AGRICULTURAL PRODUCTION OF INDIA

Objective:-

- The goal of this project is to solve the problems of various crops Cultivation or the Production in India.
- Development of the reports and the Dashboards from which one can simply understand about the area, production and yield of different crops and their cultivation and production cost by the financial year in India and as well as in different states and also that one can understand which crop and its variety is suitable for their region.

Problem Statement

- The Agriculture business domain, as a vital part of the overall supply chain, is expected to highly evolve in the upcoming years via the developments, which are taking place on the side of the Future Internet. This paper presents a novel Business-to-Business collaboration platform from the agri-food sector perspective, which aims to facilitate the collaboration of numerous stakeholders belonging to associated business domains, in an effective and flexible manner.

Benefits

- Better understanding of cultivated area of India and its states.
- Better understanding of Production and Yield of different Crops.
- Better understanding of the Production and type of the Cultivation cost of different crops which also effect in deciding the MSP of crops in India.
- Recommendation of the crops and its variety for the cultivation according to the Region.

Data Sharing Agreement :

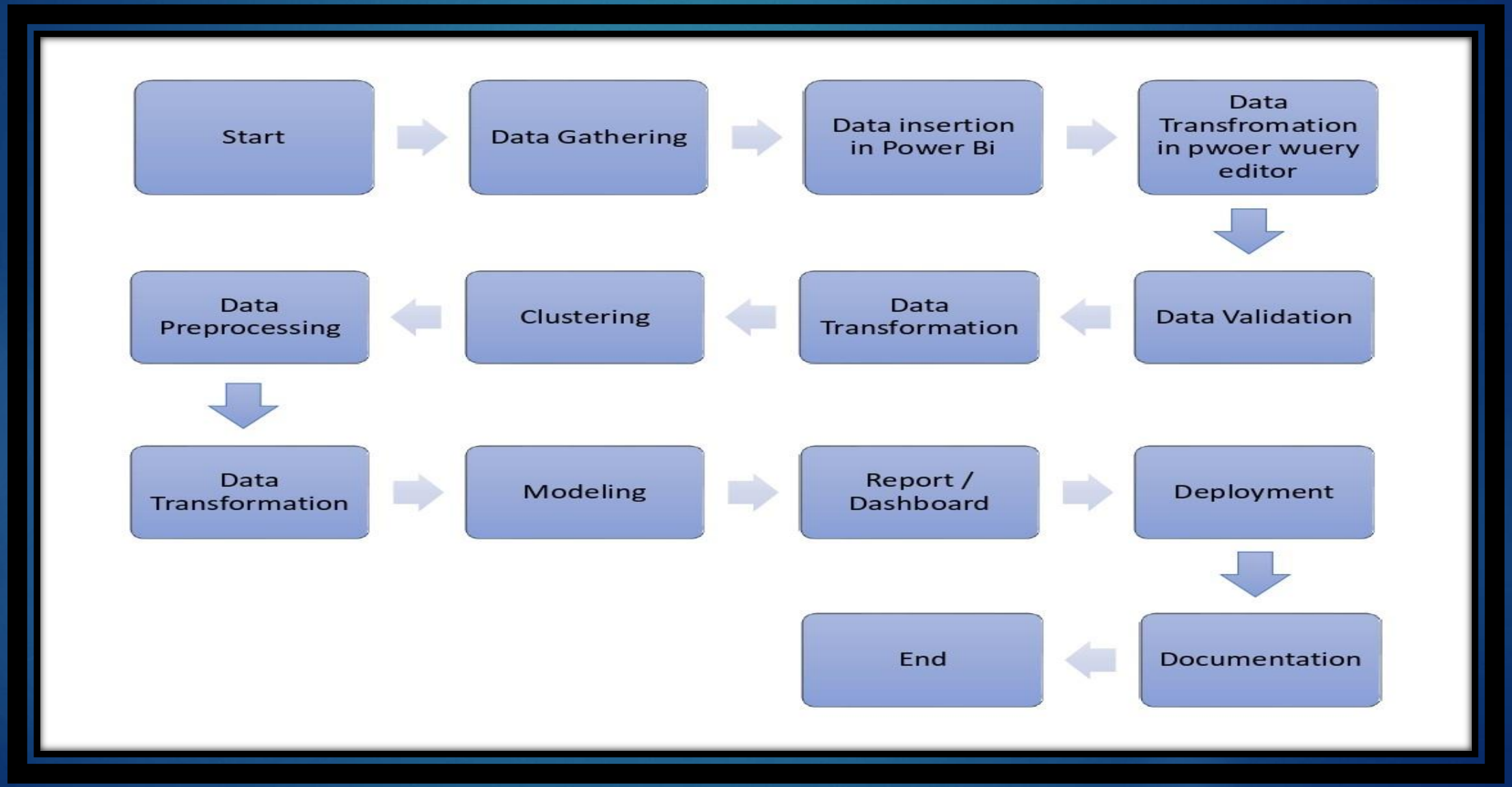
- Main File name (Agriculture Data)
- No of files under main file (5)
- Files Names under main file (Agriculture 1, Agriculture 2, Agriculture 3, Agriculture 4 and Produce)
- Length of date stamp(8 digits)
- Length of time stamp(6 digits)
- Column names (Crop, State, cost of cultivation, cost of production, area, production, yield, Financial year, Variety,
- Column data type (Integer, Decimal, Date and String)

