DATA ANALYSIS PORTFOLIO

Data Scientist

Data Analyst

Data Architect

Data Engineer

Statistician

Database Administrator

Business Analyst

Data & Analyst Manager

Professional Background

Dear Hiring Manager,

Having worked as a Data Analyst with two large organizations with expertise in data aggregation, data mining, and modeling. Providing sound analytical support to my superiors is extremely rewarding. That is why I believe I possess the versatile skillset and professional experience to be a top candidate for the job.

In my role as a customer success manager of fourteen years, I have added the following skills- customer support, team building and training, strategic leadership, call center management, continuous process improvement, and sales support functions.

In my current role, I have worked on various wide-ranging projects which allowed me to expand my data toolbox and hands-on expertise. That experience together with my customer analytics background of fourteen years has given me a deep understanding of customer analytics framework, business needs, data understanding, data preparation, and data modeling.

Some notable contributions include:

- Developing complex reports using Excel and Tableau to identify trends and potential opportunities and increase revenue for Udemy.
- Identifying and creating a new format for reporting the most paid Data science job role and giving insights into why the job roles are the highest paid.

I hope to leverage my skills and experience to help your organization to achieve further success.

Table of Content

Udemy Project Description

I was asked by my manager, Head of Curriculum at Udemy, to present the data on course revenue and was provided with data on courses from different topics to understand where opportunities lie and track the performance of courses.

This report tracks Udemy course profitability while identifying those that have a foundation for successful income development. This report aims to investigate the claim that raising the cost of "Web Development" will lead to a boost in the company's gross profits. In this report, the claim that the most popular courses are in "Web Development" is also tested.

The Data set needed to be cleaned as it had blanks and duplicates, and the headings were not consistent. The Web Development subject was not consistent with the other subjects. After the cleaning, I used the IF function to determine what course was Free or Paid. I also used the Left and Right keys to pick out the date the courses were published.

Thereafter I created Pivot tables with charts and also selected the top 20 courses using the Vlookup function.

Tableau was used in visualizing the data. Analysis of the data sets based on my findings was done. The conclusion arrived at was that a little price increase for the course Web Development would be appropriate considering the growing member base. There will undoubtedly be a rise in the number of people enrolling in these courses as a result of the expected need for Web Development.

I would also like to add that, given our desire to boost revenue, it would be a fantastic idea to slightly increase the price of Business Finance.

The Problem

The business problem was to understand where opportunities to increase revenue lie and track the performance of courses.
The project took two weeks to analyze. This was to give ample time for my manager to scan through and analyze before the meeting with the director.
The data collected were course identification numbers, the total number of subscribers per subject, price, reviews, ratings, content duration, date/ time of publishing, and subject.
The raw data was presented in excel and then pivot tables were created with charts and tableau visualization.

The questions asked were:

1. Why do we have more subscribers in Web Development than in other subjects?
2. Why are some courses free and some paid?
3. Why is the rating for some subjects higher than others?

Data Design

Data Cleaning:

In preparing the data set for analysis, the duplicates in the data sets were removed. The blanks were also removed. Headers were used to organize the data and understand what each column represents.

The find and replace function was used to make the "Web Development consistent with the other subjects.

Data Visualization:

Pivot tables gave us access to rows, columns, values, and filters. While Tableau has been used to show the key differentiating elements for each course in the data sets.

