

**DATA VISUALIZATION**

Bachelor of Technology

**Project name: DATA SCIENCE JOB SALARY DASHBOARD**

Submitted by: Abhishek Upadhyay

Registration no: 12107479

Section: K21DH

Roll no: 17

Submitted To: Tanima Thakur

LinkedIn id: www.linkedin.com/in/abhishek7upadhyay

DECLARATION

I hereby declare that the work presented in this project report title data science salary Dashboard is

original and has been carried out by me under the supervision of baljinder Mam, as part of

DATA VISUALIZATION. All sources of information and data used in this project have been

duly acknowledged and referenced. Any assistance received from individuals or organizations

has been acknowledged in the report. The views and conclusions presented in this report are

solely those of the author and do not necessarily reflect the views of any affiliated institution

or organization.

**ACKNOWLEDGEMENT**

I am deeply grateful to all those who have contributed to the successful completion of this project,

particularly in the development of the comprehensive dashboard and its multifaceted components.

First and foremost, my heartfelt appreciation goes to [Name], my esteemed project supervisor, whose

unwavering support, expert guidance, and invaluable insights have been pivotal in shaping the

trajectory of this endeavor. Their mentorship has not only enriched my understanding of the subject

matter but also inspired me to strive for excellence in every aspect of this project.

I extend my sincere gratitude to the Airbnb community for generously providing access to the wealth

of data that underpins the analysis and visualization within the dashboard. Special recognition is due

to [Name/Team/Organization], whose cooperation and assistance facilitated the acquisition and

interpretation of the data, laying the groundwork for our exploratory insights and informed conclusions.

Furthermore, I wish to acknowledge the indispensable contributions of my colleagues and peers who

have generously shared their expertise, offered constructive feedback, and provided unwavering

support at every juncture of this project's development. Their collaborative spirit and collective efforts

have significantly enriched the quality and depth of the dashboard, fostering a dynamic environment

conducive to innovation and growth.

I am also indebted to the numerous individuals within the academic and professional community whose

research, publications, and open-source contributions have served as a wellspring of knowledge and

inspiration, informing our methodologies and enriching our analytical frameworks.

Moreover, I extend my heartfelt gratitude to my friends and family for their unwavering

encouragement, patience, and understanding throughout the duration of this project. Their unwavering

support and belief in my abilities have been a constant source of motivation and resilience, sustaining

me through the inevitable challenges and triumphs of this undertaking.

In addition, I would like to express my appreciation for the technical support and resources provided

by [Institution/Organization], whose infrastructure and facilities have been instrumental in facilitating

the execution and completion of this project.

Finally, I wish to dedicate this project to [Optional dedication], whose memory and legacy continue to

inspire me to pursue excellence and make meaningful contributions to the field.

To all those mentioned above and to countless others who have contributed in ways seen and unseen, I extend my deepest gratitude and appreciation. This project stands as a testament to the collective

efforts, collaboration, and support that have made it possible

|  |  |  |
| --- | --- | --- |
| . Sr. no. | title | Page no. |
| 1 | Introduction | 06 |
| 2 | Implementation | 07-08 |
| 3 | Source of dataset | 09 -12 |
| 4 | List of analysis | 12-17 |
| 5 | Analysis of dataset | 17-21 |
| 6 | Conclusion | 22 |
| 7 | Bibiliography | 23 |
| 8 | Reference | 23 |

**INTRODUCTION**

In an era where data reigns supreme, harnessing its power to derive actionable insights is crucial. This project endeavors to provide stakeholders with a comprehensive dashboard designed to offer a nuanced understanding of data science salaries worldwide.

In the rapidly evolving landscape of data science, where opportunities span the globe, understanding the dynamics of the job market becomes essential for professionals seeking optimal career paths and policymakers aiming to foster competitive yet equitable environments.

The dashboard encompasses a range of key metrics, each shedding light on different facets of the global data science job market:

Total Job Openings by Region: This metric provides insights into the popularity of various regions among data science employers, offering valuable information for professionals seeking employment opportunities in specific areas of the world.

Average Salary per Role and Region: Salary serves as a crucial indicator of job competitiveness and the overall compensation package offered by employers. By analyzing salary data across different roles and regions, stakeholders can gauge the market value of their skills and identify areas for professional growth.

Salary by Company Size and Region: Understanding salary dynamics across different company sizes and regions provides valuable insights into market trends and employer preferences. Data science professionals can leverage this information to negotiate competitive compensation packages and make informed career decisions.

Average Salary by Experience Level and Region: This metric delves deeper into the intersection of experience levels and regions, offering granular insights into salary variations within the global data science job market.

By synthesizing these disparate data points into a user-friendly dashboard interface, this project aims to empower stakeholders with the knowledge and tools needed to navigate the complex landscape of data science salaries worldwide effectively. Whether it be professionals seeking to optimize their career trajectories on a global scale or policymakers endeavoring to foster vibrant and inclusive job markets across borders, the dashboard serves as a valuable resource for all stakeholders invested in the data science ecosystem.

**Implementation Phase: Tableau Dashboard for Data Science Salaries**

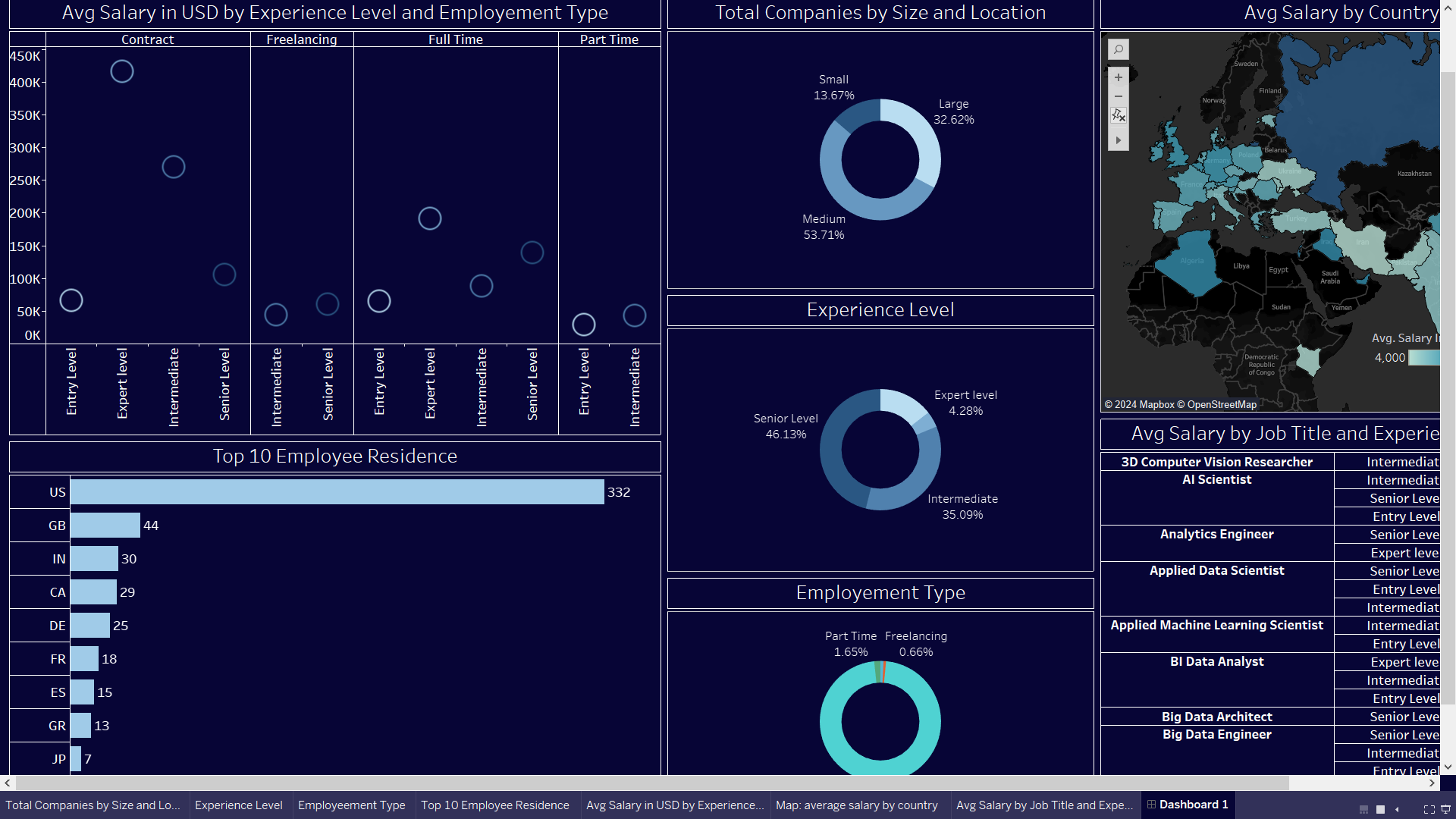
The implementation phase of this Tableau project involved the development and design of a dynamic and interactive dashboard aimed at providing stakeholders with comprehensive insights into data science salaries worldwide. Leveraging Tableau's powerful visualization capabilities, the dashboard was meticulously crafted to encompass a range of key metrics, each shedding light on different facets of the global data science job market.

