mbi <- read.csv(choose.file(),header=TRUE)

read.table(file = file, header = header, sep = sep, quote = quote, :

could not find function "choose.file"

mbi <- read.csv(choose.files(),header=TRUE)

# The file name here is 'MassBodilyInjury'

# Anything after the number sign is a comment and not run as a R command

# choose.files will go find the dataset, header=TRUE is for the 1st row

# of variable names

# R is case sensitive

mbi

rownum claims provider providerA logclaims

1 1 0.045 Other 1 -3.101

2 2 0.047 Other 1 -3.058

3 3 0.070 Other 1 -2.659

4 4 0.075 Other 1 -2.590

5 5 0.077 Other 1 -2.564

6 6 0.092 Other 1 -2.386

7 7 0.117 Other 1 -2.146

8 8 0.117 Other 1 -2.146

9 9 0.140 Other 1 -1.966

10 10 0.145 Other 1 -1.931

11 11 0.149 Other 1 -1.904

12 12 0.165 Other 1 -1.802

13 13 0.167 Other 1 -1.790

14 14 0.169 Other 1 -1.778

15 15 0.180 Other 1 -1.715

16 16 0.180 Other 1 -1.715

17 17 0.199 Other 1 -1.614

18 18 0.202 Other 1 -1.599

19 19 0.212 Other 1 -1.551

20 20 0.225 Other 1 -1.492

21 21 0.230 Other 1 -1.470

22 22 0.242 Other 1 -1.419

23 23 0.264 Other 1 -1.332

24 24 0.275 Other 1 -1.291

25 25 0.285 Other 1 -1.255

26 26 0.290 Other 1 -1.238

27 27 0.363 Other 1 -1.013

28 28 0.384 Other 1 -0.957

29 29 0.400 Other 1 -0.916

30 30 0.400 Other 1 -0.916

31 31 0.413 Other 1 -0.884

32 32 0.414 Other 1 -0.882

33 33 0.416 Other 1 -0.877

34 34 0.425 Other 1 -0.856

35 35 0.425 Other 1 -0.856

36 36 0.430 Other 1 -0.844

37 37 0.430 A 0 -0.844

38 38 0.431 Other 1 -0.842

39 39 0.450 Other 1 -0.799

40 40 0.460 Other 1 -0.777

41 41 0.486 Other 1 -0.722

42 42 0.500 Other 1 -0.693

43 43 0.500 Other 1 -0.693

44 44 0.514 A 0 -0.666

45 45 0.531 Other 1 -0.633

46 46 0.540 Other 1 -0.616

47 47 0.556 Other 1 -0.587

48 48 0.564 Other 1 -0.573

49 49 0.600 Other 1 -0.511

50 50 0.605 Other 1 -0.503

51 51 0.605 Other 1 -0.503

52 52 0.650 Other 1 -0.431

53 53 0.660 Other 1 -0.416

54 54 0.660 Other 1 -0.416

55 55 0.685 Other 1 -0.378

56 56 0.690 Other 1 -0.371

57 57 0.698 Other 1 -0.360

58 58 0.700 Other 1 -0.357

59 59 0.705 Other 1 -0.350

60 60 0.725 Other 1 -0.322

61 61 0.740 Other 1 -0.301

62 62 0.750 Other 1 -0.288

63 63 0.780 Other 1 -0.248

64 64 0.785 Other 1 -0.242

65 65 0.785 Other 1 -0.242

66 66 0.806 Other 1 -0.216

67 67 0.825 Other 1 -0.192

68 68 0.825 Other 1 -0.192

69 69 0.830 Other 1 -0.186

70 70 0.836 A 0 -0.179

71 71 0.870 Other 1 -0.139

72 72 0.900 Other 1 -0.105

73 73 0.934 Other 1 -0.068

74 74 0.950 Other 1 -0.051

75 75 0.954 Other 1 -0.047

76 76 0.956 Other 1 -0.045

77 77 0.962 Other 1 -0.039

78 78 0.970 Other 1 -0.030

79 79 0.975 Other 1 -0.025

80 80 0.988 Other 1 -0.012

81 81 1.015 Other 1 0.015

82 82 1.053 Other 1 0.052

83 83 1.058 Other 1 0.056

84 84 1.080 Other 1 0.077

85 85 1.161 Other 1 0.149

86 86 1.167 Other 1 0.154

87 87 1.195 Other 1 0.178

88 88 1.215 Other 1 0.195

89 89 1.242 Other 1 0.217

90 90 1.260 Other 1 0.231

91 91 1.295 Other 1 0.259

92 92 1.310 Other 1 0.270

93 93 1.319 Other 1 0.277

94 94 1.330 Other 1 0.285

95 95 1.340 Other 1 0.293

96 96 1.355 Other 1 0.304

97 97 1.390 Other 1 0.329

98 98 1.444 Other 1 0.367

