DAA Internship - Data Analyst Job Board Post Insights

Job seekers, are you interested in knowing which job boards post the most?

by: <u>Hafsa Masood</u>, <u>Madeeha Umar</u>, <u>Jon Ekroth</u>, <u>Shari Chin</u>

10 Key Insights

1. The most popular day job posts are uploaded is **Thursday**.
2. The most popular job post upload time is **10 AM UTC.**
3. Job listing locations were only shown for 14 states and 1 District with **Missouri** having the most job posts at **3,623**.
4. Further analysis shows MO to have **86 cities**.
5. Full-Time is the top work arrangement companies are looking for at **73%**
6. The top 5 companies that post the most data analyst jobs are **Walmart, Talentify.io, Edward Jones, Corporate, Cox Communications.**
7. The job boards with largest number of listings are **LinkedIn, Upwork, BeBee, Trabajo.org, and ZipRecruiter.**
8. The most popular time job post upload posts for LinkedIn and Upwork is **3 PM UTC**.
9. A large number of job listings are posted on Saturday and Sunday **with 23.3 %**
10. Finance is the top industry with the highest average salary at **109k.**

Project Goals

Team Goals:

- Create at least (1) unique line of python code to help automate process
- Provide value to job seekers and larger data community, so that job seekers can take a more informed approach in their job searching process
- Provide value by creating content for stakeholders
- Increase skill set using various tools and build positive relationships

Main Questions:

- Which job boards post the most data analyst jobs?
- What are the top 10 companies that post the most data analyst jobs?
- What are the most in-demand skills for the top 25 companies that post the most data analyst jobs?
- What are the most popular times that companies post jobs?
  - Do the top 2 job boards list at different times?
- What job schedule types have the highest salary potential?
- Are any job skills listed more than others?
- Which states are hiring the most?
- What industry has the highest salary?
- Which title level are companies hiring the most for?

Data Summary

Summary:

- Originally 34,045 rows, 23 columns
- Downloaded 11/02/23
- Dataset consists of job posts over 13 month period
- Data source by Luke Barousse

Tools used:

- Python - Data Cleaning, Transformation and Exploration
- Excel - Data Cleaning and Exploration
- Tableau - Analysis and Visualization

Click on links to view more

Data Summary

Limitations and Challenges:

- Small sample size with only 18% of the job postings having a salary listed
- Many company names didn't make sense, such as "Corporate" or "2,468 reviews"
- Not possible to group companies into industries with inconsistent company names (see above)
- Data cleaning/engineering took up majority of time
- Difficult to know the amount of times a job board posts their job on different platforms
- Difficult to work through the job description column without machine learning
- Very few specific job locations are present and a small number of US states are represented
- Extracting the exact time of job posting proved difficult

https://www.kaggle.com/datasets/lukebarousse/data-analyst-job-postings-google-search/

Our Team

Shari Chin: Project Lead, Data Cleaning, Data Transformation and Analysis in Python

Jon Ekroth: Data cleaning, design visuals in Excel, assist with Tableau visuals and troubleshooting