Findings

Finding 1
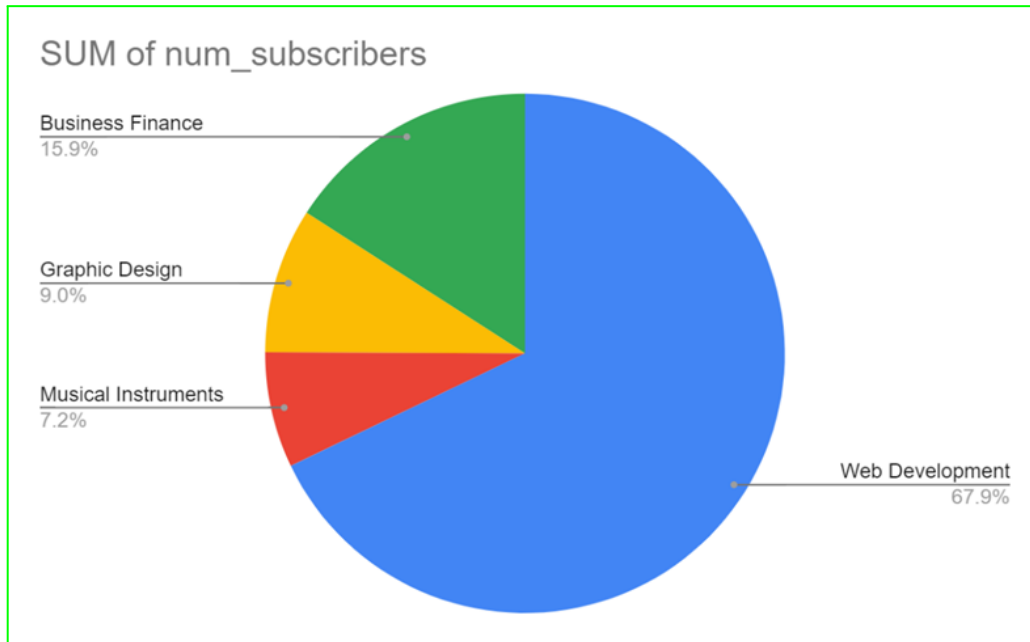
Figure 1: Sum Number of Subscribers
