# *Analyzing Data Science Job Salaries Through Interactive Visualizations*


***- Diya Patel***

***BS22DMU019***

# *CONTENTS*

# *ABSTRACT*

The domain of data science is a rapidly growing field that is driving significant changes in how businesses operate and make decisions. In this report, I have explored the data science job market by analyzing salary trends across various factors such as job title, company location, experience level, and employment type. I have chosen a rich dataset and an interactive dashboard to meticulously analyze and visualize the compensation structures within the data science profession.

This report's analysis also highlights the significant impact of experience, regional location, and organization size on wage results in the data science profession. The research presents a granular perspective of the economic value associated with data science competence. It delineates the compensation differences within the industry by deconstructing these variables through a series of interactive visualizations. This sophisticated understanding is critical for aspiring and current data scientists navigating their careers
.

# *DATASET DESCRIPTION*

| COLUMN | DESCRIPTION |
|---|---|
| **work year** | The year the paycheck 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 |
| **job title** | The role worked during the year. |
| **salary** | The total gross salary amount paid. |
| **salary currency** | The currency of the salary paid is an ISO 4217 currency code. |
| **salary_in_usd** | The salary in USD |
| **employee residence** | Employee's primary country of residence during the work year |
| **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%) |
| **company location** | The country of the employer's main office or contracting branch |
| **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) |

# *DATA READINESS*

- Data Validation:
  Before declaring the dataset appropriate for analysis, I extensively validated the data's integrity by cross-referencing the salary entries with industry norms and external compensation benchmarks. This verified that all records were within a reasonable range, allowing the analysis to be free of any potential outliers that could bias the results.

- Data Reliability:
  I standardized the naming standards across various categorical data points to preserve consistency. This includes standardizing job names, experience levels, and firm locations, all required for correct grouping and comparison within the visualizations.

- Data Presentation:
  I transformed all data into formats consistent with the software's needs to ensure a seamless connection with visualization tools. Dates were formatted according to ISO standards, and numerical values were checked for accuracy.

- Data Structuring:
  I organized the dataset into a structured format that could be easily transferred to the visualization dashboard, which included aligning the pay data with appropriate attributes such as job title and experience level to allow easy querying and filtering.

- Enriching Data:
  I supplemented the dataset with additional relevant data, such as cost of living indexes and regional economic indicators, to provide a more comprehensive view, allowing for a deeper analysis of the pay statistics about each area's living standards and economic conditions.

# *RESEARCH QUESTIONS*

I've discovered significant business questions as a data-driven professional navigating the shifting landscape of the data science industry. These are designed to capture the essence of income trends, job distributions, and market dynamics, providing a clear picture of present and emerging trends in my sector.

**RESEARCH QUESTIONS:-**

- How does the average data science wage differ by country, and what variables may impact these differences?

- What is the relationship between experience level (entry, intermediate, and senior) and salary in the data science field?

- How has the average income for data science positions changed, and are there any notable trends or patterns from prior years?

- What are the salary disparities between various data science job titles, and how does experience level affect these salaries?

- How does the type of job (full-time, part-time, contract, freelance) affect the average wage for data scientists?

- How does firm size (small, medium, large) affect data science salaries, and is there a consistent pattern across different locations?

- How are data science employment opportunities distributed across different organization sizes and locations?