**Data Preparation:**

The first step in the implementation process involved data preparation and cleansing to ensure the accuracy and integrity of the dashboard. The dataset, collected from various sources including industry surveys and job listings, was carefully curated to include relevant variables such as job roles, salaries, experience levels, company sizes, and geographic regions. Any inconsistencies or missing values were addressed through data cleaning techniques to facilitate meaningful analysis.

**Dashboard Design:**

With the dataset primed for analysis, the next step was the design of the dashboard interface. Drawing inspiration from user-centered design principles, the dashboard was crafted with usability and functionality in mind, featuring intuitive navigation and visually compelling displays.



**Key Metrics and Visualizations:**

The dashboard comprises a suite of key metrics and visualizations aimed at providing stakeholders with actionable insights into data science salaries worldwide:

Total Job Openings by Region: This visualization presents a map depicting the total number of job openings across different geographic regions, allowing users to identify areas with high demand for data science professionals.

Average Salary per Role and Region: Leveraging bar charts and heat maps, this visualization provides a comprehensive overview of salary trends across various job roles and geographic regions, enabling users to gauge the market value of different skills and roles.

Salary by Company Size and Region: A scatter plot showcases salary dynamics across different company sizes and geographic regions, allowing users to compare compensation packages offered by companies of varying sizes in different parts of the world.

Average Salary by Experience Level and Region: This visualization features line charts and histograms highlighting salary variations based on experience levels and geographic regions, facilitating deeper analysis of compensation trends within the global data science job market.

**Interactivity and Drill-Down Functionality:**

To enhance user engagement and facilitate deeper exploration of the data, the dashboard incorporates interactive features such as filter controls, tooltips, and drill-down functionality. Users can dynamically filter and drill down into specific job roles, geographic regions, company sizes, and experience levels to uncover hidden patterns and trends within the data.

**Conclusion:**

In conclusion, the implementation of this Tableau dashboard represents a significant milestone in providing stakeholders with actionable insights into data science salaries worldwide. By leveraging Tableau's robust visualization capabilities and user-centric design principles, the dashboard offers a comprehensive overview of key metrics and trends, empowering users to make informed decisions and strategic interventions within the dynamic landscape of the global data science job market.

**Implementation Hierarchy Model: Data Science Salary Dashboard**

Data Acquisition and Preparation:

* Identify and access relevant datasets from sources such as industry surveys, job listings, and government labor departments.
* Preprocess the data to clean, transform, and format it for analysis. This may involve handling missing values, aggregating data by relevant variables, and ensuring data consistency.

Data Analysis and Modeling:

* Perform exploratory data analysis (EDA) to gain insights into the dataset's characteristics, distributions, and relationships between variables.
* Apply statistical and machine learning techniques as appropriate to analyze the data and derive meaningful insights. This may include regression analysis, clustering, or time series forecasting.
* Develop models to predict data science salaries based on various factors such as job roles, experience levels, and geographic regions.

Dashboard Development:

* Design the layout and user interface of the data science salary dashboard using tools such as Tableau.
* Integrate the processed data and analysis results into the dashboard, ensuring real-time updates and interactive features for users.
* Incorporate visualizations, charts, and tables to present the analysis findings in a clear and intuitive manner. Include key metrics such as average salary trends, salary distribution by role, and geographic salary comparisons.

Testing and Validation:

* Conduct rigorous testing of the dashboard to ensure functionality, performance, and accuracy of the data visualizations and interactive features.
* Validate the analysis results against external sources such as industry reports and salary surveys to verify their reliability and consistency.
* Gather feedback from stakeholders and users to identify any usability issues or areas for improvement.

Deployment and Maintenance:

* Deploy the data science salary dashboard to the intended audience or platform, such as a web server or cloud service.
* Monitor the dashboard's performance and user engagement metrics to track its effectiveness in providing insights and informing decision-making.
* Regularly update the dashboard with new data and analysis results to ensure its relevance and timeliness in addressing evolving trends in the data science job market.

This model outlines the step-by-step process for developing and deploying your data science salary dashboard, ensuring its effectiveness in providing valuable insights to stakeholders.

**Source of Dataset:**

The dataset used in this data science salary dashboard was primarily sourced from Kaggle, a renowned platform for hosting datasets and machine learning competitions. Specifically, we utilized a comprehensive dataset containing data science job salary statistics, including job roles, experience levels, company sizes, and geographic regions. You can access the dataset using the following links:

Link : <https://www.kaggle.com/datasets/ruchi798/data-science-job-salaries>

While Kaggle datasets are generally well-curated and reliable, it's important to note that they may still have inherent limitations and biases. Therefore, rigorous validation and verification were conducted to ensure the accuracy and reliability of the data used in this analysis. Additionally, data from other reputable sources such as Geographical Analysis and the Data Science RoadMap GitHub repository were utilized to enrich the analysis and provide comprehensive insights into data science salaries worldwide.

**Title of the Dashboard:**

**Introduction:**

The analysis on the dataset for "Total Salaries by Job Roles and Experience Levels" aims to provide insights into the distribution of data science salaries across different job roles and experience levels worldwide. By examining the total salary amounts within each job role and experience level category, this analysis seeks to identify patterns and trends that can inform strategic decisions for professionals and organizations within the data science job market.

**II. General Description:**

The dataset under analysis comprises a comprehensive record of data science salaries worldwide, encompassing various attributes such as job role, experience level, salary amount, and geographic region. Each entry in the dataset represents a distinct salary record collected from various sources, providing a rich source of data for analysis.

III. Specific Requirements, Functions, and Formulas:

* To conduct the analysis on "Total Salaries by Job Roles and Experience Levels," several specific requirements, functions, and formulas are essential:
* Data Aggregation: The dataset needs to be aggregated to calculate the total salary amounts for each combination of job role and experience level. This involves grouping the data by job role and experience level and summing up the salary amounts within each group.
* Filtering: The dataset may contain irrelevant entries or outliers that need to be filtered out to focus solely on valid salary records within the scope of the analysis.

**Calculation of Total Salaries**: A formula is required to calculate the total salary amount for each job role and experience level category. This formula involves summing up the salary amounts within each group obtained through data aggregation.

**IV. Analysis Results:**

Upon conducting the analysis, the following results are obtained:

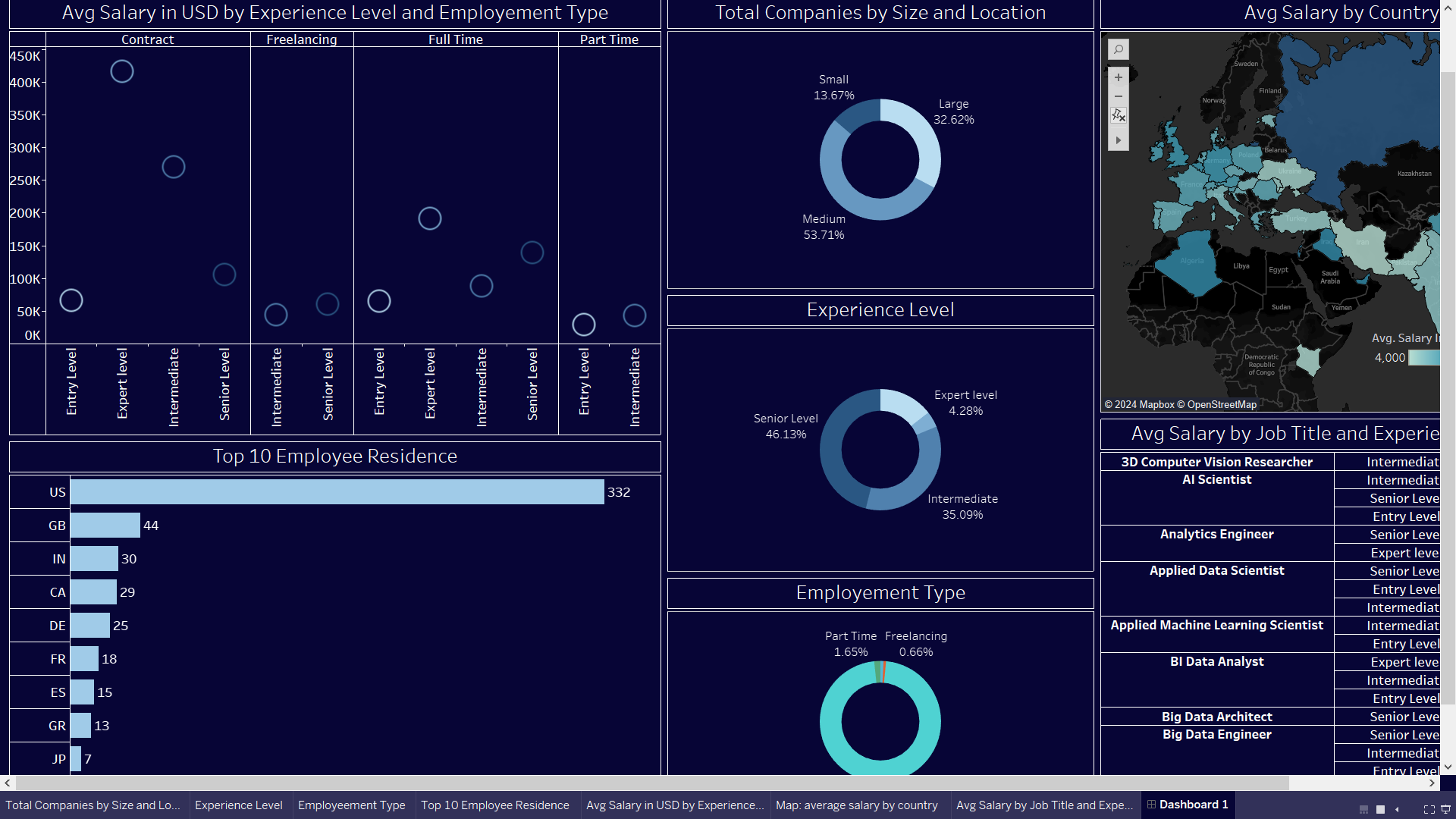
* Distribution of Salaries: The analysis reveals the distribution of salaries across different job roles and experience levels. It provides insights into which job roles and experience levels command higher salaries in the data science job market.
* Variability Across Job Roles: Disparities in salary amounts between job roles may indicate varying levels of demand and competitiveness within the data science field.