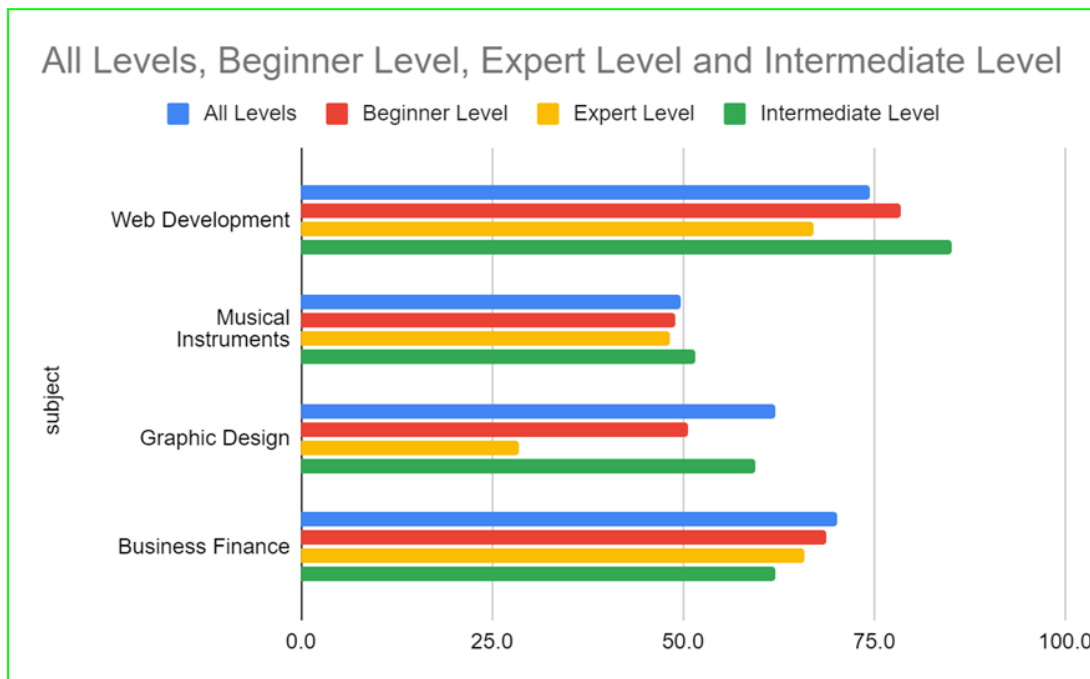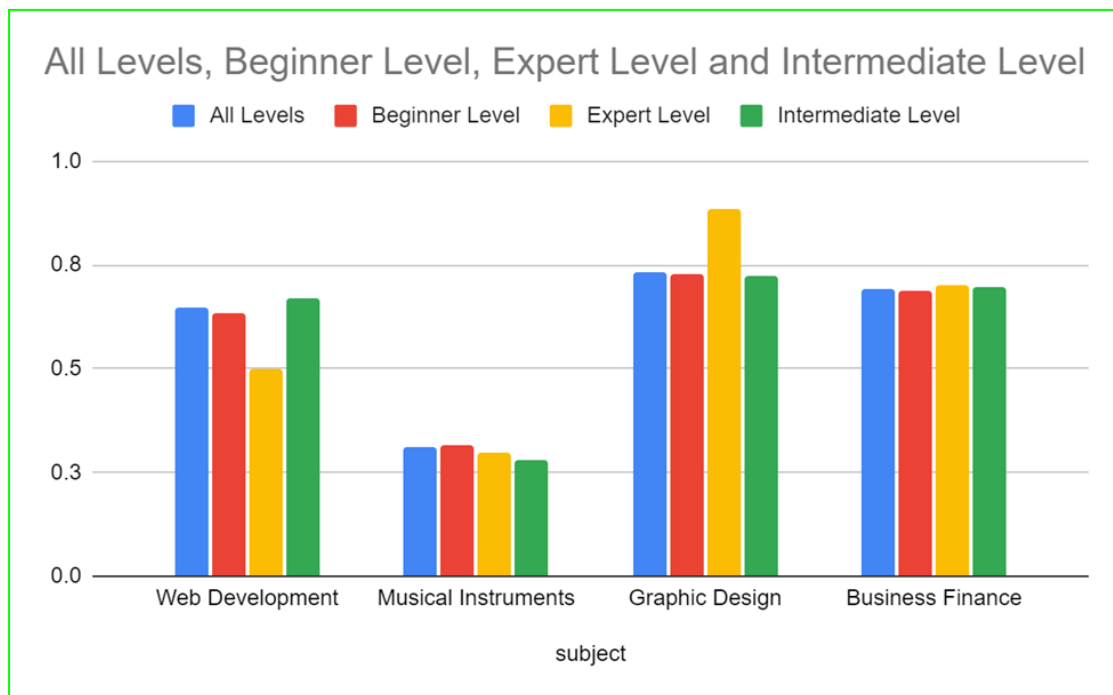