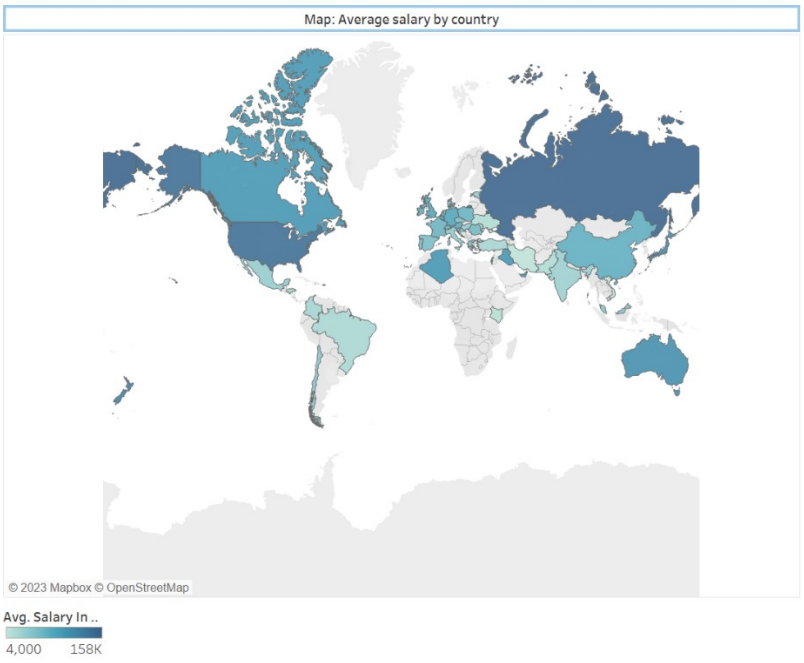
- How does the distribution of employee residences match the availability of data science employment and the accompanying salary ranges in those regions?

- How do salary progression trends in data science correlate with career advancement from entry-level to senior positions over the past few years?

- What conclusions can be derived from the employment composition by company size, particularly in data science roles and pay?

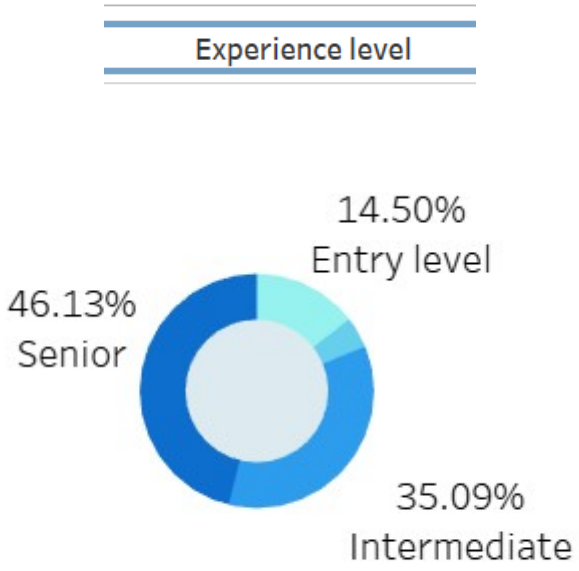
# *VISUALIZATIONS WITH INSIGHTS*

❖***How does the average data science wage differ by country, and what variables may impact these differences?***



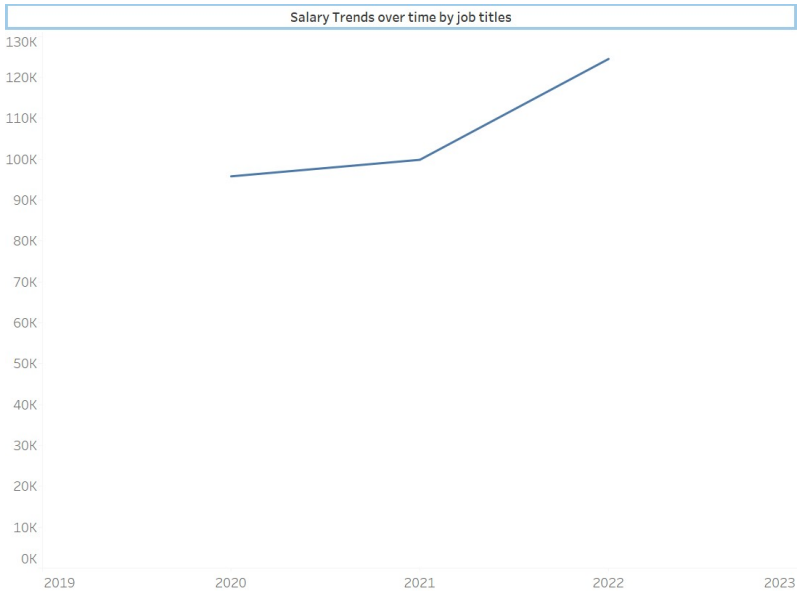

The main takeaway from this map, which illustrates average earnings by country, is a vast disparity in average salaries worldwide. The countries are tinted based on their average wage levels, with darker shades indicating more excellent average salaries. In general, countries in North America, Australia, and parts of Europe appear to have the darkest hues, meaning more excellent salaries. In contrast, lighter tones indicate lower average earnings in Africa, South Asia, and South America. This graphic depicts disparities in wealth and income levels across regions, showing worldwide economic inequality.

