

Data Visualization Assignment: Analyzing Data Science Salaries

Prashamsa Rijal
Student Number: 23140743
Total Word Count: 985
June 12, 2024
June 13, 2024

Contents

Abstract	4
1 Introduction	5
1.1 Summarise	5
1.2 Highlight	5
1.3 Aim	5
1.4 Achievements	5
1.5 Organised	6
2 Motivation and Objectives	7
2.1 Motivation	7
2.2 Description of Selected Data	7
2.3 Objectives	7
3 Questions about Dataset	8
4 Experimental Results	9
4.1 Salary Comparison of Data Scientists and Data Engineers based on expertise levels in the United States	9
4.2 Varaince of Salary in Different Employment Type	12
4.3 Salary Trend Over Time by Based on Experience Level	14
4.4 Change in company Size Over Time	15
4.5 Salary Expectations accross different Jobs	16
5 Summary	18
6 References	19
References	19

List of Figures

Abstract

The "Latest Data Science Salaries" dataset provides valuable insights into the compensation trends and variations in the field of data science from 2020 to 2024. This dataset encompasses a comprehensive collection of salary information from various industries, organizations, and geographic regions, enabling data professionals, researchers, and organizations to analyze and understand the prevailing salary landscape in the data science domain during this four-year period. By examining this dataset, one can gain a deeper understanding of the factors influencing data science salaries, such as job roles, experience levels, educational backgrounds, and geographical locations.

The dataset serves as a valuable resource for individuals seeking career guidance, companies aiming to benchmark their compensation strategies, and researchers investigating the evolving dynamics of the data science job market.

1 Introduction

1.1 Summarise

This report is based on analyzing data science salaries from 2020 to 2024. The dataset includes detailed information on various job roles, experience levels, expertise levels, salaries, company characteristics, and geographical locations. The aim is to provide a comprehensive understanding of the salary trends and factors affecting compensation in the data science field.

The report is designed for readers with a medium level of understanding of data analysis and visualization. Additionally, some advanced topics are included for more experienced readers. The report addresses several key questions and provides conclusions based on the dataset.

1.2 Highlight

Overall, there are several interesting observations in this report:

- Trends in salary growth over the four-year period.
- The impact of job roles and experience levels on salaries.
- Regional variations in data science compensation.
- The influence of company size and industry on salary levels.

These highlights provide a snapshot of the key findings from the dataset. Different readers may find different aspects more interesting, reflecting the diversity of perspectives in the field of data science.

1.3 Aim

The aim of this report is to explore and answer some of the most common questions regarding data science salaries. It investigates whether certain factors such as job role, experience level, educational background, and geographical location significantly impact compensation.

Another important goal is to dispel common misconceptions about data science salaries and provide a data-driven perspective on the job market dynamics.

1.4 Achievements

This report offers answers to several popular questions regarding data science salaries. It includes visualizations and analyses that map out salary distributions across different regions, identify the most lucrative job roles, and track salary changes over time.

Some theories about factors influencing salaries are confirmed, while others are challenged based on the data. For example, the analysis may show that experience level has a more significant impact on salary than educational background, or that salaries are higher in certain regions regardless of company size.

1.5 Organised

The structure of this report is as follows:

1. Title
 - (a) Subtitle (if applicable)
 - (b) Description
 - (c) Graphs (graphs are on their own page)

This organization ensures that the information is presented in a logical and easy-to-follow manner, allowing readers to efficiently navigate through the various sections and find the insights they are most interested in.

2 Motivation and Objectives

2.1 Motivation

The motivation behind this report stems from the growing interest in the field of data science and the increasing demand for data science professionals. Understanding salary trends and the factors that influence compensation is crucial for both job seekers and employers. This dataset, containing 3300 rows and 11 columns, offers a comprehensive overview of the salary landscape in data science from 2020 to 2024, making it an invaluable resource for detailed analysis and insights.

