

Project Based Experiential Learning Program (Nalaiya Thiran)|IBM



Estimate the Crop Yield using Data Analytics

Team ID : PNT2022TMID47431

Team Size : 4

Team Leader : NIVITHA S D

Team member : HARINI M

Team member : MANISHA S

Team member : SURIYAA

Category:

Data Analytics

Goals

Crop production in India is one of the most important sources of income and India is one of the top countries to produce crops. As per this project we will be analyzing some important visualization, creating a dashboard and by going through these we will get most of the insights of Crop production in India.

Bringing the power of Watson to farmers

Agriculture, a \$2.4 trillion industry, is an underpinning of economies around the world. Factors, for example, environmental change, populace development and food security concerns have moved the business into looking for additional creative ways to deal with safeguarding and further developing harvest yield. Subsequently, man-made brainpower is consistently arising as a component of the business' innovative development.

Along with IBM Watson and The Weather conditions Organization, groups from IBM Exploration Brazil and IBM Exploration India planned and fabricated a set-up of agribusiness instruments and answers for assist the farming business with utilizing the force of man-made intelligence to come to additional educated conclusions about their yields - the Watson Choice Stage for Horticulture.

Supporting the stage, IBM Matches GEOSCOPE processes a portion of the satellite information and fills in as capacity in the ongoing framework engineering. By amassing and investigating terabytes of multi-facet geospatial information utilizing AI and progressed examination, Matches permits us to store and run questions on the geo-referred to information.

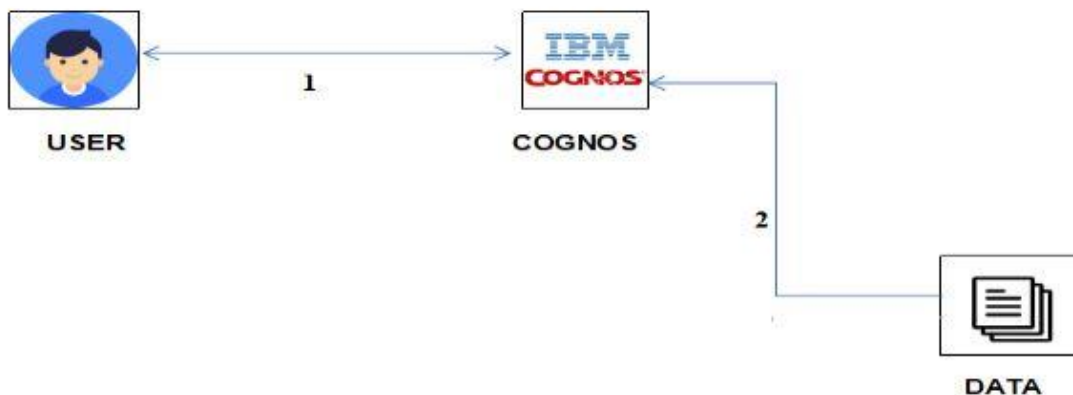
Four of the APIs remembered for this new stage come from our worldwide labs.

Specifications

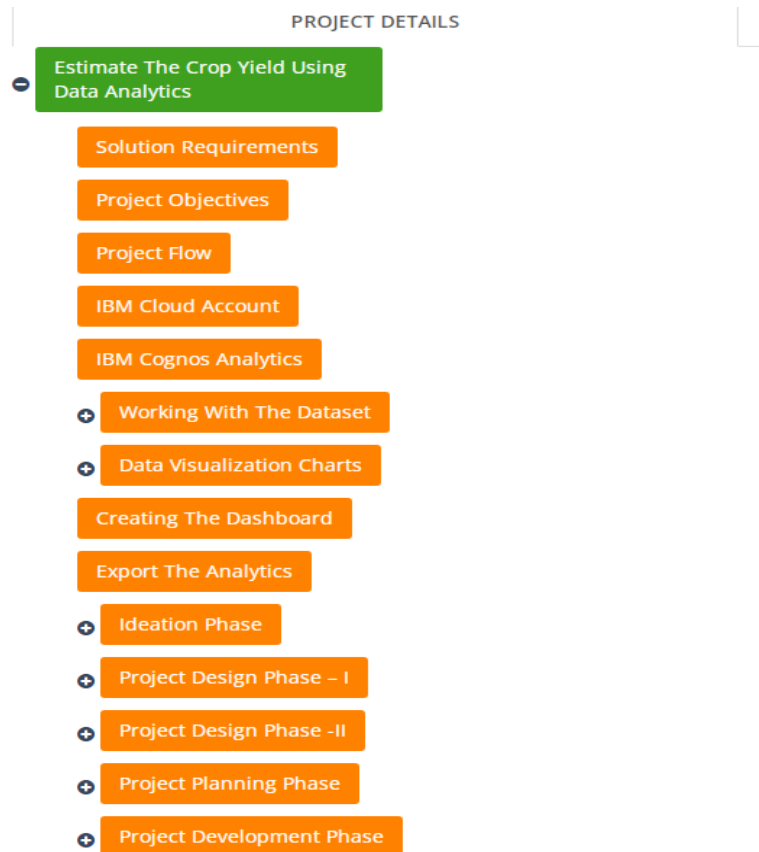
Exploratory Data Analysis, IBM Cloud

Technical Architecture

Technical Architecture:



Milestones and Project Details



Solution Requirements

Service Used: IBM Cognos Analytics.



Project Objectives

By the end of this project, you will:

- Know fundamental concepts and can work on IBM Cognos Analytics.
- Gain a broad understanding of plotting different graphs.
- Able to create meaningful dashboards

Project Flow

- Users create multiple analysis graphs/charts.
- Using the analyzed chart creation of the Dashboard is done.
- Saving and Visualizing the final dashboard in the IBM Cognos Analytics.

