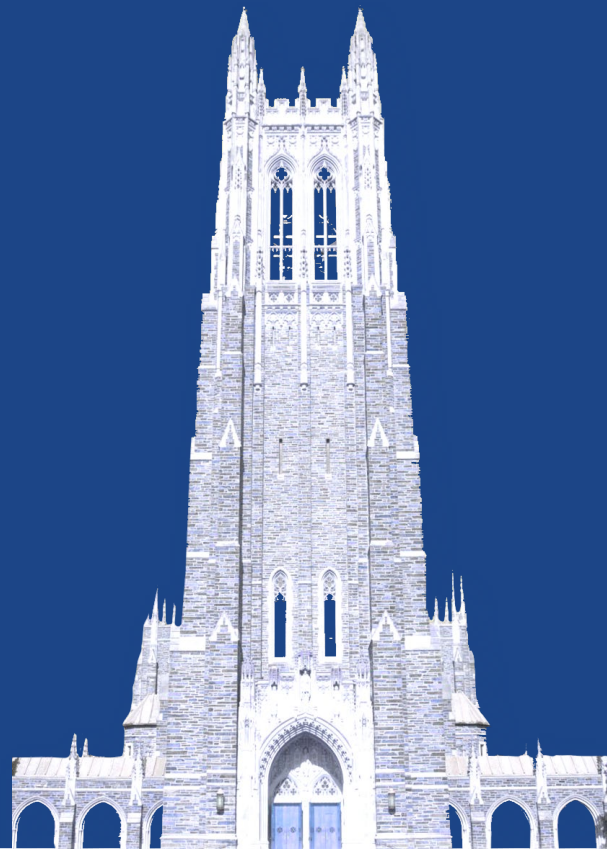


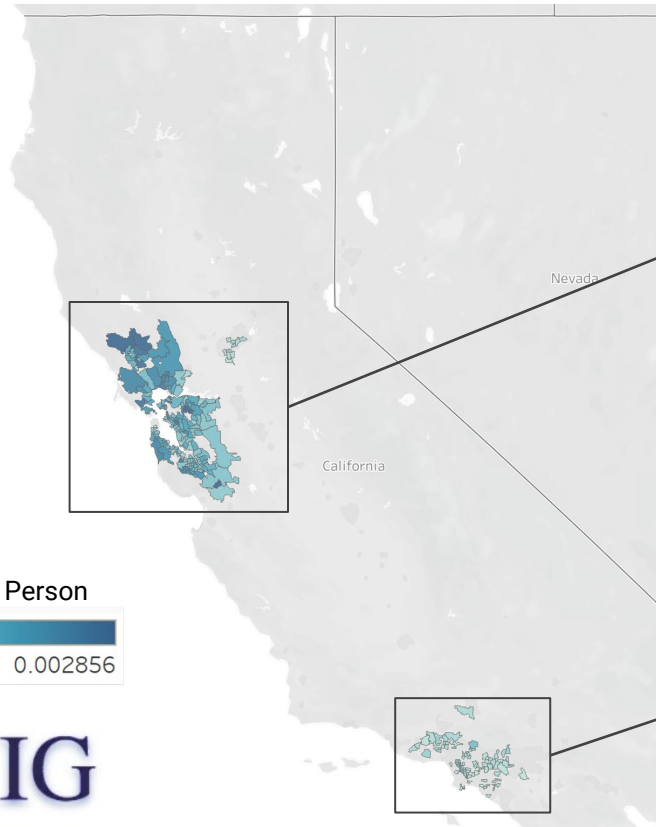


Example Deliverable

My portion of an analysis for a Duke Impact Investing Group (DIIG) Data Team Spring 21 Deliverable.