Figure 2: Average Price Per Subject

Figure 3: Average Rating Per Subject

All Levels, Beginner Level, Expert Level and Intermediate Level

■ All Levels   ■ Beginner Level   ■ Expert Level   ■ Intermediate Level

subject

Finding 2
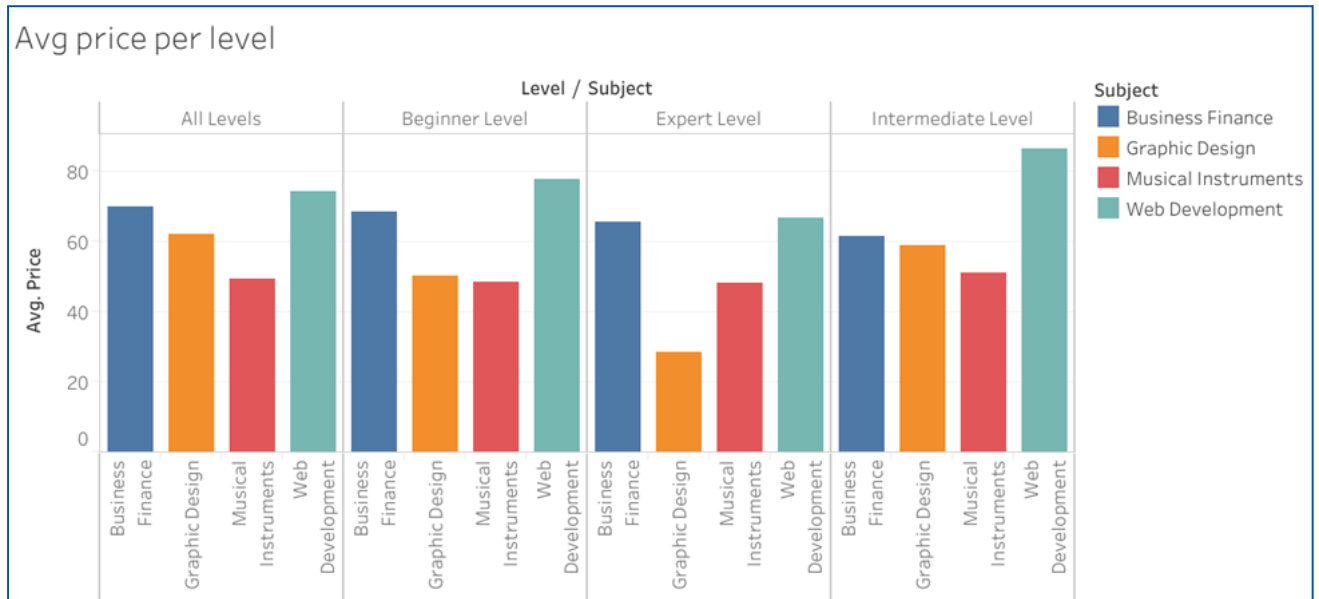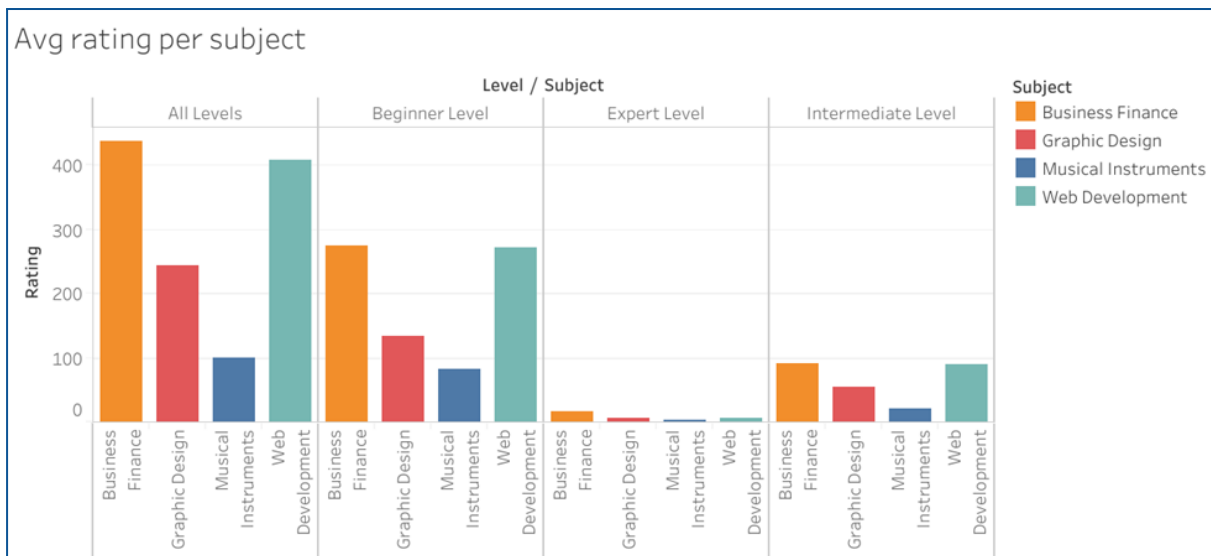
Figure 1: Avearge Price Per level

Figure 2: Average Rating Per subject On All Levels

Analysis

Based on the findings above, it is shown that Web Development is the most popular subject looking at the total number of subscribers. This was followed by Business Finance, Graphic design, and Musical Instruments. Based on the chart above, the course with the highest price and content duration is also Web Development.

