Project name: **Employee Trend Analysis -Tableau project**

The CSV file ds\_salaries.csv contains salary data for various data science jobs from 2020 to 2022. The main goal of the analysis is to understand how salaries vary based on factors like **experience level, job title, company location, and time**.

**📊 Data Analysis and Problem Statement**

**🔑 Key Findings**

The analysis of the dataset reveals several key trends and disparities in the data science job market:

| **Category** | **Finding** | **Average Salary (USD)** |
| --- | --- | --- |
| **Experience Level** | There is a clear, positive correlation between experience and salary. | **Executive-level:** $199,392.04 |
|  |  | **Senior-level:** $138,617.29 |
|  |  | **Mid-level:** $87,996.06 |
|  |  | **Entry-level:** $61,643.32 |
| **Yearly Trend** | The average salary for data science roles has increased significantly year-over-year. | **2022:** $124,522.01 |
|  |  | **2021:** $99,853.79 |
|  |  | **2020:** $95,813.00 |
| **Top-Paying Locations** | The **United States (US)** has the highest volume of high-paying jobs, though a few other countries show a higher average due to a small number of top-tier salaries. | **US:** $144,055.26 |
| **Top-Paying Job Titles** | Roles with strategic and leadership responsibility command the highest salaries. | **Principal Data Scientist:** $215,242.43 |
|  |  | **Director of Data Science:** $195,074.00 |

**⚠️ Problem Statement**

The core problem, or opportunity for optimization, is understanding and addressing the **wage disparities** across different employment factors. Specifically:

* **Global Salary Inequity:** While US-based roles dominate the high-paying jobs, the analysis should explore if the same job titles and experience levels are underpaid in other company locations.
* **Career Path and Compensation:** There's a significant salary jump from Mid-level to Senior-level (over $50,000 difference), indicating that moving up the experience ladder is the most critical factor for a raise. The question is whether specific job titles offer a faster or more lucrative path to the Senior/Executive level.
* **Remote Work Impact:** The effect of the remote\_ratio (0, 50, 100) on salary needs to be quantified to understand if companies with a higher remote presence offer competitive compensation.

**💡 Solutions and Recommendations**

Based on the analysis, here are solutions and recommendations for various stakeholders:

**For Individuals (Data Professionals)**

1. **Prioritize Experience over Lateral Moves:** Focus on gaining the experience needed for a **Senior-level** role (SE), as the average salary difference from Mid-level is the most substantial jump in the dataset.
2. **Target High-Value Roles:** Aim for strategic titles like **Principal/Director of Data Science** or **Data Architect** for maximum earning potential.
3. **Explore US-Based Remote Roles:** The **US** dominates high salaries. Professionals outside the US should target US-based companies offering **fully remote (100% remote)** positions to maximize their salary potential without relocating.

**For Companies (HR/Management)**

1. **Standardize Global Pay Bands:** Companies operating in multiple locations should conduct a **Cost of Labor** analysis to standardize salary bands for the same job and experience level, ensuring competitive and fair pay, especially for remote workers in high-cost-of-living areas.
2. **Define Clear Career Ladders:** Clearly define the responsibilities and required skills for the Mid-level to Senior-level transition to justify the significant salary increase and motivate employees.
3. **Annual Salary Review:** Given the clear **yearly salary increase trend** in the market, companies must perform annual salary adjustments to remain competitive and retain talent.

**🗃️ SQL Queries (ds\_salaries.sql)**

The following SQL queries explore the identified problem areas.

* Compare the average senior-level (SE) salary in the US vs. the rest of the world.
* Find the Top 3 highest-paying job titles for Mid-level (MI) employees.
* Analyse the salary difference between 100% remote and 0% remote jobs overall.
* Calculate the Year-over-Year (YoY) salary growth rate.
* Identify the most common job title for each company size (S, M, L).
* Find the highest paying job in each experience level for 2022.
* Count the number of unique job titles for each company size.

The output of the SQL queries are provided into a file named ds\_salaries.sql.

**📊 Tableau and Presentation**

**Tableau File** The detailed analysis and visualization are best achieved using Tableau or a similar BI tool. The following file, which is a CSV of the cleaned and analyzed data, should be used to create the visualizations. This file, **ds\_salaries\_clean.csv**, includes the more readable **experience\_level\_full** column.

**Final Presentation Outline** I've outlined a 5-slide presentation focusing on the most critical insights from the analysis.

| **Slide #** | **Title** | **Content Focus** | **Key Visualization/Data Point** |
| --- | --- | --- | --- |
| **1** | **Executive Summary: The State of Data Science Salaries (2020-2022)** | **Highest-level summary of the salary landscape.** | **Average Salary Trend (2020-2022): Show the strong YoY increase.** |
| **2** | **The $50K Bridge: Experience is the Key Multiplier** | **Analysis of salary by experience level, highlighting the jump from Mid-level to Senior-level.** | **Bar Chart: Average Salary by experience\_level\_full.** |
| **3** | **Global Compensation Disparity: The US Factor** | **Comparison of US salaries vs. the rest of the world, especially for Senior-level roles, and the impact of company location.** | **Map/Bar Chart: Average Salary by company\_location.** |
| **4** | **Targeting the Top: High-Value Job Roles** | **Deep dive into the highest-paying job titles for Mid-level and Senior-level employees.** | **Bar Chart: Top 5 job\_title by average salary.** |
| **5** | **Recommendations & Strategic Takeaways** | **Actionable advice for both employees and employers based on the analysis.** | **Bullet Points: The main recommendations (e.g., Prioritize Senior-level, Standardize Global Pay).** |

**📊 Tableau Storyboard Instructions**

The storyboard will consist of **5 main Story Points**, corresponding to the slides in the final presentation.

