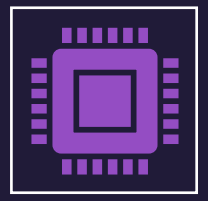
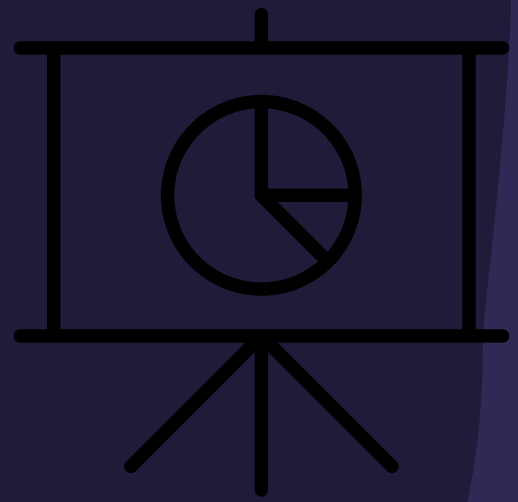
Data Analyst Job Market

Explore story telling and truth-telling with python.



We chose the “Indeed Dataset – Indeed Dataset - Data Scientist/Analyst/Engineer” from Kaggle.com

Our main goal was to familiarize ourselves with different aspects of Data Science field. For example: skills, salary, Location and so on.



Team:

Shailza Rattu
Lori Pepper
Susana Villagrana
Jennifer Rocha

Cleaning

manipulating data for analysis and visualizations

Our Data Manipulation

Original Dataset:

- 43 columns
- Lots of missing data with different values to indicate a lack of information

Dropped Columns ‘Link’, ‘Date Since Posted’, and ‘Description’

We dropped the “Description” column since it requires machine learning skills to extract useful data

Adjusted data types when necessary

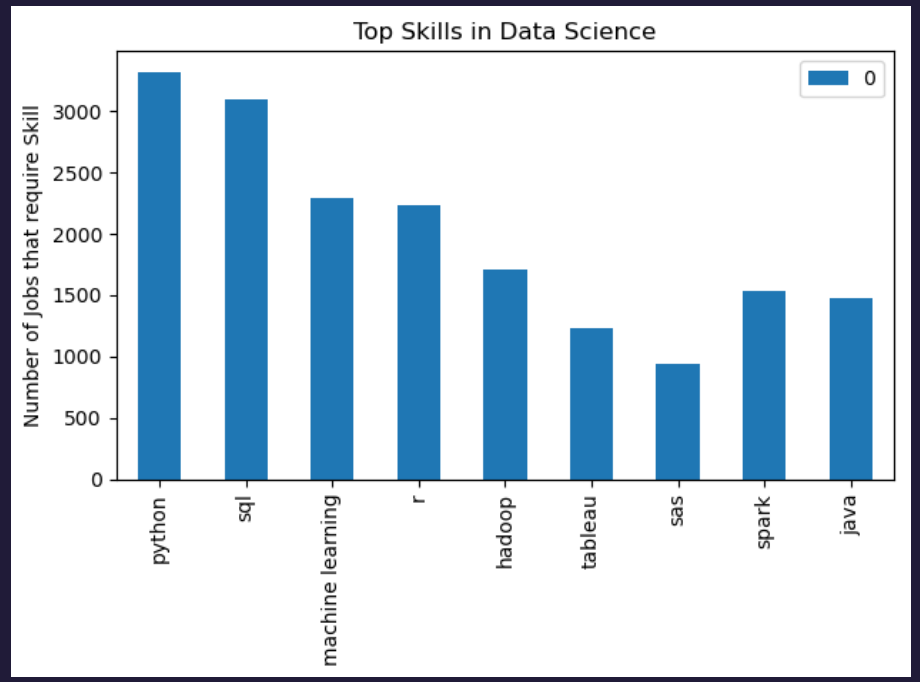
Added latitude, longitude, and a cost-of-living index for use with 'Location'

