

10alytics Internship

Data Science Jobs Group 2

Salary

Presentation



Group 2

Nwaamaka Iduwe
Oluwatunmise jiboku
Olatove Shemilore
Amaka Ihenwe



Oluwakemi
Ogebule
Group Leader



Kuti Yetunde
Researcher



Korede
Oyewaga
Resource Collector



Christiana
Oyelowo
Presenter

Introduction

Data science jobs are some of the fastest-growing, most in-demand in technology. Since 2012, Data Scientist roles have increased by 650 percent, and this rise shows no sign of stopping. The U.S. Bureau of Labor Statistics predicts that the demand for data science skills will increase another 27.9 percent by 2026. And, according to a report from McKinsey, that spells a shortage of between 140,000 and 190,000 people with analytical skills in the U.S. alone — not to mention another 1.5 million managers and analysts who will be required to understand how data analysis drives decision-making.

Data Scientist salaries have also risen with demand; Data Scientists can typically expect to make six figures. Demand also translates into an ability to relocate far more easily — from city to city, and even internationally.



Objectives

1

The First Objective

Explore every feature in the dataset

2

The Second Objective

Work Year Analysis (with Salary, Remote Ratio)

3

The Third Objective

Experience Level Analysis (with Employment Type, Top 3 Job Title, Company Size)

4

The Fourth Objective

Company Location Analysis (with Experience Level)

5

The Fifth Objective

Salary Analysis (with Work Year, Experience Level, Company Size, Job Title, Remote Ratio)

Dictionary

work_year

The year the salary was paid.

experience_level

The experience level in the job during the year with the following possible values:

- EN Entry-level
- Junior MI Mid-level
- Intermediate SE Senior-level
- Expert EX Executive-level
- Director

employment_type

The type of employment for the role:

- PT Part-time
- FT Full-time
- CT Contract
- FL Freelance

Dictionary

job_title

The role worked in during the year.

salary

The total gross salary amount paid.

salary_currency

The currency of the salary paid as an ISO 4217 currency code.

salary_in_usd

The salary in USD (FX rate divided by avg. USD rate for the respective year via fxdata.foorilla.com).

employee_residence

Employee's primary country of residence in during the work year as an ISO 3166 country code.

remote_ratio

The overall amount of work done remotely, possible values are as follows:

- 0 No remote work (less than 20%)
- 50 Partially remote
- 100 Fully remote (more than 80%)

Dictionary

company_location

The country of the employer's main office or contracting branch as an ISO 3166 country code.

company_size

The average number of people that worked for the company during the year:

- S less than 50 employees (small)
- M 50 to 250 employees (medium)
- L more than 250 employees (large)

Description of Dataset

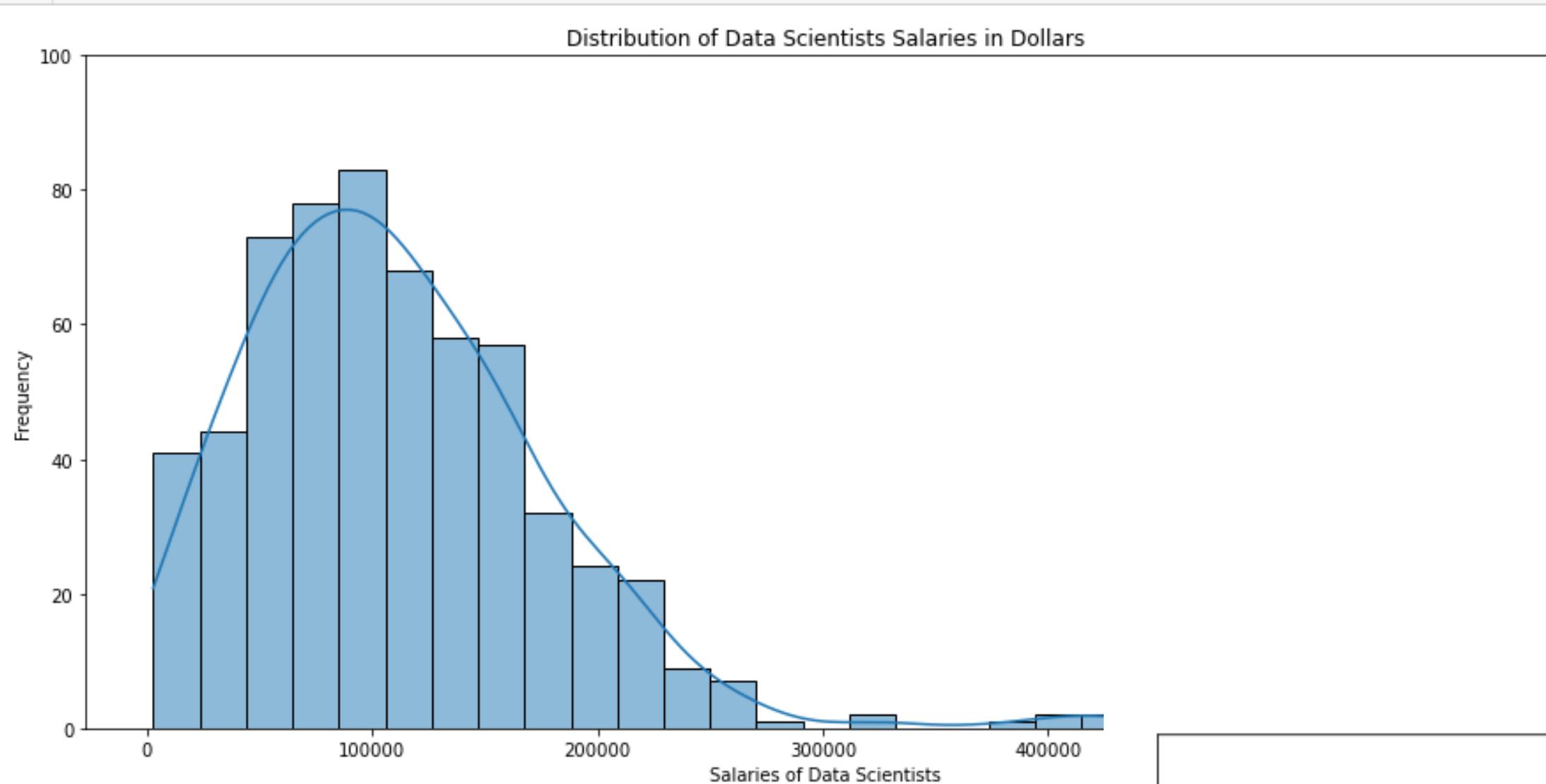
- work_year: The work year ranges from 2020 to 2022.
- experience_level: There are 4 unique experience levels. Intermediate SE Senior level tops the chart with a frequency of 280.
- employment_type: There are 4 unique employment types. Full-time(FT) constitutes the most with 588.
- job_title: There are 50 unique job titles. Data Scientist is the most common job title with the frequency 143.
- salary: The average gross salary of a data professional is 324000. The salary of a data professional ranges from 4000 to 30400000. There is a huge difference between the 75% percentile(165000) and the maximum value(30400000). This indicates that there are outliers in this column. There is a great difference between the mean and the median which means there are outliers.
- salary_currency: There are 17 different currencies with USD as the most received salary currency in the dataset.

