchart(GPA~Gender,method="stack",col=c("red","blue"),pch=19,
main="GPA versus Gender")
stripchart(SAT~Drops,method="stack",col=c("blue","red","green","lightblue"),
pch=19,main="SAT versus Drops")
stripchart(Drops~Gender,method="stack",col=c("red","blue"),pch=19,
main="Drops versus Gender")
detach(Undergrad)
```
Vacation

Number of days of paid holidays and vacation leave for sample of 35 textile workers

Description
Data for Exercise 6.46 and 6.98

Usage
Vacation

Format
A data frame with 35 observations on the following variable.

   number  a numeric vector

Source

Examples
str(Vacation)
attach(Vacation)
EDA(number)
t.test(number, mu=24)
detach(Vacation)

Vaccine

Reported serious reactions due to vaccines in 11 southern states

Description
Data for Exercise 1.111

Usage
Vaccine

Format
A data frame with 11 observations on the following 2 variables.

   State  a factor with levels Alabama Arkansas Florida Georgia Louisiana Mississippi North Carolina Oklahoma South Carolina Tennessee Texas
   number  a numeric vector

Source
Examples

str(Vaccine)
attach(Vaccine)
fn <- fivenum(number)
fn
iqr <- IQR(number)
ll <- fn[2]-1.5*iqr
ul <- fn[4]+1.5*iqr
limits <- c(ll,ul)
limits
boxplot(number)
detach(Vaccine)

Vehicle

<table>
<thead>
<tr>
<th>make</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Fatality ratings for foreign and domestic vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>domestic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>foreign</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Description

Data for Exercise 8.34

Usage

Vehicle

Format

A data frame with 2 observations on the following 6 variables.

make a factor with levels domestic foreign
A a numeric vector
B a numeric vector
C a numeric vector
D a numeric vector
E a numeric vector
F a numeric vector

Source


Examples

str(Vehicle)
attach(Vehicle)
Vehicle
chisq.test(Vehicle[,2:6])
detach(Vehicle)
Verbal test scores and number of library books checked out for 15 eighth graders

Description
Data for Exercise 9.30

Usage
Verbal

Format
A data frame with 15 observations on the following 4 variables.

number  a numeric vector
verbal   a numeric vector
SRES1    a numeric vector
FITS1    a numeric vector

Source

Examples
str(Verbal)
attach(Verbal)
model <- lm(verbal~number)
summary(model)
detach(Verbal)
remove(model)

Number of sunspots versus mean annual level of Lake Victoria Nyanza from 1902 to 1921

Description
Data for Exercise 2.98

Usage
Victoria
Viscosit

Format
A data frame with 20 observations on the following 3 variables.

- year  a numeric vector
- level a numeric vector
- sunspot  a numeric vector

Source

Examples
str(Victoria)
attach(Victoria)
plot(sunspot, level)
model <- lm(level ~ sunspot)
abline(model)
cor(sunspot, level)
model
detach(Victoria)

Viscosit  
Viscosity measurements of a substance on two different days

Description
Data for Exercise 7.44

Usage
Viscosit

Format
A data frame with 11 observations on the following 2 variables.

- first  a numeric vector
- second  a numeric vector

Source

Examples
str(Viscosit)
attach(Viscosit)
t.test(first, second)
detach(Viscosit)
Visual

Visual acuity of a group of subjects tested under a specified dose of a drug

Description
Data for Exercise 5.6

Usage
Visual

Format
A data frame with 18 observations on the following variable.
visual a numeric vector

Source

Examples
str(Visual)
attach(Visual)
stem(visual)
fivenum(visual)
boxplot(visual)
detach(Visual)

Vocab

Reading scores before and after vocabulary training for 14 employees who did not complete high school

Description
Data for Exercise 7.80

Usage
Vocab

Format
A data frame with 14 observations on the following 2 variables.
First a numeric vector
Second a numeric vector
Wastewat

Source


Examples

```r
str(Vocab)
attach(Vocab)
DIF <- Second - First
qqnorm(DIF)
qqline(DIF)
shapiro.test(DIF)
t.test(Second,First,paired=TRUE)
detach(Vocab)
remove(DIF)
```

---

**Wastewat**  
*Volume of injected waste water from Rocky Mountain Arsenal and number of earthquakes near Denver*

Description

Data for Exercise 9.18

Usage

Wastewat

Format

A data frame with 44 observations on the following 4 variables.

- **gallons**: a numeric vector
- **number**: a numeric vector
- **ln.no.**: a numeric vector
- **index**: a numeric vector

Source


Examples

```r
str(Wastewat)
attach(Wastewat)
model <- lm(number~gallons)
summary(model)
detach(Wastewat)
remove(model)
```
Weather94  

**Description**  
Data for Exercise 1.30

**Usage**  
Weather94

**Format**  
A data frame with 11 observations on the following 2 variables.  