2.2 Description of Selected Data

The "Latest Data Science Salaries" dataset provides detailed information on various aspects of data science jobs. The dataset contains 3300 rows and includes the following 11 columns:

- Job Title
- Employment Type
- Experience Level
- Expertise Level
- Salary
- Salary Currency
- Company Location
- Salary in USD
- Employee Residence
- Company Size
- Year

This extensive dataset enables a thorough analysis of the various factors that impact salaries in the data science field.

2.3 Objectives

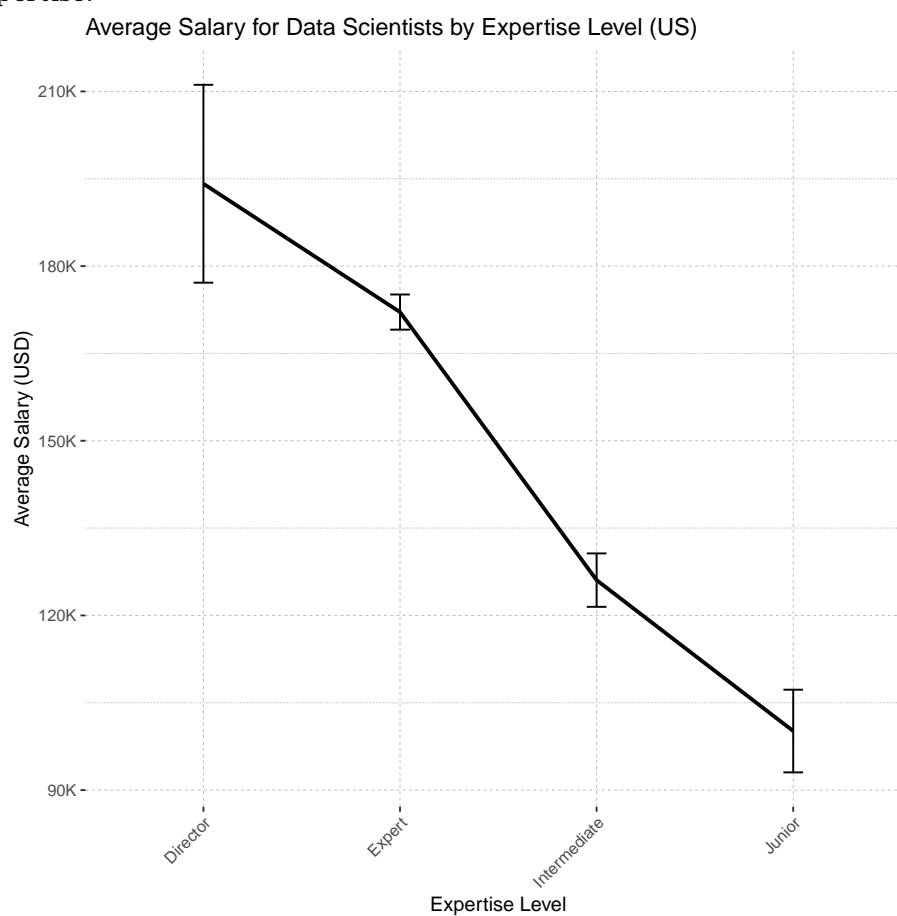
The primary objective of this report is to analyze the "Latest Data Science Salaries" dataset to uncover trends and patterns in data science compensation. Specifically, it aims to identify the most significant factors influencing salaries, examine salary evolution over the four-year period, provide insights into regional variations, and offer data-driven conclusions to dispel common misconceptions about data science salaries.

