

CROP PRODUCTION IN INDIA

Submitted

By

Dr. Megha Sharma

1 INTRODUCTION

1.1 Overview

India has the second largest arable land resources in the world, 15 significant climates of the world's 20 Agri-climatic regions, and 40 out of 60 soil types exist in India. India is the largest producer of spices, pulses, jute, etc., and the second largest producer of wheat, rice, fruits and vegetables, cotton, and oilseeds. It is vital for the growth of the Indian economy; on average, 70% of the households and 10% of the urban population depend on agriculture for their livelihood. In this project, we will visualize the crop production of 124 varieties for the period 1997- 2015 in India. IBM Cognos Analytics with Watson is used for the analysis of the data for crop production.

1.2 Purpose

This project aims to understand the fundamentals of data analytics, use IBM Cognos Analytics to understand the data and create visualization plots. Here raw datasets for crop production in India are provided in the 'CSV' format and analyzed to draw conclusions based on the information provided. Through various visualization created in the IBM Cognos Analytics, we can understand the patterns followed by the raw data and use this information to conclude.

2 LITERATURE SURVEY

2.1 Existing Problem

The amount of data collected in today's world is overwhelming. It becomes difficult to understand and access critical information. Performing this task manually is time-consuming, and collecting meaningful information and analyzing the data requires an automated system. Visual representation of the data is necessary as they are easily understood, for big data's collecting the relevant information and making it presentable is slow if done manually. With so much information to process, there is a need for a data system that can automatically collect and organize the data.

2.2 Proposed Solution

IBM Cognos Analytics is a tool that automates the analysis of big data. It is a business intelligence tool for web-based reporting and analytics. Cognos Analytics helps in finding and focusing on things that are important to the business. We can understand the latest trends, compare data and assess business performance for multidimensional analysis. Here, in this study, the crop production in India is analyzed using visualization, creating a dashboard, and getting insights into crop production using Cognos analytics. All these visualizations are discussed later in this report.

3 THEORETICAL ANALYSIS

3.1 Block diagram

The content of the data is taken from the data.world website. This dataset provides a huge amount of information on crop production in India from several years. It has 2,46,092 data points (rows) and six features (columns) describing each crop production-related details. In the dataset, State name, district name under each state, year of the crop (1997-2015), crop production season (Autumn, Kharif, Rabi, Summer, and Winter), the total area covered, and crop production. Figure 1 shows the step-by-step methodology adopted for the analysis of crop production.