Madeeha Umar: Analyzed data in Tableau, designed dashboard, created graphs and charts for insights
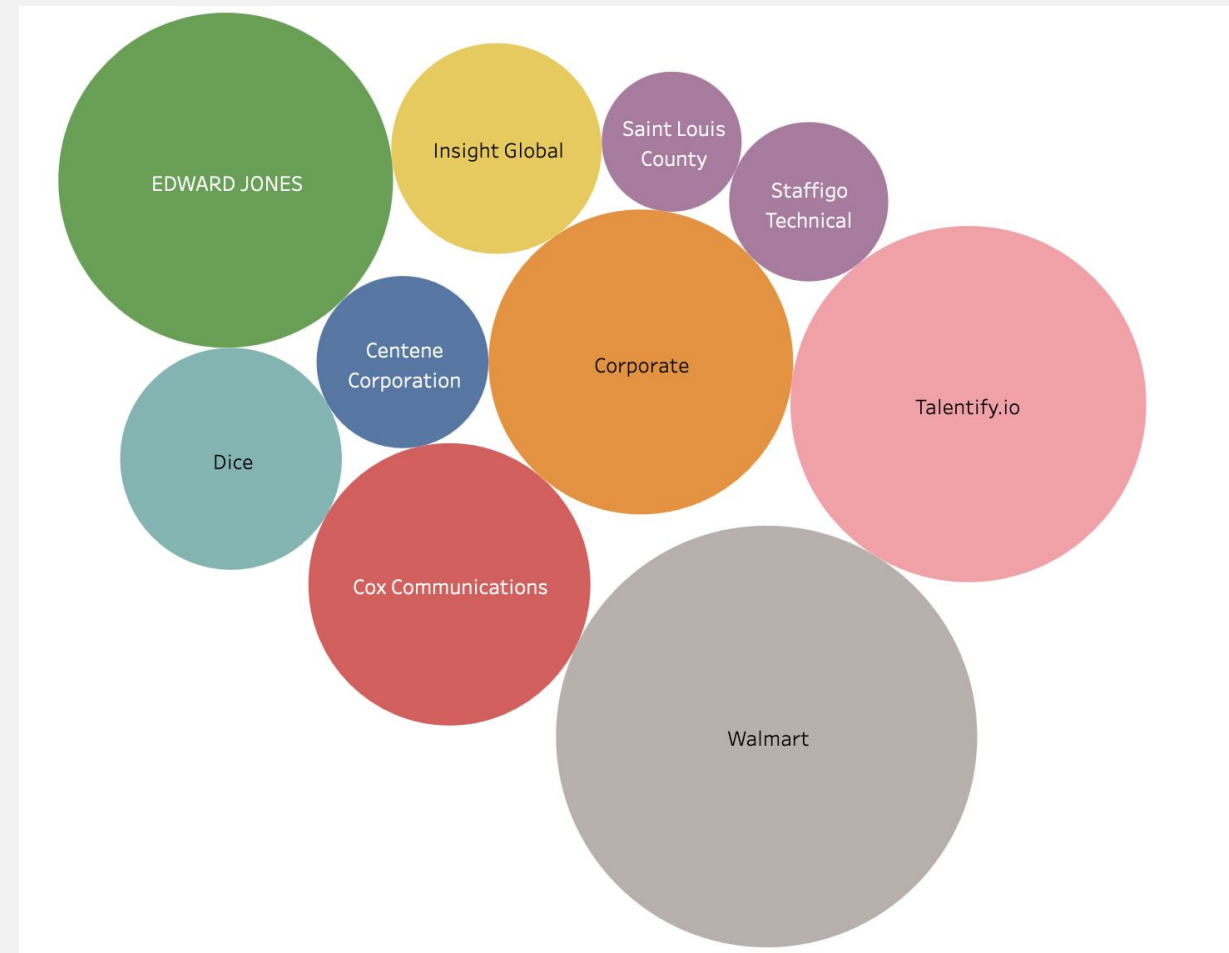
Hafsa Masood: Data Cleaning, Data Transformation, and Analysis in Python

Diving Deeper into the Insights...

