Dave Kangiser

 Scott Stavley

 Week 2 R Assignment

I expected (here!) a brief narrative introduction to the data set (and the study!). since this is the first one, however, I did not downgrade anyone. next time (and especially for the data analysis project!) you definitely need this! see the data analysis project handout, and the prologue to my solution to this assignment for ideas.

Fig. 3. The male testicular volume jitterplot? shows that testicular volumes increase dramatically! with age. Testicular volume was apparently not recorder for males (with one exception) older than about 20; for those males, testvol was missing . were a considerable number of NA's in the testicular volume data among older males in the study that were not included in the graph.

Fig. 2. Scatter plot identifying amount of igf1 hormone in test subjects by age. There is a correlation between younger age is correlated with high igf1 hormone levels. ???

PULL ME DOWN…

NO – younger subjects also had low igf1! For this population, growth hormone shows near exponential growth from birth and peaks just before age 20, and then levels off (possibly a learn near flat relationship after age 20.

BTW, I love the title of this graph, and it’s ok for informal exchange, but I’m not sure there is a relationship between igf1 and ‘young love’… in fact, it looks like that might be one topic of the study!

Fig. 1 Boxplot showing the age ranging from early childhood to advanced adulthood for subjects whotook part in the study. The majority of subjects were aged near puberty; median age is about 15. This suggests the study is targeted at subjects near puberty.

Summary of juul dataset

 age menarche sex igf1 tanner

 Min. : 0.170 No :369 M :621 Min. : 25.0 I :515

 1st Qu.: 9.053 Yes :335 F :713 1st Qu.:202.2 II :103

 Median :12.560 NA's:635 NA's: 5 Median :313.5 III : 72

 Mean :15.095 Mean :340.2 IV : 81

 3rd Qu.:16.855 3rd Qu.:462.8 V :328

 Max. :83.000 Max. :915.0 NA's:240

 NA's : 5.000 NA's :321.0

 testvol

 Min. : 1.000

 1st Qu.: 1.000

 Median : 3.000

 Mean : 7.896

 3rd Qu.: 15.000

 Max. : 30.000

 NA's :859.000

Most of this data, being categorical, is not normally distributed. Categorical data is difficult to scatter plot. Since parameter measurements were not taken on all individuals, the data are skewed right .

Apparently, the focus of this research is development rather than degeneration based on the fact that testicular volume was taken on males aged 20 or younger. The young love plot shows the correlation of age and igf1. This is the only graph that plots the two non-categorical parameters together. While a linear regression line would not fit this data set, it appears there is a correlation between age and blood igf1 concentration.