99 99 1.455 Other 1 0.375

100 100 1.463 Other 1 0.380

101 101 1.490 Other 1 0.399

102 102 1.500 Other 1 0.405

103 103 1.542 Other 1 0.433

104 104 1.598 Other 1 0.469

105 105 1.616 Other 1 0.480

106 106 1.623 Other 1 0.484

107 107 1.640 Other 1 0.495

108 108 1.645 A 0 0.498

109 109 1.650 Other 1 0.501

110 110 1.660 Other 1 0.507

111 111 1.680 Other 1 0.519

112 112 1.695 Other 1 0.528

113 113 1.700 A 0 0.531

114 114 1.758 Other 1 0.564

115 115 1.759 Other 1 0.565

116 116 1.760 Other 1 0.565

117 117 1.896 Other 1 0.640

118 118 1.920 Other 1 0.652

119 119 1.923 Other 1 0.654

120 120 1.941 A 0 0.663

121 121 1.960 Other 1 0.673

122 122 1.972 Other 1 0.679

123 123 1.990 Other 1 0.688

124 124 2.005 Other 1 0.696

125 125 2.018 Other 1 0.702

126 126 2.020 Other 1 0.703

127 127 2.020 Other 1 0.703

128 128 2.030 Other 1 0.708

129 129 2.042 A 0 0.714

130 130 2.062 Other 1 0.724

131 131 2.063 Other 1 0.724

132 132 2.080 Other 1 0.732

133 133 2.087 Other 1 0.736

134 134 2.089 Other 1 0.737

135 135 2.100 Other 1 0.742

136 136 2.115 Other 1 0.749

137 137 2.120 Other 1 0.751

138 138 2.155 Other 1 0.768

139 139 2.159 Other 1 0.770

140 140 2.161 Other 1 0.771

141 141 2.184 A 0 0.781

142 142 2.188 Other 1 0.783

143 143 2.191 Other 1 0.784

144 144 2.196 Other 1 0.787

145 145 2.224 Other 1 0.799

146 146 2.237 Other 1 0.805

147 147 2.251 A 0 0.811

148 148 2.253 Other 1 0.812

149 149 2.288 Other 1 0.828

150 150 2.295 Other 1 0.831

151 151 2.318 Other 1 0.841

152 152 2.325 Other 1 0.844

153 153 2.325 A 0 0.844

154 154 2.335 Other 1 0.848

155 155 2.341 Other 1 0.851

156 156 2.350 Other 1 0.854

157 157 2.374 Other 1 0.865

158 158 2.390 Other 1 0.871

159 159 2.406 Other 1 0.878

160 160 2.434 Other 1 0.890

161 161 2.450 Other 1 0.896

162 162 2.453 A 0 0.897

163 163 2.468 Other 1 0.903

164 164 2.468 A 0 0.903

165 165 2.480 Other 1 0.908

166 166 2.480 A 0 0.908

167 167 2.490 A 0 0.912

168 168 2.498 Other 1 0.915

169 169 2.500 Other 1 0.916

170 170 2.500 Other 1 0.916

171 171 2.500 A 0 0.916

172 172 2.510 Other 1 0.920

173 173 2.532 Other 1 0.929

174 174 2.540 Other 1 0.932

175 175 2.543 Other 1 0.933

176 176 2.559 Other 1 0.940

177 177 2.572 Other 1 0.945

178 178 2.593 Other 1 0.953

179 179 2.601 Other 1 0.956

180 180 2.616 Other 1 0.962

181 181 2.619 Other 1 0.963

182 182 2.630 Other 1 0.967

183 183 2.635 Other 1 0.969

184 184 2.635 Other 1 0.969

185 185 2.653 Other 1 0.976

186 186 2.655 Other 1 0.976

187 187 2.675 Other 1 0.984

188 188 2.679 Other 1 0.985

189 189 2.697 Other 1 0.992

190 190 2.718 Other 1 1.000

191 191 2.730 Other 1 1.004

192 192 2.734 Other 1 1.006

193 193 2.755 Other 1 1.013

194 194 2.758 Other 1 1.015

195 195 2.773 Other 1 1.020

196 196 2.775 Other 1 1.021

197 197 2.780 Other 1 1.022

198 198 2.785 A 0 1.024

199 199 2.795 Other 1 1.028

200 200 2.805 Other 1 1.031

201 201 2.805 Other 1 1.031

202 202 2.808 A 0 1.032

203 203 2.880 Other 1 1.058

204 204 2.881 Other 1 1.058

205 205 2.881 A 0 1.058

206 206 2.924 A 0 1.073

207 207 2.930 Other 1 1.075

208 208 2.934 A 0 1.076

209 209 2.940 Other 1 1.078

210 210 2.972 Other 1 1.089

211 211 2.975 Other 1 1.090