**Preference for Experience Levels:** The analysis highlights any preferences or trends in the choice of experience levels among data science professionals across different job roles.

**V. Visualization:** To effectively communicate the analysis results, visualization plays a crucial role. The following types of visualizations can be employed:

* Bar Chart: A bar chart can visually represent the total salary amounts for each job role, with each bar segmented by experience level to illustrate the distribution of salaries within each role.
* Heat Map: A heat map can provide a visual representation of salary intensity across job roles and experience levels, with color gradients indicating the relative salary amounts.
* Stacked Bar Chart: A stacked bar chart can depict the total salary amounts for each experience level, with individual bars segmented by job role to facilitate comparison and identification of trends.

These visualizations help stakeholders grasp the distribution and trends of salaries across job roles and experience levels worldwide, enabling informed decision-making and strategic planning within the data science job market.



**Average Salaries per Role and Experience Level**

**Introduction:**

This analysis aims to explore the average salaries received per role and experience level in the global data science job market. By examining salary trends across various job roles and experience levels, this analysis seeks to understand compensation levels and identify potential areas for professional growth and negotiation strategies.

**General Description:**

The dataset comprises data science salary records worldwide, including attributes such as salary amounts, job roles, experience levels, and geographic regions. Each entry represents a unique salary record reported by data science professionals.

**Specific Requirements, Functions, and Formulas:**

To conduct the analysis on "Average Salaries per Role and Experience Level," several specific requirements, functions, and formulas are essential:

* Data Aggregation: The dataset needs to be aggregated to calculate the average salary amount for each combination of job role and experience level. This involves grouping the data by job role and experience level and calculating the mean salary within each group.
* Calculation of Average Salaries: A formula is required to calculate the average salary amount for each job role and experience level category. This formula involves taking the mean of the salary amounts within each group.

**Analysis Results:**

Upon conducting the analysis, the following results are obtained:

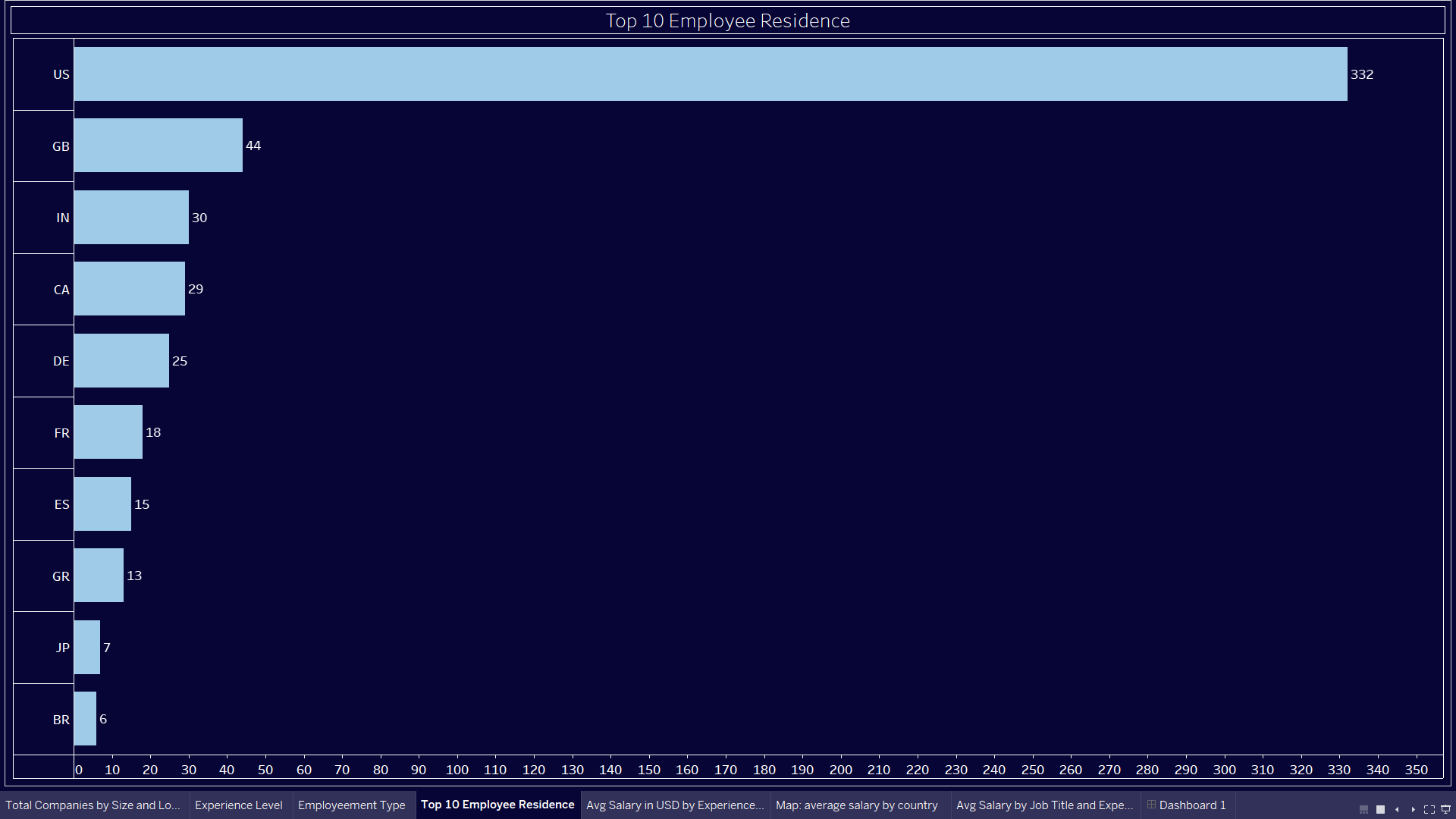
* Salary trends across different job roles and experience levels.
* Comparison of average salary amounts across job roles and experience levels.
* Identification of variations in salary levels based on experience and job role categories.