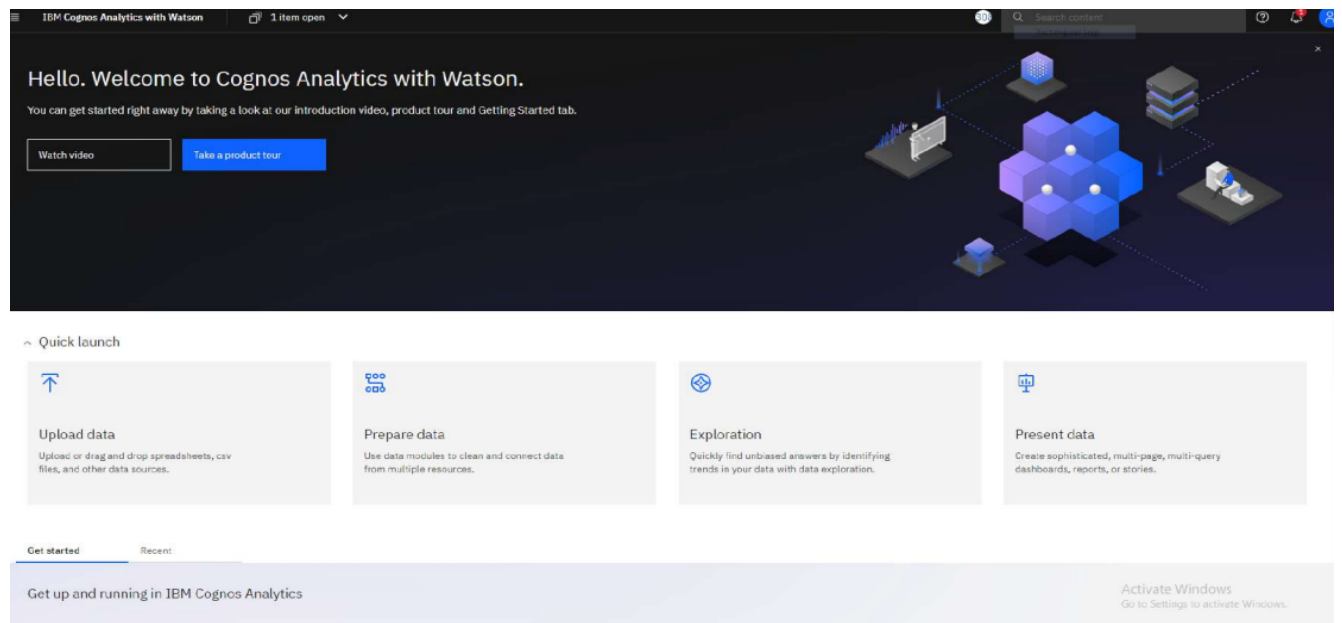
To accomplish this, we have to complete all the activities and tasks listed below

- IBM Cloud Account
- Login to Cognos Analytics
- Working with the Dataset
 - Understand the Dataset
 - Loading the Dataset
- Data visualization charts
 - Seasons with average productions

- With years usage of Area and Production
- Top 10 States with most area
- State with crop production
- States with the crop production along with season (Text Table)
- Dashboard Creation
- Export the Analytics

IBM Cognos Account

Create and login to IBM Account.



Working With The Dataset

To work on the given dataset, you need to first Understand the Dataset and the Load it to Cloud platform then Build the required Visualizations to provide various visual analytical solutions.

Understanding Dataset:

Dataset link: [Crop Production in India | Kaggle](#)

This project is based on an understanding of the crop production of India .Download the dataset from the below link. It has 2,46,092 data points (rows) and 6 features (columns) describing each crop production related details.

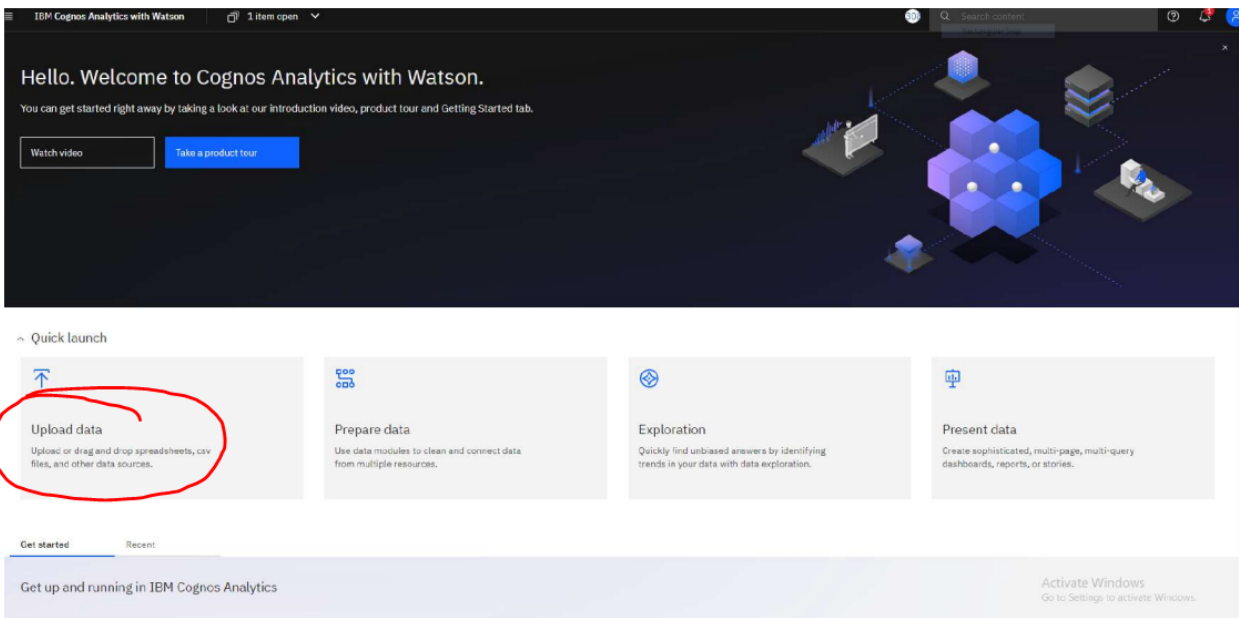
Dataset Link :[Dataset](#)

Let's understand the data we're working with and give a brief overview of what each feature represents or should represent