- `Weather.Type` a factor with levels: Extreme Temp Flash flood Fog High wind Hurricane Lightning Other River flood Thunderstorm Tornado Winter weather  
- `Number` a numeric vector

**Source**  

**Examples**  
str(Weather94)  
attach(Weather94)  
names(Number) <- Weather.Type  
barplot(Number,col="lightblue",las=2,cex.names=.65,main="Problem 1.3/zero.noslash")  
# las=2 places bar names vertically  
detach(Weather94)

---

Wheat  

**Description**  
Price of a bushel of wheat versus the national weekly earnings of production workers

**Usage**  
Wheat

**Format**  
A data frame with 19 observations on the following 3 variables.  
- `year` a numeric vector  
- `earnings` a numeric vector  
- `price` a numeric vector
Windmill

Source


Examples

```r
str(Wheat)
attach(Wheat)
par(mfrow=c(1,2))
plot(year,earnings)
plot(year,price)
par(mfrow=c(1,1))
detach(Wheat)
```

```
Windmill

Direct current produced by different wind velocities
```

Description

Data for Exercise 9.34

Usage

Windmill

Format

A data frame with 25 observations on the following 7 variables.

- `velocity`: a numeric vector
- `output`: a numeric vector
- `SRES1`: a numeric vector
- `FITS1`: a numeric vector
- `X1.velocity`: a numeric vector
- `SRES2`: a numeric vector
- `FITS2`: a numeric vector

Source


Examples

```r
str(Windmill)
attach(Windmill)
summary(lm(output~velocity))
detach(Windmill)
```
**Window**  
*Wind leakage for storm windows exposed to a 50 mph wind*

**Description**  
Data for Exercise 6.54

**Usage**  
Window

**Format**  
A data frame with 9 observations on the following 2 variables.

- Window  a numeric vector
- Leakage  a numeric vector

**Source**  

**Examples**
```r
str(Window)
attach(Window)
SIGN.test(Leakage,md=.125,alternative="greater")
detach(Window)
```

---

**Wins**  
*Baseball team wins versus 7 independent variables for National league teams in 1990*

**Description**  
Data for Exercise 9.23

**Usage**  
Wins

**Format**  
A data frame with 12 observations on the following 9 variables.

- team  a factor with levels Atlanta Chicago Cincinnati Houston Los Angeles Montreal New York Philadelphia Pittsburgh San Diego San Francisco St. Louis
- wins  a numeric vector
- batavg  a numeric vector
- rbi  a numeric vector
Wool

stole a numeric vector
strkout a numeric vector
caught a numeric vector
errors a numeric vector
era a numeric vector

Source

Examples
str(Wins)
attach(Wins)
plot(era,wins)
model <- lm(wins~era)
abline(model)
summary(model)
detach(Wins)
remove(model)

---

Wool

*Strength tests of two types of wool fabric*

Description
Data for Exercise 7.42

Usage
Wool

Format
A data frame with 10 observations on the following 2 variables.

Type.1 a numeric vector
Type.2 a numeric vector

Source

Examples
str(Wool)
attach(Wool)
t.test(Type.1,Type.2,var.equal=TRUE)
detach(Wool)
## Description
Data for Exercise 2.7

## Usage
Yearsunspot

## Format
A data frame with 252 observations on the following 24 variables.

- X1979 a numeric vector
- X1980 a numeric vector
- X1981 a numeric vector
- X1982 a numeric vector
- X1983 a numeric vector
- X1984 a numeric vector
- X1985 a numeric vector
- X1986 a numeric vector
- X1987 a numeric vector
- X1988 a numeric vector
- X1989 a numeric vector
- X1990 a numeric vector
- X1991 a numeric vector
- X1992 a numeric vector
- X1993 a numeric vector
- X1994 a numeric vector
- X1995 a numeric vector
- X1996 a numeric vector
- X1997 a numeric vector
- X1998 a numeric vector
- X1999 a numeric vector
- X2000 a numeric vector
- SSN a numeric vector
- year a numeric vector

## Source
z.test

Examples

str(Yearsunspot)
attach(Yearsunspot)
boxplot(SSN~year,main="Exercise 2.7",col="lightblue")
detach(Yearsunspot)

---

z.test Z-test

Description

This function is based on the standard normal distribution and creates confidence intervals and tests hypotheses for both one and two sample problems.

Usage

z.test(x, y = NULL, alternative = "two.sided", mu = 0, sigma.x = NULL, sigma.y = NULL, conf.level = 0.95)

Arguments

x numeric vector; NAs and Infs are allowed but will be removed.
y numeric vector; NAs and Infs are allowed but will be removed.
alternative character string, one of "greater", "less" or "two.sided", or the initial letter of each, indicating the specification of the alternative hypothesis. For one-sample tests, alternative refers to the true mean of the parent population in relation to the hypothesized value mu. For the standard two-sample tests, alternative refers to the difference between the true population mean for x and that for y, in relation to mu.
mu a single number representing the value of the mean or difference in means specified by the null hypothesis
sigma.x a single number representing the population standard deviation for x
sigma.y a single number representing the population standard deviation for y
conf.level confidence level for the returned confidence interval, restricted to lie between zero and one

Details

If y is NULL, a one-sample z-test is carried out with x. If y is not NULL, a standard two-sample z-test is performed.

