AI Job Market Analysis Project – Summary Report

Author: Sidhant Kumar
Project Duration: Day 1–6
Tools Used: Excel, SQLite, Power BI, Python (Colab), Pandas, Matplotlib, Seaborn

Overview

This project presents a comprehensive analysis of the **global AI job market**, combining data cleaning, database design, visualization, and predictive analysis. The workflow was designed to simulate real-world business intelligence and data science practices — starting from raw data handling and ending with insight-driven dashboards.

The dataset includes **job listings** with details like titles, required skills, experience level, company location, remote ratio, salary (USD), and industry type.
The goal was to uncover **salary patterns, job demand trends, and remote work insights** within the AI and tech industry.

Key Objectives

1. **Data Cleaning & Preparation** – Handle duplicates, missing values, inconsistent text formats, and unify columns across two raw datasets.
2. **Database Integration** – Create relational tables in SQLite, define keys, and join datasets into a clean consolidated table (`job_analysis`).
3. **Visualization** – Build an interactive Power BI dashboard with KPIs, maps, and slicers to explore global AI job trends.
4. **Advanced Analysis** – Use Python for statistical exploration and visualization (salary distributions, industry averages, correlation analysis).

Key Insights

- **Highest Paying Roles:** AI Research Scientist, ML Engineer, and Data Architect dominate salary brackets.
- **Top Locations:** The United States, Germany, and the UK lead in average AI salaries.
- **Remote Work:** Remote and hybrid roles are rapidly increasing, often offering salaries comparable to on-site jobs.
- **Industry Trends:** Tech and finance industries lead in hiring AI professionals with competitive pay.

Outcomes

1. Built a **complete data pipeline** from raw files to insights.
2. Created **interactive dashboards** and **Python-based analytics**.
3. Delivered **business-ready KPIs** and visual storytelling.
4. Enhanced personal portfolio with a **multi-tool project** showcasing both engineering and analytical thinking.

Future Improvements

- Add automated ETL pipelines using Python.
- Connect Power BI to a live SQL database for real-time updates.
- Integrate predictive models to estimate salaries for new job listings.

This project demonstrates full-stack data capability — **Data Engineering + Business Intelligence + Data Science**.