References:

1) Test for the normal distribution of a sample

http://www.had2know.com/academics/test-for-normality-statistics.html

2) How Mann Whitney Test Works

http://www.graphpad.com/guides/prism/6/statistics/index.htm?how_the_mann-whitney_test_works.htm

3) OLS Linear Regression

https://datajobs.com/data-science-repo/OLS-Regression-[GD-Hutcheson].pdf

4) Scatter Plot

http://docs.ggplot2.org/0.9.3.1/geom_point.html

5) GGPlot for python

http://blog.yhathq.com/posts/ggplot-for-python.html

http://docs.ggplot2.org/current/scale_continuous.html

https://pypi.python.org/pypi/ggplot/

6) Copy of pandas Dataframe

http://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.copy.html http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy

7) Disadvantage of statistical analysis

http://www.ehow.com/info_8471854_disadvantages-statistical-analysis.html

8) Disadvantage of linear regression

http://www.ehow.com/info_8562780_disadvantages-linear-regression.html

9) One sided vs two sided p value

http://www.graphpad.com/guides/prism/6/statistics/index.htm?one-tail_vs_two-tail_p_values.htm

10) QQPlot

https://en.wikipedia.org/wiki/Q%E2%80%93Q_plot

11) How to call function from other python file.

http://stackoverflow.com/questions/7701646/how-to-call-a-function-from-another-file http://stackoverflow.com/questions/2325923/how-to-fix-importerror-no-module-named-error-in-python