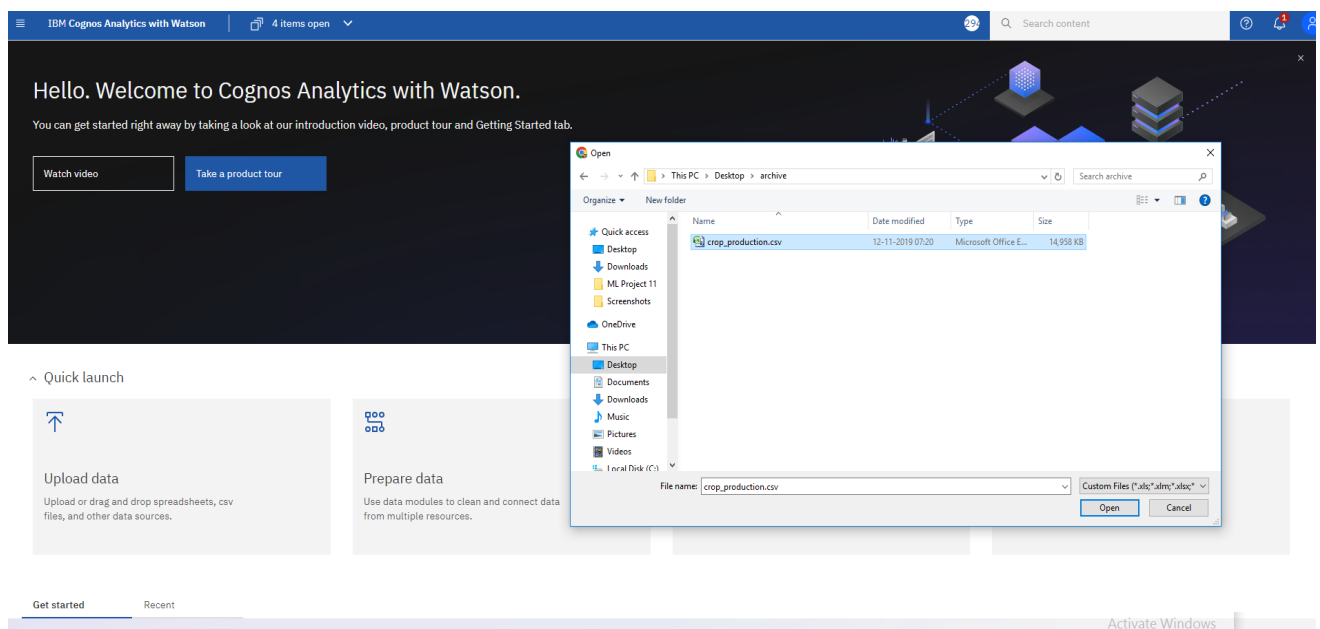
1. State Name - All the Indian State names.
2. District Name -Different District names.
3. Crop Year- contains the crop years.
4. Season – Different seasons for crop production.
5. Area- Total number of areas covered.
6. Production- production of crops.

Loading Data Set to Cognos:

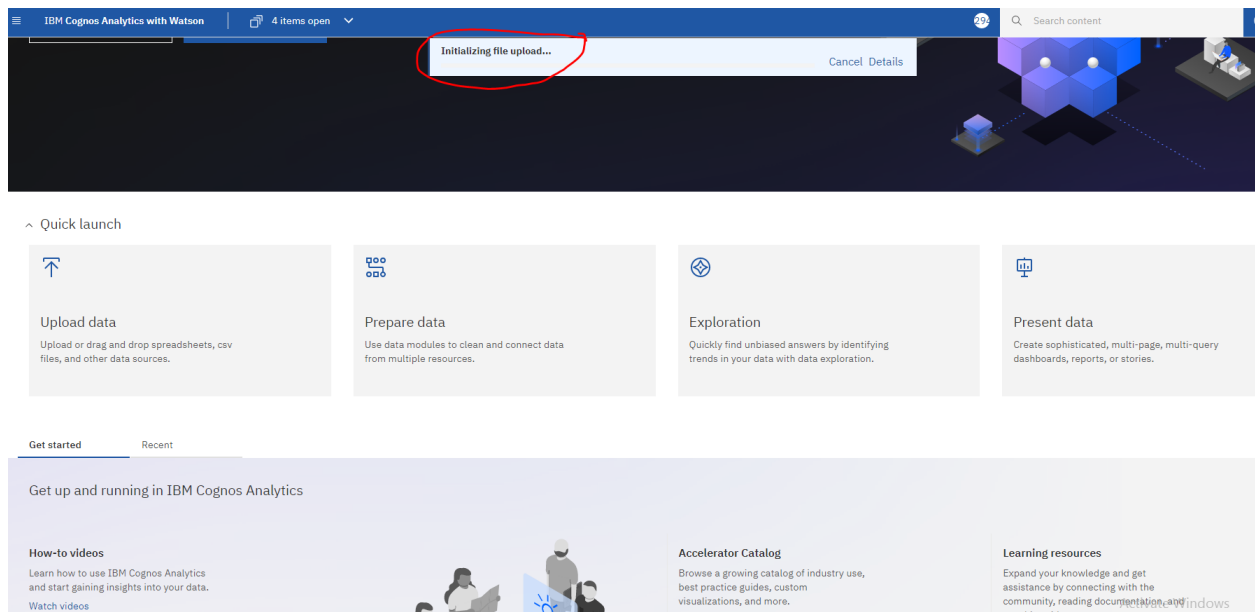
Step:1



Step:2



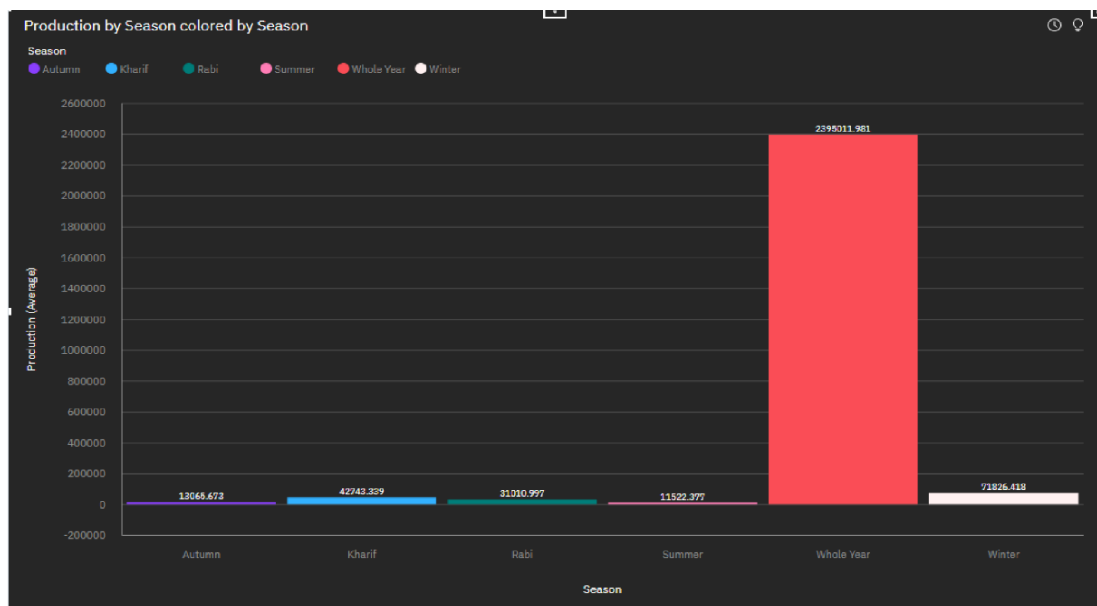
Step:3



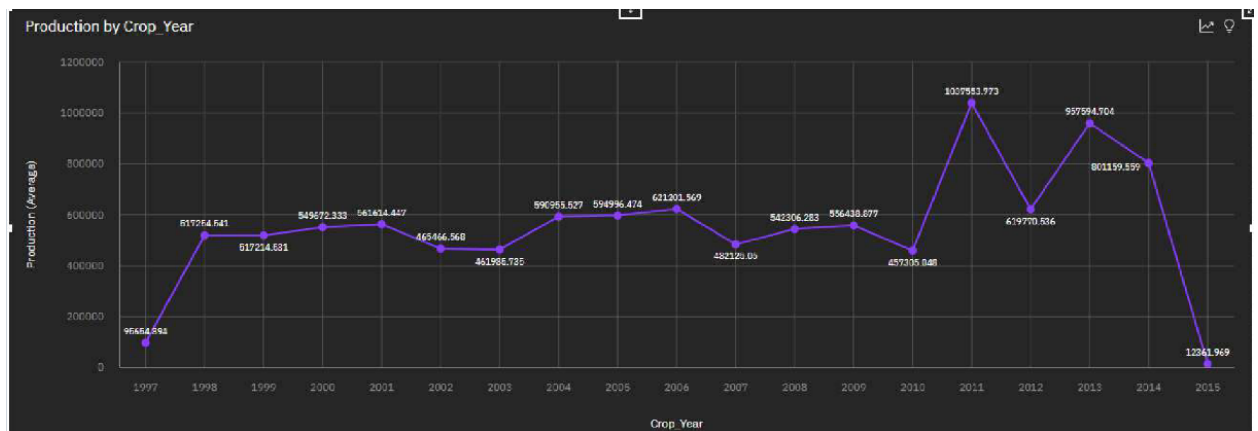
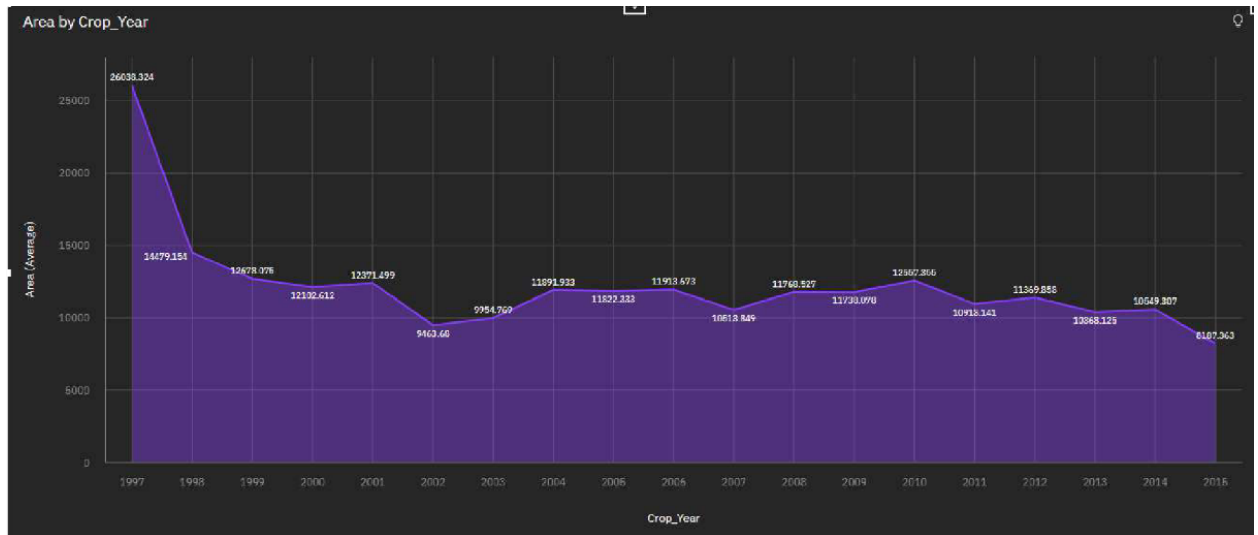
Data Visualization Charts

