



Fortune 500 Statistics

Project by Aaron, Tyler, Tyson



Research Questions

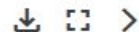
- Which companies and sectors produced the most revenue?
- How does market value and profit compare with the revenue generated by these companies?
- Which companies should be considered for potential investments?



Data Collection

- Kaggle dataset containing data from the Fortune 500 List spanning 1996-2023
- 16 columns x 27884 rows
- All data was from publicly traded companies

Fortune 500 Companies.csv (1.41 MB)



Detail Compact Column

10 of 16 columns ▾

About this file

This is the cleaned csv file that was made from the original web scraped data.

Data Cleaning

- 11 columns x 4500 rows
- Dataset prior to 2015 was missing a LOT of data including the sector

```
final_df = renamed_df.loc[renamed_df["year"]>= 2015]
final_df.reset_index(drop=True,inplace=True)
final_df
```

| | name | rank | year | industry | sector | headquarters_state | headquarters_city | market_value_millions | revenue_millions | profit_millions | employees |
|------|---------------------------|------|------|--|-------------------|--------------------|-------------------|-----------------------|------------------|-----------------|-----------|
| 0 | Walmart | 1 | 2015 | General Merchandisers | Retailing | AR | Bentonville | 265344.0 | 485651.0 | 16363.0 | 2200000.0 |
| 1 | Exxon Mobil | 2 | 2015 | Petroleum Refining | Energy | TX | Irving | 356549.0 | 382597.0 | 32520.0 | 83700.0 |
| 2 | Chevron | 3 | 2015 | Petroleum Refining | Energy | CA | San Ramon | 197381.0 | 203784.0 | 19241.0 | 64700.0 |
| 3 | Berkshire Hathaway | 4 | 2015 | Insurance: Property and Casualty (Stock) | Financials | NE | Omaha | 357344.0 | 194673.0 | 19872.0 | 316000.0 |
| 4 | Apple | 5 | 2015 | Computers, Office Equipment | Technology | CA | Cupertino | 724773.0 | 182795.0 | 39510.0 | 97200.0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 4495 | KKR | 496 | 2023 | Securities | Financials | NY | New York | 45225.0 | 7273.0 | 841.0 | 4150.0 |
| 4496 | Equinix | 497 | 2023 | Real Estate | Financials | CA | Redwood City | 66873.0 | 7263.0 | 704.0 | 12097.0 |
| 4497 | Sonoco Products | 498 | 2023 | Packaging, Containers | Materials | SC | Hartsville | 5975.0 | 7251.0 | 466.0 | 22000.0 |
| 4498 | ServiceNow | 499 | 2023 | Computer Software | Technology | CA | Santa Clara | 94338.0 | 7245.0 | 325.0 | 20433.0 |
| 4499 | Robert Half International | 500 | 2023 | Temporary Help | Business Services | CA | Menlo Park | 8677.0 | 7238.0 | 658.0 | 16300.0 |

4500 rows x 11 columns



Data Transformation

- Columns with values in billions were added

| market_value_millions | market_value_billions | revenue_millions | revenue_billions | profit_millions | profit_billions |
|-----------------------|-----------------------|------------------|------------------|-----------------|-----------------|
| 265344.0 | 265.3 | 485651.0 | 485.7 | 16363.0 | 16.4 |
| 356549.0 | 356.5 | 382597.0 | 382.6 | 32520.0 | 32.5 |
| 197381.0 | 197.4 | 203784.0 | 203.8 | 19241.0 | 19.2 |
| 357344.0 | 357.3 | 194673.0 | 194.7 | 19872.0 | 19.9 |
| 724773.0 | 724.8 | 182795.0 | 182.8 | 39510.0 | 39.5 |



