Introduction –

Data –

The dataset **ds_salaries** dataset contains comprehensive data on salaries, employment statistics, demographic information, and business details in the domain of data science for the period spanning from 2020 to 2023. The dataset consists of 11 fields and 3755 rows.

Here is the link for the dataset - https://www.kaggle.com/datasets/arnabchaki/data-science-salaries-2023

The data dictionary demonstrates the key original variable below.

Data Dictionary –

Name	Description	Domain
Work Year	The specific years the employee worked	Categorical (Nominal)
Experience Level	The level of experience of an employee	Categorical (Ordinal)
Employment Type	Type of employment like part-time or full time etc.	Categorical (Nominal)
Job Title	Name of the job title or job roles	Categorical (Nominal)
Salary	Salary of an employee	Numerical (Integer)
Salary Currency	Currency of the salary whether it is Eur or USD	Categorical (Nominal)
Salary in Usd	Salary of an employee in USD	Numerical (Integer)
Employee Residence	Name of the country the employees reside	Categorical (Nominal)
Remote Ratio	The total value of the remote work	Numerical (Discrete)
Company Location	The country or region where the company is located	Categorical (Nominal)
Company Size	The size of the company (small, medium, or Large)	Categorical (Ordinal)

Persona and Questions –

The user, an HR recruiter, seeks to get precise information on data science and other job salaries. Specifically, he is interested in exploring variations in pay based on country location, distinct data science job roles, and firm locations. He has some specific questions that include 3 simple questions and 1 Complex question. The questions are given below respectively.

Simple Question –

1. What are the average salaries in Usd based on different data science job roles and experience levels?

2. What is the total sum of salaries in Usd based on different job roles and company locations?

3. What are the top 5 employee residences based on the company size?

Complex Question –

4. What are the highest average salary and total remote ratio based on different experience levels, company size, and different job titles?

Requirements –

Here, we outline the primary needs in terms of the relationships that need to be shown to provide an answer to each question. Brief design concepts for interactions and representations are then provided, indicated by light blue sub-bullets.

Q1: “What are the average salaries in Usd based on different data science job roles and experience levels?”

R1: To answer Q1 the user has to visualize the variables AVG (Salary in Usd), Experience Level, and Job Title in a way that visualizes the relationship between them.

- A good view would be a horizontal bar graph by placing the Salary in Usd in Columns and the Experience Level and Job Title in Rows.
- By clicking on the Salary in Usd and taking the measure as Average, the average of the Salary in Usd can be created.
- Also, for a better view The AVG (Salary in Usd) could be dragged by pressing the ctrl and left key and place it on the color and label(T) in the Marks area.
- The user will be able to see the differences in average salary across the different Experience Level and Job Title which will be represented by each bar.

Q2: “What is the total sum of salaries in Usd based on different job roles and company location?”

R2: To answer Q2 the user needs to visualize the variables SUM (Salary in Usd), Company Location, and Job Title which will demonstrate the total sum of salaries based on the various job roles and company location.

- A map chart will be the perfect view to visualize this by clicking the Company Location and putting Longitude in Columns and Latitude in Rows.
- Apply filter on both Longitude and Latitude.
- Include the SUM (Salaries in Usd) by colouring it and labelling it in Marks.
- Put the Job Title in Text (Label).
- The user will be able to see which countries have their specific major job roles and the total sum of salaries on that job role.

Q3: “What are the Top 5 Employee Residence Based on Company Size?”

R3: To answer Q3 the user needs to visualize the variables Employee Residence and CNT (Employee Residence) in a way that delineates the top 5 employee residences.

- Select Employee Residence in Rows and also drag the Employee Residence (press Ctrl and left click together) and put it in Text.
- Take the measure as Count for Employee Residence Variable and place it in Columns
- Filter Employee Residence by taking the top by field 5.
- Put Company Size in Colour and Text (Label).
- This will help the user to find the top 5 employee residences based on the different company size.

Q4: “What are the highest average salary and total remote ratio based on different experience levels, company size, and different job titles?”

R4: To answer Q4 the user needs to visualize the SUM (Remote Ratio) and AVG (Salary in Usd) both based on the Variables Company Size, Job Title and Experience Level.

- Line graph would be nice view for the visualization, by doing a dual axis between the Remote Ratio and AVG (Salary in Usd).
- Put Experience Level in Column and AVG (Salary in Usd) and Sum (Remote Ratio) in Row.
- Filter could be used on Job Title and Company Size.
- Put AVG (Salary in Usd) in Text and Colour.
- Put SUM (Remote Ratio) in Text and Colour.
- Click Dual Axis on Remote Ratio.
- This will help the user to find out how remote ratio and highest average salary changes when he uses the filters to change Company Size and Job Title.

