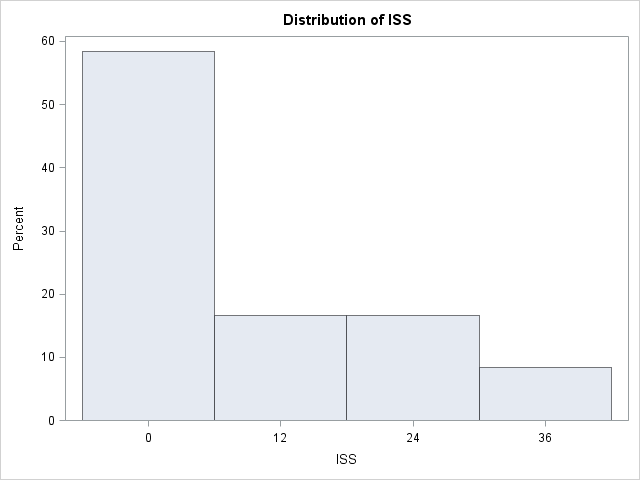
Dominic LaRoche

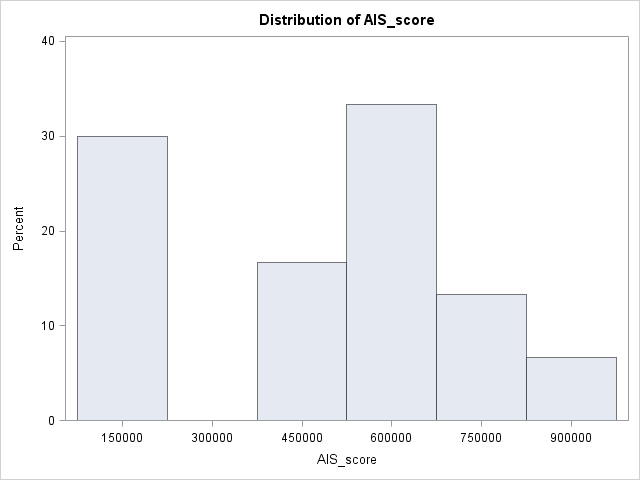
10/20/14- Assignment 6

CPH576D

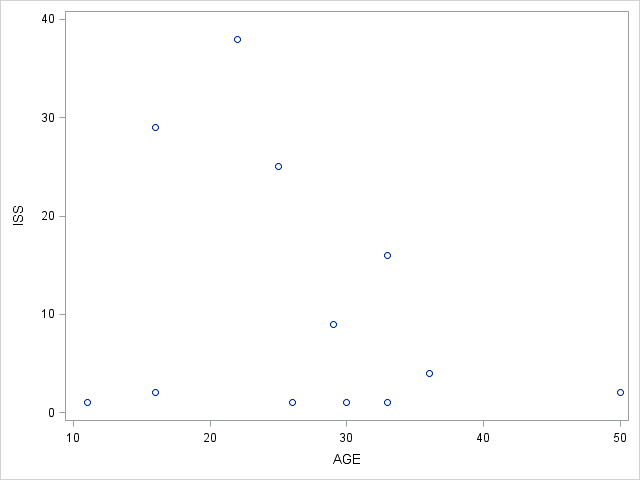
1. There are many problems with this data set! There are missing variables which assign data to a case ID. There is a composite field, AIS, that contains information that should be in two fields. Also, there are some rows which have no information at all.

6) Since a patient has only one ISS score we must be careful not to include the repeated scores that we have just filled in. This would create artificial density for ISS scores of patients with multiple AIS scores. I will only use the first ISS score for each patient ID.





7) Evaluating the scatter plot of age and ISS score shows that there might be a negative correlation between age and injury severity. Although, the relationship may actually be non-linear (and therefore the correlation would be close to 0) since it appears to be low at low ages, high in the late teens to late twenties and then low above about 35. We would need much more information to support such a non-linear trend.