**I created and presented the following 10 slides
(originally from larger, four-person presentation)**

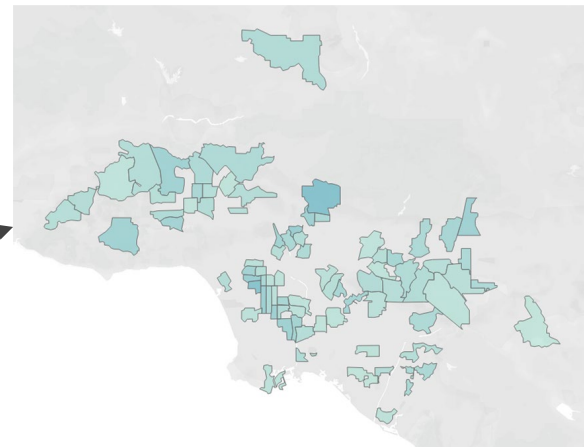
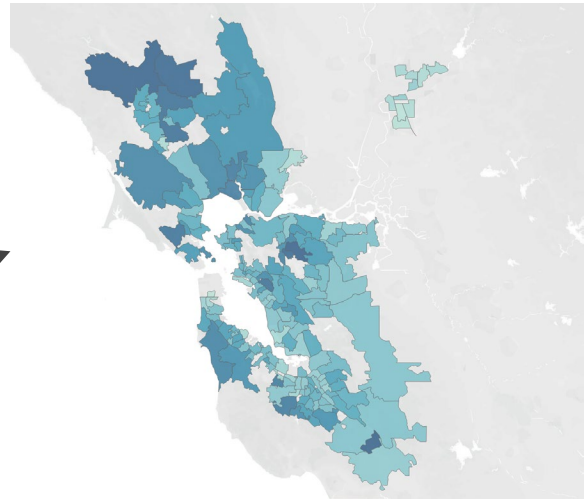


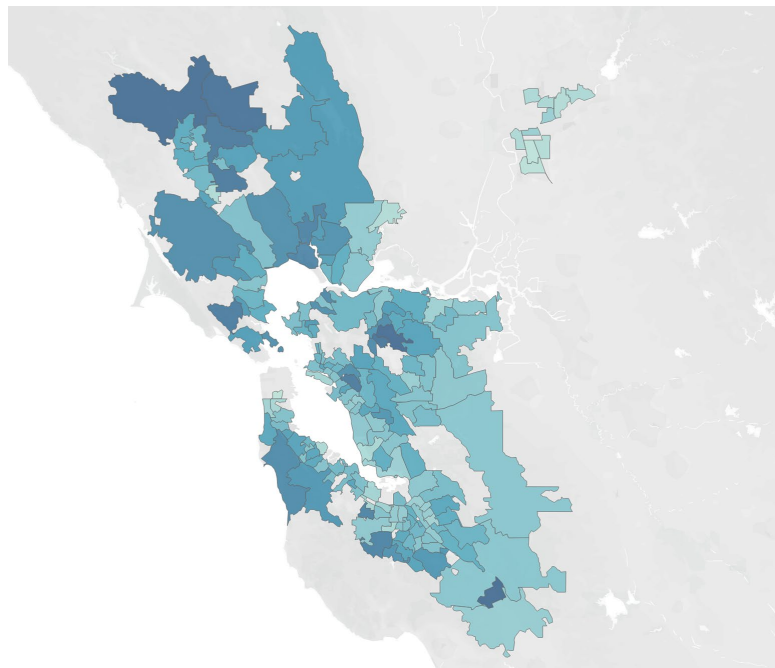


Nevada

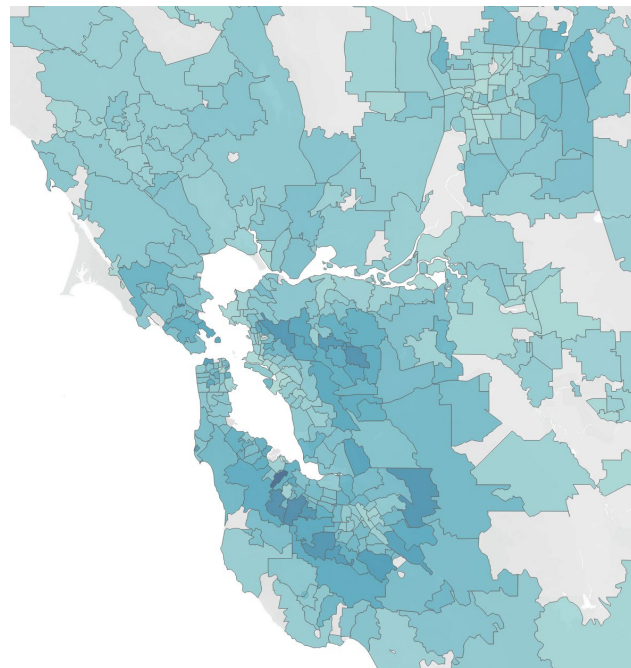
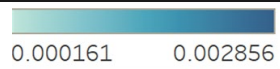
California