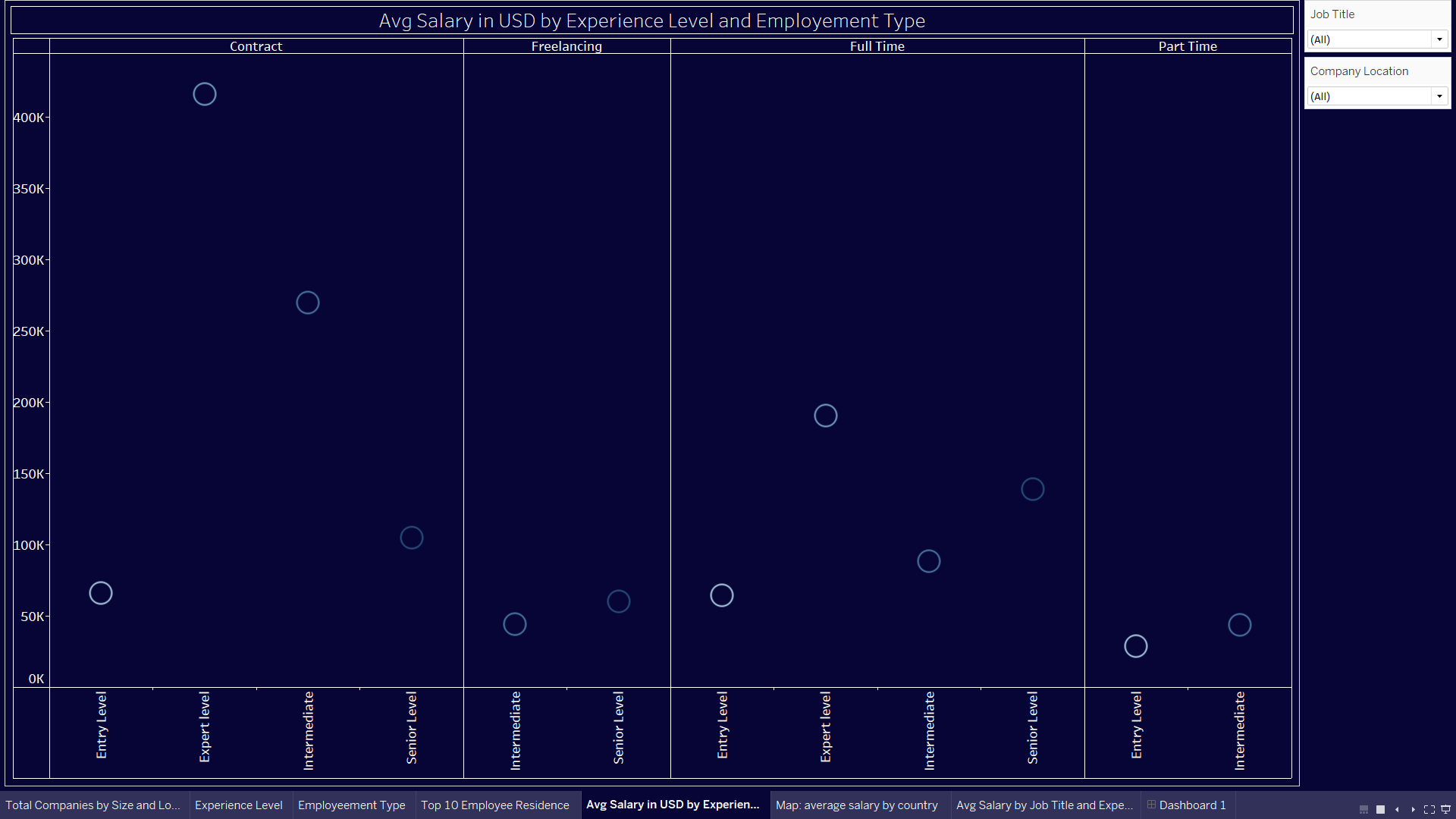
**Visualization:**

To effectively communicate the analysis results, visualization plays a crucial role. The following types of visualizations can be employed:

* Bar Chart: A bar chart can visually represent the average salary amounts for each job role, with each bar segmented by experience level to illustrate the distribution of salaries within each role.
* Heat Map: A heat map can provide a visual representation of salary intensity across job roles and experience levels, with color gradients indicating the relative salary amounts.

These visualizations help stakeholders grasp the distribution and trends of salaries across job roles and experience levels worldwide, enabling informed decision-making and strategic planning within the data science job market





**Average Salary by Geographic Region - Country/Region: Country/Region**

**Introduction:**

This analysis aims to examine the average salary of data science professionals within different geographic regions worldwide. By comparing salary dynamics across various countries and regions, this analysis seeks to understand salary disparities and trends within the global data science job market.

**General Description:**

The dataset includes data science salary records from various countries and regions, with attributes such as salary amounts, job roles, experience levels, etc.

**Specific Requirements, Functions, and Formulas:**

To conduct the analysis on "Average Salary by Geographic Region," several specific requirements, functions, and formulas are essential:

* Data Aggregation: The dataset needs to be aggregated to calculate the average salary of data science professionals within each geographic region. This involves grouping the data by country or region and calculating the mean salary within each group.
* Calculation of Average Salary: A formula is required to calculate the average salary amount for each country or region. This formula involves taking the mean of the salary amounts within each group.

**Analysis Results:**

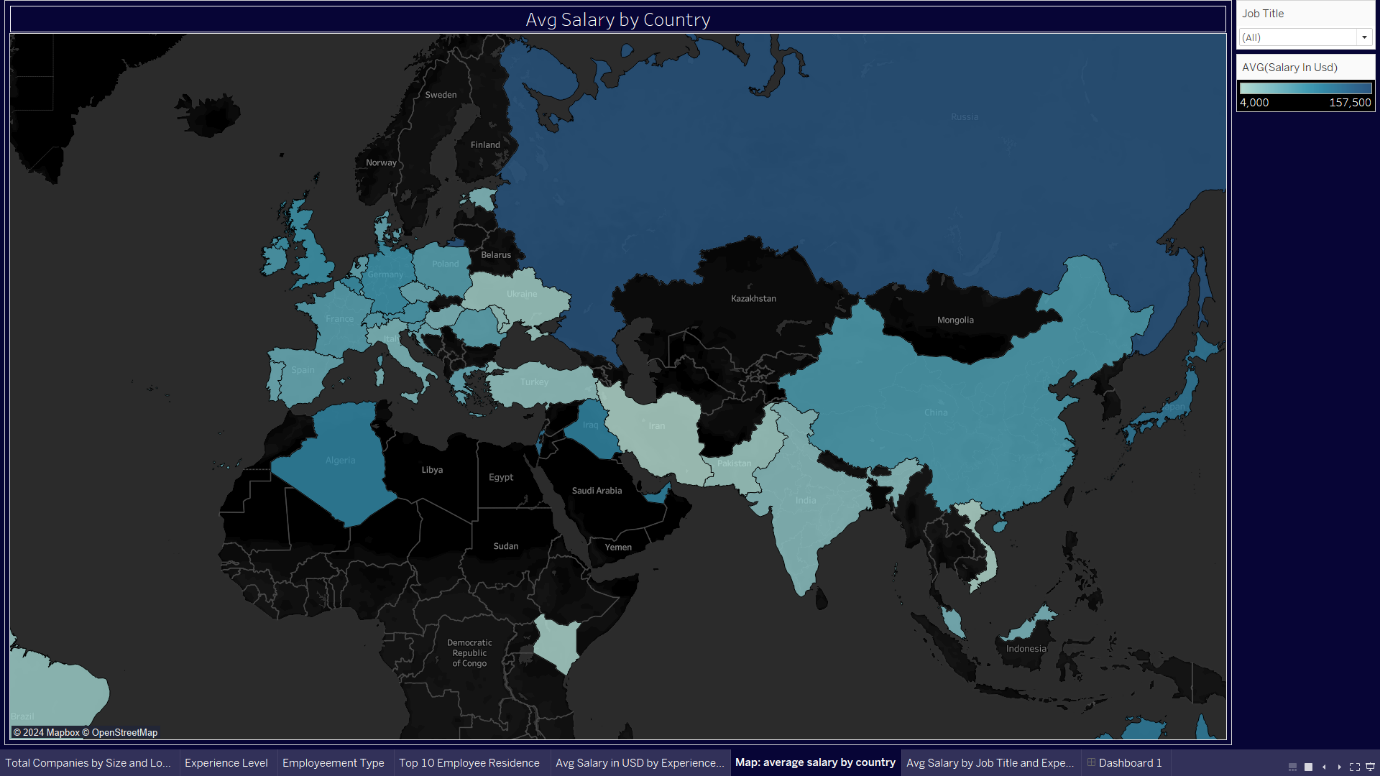
* Upon conducting the analysis, the following results are obtained:
* Comparison of average salaries across different countries and regions.
* Identification of salary disparities between different geographic areas.
* Analysis of salary trends and patterns within the global data science job market.