Modified data format for use in visualizations

Merged the four cleaned .csv files into one dataset to use for visualizations

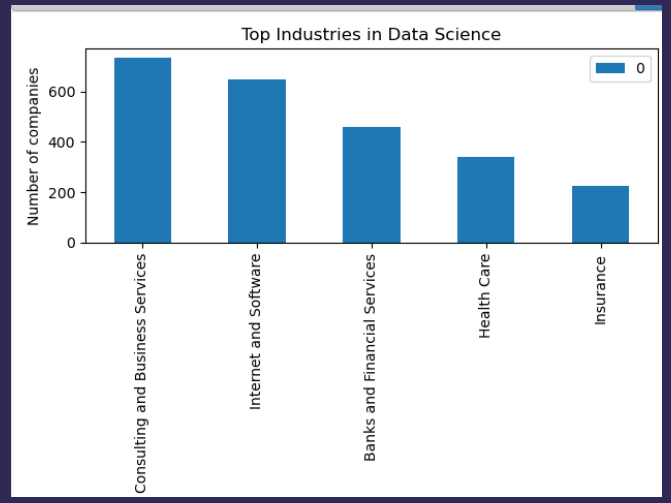
Overview

Popular Skills



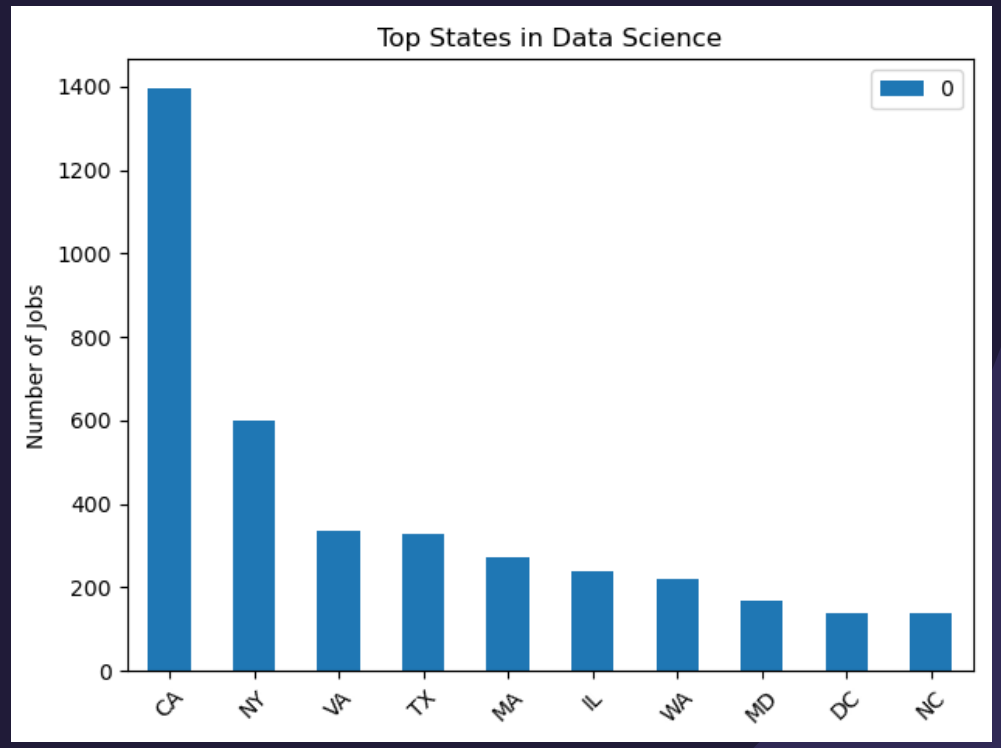
Python	SQL	Machine learning
r	Hadoop	Tableau
Sas	Spark	java

Top Industries



- Educating clients on varying aspects of data and modern technology.
- Manage algorithmic trading, fraud detection, customer management, risk analytics
- Develop forecasting and modeling programs designed to form analyses of medical records or other forms of healthcare information
- Combine analytical applications – e.g., behavioral models based on customer profile data – with a continuous stream of real-time data – e.g., satellite data, weather reports, vehicle sensors – to create detailed and personalized assessments of risk.