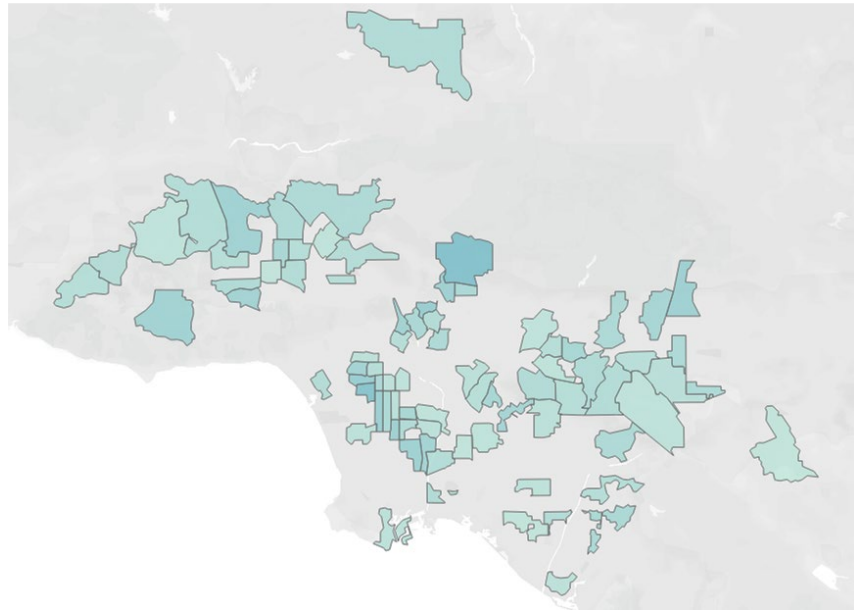
Entries per Person



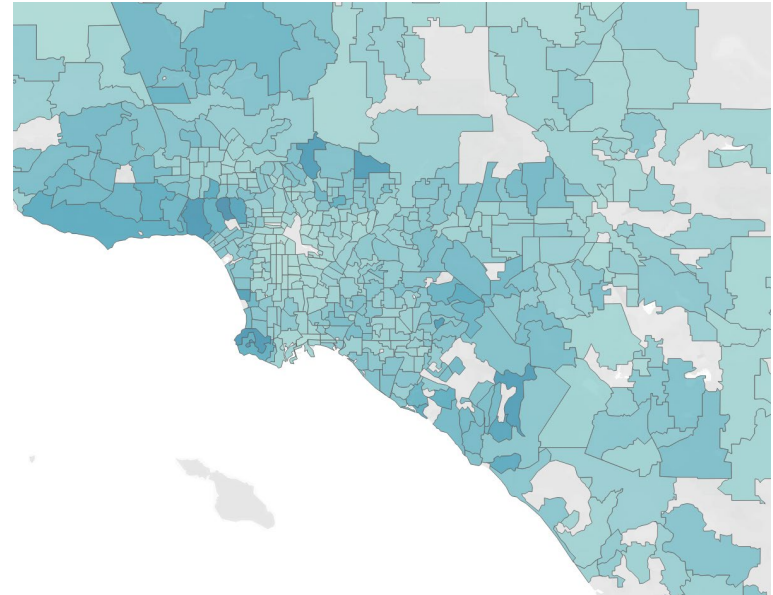
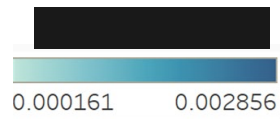


