1. Yahoo Finance API was used. This was done because it is easier to connect with PDR and YF fix package to retrieve the data. There were problems connecting to Google Finance, but the YF fix shines and retrieves data correctly. Plus it has additional methods to replace the close with adj. close etc.

2. A list of five countries was chosen. One American, two European and two Asian. This was done so that a healthy mix of continental countries was taken, in order to find insights about movements of markets all around the world, and whether they were correlated or not. The countries are as follows:

- US
- UK
- France
- Hongkong
- Japan

3. Upto 10 years of data is loaded for the aforementioned markets, and their monthly returns are calculated.

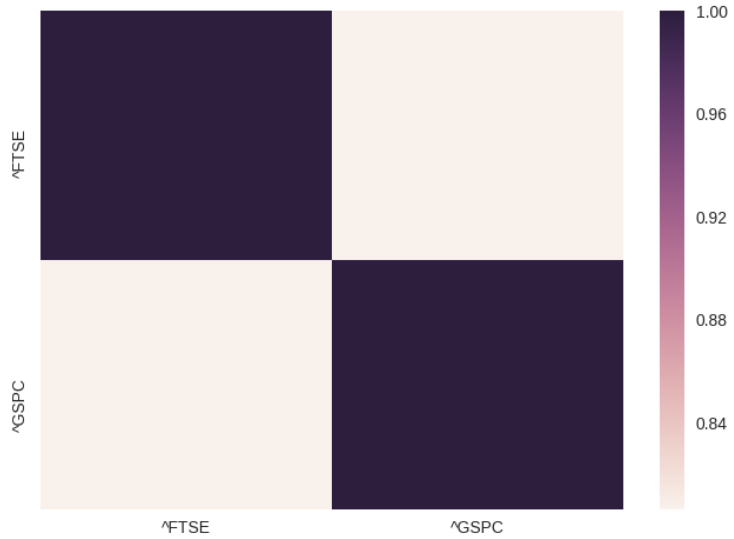
4. A correlation (Pearsons Correlation) is then calculated. There are two views which the user gets by the end of the program

- Granular View – This represents a correlation between two markets chosen from the aforementioned five by the user. This helps in analyzing the correlation between the chosen markets more closely
- Birds-Eye View – This gives an overall picture of all the five markets

4. There are 2 types of plots that are plotted after the program runs.

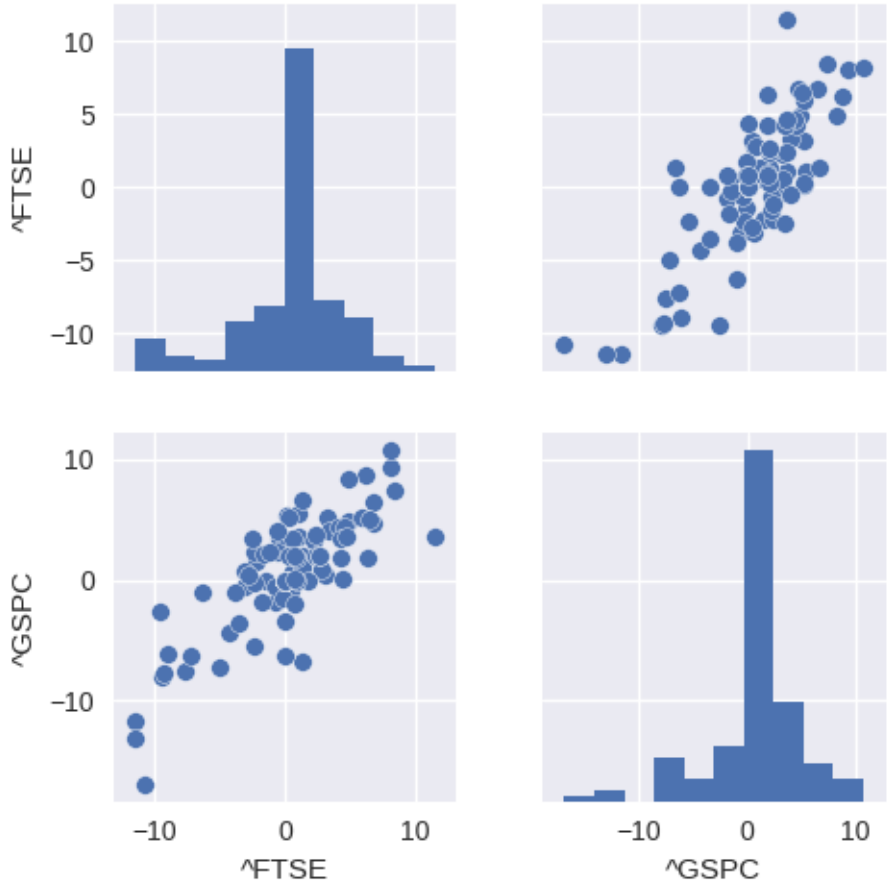
- Heatmap – A heatmap best represents how correlated the data is
- Scatter plot with Distribution – A scatter plot of the data points of and the Distribution (expected to be Normal for Pearsons Correlation)