Description of Dataset

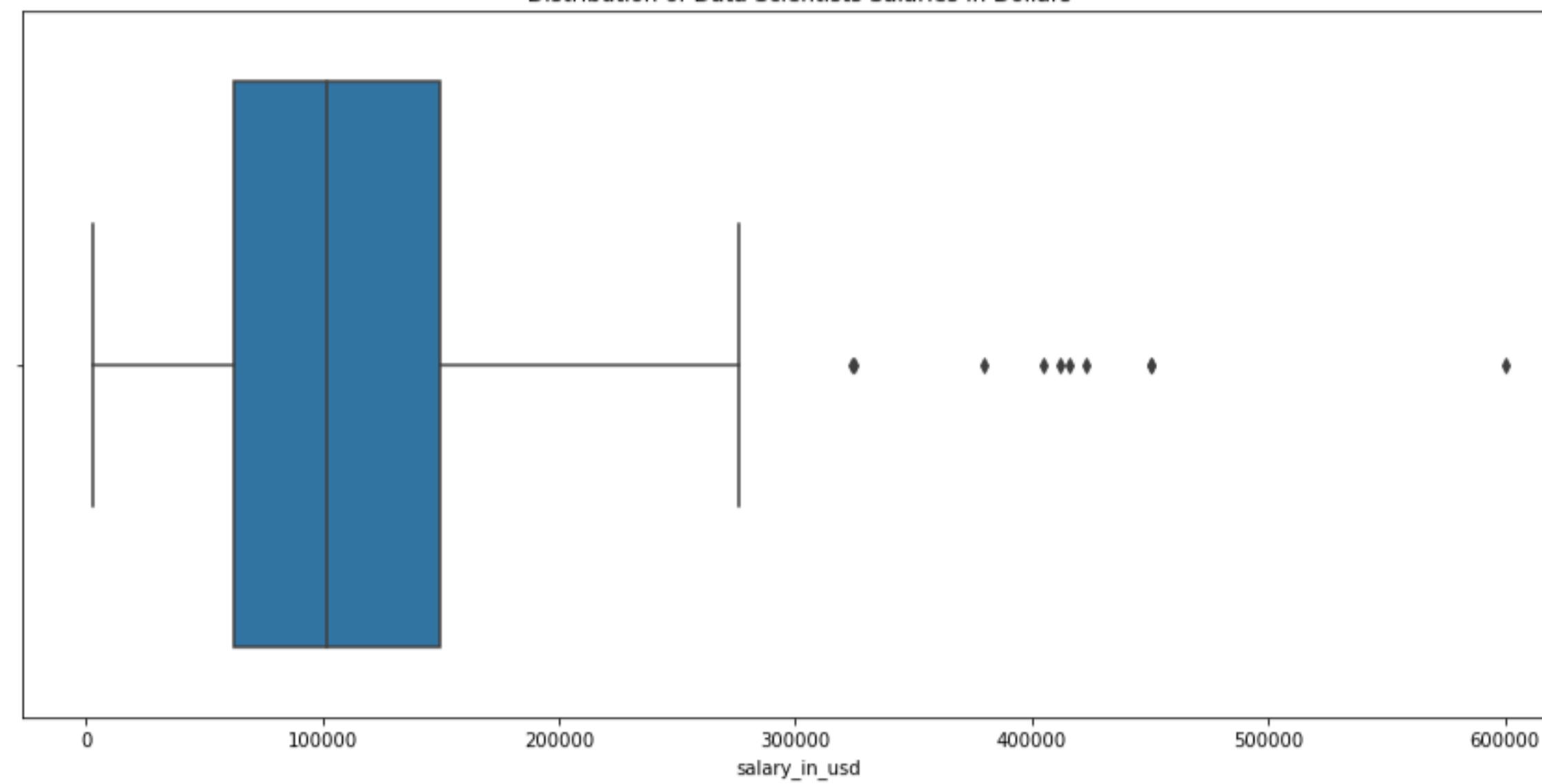
- salary_in_usd: The average salary in usd is 112297. The salary in dollars ranges from 2859 to 600000. There is not much difference between the average and the median salary in dollars but there is a huge difference between the 75% percentile(150000) and the maximum usd salary(600000) which indicates the presence of outliers in this column.
- employee_residence: United States(US) has the highest count of data professionals.
- remote_ratio: Half of the workers are working fully remote.
- company_location: Companies located in the United States(US) are more than those in other countries.
- company_size: There are 3 unique company sizes. There are more medium-sized companies in the dataset.