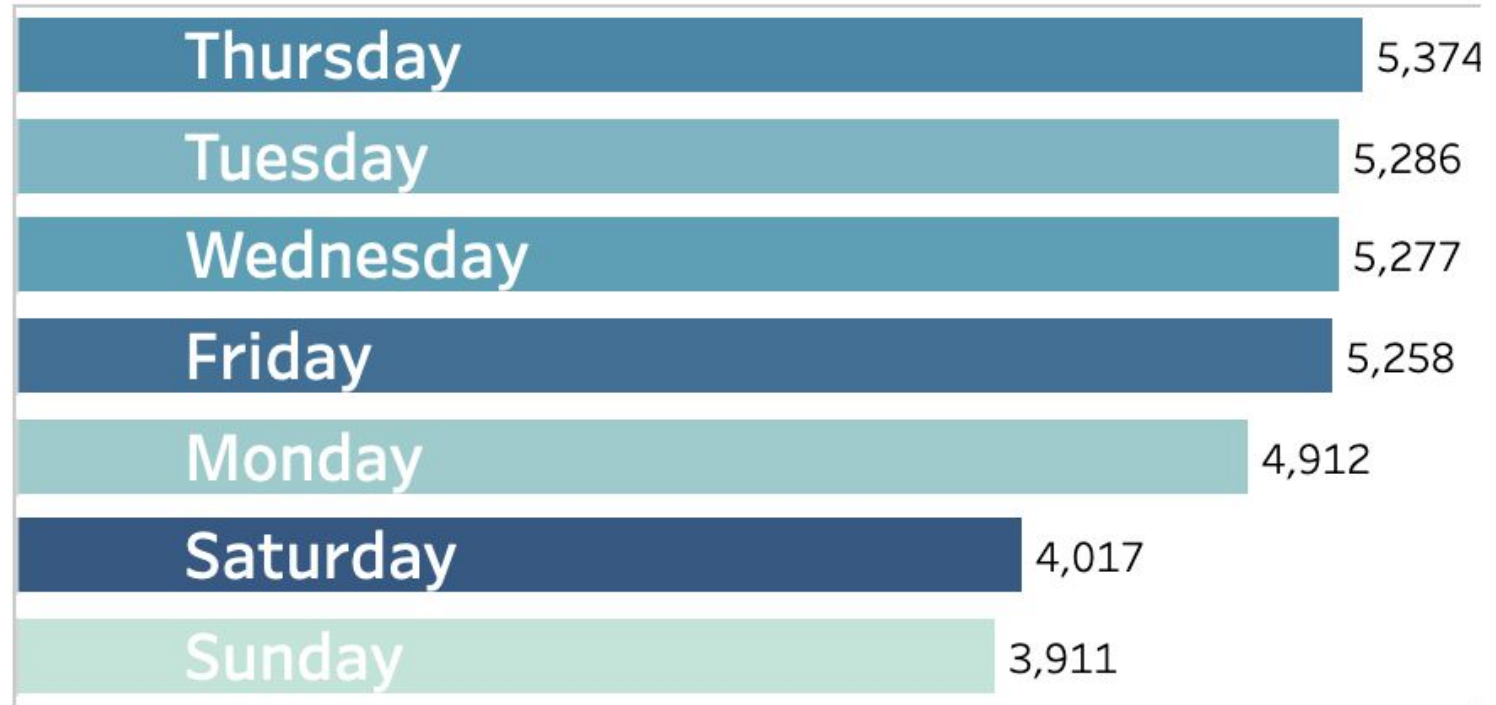# Top 10 Companies that post Data Analyst Jobs

(In descending order)