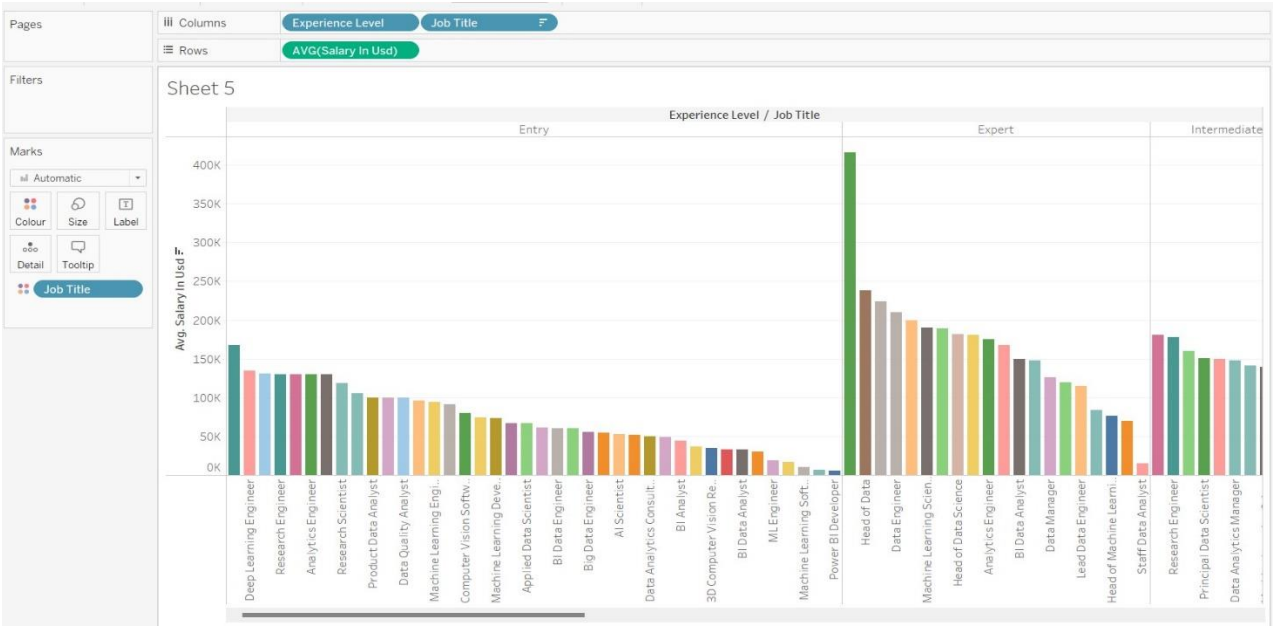
Design –

According to Deniz Oktay Tuncay (2023), To design a Tableau dashboard, understand its purpose and integration within the broader BI plan. Determine the intended recipients and data needs, and select appropriate visualizations for performance indicators and data exploration.

Sketching is crucial for creating a dashboard, as it helps position charts appropriately and aids comprehension. Organizing everything before starting is time-consuming. Sketching helps construct the narrative or flow, allowing viewers to navigate through the dashboard effectively (Thedataschool.co.uk, 2021).

The first question inquiries about, "What are the average salaries in Usd based on different data science job roles and experience levels?"

In this question, I have used the average of the salaries in Usd depending on the different jobs and experience levels. I have considered using a Horizontal bar chart by placing the Experience Level and Job Title in Columns and the average salaries in Usd in Rows. But the main problem is using the graph is cluttered for the job titles as it has many values. The below image it can be seen how the visualization look like-

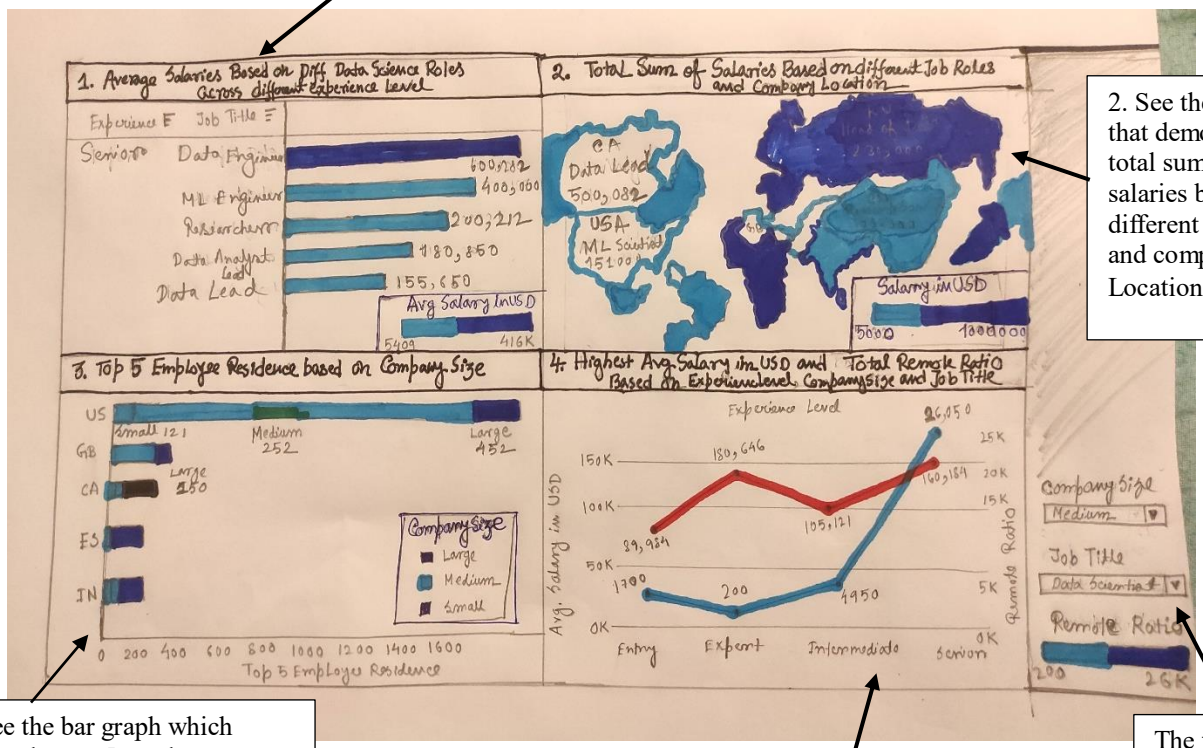


Hence the best option is to demonstrate the first question to interchange the position of Experience Level, Job Title, and AVG(Salary in Usd). That means Placing the Experience Level and Job Title in rows and AVG (Salary in Usd) in columns which would be a better visualization. Also putting down the AVG (Salary in Usd) in text and color would enhance the visualization further.

The paper landscape will show the initial idea of how the final dashboard would look like.