The first plot that is encountered when the program is run is a heatmap. For this example, I took US and UK Markets, to see whether they are correlated:

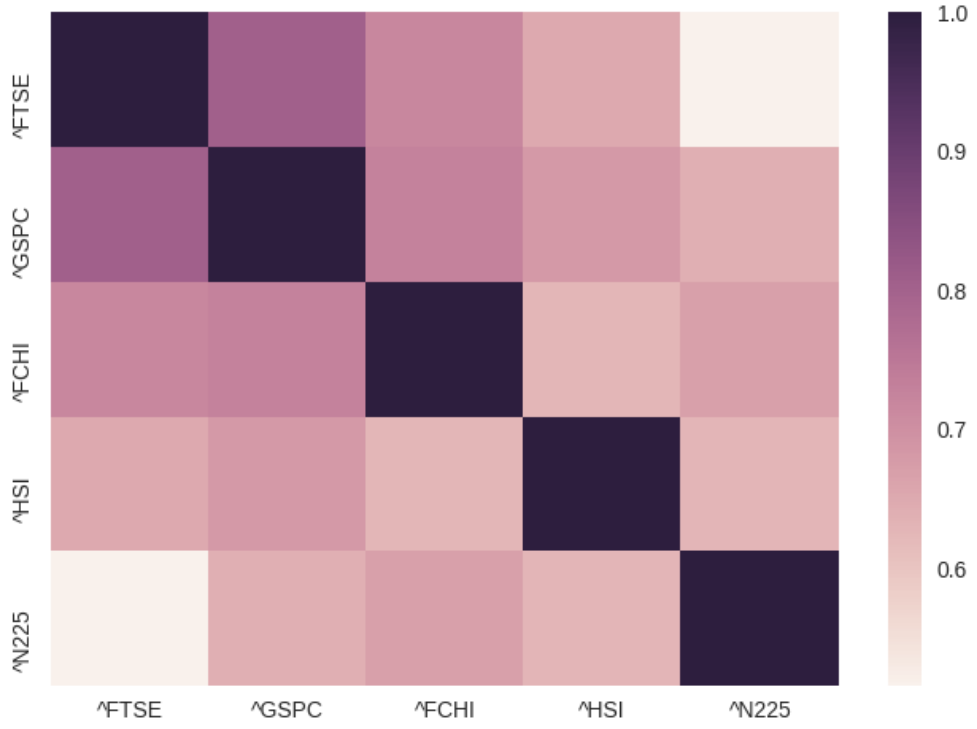


From the heatmap, if one looks at say, the bottom-left hand side, one can see that it's tan-ish in color, and the scale on the right indicates it's correlational strength. Here, it's around less than 0.84. One must not be fooled by the numbers, since this is just a 2x2 matrix plot, the Seaborn plot is showing a low 'color' one the scale, but in reality, the two markets move nearly 84% of the time in tandem. Which means that they are highly correlated.

The second plot which greets us is the scatter plot and the distribution plot. By looking at the distributions we can state that they are normally distributed, and by looking at the scatter we can say that they are somewhat linear (because of being nearly 84% correlated). An ^FTSE and ^FTSE scatter would be produce +1 as a correlational coefficient, and hence, would have be a linear diagonal line in scatter plot.