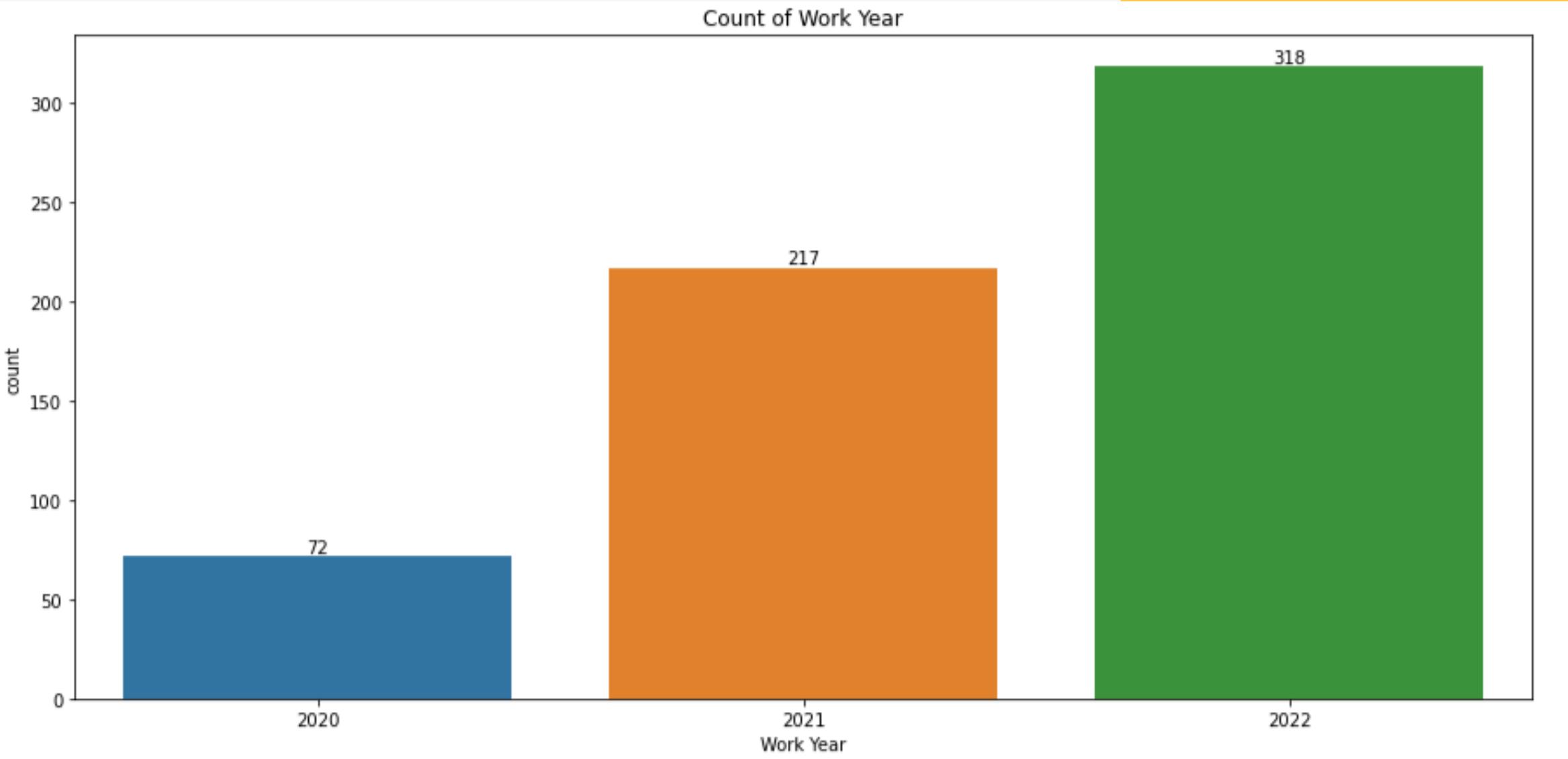
Visualizations and summaries

Distribution of Data Scientists Salaries in Dollars



- The distribution is skewed towards the right as a large number of occurrences appear on the left side of the plot.
- Majority of data scientists earn between 60,000usd and 160,000usd.
- We can see that as the salaries increase from 200,000 usd the number of data scientists decrease.
- There are outliers present in this column.
- The median salary in dollars is around 100,000 usd. 75% of data scientists earn less than or equal to 150,000 usd.



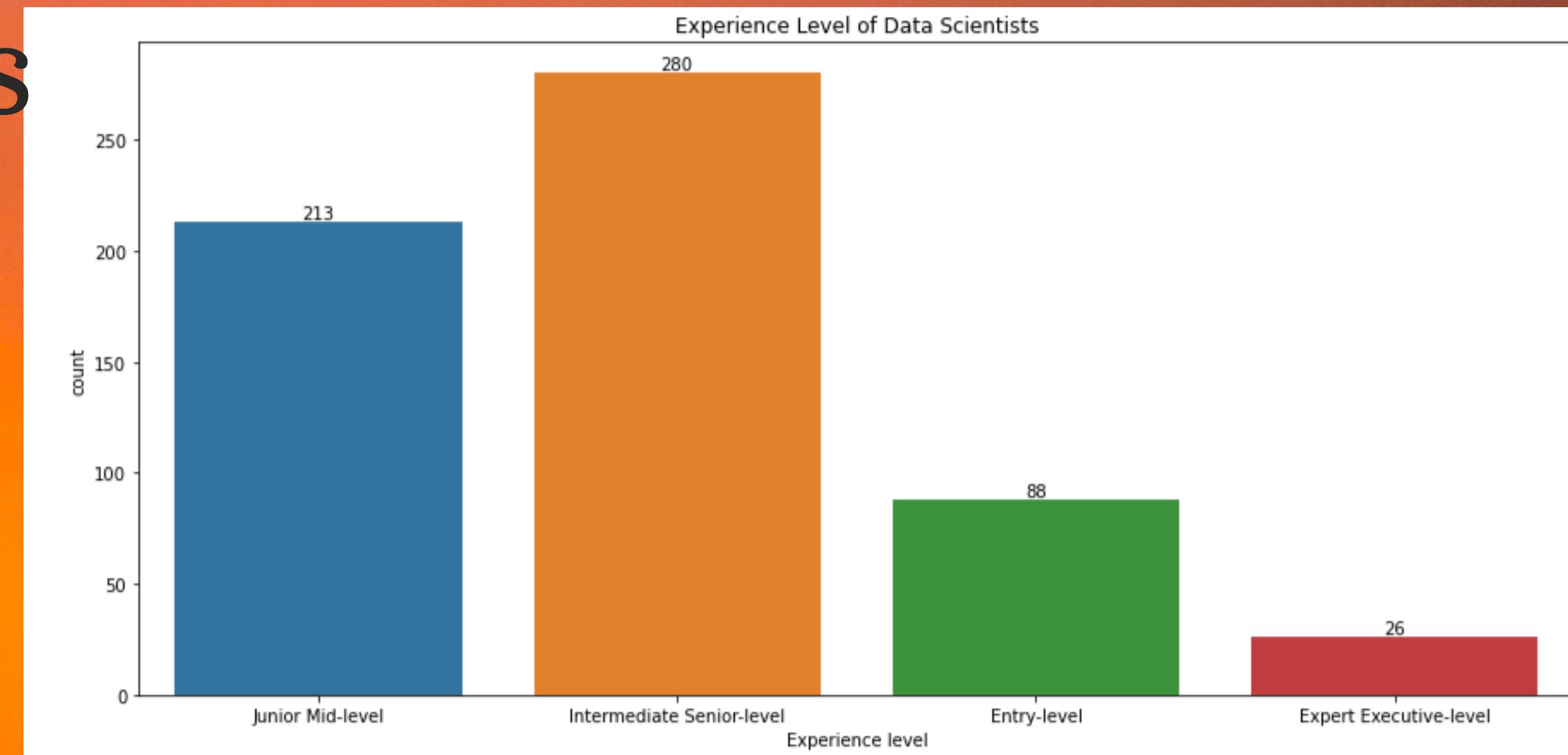


Count of work year

- In 2020, the number of salaries paid were few. This could be as a result of the Covid Pandemic.
- In 2021, the salaries increased, which means that the companies were gradually recovering from the pandemic and in 2022, more salaries were paid compared to 2020.