Paper Landscape-

1. See the bar graph which delineates the trend between Average salaries, Experience Level and Job Title(Q1)



2. See the Map that demonstrates total sum of salaries based on different job roles and company Location (Q2)

3. See the bar graph which shows the top 5 employee residence based on the company size (Q3)

4. See the two Line graphs which denotes one the remote ratio and another one is the Average Salaries (Q4)

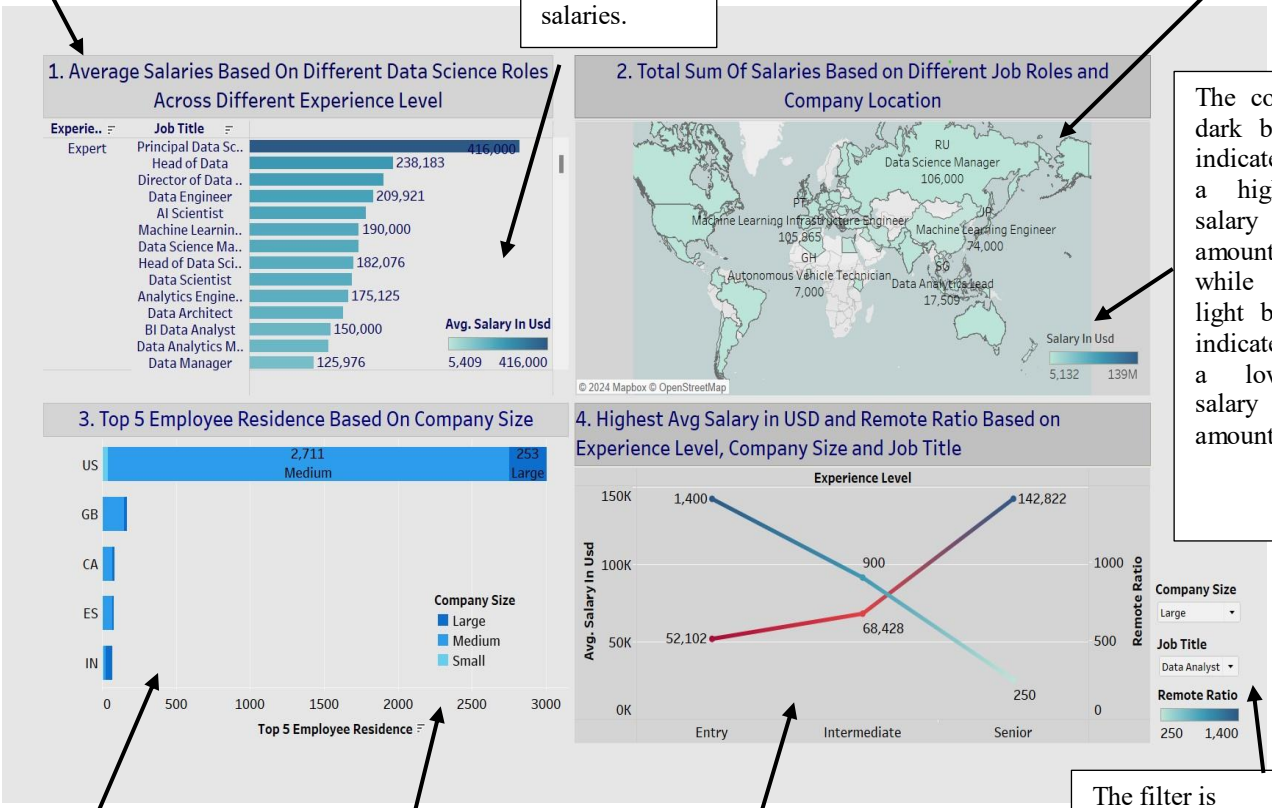
The filter is used for Q4 where a user can change the company size and job title and see the change in the line graph

Tableau Dashboard -

1. The bar Graph represents how it differs based on the different experience level and different job title (Q1)

The bar graphs' color indicates the average salary, with darker blue indicating higher salaries, and less blue indicating lower average salaries.

2. The map delineates about which country have specific job roles with the total sum of salaries for that specific job title (Q2)



The color dark blue indicates a higher salary amount, while light blue indicates a lower salary amount.

3. The Bar Graph shows the Top 5 Employee Residence based on the Company Size (Q3)

The colour distribution shows the size of the company

4. The Two Line graphs one for the Remote Ratio and another for the Average salaries (Q4)

The filter is utilized to display the differences in visualization between a company size and job title.

Analysis of the Tableau Dashboard –

There are few differences between the paper landscape and the tableau dashboard. But in one thing I faced the problem on the Tableau dashboard was the coloring of the map. In the map of the paper landscape, I have used the blue palette while the dark blue represents the region having the highest salary and the light blue represents the region having the lowest salaries. Although in the dashboard from the card, it can be seen the color range represents the amount of the salary on the map, but on the countries, there is only light blue. The main reason the map presents all the countries having different job roles and the total sum of salaries for that job role. I believe the whole visualization is good but the coloring needs some work. In the future, while creating a map I would like to highlight the specific countries that have more job roles. Also, in Graph 4, I have shown the card for the Remote

Ratio, but I was unable to show the card for the average salaries in the dashboard because there is not enough space for it in the layout. Otherwise, this dashboard will help the user mainly the HR recruiters to gain insights about the market of the data science and other jobs.

Implementation –

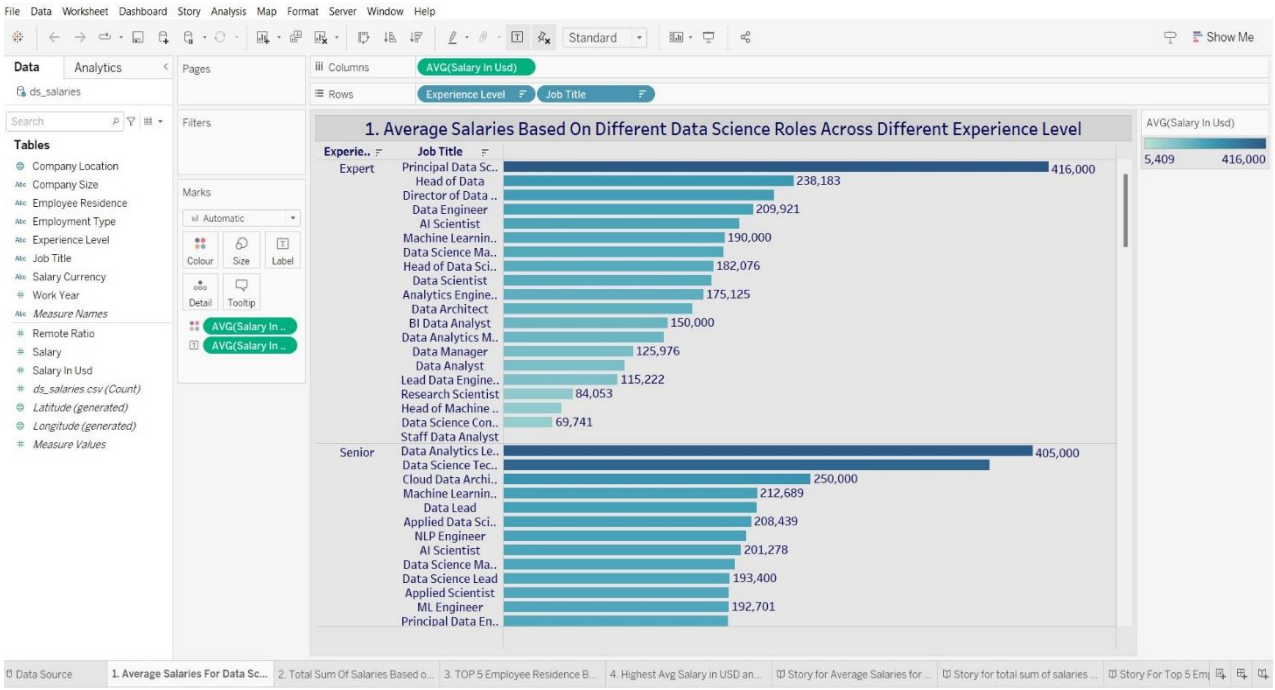
There are 4 different views and 3 different chart types