3





4



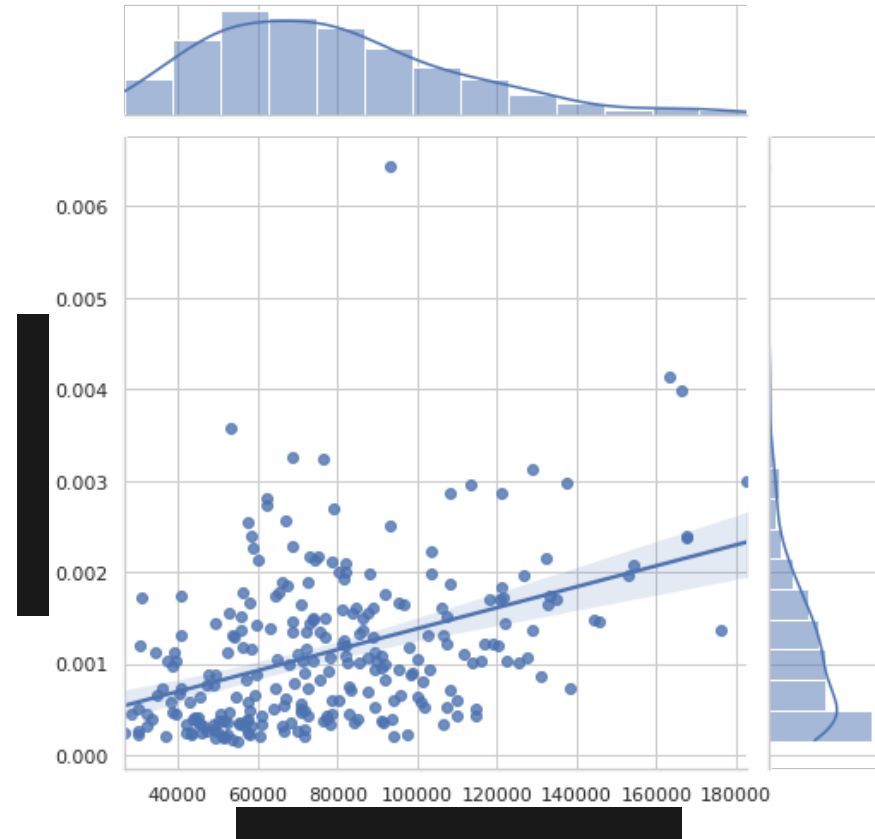
Summary Statistics

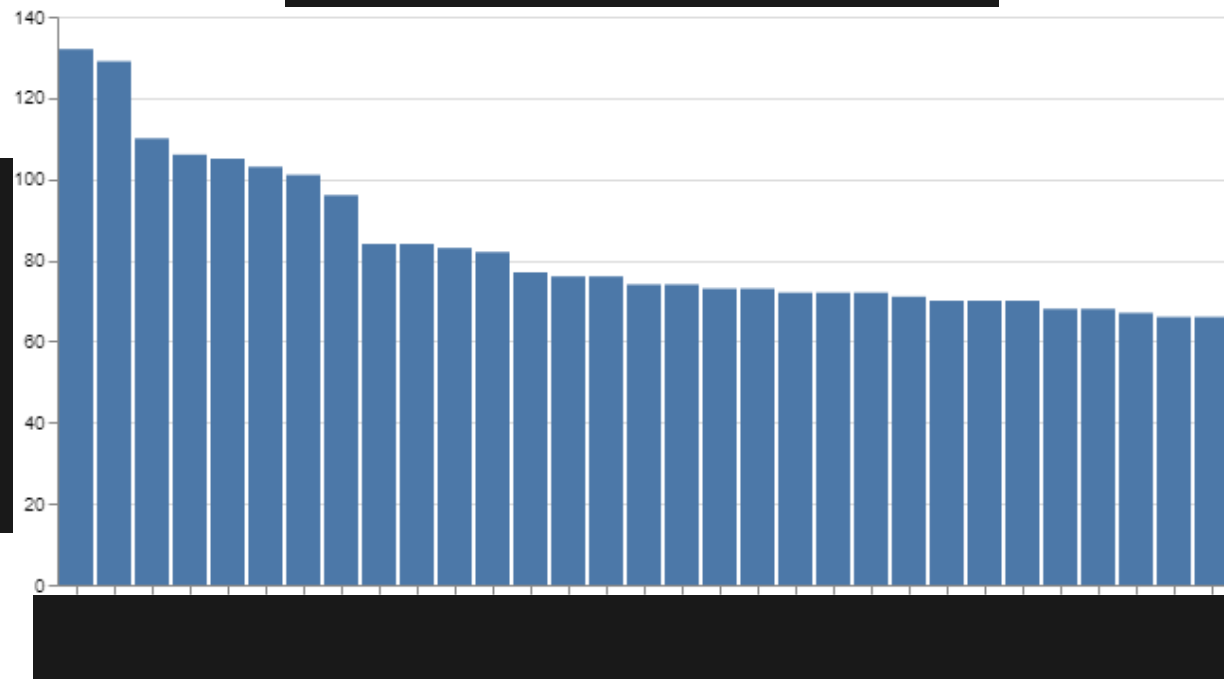
- R-squared: 0.176
- R: 0.42
- On average, an increase in \$---- of ---- correlates to 1.145 more entries per 10,000 ---- in a given ----.