❖***What is the relationship between experience level (entry, intermediate, and senior) and salary in the data science field?***




The donut graphic depicts the distribution of data science workers at various levels of experience. Senior-level professionals make up most of the workforce, implying that more excellent salaries are possible at this level due to acquired skills and industry knowledge. Intermediate-level professionals account for a sizable fraction, indicating sustained career advancement chances within the area. Entry-level roles have the smallest share, possibly due to competitive entry hurdles or fewer openings in the business for newcomers.

## ❖*How has the average income for data science positions changed, and are there any notable trends or patterns from prior years?*



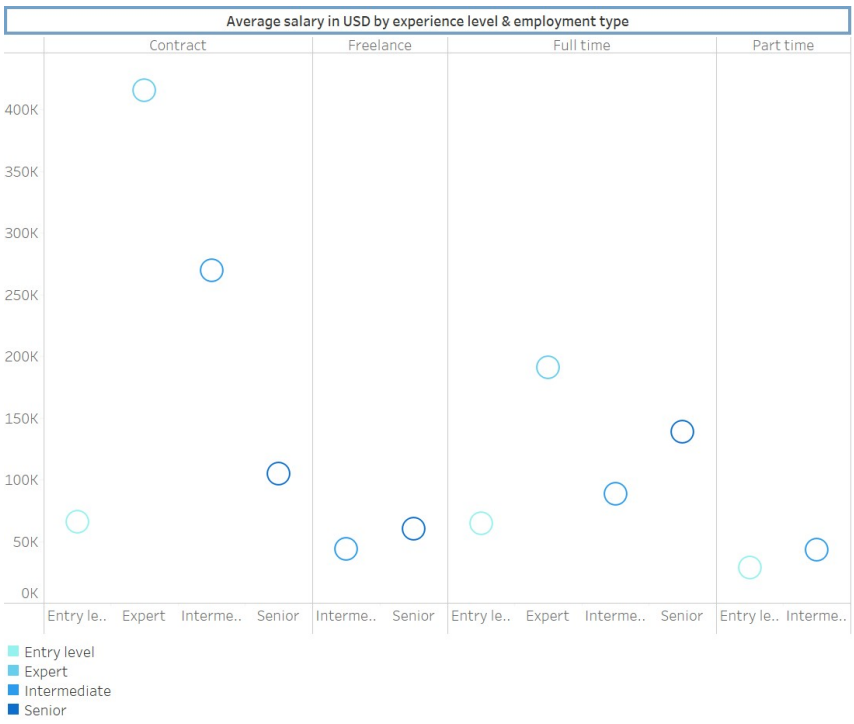

From 2019 to 2023, the average salary for data science job titles shows a favorable trend. Notably, there is a significant upward change between 2021 and 2022, indicating that there may be an increase in demand for data science abilities or a market adjustment to the value of these professions. This development may show a growing appreciation for the value of data science in driving business choices and technological advancements. The constant upward trend supports the idea that data science talent is becoming more critical in the employment market.

## ❖*What are the salary disparities between various data science job titles, and how does experience level affect these salaries?*