Top States



California

New York

Virginia

Texas

Massachusetts

Illinois

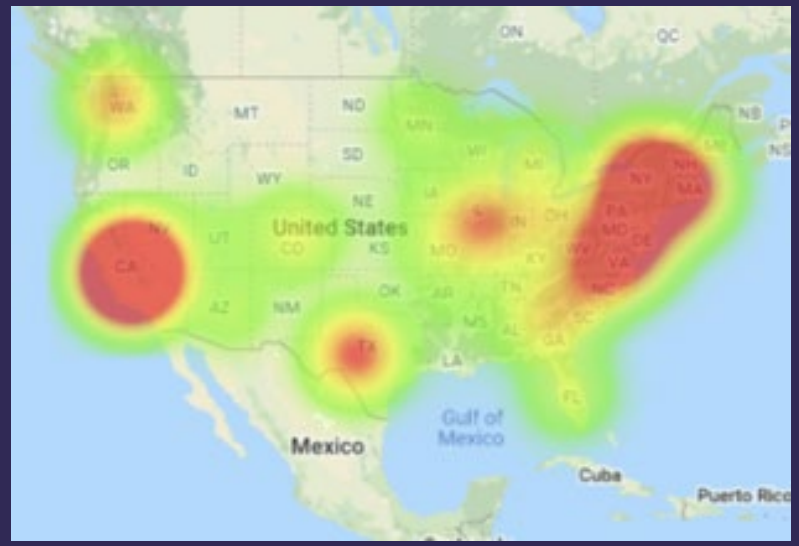
Washington

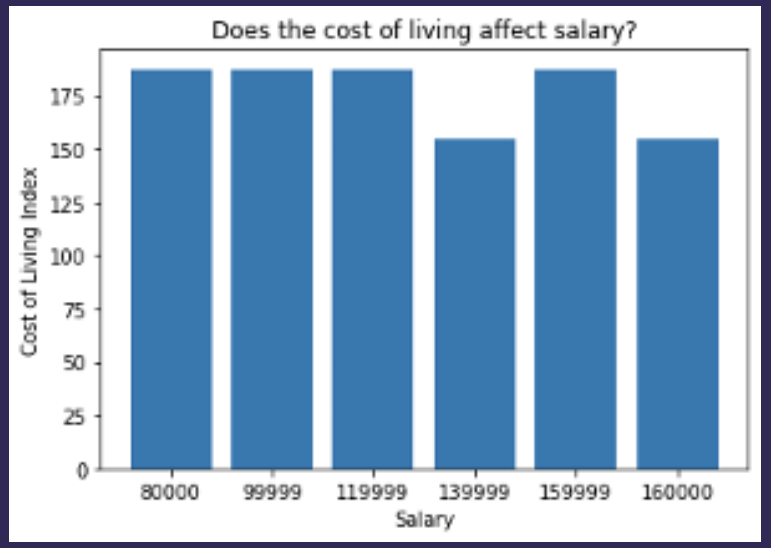
Maryland

DC

North Carolina

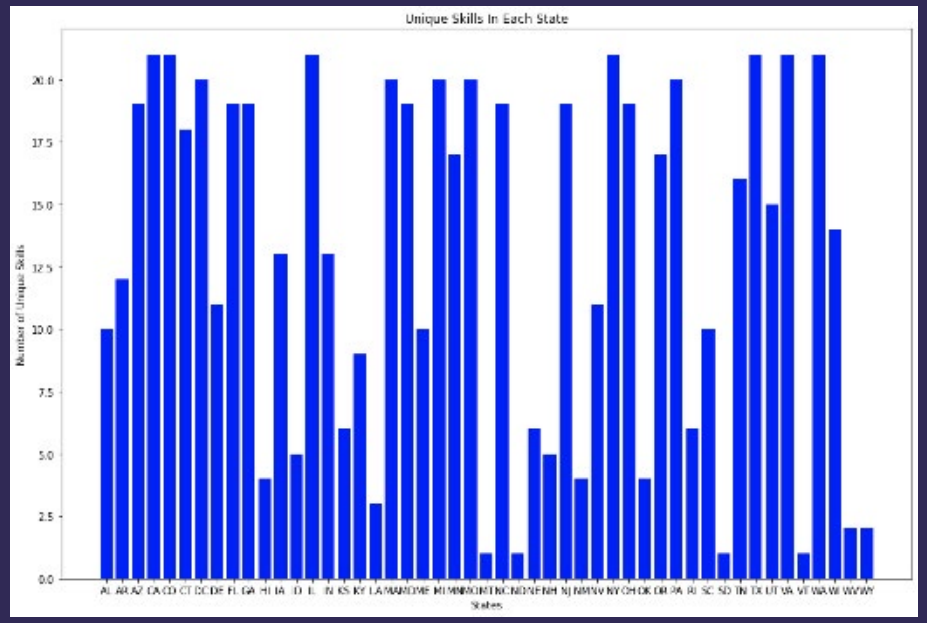
Location Insights





Where are the higher-paying jobs??

- The heatmap shows higher salaries in areas that have established or emerging technology centers.
- Examples of emerging areas include Utah (Silicon Slopes), South Carolina (Silicon Harbor)
- The darker areas tend to also have a higher cost of living, but comparing the salary to the Cost of Living Index didn't show a significant relationship between the two.



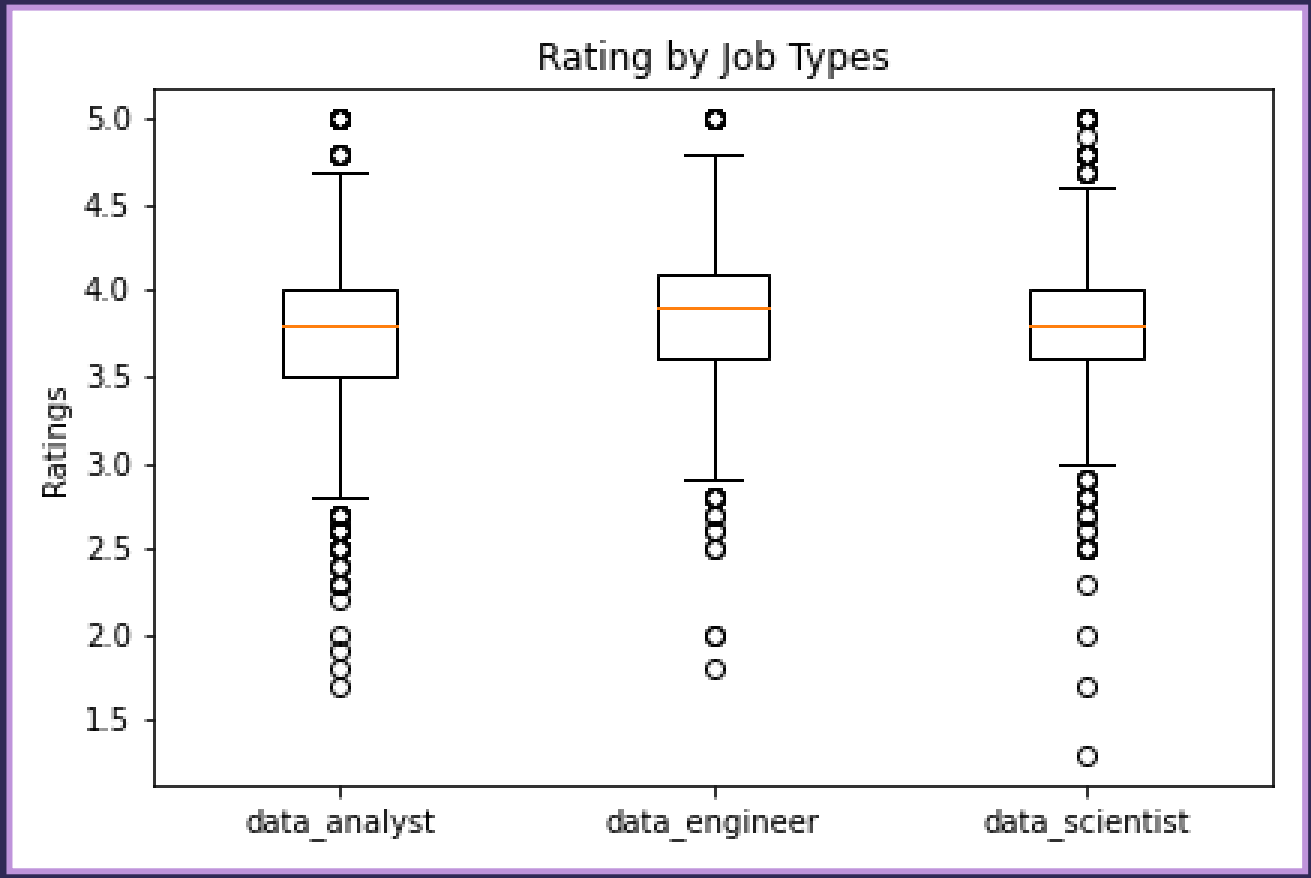
Are there areas of the country that require more data analytics skills?

