

Salary Data Analysis

DS 311

Team ANOVA:

Vinny Giannantonio, Alyssa Gacos,
William Lin, Tristan Lampley, and Miles
Mendez



Introduction to the Dataset





Dataset: US Department of Labor's Office of Foreign Labor Certification

- Assembly of 167,278 records from the Labor Condition Applications and permanent resident applications from 2008 to 2015 (H1-B and specialty visas)
- Records display individual immigrant employee salaries
 - Unique view of what salary entails for data-based jobs
- Dataset Purposes
 - Tracks immigrant entrants for work-related reasons
 - Ensures financial equity of immigrant workers
 - Unique view of what salary entails for data-based jobs

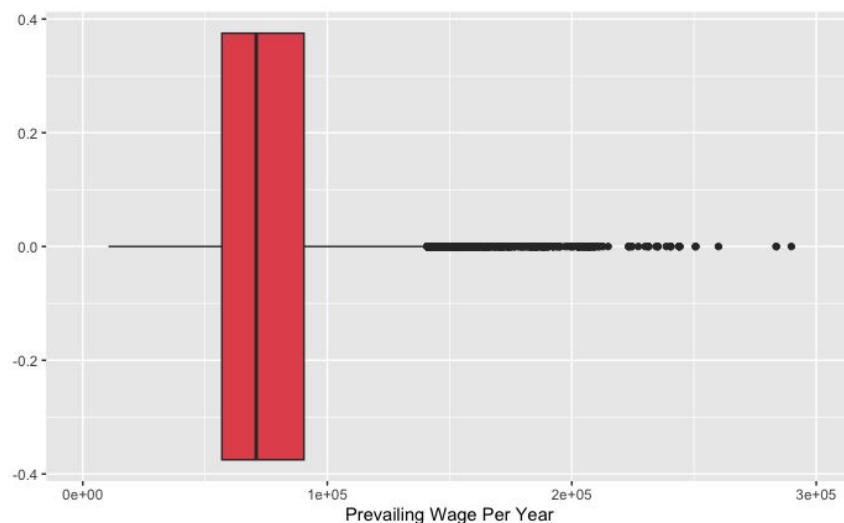
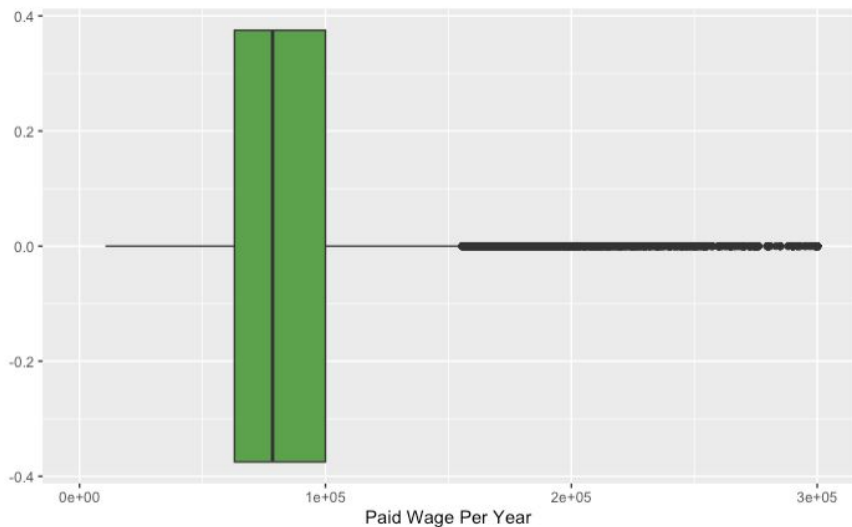
What is our KPI?

Paid Wage Per Year	<p>Paid salary per year for an individual employee</p> <ul style="list-style-type: none">• Accounts for “Paid Wage Submitted” and “Paid Wage Submitted Unit”
---------------------------	--



Summary Statistics

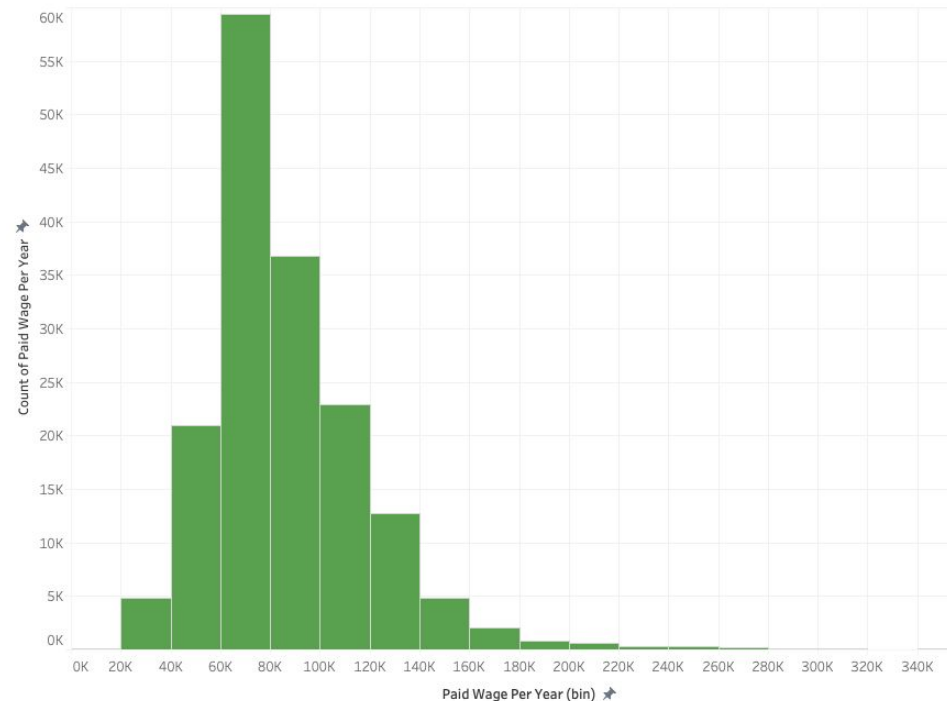
Variable <chr>	Min <dbl>	Q25 <dbl>	Median <dbl>	Mean <dbl>	Q75 <dbl>	Q99 <dbl>	Max <dbl>
Paid Wage Per Year (\$)	10500	63000	78600	85532.77	100006	200000.0	2500000
Prevailing Wage Per Year (\$)	10504	56880	70928	74274.87	90459	139742.1	320000