1. Two Bar Graphs

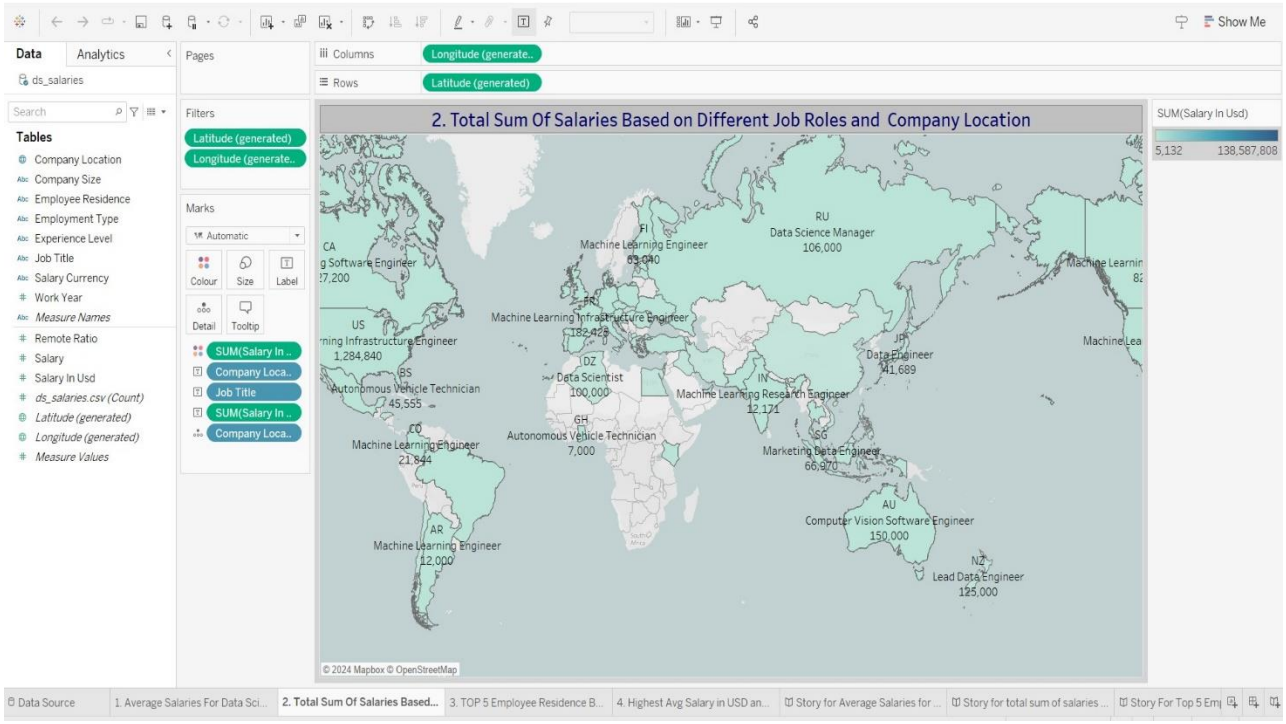
2. One Map

3. One Line Graph

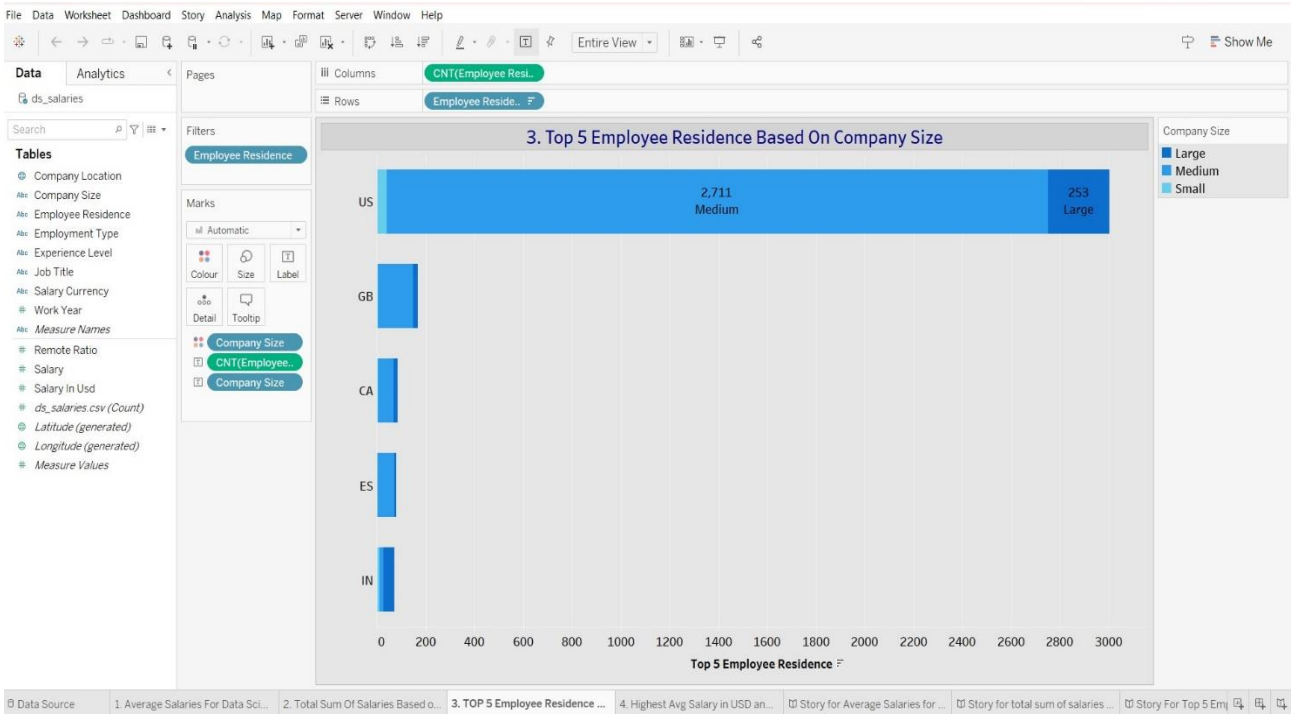
The first bar graph is created by displaying AVG (Salary in USd) in columns and Experience Level and Job Title in rows. By dragging and coloring the AVG in text, the graph's appearance is enhanced, with darker blue indicating higher salaries and lighter blue indicating lower salaries.



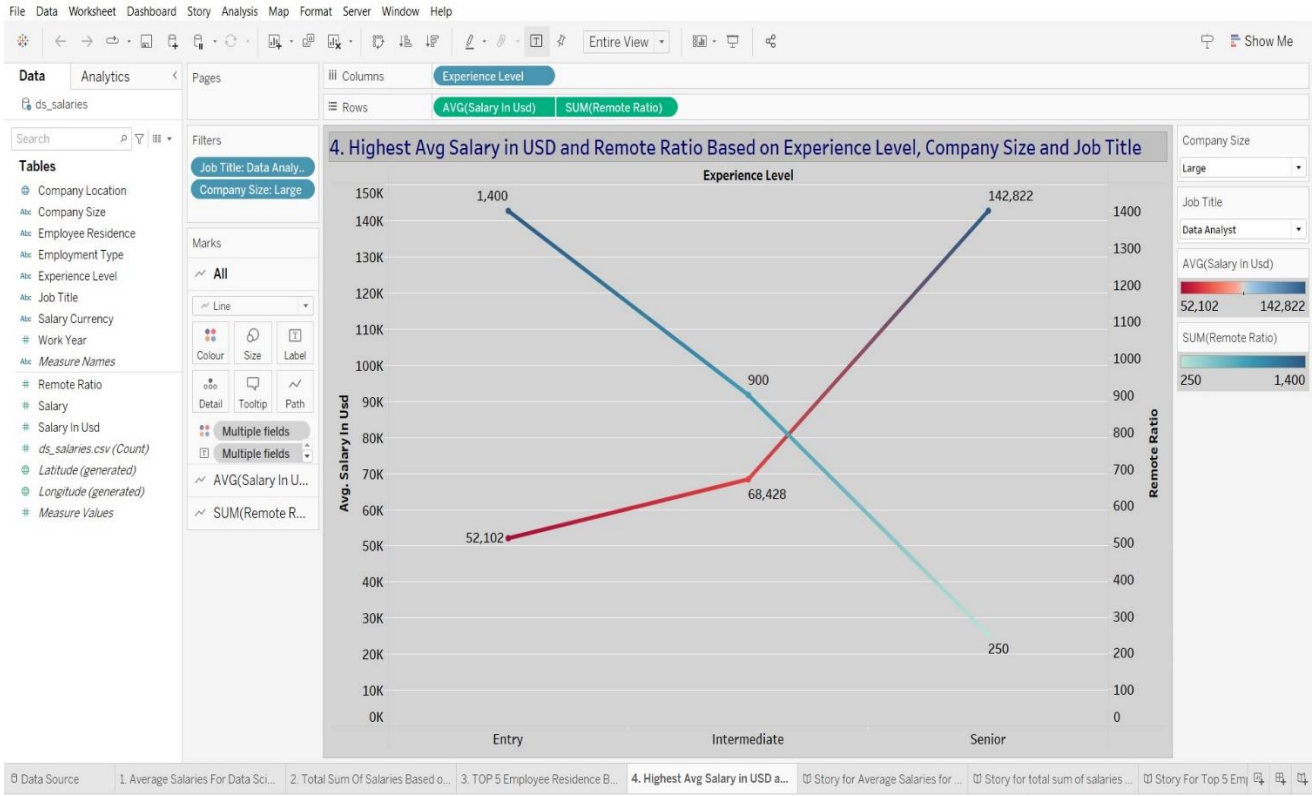
The map displays latitude and longitude in columns and rows, with the SUM (Salary in Usd) displayed in color to show salary differences based on company location and job roles. The Company Location, Job Title, and SUM are also highlighted in text for easy visualization. The color section shows the highest sum of salaries, with darker blue indicating higher salaries and lighter blue indicating decreasing total salaries.



The second bar graph displays the top 5 employee residences based on company size, with dark blue for large, medium dark for medium, and light blue for small. The count of Employee Residences (CNT) is filtered and labeled in columns, along with the company size, for easy understanding. The total employee residence and company size are labeled on the bar graph for easy comprehension.