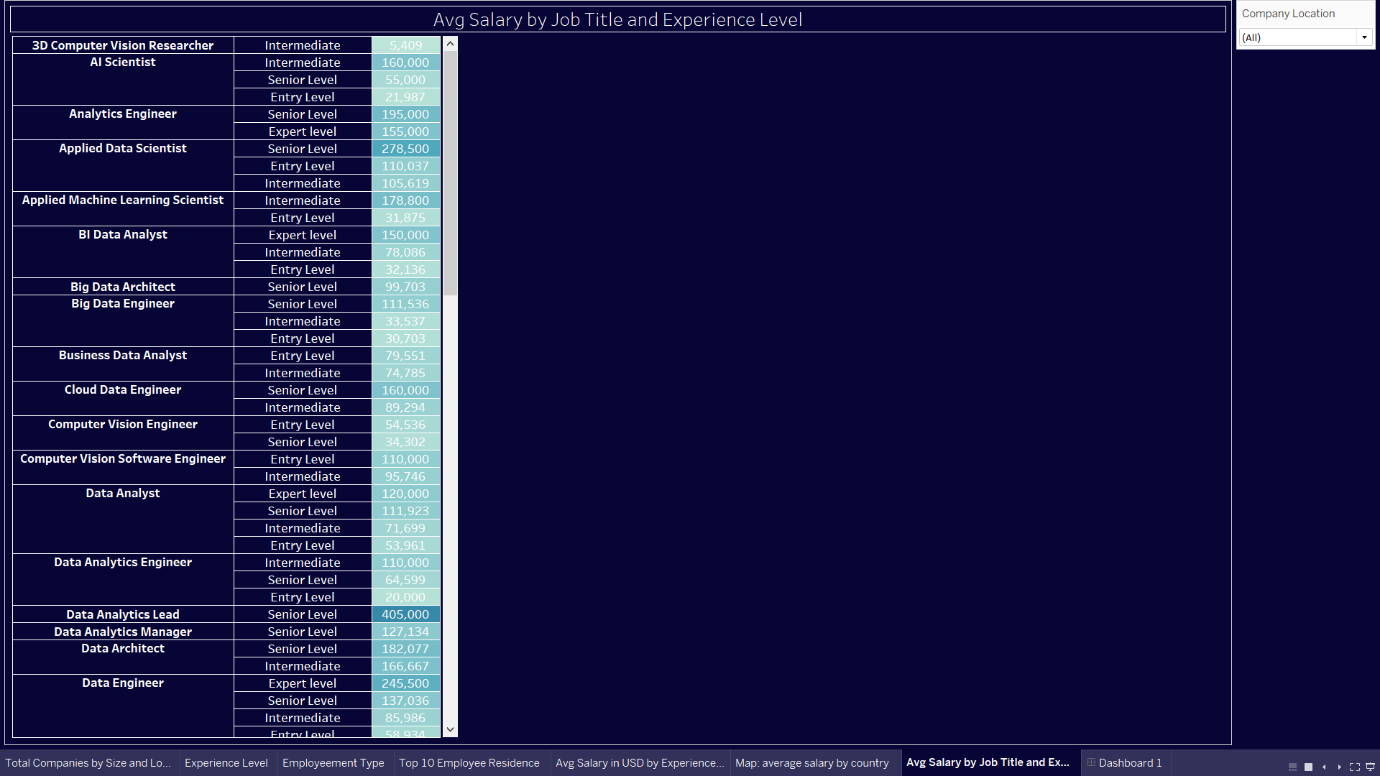
**Visualization:**

To effectively communicate the analysis results, visualization plays a crucial role. The following types of visualizations can be employed:

* Choropleth Map: A choropleth map can visually represent the average salary of data science professionals within each country or region, with color gradients indicating the relative salary amounts.
* Bar Chart: A bar chart can display the average salary amounts for each country or region, allowing stakeholders to compare salary levels across geographic areas.

These visualizations help stakeholders grasp the distribution and trends of salaries across different countries and regions worldwide, enabling informed decision-making and strategic planning within the global data science job market.





**Average Salary by Job Role and Experience Level**

**Introduction:**

This analysis aims to examine the average salary of data science professionals within each job role and experience level category. By analyzing salary variations across different segments, this analysis seeks to understand the factors influencing compensation decisions and preferences within the global data science job market.

**General Description:**

The dataset includes data science salary records worldwide, with attributes such as salary amounts, job roles, experience levels, etc.

**Specific Requirements, Functions, and Formulas:**

* To conduct the analysis on "Average Salary by Job Role and Experience Level," several specific requirements, functions, and formulas are essential:
* Data aggregation to calculate the average salary of data science professionals for each combination of job role and experience level.

Calculation of average salary using the formula: Average Salary = Total Salary / Total Professionals