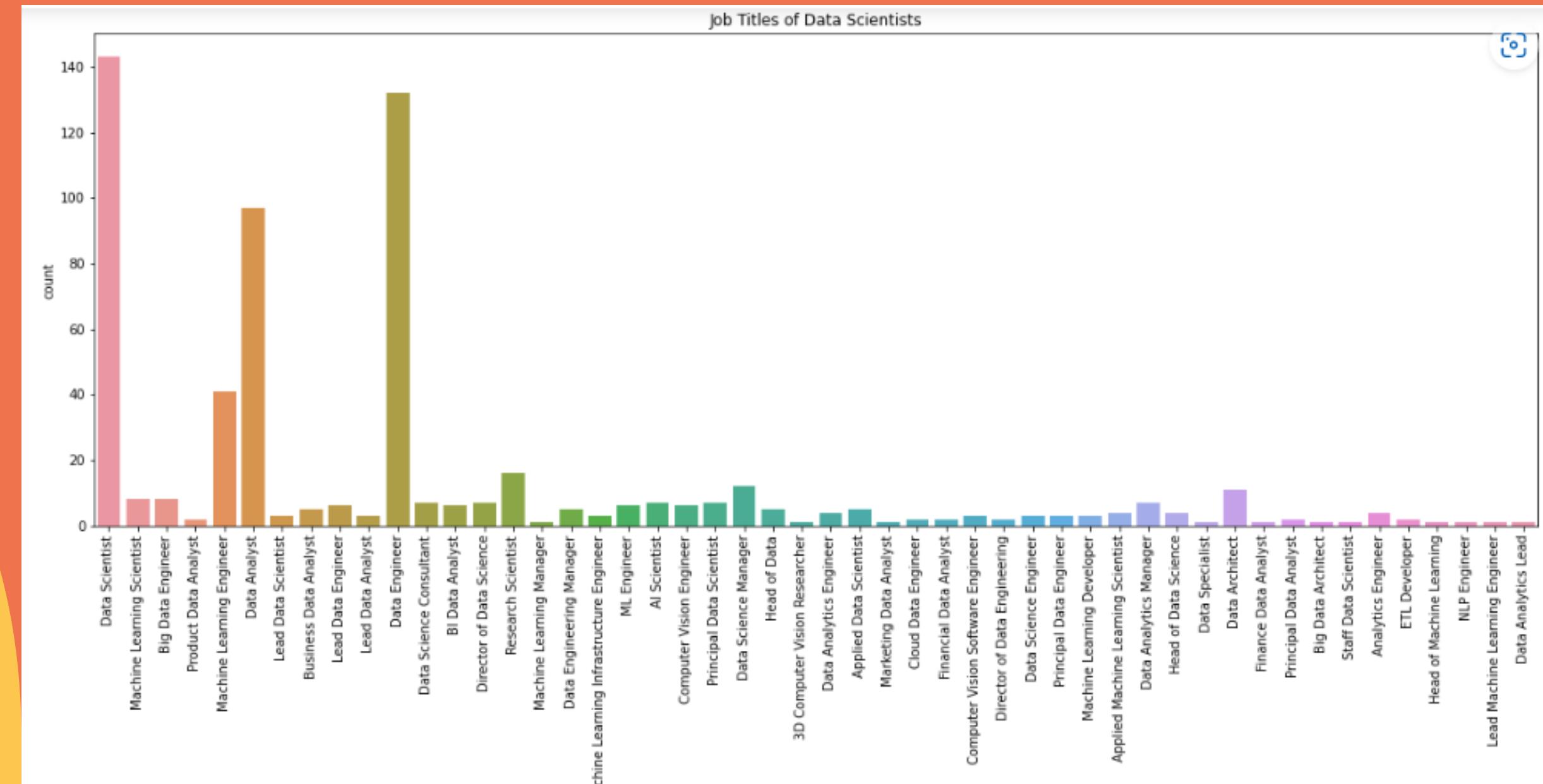
Experience Level of Data Scientists

- There are more data scientists in Intermediate Senior-level (SE).
- There are few data scientists in Expert Executive-level.



