**Data Analysis Project: Salary Analysis in Data Science**

**Project Description:** In this project, you will analyze a dataset containing information on salaries in the data science field. The dataset includes details such as work experience, job titles, employment types, locations, and remote working conditions. Your objective is to explore salary trends, understand factors that influence salary in the data science industry, and provide insights on how location, experience, and company size affect compensation.

You will use **Python** to clean, analyze, and visualize the data. Libraries like **pandas**, **NumPy**, **Matplotlib**, and **Seaborn** should be used for efficient data handling and visualization.

**Dataset Overview:** The dataset contains the following columns:

* **work\_year**: The year in which the salary was recorded.
* **experience\_level**: The level of experience of the employee (e.g., Junior, Mid, Senior).
* **employment\_type**: Type of employment (e.g., Full-time, Contract).
* **job\_title**: The specific job title (e.g., Data Scientist, Machine Learning Engineer).
* **salary**: The salary amount paid to the employee.
* **salary\_currency**: The currency in which the salary is paid.
* **salary\_in\_usd**: The salary converted to USD.
* **employee\_residence**: The location where the employee resides.
* **remote\_ratio**: The percentage of remote work (0 = no remote, 50 = hybrid, 100 = fully remote).
* **company\_location**: The location of the company.
* **company\_size**: The size of the company (Small, Medium, Large).

**Project Objectives:**

1. Understand salary distributions in the data science field.
2. Explore how experience, job titles, and employment types influence salaries.
3. Analyze the impact of remote work on salary.
4. Investigate geographic differences in salaries, including employee residence vs. company location.
5. Provide insights on how company size affects salary.

**Analysis Tasks & Questions:**

1. **Descriptive Analysis:**
   * What is the distribution of salaries in the data science field? (Use histograms or box plots)
   * What is the average salary for different experience levels (e.g., Junior, Mid, Senior)?
   * How do salaries vary between different job titles (e.g., Data Scientist, Data Engineer)?
2. **Salary vs. Experience and Employment Type:**
   * Is there a clear trend in salary increases with more years of experience?
   * How do salaries differ between full-time, part-time, contract, and internship employment types?
3. **Geographic Insights:**
   * How do salaries differ based on employee residence vs. company location? (Compare using bar charts or box plots)
   * Do certain countries or regions offer higher salaries for data science roles?
   * How does the remote ratio impact salary? Are fully remote jobs compensated differently from hybrid or on-site jobs?
4. **Remote Work Analysis:**
   * What is the distribution of remote work (0%, 50%, 100%) in the data science field?
   * Does working remotely (partial or full-time) affect salary levels?
5. **Company Size and Salary:**
   * How does company size (small, medium, large) influence salaries in the data science field?
   * Are larger companies paying higher salaries than smaller companies?
6. **Recommendations:**
   * Based on your analysis, provide insights for data science professionals on what factors can lead to higher salaries (e.g., geographic location, company size, job title).
   * Suggest strategies for negotiating better compensation, including the impact of experience, employment type, and remote work.

**Deliverables:**

* A **Python notebook** with:
  + Data cleaning and preprocessing steps.
  + Descriptive and exploratory analysis (with appropriate visualizations).
  + Insights and conclusions based on your analysis.
* A final report or presentation summarizing your findings and actionable recommendations for professionals in the data science field.