**Analysis Results:**

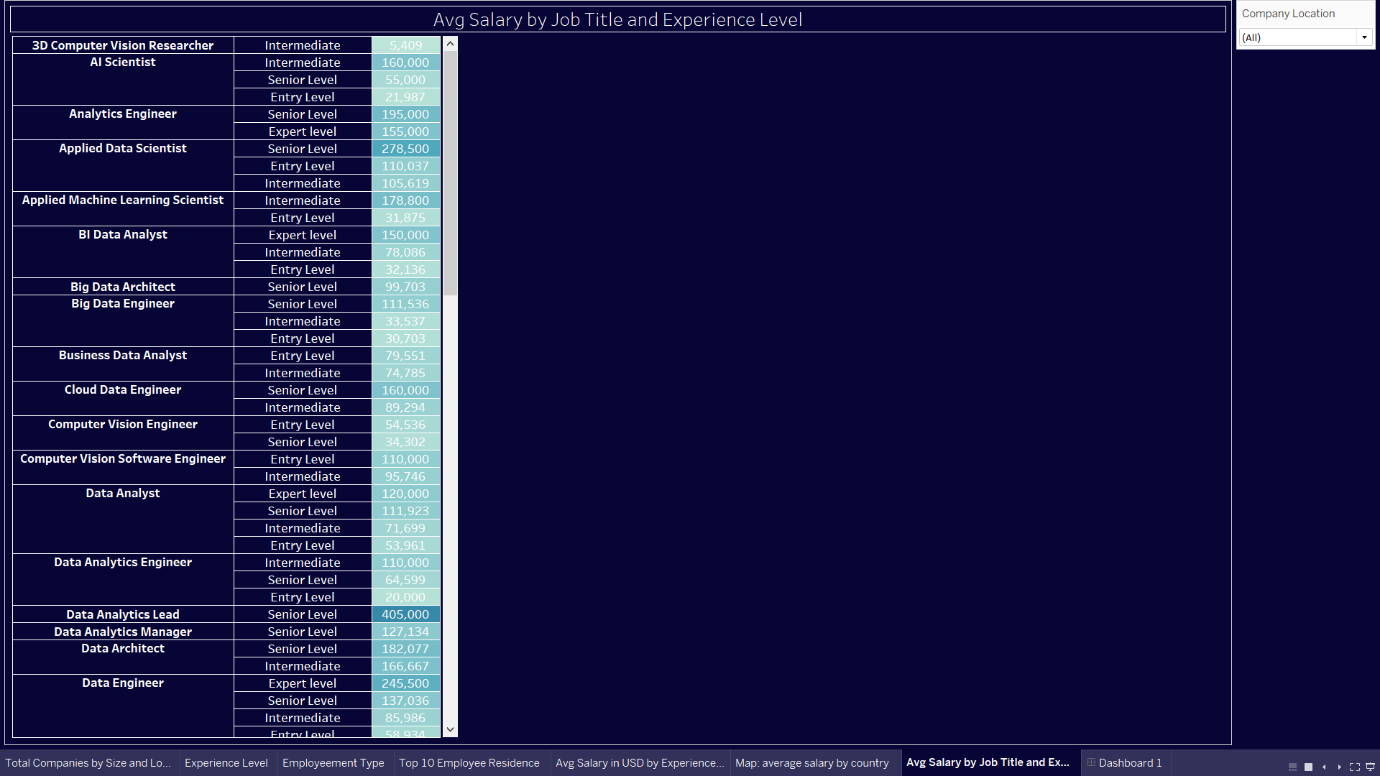
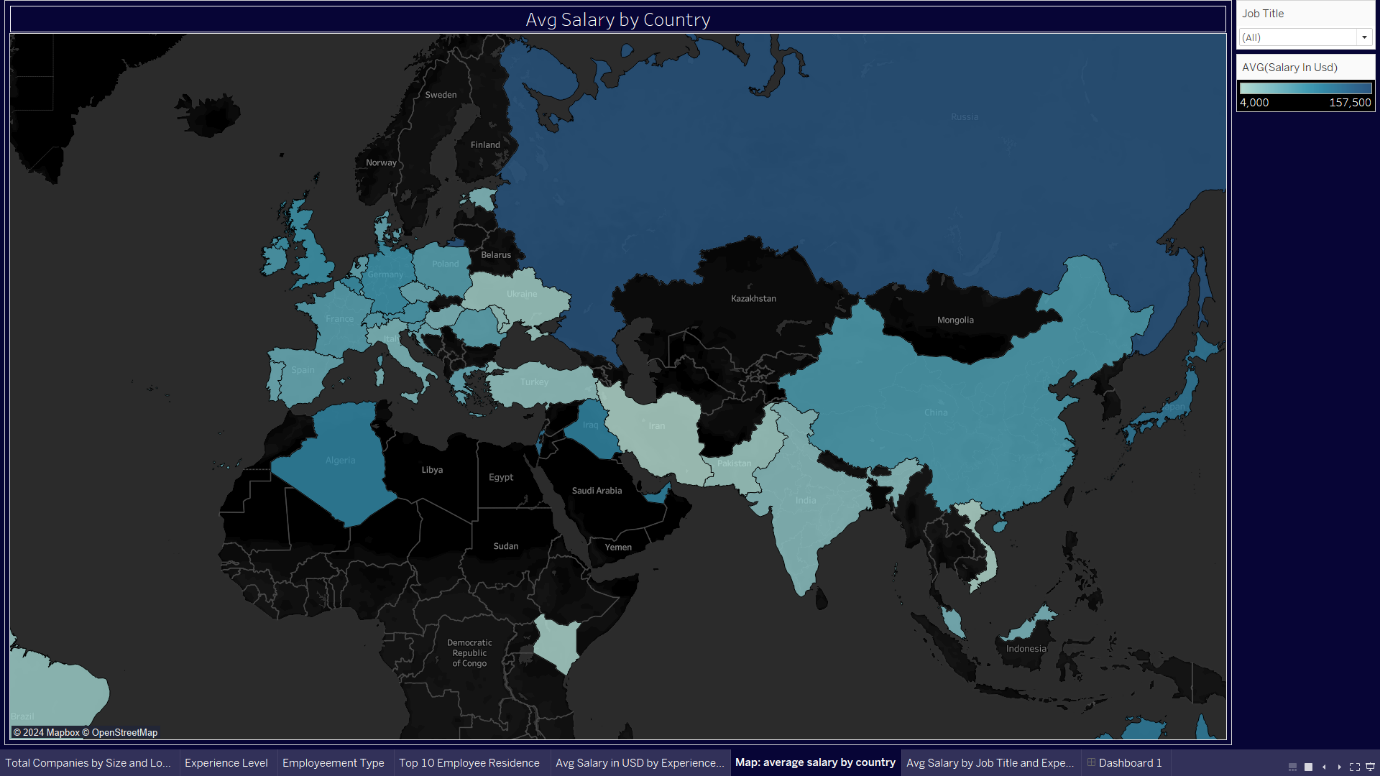
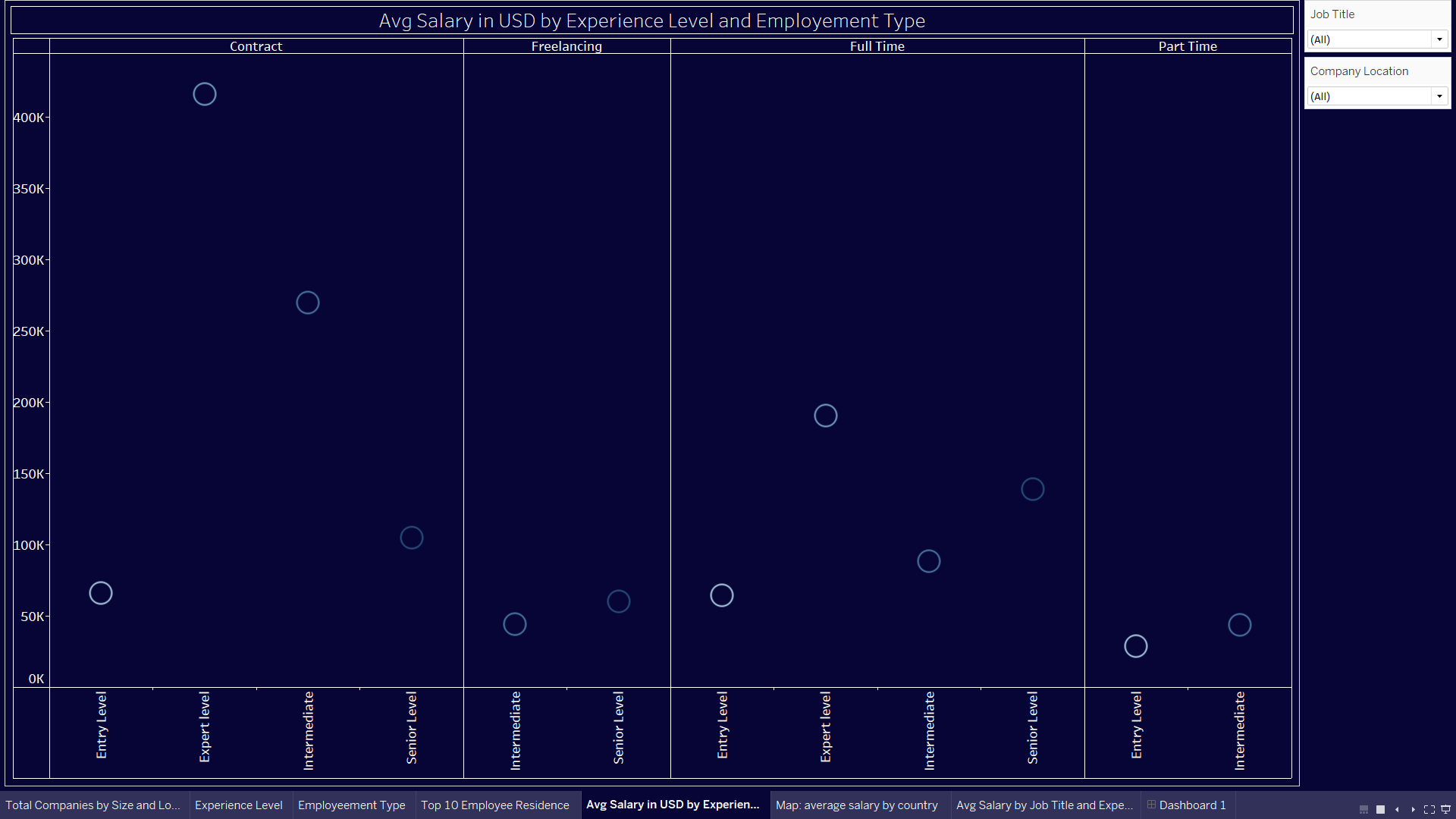
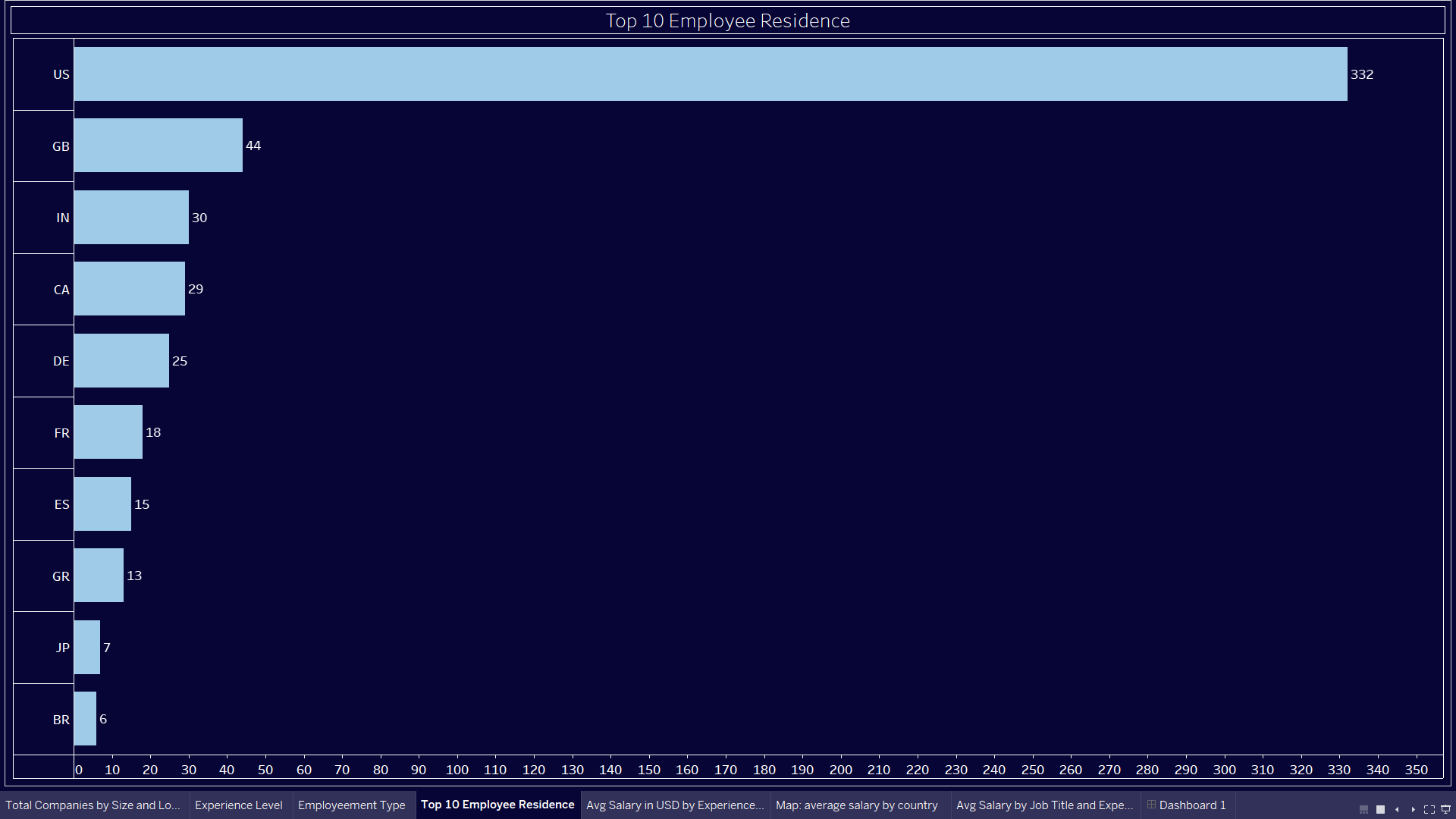
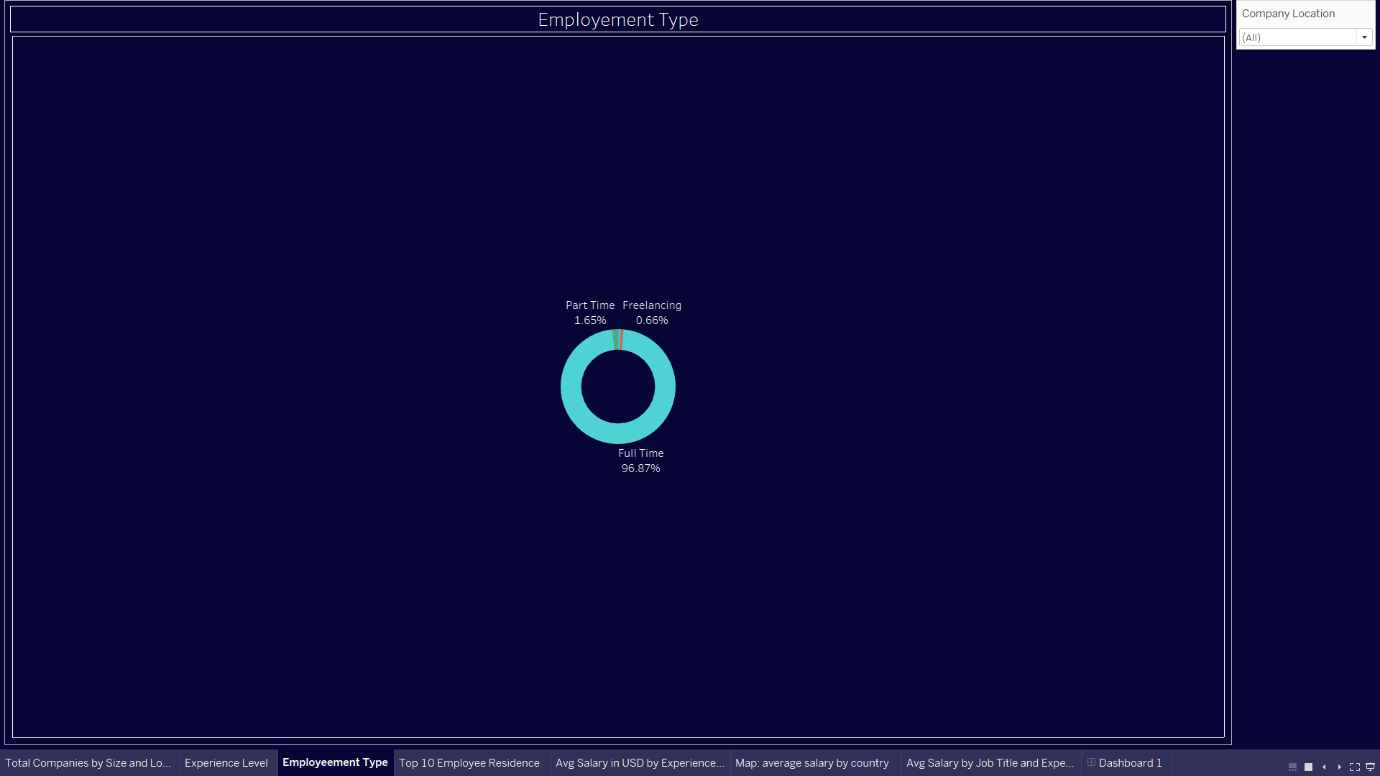
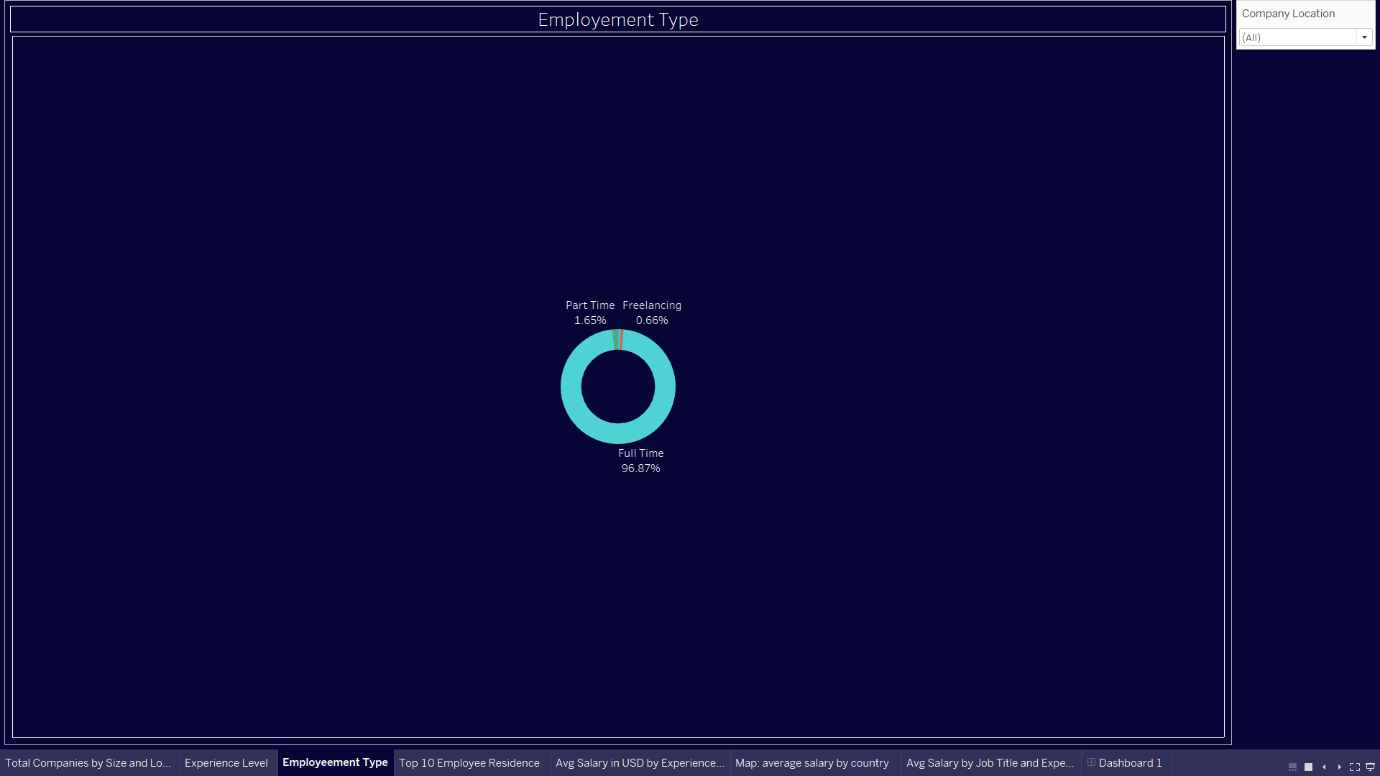
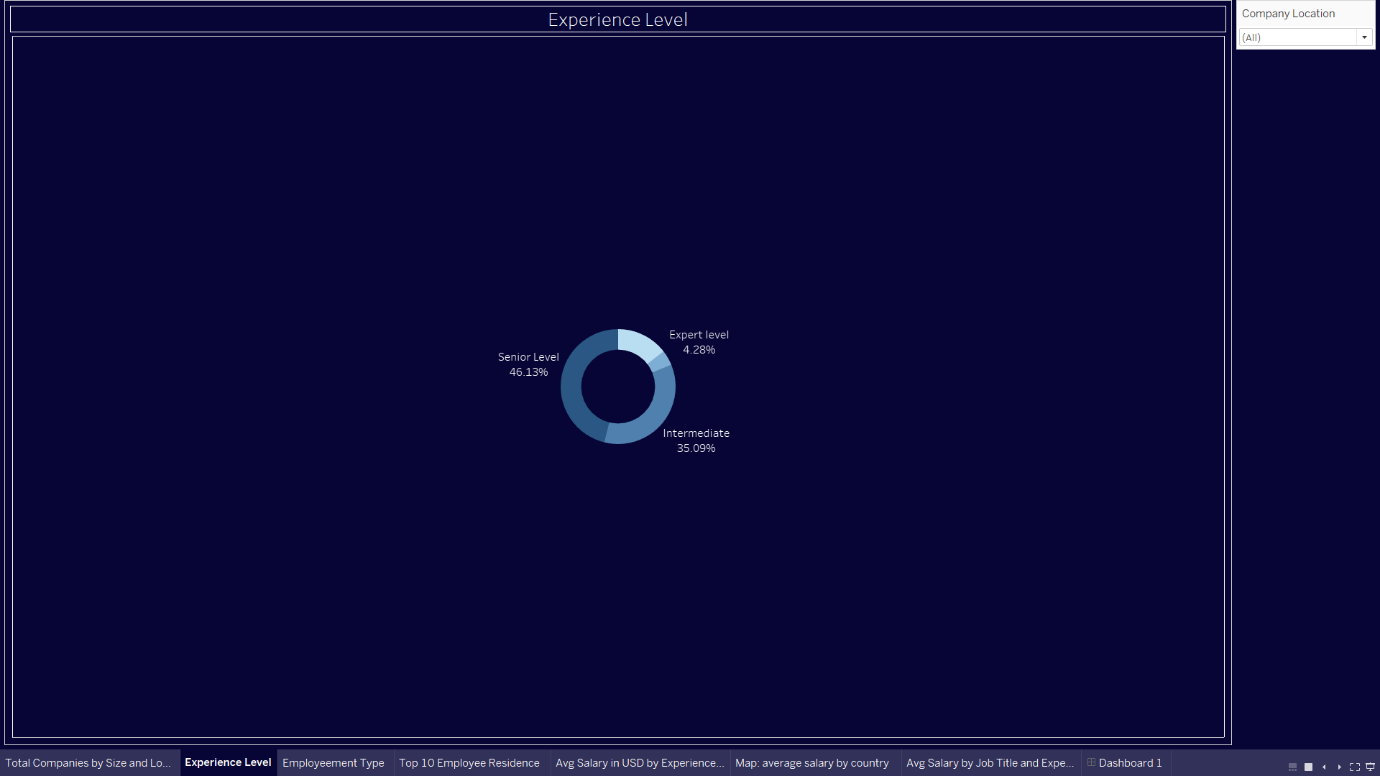