OLS Regression Results

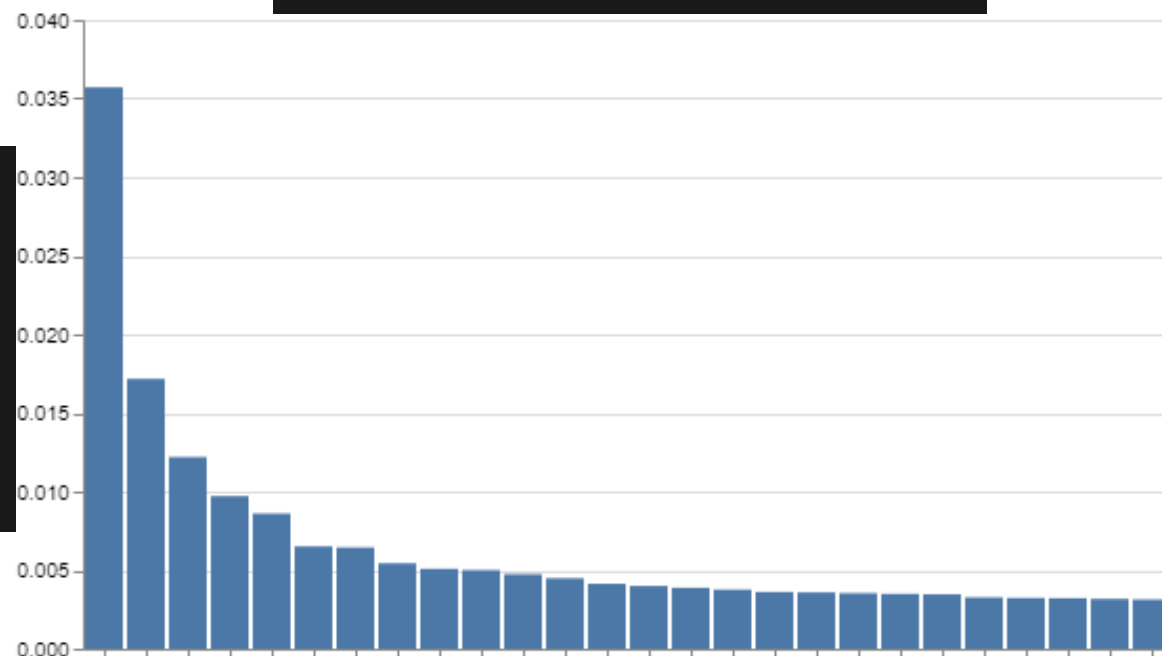
Dep. Variable: [REDACTED] **R-squared:** 0.176
Model: OLS **Adj. R-squared:** 0.173
Method: Least Squares **F-statistic:** 55.66
Date: Sat, 15 May 2021 **Prob (F-statistic):** 1.29e-12
Time: 15:58:01 **Log-Likelihood:** 1508.4
No. Observations: 262 **AIC:** -3013.
Df Residuals: 260 **BIC:** -3006.
Df Model: 1
Covariance Type: nonrobust

| | coef | std err | t | P> t | [0.025 | 0.975] |
|----------------|-----------|-------------------|-----------|-------|-----------|----------|
| const | 0.0002 | 0.000 | 1.840 | 0.067 | -1.66e-05 | 0.000 |
| Income | 1.145e-08 | 1.54e-09 | 7.460 | 0.000 | 8.43e-09 | 1.45e-08 |
| Omnibus: | 128.445 | Durbin-Watson: | 1.395 | | | |
| Prob(Omnibus): | 0.000 | Jarque-Bera (JB): | 774.589 | | | |
| Skew: | 1.903 | Prob(JB): | 6.31e-169 | | | |
| Kurtosis: | 10.515 | Cond. No. | 2.28e+05 | | | |

