

TABLE OF CONTENTS

1

Project Overview

Key User Attributes: Work Year, Experience Level, Employment Type, Job Title, Salary, Salary Currency, Salary in USD, Employee Residence, Remote Ratio, Company Location, Company Size

2

Libraries and Data Handling

Libraries Used: Pandas, Seaborn, NumPy, Matplotlib
Data Loading and Preprocessing: Loading from CSV, Handling Missing Values, Lowercasing Column Names, One-Hot Encoding, Binning

3

Data Analysis Technique

Exploratory Data Analysis: Grouping and Aggregation, Column Renaming, Unstacking, Data Summarization, Groupby Aggregation
Data Visualization: Histogram, Bar Plots, Pair Plots, FacetGrid with Barplot, Heatmap, Line Plots

4

Key Findings

Average Salary: Average Salary based on multiple attributes

5

Advanced Analysis

Geographical Analysis: Categorization into Regions, and Regional Analysis
Temporal Trends: Average Salary Trends Over the Years by Region

6

Visual Insights

Salary Distribution in Relation to Various Attributes

7

Conclusion

Summary of insights derived , implications for future strategic decisions

8

Appendix

Code Snippets: Provided Python code used for loading, cleaning, transforming data, and generating visualizations.

Datasets: Sample dataset of Data Science Job Salaries

Github Website Link

<https://nickcore18.github.io/CSEL302-Finals/>