- Seasons with average productions

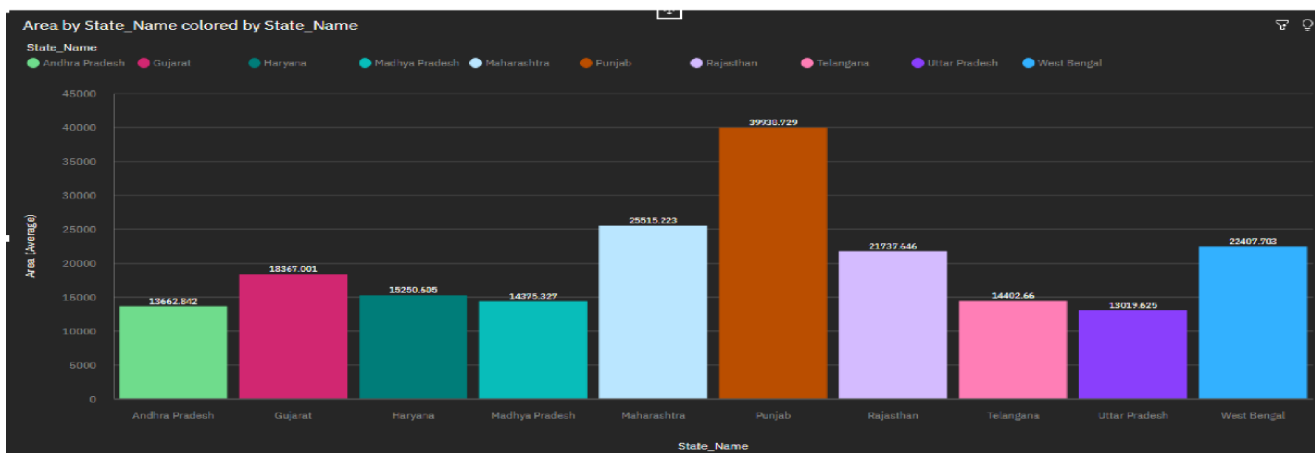
As production of crops depends on different seasons, so let's plot the graphs to visualize the average production based on different seasons.



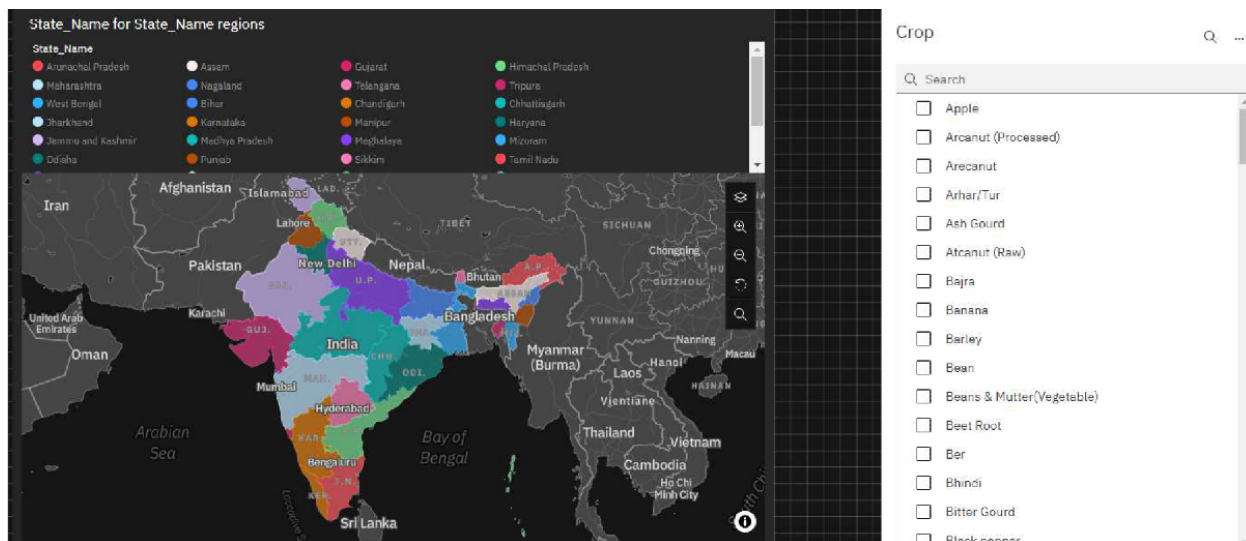
● With Years Usage Of Area And Production



● Top 10 States With Most Area



● State With Crop Production



● States With The Crop Production Along With Season (Text Table)

| State_Name and Crop | | Crop | | | |
|---------------------|-----------------------------|--------|-----------------|-----------------|----------------|
| Crop | State_Name | Season | Crop | Crop | Crop |
| Apple | Tamil Nadu | Autumn | Arhar/Tur | Arhar/Tur | Arhar/Tur |
| Arcanut (Processed) | Karnataka | | Banana | Banana | Banana |
| Arcanut | Andaman and Nicobar Islands | | Cotton(lint) | Cotton(lint) | Cotton(lint) |
| | Andhra Pradesh | | Dry chillies | Dry chillies | Dry chillies |
| | Assam | | Dry ginger | Dry ginger | Dry ginger |
| | Goa | | Groundnut | Groundnut | Groundnut |
| | Karnataka | | Jowar | Jowar | Jowar |
| | Kerala | | Jute | Jute | Jute |
| | Meghalaya | | Maize | Maize | Maize |
| | Madhucherry | | Moong(Green... | Moong(Green... | Moong(Gree... |
| | Tamil Nadu | | Onion | Onion | Onion |
| | West Bengal | | Paddy | Paddy | Paddy |
| | Andaman and Nicobar Islands | | Peas & beans... | Peas & beans... | Peas & bean... |