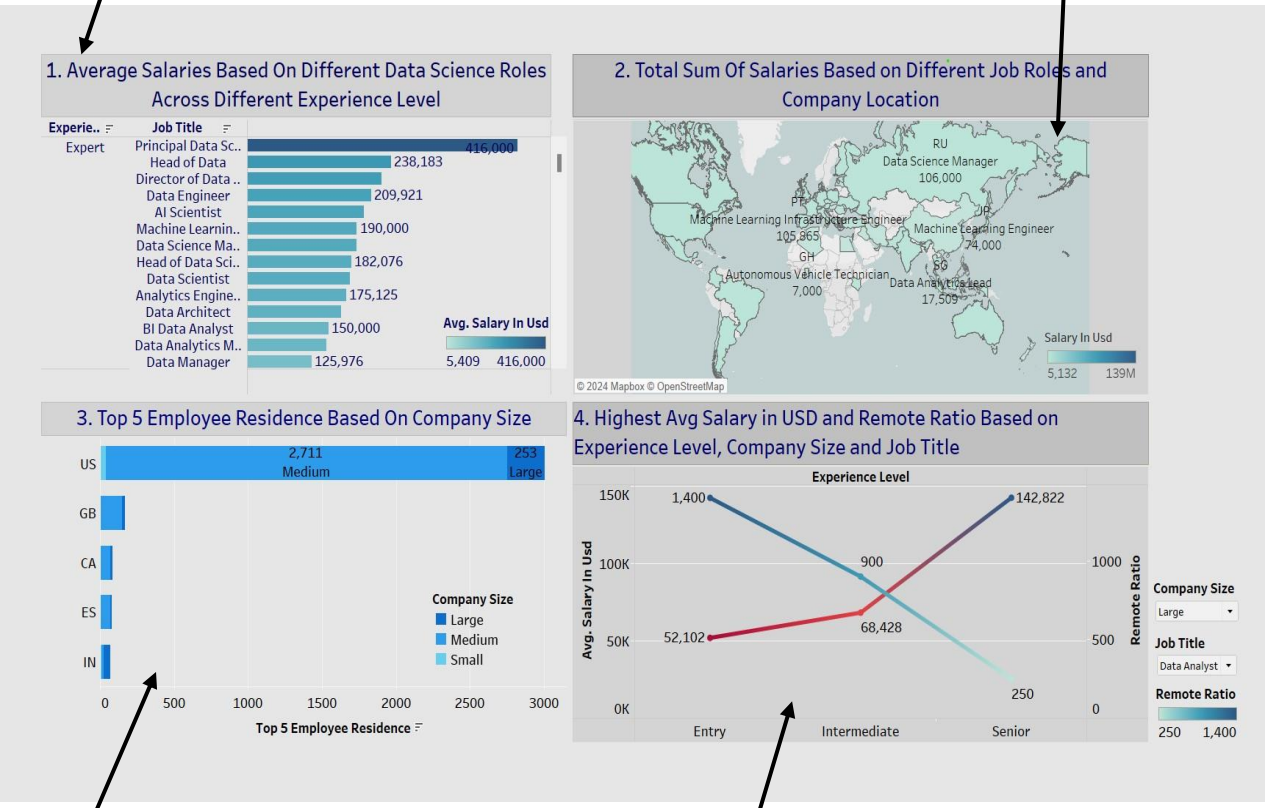
Two-line graphs are presented for the Remote Ratio and Average salaries in USD. The graphs are created by placing Experience Level in columns, AVG (Salary in USD) and SUM (Remote Ratio) on rows, and selecting the dual axis on the SUM. Average Salaries and Remote Ratio are colored and labeled, and the job title and company size are filtered. For example, a large company with a Data Analyst job title shows the highest average salary of 142,822 for Senior experience and 1400 for Entry level experience. The user can change the filter by changing the Company Size and Job Title to see differences in the line graph and shape. The AVG (Salary in USD) card shows more salaries, while the SUM (Remote Ratio) card shows darker blue and lighter blue.



The Dashboard -

1. The bar Graph shows the trend differs between experience level and different job title (Q1)

2. The map shows country have specific job roles with the total sum of salaries(Q2)



3. The Bar Graph shows the Top 5 Employee Residence based on the Company Size (Q3)

4. The Two-Line graphs one for the Remote Ratio and another for the Average salaries (Q4)

The proposed design and Tableau dashboards have similar differences, but one is the map coloration. In the proposed design, countries with the highest total salaries are presented in dark blue, while the dashboard shows the lowest salaries in lighter blue. This is due to the lack of filtering. If filtering was applied based on specific job titles and countries, the Tableau dashboard could have focused on specific regions and shown the company location with the highest salaries.

Walkthrough –

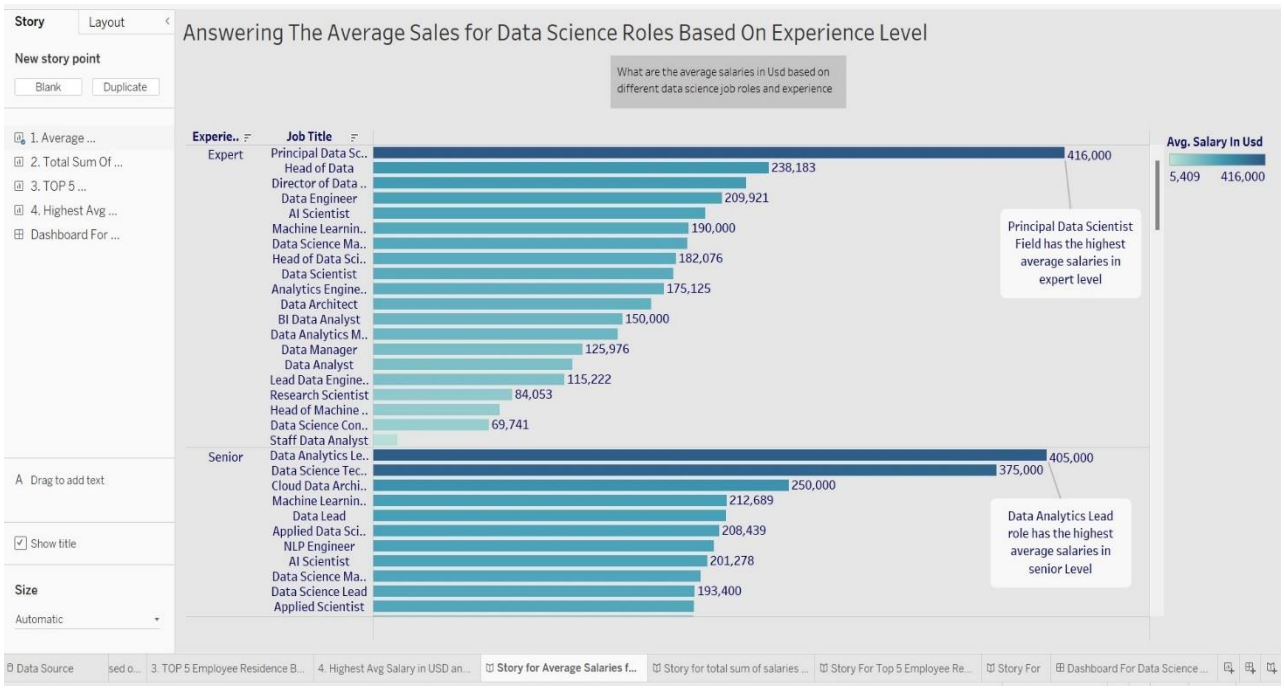
A Tableau story is a data analysis tool that uses narrative points from a Tableau worksheet to showcase important insights through logically connected visuals (Tableau.com, 2024b).