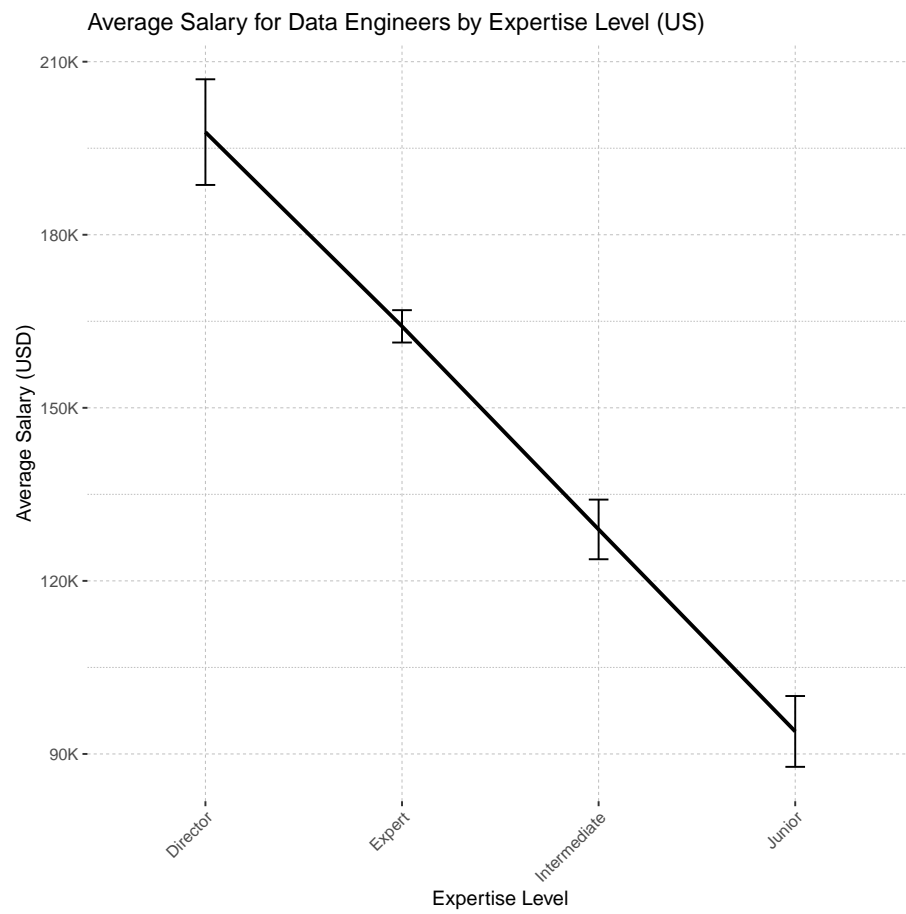
3 Questions about Dataset

4 Experimental Results

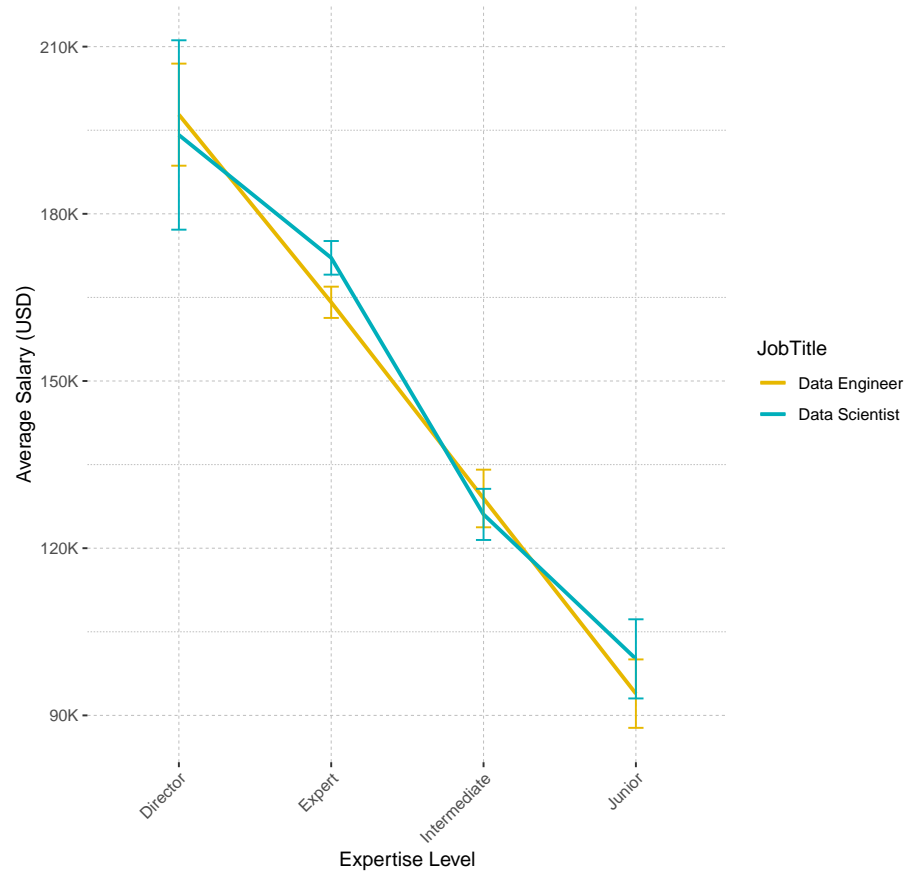
4.1 Salary Comparison of Data Scientists and Data Engineers based on expertise levels in the United States