Crop

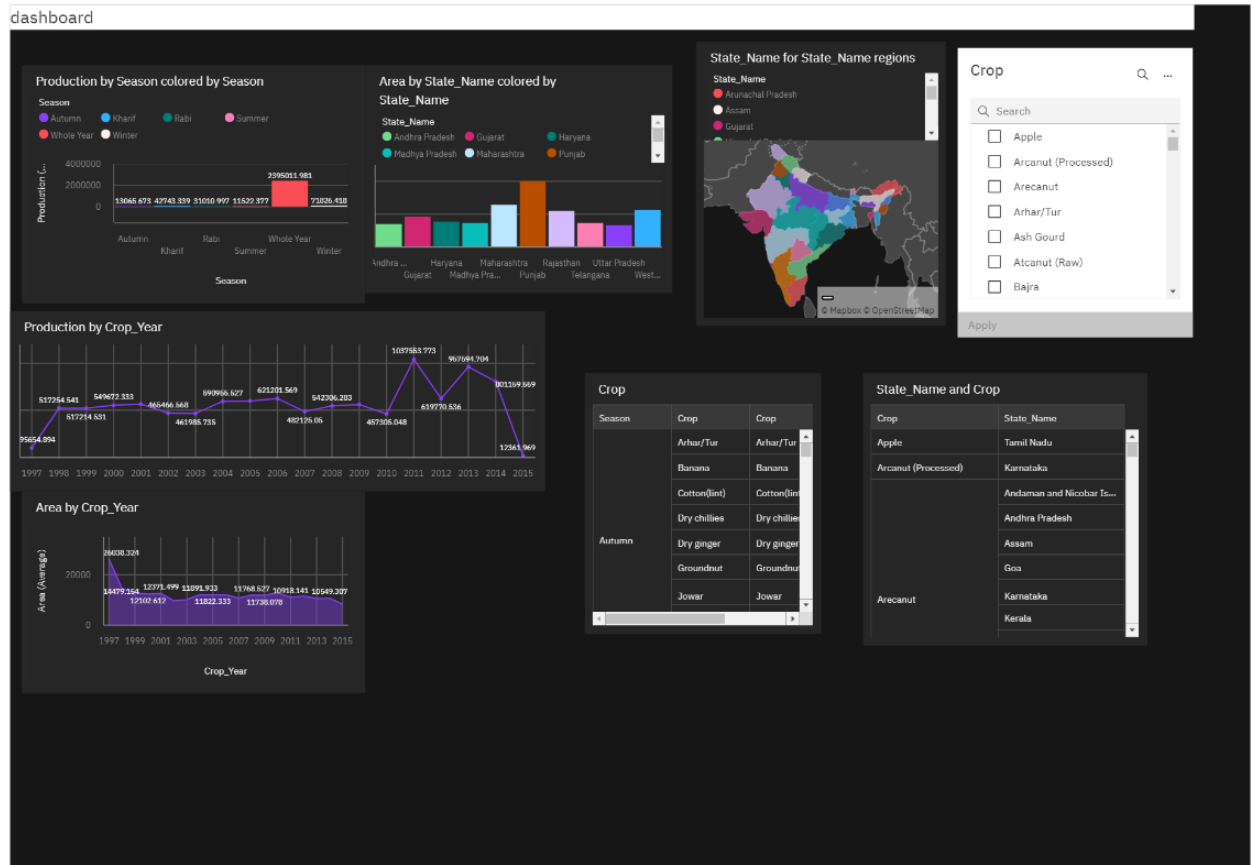
Q Search

- ☐ Apple
- ☐ Arcanut (Processed)
- ☐ Arcanut
- ☐ Arhar/Tur
- ☐ Ash Gourd
- ☐ Atcanut (Raw)
- ☐ Bajra
- ☐ Banana
- ☐ Barley
- ☐ Bean
- ☐ Beans & Mutter(Vegetable)
- ☐ Beet Root
- ☐ Ber
- ☐ Bhindi
- ☐ Bitter Gourd
- ☐ Black pepper