Using text annotations and filters, the narrative function creates connected points with viewpoints, control panels, or written content to guarantee a coherent story (Tableau.com, 2024a).

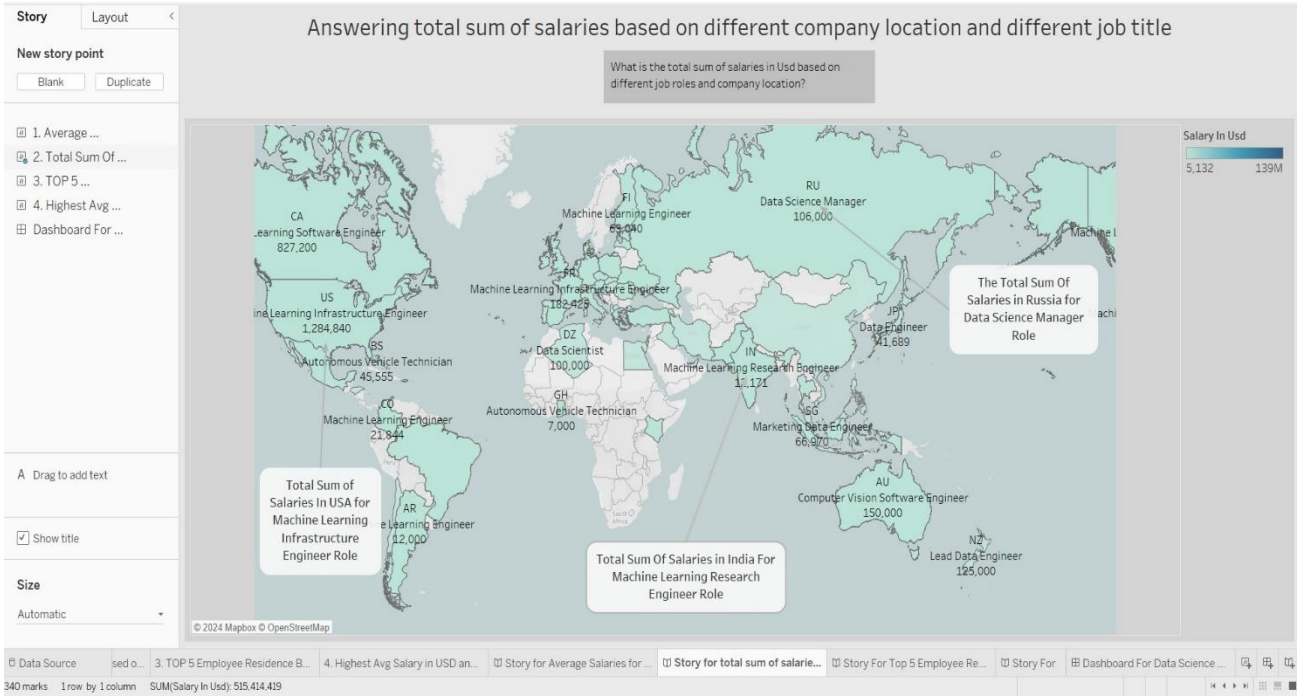
These stories are meant to set the scene and show how decisions have an impact on outcomes, strengthening the case you're making (McKay, 2023).

The HR recruiter has four questions. The first question has been asked, **"What are the average salaries in Usd based on different data science job roles and experience levels?"**