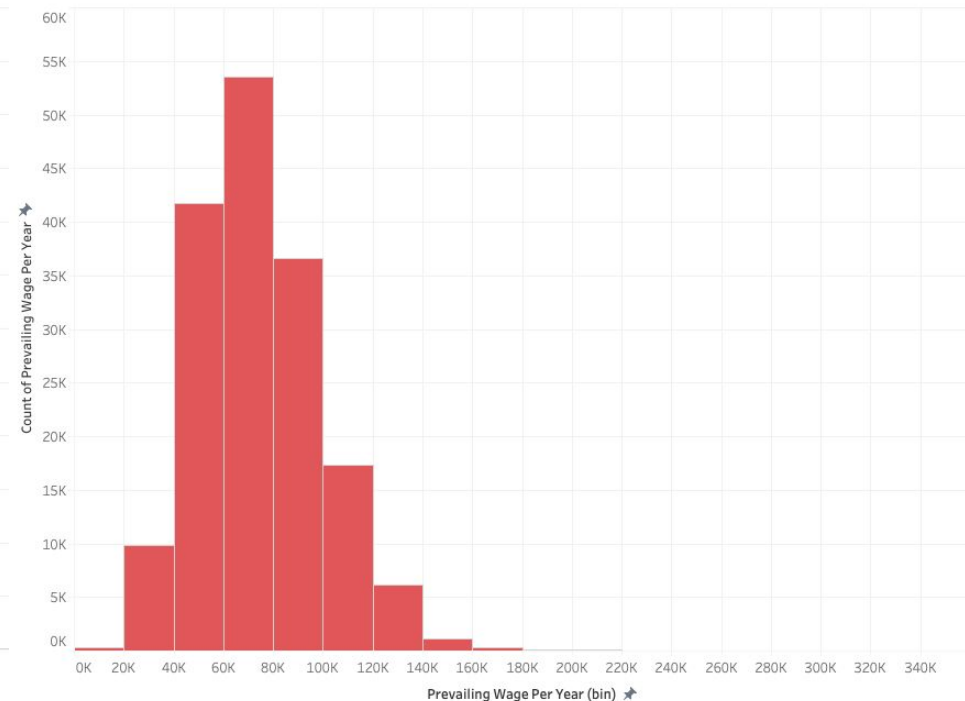


Visualization of Wage Distributions

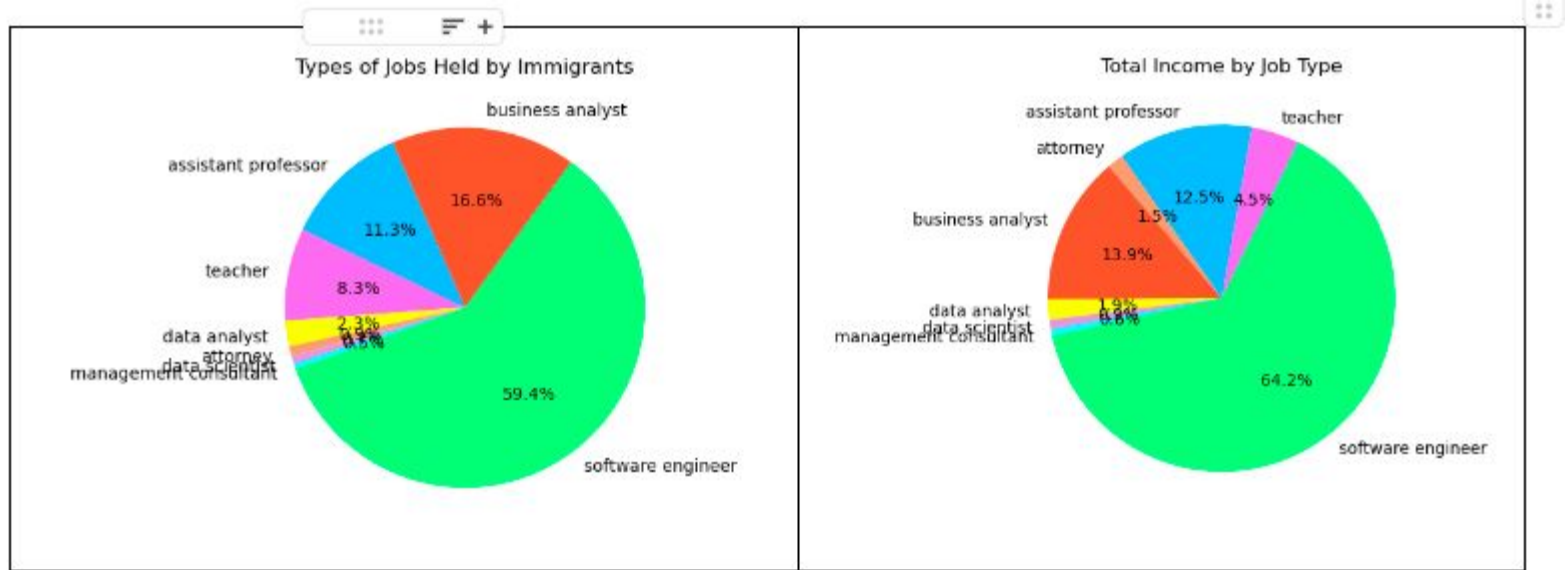
<Distribution of Paid Wage Per Year>



<Distribution of Prevailing Wage Per Year>




Pie Charts of Job Subtypes





Q1

**Do specific sub-types of
data-related jobs have higher or
lower salaries than others?**

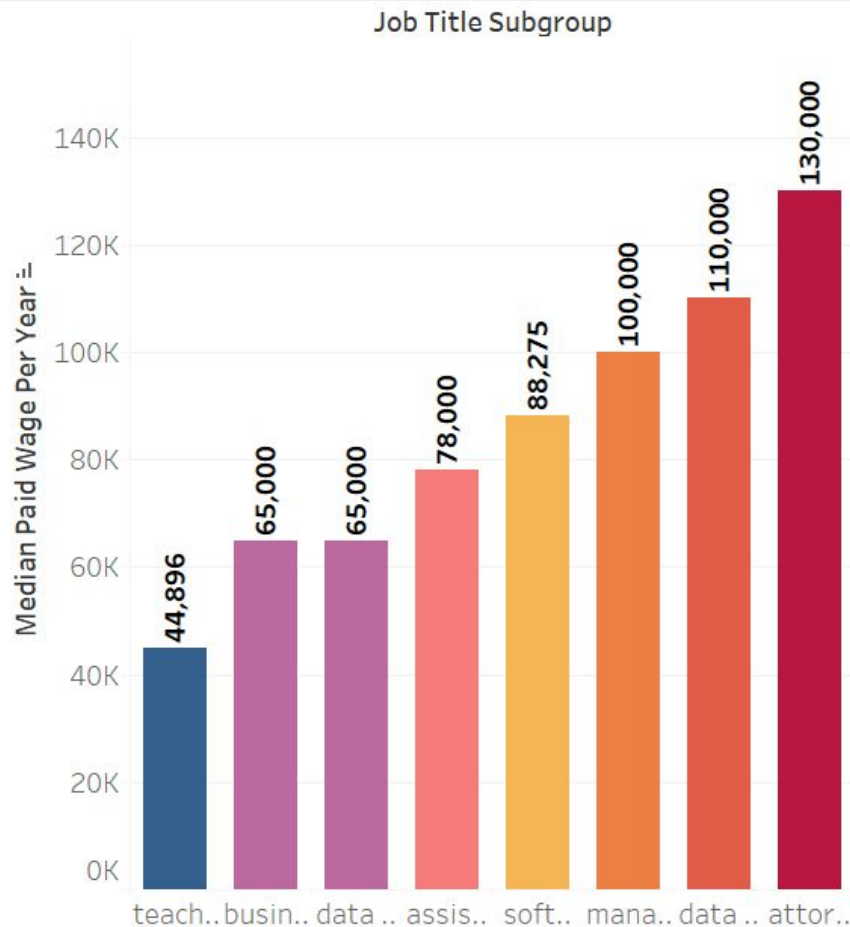




1a. What are the highest median salaries by subtypes?

1. *Attorney: \$130,000*
2. *Data Scientist: \$110,000*
3. *Management Consultant: \$100,000*

Salaries by Job Subtypes







1b. Which companies have the highest salaries for those sub-types?

JOB_TITLE_SUBGROUP <chr>	EMPLOYER_NAME <chr>	AvgSalary <dbl>
attorney	HUGHES HUBBARD & REED LLP	2400000.0
business analyst	OFFICEMAX INCORPORATED	603711.7
software engineer	SIGMATEK SYSTEMS, LLC	600000.0
assistant professor	DIXIE STATE COLLEGE OF UTAH	492250.0
data analyst	INTUIT	433161.5
management consultant	BAKER BOTTS L.L.P.	350000.0
teacher	THE CHILDREN'S SCHOOL, INC.	283500.0
data scientist	NETFLIX	220000.0