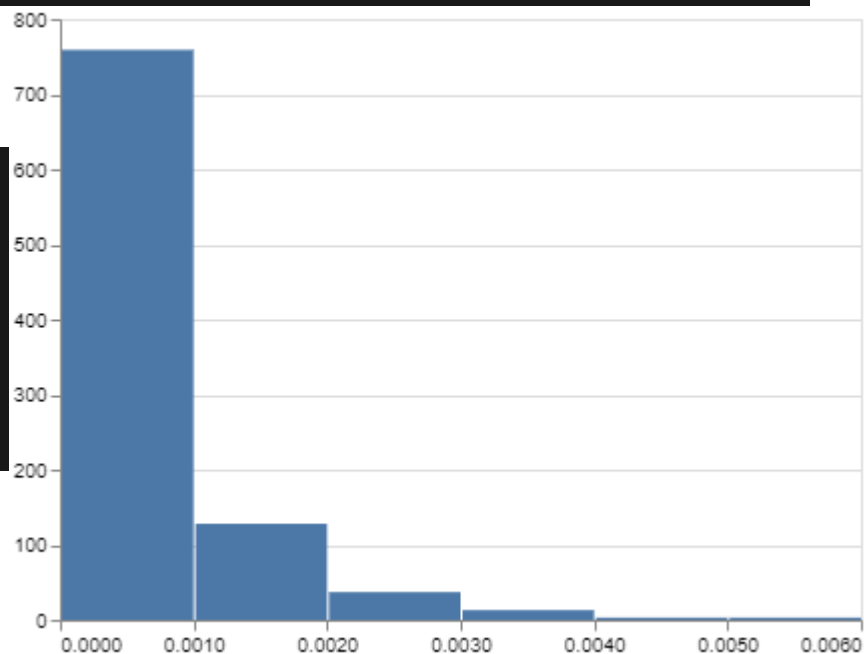
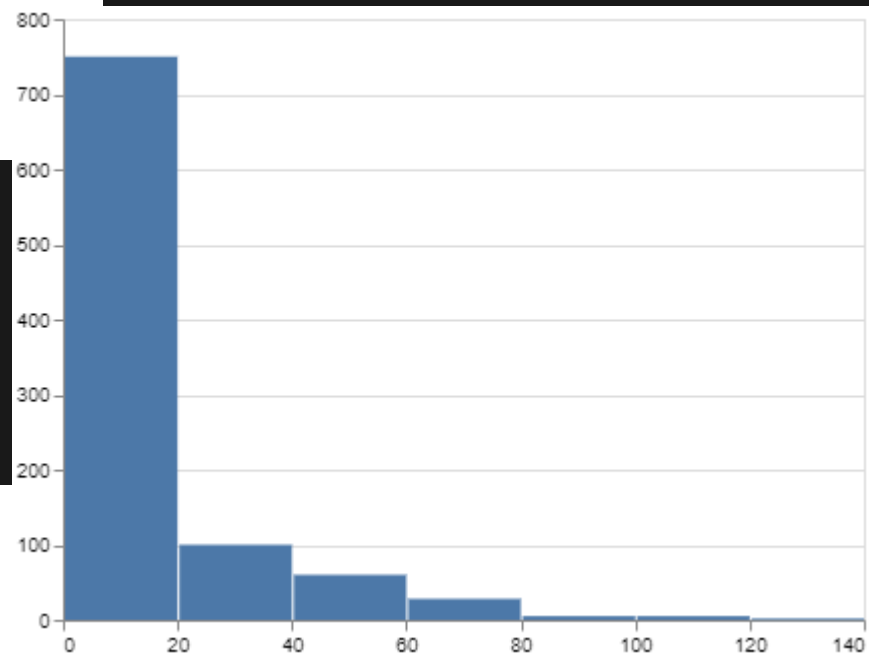


