Analysis of data from World Bank on Cereal Yield and Production:

Mohammed Farhan Ahmed (Student ID 21073161)

Link for data set: https://data.worldbank.org/indicator/AG.PRD.CREL.MT
https://data.worldbank.org/indicator/AG.YLD.CREL.KG

My GitHub repository link: https://github.com/Farhan608/Rework-for-Assignment-2-Statistic-Trends

Introduction:

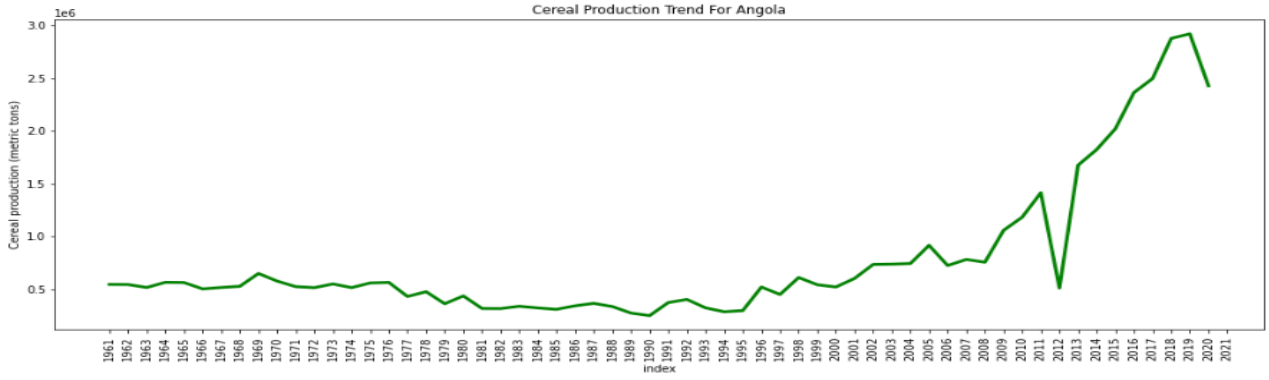
In this project we are working with the Cereal data collected from the Word Data Bank. The data for cereal production and cereal yield for years 1960 to 2021 is collected for all the countries. The data has been obtained in two different files and thus having the same structure. The two files are merged to create a master dataset and the final dataset is used to further analyze the data. The main objectives that we have worked on in this project is to see the trend of cereal production and cereal yield for country "Angola", the correlation between cereal production and cereal yield.

Analysis & Results:

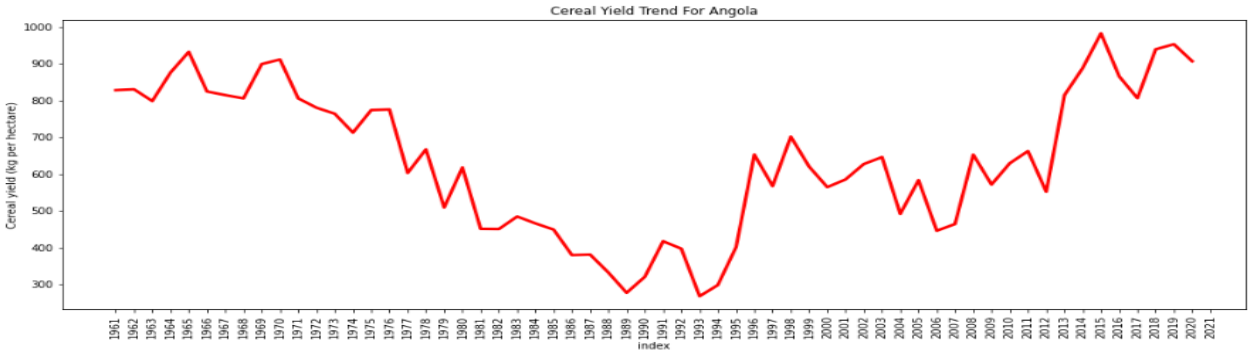
The correlation matrix attached below shows us that the correlation between production of cereal and the cereal yield is very low. Having such low correlation is sometimes negligible. Hence we can say that there is no correlation between cereal production and cereal yield. One possibility might be having different measure units for cereal production and cereal yield.

