World's Largest Companies 2019

Introduction

- 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

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,

Data acquisition and cleaning

- 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large companies such as Vitol, Cargill, Koch Industries, Schwarz Group and Kuwait Petroleum Corporation because they do not publish financial data

Data cleaning

- 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.

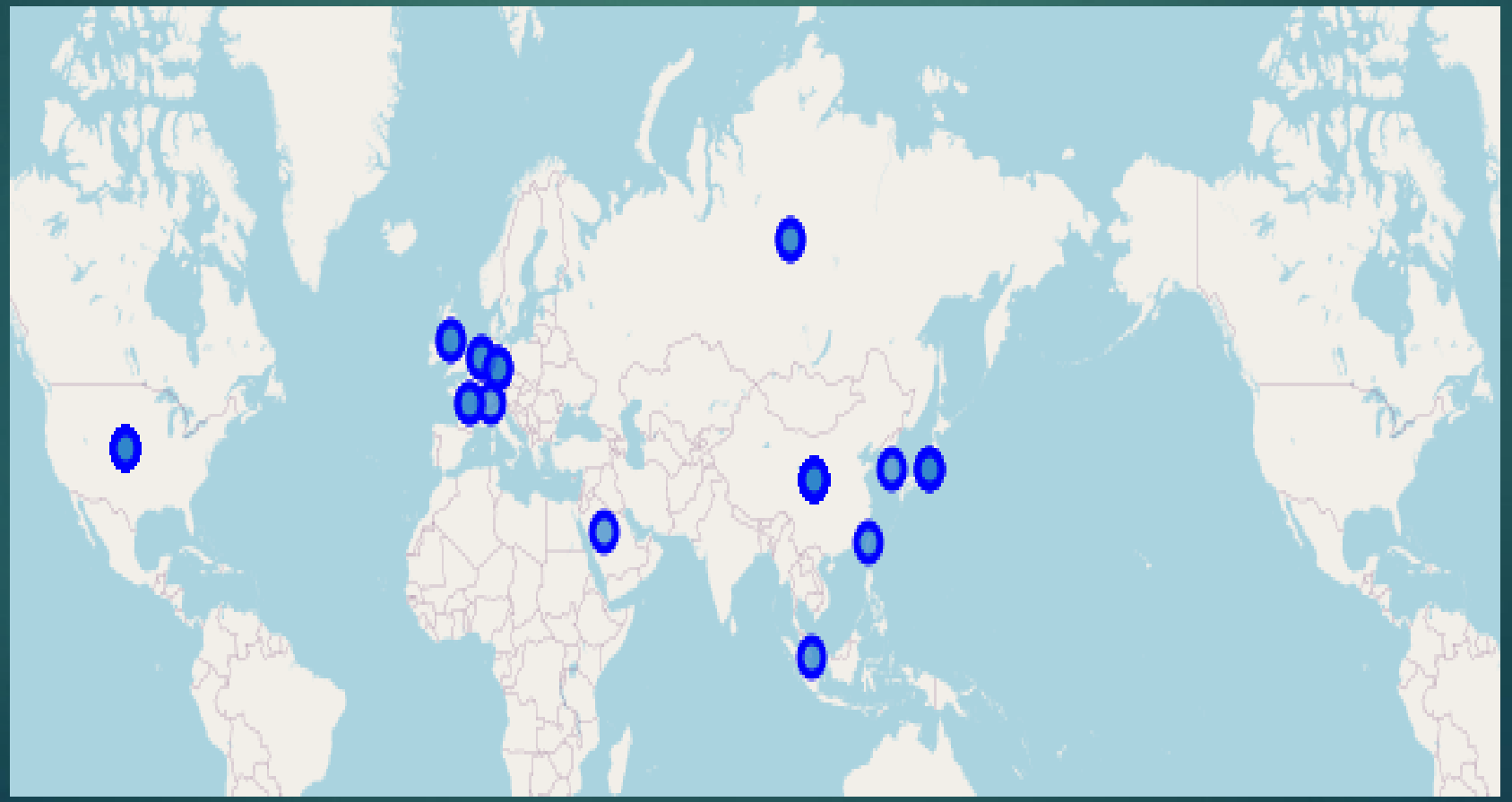
Old column's name	New column's name
Revenue (USD millions)	Revenue_USDM
Profit (USD millions)	Profit_USDM

Data cleaning

- Third, we will remove the duplicated countries names in the column “Country”.
- Forth, we have the value "Netherlands / United Kingdom" separated by "/", let's duplicate the row for each county.
- In the last step of data cleaning, the revenue column type is "object", we need to convert it to int.

