STATS 'N SWAG

Dalton Hopper, Matt Lee, Zack Stone, Autumn Ward, Jonray Wilson

ISSUES WITH THE DATA SET

- There is a perfect linear relationship between the variables "Applies" and "Clicks"

$$r = 1 \qquad y = 4x + 11$$

- Variables that would have been useful if included:
  - Salary / Hourly Wages
  - Prioritization of Job Listings
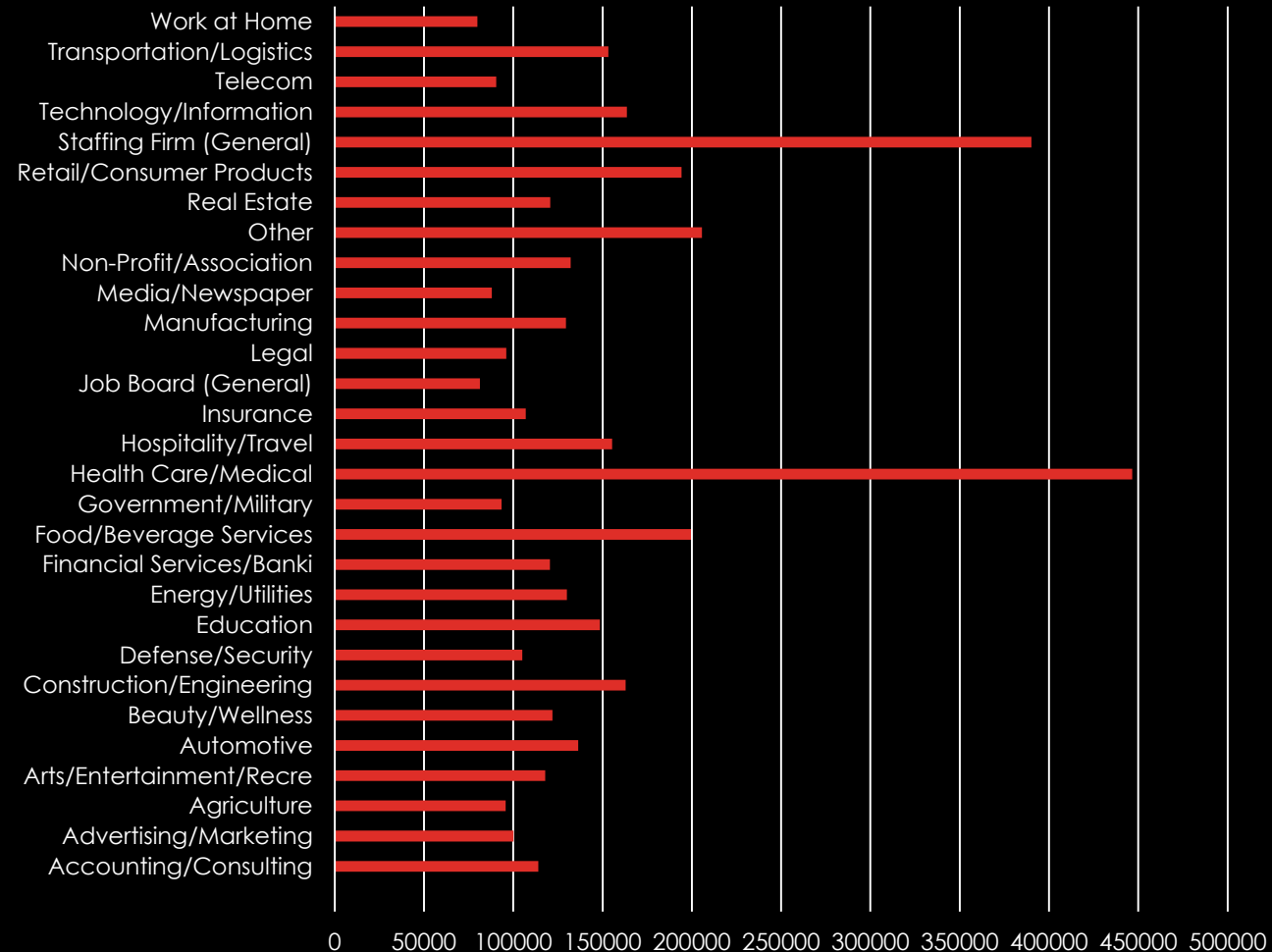 - Weekly Hours (part-time, full-time, overtime)

