

Data-Related Career Path

(Intro) - BRIAN

To help narrow down the choices of a data-related career path, our group examined a Kaggle dataset that lists the different jobs and salaries in data science. The dataset includes information on job titles, job categories, salaries, employment type, and work setting. By analyzing the data, we can see the different career opportunities in data science, both by field and specific job titles, as well as the financial expectations of different roles.

To make the data easier to analyze, we cleaned the dataset by removing excessive rows and columns that did not fit our focus. To start, we only wanted to examine data from 2023, as it is the most recent year and had by far the most information. Since we only wanted to look at data from the United States, we also removed all European countries from the Employee Residence column, as well as dropped the Salary In USD column. This shrunk the dataset from 9,355 jobs to 6,644.

(General Statistics) - BRIAN

After cleaning the dataset, we can see that the jobs fall under ten categories: BI and Visualization (285 jobs), Cloud and Database (3 jobs), Data Analysis (975 jobs), Data Architecture and Modeling (189 jobs), Data Engineering (1,527 jobs), Data Management and Strategy (44 jobs), Data Quality and Operations (38 jobs), Data Science and Research (2,179 jobs), Leadership and Management (369 jobs), and Machine Learning and AI (1,035 jobs). The four experience levels are Entry-level, Mid-level, Senior and Executive, and the three settings are In-person, Remote and Hybrid.

The jobs range from a minimum annual salary of \$24,000 to a maximum of \$405,000. The means of the salaries range from \$99,050 to \$195,515. By examining the min, max and mean of each job category, we can get a sense of how much one could expect to make in an entry-level position, and what they could expect to make in the future as a senior or executive. A person using this data could be underwhelmed at the idea of making \$25,500 in Machine Learning and AI, but inspired to find out the average salary in this category is \$195,515, with a max salary of \$392,000. They could also see that while a career in Cloud and Database is profitable, the fact that only three jobs are displayed in the data could indicate that there will not be many openings.

(Salary and Experience Level) - MEDHA

Looking into the data further and comparing salaries by experience level, we can see how the overall data is spread amongst the various levels. If someone is trying to find a job in Data Science/Analysis with Entry-Level experience, they can expect an average salary of \$105,290, where the lowest salary they can make is \$24,000 and the highest is \$281,700. If this person gains more years of experience, as a Mid-Level, the average salary increases to \$131,202, where the range of the salary can fall as low as \$25,500 and as high as \$405,000. The Senior level salary averages to \$168,754, and ranges from \$38,000 to \$392,000. As an Executive, on average their salary is \$192,222 and ranges from as low as \$55,000 to as high as \$375,500.

Based on the averages alone, you can see that as you gain years of experience, the salary pay also increases significantly. However, the odd factor that can be noticed is when looking into the maximum salary, someone with Mid-Level experience seems to be able to make more than if they were a Senior or an Executive. This can be due to the specific job or title this person may have. If we look further into the data and see the top paying job titles based on years of experience, we can see that having the title of Data Architect (AI/AWS) or a title in relation to Machine Learning (ML Engineer) will provide a high salary amongst all levels of experience.

(Salaries and Job Titles/Categories) - SARAH

The dataset includes information about job titles, job categories, and corresponding salaries, as well as other factors analyzed in different sections. The goal of this section is to understand the average salaries associated with different job titles and categories in the tech field. Before calculating the average salary of someone working in a specific job title, we set a parameter to display the top 5 most popular jobs. Choosing this parameter is helpful because we can gain a larger understanding of the trends that occur in the tech field. After filtering for the average salary of the most popular job titles, we see that the salaries the results indicated that:

- Applied Scientist \$191591.58
- Data Analyst \$114290.58
- Data Engineer \$155378.35
- Data Scientist \$170201.76
- Machine Learning Engineer \$195183.65

In addition we broke down the average salary within each category:

- BI and Visualization \$135000.0
- Cloud and Database \$140000.0
- Data Analysis \$109900.0
- Data Architecture and Modeling \$150000.0
- Data Engineering \$146115.0
- Data Management and Strategy \$88750.0
- Data Quality and Operations \$95000.0

- Data Science and Research \$167441.0
- Leadership and Management \$140000.0
- Machine Learning and AI \$189360.0

After creating a library to group our data by job category and title, it printed the top paying job within each category and the corresponding amount. Ultimately the data concluded that the highest paying category and job title was Data Science and Research with an average of \$267475. What this tells us is that the potential for growth in this field is astounding. Not only do we know that there are a significant amount of entry-level to mid-level jobs for data science roles, but that one has the potential to double their salary to over 250k. One thing we are unaware of is the potential skew of the respondents in this dataset. The average of the most popular jobs alone feel fairly high, from this data alone we are unable to conclude how representative these jobs are of the whole data science field.

(Salaries and Company Size) - BRIANNA

The analysis reveals a clear distribution of firms within the data science industry, with the majority categorized as medium-sized (90.3%), followed by large firms (8%), and a smaller proportion classified as small firms (1.7%). Average salaries vary significantly across firm sizes, with medium-sized firms offering the highest average salary (\$151,623.61), followed by large firms (\$142,949.26), and small firms offering the lowest average salary (\$92,666.55). This distribution highlights a positive correlation between firm size and salaries in data science roles, suggesting that larger organizations tend to provide higher compensation packages compared to smaller counterparts. These findings emphasize the importance of considering firm size when assessing salary expectations and negotiating compensation within the data science industry, providing valuable insights for both job seekers and employers.