A further look into the root cause analysis of why Web Development is the most popular is because people tend to see it as a gateway to the technology industry as this seems to be a high-paying skillset in the technology industry. Being the course with the highest content duration and price point, this indicates that it has a higher quality compared to the other courses and so people tend to go for it more.

Another observation was that Business Finance seems to have the highest rating on all levels. That indicates that though Web Development is the most popular, it still does not have a very good rating. This would imply that the users are getting more value for it that the other courses.

Conclusion

The submissions made above indicate that a little price increase for the course, Web Development would be appropriate considering the growing member base. There will undoubtedly be a rise in the number of people enrolling in these courses as a result of the expected need for Web Development.

I would also like to add that, given our desire to boost revenue, it would be a fantastic idea to slightly increase the price of Business Finance.

LinkedIn Project Description

This report analyses Data Scientist salaries while highlighting the career path that a recent American graduate might want to follow to increase their earnings in the tech sector. The purpose of this report is to examine the assertion that "Data Analyst" is the best paying job title. The highest-paid Data Science job title between 2020 and 2022 is also typically seen in this research.

Using the Find and Replace function on google sheets to create consistent data.

Pivot tables were used to create charts and graphs to visualize the data by rows and columns. While Tableau has been used to show the key differentiating elements for each job title and the salaries in the data sets.

Analysis of the data sets based on my findings was done. The conclusion arrived at was that getting a remote full-time job would also increase one's ability to earn a higher income. For aspiring data scientists looking to gain work experience or to switch careers to data science, the opportunities for remote work have never been greater.

Data Design

Data Cleaning

Using the Find and Replace function on google sheet to create consistent data with the below terms;

Experience level, for instance, EN for Entry Level; employment type- FT for Full Time, employee residence and company location for instance US for the United States, were all written in abbreviations.

Data Visualisation

Pivot tables gave us access to rows, columns, values, and filters. While Tableau has been used to show the key differentiating elements for each job title and the salaries in the data sets.

