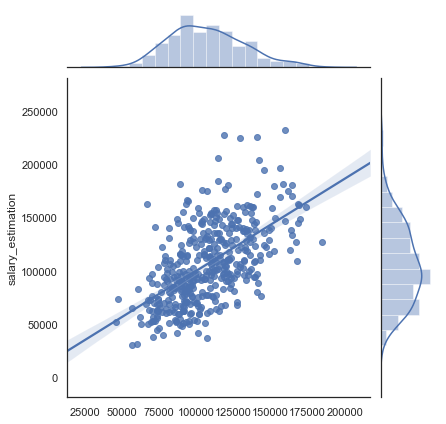
**Glassdoor Salary Estimation**

**Progress Model**

Brendon Happ

My model uses data related job descriptions from Glassdoor to predict salary estimations. I split the features up into programming languages, different skills (presentation, data mining, stats ect.), job title (jr, sr, director), and other values related to the companies.

My model is using only categorical features so a regression with only one variable wouldn’t show much. This model shows the actual salaries vs. the predicted salaries using all of my features. Overall, the feature with the biggest indicator of salary is the title of director, unsurprisingly.



Next I plan on using K-means and applying lasso regression to handle the overfitting and the large quantity of features.

Some of the features I chose may relate to one another so I may also go back and regroup them to account for multicollinearity.