| Average salary by job title and experience level | | |
|---|---|---|
| 3D Computer Vision Resea.. | Intermediate | 5,409 |
| AI Scientist | Intermediate | 160,000 |
| | Senior | 55,000 |
| | Entry level | 21,987 |
| Analytics Engineer | Senior | 195,000 |
| | Expert | 155,000 |
| Applied Data Scientist | Senior | 278,500 |
| | Entry level | 110,037 |
| | Intermediate | 105,619 |
| Applied Machine Learning Scientist | Intermediate | 178,800 |
| | Entry level | 31,875 |
| BI Data Analyst | Expert | 150,000 |
| | Intermediate | 78,086 |
| | Entry level | 32,136 |
| Big Data Architect | Senior | 99,703 |
| Big Data Engineer | Senior | 111,536 |
| | Intermediate | 33,537 |
| | Entry level | 30,703 |
| Business Data Analyst | Entry level | 79,551 |
| | Intermediate | 74,785 |
| Cloud Data Engineer | Senior | 160,000 |
| | Intermediate | 89,294 |
| Computer Vision Engineer | Entry level | 54,536 |
| | Senior | 34,302 |
| Computer Vision Software Engineer | Entry level | 110,000 |
| | Intermediate | 95,746 |
| Data Analyst | Expert | 120,000 |
| | Senior | 111,923 |
| | Intermediate | 71,699 |
| | Entry level | 53,961 |
| Data Analytics Engineer | Intermediate | 110,000 |
| | Senior | 64,599 |
| | Entry level | 20,000 |
| Data Analytics Lead | Senior | 405,000 |
| Data Analytics Manager | Senior | 127,134 |
| Data Architect | Senior | 182,077 |
| | Intermediate | 166,667 |
| Data Engineer | Expert | 245,500 |
| | Senior | 137,036 |
| | Intermediate | 85,986 |
| | Entry level | 58,934 |
| Data Engineering Manager | Senior | 159,000 |
| | Expert | 79,833 |
| | Intermediate | 59,303 |

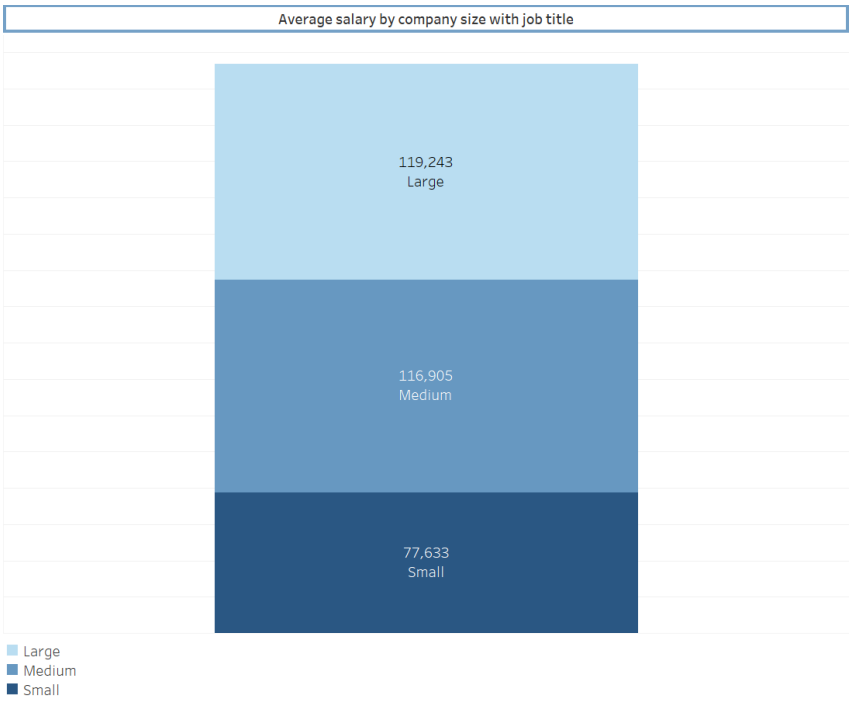
The table depicts a complete analysis of average earnings in the data science industry across various job titles and experience levels. It demonstrates a significant pay disparity, with entry-level posts starting at low salaries and expert and senior roles commanding significantly more excellent pay, reflecting the premium on experience and specialized abilities. Positions such as Principal Data Engineer and Data Scientist have some of the highest pay scales, emphasizing the importance of advanced skill and leadership in data-driven roles. This detailed analysis aids in understanding the value of each role in the job market and the salary progression with more expertise.