1c. Does changing location change the ranking of job sub-types?

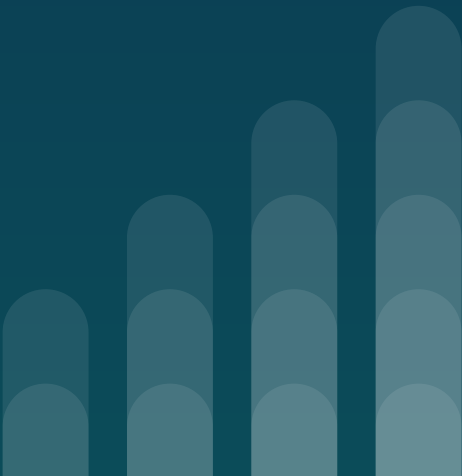
JOB_TITLE_SUBGROUP <chr>	CA_MedianSalary <dbl>
attorney	146221.5
data scientist	119800.0
software engineer	105972.5
management consultant 	97446.0
assistant professor	81353.5
data analyst	75046.0
business analyst	70000.0
teacher	48150.0

JOB_TITLE_SUBGROUP <chr>	MA_MedianSalary <dbl>
data scientist	102500
management consultant 	97000
attorney	93850
software engineer	88691
assistant professor	88500
business analyst	67000
data analyst	65000
teacher	48000

Conclusion: Yes!



Q2: What states (of those that I am willing to move to) have the highest paying data-related salaries?

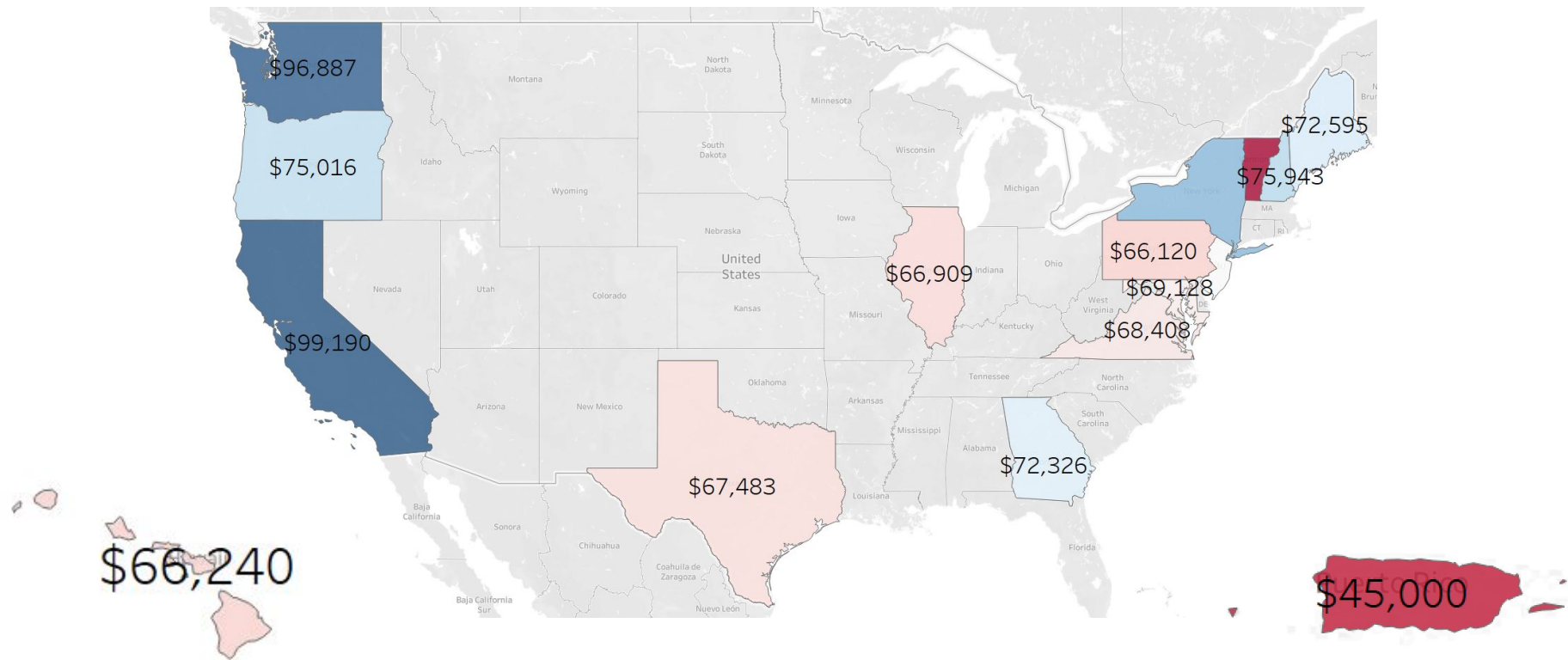




Top 10 and bottom 10 states for data jobs, ranked by average salary

	State	Average Salary
0	California	99190.31
1	Washington	96887.08
2	New York	81722.32
3	New Hampshire	75943.25
4	Oregon	75016.47
5	Maine	72595.25
6	Georgia	72325.67
7	New Jersey	70514.90
8	Maryland	69127.53
9	Virginia	68407.99

