AI Companies Research Project

Overview

In order to strengthen my understanding of data modeling and visualization, I conducted a research project where I collected data on AI companies, processed and visualized it, and extracted insights. This project involved web scraping, data cleaning, and feature engineering, followed by graphing and analysis.

Data Collection

To collect the data, I used web scraping techniques with `BeautifulSoup` and `requests`. The dataset was extracted from Forbes' AI 50 list, which includes various AI startups and companies. The scraped information included:

- Company Name
- What the company does
- Funding Amount
- Headquarters
- CEO Name
- Year Founded
- Number of Employees

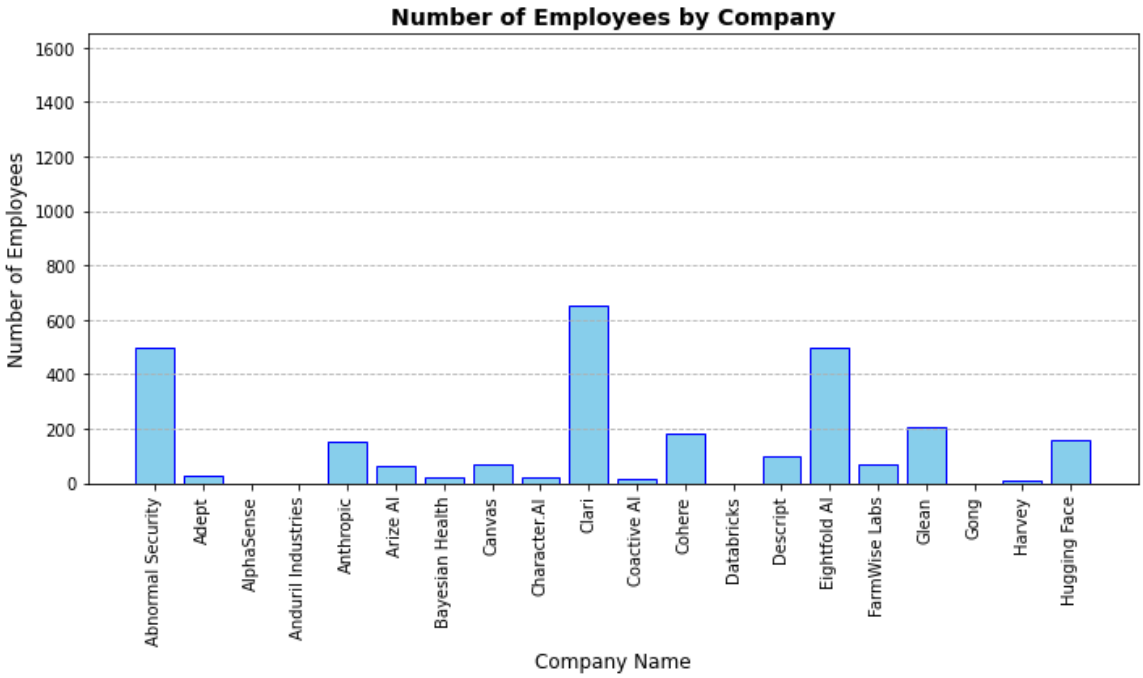
The collected data was then stored in a structured Pandas DataFrame and exported as a CSV file for further analysis.

Data Processing and Visualization

Once the data was collected, I performed data cleaning and feature engineering to ensure accurate analysis.

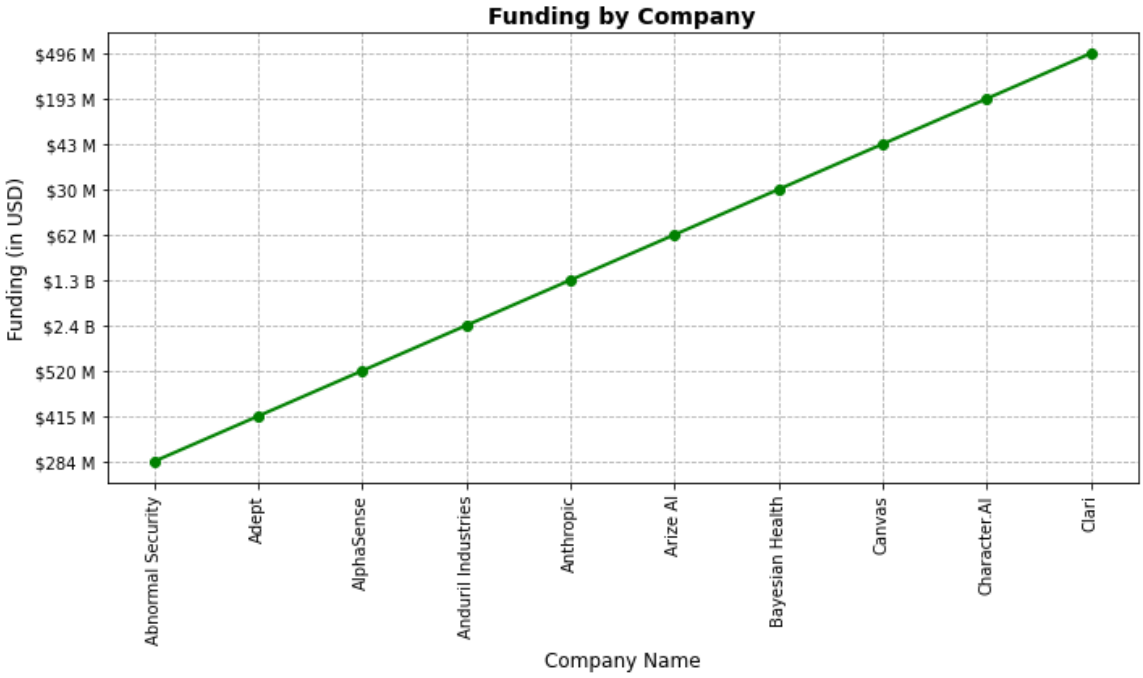
Employee Count Bar Chart

A bar chart was created to compare the number of employees across different AI companies.



Funding Line Chart

A line chart was created to visualize the funding amounts of the top AI companies.



Findings and Analysis

After analyzing the dataset, I found several interesting patterns:

Funding Disparities: The smallest company, Bayesian Health, received only $30 million in funding, while the largest, Anduril Industries, received $2.4 billion. This suggests a large variation in investment interest within the AI industry.

Employee Distribution: Companies based in San Francisco tend to have a higher number of employees compared to those in other locations, indicating that the city remains a major hub for AI development.

Investment Trends: The companies with higher funding tend to focus on areas like security, automation, and AI-driven infrastructure, making them more attractive to investors.

Conclusion

This research project allowed me to apply web scraping, data cleaning, and visualization techniques to a real-world dataset. By analyzing AI companies, I was able to gain insights into investment trends and employment patterns within the industry. The project highlighted how data science techniques can be used to extract meaningful insights from raw data.