## ❖*How does the type of job (full-time, part-time, contract, freelance) affect the average wage for data scientists?*



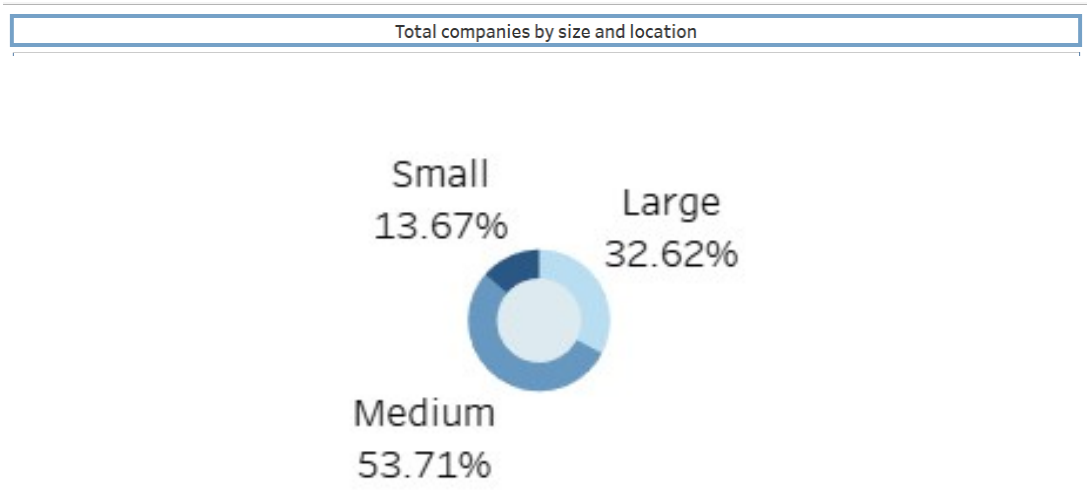

The graph shows that full-time data science roles have the highest average compensation across all experience levels, offering the value put on steady, ongoing employment. Contract employment also has a high earning potential, especially for senior-level experts, indicating that corporations are willing to invest heavily in skilled people for specific projects. The salary range for freelance and part-time work is minor, meaning the trade-offs between flexibility and remuneration. This visualization emphasizes the effect of job type on salary in the data science sector.

## ❖*How does firm size (small, medium, large) affect data science salaries, and is there a consistent pattern across different locations?*