	State	Average Salary
38	Arkansas	62376.48
39	Florida	62016.61
40	South Carolina	61402.00
41	Louisiana	59692.31
42	Oklahoma	55498.40
43	Nebraska	52010.07
44	Idaho	48699.75
45	North Dakota	47714.67
46	Puerto Rico	45000.00
47	Vermont	42000.00



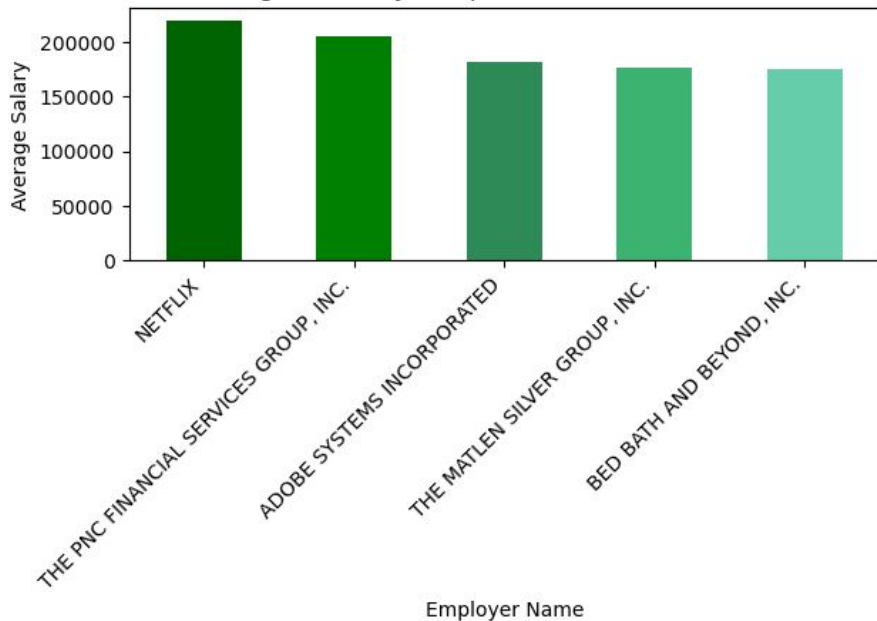
2a. Which is the best state to work in for each job?

JOB_TITLE_SUBGROUP <chr>	WORK_STATE <chr>	MedSalary <dbl>
attorney	District of Columbia	160000.0
attorney	New York	160000.0
management consultant	Oklahoma	141847.5
data scientist	California	119102.5
software engineer	California	105000.0
software engineer	Washington	105000.0
assistant professor	Arkansas	97000.0
business analyst	Montana	85000.0
data analyst	Delaware	77700.0
teacher	Maryland	69664.0

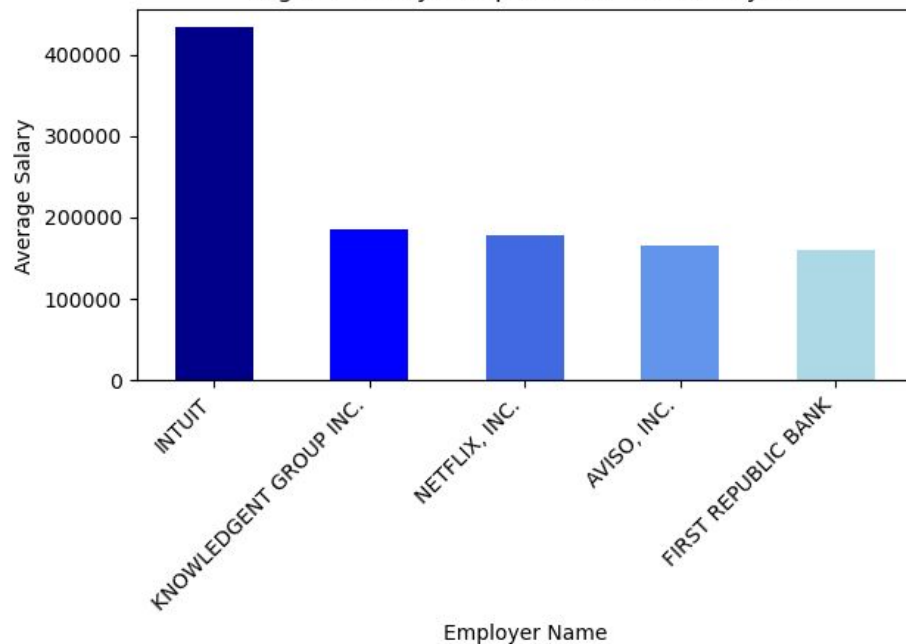


What companies are good to work for?

Highest Salary Companies for Data Scientists



Highest Salary Companies for Data Analysts






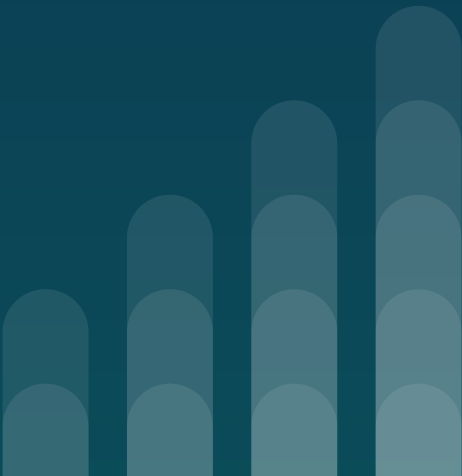
Comparing Costs of Living

- COL Ratio = Average Salary/COL index
- Washington wages are better than California when considering cost-of-living

State	Average Salary	Cost of Living Index	COL Ratio
Washington	96887.08	115.7	837.40
Georgia	72325.67	90.9	795.66
Illinois	66909.28	91.8	728.86
California	99190.31	136.4	727.20
Texas	67483.28	93.0	725.63
Pennsylvania	66120.19	94.8	697.47
Virginia	68407.99	101.4	674.64
New Hampshire	75943.25	115.6	656.95
Maine	72595.25	110.7	655.78
Oregon	75016.47	114.7	654.02
New York	81722.32	126.5	646.03
New Jersey	70514.90	110.3	639.30
Maryland	69127.53	114.8	602.16
Hawaii	66240.33	179.2	369.64
Vermont	42000.00	114.9	365.54
Puerto Rico	45000.00	NaN	NaN