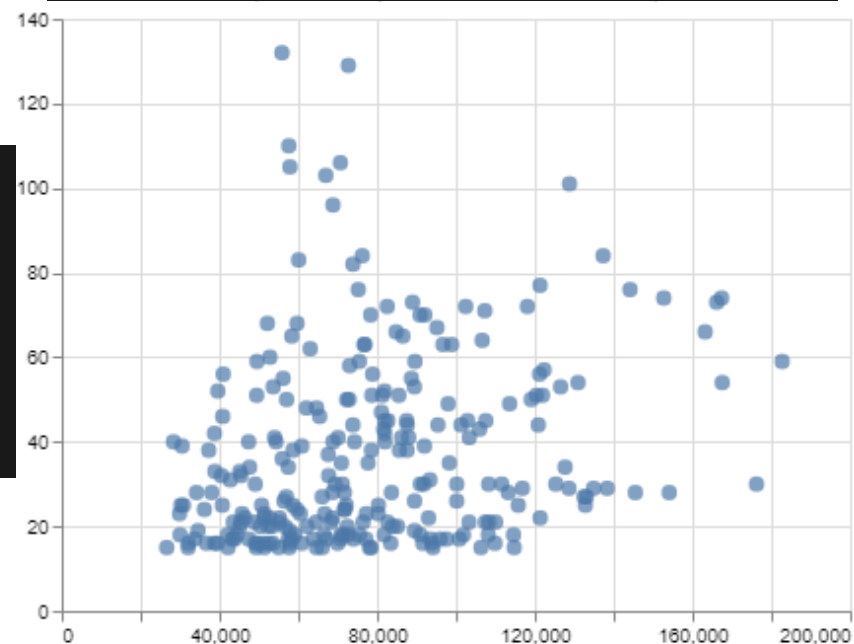
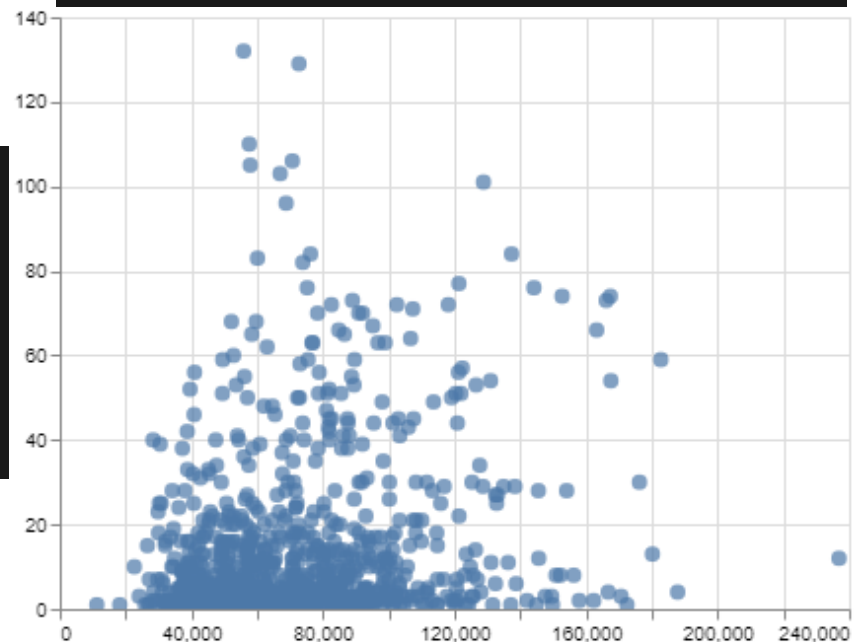