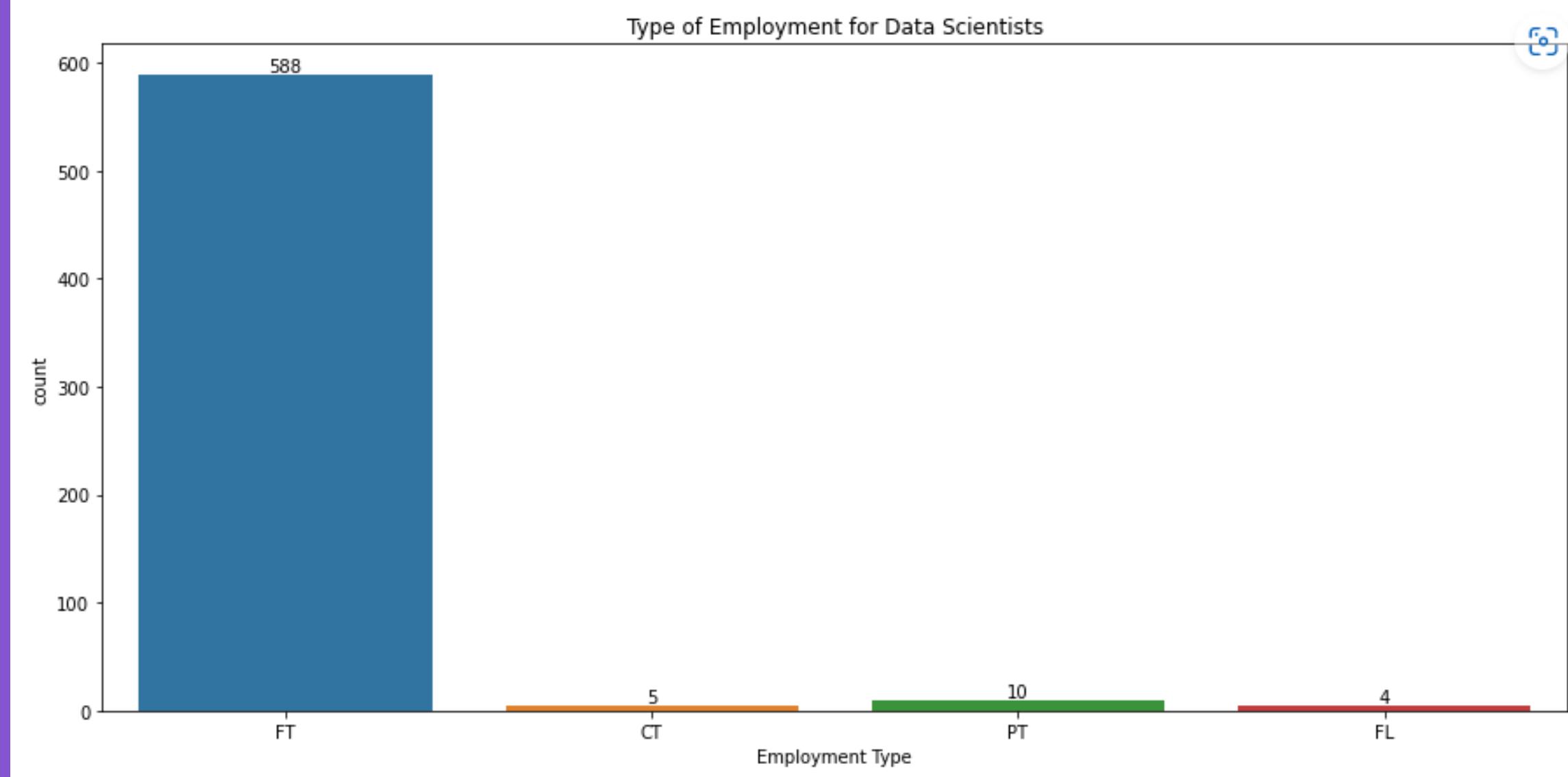
Job Titles of Data Scientists

- The top 4 job titles are Data Scientist, Data Engineer, Data Analyst and Machine Learning Engineer



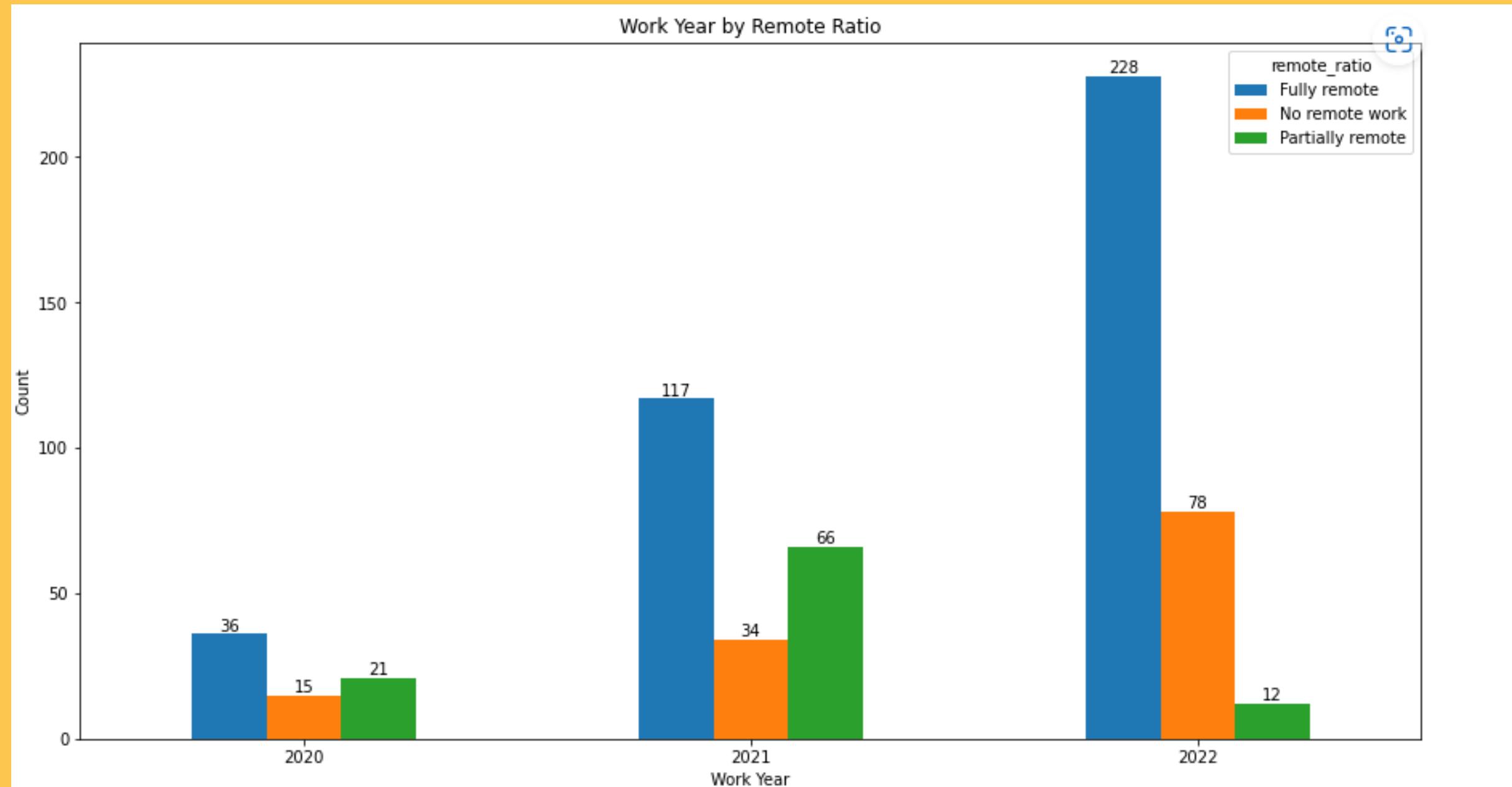
Type of Employment for Data Scientists

- There are more data professionals who work full time (almost every data professional works full time).