1. Walmart
2. Talentify.io
3. Edward Jones
4. Corporate
5. Cox Communications
6. Dice
7. Insight Global
8. Centene Corporation
9. Staffigo Technical Services
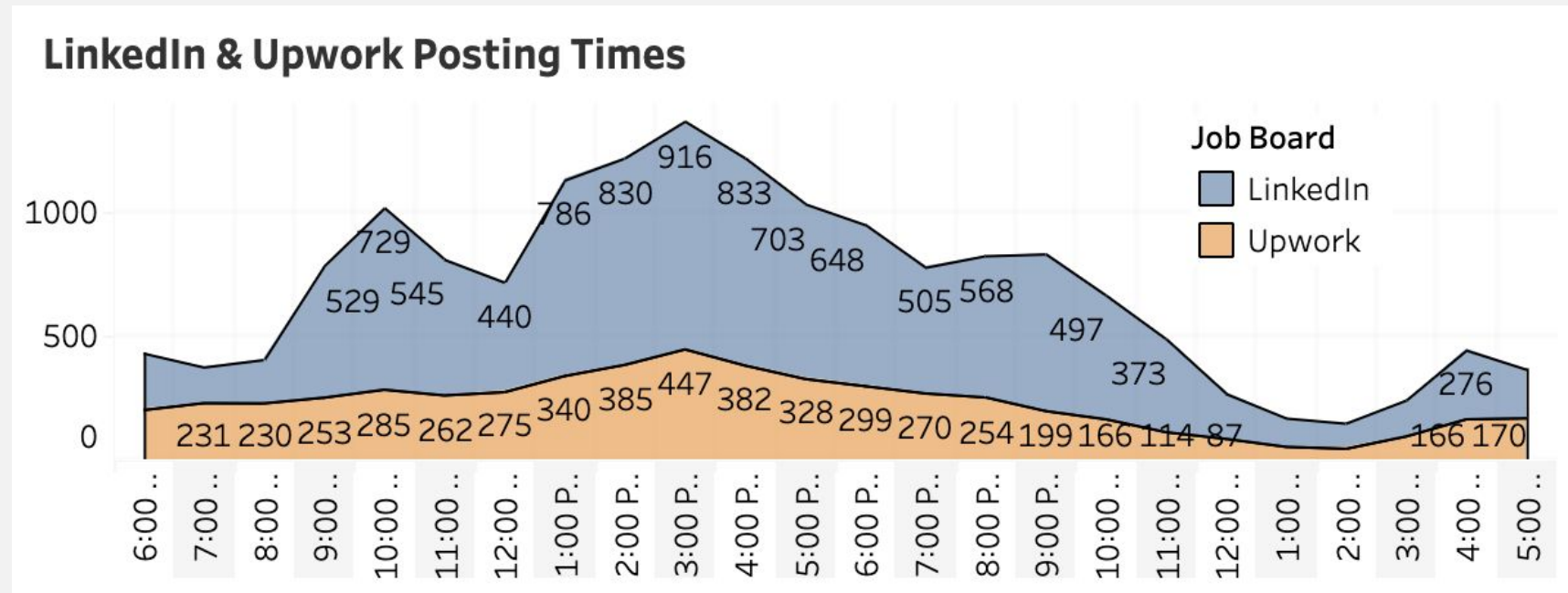10. Saint Louis County Clerks Office