This visualization explores the relationship between expertise level and salary for Data Scientists and Data Engineers in the United States. As the expertise level is good the salary is seen to be improved as well in both of the cases. While comparing the Data Scientist Job is seems to have high income in case of expert and intermediate expertise.





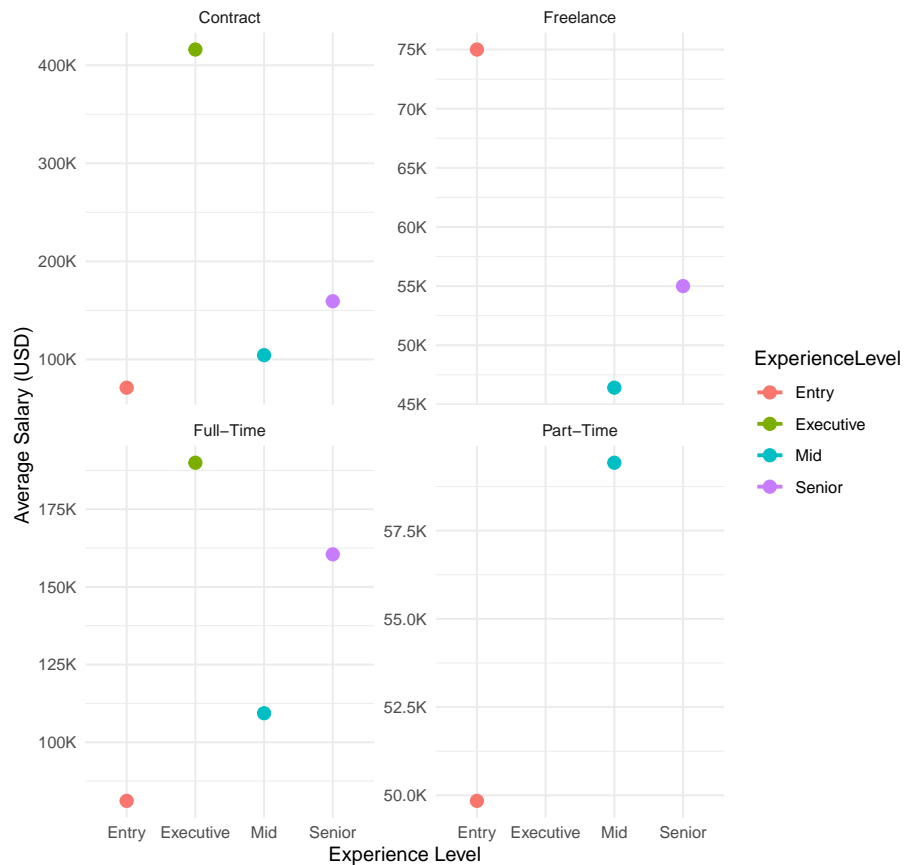
Average Salary for Data Scientists and Data Engineers by Expertise Level (US)