Findings

Findings 1

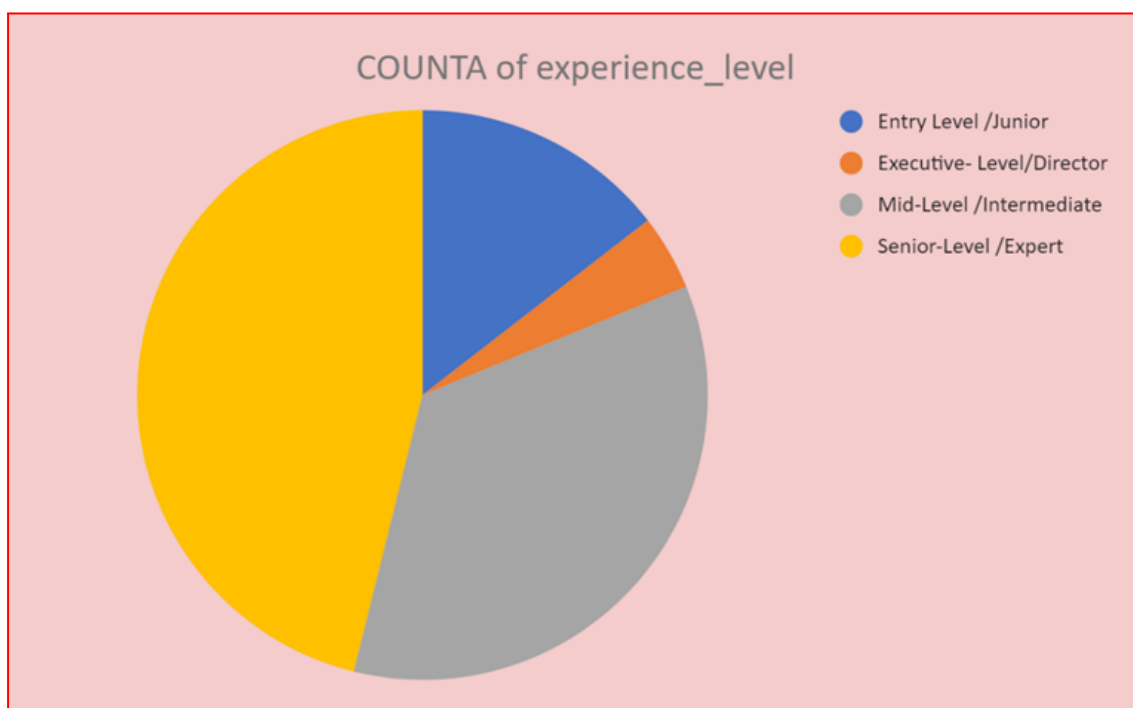*Figure1: Count of Experience Level*

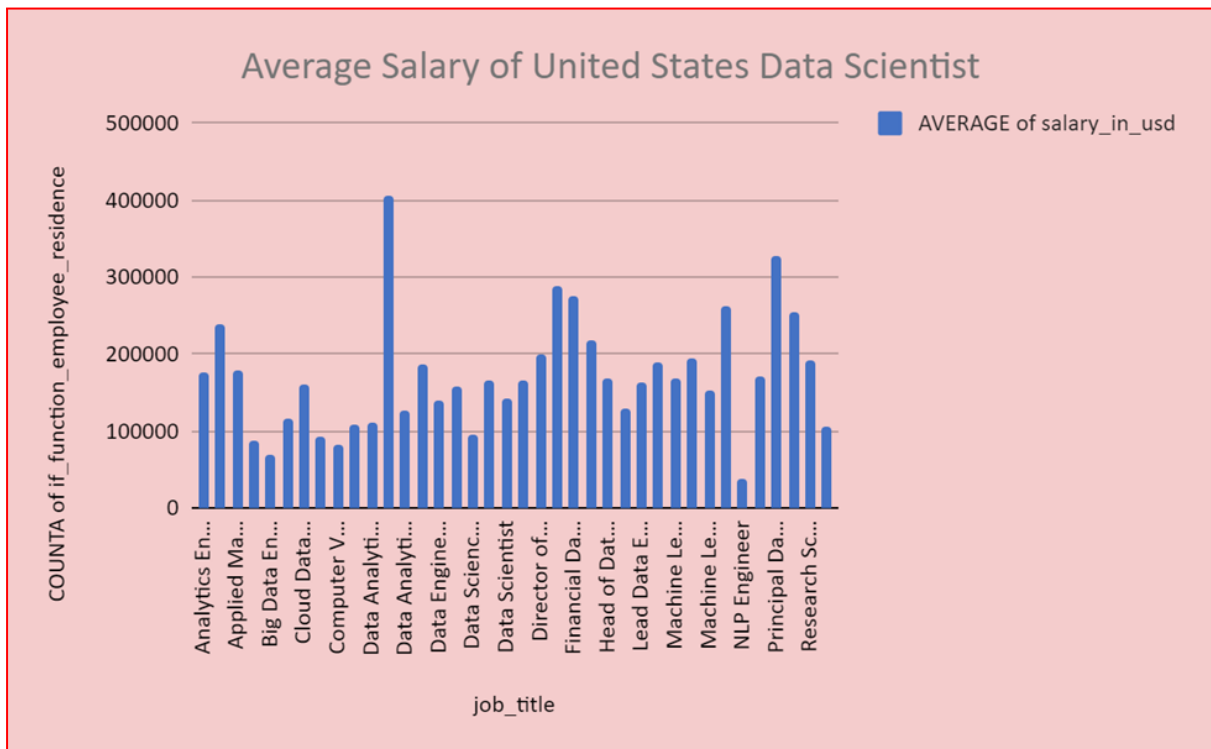*Figure 