- The states that house jobs that require a higher number of data analytics skills generally align with the areas that include jobs with higher salaries.
- Courses that teach multiple coding languages and skill sets are helpful in being qualified for higher paying jobs.

Employer Data

How happy are employees with these businesses?

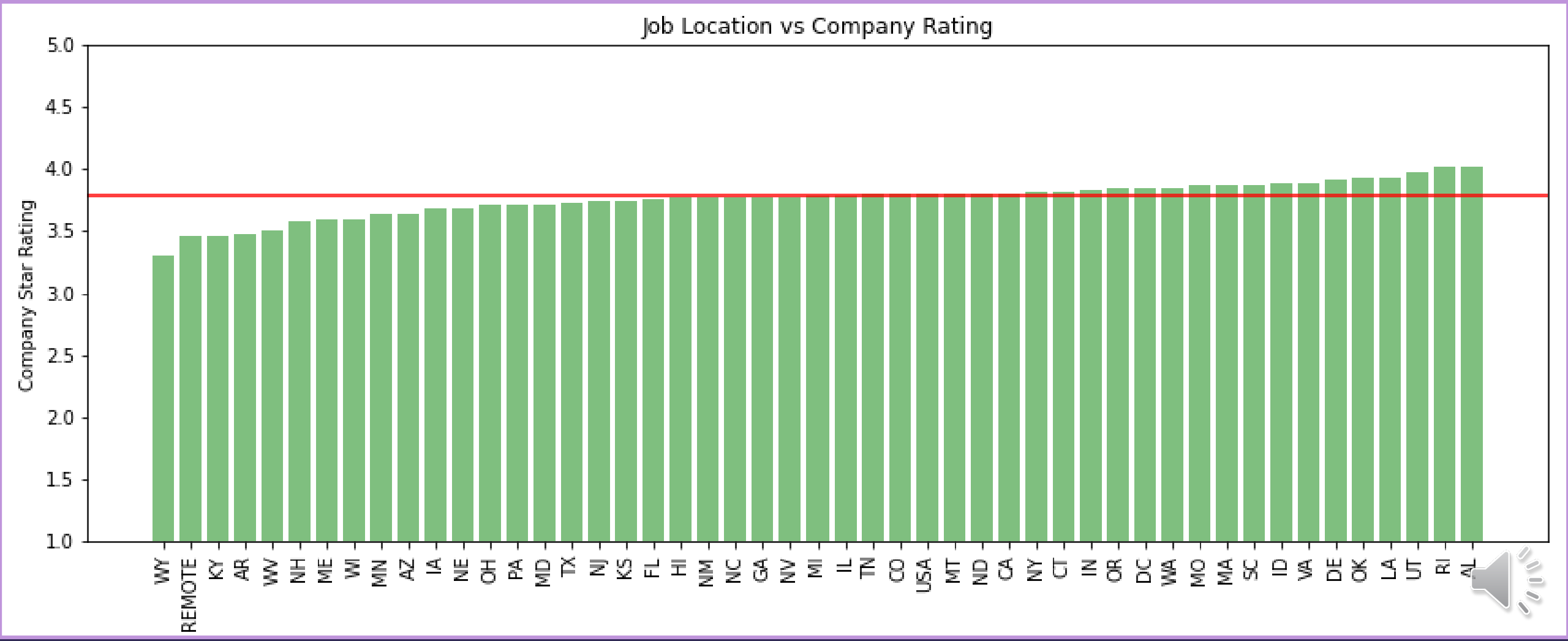
Company Ratings by Job Types



- Overall lowest rating: 1.30
- Overall mean: 3.78
- Overall highest: 5.0
- Excluding outliers, most data ranged between 2.7 to 4.7