212 212 3.000 Other 1 1.099

213 213 3.000 A 0 1.099

214 214 3.025 Other 1 1.107

215 215 3.058 Other 1 1.118

216 216 3.082 A 0 1.126

217 217 3.085 Other 1 1.127

218 218 3.095 Other 1 1.130

219 219 3.100 Other 1 1.131

220 220 3.102 A 0 1.132

221 221 3.106 Other 1 1.133

222 222 3.135 Other 1 1.143

223 223 3.170 Other 1 1.154

224 224 3.187 Other 1 1.159

225 225 3.192 A 0 1.161

226 226 3.193 Other 1 1.161

227 227 3.200 Other 1 1.163

228 228 3.210 Other 1 1.166

229 229 3.230 Other 1 1.172

230 230 3.230 Other 1 1.172

231 231 3.230 A 0 1.172

232 232 3.232 Other 1 1.173

233 233 3.235 Other 1 1.174

234 234 3.243 A 0 1.176

235 235 3.248 A 0 1.178

236 236 3.249 Other 1 1.178

237 237 3.260 Other 1 1.182

238 238 3.261 Other 1 1.182

239 239 3.272 A 0 1.185

240 240 3.290 Other 1 1.191

241 241 3.295 Other 1 1.192

242 242 3.304 Other 1 1.195

243 243 3.332 A 0 1.204

244 244 3.333 Other 1 1.204

245 245 3.338 Other 1 1.205

246 246 3.340 Other 1 1.206

247 247 3.341 A 0 1.206

248 248 3.349 A 0 1.209

249 249 3.349 A 0 1.209

250 250 3.349 A 0 1.209

251 251 3.353 A 0 1.210

252 252 3.360 Other 1 1.212

253 253 3.378 A 0 1.217

254 254 3.385 A 0 1.219

255 255 3.387 A 0 1.220

256 256 3.416 Other 1 1.228

257 257 3.429 A 0 1.232

258 258 3.438 A 0 1.235

259 259 3.444 A 0 1.237

260 260 3.469 A 0 1.244

261 261 3.473 A 0 1.245

262 262 3.473 A 0 1.245

263 263 3.475 A 0 1.246

264 264 3.477 A 0 1.246

265 265 3.505 Other 1 1.254

266 266 3.517 A 0 1.258

267 267 3.518 Other 1 1.258

268 268 3.527 A 0 1.260

269 269 3.535 A 0 1.263

270 270 3.547 A 0 1.266

271 271 3.550 Other 1 1.267

272 272 3.552 Other 1 1.268

273 273 3.567 A 0 1.272

274 274 3.570 Other 1 1.273

275 275 3.575 Other 1 1.274

276 276 3.580 Other 1 1.275

277 277 3.583 A 0 1.276

278 278 3.590 A 0 1.278

279 279 3.603 A 0 1.282

280 280 3.615 A 0 1.285

281 281 3.623 A 0 1.287

282 282 3.647 A 0 1.294

283 283 3.655 Other 1 1.296

284 284 3.655 A 0 1.296

285 285 3.658 Other 1 1.297

286 286 3.675 Other 1 1.302

287 287 3.675 Other 1 1.302

288 288 3.687 A 0 1.305

289 289 3.720 Other 1 1.314

290 290 3.720 Other 1 1.314

291 291 3.742 Other 1 1.320

292 292 3.757 A 0 1.324

293 293 3.765 Other 1 1.326

294 294 3.800 A 0 1.335

295 295 3.809 Other 1 1.337

296 296 3.848 A 0 1.348

297 297 3.857 A 0 1.350

298 298 3.867 Other 1 1.352

299 299 3.867 A 0 1.352

300 300 3.870 Other 1 1.353

301 301 3.883 Other 1 1.357

302 302 3.890 Other 1 1.358

303 303 3.905 A 0 1.362

304 304 3.907 A 0 1.363

305 305 4.000 Other 1 1.386

306 306 4.011 Other 1 1.389

307 307 4.039 A 0 1.396

308 308 4.065 A 0 1.402

309 309 4.095 Other 1 1.410

310 310 4.134 Other 1 1.419

311 311 4.147 Other 1 1.422

312 312 4.155 A 0 1.424

313 313 4.170 Other 1 1.428

314 314 4.179 A 0 1.430

315 315 4.200 Other 1 1.435

316 316 4.215 Other 1 1.439

317 317 4.257 A 0 1.449

318 318 4.300 Other 1 1.459

319 319 4.489 A 0 1.502

320 320 4.593 A 0 1.525

321 321 4.595 Other 1 1.525

322 322 4.630 A 0 1.533

323 323 4.653 Other 1 1.538

324 324 4.700 A 0 1.548

325 325 4.731 Other 1 1.554

326 326 4.741 A 0 1.556

327 327 4.750 Other 1 1.558

328 328 4.761 Other 1 1.560

329 329 4.810 Other 1 1.571