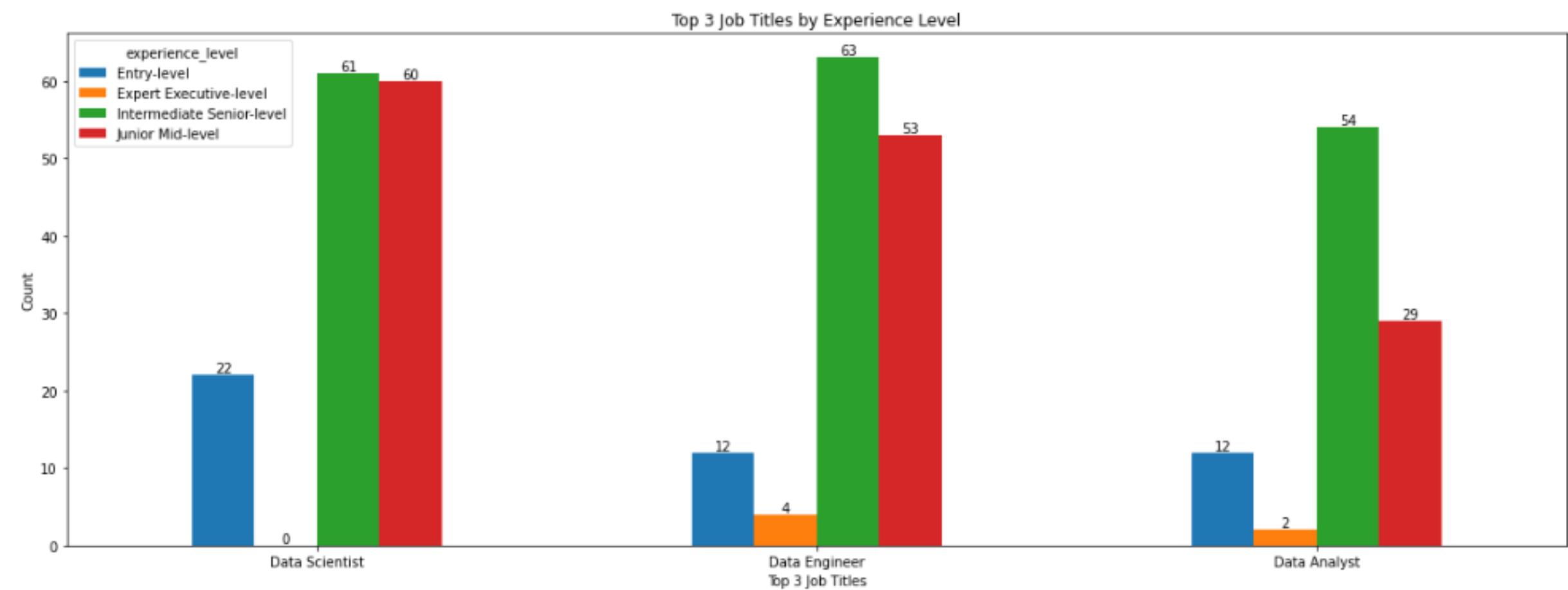


Work Year by Remote Ratio

- There is a huge increase in the count of fully remote jobs across the years. This indicates that data professionals prefer to work fully remote.
- There is also a slight increase in the number of 'No remote' jobs across the years.



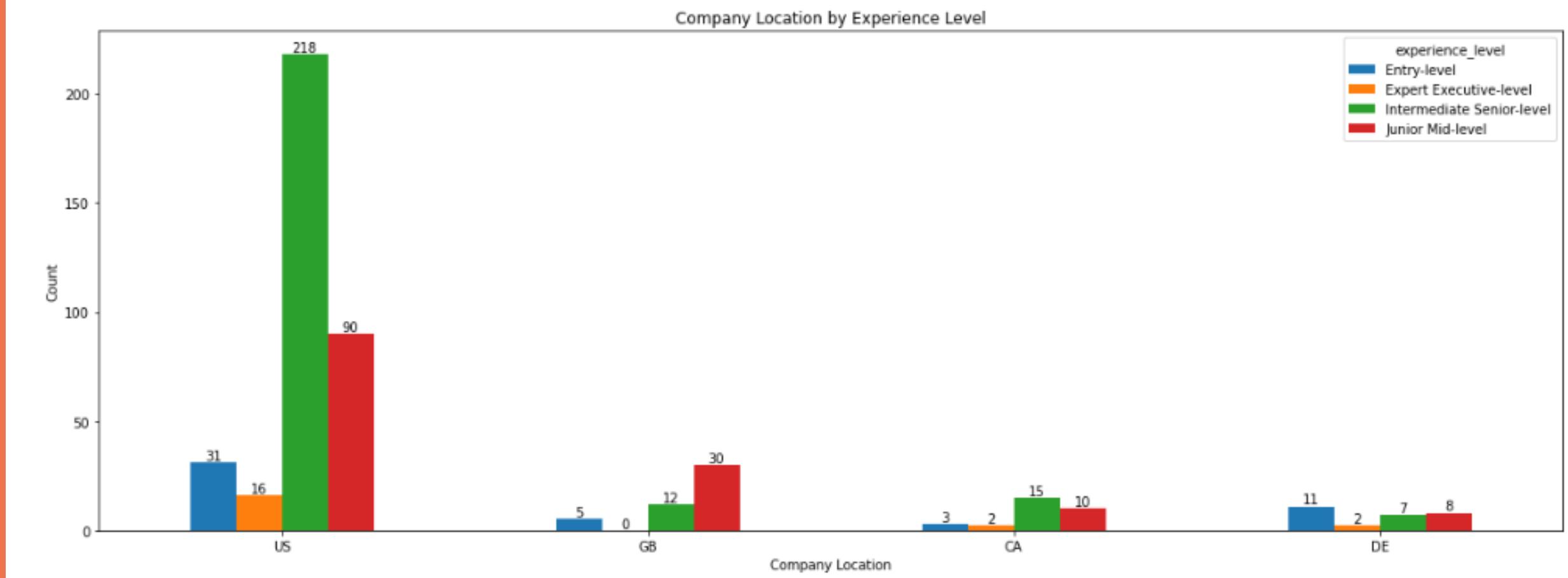
Top 3 Job Titles by Experience Level



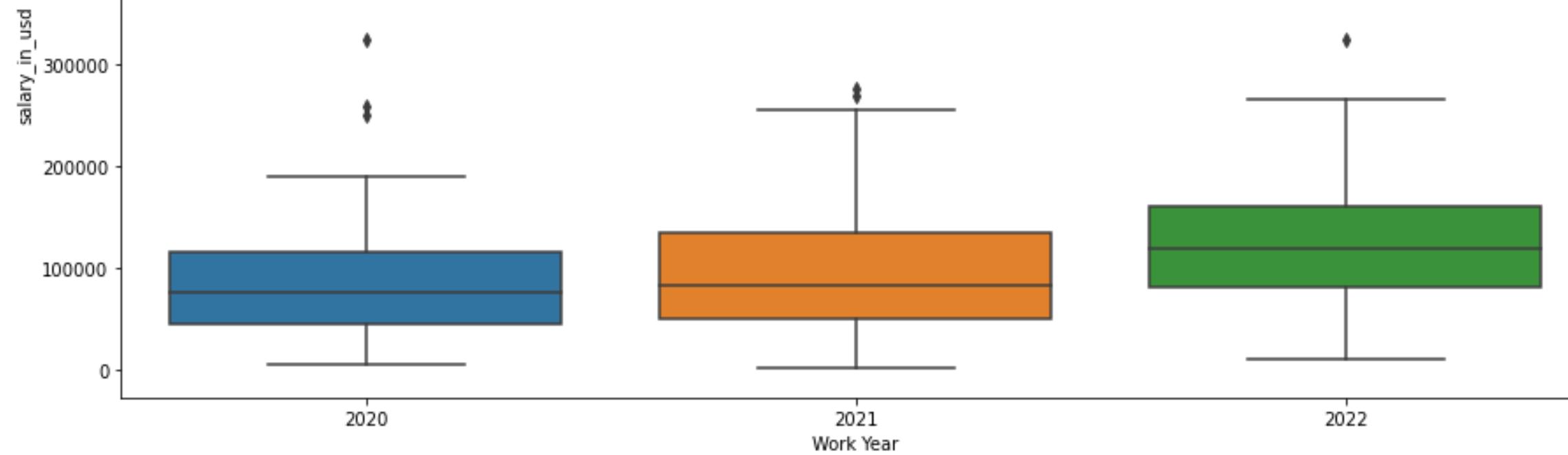
- The top 3 job titles are Data Scientist, Data Engineer and Data Analyst.
- Junior mid level(MI) and Intermediate Senior Level(SE) generated the highest count across the job titles.
- The number of experts Executive-level(EX) is extremely