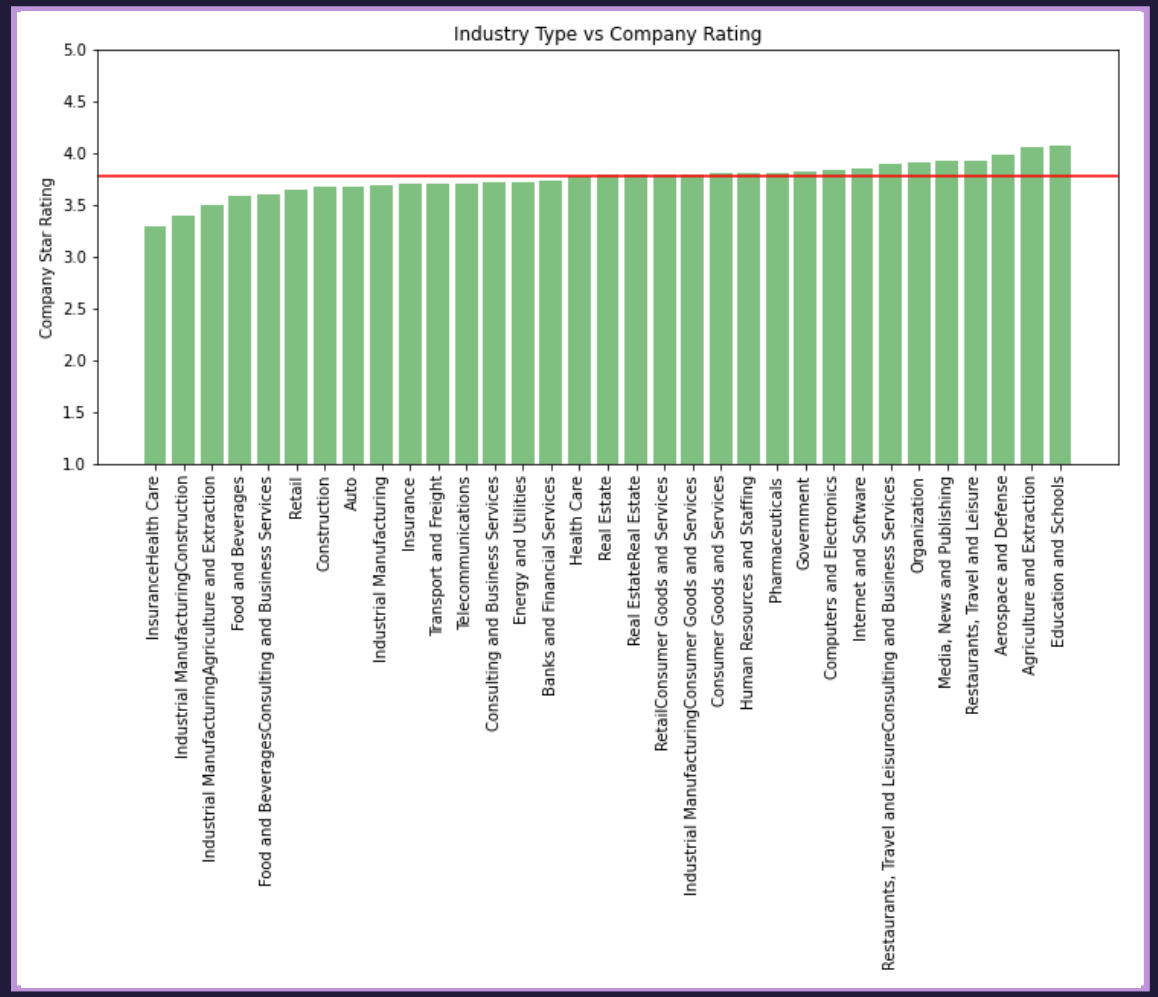
Employee happiness by position location

- Lowest company rating: Wyoming (3.30)
- Mean company rating (3.78): Michigan & Illinois
- Highest company rating: Alabama (4.03)