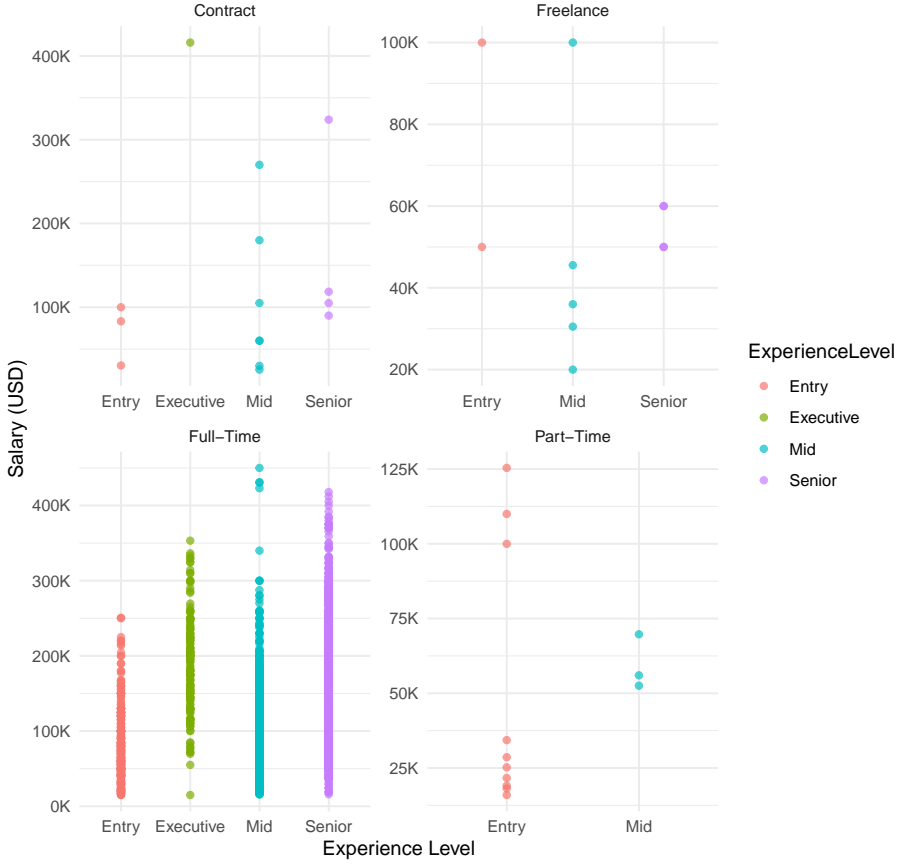
4.2 Varaince of Salary in Different Employment Type

This visualization shows how much data professionals earn based on their experience and the type of work they do (full-time, part-time, etc.). The plots show that the people who do full time earn the highest if they are executive or mid level. In other cases (contract and freelance) the mid experience is the most earning.