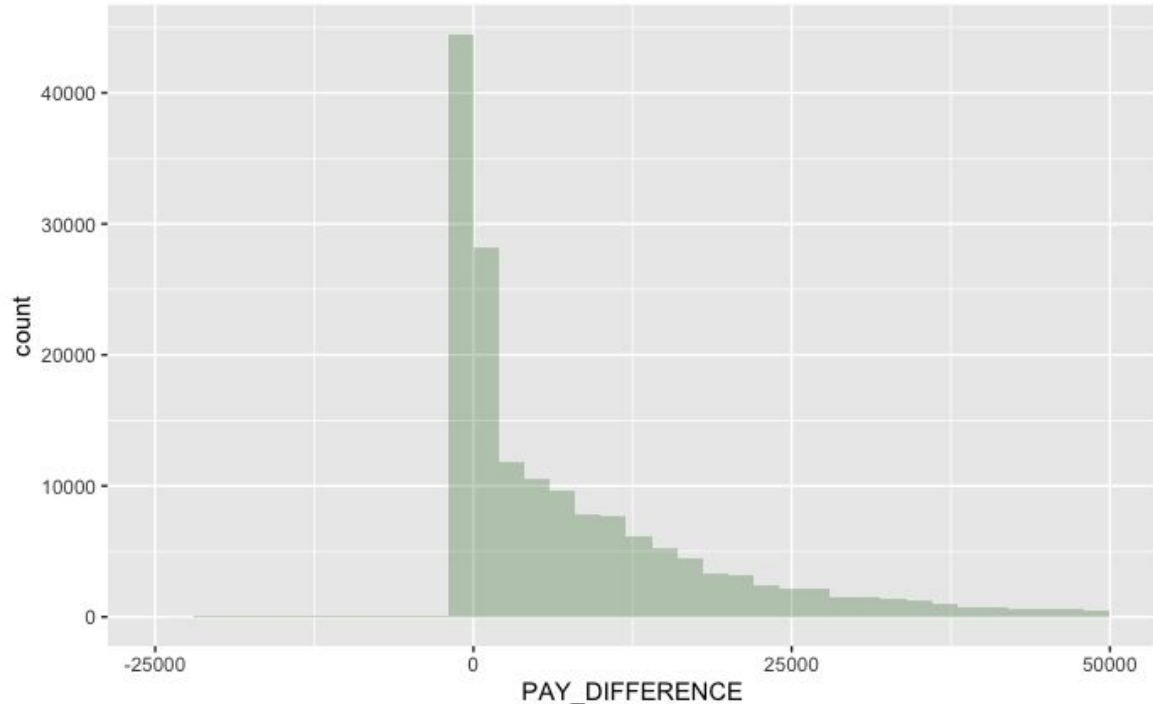


**Q3: How do offered
salaries compare to the
prevailing wage?**



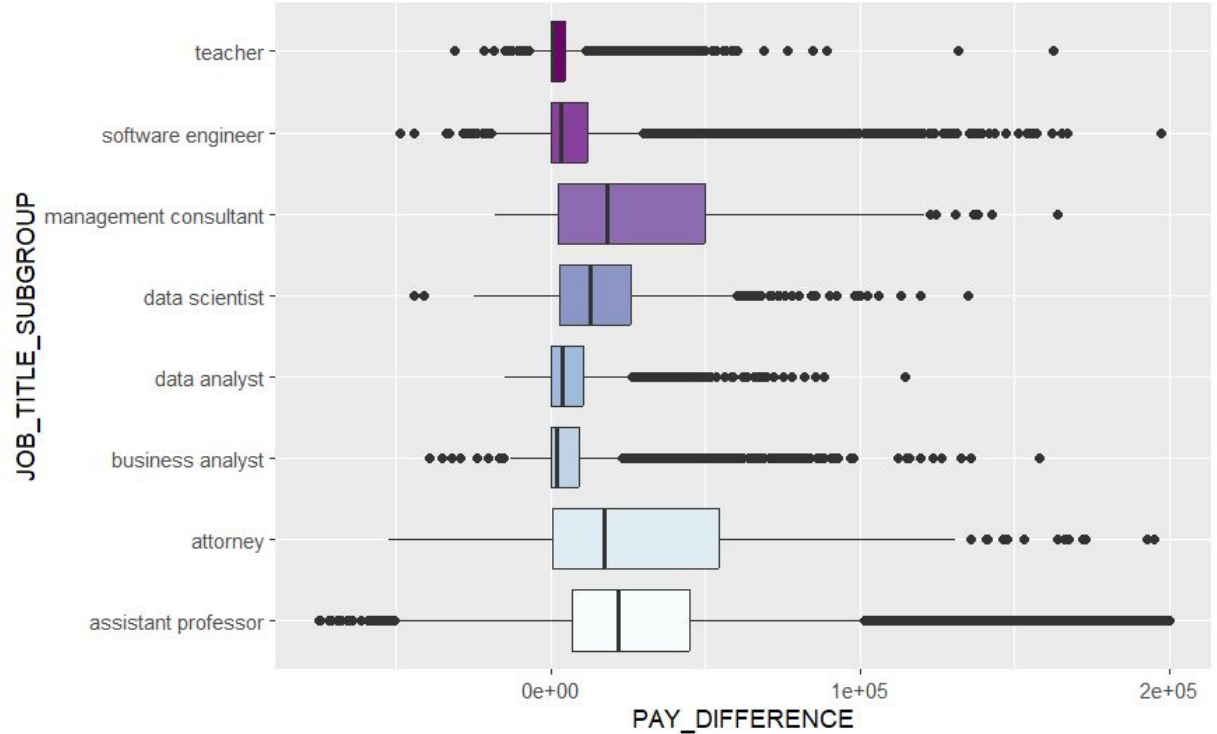
Paid Wages minus Prevailing Wages

- Offered salaries are mostly close to or higher than prevailing wage (difference > 0)
- ~20% of jobs are underpaid by \$0-\$2,000 per year (one bin width)