2: Average Salary of United States Data Scientist*
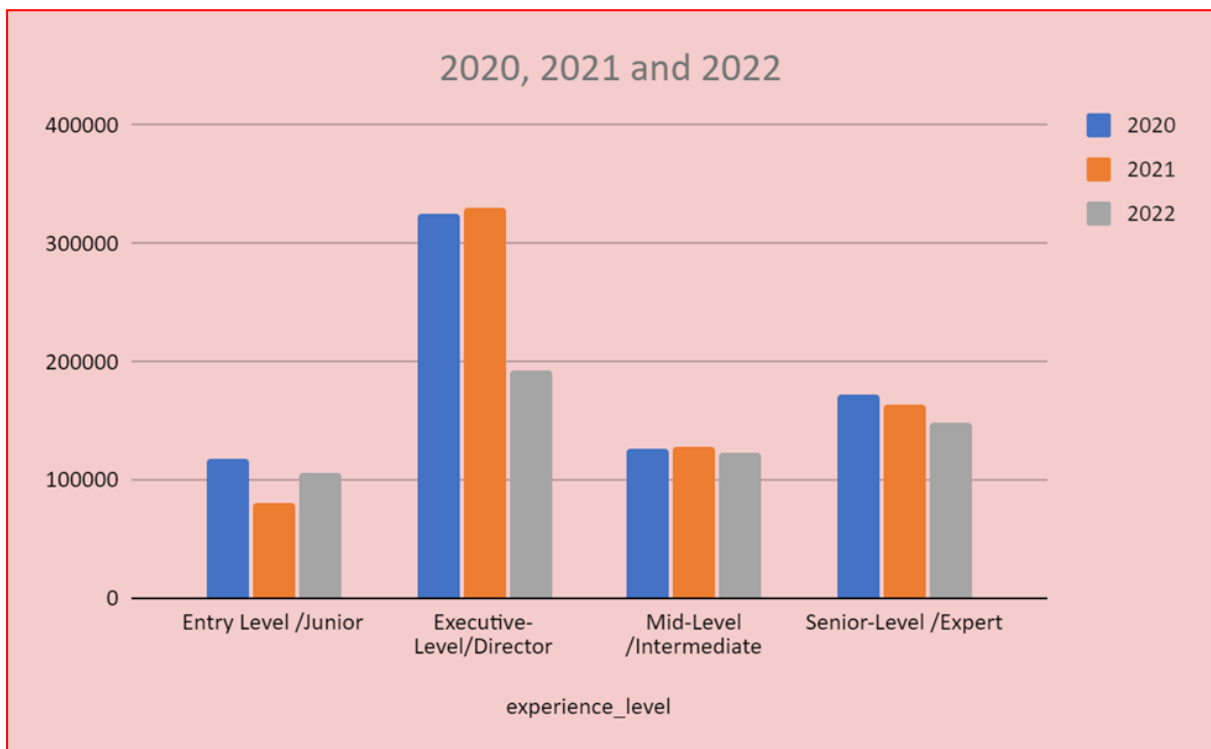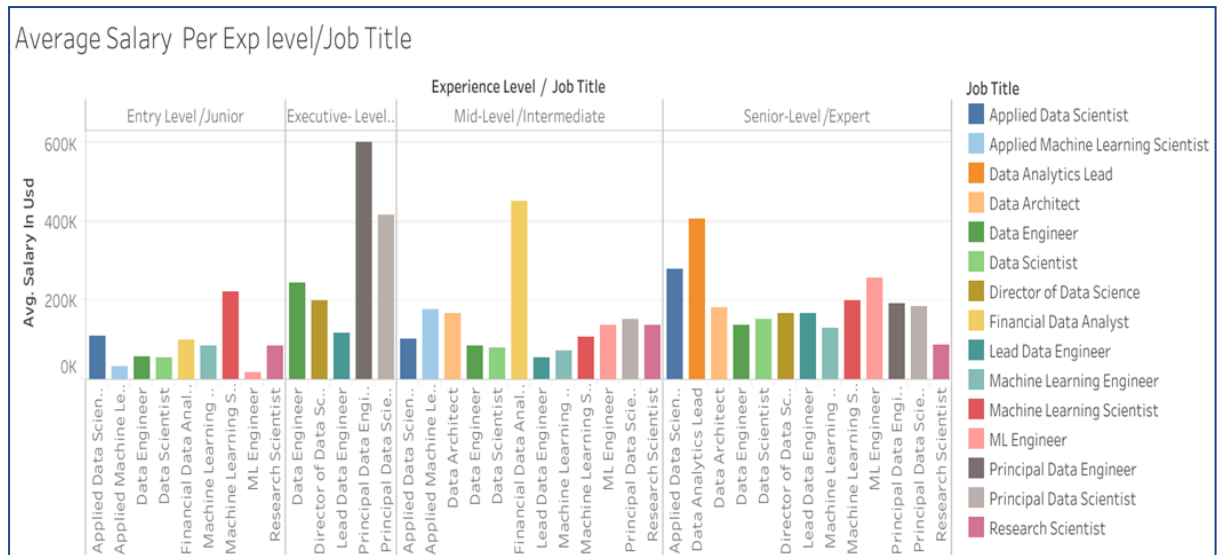