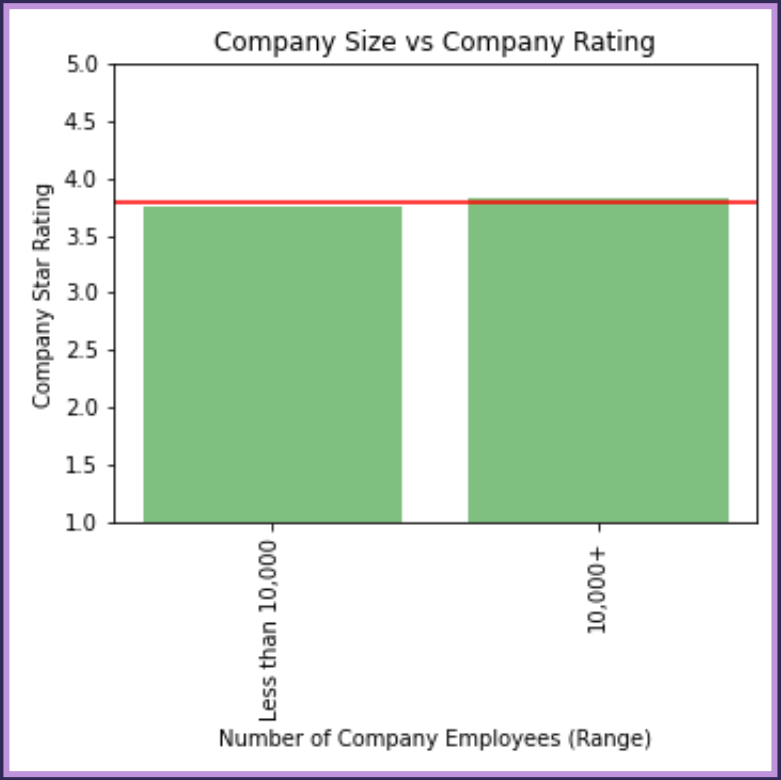


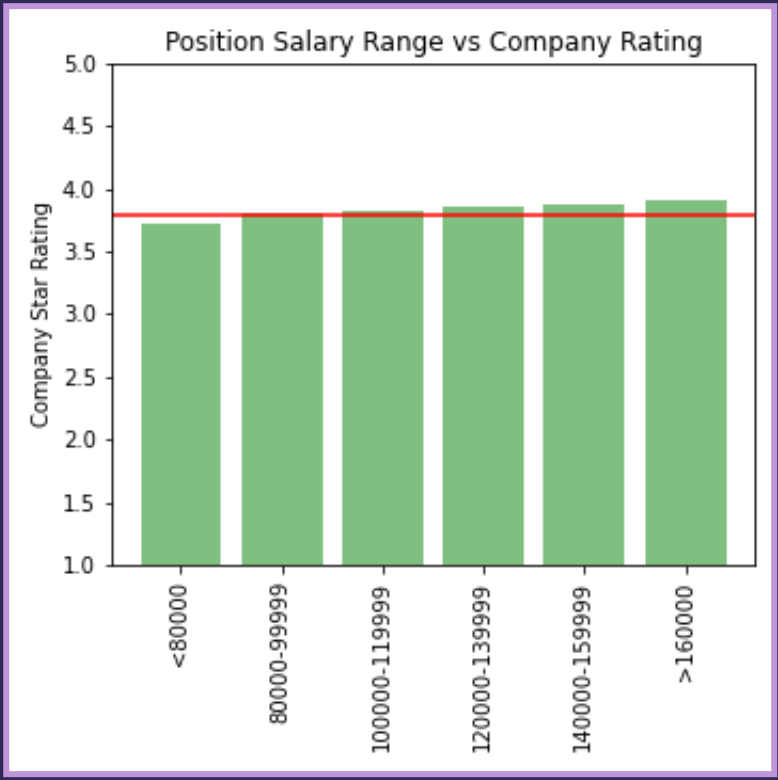
Employee happiness by Company Industry

- Lowest company rating: Insurance Health Care (3.30)
- Mean company rating (3.78): Health Care
- Highest company rating: Education and Schools (4.08)