Data visualization

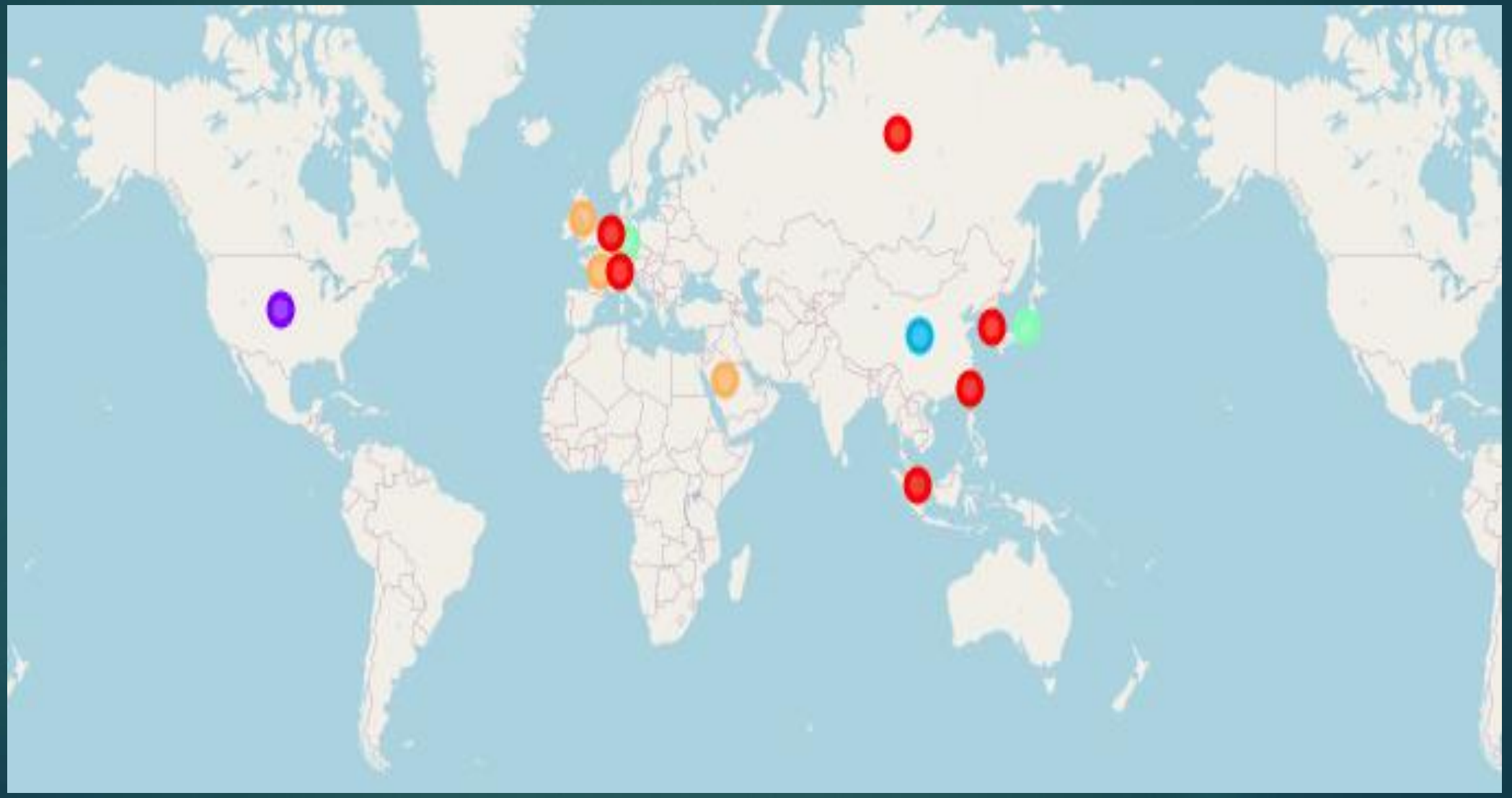
- we will create a world map with the localization of each country



Cluster countries based on the revenues of its companies

Cluster Labels	Country	index	Latitude	Longitude	Revenue_USDM
0	Netherlands	5	52.500170	5.748082	175009
	Russia	7	64.686314	97.745306	250447
	Singapore	9	1.357107	103.819499	180744
	South Korea	10	36.638392	127.696119	221579
	Switzerland	11	46.798562	8.231974	219754
	Taiwan	12	23.598298	120.835363	175617
1	United States	14	39.783730	-100.445882	4169049
2	China	1	35.000074	104.999927	2263520
3	Germany	3	51.083420	10.423447	602635
	Japan	4	36.574844	139.239418	561157
4	United Kingdom	0	54.702354	-3.276575	396556
	France	2	46.603354	1.888334	309684
	Saudi Arabia	8	25.624262	42.352833	355905

Clustering Visualization with folium



Conclusion

- To conclude, we have to say that we were able in this project to cluster the World's Largest Public Companies contributions around the globe to know countries that are really powerful economically.