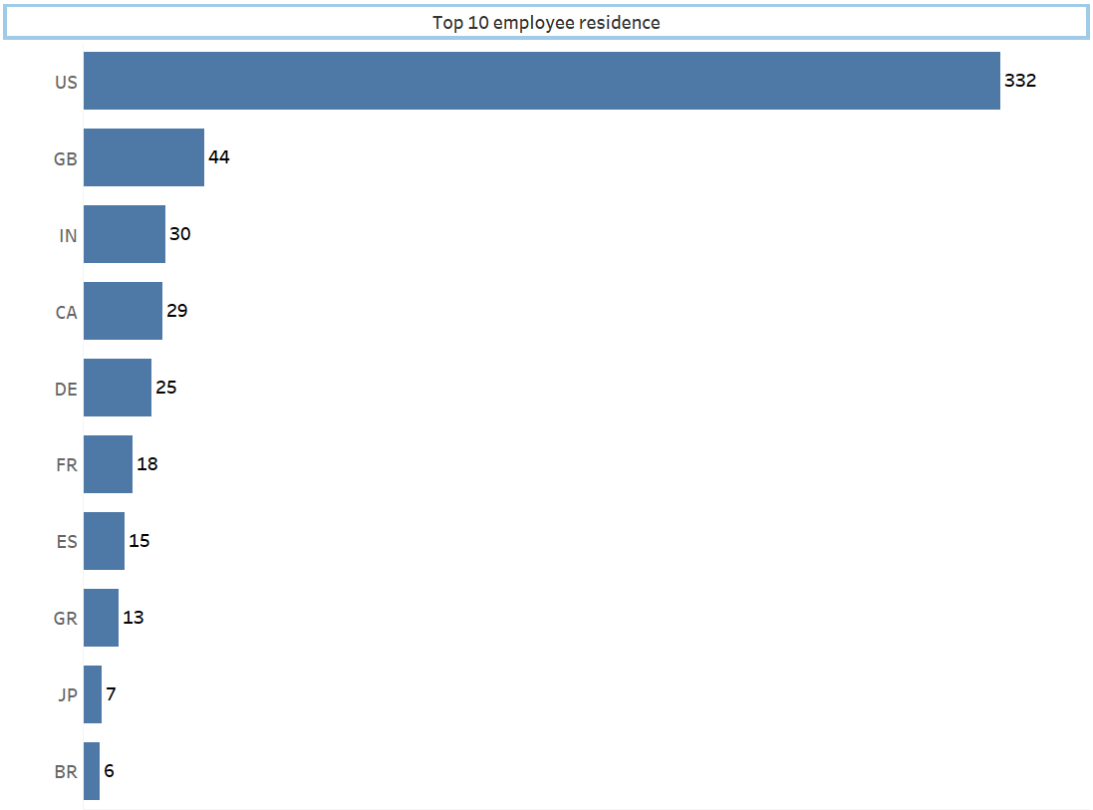

The bar graph depicts the average salary for data science positions in small, medium, and enormous businesses. Large corporations pay the highest wages, which may be attributable to their resources and the magnitude of their operations. Medium-sized businesses closely follow, implying competitive pay to recruit talent in a market where they compete with larger organizations. Small companies typically pay significantly lower wages, which may reflect budget restrictions or a planned allocation of resources to growth and development in other areas. This visualization shows how the size of a corporation affects salary in the field of data science.

## ❖*How are data science employment opportunities distributed across different organization sizes and locations?*



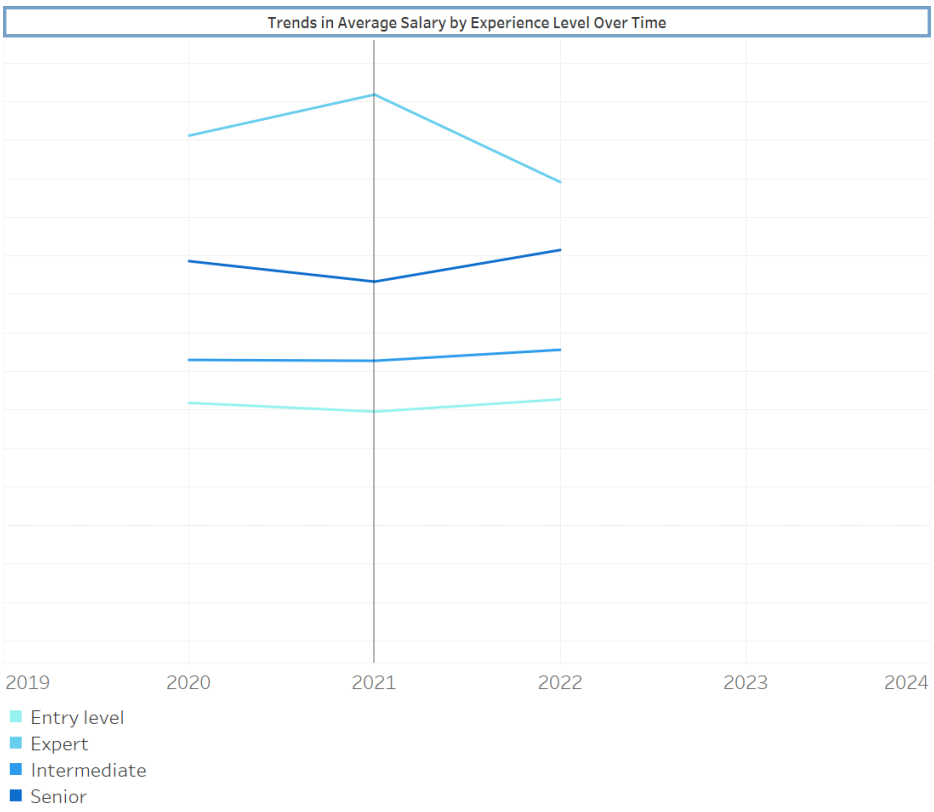

According to the pie chart, medium-sized businesses dominate the data science job market, accounting for over half of all employers. Large corporations also account for a considerable percentage of the job market, whereas small businesses account for a smaller portion. This shows that, while possibilities exist across all firm sizes, medium-sized businesses are the most active employers of data science specialists, balancing resource availability and agility in implementing data-driven strategies.