Number of Jobs Posted by Day of the Week

Count of Job Board vs Day

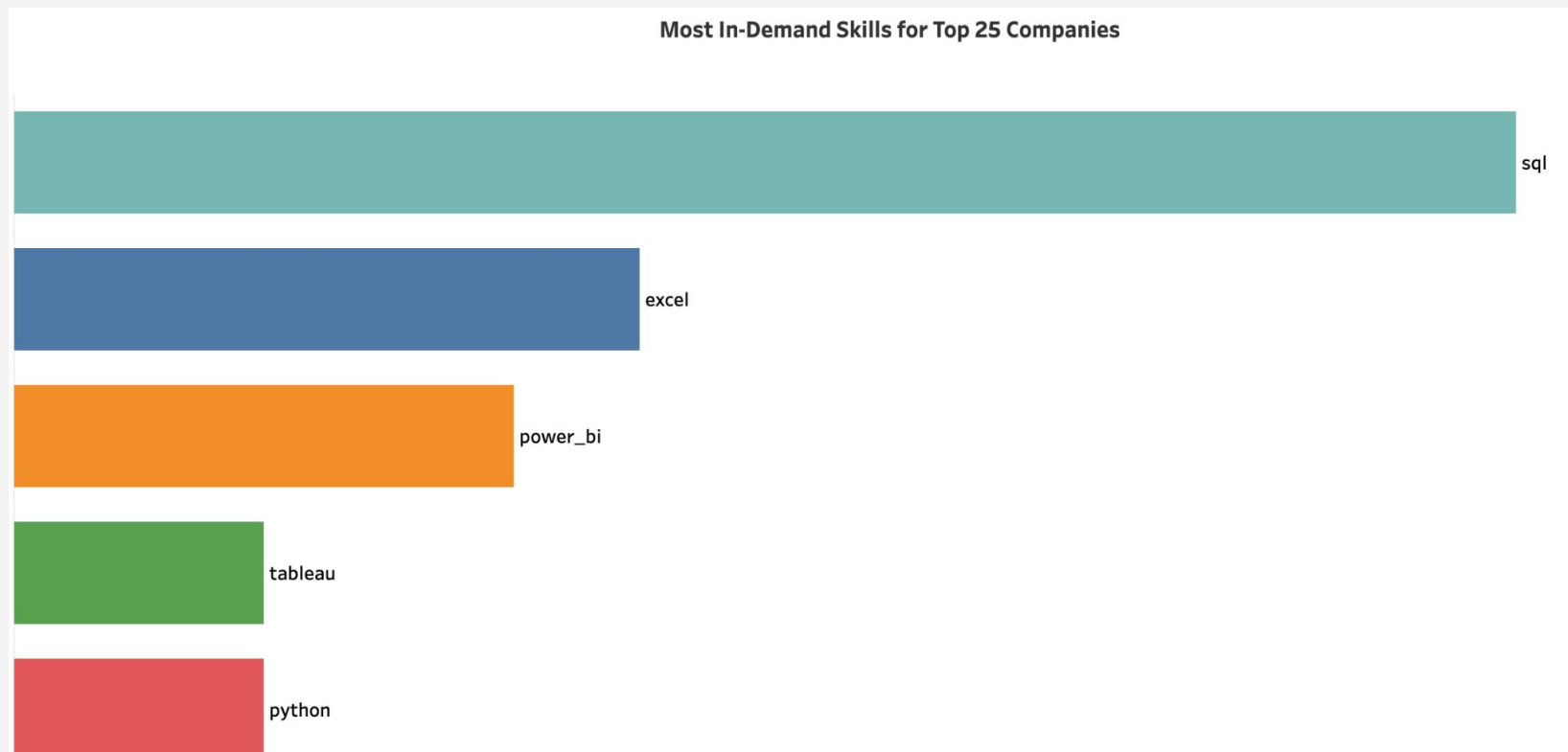| Day	Count
Thursday	5,374
Tuesday	5,286
Wednesday	5,277
Friday	5,258
Monday	4,912
Saturday	4,017
Sunday	3,911

Thursday is the most popular day to post jobs.