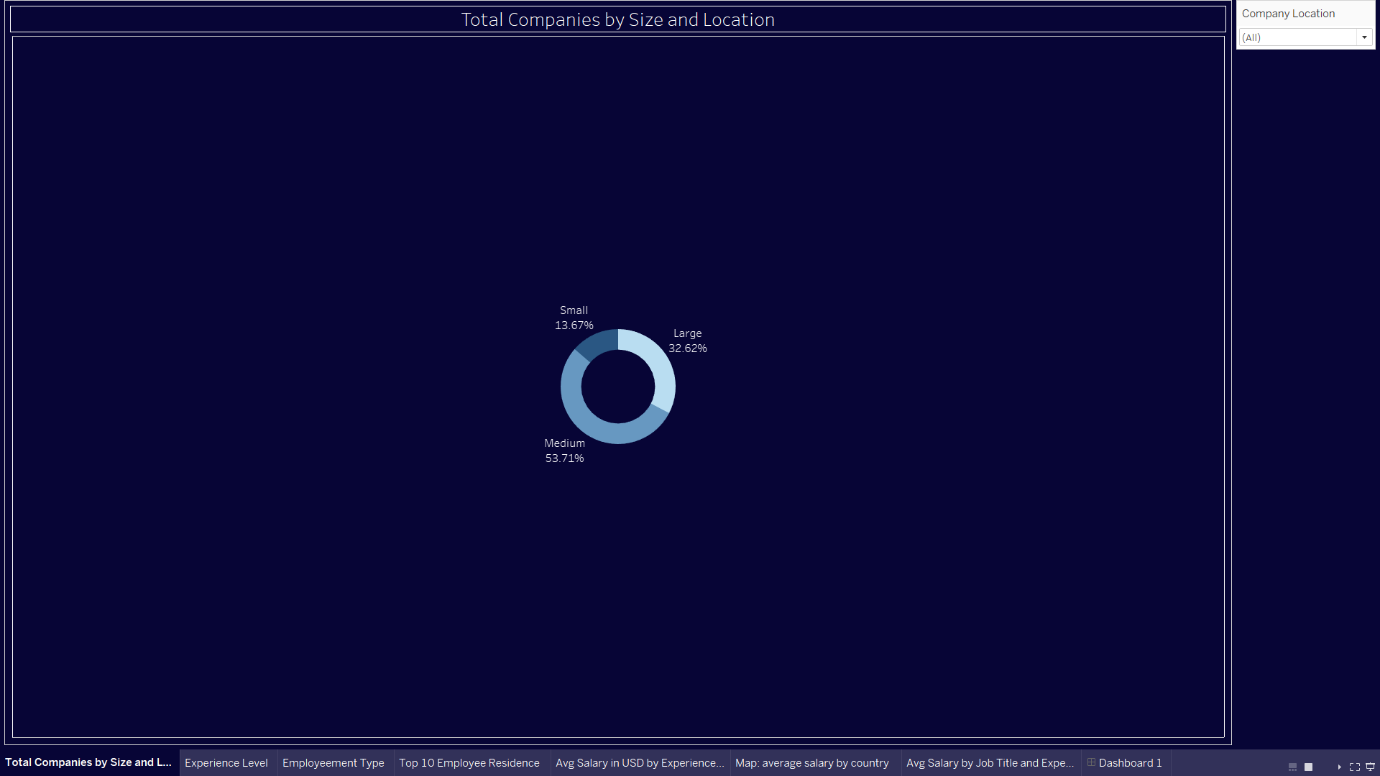
* Upon conducting the analysis, the following results are obtained:
* Comparison of average salaries across different combinations of job roles and experience levels.
* Identification of salary variations and trends within the global data science job market.