Employee Happiness Analysis without Results

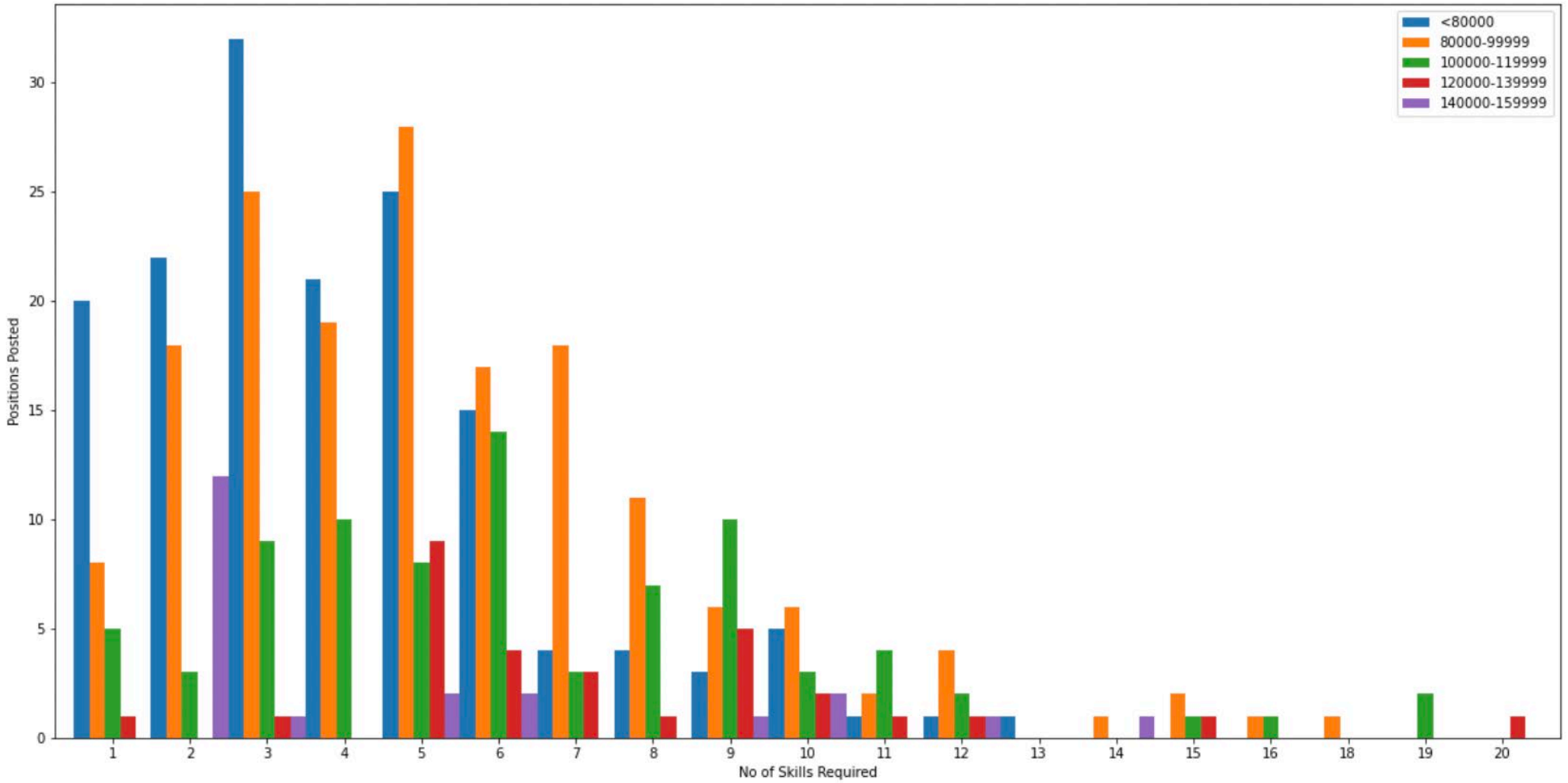




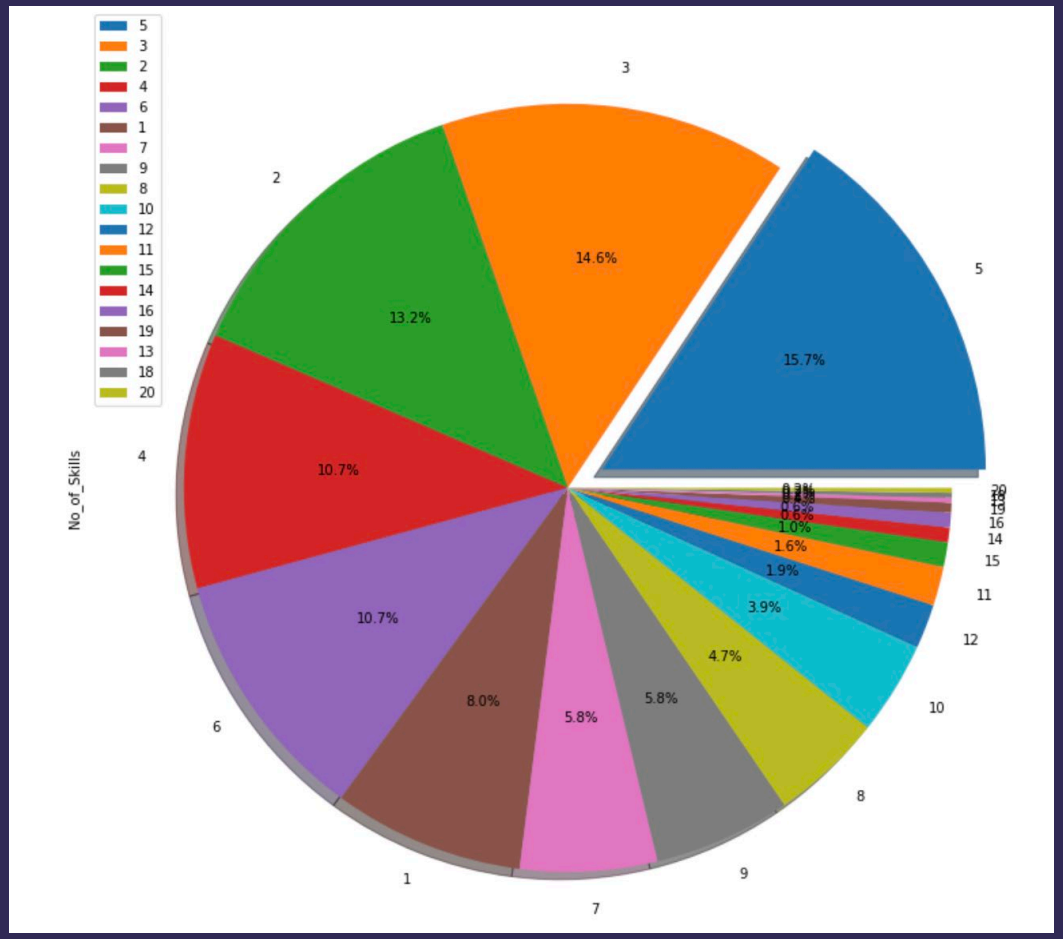
Salary Insights

Number of Skills Desired per Salary Range

- One to Five skills represented a large portion of the positions posted.
- The two higher dollar brackets had far fewer posting. A bracket for "over $160 existing for the Data Engineer and Data Scientist positions but the top end for Data Analytics was the $140k-$159,999