Salary vs. Experience by Employment Type



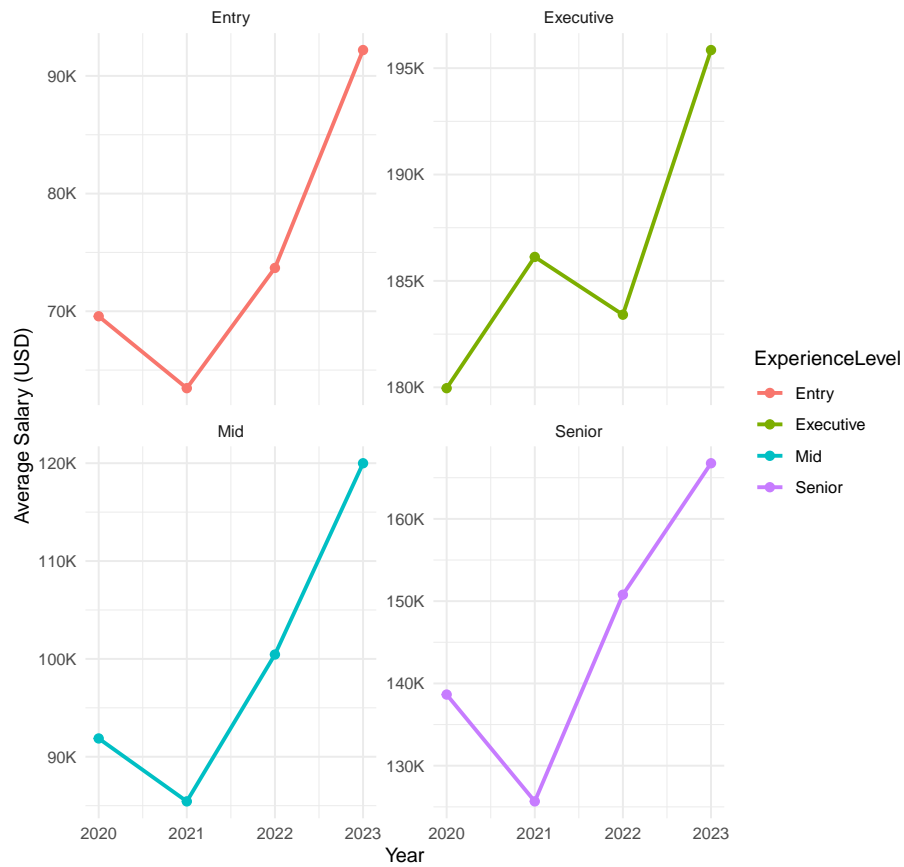
Salary vs. Experience by Employment Type



4.3 Salary Trend Over Time by Based on Experience Level

This visualization tracks how average salaries have changed over the years for data science professionals, grouped by how much experience they have. From visualization we can observe that there is increment in money over the time after 2021 in 2021 there was a fluctation in salary for all the job holders.