Value

A list of class htest, containing the following components:

statistic the z-statistic, with names attribute "z"
p.value the p-value for the test
conf.int is a confidence interval (vector of length 2) for the true mean or difference in means. The confidence level is recorded in the attribute conf.level. When alternative is not "two.sided", the confidence interval will be half-infinite, to reflect the interpretation of a confidence interval as the set of all values $k$ for which one would not reject the null hypothesis that the true mean or difference in means is $k$. Here infinity will be represented by Inf.

estimate vector of length 1 or 2, giving the sample mean(s) or mean of differences; these estimate the corresponding population parameters. Component estimate has a names attribute describing its elements.

null.value is the value of the mean or difference in means specified by the null hypothesis. This equals the input argument mu. Component null.value has a names attribute describing its elements.

alternative records the value of the input argument alternative: "greater", "less" or "two.sided".

data.name a character string (vector of length 1) containing the actual names of the input vectors x and y

Null Hypothesis

For the one-sample z-test, the null hypothesis is that the mean of the population from which x is drawn is mu. For the standard two-sample z-tests, the null hypothesis is that the population mean for x less that for y is mu.

The alternative hypothesis in each case indicates the direction of divergence of the population mean for x (or difference of means for x and y) from mu (i.e., "greater", "less", "two.sided").

Test Assumptions

The assumption of normality for the underlying distribution or a sufficiently large sample size is required along with the population standard deviation to use Z procedures.

Confidence Interval

For each of the above tests, an expression for the related confidence interval (returned component conf.int) can be obtained in the usual way by inverting the expression for the test statistic. Note that, as explained under the description of conf.int, the confidence interval will be half-infinite when alternative is not "two.sided"; infinity will be represented by Inf.

Author(s)

Alan T. Arnholt

References


zsum.test

Summarized z-test

Description

This function is based on the standard normal distribution and creates confidence intervals and tests hypotheses for both one and two sample problems based on summarized information the user passes to the function. Output is identical to that produced with z.test.

Usage

zsum.test(mean.x, sigma.x = NULL, n.x = NULL, mean.y = NULL, sigma.y = NULL, n.y = NULL, alternative = "two.sided", mu = 0, conf.level = 0.95)

Arguments

mean.x a single number representing the sample mean of x
sigma.x a single number representing the population standard deviation for x
n.x a single number representing the sample size for x
mean.y a single number representing the sample mean of y

Examples

x <- rnorm(12)
z.test(x, sigma.x=1)
  # Two-sided one-sample z-test where the assumed value for
  # sigma.x is one. The null hypothesis is that the population
  # mean for 'x' is zero. The alternative hypothesis states
  # that it is either greater or less than zero. A confidence
  # interval for the population mean will be computed.

x <- c(7.8, 6.6, 6.5, 7.4, 7.3, 7., 6.4, 7.1, 6.7, 7.6, 6.8)
y <- c(4.5, 5.4, 6.1, 6.1, 5.4, 5., 4.1, 5.5)
z.test(x, sigma.x=0.5, y, sigma.y=0.5, mu=2)
  # Two-sided standard two-sample z-test where both sigma.x
  # and sigma.y are both assumed to equal 0.5. The null hypothesis
  # is that the population mean for 'x' less that for 'y' is 2.
  # The alternative hypothesis is that this difference is not 2.
  # A confidence interval for the true difference will be computed.

z.test(x, sigma.x=0.5, y, sigma.y=0.5, conf.level=0.90)
  # Two-sided standard two-sample z-test where both sigma.x and
  # sigma.y are both assumed to equal 0.5. The null hypothesis
  # is that the population mean for 'x' less that for 'y' is zero.
  # The alternative hypothesis is that this difference is not
  # zero. A 90% confidence interval for the true difference will
  # be computed.

rm(x, y)
**zsum.test**

sigma.y  
a single number representing the population standard deviation for y

n.y  
a single number representing the sample size for y

alternative  
is a character string, one of "greater", "less" or "two.sided", or the initial letter of each, indicating the specification of the alternative hypothesis. For one-sample tests, alternative refers to the true mean of the parent population in relation to the hypothesized value mu. For the standard two-sample tests, alternative refers to the difference between the true population mean for x and that for y, in relation to mu.

mu  
a single number representing the value of the mean or difference in means specified by the null hypothesis

conf.level  
confidence level for the returned confidence interval, restricted to lie between zero and one

**Details**

If y is NULL, a one-sample z-test is carried out with x. If y is not NULL, a standard two-sample z-test is performed.

**Value**

A list of class htest, containing the following components:

- **statistic**  
  the z-statistic, with names attribute z.

- **p.value**  
  the p-value for the test

- **conf.int**  
  is a confidence interval (vector of length 2) for the true mean or difference in means. The confidence level is recorded in the attribute conf.level. When alternative is not "two.sided", the confidence interval will be half-infinite, to reflect the interpretation of a confidence interval as the set of all values k for which one would not reject the null hypothesis that the true mean or difference in means is k. Here, infinity will be represented by Inf.