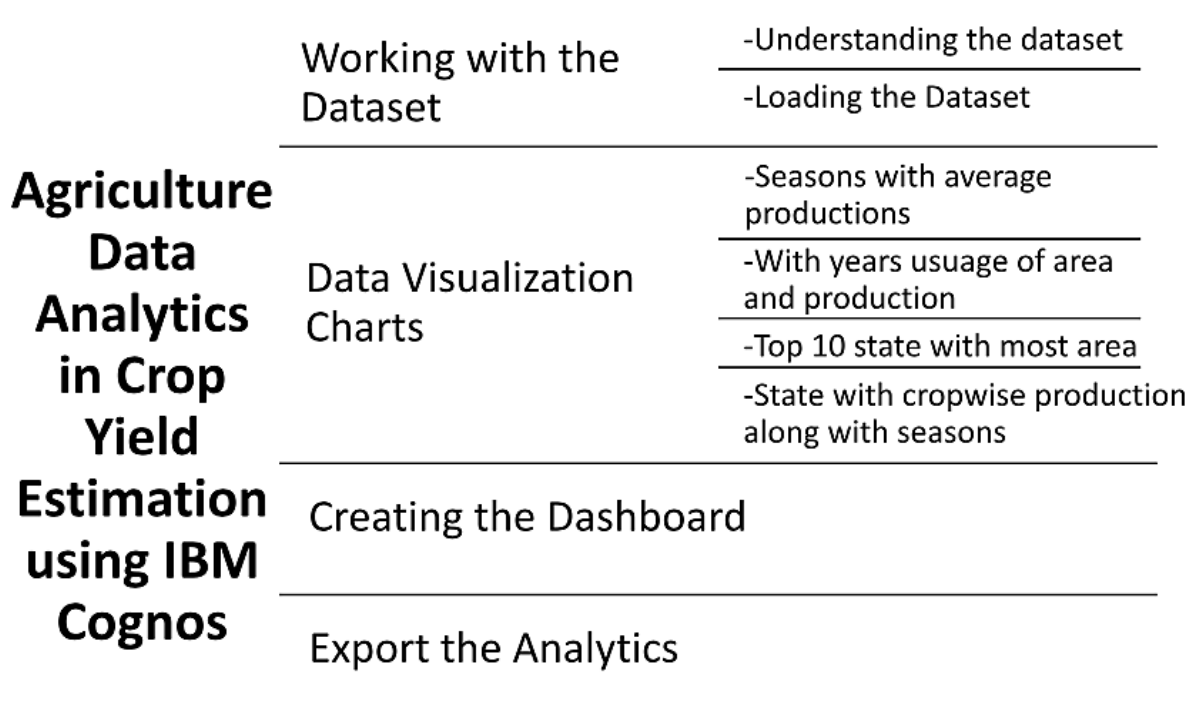


Figure 1 Overview of the Project.

3.2 Software Designing

IBM Cognos is a business intelligence tool for web-based reporting and analytics. Various features of data aggregation and user-friendly detailed reports are performed using this software. Cognos offers the following features which help us in data analysis:

- Offering In-memory streaming analytics
- Real-time events, alerts, and notifications.
- Personalized and progressive interaction
- Drag-and-drop, free-form assembly, and search-assisted authoring
- Scenario modeling, real-time monitoring, and predictive analytics
- Editing existing data.
- It is platform-independent, scalable, and reliable.

This software can do comparisons, trend analysis, and analysis of the top and bottom performers, allowing you to share your research with others. It provides the filtering, calculating, and sorting support you need for analysis. Next section, we will discuss the study done using IBM Cognos.

4 EXPERIMENTAL INVESTIGATION

In Figure 2, the interface of IBM Cognos Analytics is illustrated, consisting of several sections. On the left side, we have a selection pane with all the data of the “crop_production.csv.” State Name, District Name, Crop year, Season, Crop Area, and Crop Production data is stored, and by selecting those names, we can upload the data to the work area in the center. The work area contains the charts to perform the analysis. You can display analysis in the form of graphs or tables or maps, or a combination.

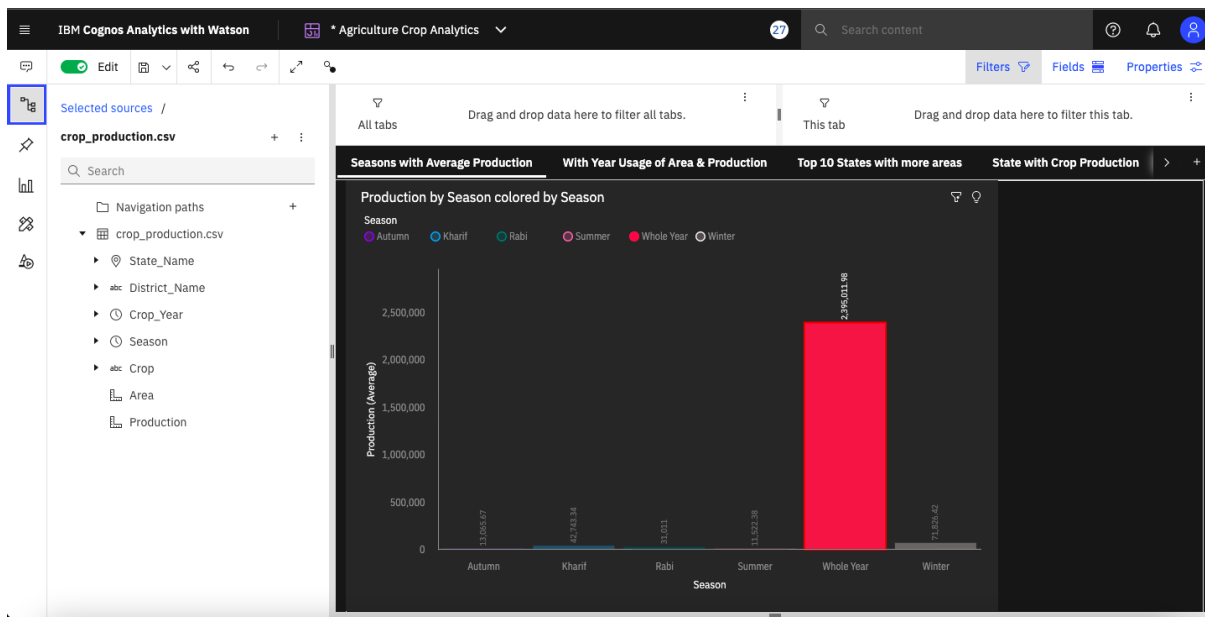


Figure 2 Interface of IBM Cognos Analytics.