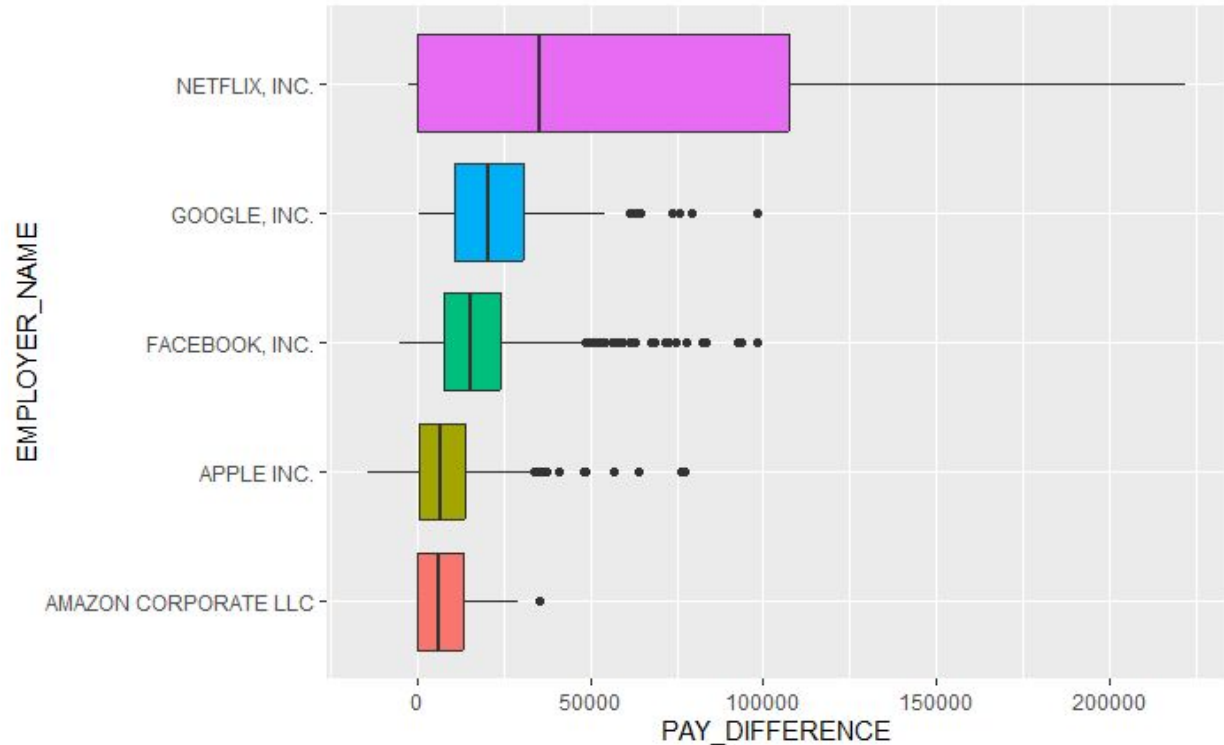


3a. Are there job sub-categories that tend to get overpaid or underpaid?

1. ~25% of every job is underpaid (in dataset)
2. Assistant professor, Attorney, and management consultant highest median pay relative to prevailing wage



3b. Are there companies that tend to overpay or underpay?



- Amazon is the only company that does not underpay
- Netflix overpays the most
- Apple underpays the most