Indicator Name	Cereal production (metric tons)	Cereal yield (kg per hectare)
Indicator Name		
Cereal production (metric tons)	1.000000e+00	5.597393e-18
Cereal yield (kg per hectare)	5.597393e-18	1.000000e+00

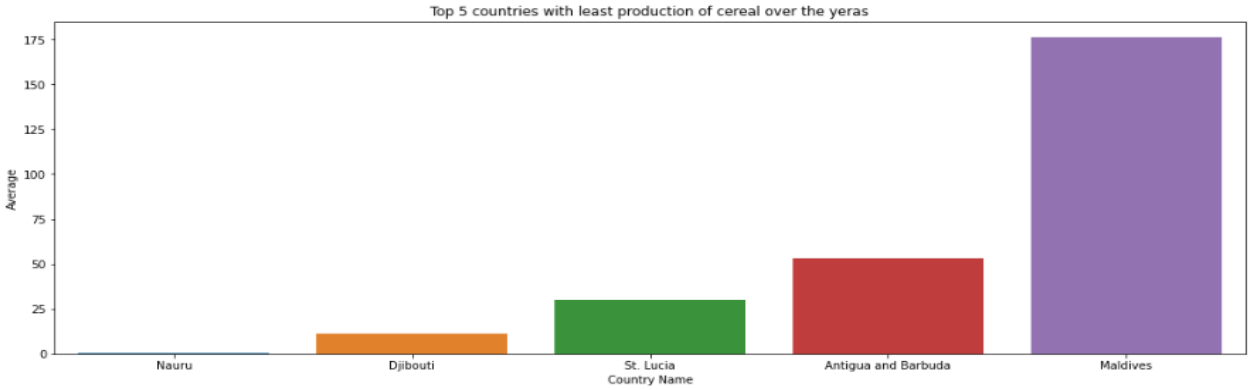
Moving forward, a graph has been created to show the cereal production for Angola and is attached below. It can be seen that the overall trend of cereal production is increasing from 1961 till now. A sudden drop in cereal production could be observed in year 2012 but after that we are seeing a continuous increase except 2021 where we were having most of the values empty.



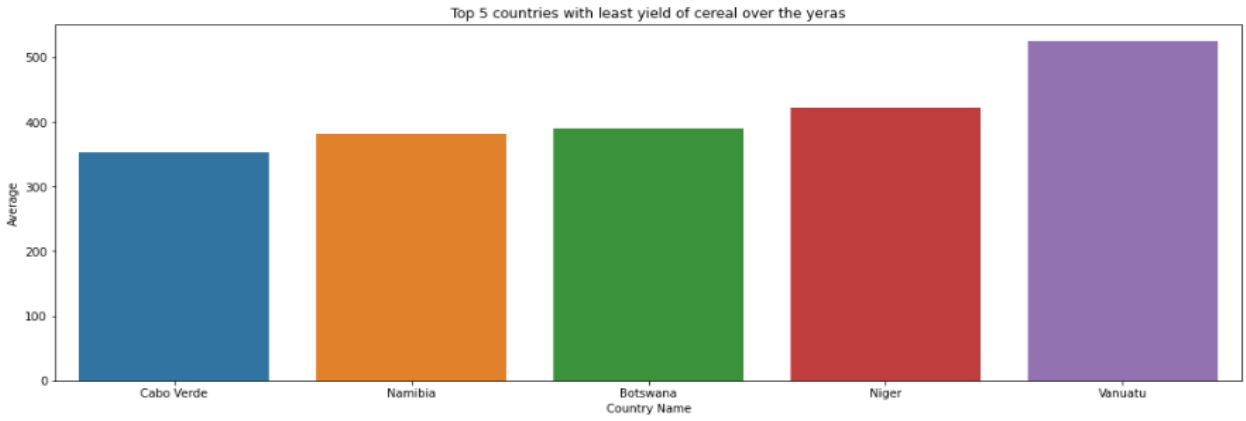
Similarly, another trend plot is created for cereal yield in Angola and is attached below. Here the trend is not steady. It is decreasing from 1961 till 1993 and after that its overall increasing but still if we do an year over year comparison it sometime increases and sometime it decreases. So we can say that in early years till 1990's the cereal yield was very low compared to the yield after 1990's.



Next, the top 5 least producing cereal countries are find out by taking the average production over the years and a bar chart is created. It can be seen that Nauru and Dajibouti are the two countries where the production of cereals is the lowest.



Similarly, a graph is created for cereal yield as well and the top 5 countries with least cereal yield are shown where we have Cobo Verde on first and Naimibia on second number.



Conclusion:

Following the completion of this investigation, we came to the conclusion that there is no association between the production of cereal and the yield of cereal. We saw that the general trend of cereal production is one of expansion, but in terms of cereal yield, we noticed that cereal yield fluctuated from year to year, with some years seeing increases and others seeing losses. In addition, we noticed that Nauru has the lowest grain output, and Cobo Verde has the lowest cereal yield. Both of these findings were based on our observations.