330 330 5.072 Other 1 1.624

331 331 5.161 Other 1 1.641

332 332 5.240 Other 1 1.656

333 333 5.640 Other 1 1.730

334 334 5.779 Other 1 1.754

335 335 6.166 Other 1 1.819

336 336 6.406 Other 1 1.857

337 337 6.725 Other 1 1.906

338 338 7.717 A 0 2.043

339 339 8.000 Other 1 2.079

340 340 9.500 Other 1 2.251

341 341 10.295 Other 1 2.332

342 342 12.533 Other 1 2.528

343 343 12.688 Other 1 2.541

344 344 16.043 Other 1 2.775

345 345 18.847 Other 1 2.936

346 346 19.500 Other 1 2.970

347 347 20.827 Other 1 3.036

348 348 50.000 Other 1 3.912

attach(mbi)

#u have to attach to tell R which dataset to run all future command on

summary(mbi)

rownum claims provider providerA

Min. : 1.00 Min. : 0.045 A : 76 Min. :0.0000

1st Qu.: 87.75 1st Qu.: 1.210 Other:272 1st Qu.:1.0000

Median :174.50 Median : 2.542 Median :1.0000

Mean :174.50 Mean : 2.830 Mean :0.7816

3rd Qu.:261.25 3rd Qu.: 3.473 3rd Qu.:1.0000

Max. :348.00 Max. :50.000 Max. :1.0000

logclaims

Min. :-3.1010

1st Qu.: 0.1908

Median : 0.9325

Mean : 0.6225

3rd Qu.: 1.2450

Max. : 3.9120

# summary(of entire dataset)

summary(claims)

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.045 1.210 2.542 2.830 3.473 50.000

sd(claims)

[1] 3.52524

# sd for standard dev

mean(claim)

Error in mean(claim) : object 'claim' not found

min(claim)

Error: object 'claim' not found

means(claim)

Error: could not find function "means"

mean(claims)

[1] 2.830006

min(claims)

[1] 0.045

max(claims)

[1] 50

quantile(claims)

0% 25% 50% 75% 100%

0.0450 1.2100 2.5415 3.4730 50.0000

quantile(claims,.25)

25%

1.21

quantile(claims,.75)

75%

3.473

mu = mean(claims)

mu

[1] 2.830006

sd(mbi)

Error in is.data.frame(x) :

(list) object cannot be coerced to type 'double'

sd(mbi)

Error in is.data.frame(x) :

(list) object cannot be coerced to type 'double'

sd(mbi)

Error in is.data.frame(x) :

(list) object cannot be coerced to type 'double'

quantile(claims, .75

+

+

+ )

75%

3.473

Or hit the ESC key to force a line of code to end and start a new one

Error: unexpected symbol in "Or hit"

quantile(claims, .75

hist(claims)

# this is right skewed,which is not good

#This is figure 1.2 in the book

boxplot(claims)

boxplot(claims, horizontal=TRUE)

par(mfrow=c(1,2))

# this makes a graphics window with(1 row 2 col)

boxplot(claims)

boxplot(claims, horizontal=TRUE)

boxplot(claims, horizontal=TRUE)

boxplot(claims, horizontal=TRUE)

qqnorm(claims)

qqline(claims)

#This makes the QQ plot and puts the reference line on it

#We see again that claims data is not normally distributed

qqnorm(claims)

qqline(claims)

> mbi <- read.csv(choose.files(),header=TRUE)

> # "MassBodilyInjury'file

> attach(mbi)

> hist(claims)

> hist(log(claims))

> logclaims = log(claims)

> # logclaims variable

> hist(log10(claims))

> # log10 command is the log and the log command takes the ln

> # both of these histogramsare much better than the original claims

> his((sqrt(claims)

+ his((sqrt(claims))

Error: unexpected symbol in:

"his((sqrt(claims)

his"

> hist(sqrt(claims))

> hist(1/(claims))

> hist(-1/(claims))

> par(mfrow=c(1,2))

> boxplot(claims)

> boxplot(log(claims))