Evan Hayduk Ryan Kruse

April 10, 2011

QQMethods

Juul Data Analysis

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 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

This is a summary of the dataset juul, we converted the categories 'menarche' and 'sex' from numerical data to categorical, since the latter fits better with what the data represents. Below are graphs and plots that, in our opinion, summarize the data and how it relates to the data set as a whole. Enjoy!

Min. : 1.000

1st Qu.: 1.000

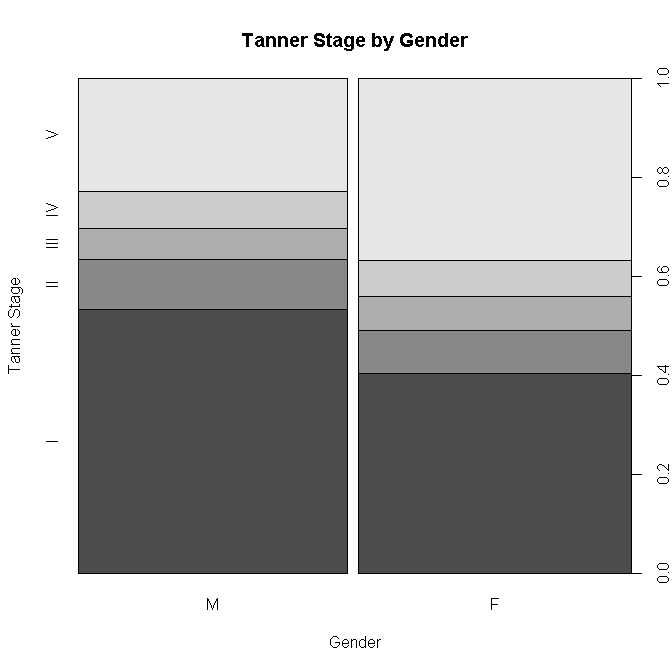
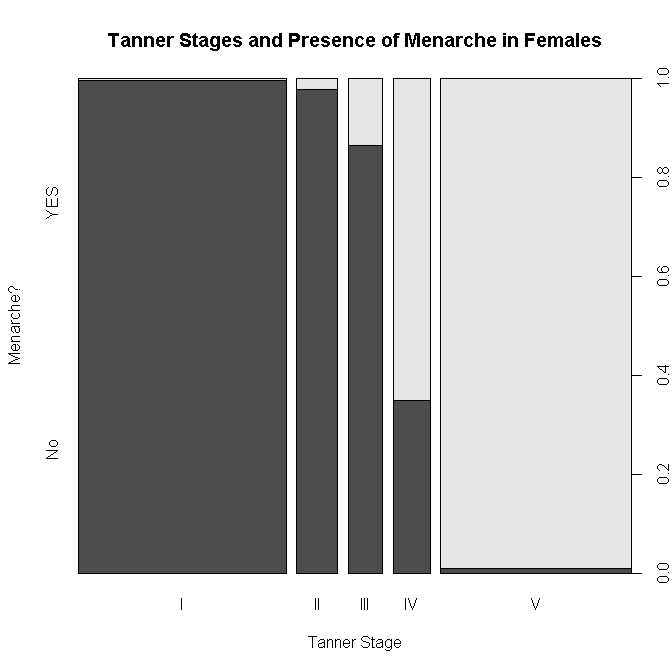
Median : 3.000