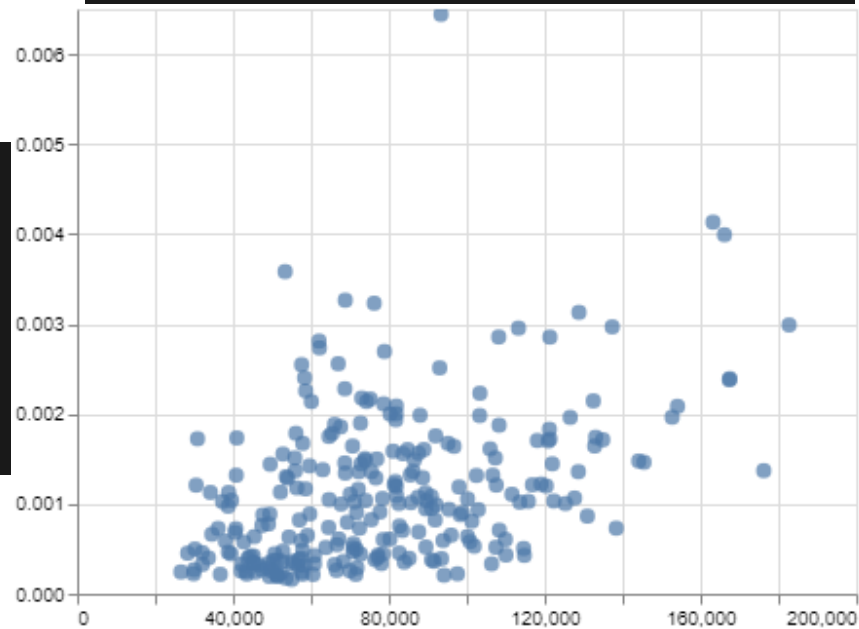
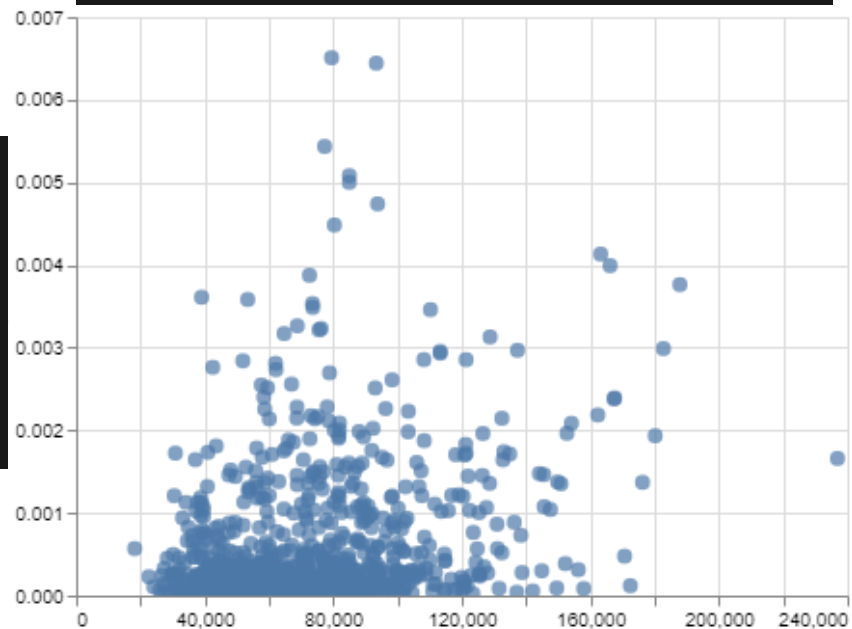
| | |
|-------|------------|
| count | 950.000000 |
| mean | 13.620000 |
| std | 19.009393 |
| min | 1.000000 |
| 25% | 2.000000 |
| 50% | 6.000000 |
| 75% | 16.000000 |
| max | 132.000000 |



| | |
|-------|------------|
| count | 950.000000 |
| mean | 0.000666 |
| std | 0.001597 |
| min | 0.000012 |
| 25% | 0.000102 |
| 50% | 0.000255 |
| 75% | 0.000727 |
| max | 0.035714 |







- ----- displays a modest positive, linear correlation with the number of ----- per ----- (excluding ----- with < 15 entries).
- If we assume median -----s display a proportional relationship with median -----, this **confirms the preexisting ranking value** for -----.
- Our R^2 value is **0.173** which approximately equals the **15%** value used in the ranking model to account for -----.
- Precisely, excluding ----- with less than 15 entries, median ----- accounts for 17.3% of the variability in number of potential ----- entries.
- We thus conclude it is best to focus marketing in higher -----.