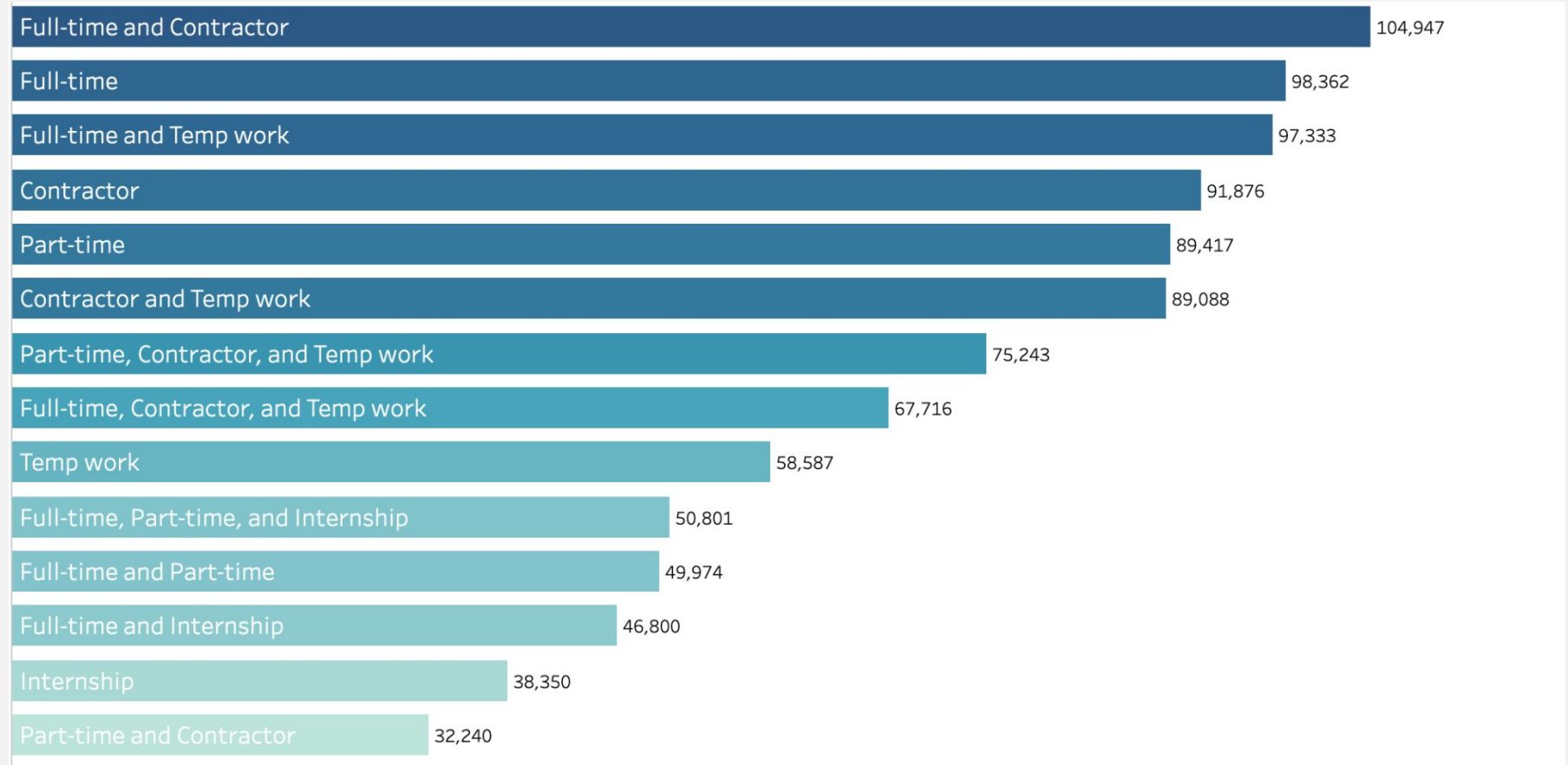# Time of Day that jobs are posted on LinkedIn & Upwork

LinkedIn and Upwork have the most jobs posted on them at 3 PM UTC.

In-Demand Skills for the Top 25 Companies that Post Data Analyst Jobs

SQL is the most in-demand skill for companies that post the most data analyst jobs.

Salaries by Schedule Type

Schedule Type	Salary
Full-time and Contractor	104,947
Full-time	98,362
Full-time and Temp work	97,333
Contractor	91,876
Part-time	89,417
Contractor and Temp work	89,088
Part-time, Contractor, and Temp work	75,243
Full-time, Contractor, and Temp work	67,716
Temp work	58,587
Full-time, Part-time, and Internship	50,801
Full-time and Part-time	49,974
Full-time and Internship	46,800
Internship	38,350
Part-time and Contractor	32,240

Data Analysts who are full-time and contractors make the highest salaries.

Data Cleaning Summary

- Removed duplicates
- Recognized outliers in dataset by filtering in tables
- Removed (8) columns not needed for analysis:
  - Unnamed, job id, salary pay, search location, commute time, search term
  - See code for more
- Filled null values
- Removed punctuation
- Transformed and added (7) new columns:
  - Title level, title type, day job board posted, time job board posted, job location by state, industry column, cleaned schedule type
- Grouped categorical data
  - Title level grouped by keywords into 3 levels: EN, MD, SR and ANY
  - Title type grouped by keywords into (8) categories: Engineer, Architect, Scientist, Analyst, Developer Programmer, Specialist, Writer
  - Schedule type grouped into (4) categories: Full-time, Part-time, Contractor and Other
  - Description tokens column separated skills per job posting for the largest 25 companies
  - See code for more
- Changed data types strings and decimals

Lessons Learned

If we could go back in time, what could we have done different? Where could we make improvements?
- Deciding on a data rubrik within the group prior to starting i.e deciding on rules for outliers
- Diving deeper into the insights with more analysis on the salary distribution compared to a larger sample size

Thank you!

Fonts used: Barlow and **Anton**