Overpay

Netflix

Google

Facebook

Apple

Amazon

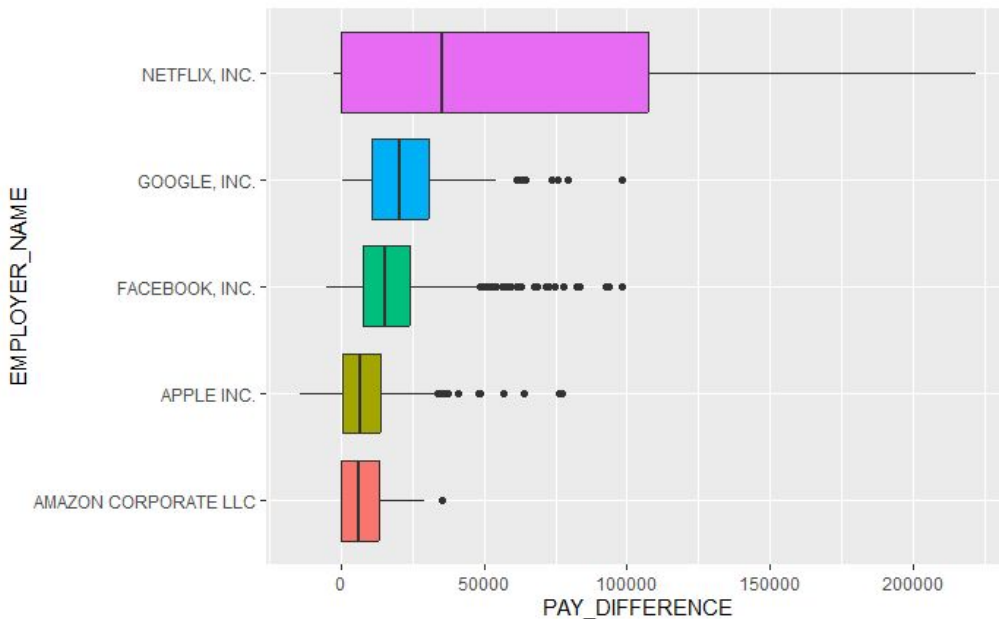
Underpay


Netflix

Google

Facebook

Apple

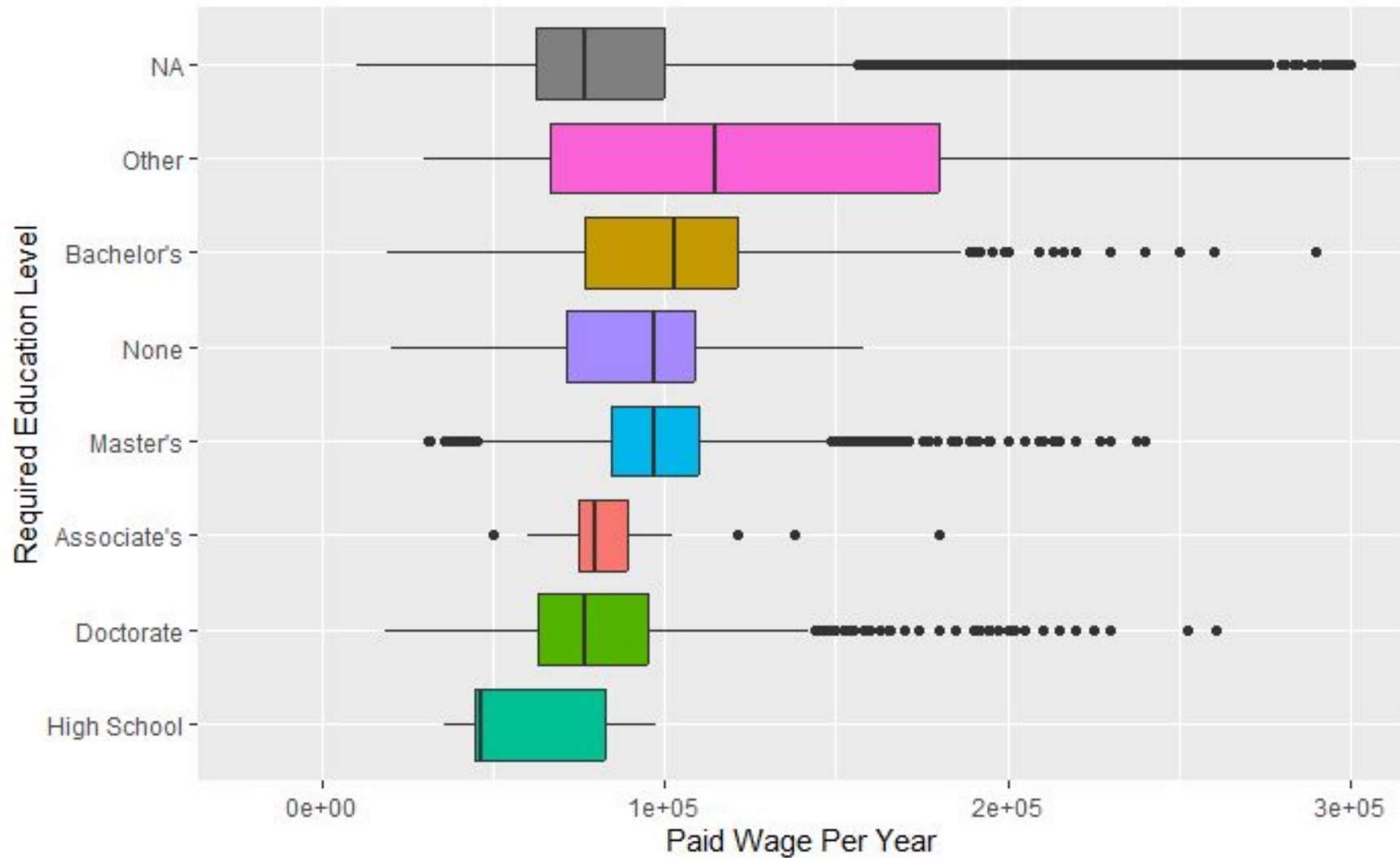




Q4


**How is wage correlated
with education required?**



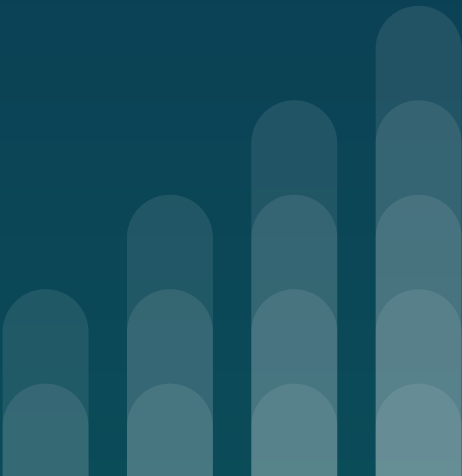


4. How is wage correlated with education required?

1. High school required jobs pay less
2. Bachelor degree required jobs pay the most
3. No requirement and master's are similar

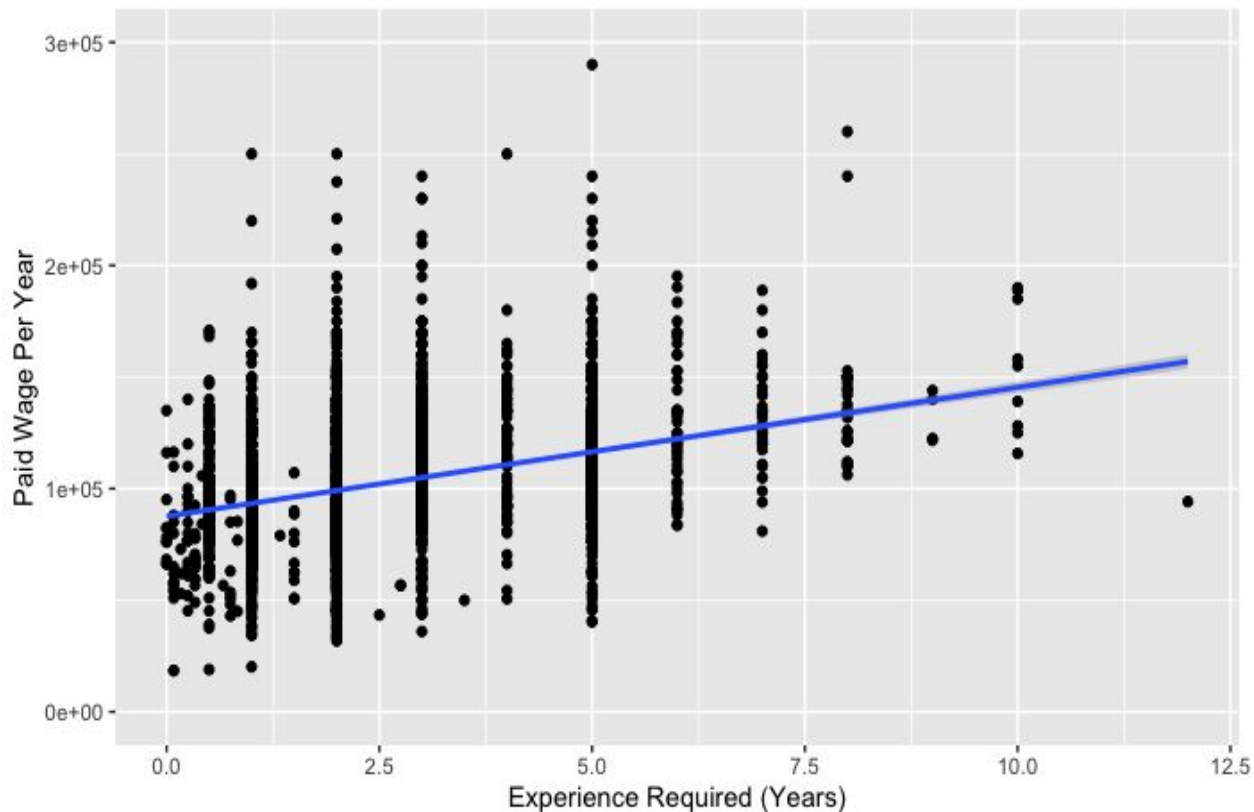



**Q5 How is wage
correlated with
experience required?**



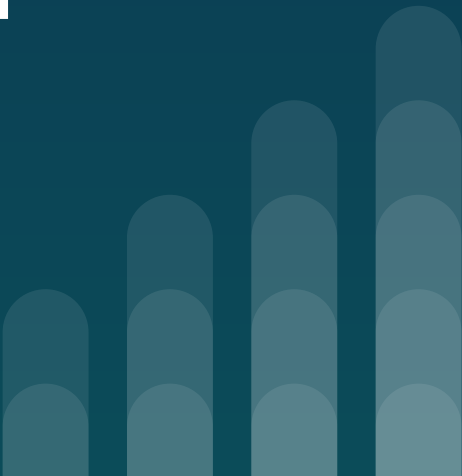
5. How is wage correlated with experience required?