Company Location Analysis

- Companies located in the United States(US) tend to be higher than those in other countries.
- Data Professionals with Intermediate SE senior-level have the highest count in US and Canada.



Salary by work year

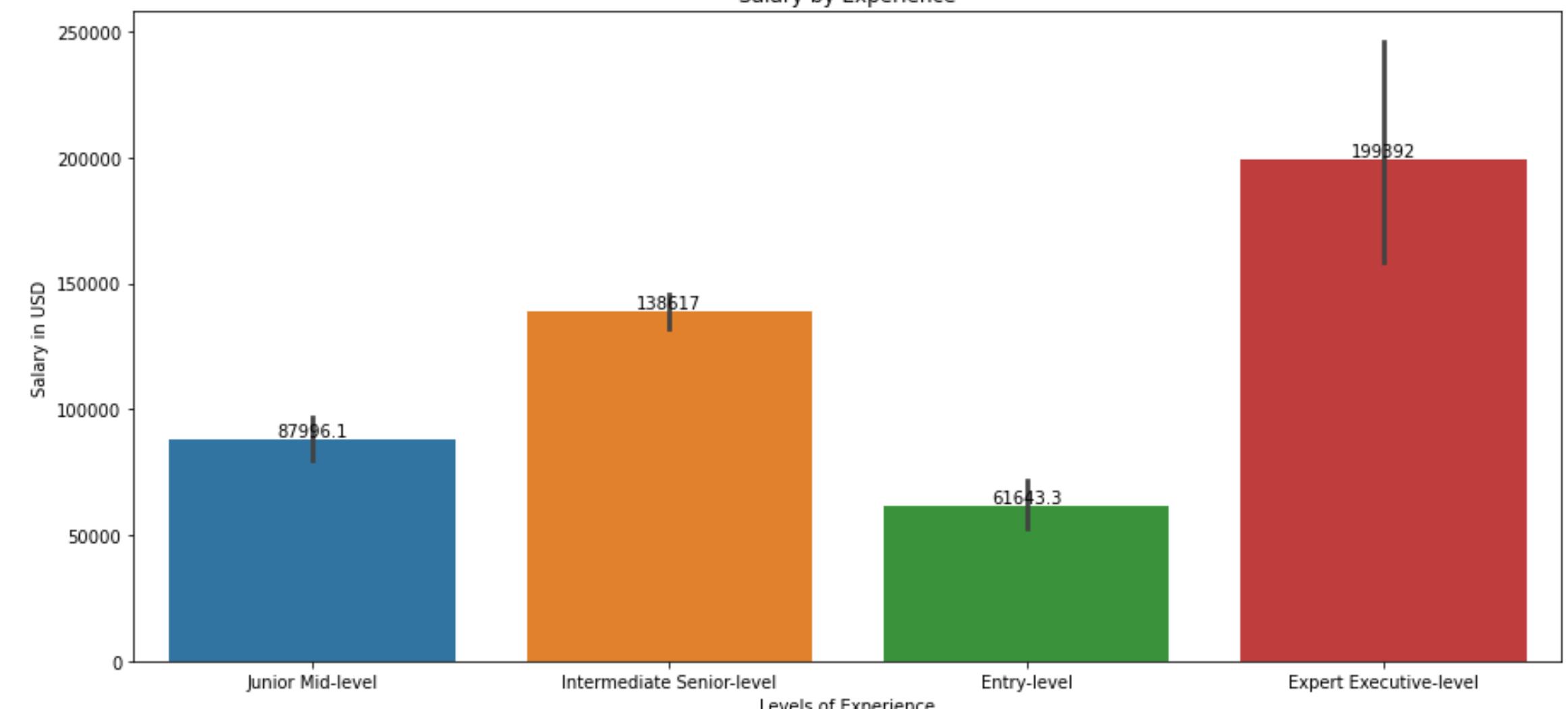
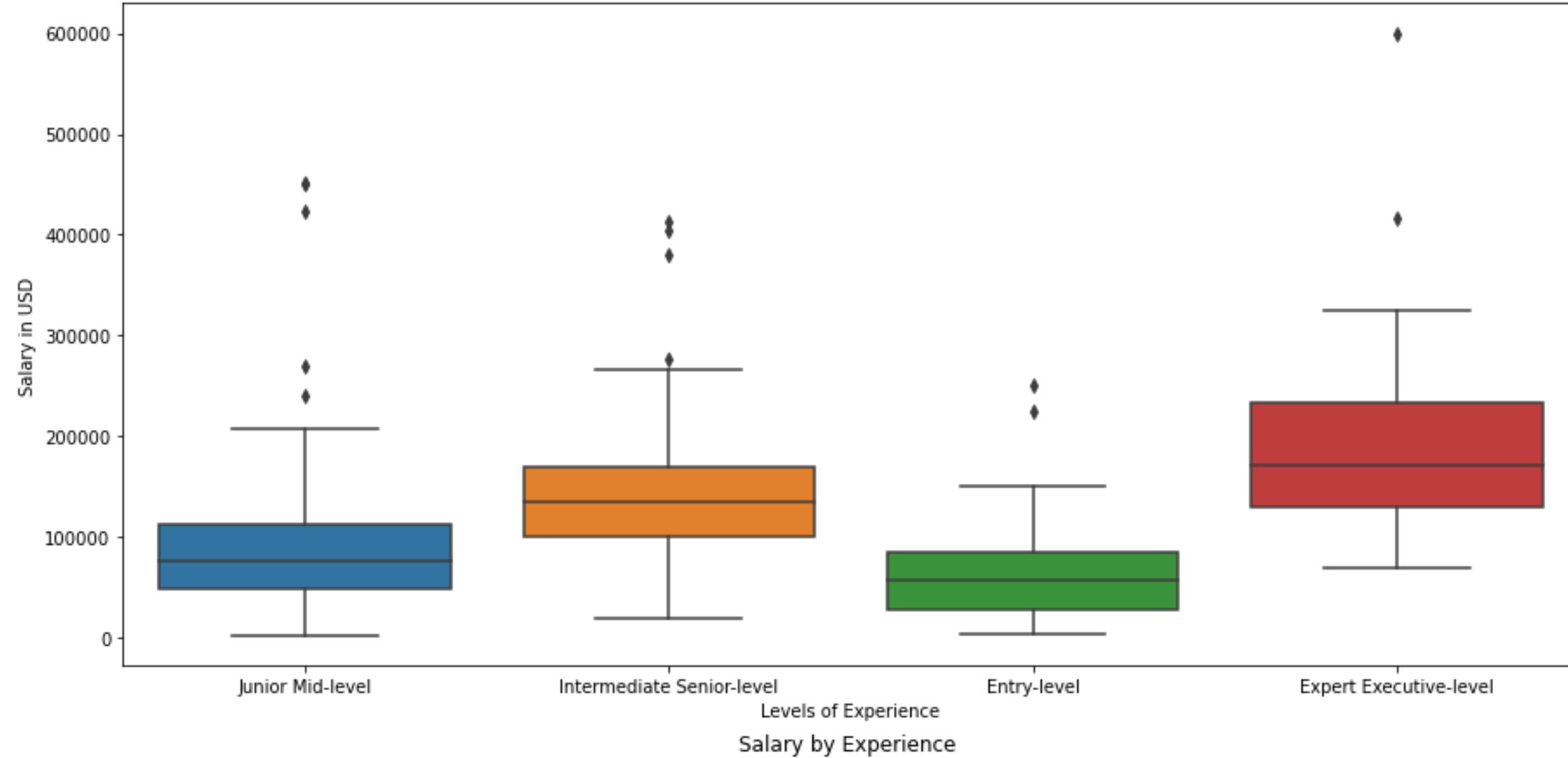