- **estimate**  
  vector of length 1 or 2, giving the sample mean(s) or mean of differences; these estimate the corresponding population parameters. Component estimate has a names attribute describing its elements.

- **null.value**  
  the value of the mean or difference in means specified by the null hypothesis. This equals the input argument mu. Component null.value has a names attribute describing its elements.

- **alternative**  
  records the value of the input argument alternative: "greater", "less" or "two.sided".

- **data.name**  
  a character string (vector of length 1) containing the names x and y for the two summarized samples

**Null Hypothesis**

For the one-sample z-test, the null hypothesis is that the mean of the population from which x is drawn is mu. For the standard two-sample z-tests, the null hypothesis is that the population mean for x less that for y is mu.

The alternative hypothesis in each case indicates the direction of divergence of the population mean for x (or difference of means of x and y) from mu (i.e., "greater", "less", "two.sided").
Test Assumptions

The assumption of normality for the underlying distribution or a sufficiently large sample size is required along with the population standard deviation to use Z procedures.

Confidence Intervals

For each of the above tests, an expression for the related confidence interval (returned component \textit{conf.int}) can be obtained in the usual way by inverting the expression for the test statistic. Note that, as explained under the description of \textit{conf.int}, the confidence interval will be half-infinite when alternative is not "two.sided"; infinity will be represented by Inf.

Author(s)

Alan T. Arnholt

References


See Also

\textit{z.test}, \textit{tsum.test}

Examples

\begin{verbatim}
zsum.test(mean.x=56/30,sigma.x=2, n.x=30, alternative="greater", mu=1.8)
  # Example 9.7 part a. from PASWR.
x <- rnorm(12)
zsum.test(mean(x), sigma.x=1, n.x=12)
  # Two-sided one-sample z-test where the assumed value for
  # sigma.x is one. The null hypothesis is that the population
  # mean for 'x' is zero. The alternative hypothesis states
  # that it is either greater or less than zero. A confidence
  # interval for the population mean will be computed.
  # Note: returns same answer as:
z.test(x, sigma.x=1)
  #
x <- c(7.8, 6.6, 6.5, 7.4, 7.3, 7.0, 6.4, 7.1, 6.7, 7.6, 6.8)
y <- c(4.5, 5.4, 6.1, 6.1, 5.4, 4.1, 5.5)
zsum.test(mean(x), sigma.x=0.5, n.x=11, mean(y), sigma.y=0.5, n.y=8, mu=2)
  # Two-sided standard two-sample z-test where both sigma.x
  # and sigma.y are both assumed to equal 0.5. The null hypothesis
  # is that the population mean for 'x' less that for 'y' is 2.
  # The alternative hypothesis is that this difference is not 2.
  # A confidence interval for the true difference will be computed.
  # Note: returns same answer as:
z.test(x, sigma.x=0.5, y, sigma.y=0.5)
  #\end{verbatim}
zsum.test(mean(x), sigma.x=0.5, n.x=11, mean(y), sigma.y=0.5, n.y=8, conf.level=0.90)

# Two-sided standard two-sample z-test where both sigma.x and
# sigma.y are both assumed to equal 0.5. The null hypothesis
# is that the population mean for 'x' less that for 'y' is zero.
# The alternative hypothesis is that this difference is not
# zero. A 90% confidence interval for the true difference will
# be computed. Note: returns same answer as:
z.test(x, sigma.x=0.5, y, sigma.y=0.5, conf.level=0.90)