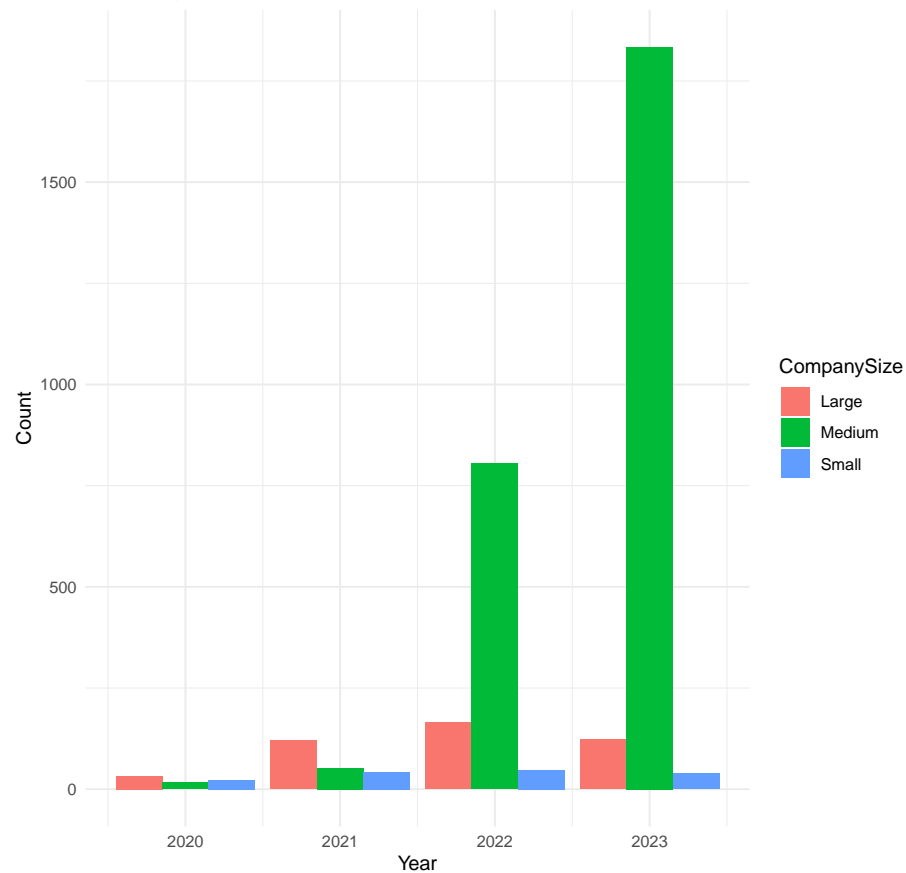
Salary Trends by Experience Level



4.4 Change in company Size Over Time

This visualization shows how the number of data science companies of different sizes has changed over the years. The bar chart shows how many companies are in each size category for each year. We can see that medium sized company has increased alot compared to others over the year although at first there were more large companies.