Data Description:-

- Agricultural data contains mainly Area, production and yield and financial year of different crops and other columns as like cost, recommended zone and variety.
- Crops: Different types of crops.
- Production: Agricultural production is the use of crops and animal products to enhance human life sustainably.
- Area: Crops are cultivated at this area.
- 4. Yield: It is also called Production per unit area.
- 5. Year of Financial year: Crops grown or produce in that specific time.
- 6. Cost of Cultivation (A2+FL): The A2 + FL cost includes all cash transactions and payments made by the farmer, including the cost of family labour It also includes the rental value of the leased land.

- Cost of Cultivation (C2): C2 includes A2 + FL cost as well as the rent of owned land and interest on owned capital. Hence, the MSP calculated on the basis of C2 cost is much higher as compared to A2 + FL.
- Cost of Production (C2): It is C2 type production cost.
- Variety: It includes the variety of Crops.
- Recommended zone: Suitable region for a better production for the crops

Architecture



Data Insertion in Microsoft Powe Bi:

- All the files of Agricultural data was in csv format. So, after opening the Microsoft Power Bi, I have imported all the files in the Microsoft Power Bi.
- Data was unstructured so I have transferred the data in Power Query Editor for the further process.

Data Validation and Data Transformation :

- First I have done some basic transformation of the data such as- removing error, removing null values, getting first row as header if required in the Power Query Editor.
- The data was so messy so I have extracted the important data from the files and transformed it according to the requirement.
- In the transformation process I have done different transformation such as- pivoting and unpivoting columns, removing error, removing null values, creation of new tables.
- Validation of date column:- I have validated the columns data type the most unstructured columns was date type data.

Report and dashboard:

I have made some report and dashboard according to the requirement-

1. Production and yield of Rice of Indian states by their cultivated area and financial year.
2. Production and yield of Wheat of Indian states by their cultivated area and financial year.
3. Production and yield of Coarse Cereals of Indian states by their cultivated area and financial year.
4. Cost of Cultivation and production of different crops and their Yields by their Area.
5. Production of Food Grains and Oilseeds of different Crops by their financial Year.
6. Comparison of Production of different Crops by their Financial year Crop recommendation for the Cultivation.

Deployment:-

After making the Report and dashboards I have published it on the Microsoft Power Bi Service from Microsoft Power Bi Desktop.

Insights:-

- Some measure insights of this project:--
- From year 2004-2013 the top states in crops area, production and yield is given below-

Parameters	Top State for the Crops		
	Rice	Wheat	Coarse Cereals
Area	Uttar Pradesh	Uttar Pradesh	Rajasthan
Volume	Best Bengal	Uttar Pradesh	Rajasthan
Yield	Punjab	Punjab	Punjab

- In the yield Punjab is the top state for the crops of Rice, Wheat and Coarse Cereals. It means that Punjab is producing more rice, wheat and coarse cereals in the less area than the other state.
- Cost of cultivation is higher than the cost of cultivation.
- Cost of cultivation(C2) is higher than the Cost of Cultivation(A2+FL).
- Production of Food grains followed uptrend from 2005-2014.
- Production of Oilseeds overall followed uptrend but from 2005-2010 it followed a little bit uptrend downtrend in consecutive years.
- Andhra Pradesh's cultivation cost is higher than the other states.
- -Karnataka is the top yielder for the Sugarcane.
- There are a lot of insights from this Projects Dashboard. According to the requirement we can easily get that information.

Q & A:

Q1) What's the source of data?

The data is provided by the Ineuron in the format of google drive link.

Q 2) What was the type of data?

The data was the combination of numerical and Categorical values.

Q 3) What's the complete flow you followed in this Project?

Refer slide 5th for better Understanding

Q 4) After the File validation what you do with incompatible file or files which didn't pass the validation?

Files like these are moved to the Achieve Folder and a list of these files has been

shared with the client and we removed the bad data folder.

Q 5) How logs are managed?

We are using different logs as per the steps that we follow in validation and modeling like File validation log , Data Insertion ,Model Training log , prediction log etc.

Q 6) What techniques were you using for data pre-processing?

- Removing unwanted attributes
- Visualizing relation of independent variables with each other and output variables
- Checking and changing Distribution of continuous values
- Removing outliers
- Cleaning data and imputing if null values are present.
- Converting categorical data into numeric values.
- Scaling the data

Q 9) What are the different stages of deployment?

I have deployed it on the Microsoft Power Bi Service.

Thankyou