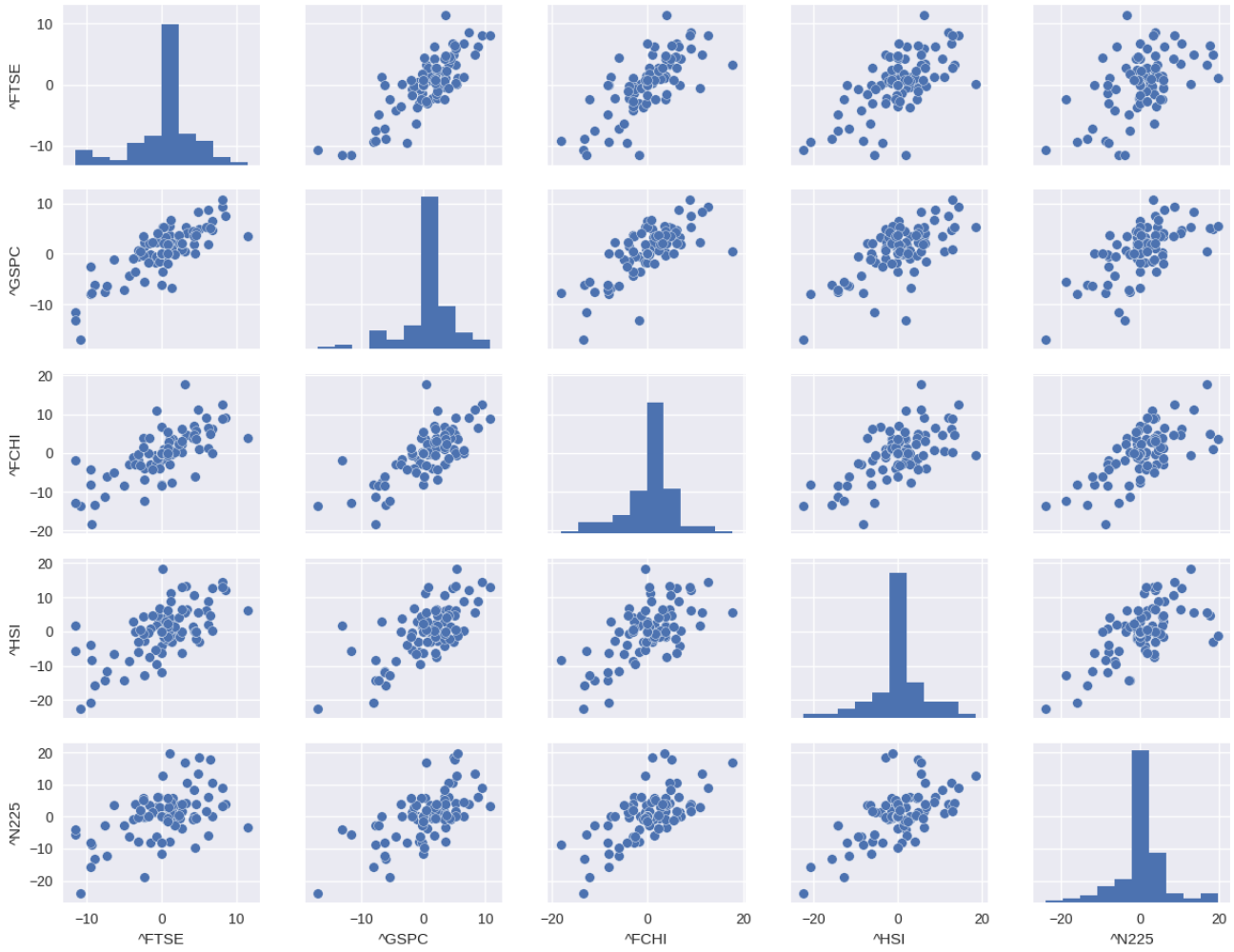


The third plot consists of a heatmap of all the 5 markets



By observing this plot, we come to know that the western markets are more correlated to each other than the easter markets. At the same time the eastern markets are more correlated to each other than their western counterparts. Like the FTSE and N225 are not as heavily correlated as, say, HSI and N225.

The forth plot is that of a scatter and distribution plot. This shows the distributions and the scatter is nearly linear in terms of highly correlated markets, and not so linear in terms of those that are not so highly correlated



Pearsons Correlation:

Pearsons correlation was used to determine the correlation between the markets. Pearson's correlation requires that each dataset be normally distributed. Like other correlation coefficients, this one varies between -1 and +1 with 0 implying no correlation. Correlations of -1 or +1 imply an exact linear relationship. Positive correlations imply that as x increases, so does y. Negative correlations imply that as x increases, y decreases. The p-value roughly indicates the probability of an uncorrelated system producing datasets that have a Pearson correlation at least as extreme as the one computed from these datasets