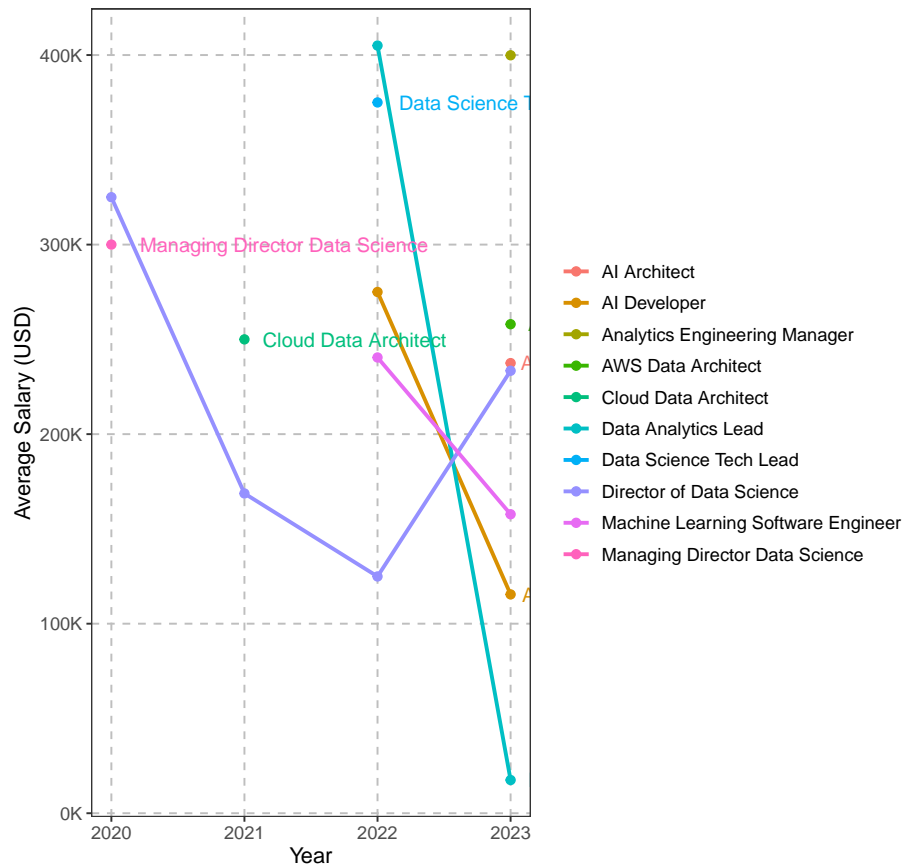
Company Size Distribution Over Time



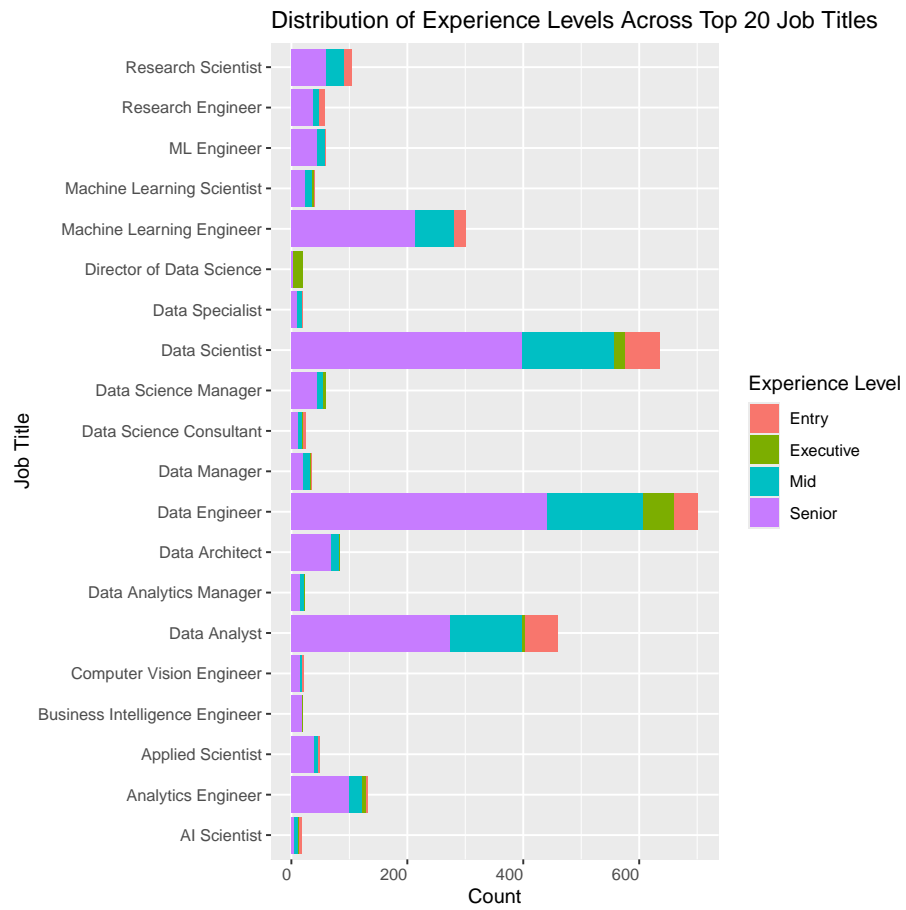
4.5 Salary Expectations accross different Jobs

This visualization shows how salaries for the 10 highest-paying data science job roles have changed over the years. Some jobs were highest paying for a year only then it was no more seen in other years whereas the one job that is seen from 2020 till 2023 is Director of Data Science. Compared to 2022 the income has decreased for many of the job roles.

Salary Trends for Top 10 Data Science Job Roles



FALSE 'summarise()' has grouped output by 'JobTitle'. You can override using the FALSE '.groups' argument.



5 Summary

6 References

- KaggleDataset2023 Kaggle. (2023). Data Science Salaries 2023. Kaggle. Available at: <https://www.kaggle.com/datasets/iamsouravbanerjee/data-science-salaries-2023>
- AIJobsNet2023 AI-Jobs.net. (2023). Data Science Salaries. Available at: <https://ai-jobs.net/>
- Ramachandran2024 Ramachandran, K. K. (2024). Data Science in the 21st Century: Evolution, Challenges, and Future Directions. *International Journal of Business and Data Analytics*, 1(1), 1-15.