Fortune 100 Companies

Business & Revenue Analysis

An end-to-end analytics project examining America's largest companies through web scraping, SQL analysis, and interactive dashboards.

Project Overview

Data Collection

Scraped Fortune 100 data from Wikipedia using Python, BeautifulSoup, and requests library.

Data Preparation

Cleaned and standardized industry names, revenue figures, and employee counts.

SQL Analysis

Validated KPIs using PostgreSQL queries for aggregation and ranking logic.

Visualization

Built interactive Power BI dashboard with drill-down capabilities and geographic mapping.

Web Scraping Methodology

Rather than using ready-made datasets, data was collected from scratch using Python to simulate real-world analytics scenarios.

01

Fetch Webpage

Used Python's requests library to retrieve HTML content from Wikipedia.

02

Parse HTML

Extracted structured tables using BeautifulSoup for data identification.
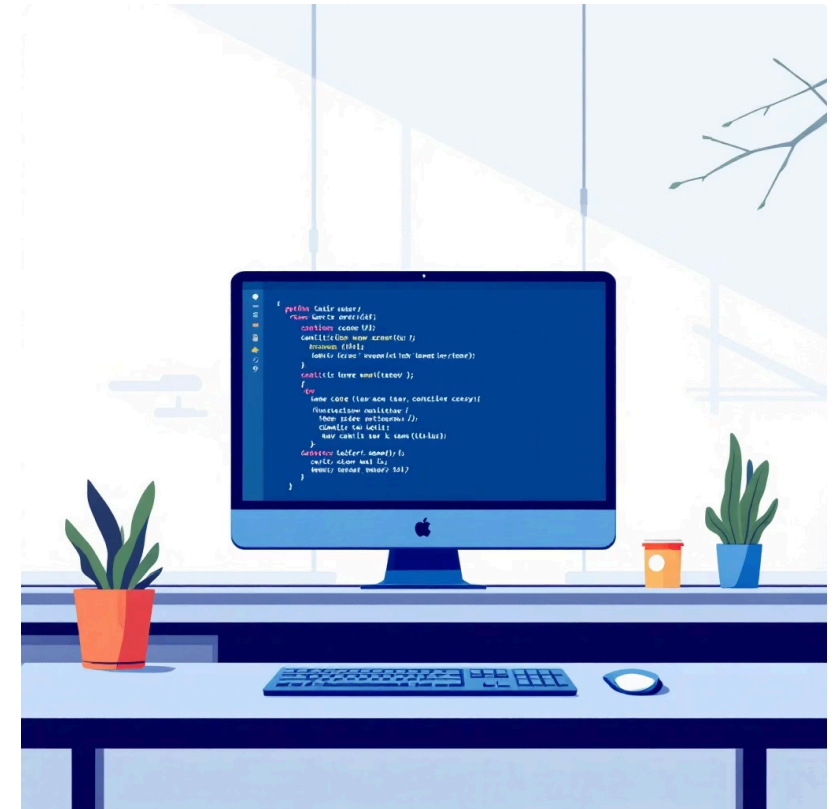
03

Extract Fields

Collected company name, revenue, industry, employees, and headquarters location.

04

Export Data

Stored scraped data into structured CSV format for downstream analysis.

Dataset at a Glance

100
Companies
Fortune 100 firms analyzed

$13.0T
Total Revenue
Combined annual revenue

16.2M
Employees
Total workforce size

7.69%
Avg Growth
Revenue growth rate

The dataset includes company name, industry classification, revenue, employee count, growth rate, and headquarters state across all Fortune 100 companies.

Technology Stack

Python Ecosystem

Web scraping with BeautifulSoup and Requests. Data transformation using Pandas and NumPy for cleaning and aggregation.
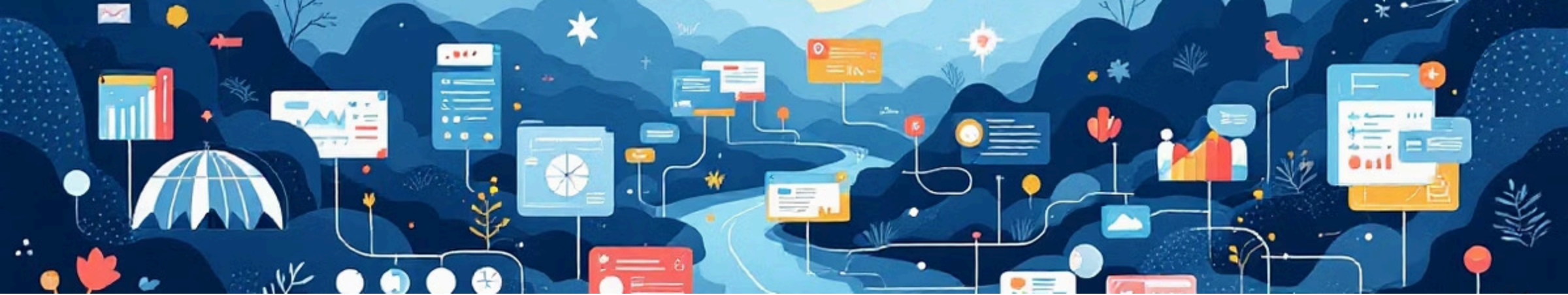
PostgreSQL

SQL querying for metric validation, aggregation logic, and ranking analysis of company performance.

Power BI

Interactive dashboard creation with KPI cards, drill-down filters, and geographic mapping capabilities.

Analysis Workflow

Scrape

Extract data from Wikipedia tables

Clean

Standardize formats and handle missing values

Query

Validate metrics using SQL

Visualize

Build interactive dashboards

Dashboard Highlights

Key Visualizations

- Top 10 Fortune 100 Companies by Revenue

- Revenue Contribution by Industry

- Revenue vs Growth Scatter Plot

- U.S. Map of Company Headquarters

The dashboard enables drill-down analysis using slicers for Industry and HQ State, providing interactive exploration of corporate performance metrics.

Key Insights

Revenue Concentration

Revenue is highly concentrated among top companies, with the largest firms dominating market share.

Industry Leaders

Financials, Healthcare, and Retail are the largest contributors to Fortune 100 revenue.

Growth Potential

Several mid-revenue companies exhibit strong growth rates, indicating future market opportunities.

Geographic Clusters

Headquarters are concentrated in major economic hubs: California, Texas, and New York.

Business Value

Strategic Insights

Provides clear view of corporate revenue concentration and competitive landscape positioning.

Industry Intelligence

Identifies dominant and high-growth industries for investment and market analysis.

Benchmarking

Supports strategic analysis and performance comparison across sectors and regions.

Skills Demonstration

Showcases practical analytics capabilities aligned with real business use cases.

Conclusion

Fortune 100 revenue is concentrated among leading companies, with Financial Services, Healthcare, and Retail dominating overall contribution.

While large firms lead in scale, several mid-sized companies exhibit higher growth rates, indicating future potential.

Geographic clustering of headquarters highlights key economic hubs, demonstrating the concentration of corporate power in strategic U.S. locations.