EDA - stats and graphs

This assignment will help you learn some of the techniques commonly used in exploratory data analysis. The assignment can be done within a jupyter notebook, or submitted as a text document writeup (e.g. in Word, Google docs, etc, with summary stats and graphics) along with code you used to create the plots and statistics. If using a Jupyter notebook, you can change some of the cells from 'code' to 'markdown' to add discussion. Here's a cheatsheet for markdown, it's pretty easy to learn.

You can also post the assignment on Github, and send me the link. Github will naturally display jupyter notebooks and markdown within web browsers.

The written part of the assignment should not be very long, at most a few paragraphs for each of the criteria listed below.

Please also link to your data source you used for the assignment in your writeup/notebook.

Criteria	Specifications	Percent of assignment grade
Summary Stats	At least one of each of the following summary statistics has been shown, and have been discussed in writing: mean median max min standard deviation These summary stats should be from data that has been grouped (for example, grouped by loan grade for lendingclub data).	50
Plotting	At least one of each type of plot has been made: histogram bar boxplot scatter plot and at least one of the plots has been discussed in writing.	50
Bonus Statistics	Higher order summary statistics are calculated and discussed: skewness, kurtosis, L-moments, quartiles, correlation coefficients (Pearson, Spearman, etc)	5% extra credit
Bonus Plots	One or more of these plots is shown and discussed: treeplots, violin plots, heatmaps, 3d plots, geographic maps	5% extra credit