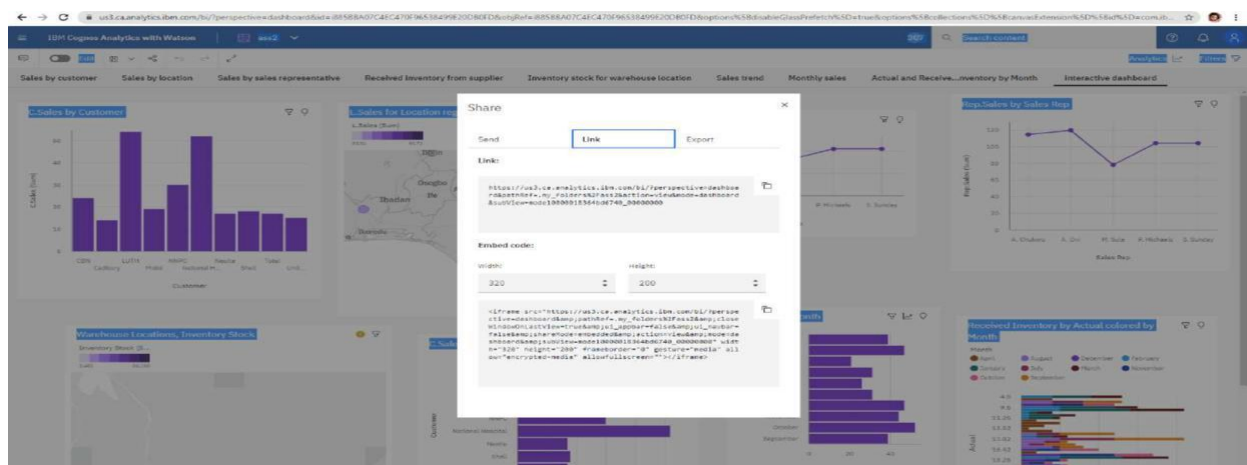
Apply

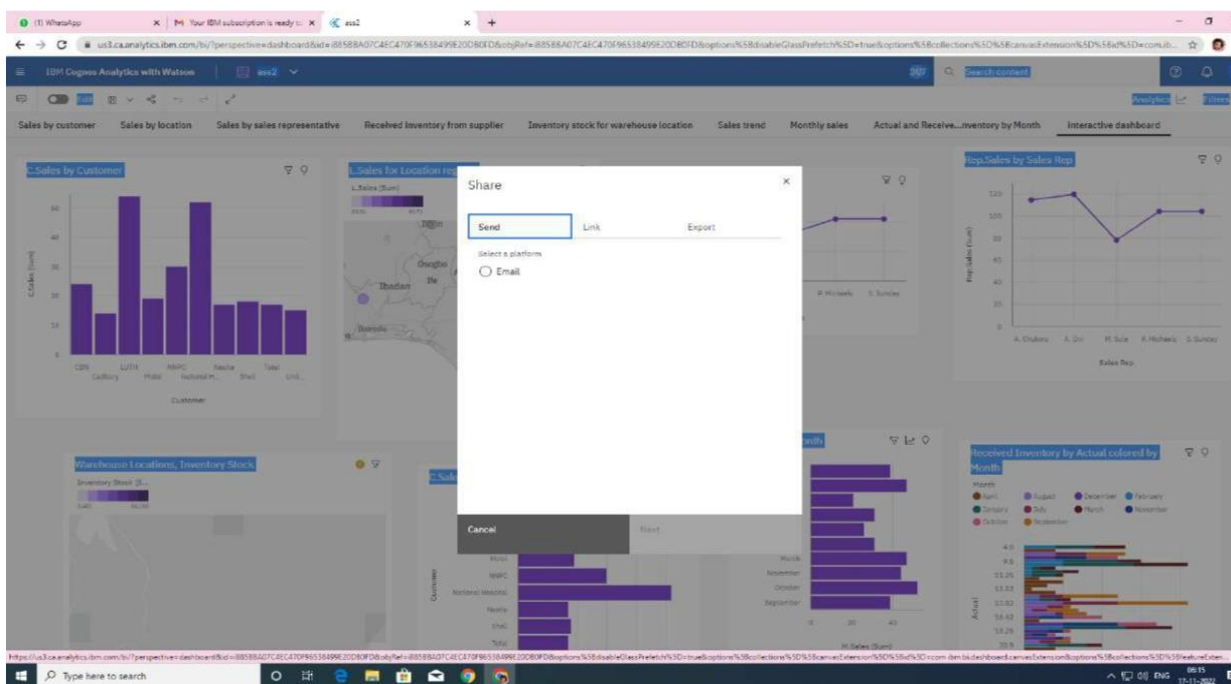
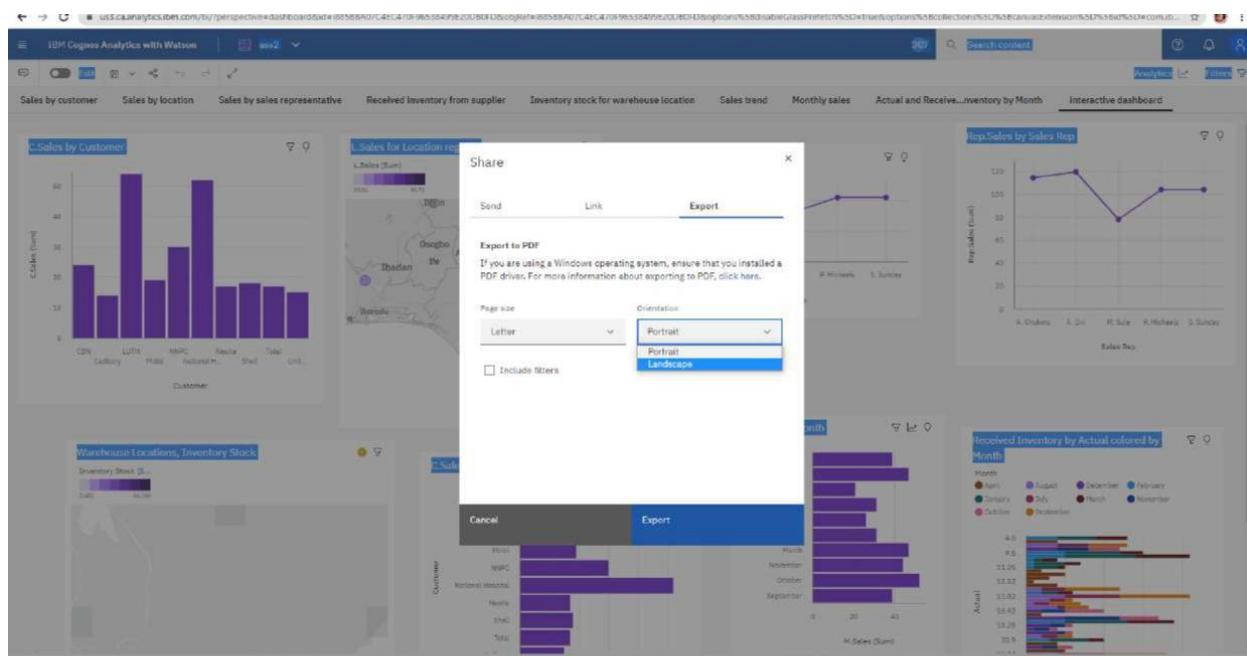
Creating The Dashboard