First, upload the data file in “.csv” format to start with an analysis. Next, create a new dashboard and start a further study. Drag and drop the elements from the selection pane to the workspace and select the visualization needed.

5 RESULTS AND DISCUSSION

This section will discuss different visualizations made using the crop production data. Figure 3 to 7 shows the visualization created for this analysis; we will discuss one by one each visualization.

Figure 3 shows the variation of average crop production for Autumn, Kharif, Rabi, Summer, and Winter. The visualization used for this analysis is a Column, with seasons on the horizontal axis and average production on the vertical axis. As the season varies, the crop production changes since not every crop can sustain all the weather. As per the data, the highest crop production is for winter with 71,826 tons/year, and the lowest is for summer with 11,522 tons/ year.

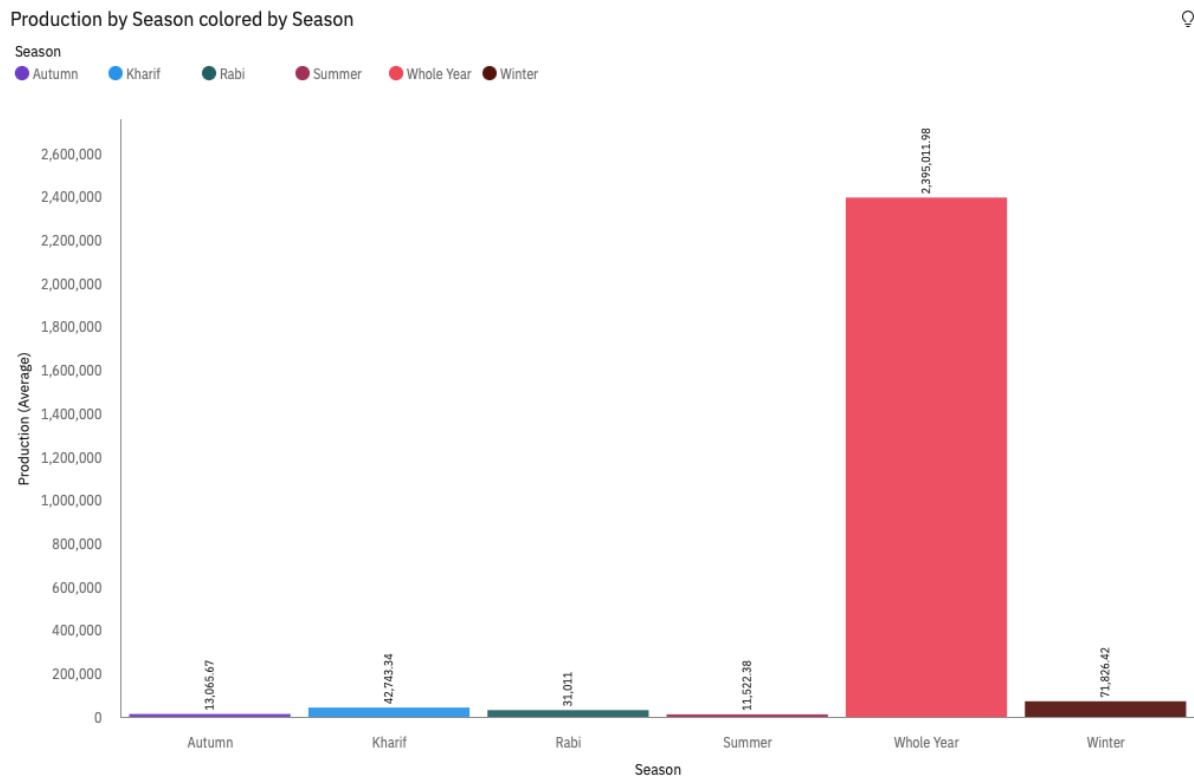


Figure 3 Average Production by Season.