**1. Data Preparation in Tableau**

1. **Connect to Data:** Open Tableau and connect to the **ds\_salaries\_clean.csv** file.
2. **Verify Data Types:** Ensure the following data types are correct:
   * work\_year: **Number (Whole)** or **Date** (if converted for time series).
   * salary\_in\_usd: **Number (Decimal)**.
   * remote\_ratio: **Number (Whole)**.
   * experience\_level\_full, job\_title, company\_location, etc.: **String**.

**2. Story Point 1: Executive Summary - The Market Trend**

**Goal:** Show the significant Year-over-Year (YoY) salary growth.

| **Step** | **Action** | **Tableau Fields & Settings** |
| --- | --- | --- |
| **1. Create Sheet** | Name the sheet: **"YoY Salary Growth"** |  |
| **2. Add Time** | Drag work\_year to the **Columns** shelf. | Convert to **Discrete** if it's not already. |
| **3. Add Salary** | Drag salary\_in\_usd to the **Rows** shelf. | Ensure the measure is set to **Average**. |
| **4. Chart Type** | Set the **Marks** type to **Line**. |  |
| **5. Formatting** | Add labels to the line (Average Salary). Set the **Title** to: **"Average Data Science Salary ($$$USD) by Year"**. |  |

**3. Story Point 2: The $50K Bridge - Experience is the Key Multiplier**

**Goal:** Highlight the clear progression and the large jump in salary between Mid-level and Senior-level.

| **Step** | **Action** | **Tableau Fields & Settings** |
| --- | --- | --- |
| **1. Create Sheet** | Name the sheet: **"Salary by Experience Level"** |  |
| **2. Add Experience** | Drag experience\_level\_full to the **Columns** shelf. | Ensure the order is logical (Entry-level $\rightarrow$ Executive-level). |
| **3. Add Salary** | Drag salary\_in\_usd to the **Rows** shelf. | Ensure the measure is set to **Average**. |
| **4. Chart Type** | Set the **Marks** type to **Bar**. |  |
| **5. Formatting** | Sort the bars by Average Salary (Descending). Add labels for the average salary on each bar. Set the **Title** to: **"Average Salary by Experience Level"**. |  |

**4. Story Point 3: Global Compensation Disparity**

**Goal:** Show where the highest-paying companies are located, emphasizing the US's dominance.

| **Step** | **Action** | **Tableau Fields & Settings** |
| --- | --- | --- |
| **1. Create Sheet** | Name the sheet: **"Avg Salary by Company Location"** |  |
| **2. Create Map** | Drag company\_location to the **Detail** or **Color** shelf. Tableau should automatically recognize it as a geographic field and switch to a map view. |  |
| **3. Color by Salary** | Drag salary\_in\_usd to the **Color** shelf. | Ensure the measure is set to **Average**. This will color the countries based on average salary. |
| **4. Add Details** | Drag company\_location and Average salary\_in\_usd to the **Tooltip** to show the exact value on hover. |  |
| **5. Filtering (Optional)** | *For a cleaner view, you may filter out countries with fewer than 5 job entries.* |  |

**5. Story Point 4: Targeting the Top - High-Value Job Roles**

**Goal:** Identify the specific job titles that offer the highest compensation.

| **Step** | **Action** | **Tableau Fields & Settings** |
| --- | --- | --- |
| **1. Create Sheet** | Name the sheet: **"Top 10 High-Value Jobs"** |  |
| **2. Add Job Title** | Drag job\_title to the **Rows** shelf. |  |
| **3. Add Salary** | Drag salary\_in\_usd to the **Columns** shelf. | Ensure the measure is set to **Average**. |
| **4. Sorting & Filtering** | **Sort** the job\_title by the Average salary\_in\_usd (Descending). **Filter** the job\_title to show only the **Top 10** based on the Average salary\_in\_usd. |  |
| **5. Formatting** | Set the **Title** to: **"Top 10 Highest Paying Data Science Job Titles"**. |  |

**6. Creating the Final Storyboard**

1. Click the **New Story** icon ($\text{\textcircled{S}}$).
2. **Add Story Points:** Drag the four sheets created above onto the story area in the following order:
   1. **YoY Salary Growth**
   2. **Salary by Experience Level**
   3. **Avg Salary by Company Location**
   4. **Top 10 High-Value Jobs**
3. **Create Final Summary Point (Text):**
   1. Create a new blank **Dashboard** named **"Recommendations"**.
   2. Add a **Text Object** to the dashboard.
   3. Type the key recommendations from the analysis (e.g., *Target Senior Level*, *Explore US-Based Remote Roles*, *Implement Standardized Global Pay*).
   4. Drag this **"Recommendations"** dashboard as the **5th Story Point**.
4. **Final Polish:** Add descriptive captions for each story point to summarize the key insight.