Data Collection

- To use geographical data for visualizations, we used an API to gather coordinates

```
for i in city_state:
    if i in city or i == None or i in skip_list:
        pass
    else:
        try:
            location = geolocator.geocode(i)
            lat.append(location.latitude)
            lng.append(location.longitude)
            city.append(i)
            print(f'Adding coordinates for {i} to list')
            time.sleep(1)
        except:
            print(f'Skipping {i}')
            skip_list.append(i)
            time.sleep(1)
```

```
Adding coordinates for ('Bentonville', 'AR') to list
Adding coordinates for ('Irving', 'TX') to list
Adding coordinates for ('San Ramon', 'CA') to list
Adding coordinates for ('Omaha', 'NE') to list
Adding coordinates for ('Cupertino', 'CA') to list
Adding coordinates for ('Detroit', 'MI') to list
```



Data Transformation

- All of our data was merged into a final dataset

| name | rank | year | industry | sector | headquarters_state | headquarters_city | latitude | longitude | employees | market_value_millions | market_value_billions |
|---------------------------|------|------|--|-------------------|--------------------|-------------------|-----------|-------------|-----------|-----------------------|-----------------------|
| Walmart | 1 | 2015 | General Merchandisers | Retailing | AR | Bentonville | 36.372854 | -94.208817 | 2200000.0 | 265344.0 | 265.3 |
| Exxon Mobil | 2 | 2015 | Petroleum Refining | Energy | TX | Irving | 32.829518 | -96.944218 | 83700.0 | 356549.0 | 356.5 |
| Chevron | 3 | 2015 | Petroleum Refining | Energy | CA | San Ramon | 37.764802 | -121.954439 | 64700.0 | 197381.0 | 197.4 |
| Berkshire Hathaway | 4 | 2015 | Insurance: Property and Casualty (Stock) | Financials | NE | Omaha | 41.258746 | -95.938376 | 316000.0 | 357344.0 | 357.3 |
| Apple | 5 | 2015 | Computers, Office Equipment | Technology | CA | Cupertino | 37.322893 | -122.032290 | 97200.0 | 724773.0 | 724.8 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| KKR | 496 | 2023 | Securities | Financials | NY | New York | 40.712728 | -74.006015 | 4150.0 | 45225.0 | 45.2 |
| Equinix | 497 | 2023 | Real Estate | Financials | CA | Redwood City | 37.486324 | -122.232523 | 12097.0 | 66873.0 | 66.9 |
| Sonoco Products | 498 | 2023 | Packaging, Containers | Materials | SC | Hartsville | 34.374043 | -80.073401 | 22000.0 | 5975.0 | 6.0 |
| ServiceNow | 499 | 2023 | Computer Software | Technology | CA | Santa Clara | 37.233325 | -121.684635 | 20433.0 | 94338.0 | 94.3 |
| Robert Half International | 500 | 2023 | Temporary Help | Business Services | CA | Menlo Park | 37.451967 | -122.177992 | 16300.0 | 8677.0 | 8.7 |



Data Storage

- Our final dataset was stored in a SQL database
- SQL data was converted to JSON format in order to be utilized for JavaScript applications

| name character varying (255) | rank integer | year integer | industry character varying (255) | sector character varying (255) | headquarters_state character varying (5) |
|---------------------------------|-----------------|-----------------|--|-----------------------------------|---|
| Walmart | 1 | 2015 | General Merchandisers | Retailing | AR |
| Exxon Mobil | 2 | 2015 | Petroleum Refining | Energy | TX |
| Chevron | 3 | 2015 | Petroleum Refining | Energy | CA |
| Berkshire Hathaway | 4 | 2015 | Insurance: Property and Casualty (Stock) | Financials | NE |
| Apple | 5 | 2015 | Computers, Office Equipment | Technology | CA |
| General Motors | 6 | 2015 | Motor Vehicles & Parts | Motor Vehicles & Parts | MI |
| Phillips 66 | 7 | 2015 | Petroleum Refining | Energy | TX |
| General Electric | 8 | 2015 | Diversified Financials | Financials | CT |
| Ford Motor | 9 | 2015 | Motor Vehicles & Parts | Motor Vehicles & Parts | MI |
| CVS Health | 10 | 2015 | Food & Drug Stores | Food & Drug Stores | RI |



Visualizations

[Link:](#)