❖***How does the distribution of employee residences match the availability of data science employment and the accompanying salary ranges in those regions?***




According to the bar chart, the United States has the most data science experts among the top ten employee households, with much more than any other country. The United Kingdom and India are following, both of which are significant hubs for data science knowledge. This distribution indicates a high correlation between job availability in these regions and anticipated competitive compensation ranges provided, showing the concentration of tech industries and the demand for data science skills in these areas.

## ❖*How do salary progression trends in data science correlate with career advancement from entry-level to senior positions over the past few years?*




According to the graph, earnings in the data science sector have evolved based on expertise level. The average wage for senior posts has risen significantly, reflecting the increased need for experienced data scientists who can lead projects and teams. Intermediate occupations exhibit a moderate increase, while entry-level positions remain relatively stable. This suggests that as data scientists advance in their careers, their pay increases significantly, underscoring the importance of industry experience and knowledge.

❖***What conclusions can be derived from the employment composition by company size, particularly in data science roles and pay?***

| Employment Composition by Company Size | | | | |
|---|---|---|---|---|
| | Contract | Freelance | Full time | Part time |
| Large | 0.33% | 0.16% | 31.80% | 0.33% |
| Medium | 0.33% | 0.33% | 52.39% | 0.66% |
| Small | 0.16% | 0.16% | 12.69% | 0.66% |

% of Total Cou..

0.16% 52.39%

The grid shows that full-time positions are the most common type of employment across all firm sizes, particularly in medium-sized businesses, which account for more than half of such positions. Large organizations also prefer full-time positions, albeit in a smaller proportion than medium-sized companies. Contract and freelance work are uncommon across the board, and part-time positions are the least popular in data science, regardless of organization size. This indicates a high demand for stable full-time employment in the data science market.

# *INTERACTIVE DASHBOARD*

I've included various interactive components to allow a more meaningful review of the data in my interactive dashboard on Data Science job salaries:

- Customizable Views: I may personalize the information I view using the job title, company location, and experience level filters. This enables me to separate variables and comprehend their impact on pay in various circumstances.

- Dynamic Data Exploration: The dashboard's interactive features, such as parts of the graphs or the map, allow me to go further into individual data points. For example, by clicking on a country on the globe, I can quickly see the average income for that place, making data analysis more interactive and exciting.

- Trend Analysis: Salary by experience level trend lines over time visually show an increase or decline. I may engage with these patterns to see which experience levels are seeing the highest compensation rises and why this is happening.

- Comparative Analysis: Using the dashboard, I can compare incomes based on firm size, contract type, and employment type. This is not a static comparison; I can interact with the data to change what I'm comparing nearly on the fly.

- Interactive Statistical Summaries: The dashboard also allows me to hover over individual sections to obtain immediate statistical data, such as the exact percentage of workers employed in a particular role.

THE INTERACTIVE DASHBOARD ATTACHED IN THE TABLEAU FILE.

# *CONCLUSION*

In this research, I thoroughly investigated the various elements influencing compensation trends in the data science employment market. Through rigorous examination, I've demonstrated how factors such as professional experience, geographical location, and the employing firm's size may all significantly impact a data scientist's salary. My interactive dashboard exemplifies the potential of dynamic visualization technologies, exhibiting their capacity to bring complex data to life in an approachable and entertaining manner.

This study and its associated dashboard are intended to serve as a practical reference for aspiring data scientists. By interacting with the data, they can better understand which segments within the field are the most lucrative and which abilities they may need to develop to fetch higher salaries.

Furthermore, the paper emphasizes the significance of ongoing learning and adaptability in a continually evolving industry. Those prepared to spend time improving their skills, particularly in growing areas of data science, may be better positioned for high-paying employment.

Finally, this paper aims to provide knowledge to those on the verge of a career in data science. Understanding that can assist them in making strategic decisions, identifying development areas, and seeing a future in which they contribute to and prosper within the profession. I intend to inspire a data-driven approach to career planning within the vibrant and ever-changing field of data science by giving a detailed, interactive look at wage statistics.

REFERENCES:-

Source: https://www.kaggle.com/datasets/ruchi798/data-science-job-salaries

ai-jobs.net. https://ai-jobs.net/salaries/download/