- There is a slight increase in the average salary from 2020(95,813) to 2021(99,854). However, the mean salary increased by approximately 25% from 2021 to 2022.
- In 2020 and 2021, half of the data professionals earn less than 100,000.
- In 2021, there are more scattered data with outliers of 600,000.
- In 2022, 75% of data professionals earn less than 180,000

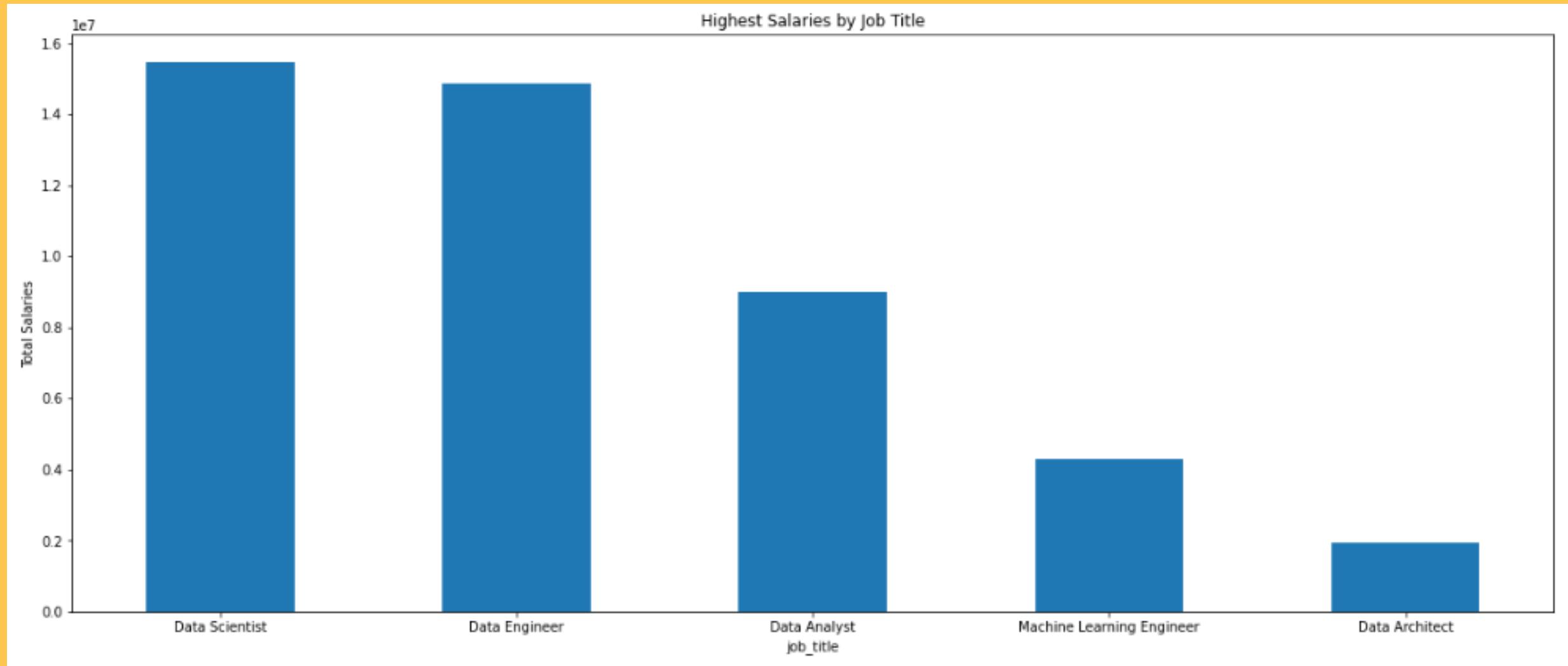


Salary by experience

- The average salary increases with the level of experience.
- Entry level(EN):The average salary is a little above 60,000. 75% of data scientists earn less than 100,000.
- Junior Mid-level(MI): The mean salary is around 88,000. While 50% of data scientists earn less than 100,000.
- Intermediate Senior Level(SE): The mean salary is close to 140,000. 75% of data scientists earn less than 200,000.
- Expert Executive Level(EX): The average salary is around 200,000. We can see that the outlier of 600,000 is earned by an expert in the Executive level and this is expected as a data professional increases in levels of experience, the salary also increases.



Highest salary by job titles



- Data Scientists have the highest salaries, followed by Data Engineers and then Data Analysts.

Recommendations



- I would recommend that information be given about the industries the data professionals work in. For example, Technology, Health, Government, Banking and Finance, Education etc
- Companies in other countries should deploy the use of data in taking decisions like the U.S because this will greatly help their country's economy to grow.

THANK YOU!