SAS Code:

**data** ret;

set class.test\_trauma;

run;

/\*2)-4) separating out the score from the text \*/

**data** ret;

set ret;

AIS\_score= **.**;

if AIS ne "" then AIS\_score = substr(AIS,**1**,**6**);

if AIS ne "" then AIS\_txt = substrn(AIS,**8**);

drop AIS;

run;

**data** retc;

set ret;

retain xmrn xage xsex xrace xiss;

if AIS\_score = **.** then delete;

/\*Create retained variables\*/

if age ne **.** then do;

xmrn = mrn;

xage = age;

xsex = sex;

xrace = race;

xiss = iss;

end;

/\*Fill in missing information from retained variables\*/

if mrn = **.** then mrn = xmrn;

if age = **.** then age = xage;

if sex = "" then sex = xsex;

if race = "" then race = xrace;

if iss = **.** then iss = xiss;

drop xmrn xage xsex xrace xiss;

run;

/\*need to use only the first value of iss to avoid excess density at multiple obs\*/

**data** retfirst;

set retc;

by mrn;

if first.mrn;

run;

**proc** **univariate** data = retfirst;

var iss;

histogram;

**run**;

**proc** **univariate** data = retc;

var ais\_score;

histogram;

**run**;

/\*For the sgplot we must also use the first record data set\*/

**proc** **sgplot** data = retfirst;

scatter x=age y=iss;

**run**;

SAS Log File:

247 data ret;

248 set class.test\_trauma;

249 run;

NOTE: There were 41 observations read from the data set CLASS.TEST\_TRAUMA.

NOTE: The data set WORK.RET has 41 observations and 6 variables.

NOTE: DATA statement used (Total process time):

real time 0.01 seconds

cpu time 0.01 seconds

250 /\*2)-4) separating out the score from the text \*/

251 data ret;

252 set ret;

253 AIS\_score= .;

254 if AIS ne "" then AIS\_score = substr(AIS,1,6);

255 if AIS ne "" then AIS\_txt = substrn(AIS,8);

256 drop AIS;

257 run;

NOTE: Character values have been converted to numeric values at the places given by:

(Line):(Column).

254:31

NOTE: There were 41 observations read from the data set WORK.RET.

NOTE: The data set WORK.RET has 41 observations and 7 variables.

NOTE: DATA statement used (Total process time):

real time 0.01 seconds

cpu time 0.01 seconds

258

259 data retc;

260 set ret;

261 retain xmrn xage xsex xrace xiss;

262 if AIS\_score = . then delete;

263 /\*Create retained variables\*/

264 if age ne . then do;

265 xmrn = mrn;

266 xage = age;

267 xsex = sex;

268 xrace = race;

269 xiss = iss;

270 end;

271 /\*Fill in missing information from retained variables\*/

272 if mrn = . then mrn = xmrn;

273 if age = . then age = xage;

274 if sex = "" then sex = xsex;

275 if race = "" then race = xrace;

276 if iss = . then iss = xiss;

277 drop xmrn xage xsex xrace xiss;

278 run;

NOTE: Character values have been converted to numeric values at the places given by:

(Line):(Column).

272:4

NOTE: There were 41 observations read from the data set WORK.RET.

NOTE: The data set WORK.RETC has 30 observations and 7 variables.

NOTE: DATA statement used (Total process time):

real time 0.01 seconds

cpu time 0.01 seconds

279

280 /\*need to use only the first value of iss to avoid excess density at multiple obs\*/

281 data retfirst;

282 set retc;

283 by mrn;

284 if first.mrn;

285 run;

NOTE: There were 30 observations read from the data set WORK.RETC.

NOTE: The data set WORK.RETFIRST has 12 observations and 7 variables.

NOTE: DATA statement used (Total process time):

real time 0.01 seconds

cpu time 0.00 seconds

286

287

288 proc univariate data = retfirst;

289 var iss;

290 histogram;

291 run;

NOTE: PROCEDURE UNIVARIATE used (Total process time):

real time 0.18 seconds

cpu time 0.06 seconds

292

293 proc univariate data = retc;

294 var ais\_score;

295 histogram;

296 run;

NOTE: PROCEDURE UNIVARIATE used (Total process time):

real time 0.16 seconds

cpu time 0.06 seconds

297

298 /\*For the sgplot we must also use the first record data set\*/

299 proc sgplot data = retfirst;

300 scatter x=age y=iss;

301 run;

WARNING: SASUSER.TEMPLAT is not a template store! It will be ignored.

NOTE: PROCEDURE SGPLOT used (Total process time):

real time 0.15 seconds

cpu time 0.04 seconds

NOTE: There were 12 observations read from the data set WORK.RETFIRST.