**Visualization:**

To effectively communicate the analysis results, visualization plays a crucial role. The following types of visualizations can be employed:

* Stacked Bar Chart: A stacked bar chart can visually represent the average salary of data science professionals for each combination of job role and experience level, allowing stakeholders to compare salary levels across different segments.
* Heat Map: A heat map can provide a visual representation of salary intensity across job roles and experience levels, with color gradients indicating the relative salary amounts.

These visualizations help stakeholders grasp the distribution and trends of salaries across different job roles and experience levels worldwide, enabling informed decision-making and strategic planning within the data science job market.



**Conclusion:**

The development and implementation of the data science salary dashboard represent a significant milestone in providing stakeholders with actionable insights into the global data science job market. Through meticulous data analysis and visualization, the dashboard offers a comprehensive overview of key metrics and trends, empowering users to make informed decisions and strategic interventions within the dynamic landscape of data science employment worldwide.

The dashboard provides a range of functionalities and visualizations, allowing users to explore various aspects of the data science job market, including average salaries, salary distribution across job roles and experience levels, and geographic salary comparisons. By leveraging interactive features such as filters, tooltips, and drill-down functionality, users can delve deeper into the data, uncovering hidden patterns and trends that inform their career decisions and negotiation strategies.

One of the dashboard's primary strengths lies in its ability to provide granular insights into salary dynamics across different job roles and experience levels. Whether it's identifying lucrative roles, understanding salary trends over time, or comparing compensation packages across regions, the dashboard equips users with the tools and information needed to navigate the complexities of the data science job market effectively.

Moreover, the dashboard serves as a valuable resource for a diverse range of stakeholders, including data science professionals, hiring managers, educators, and policymakers. Data science professionals can utilize the dashboard to benchmark their salaries, explore career advancement opportunities, and negotiate competitive compensation packages. Hiring managers can leverage the insights provided by the dashboard to attract top talent, benchmark salary offerings against industry standards, and tailor recruitment strategies to meet hiring needs. Educators can utilize the dashboard to inform curriculum development, identify emerging skills in demand, and prepare students for the evolving landscape of data science employment. Policymakers can leverage the insights provided by the dashboard to inform workforce development initiatives, support talent retention efforts, and foster a competitive environment for data science innovation.

In conclusion, the data science salary dashboard stands as a testament to the power of data-driven decision-making and visualization in unlocking actionable insights within the global data science job market. By providing stakeholders with the tools and information needed to navigate career decisions, negotiate competitive compensation packages, and drive strategic interventions, the dashboard contributes to the advancement and sustainability of the data science profession worldwide.

**BIBLIOGRAPHY**

• Kaggle

• coursera

• YouTube

**References:**

Kaggle. (n.d.). Retrieved from https://www.kaggle.com/

Our World in Data. (n.d.). Retrieved from https://ourworldindata.org/

GitHub. (n.d.). Retrieved from https://github.com/

Sudalairajkumar. (n.d.). Retrieved from https://www.kaggle.com/datasets/sudal...

Anshuls2382. (n.d.). Retrieved from https://www.kaggle.com/code/anshuls23...

Geographical Analysis. (n.d.). Retrieved from https://geographicalanalysis.com/downloads/