U.S. Software Developer Salaries

Description

Data Quantity

The dataset, encapsulating information on software developer salaries, is formatted in a CSV file. This dataset is the result of the processing and aggregation of various datasets, each providing different perspectives and insights. The database comprises 12 columns that represent various aspects of income and expenditure and 77 rows, each detailing the data for a specific city.

Data Quality

The dataset is characterized by a high degree of relevance to the business questions at hand, including critical attributes such as purchasing power, rent, and house prices, which are pivotal in analyzing the economic landscape for software developers. The dataset is composed of two primary data types: symbolic and numeric. The symbolic data represents city names and their respective states, providing a geographical context to the analysis. The numeric data indicates critical financial metrics such as salaries and prices, offering quantitative insight to evaluate the industry.

Upon computing basic statistics for key attributes, I noticed that a significant insight emerged regarding the rent-to-salary ratio across different career paths. It was observed that the rent constitutes approximately 21.4% of the total salary for a software developer, which contrasts with the 39.5% observed for average salaries across all careers. It underscored a notable discrepancy in the salary structures between software engineers and other professions, highlighting the financial advantages associated with a career in software development.

Usage

salary_data <- read_csv("SweSalaries.csv")

Format

Metro	The metropolitan area of the city. (String)
Mean Software Developer Salary (adjusted)	The average salary for software developers adjusted for cost of living differences between cities. (Number)
Mean Software Developer Salary (unadjusted)	The average salary for software developers without adjusting for cost of living differences between cities. (Number)
Mean Unadjusted Salary (all occupations)	The average salary for all occupations without adjusting for cost of living differences between cities. (Number)
Number of Software Developer Jobs	The number of software developer jobs in the city. (Number)
Median Home Price	The median home price in the city. (Number)
City	The name of the city. (String)
Cost of Living avg	The average cost of living in the city. (Number)
Rent avg	The average rent in the city. (Number)
Cost of Living Plus Rent avg	The average cost of living plus rent in the city. (Number)
Local Purchasing Power avg	The average local purchasing power in the city. (Number)

Source

Kaggle. "U.S. Software Developer Salaries," February 11, 2023.
https://www.kaggle.com/datasets/thedevastator/u-s-software-developer-salaries.

Seveso, Bruno Bonfrisco Franco. "Sofware Developer Income/Expenses per USA City."
Zenodo, December 7, 2022. https://doi.org/10.5281/zenodo.7412091.

Example

```
# Read the data from CSV file

salary_data <- read_csv("SweSalaries.csv")


# Calculate the average rent ratio for SWE careers and all careers

average_ratios <- salary_data %>%

  summarize(

    Avg_Rent_Ratio_SWE = mean(Rent * 12/ MeanSalaryAdjusted, na.rm = TRUE),

    Avg_Rent_Ratio_All_Careers = mean(Rent * 12/ MeanSalary, na.rm = TRUE)

  )

print(average_ratios)
```