rm(x, y)
Index

**Topic** datasets
Abbey, 8
Abc, 9
Abilene, 10
Ability, 10
Abortion, 11
Absent, 12
Achieve, 13
Adsales, 13
Aggress, 14
Aid, 15
Aids, 15
Airdisasters, 16
Airline, 17
Alcohol, 18
Allergy, 18
Anesthet, 19
Anxiety, 20
Apolipop, 20
Append, 21
Appendec, 22
Aptitude, 22
Archeao, 23
Arthriti, 24
Artifici, 25
Asprin, 25
Asthmati, 26
Attorney, 27
Autogear, 28
Backtoback, 28
Bbsalaries, 29
Bigten, 30
Biology, 30
Birth, 31
Blackedu, 32
Blood, 33
Board, 33
Bones, 34
Books, 35
Bookstor, 35
Brain, 36
Bumpers, 37
Bus, 37
Bypass, 38
Cabinets, 39
Cancer, 40
Carbon, 40
Cat, 41
Censored, 42
Challeng, 42
Chemist, 43
Chesapea, 44
Chevy, 45
Chicken, 45
Chipavg, 46
Chips, 47
Cigar, 48
Cigarett, 48
Citrus, 50
Clean, 51
Coaxial, 52
Coffee, 52
Coins, 53
Commute, 54
Concept, 55
Concrete, 56
Corn, 56
Correlat, 57
Counsel, 58
Cpi, 58
Crime, 59
Darwin, 60
Dealers, 60
Defectiv, 61
Degree, 61
Delay, 62
Depend, 63
Detroit, 64
Develop, 64
Devmath, 65
Dice, 65
Diesel, 66
Diplomat, 67
Disposal, 68
Dogs, 68
Domestic, 69
Dopamine, 70
Dowjones, 70
Drink, 71
Drug, 72
Dyslexia, 72
Earthqk, 73
Educat, 75
Eggs, 75
Elderly, 76
Energy, 77
Engineer, 78
Entrance, 78
Epaminicompact, 79
Epatwoseater, 80
Executiv, 81
Exercise, 81
Fabric, 82
Faithful, 83
Family, 83
Ferraro1, 84
Ferraro2, 85
Fertility, 85
Firstchi, 86
Fish, 87
Fitness, 87
Florida2000, 88
Fluid, 89
Food, 90
Framingh, 90
Freshman, 91
Funeral, 92
Galaxie, 92
Gallup, 93
Gasoline, 94
German, 95
Golf, 95
Governor, 96
Gpa, 97
Grades, 97
Graduate, 98
Greenriv, 99
Grrriv2, 99
Groupabc, 100
Groups, 101
Gym, 101
Habits, 102
Haptologo, 103
Hardware, 103
Hardwood, 104
Heat, 104
Heating, 105
Hodgkin, 106
Homes, 107
Homework, 108
Honda, 108
Hostile, 109
Housing, 110
Hurrican, 111
Iceberg, 112
Income, 112
Independent, 113
Indian, 114
Indiapol, 115
Indy500, 115
Inflatio, 116
Inletoil, 117
Inmate, 117
Inspect, 118
Insulate, 119
Iqgpa, 119
Irises, 120
Jdpower, 121
Jobsat, 121
Kidsmoke, 122
Kilowatt, 123
Kinder, 123
Laminect, 124
Lead, 125
Leader, 125
Lethal, 126
Life, 127
Lifespan, 127
Ligntmonth, 128
Lodge, 129
Longtail, 130
Lowabil, 130
Magnesiu, 131
Malpract, 132
Manager, 132
Marked, 133
MATH, 134
Mathcomp, 134
Mathpro, 135
Maze, 136
Median, 136
Mental, 137
Mercury, 138
Metrent, 138
Miller, 139
Miller1, 139
Moisture, 140
Monoxide, 141
Movie, 141
Music, 142
<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name, 143</td>
<td></td>
</tr>
<tr>
<td>Nascar, 143</td>
<td></td>
</tr>
<tr>
<td>Nervous, 144</td>
<td></td>
</tr>
<tr>
<td>Newsstand, 145</td>
<td></td>
</tr>
<tr>
<td>NFLdraft2, 145</td>
<td></td>
</tr>
<tr>
<td>NFLdraft, 146</td>
<td></td>
</tr>
<tr>
<td>Nicotine, 147</td>
<td></td>
</tr>
<tr>
<td>Orange, 150</td>
<td></td>
</tr>
<tr>
<td>Orioles, 151</td>
<td></td>
</tr>
<tr>
<td>Oxytocin, 151</td>
<td></td>
</tr>
<tr>
<td>Parented, 152</td>
<td></td>
</tr>
<tr>
<td>Patrol, 153</td>
<td></td>
</tr>
<tr>
<td>Pearson, 153</td>
<td></td>
</tr>
<tr>
<td>Phone, 154</td>
<td></td>
</tr>
<tr>
<td>Poison, 155</td>
<td></td>
</tr>
<tr>
<td>Politic, 155</td>
<td></td>
</tr>
<tr>
<td>Pollutio, 156</td>
<td></td>
</tr>
<tr>
<td>Porosity, 156</td>
<td></td>
</tr>
<tr>
<td>Poverty, 157</td>
<td></td>
</tr>
