World's Largest Companies 2019

1.1 Introduction

1.1.1 Background

The largest companies in the world are getting bigger. The stock valuation of a company can changes quickly. Microsoft briefly dethroned Apple as the most valuable enterprise in the world in 2018, Amazon beat both companies for the top slot in January 2019, and then Microsoft's valuation surged past rivals to cross the $1 trillion threshold for the first time in April 2019. How did that happen? While Apple—which is trying to move beyond hardware to services— suffered due to poor iPhone and MacBook sales, Microsoft's business model is centered around steadily growing streams of recurring revenues. You might not need a new smartphone or laptop every year, but if you purchased a software license, a cloud package or a videogame subscription, you are likely to buy it again in the future.

However, the total dollar value of a company's outstanding shares, its market capitalization can be affected by a myriad of unpredictable factors. For example, stock markets have repeatedly lost value due to US President Donald Trump's tweets, including companies that have little or nothing to do with China.

Market capitalization indicates the cost of owning a piece of a company today. But does it reflect the company's objective value or inherent worth? Warren Buffett once famously said, "Nothing is further from the truth." Still, as the founder of Berkshire Hathaway, the sixth most valuable business in the world by market capitalization, Buffett knows that behind such numbers always lies some truth.

This list comprises the world's largest companies by consolidated 2018 revenue, according to the Fortune Global 500 latest tally published on July 22, 2019. American retail corporation Walmart has been the world's largest company by revenue since 2014, with US$514 billion in revenue in 2018. Walmart was also the largest company in the world by revenue from 2002 to 2005, from 2007 to 2008, and from 2010 to 2017.

1.1.2 Problem

Data that might contribute to determining companies improvement might include the revenues, profits, the countries witch government owns 50% or more, and the domain of the industry. This project aims to cluster the he World's Largest Public Companies around the globe to know countries that are really powerful economically.

1.2 Data acquisition and cleaning

1.2.1 Data sources

Our data frame is from Wikipedia, the dataset is limited to the top 50 companies, all of which have annual revenues exceeding US$110 billion. 32 out of the 50 companies are either from the United States or from China. Only companies that publish financial data and report figures to a government agency are included. Therefore, this list is incomplete, as it excludes large companies such as Vitol, Cargill, Koch Industries, Schwarz Group and Kuwait Petroleum Corporation because they do not publish financial data.

1.2.2 Data cleaning

The data frame contains the columns Rank, Name, Industry, Revenue (USD millions), Profit (USD millions), Employees, Country and Ref.

First of all, we are going to delete the column Ref cause it's for the reference information in Wikipedia and we don't need it in our analysis.

Second step is to change some column names to an easiest manipulation.

Old column's name	New column's name
Revenue (USD millions)	Revenue_USDM
Profit (USD millions)	Profit_USDM