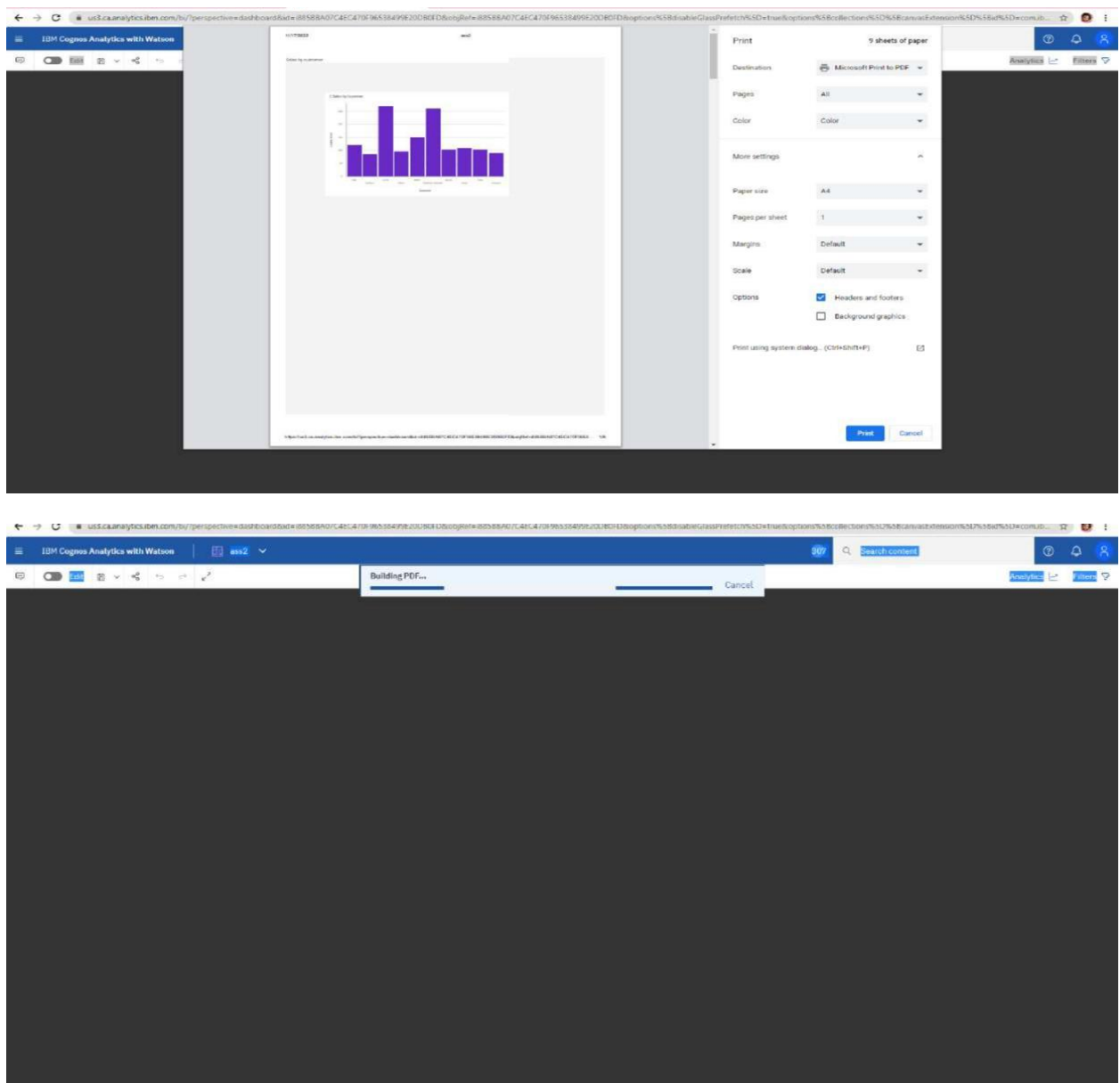
Once we've created views on different tabs in Cognos analytics, we can pull them into a dashboard.



Export The Analytics







We can export the analytics by clicking the share icon on the top left corner of the cognac screen.


Conclusion:

Many data sources operate in real time, producing data streams that can overwhelm data analysis pipelines. Moreover, there is often a desire to make decisions rapidly, perhaps also in real time. These temporal issues provide a particularly clear example of the need for further dialog between

statistical and computational researchers. Statistical research has rarely considered constraints due to real-time decision-making in the development of data analysis algorithms, and computational research has rarely considered the computational complexity of algorithms for managing statistical risk.

References:

- [1] Dhivya B H, Manjula R, Siva Bharathi S, Madhumathi R. A Survey on Crop Yield Prediction based on Agricultural Data, *International Journal of Innovative Research in Science, Engineering and Technology*. 2017; 6(3).
- [2] Jharna Majumdar, Sneha Naraseeyappa, Shilpa Ankalaki. Analysis of agriculture data using datamining techniques: application of big data. *Journal of Big data*. 2017.
- [3] Majumdar J, Ankalaki S. Comparison of clustering algorithms using quality metrics with invariant features extracted from plant leaves. *International Conference on Computational Science and Engineering*. 2016.
- [4] D Ramesh, B Vishnu Vardhan. Data Mining Techniques and Applications to Agricultural Yield Data. *International Journal of Advanced Research in Computer and Communication Engineering*. 2013; 2(9).
- [5] Swarupa Rani. The Impact of Data Analytics in Crop Management based on Weather Conditions. *International Journal of Engineering Technology Science and Research*. 2017; 4(5):299-308.
- [6] F K Van Evert, S Fountas, D Jakovetic, V Crnojevic, I Travlos, C Kempenaar. Big Data for weed control and crop protection. John Wiley & Sons Ltd on behalf of European Weed Research Society, 2017: 218–233.
- [7] Wu Fan, Chen Chong, Guo Xiaoling, Yu Hua. Prediction of crop yield using Big Data. 8th International Symposium on Computational Intelligence and Design. 2015.
- [8] Dakshayini Patil, M .S, Shirdhonkar. Rice Crop Yield Prediction using Data Mining Techniques: An Overview. *International Journal of Advanced Research in Computer Science and Software Engineering*, 2017; 7(5):427-431.
- [9] Dhivya B H, Manjula R, Siva Bharathi S, Madhumathi



R. A Survey on Crop Yield Prediction based on Agricultural Data, International Journal of Innovative Research in Science, Engineering and Technology. 2017; 6(3):4177-4182. [10] Yogesh Gandge, Sandhya. A Study on Various Data Mining Techniques for Crop Yield Prediction, International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques, IEEE, 2017;420-423 [11] R. Sujatha, P.Isakki Devi. A Study on Crop Yield Forecasting Using Classification Techniques, IEEE, 2016. [12] V. Sellam and E. Poovammal. Prediction of Crop Yield using Regression Analysis, Indian Journal of Science and Technology, 2016; 9(38). [13] Patricio Grassinia, Lenny G.J. van Bussel, Justin Van Warta, Joost Wolf, Lieven Claessens, d, Haishun Yanga, Hendrik Boogaarde, Hugo de Groote, Martin K. van Ittersumb, Kenneth G. Cassman. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crops Research. 2015; 49–63. [14] David B. Lobell, The use of satellite data for crop yield gap analysis, Field Crops Research-143, 2013; 56–64. [15] Martin K. van Ittersuma, Kenneth G. Cassmanb, Patricio Grassinib, Joost Wolfa, Pablo Tittone, Zvi Hochman. Yield gap analysis with local to global relevance-A review. Field Crops Research – 143, 2013; 4–17