<tr>
<td>Precinct, 158</td>
<td></td>
</tr>
<tr>
<td>Prejudic, 159</td>
<td></td>
</tr>
<tr>
<td>Presiden, 159</td>
<td></td>
</tr>
<tr>
<td>Press, 160</td>
<td></td>
</tr>
<tr>
<td>Prognost, 161</td>
<td></td>
</tr>
<tr>
<td>Program, 161</td>
<td></td>
</tr>
<tr>
<td>Psat, 162</td>
<td></td>
</tr>
<tr>
<td>Psych, 163</td>
<td></td>
</tr>
<tr>
<td>Puerto, 163</td>
<td></td>
</tr>
<tr>
<td>Quail, 164</td>
<td></td>
</tr>
<tr>
<td>Quality, 164</td>
<td></td>
</tr>
<tr>
<td>Rainks, 165</td>
<td></td>
</tr>
<tr>
<td>Randd, 166</td>
<td></td>
</tr>
<tr>
<td>Rat, 167</td>
<td></td>
</tr>
<tr>
<td>Ratings, 167</td>
<td></td>
</tr>
<tr>
<td>Reaction, 168</td>
<td></td>
</tr>
<tr>
<td>Reading, 169</td>
<td></td>
</tr>
<tr>
<td>Readiq, 169</td>
<td></td>
</tr>
<tr>
<td>Referend, 170</td>
<td></td>
</tr>
<tr>
<td>Region, 171</td>
<td></td>
</tr>
<tr>
<td>Register, 171</td>
<td></td>
</tr>
<tr>
<td>Rehab, 172</td>
<td></td>
</tr>
<tr>
<td>Remedial, 173</td>
<td></td>
</tr>
<tr>
<td>Rentals, 174</td>
<td></td>
</tr>
<tr>
<td>Repair, 174</td>
<td></td>
</tr>
<tr>
<td>Retail, 175</td>
<td></td>
</tr>
<tr>
<td>Ronbrown1, 175</td>
<td></td>
</tr>
<tr>
<td>Ronbrown2, 176</td>
<td></td>
</tr>
<tr>
<td>Rural, 177</td>
<td></td>
</tr>
<tr>
<td>Salary, 178</td>
<td></td>
</tr>
<tr>
<td>Salinity, 178</td>
<td></td>
</tr>
<tr>
<td>Sat, 179</td>
<td></td>
</tr>
<tr>
<td>Saving, 180</td>
<td></td>
</tr>
<tr>
<td>Scales, 180</td>
<td></td>
</tr>
<tr>
<td>Schizop2, 181</td>
<td></td>
</tr>
<tr>
<td>Schizoph, 182</td>
<td></td>
</tr>
<tr>
<td>Seatbelt, 182</td>
<td></td>
</tr>
<tr>
<td>Selfdefe, 183</td>
<td></td>
</tr>
<tr>
<td>Senior, 184</td>
<td></td>
</tr>
<tr>
<td>Sentence, 184</td>
<td></td>
</tr>
<tr>
<td>Shkd, 185</td>
<td></td>
</tr>
<tr>
<td>Shock, 186</td>
<td></td>
</tr>
<tr>
<td>Shoplift, 186</td>
<td></td>
</tr>
<tr>
<td>Short, 187</td>
<td></td>
</tr>
<tr>
<td>Shuttle, 188</td>
<td></td>
</tr>
<tr>
<td>Simpson, 191</td>
<td></td>
</tr>
<tr>
<td>Situp, 192</td>
<td></td>
</tr>
<tr>
<td>Skewed, 192</td>
<td></td>
</tr>
<tr>
<td>Skin, 193</td>
<td></td>
</tr>
<tr>
<td>Slc, 194</td>
<td></td>
</tr>
<tr>
<td>Smokyph, 194</td>
<td></td>
</tr>
<tr>
<td>Snore, 195</td>
<td></td>
</tr>
<tr>
<td>Snow, 196</td>
<td></td>
</tr>
<tr>
<td>Soccer, 196</td>
<td></td>
</tr>
<tr>
<td>Social, 197</td>
<td></td>
</tr>
<tr>
<td>Sophomor, 198</td>
<td></td>
</tr>
<tr>
<td>South, 198</td>
<td></td>
</tr>
<tr>
<td>Speed, 199</td>
<td></td>
</tr>
<tr>
<td>Spellers, 200</td>
<td></td>
</tr>
<tr>
<td>Spelling, 200</td>
<td></td>
</tr>
<tr>
<td>Sports, 201</td>
<td></td>
</tr>
<tr>
<td>Spouse, 202</td>
<td></td>
</tr>
<tr>
<td>Stable, 203</td>
<td></td>
</tr>
<tr>
<td>Stamp, 204</td>
<td></td>
</tr>
<tr>
<td>Statclas, 204</td>
<td></td>
</tr>
<tr>
<td>Statelaw, 205</td>
<td></td>
</tr>
<tr>
<td>Statisti, 206</td>
<td></td>
</tr>
<tr>
<td>Step, 206</td>
<td></td>
</tr>
<tr>
<td>Stress, 207</td>
<td></td>
</tr>
<tr>
<td>Study, 208</td>
<td></td>
</tr>
<tr>
<td>Submarin, 208</td>
<td></td>
</tr>
<tr>
<td>Subway, 209</td>
<td></td>
</tr>
<tr>
<td>Sunspot, 209</td>
<td></td>
</tr>
<tr>
<td>Superbowl, 210</td>
<td></td>
</tr>
<tr>
<td>Supercar, 211</td>
<td></td>
</tr>
<tr>
<td>Tablrock, 212</td>
<td></td>
</tr>
<tr>
<td>Teacher, 213</td>
<td></td>
</tr>
<tr>
<td>Tenness, 213</td>