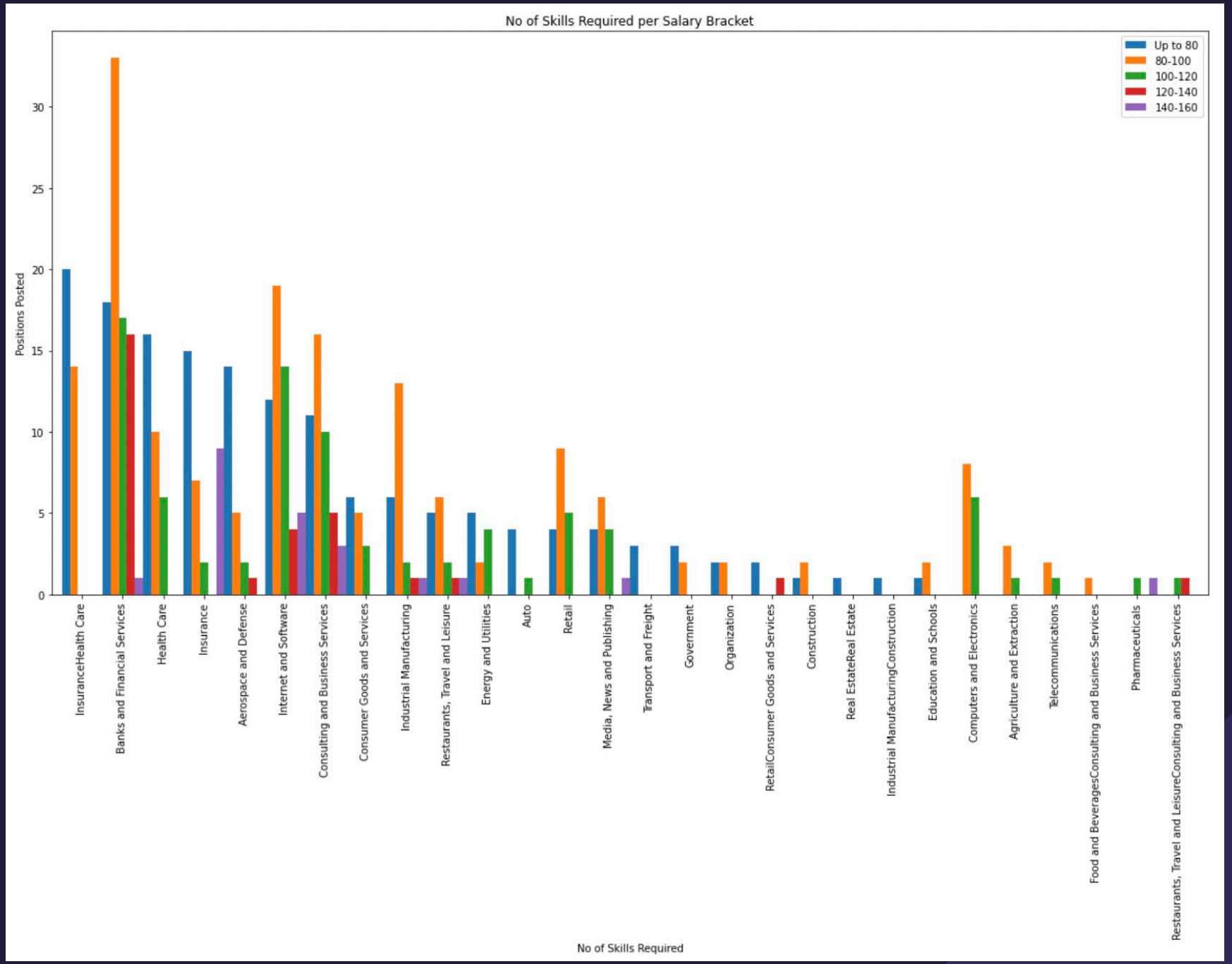


Number of Skills Represented as a Percentage of Job Postings



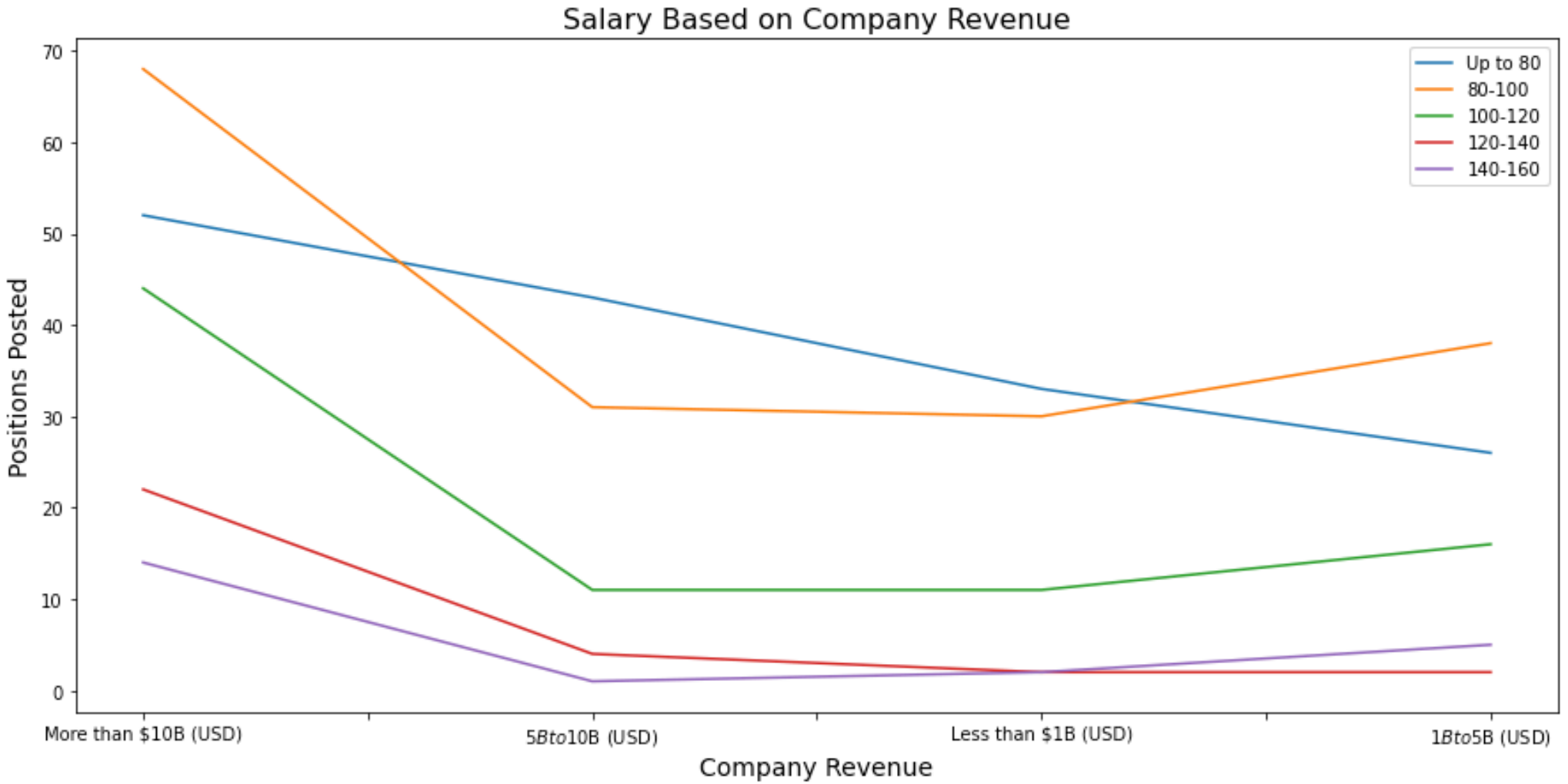
- As seen in the previous slide the number of skills desired with the highest percentage was three skills.
- As mentioned earlier in the presentation, the three skills most commonly sought by employers were Python, Sequel, and Machine learning, r, and Hadoop.

Salary based on Industry



Salary Range Based on Company Revenue

- A higher number of jobs are posted by companies with a company revenue of "More than $10B."
- It also seems, based on the data, companies with the higher revenues seem to be more likely to post positions with a higher salary range.



Thank You

Any questions??