Figure 3: Average Salary per Experience level per year

Findings 2

Figure 1: Average Salary Per Experience Level/Job Title

Analysis

The Senior-Level/Expert appears to have the highest count based on the data above. When compared to other nations, the United States has the highest average salary in USD. Additionally, compared to 2021 and 2020, salaries paid in 2022 are higher.

More people worked remotely in 2022 compared to the other years after the pandemic. With technological advancements and, of course, the pandemic, remote working has become the new norm among millennials and experienced professionals looking for more flexibility in their career paths.

The Executive-level/Director has the highest paid salary. The Data Analyst Lead received the highest pay, according to Tableau.

According to data from the United States, the Director of Data Science is the position with the highest pay, followed by the Principal Data Engineer, the Data Analyst Lead, and the Applied Machine Learning Scientist.

The employees that worked Full time in all sizes of the company be it small, medium, or large were paid higher than the Contract, Freelance, and Part-time. Employers have started to understand that, by giving their employees more flexibility with their working hours and not binding them within the four walls of the office, employees can be more productive. In effect, employees have more time to work and can adjust their personal, physical and mental well-being. Many organizations found they were able to reduce costs by a large amount per annum just by promoting work from home and cutting out the need for office space and other fixed costs.

Conclusion

The career path that an American graduate can take is:

Entry level- Machine Learning Scientist.

Mid-Level/Intermediate- Financial Data Analyst.

Senior-Level/Expert- Data Analyst lead.

Executive- Level/Director – Principal Data Engineer.

The various levels of average incomes paid for each job title led to this conclusion.

Getting a remote full-time job would also increase one's ability to earn a higher income. For aspiring data scientists looking to gain work experience or to switch careers to data science, the opportunities for remote work have never been greater.

Appendix

Google Sheet data Set for Udemy Project
Click Here

Tableau Data Visualization For Udemy Project
Click Here

Google Sheet data Set for LinkedIn Project
Click Here

Tableau Data Visualization For LinkedIn Project
Click Here