Figure 4 shows the year-wise average crop area and crop production variations. The visualization used for average crop area is an area plot, while for crop production line graph is used. We can observe from the graph that for the year 1997, even though the average crop area is very high, the production is low; this may be due to technological

advancement or natural factors affecting crop production. After 1997, crop area is almost even till 2015 with few lows in 2002 and 2007. Average crop production increased significantly in 1998 and 2011 compared to the previous years. We can also observe a sudden drop in the average crop production at the end of the year 2014.

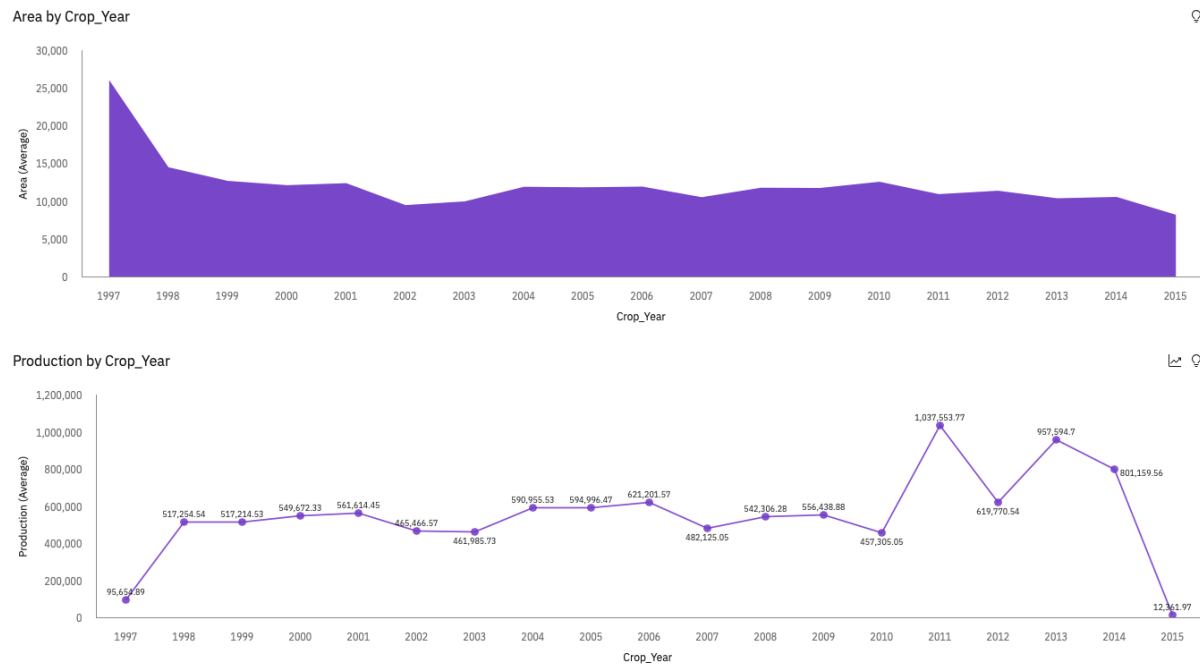


Figure 4 Average Area and Production over the Years.

A column chart visualization was adopted for Figure 5, illustrating the top 10 states with maximum crop area. Haryana, Telangana, Madhya Pradesh, Andhra Pradesh, and Uttar Pradesh have almost the same area under cultivation. Punjab has the maximum crop area compared to any other state.

After applying the visualization technique, we can create a dashboard, as shown in Figure 7. This dashboard shows the summary for all the visualization in one place, which helps understand the data better.