<td></td>
</tr>
<tr>
<td>Tensile, 214</td>
<td></td>
</tr>
<tr>
<td>Testi, 215</td>
<td></td>
</tr>
<tr>
<td>Thermal, 215</td>
<td></td>
</tr>
<tr>
<td>Tiaa, 216</td>
<td></td>
</tr>
<tr>
<td>Ticket, 217</td>
<td></td>
</tr>
<tr>
<td>Toaster, 218</td>
<td></td>
</tr>
<tr>
<td>Tonsils, 218</td>
<td></td>
</tr>
</tbody>
</table>
Tort, 219
Toxic, 220
Track, 221
Track15, 222
Treatments, 222
Trees, 223
Trucks, 224
Tv, 227
Twin, 228
Undergrad, 229
Vacation, 230
Vaccine, 230
Vehicle, 231
Verbal, 232
Victoria, 232
Viscositis, 233
Visual, 234
Vocab, 234
Wastewater, 235
Weather94, 236
Wheat, 236
Windmill, 237
Window, 238
Wins, 238
Wool, 239
Yearsunspot, 240

*Topic distribution
  CIsim, 49
  Combinations, 54
  normarea, 148
  ntester, 149
  SRS, 202

*Topic htest
  SIGN.test, 188
  tsum.test, 224
  z.test, 241
  zsum.test, 243

*Topic package
  BSDA-package, 8

*Topic univar
  EDA, 74
  nsizese, 148

Abbey, 8
Abc, 9
Abilene, 10
Ability, 10
Abortion, 11
Absent, 12
Achieve, 13
Adsales, 13
Aggress, 14
Aid, 15

Aids, 15
Airdisasters, 16
Airline, 17
Alcohol, 18
Allergy, 18
Anesthesia, 19
Anxiety, 20
Apolipop, 20
Append, 21
Appendec, 22
Aptitude, 22
Archeo, 23
Arthritis, 24
Artifici, 25
Aspin, 25
Asthmati, 26
Attorney, 27
Autogear, 28
Backtoback, 28
Bbsalararies, 29
Bigten, 30
Biology, 30
Birth, 31
Blackedu, 32
Blood, 33
Board, 33
Bones, 34
Books, 35
Bookstor, 35
Brain, 36
BSDA (BSDA-package), 8
BSDA-package, 8
Bumpers, 37
Bus, 37
Bypass, 38
Cabinets, 39
Cancer, 40
Carbon, 40
Cat, 41
Censored, 42
Challeng, 42
Chemist, 43
Chesapeake, 44
Chevy, 45
Chicken, 45
Chipavg, 46
Chips, 47
Cigar, 48
Cigarette, 48
CIsim, 49
Citrus, 50
INDEX

Clean, 51
Coaxial, 52
Coffee, 52
Coins, 53
Combinations, 54, 203
Commute, 54
Concept, 55
Concrete, 56
Corn, 56
Correlat, 57
Counsel, 58
Cpi, 58
Crime, 59
Darwin, 60
Dealers, 60
Defectiv, 61
Degree, 62
Delay, 62
Depend, 63
Detroit, 64
Develop, 64
Devmath, 65
Dice, 65
Diesel, 66
Diplomat, 67
Disposal, 68
Dogs, 68
Domestic, 69
Dopamine, 70
Dowjones, 70
Drink, 71
Drug, 72
Dyslexia, 72
Earthqk, 73
EDA, 74
Educat, 75
Eggs, 75
Elderly, 76
Energy, 77
Engineer, 78
Entrance, 78
Epaminicompact, 79
Epatwoseater, 80
Executiv, 81
Exercise, 81
Fabric, 82
Faithful, 83
Family, 83
Ferraro1, 84
Ferraro2, 85
Fertility, 85
Firstchi, 86
Fish, 87
Fitness, 87
Florida2000, 88
Fluid, 89
Food, 90
Framingh, 90
Freshman, 91
Funeral, 92
Galaxie, 92
Gallup, 93
Gasoline, 94
German, 95
Golf, 95
Governor, 96
Gpa, 97
Graduate, 98
Greenriv, 99
Grrriv2, 99
Groupabc, 100
Groups, 101
Gym, 101
Habits, 102
Haptologo, 103
Hardware, 103
Hardwood, 104
Heat, 104
Heating, 105
Hodgkin, 106
Homes, 107
Homework, 108
Honda, 108
Hostile, 109
Housing, 110
Hurrican, 111
Iceberg, 112
Income, 112
Independent, 113
Indian, 114
Indiapol, 115
Indy500, 115
Inflatio, 116
Inletoil, 117
Inmate, 117
Inspect, 118
Insulate, 119
Iqgpa, 119
Irises, 120
INDEX

Jdpower, 121
Jobsat, 121
Kidsmoke, 122
Kilowatt, 123
Kinder, 123
Laminect, 124
Lead, 125
Leader, 125
Lethal, 126
Life, 127
Lifespan, 127
Lightmonth, 128
Lodge, 129
Longtail, 130
Lowabil, 130
Magnesiu, 131
Malpract, 132
Manager, 132
Marked, 133
MATH, 134
Mathcomp, 134
Mathpro, 135
Maze, 136
Median, 136
Mental, 137
Mercury, 138
Metrent, 138
Miller, 139
Miller1, 139
Moisture, 140
Monoxide, 141
Movie, 141
Music, 142
Name, 143
Nascar, 143
Nervous, 144