Mean : 7.896

3rd Qu.: 15.000

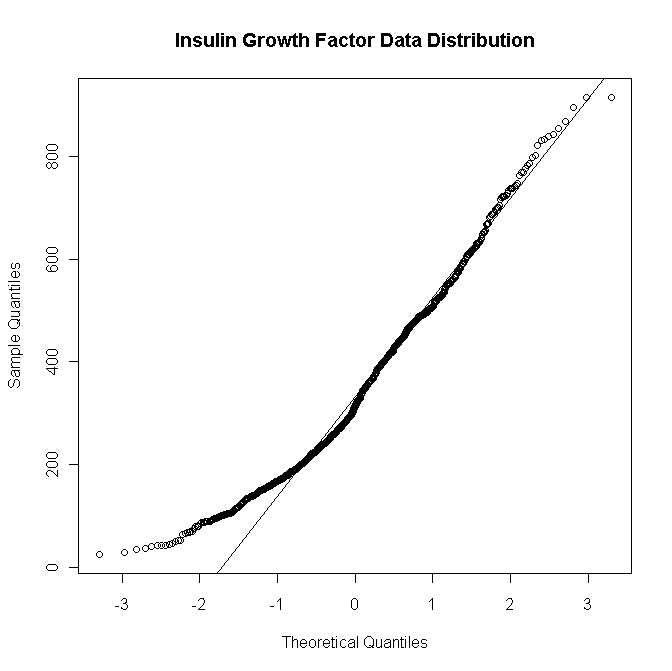
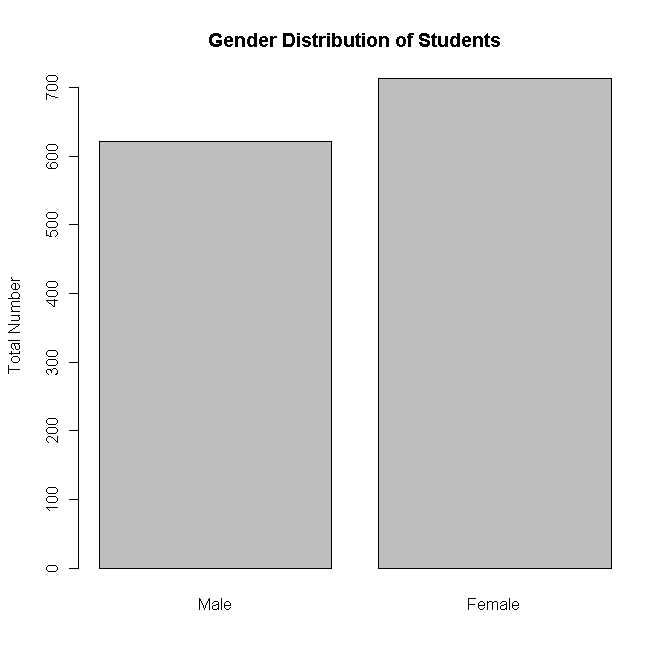
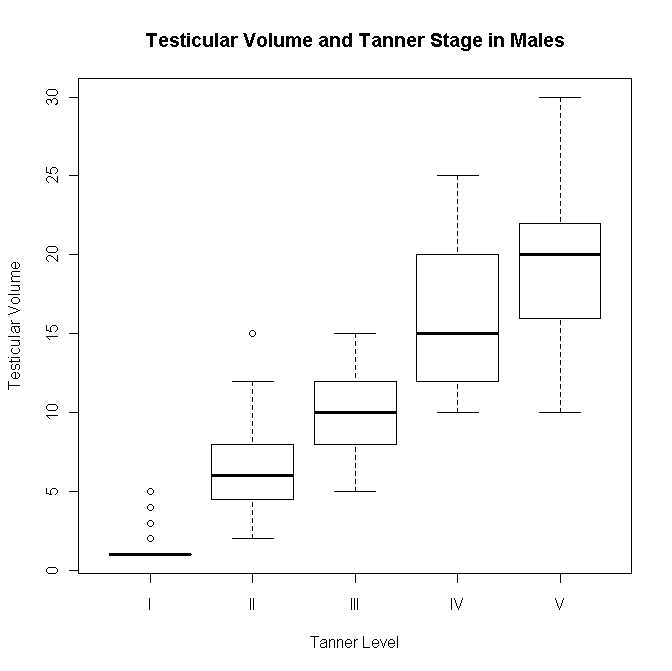
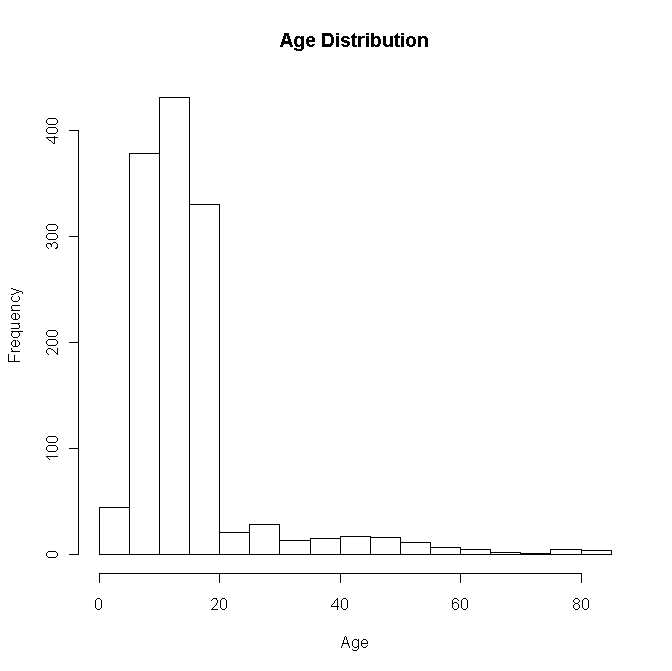
Max. : 30.000

NA's :859.000



This graph suggests that females mature (reach higher levels of tanner) earlier than males. However, this assumes? that the age distribution for the two groups were similar.

This graph shows that a higher percentage of females have reached menarche through each tanner stage. these two variables are correlated and co-predict. At tanner stage 1 almost no females have reached puberty; at stave V almost all have.



More females were tested than males

This QQ plot shows that insulin growth factor data seems to be close to normal distribution, with some skew at low igf1 values. The skew suggests that there are fewer individuals with low igf1 values in the sample population than one would expect, if the igf1 levels were normally distributed. This is expected, since there are fewer individuals in the lowest and highest age category, where one expect lower igf1 values. to see if this were true, judy plotted igf1 by age (at right).

The distribution of the data for age of participants is not a normal distribution.

When comparing tanner stage and testicular volume the result is what is to be expected