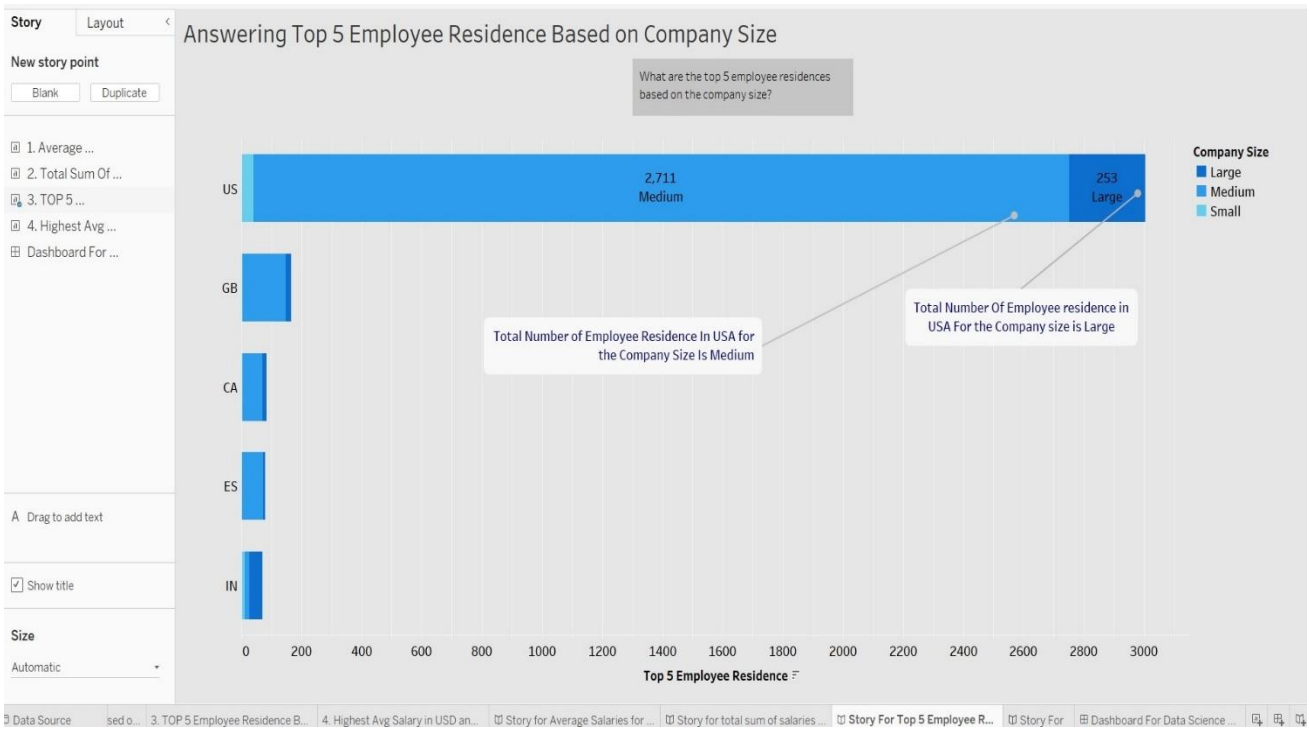
Here is a screenshot of the story for the first question that has been provided. The question can be answered using the horizontal bar graph. The Experience Level and the Job Title are in Rows and the Average Salaries in Usd are in Columns. **One thing the user would notice here the darkest blue indicates the average salary is more than 400,000 and the lightest blue indicates the salary is nearly 5,500.** For the Experience level expert, the Principal Data Scientist has the highest average salary, which is presented in the darkest blue color. **The Lighter the blue** means the salary is getting less. The bar graph is in descending order. The Staff Data Analyst has the lowest average salary for the expert level. Similarly, At the Senior level, the dark blue color indicates that the job position Data Analytics Lead has the highest average salary followed by Data Science Tech Lead.



The second question of the HR Recruiter is, **"What is the total sum of salaries in Usd based on different job roles and company locations?**" It can be answered using the Map. The Map denotes the different countries where the company is located with the specific job title and the total sum of salaries for that job role. Here from the Map, we can see that in Russia the major job role is Data Science Manager with a total sum of salaries 106,000 USD. Similarly, For US has a Machine Learning Infrastructure Engineer as a major Job Role with a total sum of salaries of about 1,284,840 USD.



The Third Question has been asked by the HR Recruiter is, **"What are the top 5 employee residences based on the company size?"** This question also can be answered by using a horizontal bar chart. The company sizes in the color box denote the color changes depending on the company size. The Employee residences have been filtered by taking the top 5. From the below graph, we can see that the US is the country which came at the top where for the Large size company number of employees is 253 and for the medium-sized company the number of employees is 2711.



The last question from the HR recruiter is, **"What are the highest average salary and total remote ratio based on different experience levels, company size, and different job titles?"**

The Line graphs are perfect for this type of question. The Average Salary in Usd and SUM (Remote Ratio) has been placed in Rows while Experience Level has been placed in Columns. While Filter has been applied to Job Title and Company Size. The dual axis has been applied on Remote Ratio. From the below visualization, we can see that for the Senior Level, The Highest average salary is 159,267 Usd and the Remote Ratio for the Intermediate experience level is 1800 when the Company Size is Large for the Data Engineer Job Roles. The user can change the Company Size and Job Title and see the differences. Also, for the Average salary in Usd the highest salary is indicated by Dark blue color while the lowest one is indicated in Red color. In the case, of Remote Ratio the highest one is indicated by dark blue while lighter blue indicates the lowest Remote Ratio.

Discussion –

The Dashboard is clear and the visualization for every question is answered so the HR recruiter can easily understand how the salary differs based on the different regions, job titles, job types, and experience level. One thing I have found in my Tableau Dashboard is that while making the map the coloration problem is the hindrance I have faced. What I have made in the paper landscape, as per that design only the coloration of the map is the biggest challenge and has made a difference. Maybe it could have been solved if the dataset contained more data and used the filtering so It could have focused on specific regions according to Job Titles.

Overall the design of the dashboard is good, and gives the answers perfectly what the HR recruiter asked.

Conclusion –

The project aims to educate HR recruiters on salary differences based on job title, experience level, and company location using various charts. Key outcomes include designing and implementing graphics techniques that work with various data sources, resulting in useful insights and smart decisions. Critical reflection on the usefulness of visualization methods ensures clear, relevant, and actionable results.

References –

Deniz Oktay Tuncay (2023). *10 Things You Need to Know About Tableau Before Designing a Dashboard*. [online] Medium. Available at: https://medium.com/@deniz.oktay.tuncay/10-things-you-need-to-know-about-tableau-before-designing-a-dashboard-12a567ca268e [Accessed 15 Apr. 2024].

McKay, S. (2023). *What is a Tableau Story? A Complete Guide | Master Data Skills + AI*. [online] Master Data Skills + AI | Insights and Strategies from the Enterprise DNA Blog. Available at: https://blog.enterprisedna.co/what-is-a-tableau-story/ [Accessed 15 Apr. 2024].

Tableau.com. (2024a). *Best Practices for Telling Great Stories*. [online] Available at: https://help.tableau.com/current/pro/desktop/en-us/story_best_practices.htm [Accessed 15 Apr. 2024].

Tableau.com. (2024b). *Stories*. [online] Available at: https://help.tableau.com/current/pro/desktop/en-us/stories.htm [Accessed 15 Apr. 2024].

Thedataschool.co.uk. (2021). *The Data School - 5 things to do before creating your dashboard in Tableau*. [online] Available at: https://www.thedataschool.co.uk/rahul-lalwani/5-things-to-do-before-creating-your-dashboard-in-tableau/ [Accessed 15 Apr. 2024].