Data Science Salaries Analysis Project

Project Overview

In this project, you will use SQL to analyze a dataset containing information about salaries in the data science field. This will help you understand the salary trends, variations across regions, experience levels, and company sizes. By the end of the project, you will have a solid grasp of SQL query construction and be able to derive meaningful insights from real-world data.

Objectives

- Understand how to retrieve, filter, and aggregate data using SQL.
- Explore salary trends and patterns in the data science industry.
- Practice SQL join operations, subqueries, and advanced aggregation techniques.
- Develop insights that can inform career decisions or business strategies.

Dataset Description

The dataset contains the following columns:

- **work_year**: The year of the salary data.
- **experience_level**: The experience level of the employee (e.g., Junior, Mid, Senior, Executive).
- **employment_type**: The type of employment (e.g., Full-time, Part-time).
- **job_title**: The job title (e.g., Data Scientist, BI Data Analyst).
- **salary**: The salary amount in the local currency.
- **salary_currency**: The currency of the salary.
- **salary_in_usd**: The salary converted to USD.
- **employee_residence**: The country where the employee resides.
- **remote_ratio**: The percentage of remote work.
- **company_location**: The country where the company is located.
- **company_size**: The size of the company (Small, Medium, Large).

Tasks

Task 1: Basic Queries

1. Write a query to retrieve the job titles and salaries in USD from the dataset.
2. Write a query to filter and display data for the year 2022, for employees at the Senior experience level, working full-time.

Task 2: Aggregation and Grouping

1. Calculate the average salary in USD for each job title.
2. Group the data by company location and calculate the average salary in each region.
3. Count the total number of employees working remotely vs. on-site.

Task 3: Join Operations (Optional)

1. If provided with an additional table of currency exchange rates, write a query to join this table with the salaries dataset to analyze trends in local vs. USD salaries.

Task 4: Advanced Analysis

1. Identify the top 5 job titles with the highest average salaries.
2. Analyze salary trends over the years.
3. Examine the relationship between company size and average salary.

Task 5: Subqueries

1. Write a query to find the top 5 highest-paying jobs in the dataset.
2. Write a query to find the most common job titles in different regions.