- Weak positive association
- Line is a bad fit \Rightarrow experience required alone is not a good predictor of wages



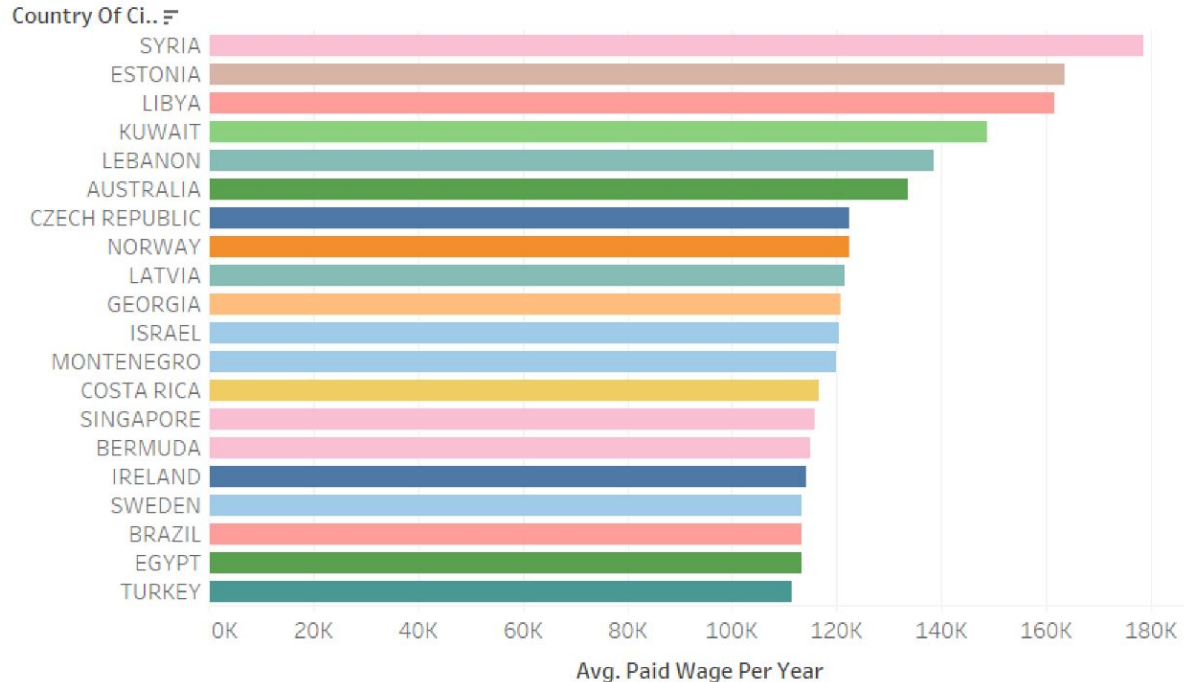


**Q6: What is the relationship
between the Country of
Citizenship and Paid Wage per
Year?**



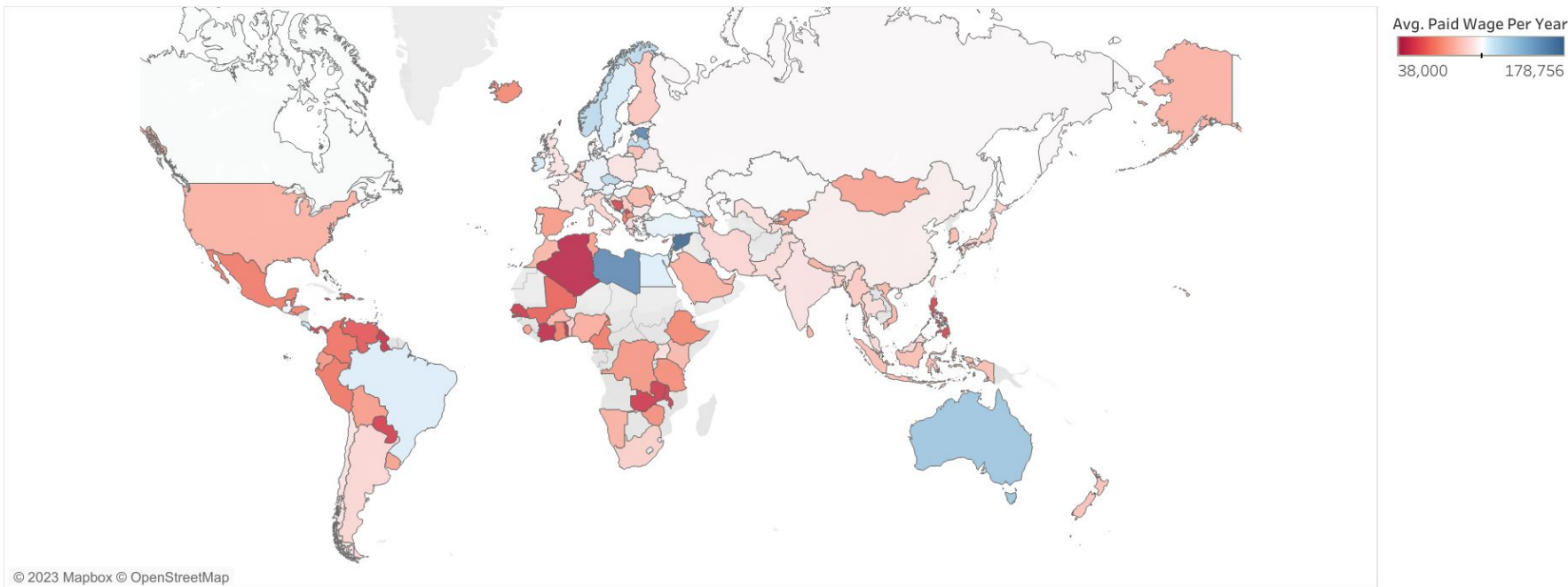
What are the countries of citizenship w/ Highest Avg Paid Wage?

Country of Citizenship and Avg Paid Wage



Continental display of countries of citizenship based on average paid wage per year

Q6 Map





Conclusions

- Median wage: \$78,600; IQR: (\$63,000-\$100,000)
 - In California: median = \$101,000
 - Washington outperforms California when considering cost of living
- Software engineer most prevalent job (63%)
- ~75% of salaries are equal to or above prevailing wages
- Education and experience required are not good predictors of wages by themselves



Github repository

<https://github.com/jvincent1113/Team-Anova-Project>