Statistical analysis

There are two python files...Network_locum.py and predict_filled.py

In Network_locum.py, the 3 data sets (sessions, ccgs and practices) are read as pandas dataframes, followed by data selection and merging.

predict_filled.py is a code to predict the likelihood of a job being filled

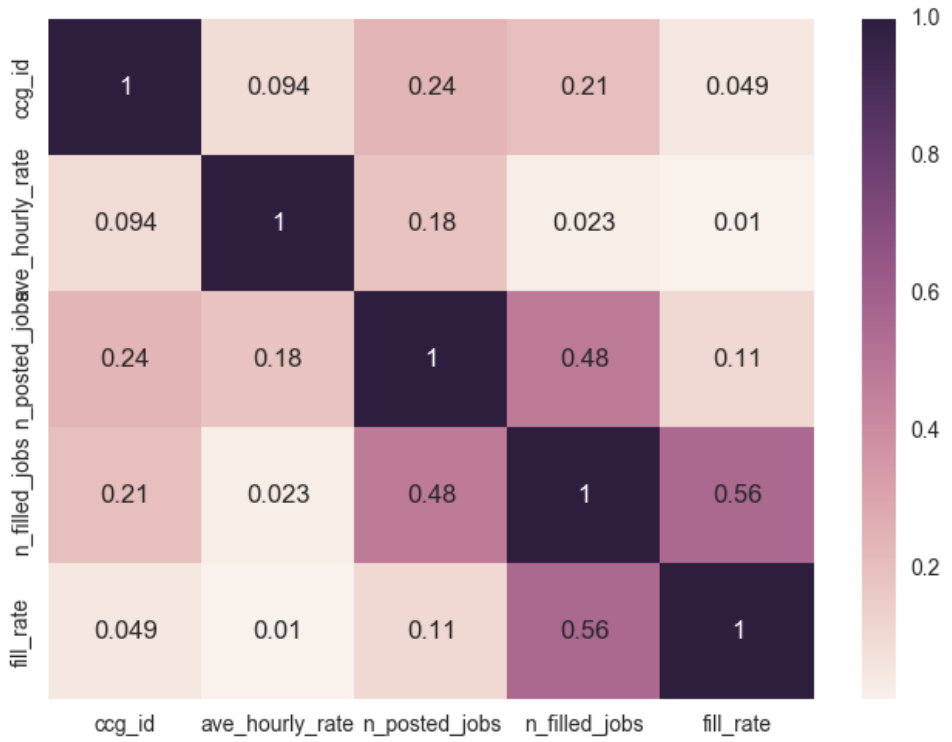
-To calculate the fill rate, I used

```
data['fill_rate'] = (data['n_filled_jobs']/data['n_posted_jobs'])*100 #calculate fill rate
```

Question 1: Does a higher average hourly rate result in a higher fill rate?

This question seeks to know whether there is a strong positive relationship between the two variables. The bivariate correlation heatmap below shows that the value of the correlation is only positive 0.01. The correlation is almost non-existent.

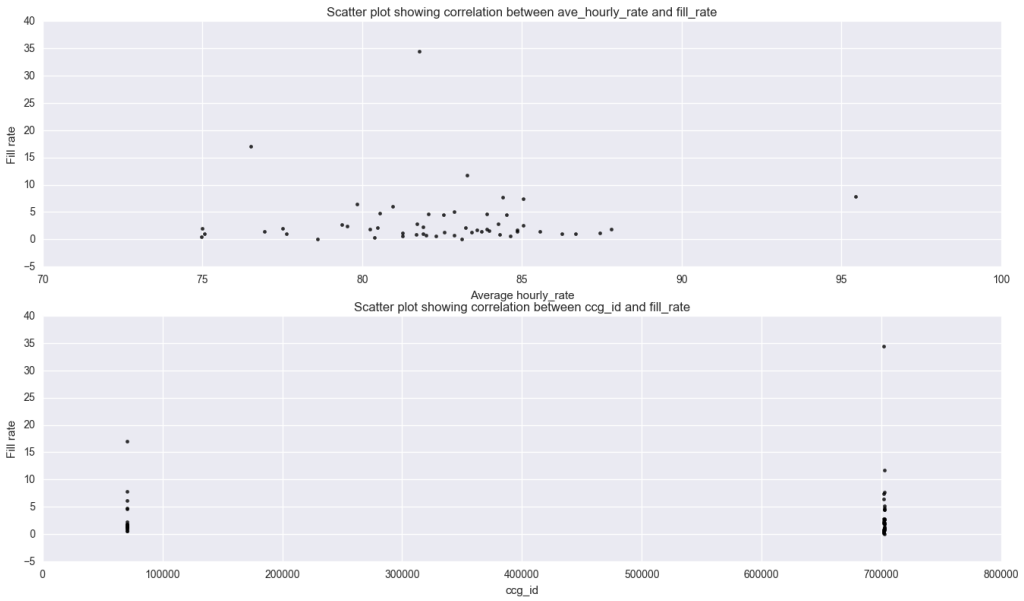
Figure 1:



The correlation can also be seen using a scatter plot of the two variables as shown below in figure 2. The first subplot shows the strength of the correlation.

Further analysis using Pearson's r coefficient confirms this conclusion

Figure 2: **corr_plots.png**



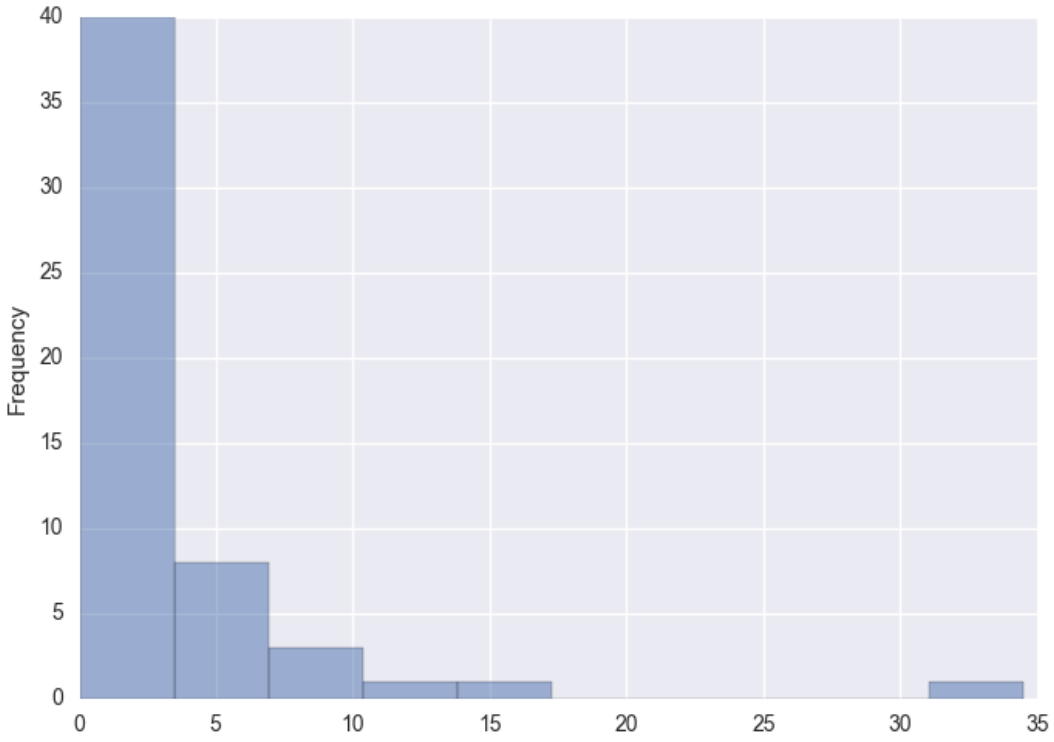
The coefficient of determination (r-squared) for this relationship is 0.000102388502229. R-squared value of 0 means that the fill rate cannot be predicted from the average hourly rate. R-squared value of 0.000102388 means that only 0.01% of fill rate is predictable from average hourly rate.

Question 2: What is the variability of this relationship across CCGs?

data['fill_rate'].describe()

	ccg_id	fill_rate
count	54.000000	54.000000
mean	468135.592593	3.394070
std	308035.313933	5.286650
min	70210.000000	0.068871
25%	70252.250000	1.037831
50%	702123.500000	1.789527
75%	702201.500000	4.050626
max	702285.000000	34.463895

The standard deviation is used to measure the amount of variation or dispersion of a set of data values. The standard deviation and the mean for the fill rate are 5.286650 and 3.394070 respectively. The maximum fill rate is 34.463895, 75% of the fill rate is 4.050626 and the minimum is 0.068871. The histogram below helps to visualize the variation.



Isaac Alabi