Analysis of Data Analyst and Data Scientist Salary Trends Across Global Markets

Project Objectives:

The primary goal of this project is to analyze salary trends for data-related roles across different companies, countries, and rating levels, with a focus on identifying patterns and deriving insights that can help professional data make informed career decisions and aid companies in competitive compensation strategies.

Key Project Questions

1. Salary Distribution Analysis

 - What is the overall distribution of salaries (median, low, high) for data-related roles?

  - Objective: To understand the range and central tendency of data-related salaries, which can inform you about the overall market pay scale for data professionals.

 - Are there outliers or exceptionally high/low salary figures, and what factors might contribute to these anomalies?

  - Objective: To identify if any role, company, or country has unusually high or low salaries and explore potential reasons (e.g., company reputation, demand for specific roles).

2. Role-Specific Salary Comparison

 - How do salaries differ between the `Data Analyst` and `Data Scientist` roles?

  - Objective: To evaluate how job roles impact salary, which can help professionals understand the potential salary benefits of transitioning from one role to another.

 - Within each role, is there a significant difference between entry-level, mid-level, and senior-level positions?

  - Objective: To determine if experience levels or job seniority significantly impact compensation, which can aid in benchmarking expected salaries at various career stages.

3. Country-Based Salary Patterns

   - What are the salary ranges for data roles across different countries, and which countries offer the highest/lowest salaries?

   - Objective: To identify geographic trends in salaries, helping data professionals understand where compensation is most competitive.

   - Is there a correlation between the `company_rating` and `country` for data-related salaries?

   - Objective: To explore if countries with generally high ratings also offer competitive salaries, or if this trend varies.

4. Company-Level Salary Insights

   - How do salaries vary across different companies?

   - Objective: To analyze company-level salary data to identify which companies are market leaders in compensation.

   - What is the impact of `company_rating` on salary for data-related roles?

   - Objective: To assess if higher-rated companies offer higher salaries, which can be valuable for professionals prioritizing workplace culture alongside compensation.

5. Unit-Based Salary Trends

   - What percentage of salaries are offered `per year`, `per month`, and `per hour`, and how do these impact overall compensation?

   - Objective: To explore how different salary units affect total compensation and assess if certain countries or roles tend to favor particular units.

   - Are certain salary units (e.g., `per hour`) associated with higher or lower ratings?

   - Objective: To evaluate if companies with specific pay structures (such as hourly pay) differ significantly in terms of company rating.

6. Correlation and Predictive Analysis

   - Is there a significant correlation between `company_rating` and any salary metric (e.g., `median`, `high`, `low`)?

     - Objective: To assess if better-rated companies tend to pay more, which can inform job seekers about trade-offs between rating and compensation.

   - Can we predict missing `company_rating` values based on salary metrics or `country`?

     - Objective: To use regression or imputation techniques for filling in missing `company_rating` values, enhancing data completeness and reliability.

   - Can we build a model to predict the `high` salary based on other metrics, such as `median` and `low`?

     - Objective: To create a salary predictor that could assist professionals in estimating potential earnings based on available data.

7. Salary Trends Over Time (If Time-Based Data is Available)

   - How have data-related salaries changed over time in different countries and for different roles?

     - Objective: To identify if there are any upward or downward trends in salaries, helping to forecast future compensation trends.

   - Are certain countries experiencing faster growth in data-related salaries compared to others?

     - Objective: To understand geographic trends in salary growth, offering insights for professionals considering relocation or companies setting global compensation strategies.

 Expected Deliverables

1. Descriptive Report: Summary of findings on overall salary distribution, role-specific insights, geographic patterns, and company-level comparisons.

2. Data Visualizations:

   - Salary distribution histograms

- Boxplots for salary comparisons across roles, countries, and companies

- Scatter plots to show correlation between `company_rating` and salary metrics

3. Predictive Models:

- Imputation model for `company_rating` where data is missing

- Salary prediction model for estimating `high` salary values based on other predictors

4. Presentation: A slide deck summarizing key insights for stakeholders, highlighting salary trends, company-level analysis, and country-specific observations.

5. Executive Summary: A concise, high-level summary of findings and recommendations for data professionals and hiring managers.

Conclusion

This analysis will provide a comprehensive view of data-related salaries across different roles, companies, and countries, helping stakeholders make informed decisions on compensation, hiring strategies, and career planning.