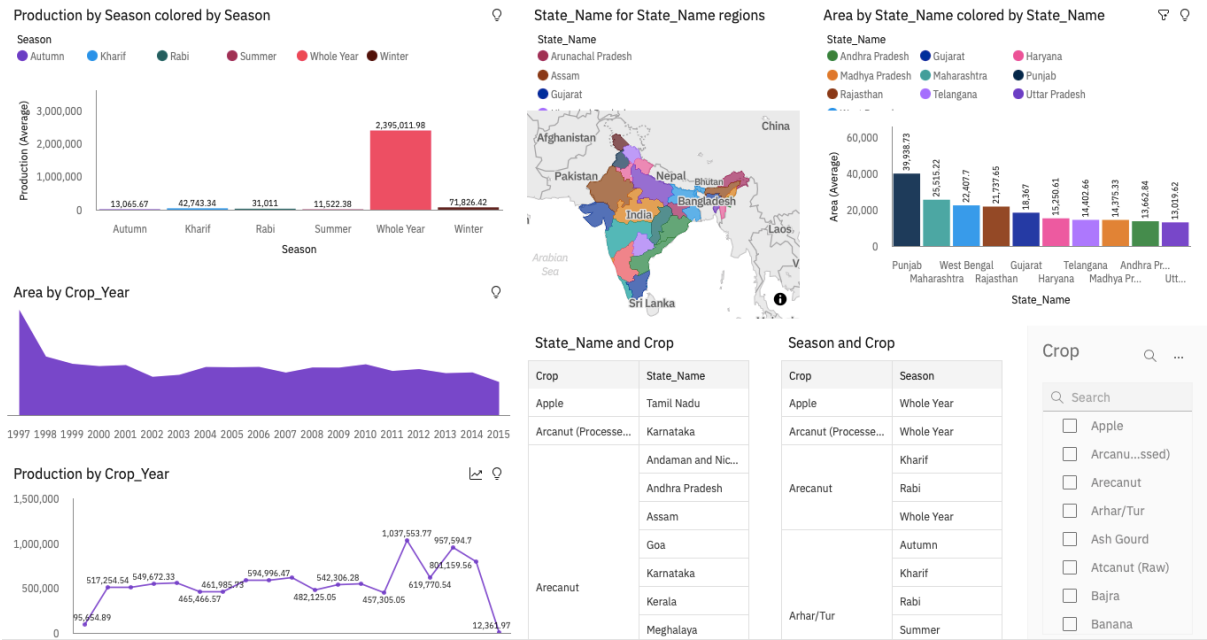


Figure 7 Dashboard for Crop Production in India.

6 ADVANTAGES & DISADVANTAGES

Here, using IBM Cognos, our analysis became easier to understand. It is hard to comprehend the tabular data while conducting the research. Following are some of the advantages of using IBM Cognos for data analytics:

1. We can view, assemble and personalize the information quickly.
2. It helps in understanding the analysis and anticipating the implications.
3. Visual representation of data helps in collaborating with the decision makes.
4. Time is reduced in analyzing and assessing the data.
5. Communication and coordination of the task among the right people.
6. It makes the analysis much simpler to understand and implement.

The analysis considers this a good platform, but training is required to use data analytics at departmental levels.

7 APPLICATIONS

Data analysis helps businesses and organizations succeed by interacting with information technology, statistics, and business. Data analytics applications increase efficiency and improve performance by discovering patterns in data. Transportation, managing risk, planning, healthcare, etc., are the areas where data analytics applications will enhance their functionality.

8 CONCLUSIONS

Here is the summarized observation from the crop data production using IBM Cognos:

1. The highest crop production is for winter with 71,826 tons/year, and the lowest is for summer with 11,522 tons/ year.
2. In 1997 the average crop area was very high, but the production was low; this may be due to technological advancement or natural factors affecting crop production.
3. Average crop production increased significantly in 1998 and 2011, and a sudden drop in the average crop production at the end of 2014.
4. Haryana, Telangana, Madhya Pradesh, Andhra Pradesh, and Uttar Pradesh have almost the same area under cultivation. Punjab has a maximum crop area compared to any other state.

Data Link:

https://us1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FAgriculture%2BCrop%2BAnalytics&action=view&mode=dashboard&subView=model00000182013b62fe_00000002

9 BIBILOGRAPHY

1. <https://www.youtube.com/embed/2gWcJy7wF9E>
2. <https://www.youtube.com/embed/Mqw9YfLtQIk>
3. <https://www.youtube.com/embed/q36JRI4E6oY>
4. <https://www.youtube.com/embed/3H-3DmRKQP8>
5. <https://www.youtube.com/embed/blgUU-yZo6A>
6. <https://www.youtube.com/embed/SetaNqDvsV8>
7. <https://www.youtube.com/embed/5rYSaeUfluc>
8. <https://www.guru99.com/cognos-tutorial.html#8>
9. https://www.tutorialspoint.com/cognos/cognos_analysis_studio.htm
10. <https://data.world/thatzprem/agriculture-india>