Newsstand, 145
Nfldraft2, 145
Nfldraft, 146
Nicotine, 147
normarea, 148
nsize, 148
ntester, 149
Orange, 150
Orioles, 151
Oxytocin, 151
Parented, 152
Patrol, 153
Pearson, 153
Phone, 154
Poison, 155
Politc, 155
Pollutio, 156
Porosity, 156
Poverty, 157
Precinct, 158
Prejudic, 159
Presiden, 159
Press, 160
Prognost, 161
Program, 161
Psat, 162
Psych, 163
Puerto, 163
Quail, 164
Quality, 164
Rainks, 165
Randd, 166
Rat, 167
Ratings, 167
Reaction, 168
Reading, 169
Readiq, 169
Referend, 170
Region, 171
Register, 171
Rehab, 172
Remedial, 173
Rentals, 174
Repair, 174
Retail, 175
Ronbrown1, 175
Ronbrown2, 176
Rural, 177
Salary, 178
Salinity, 178
Sat, 179
Saving, 180
Scales, 180
Schizop2, 181
Schizoph, 182
Seatbelt, 182
Selfdefe, 183
Senior, 184
Sentence, 184
Shkdrug, 185
Shock, 186
Shoplift, 186
<table>
<thead>
<tr>
<th>INDEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short, 187</td>
</tr>
<tr>
<td>Shuttle, 188</td>
</tr>
<tr>
<td>SIGN.test, 188</td>
</tr>
<tr>
<td>Simpson, 191</td>
</tr>
<tr>
<td>Situp, 192</td>
</tr>
<tr>
<td>Skewed, 192</td>
</tr>
<tr>
<td>Skin, 193</td>
</tr>
<tr>
<td>S1c, 194</td>
</tr>
<tr>
<td>Smokyp, 194</td>
</tr>
<tr>
<td>Snore, 195</td>
</tr>
<tr>
<td>Snow, 196</td>
</tr>
<tr>
<td>Soccer, 196</td>
</tr>
<tr>
<td>Social, 197</td>
</tr>
<tr>
<td>Sophomor, 198</td>
</tr>
<tr>
<td>South, 198</td>
</tr>
<tr>
<td>Speed, 199</td>
</tr>
<tr>
<td>Spellers, 200</td>
</tr>
<tr>
<td>Spelling, 200</td>
</tr>
<tr>
<td>Sports, 201</td>
</tr>
<tr>
<td>Spouse, 202</td>
</tr>
<tr>
<td>SRS, 54, 202</td>
</tr>
<tr>
<td>Stable, 203</td>
</tr>
<tr>
<td>Stamp, 204</td>
</tr>
<tr>
<td>Statclas, 204</td>
</tr>
<tr>
<td>Statelaw, 205</td>
</tr>
<tr>
<td>Statisti, 206</td>
</tr>
<tr>
<td>Step, 206</td>
</tr>
<tr>
<td>Stress, 207</td>
</tr>
<tr>
<td>Study, 208</td>
</tr>
<tr>
<td>Submarin, 208</td>
</tr>
<tr>
<td>Subway, 209</td>
</tr>
<tr>
<td>Sunspot, 209</td>
</tr>
<tr>
<td>Superbowl, 210</td>
</tr>
<tr>
<td>Supercar, 211</td>
</tr>
<tr>
<td>Tablrock, 212</td>
</tr>
<tr>
<td>Teacher, 213</td>
</tr>
<tr>
<td>Tenness, 213</td>
</tr>
<tr>
<td>Tensile, 214</td>
</tr>
<tr>
<td>Test1, 215</td>
</tr>
<tr>
<td>Thermal, 215</td>
</tr>
<tr>
<td>Tiaa, 216</td>
</tr>
<tr>
<td>Ticket, 217</td>
</tr>
<tr>
<td>Toaster, 218</td>
</tr>
<tr>
<td>Tonsils, 218</td>
</tr>
<tr>
<td>Tort, 219</td>
</tr>
<tr>
<td>Toxic, 220</td>
</tr>
<tr>
<td>Track, 221</td>
</tr>
<tr>
<td>Track15, 222</td>
</tr>
<tr>
<td>Treatments, 222</td>
</tr>
<tr>
<td>Trees, 223</td>
</tr>
<tr>
<td>Trucks, 224</td>
</tr>
<tr>
<td>tsum.test, 190, 224, 243, 245</td>
</tr>
<tr>
<td>Tv, 227</td>
</tr>
<tr>
<td>Twin, 228</td>
</tr>
<tr>
<td>Undergrad, 229</td>
</tr>
<tr>
<td>Vacation, 230</td>
</tr>
<tr>
<td>Vaccine, 230</td>
</tr>
<tr>
<td>Vehicle, 231</td>
</tr>
<tr>
<td>Verbal, 232</td>
</tr>
<tr>
<td>Victoria, 232</td>
</tr>
<tr>
<td>Viscosit, 233</td>
</tr>
<tr>
<td>Visual, 234</td>
</tr>
<tr>
<td>Vocab, 234</td>
</tr>
<tr>
<td>Wastewat, 235</td>
</tr>
<tr>
<td>Weather94, 236</td>
</tr>
<tr>
<td>Wheat, 236</td>
</tr>
<tr>
<td>Windmill, 237</td>
</tr>
<tr>
<td>Window, 238</td>
</tr>
<tr>
<td>Wins, 238</td>
</tr>
<tr>
<td>Wool, 239</td>
</tr>
<tr>
<td>Yearsunspot, 240</td>
</tr>
<tr>
<td>z.test, 190, 227, 241, 245</td>
</tr>
<tr>
<td>zsum.test, 190, 227, 243, 243</td>
</tr>
</tbody>
</table>