- Data set was limited to one year

License	NoEd	High School	Higher Education	Counts
0	0	0	0	4125157
0	0	0	1	18179
0	0	1	0	13534
0	0	1	1	28
0	1	0	0	72855
0	1	0	1	359
0	1	1	0	326
0	1	1	1	5
1	0	0	0	768
1	0	0	1	6859
1	0	1	0	11798
1	0	1	1	56
1	1	0	0	28527
1	1	0	1	231
1	1	1	0	317
1	1	1	1	1

- Required Qualifications
  - 0.6% Higher Education
  - 0.61% High School Education
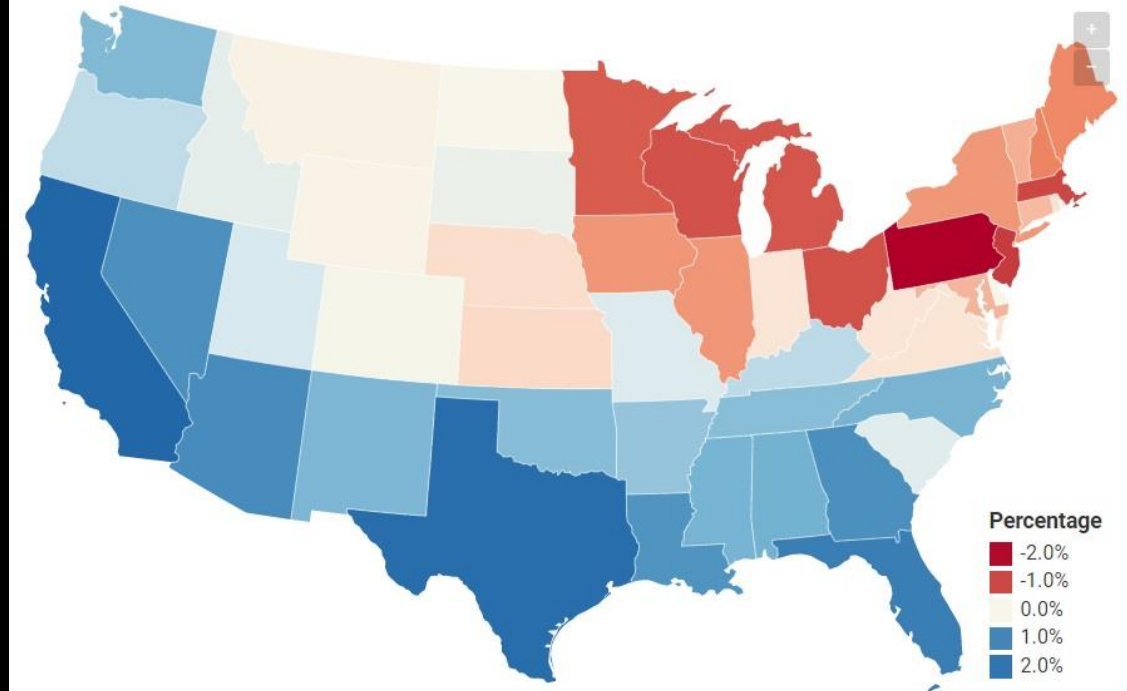 - 2.4% No Education
  - 1.13% Licensing

INDUSTRY WITHIN REGION

Industry Distribution

Difference in % US Pop. and % Indeed Job Listings

Difference = %US Population - % Indeed Job Listings

- How are the industries within each region distributed?

- How well do Indeed job postings represent the state populations and the amount of jobs listed in each state?

- Which states seem to not use Indeed?