

Yield Analysis for Telangana for years 2016 - 2018

Authors:

Bharati Panigrahi : bharatipanigrahi9901@gmail.com

Manasi Murarka : murarkamanasi@gmail.com

UMassAmherst

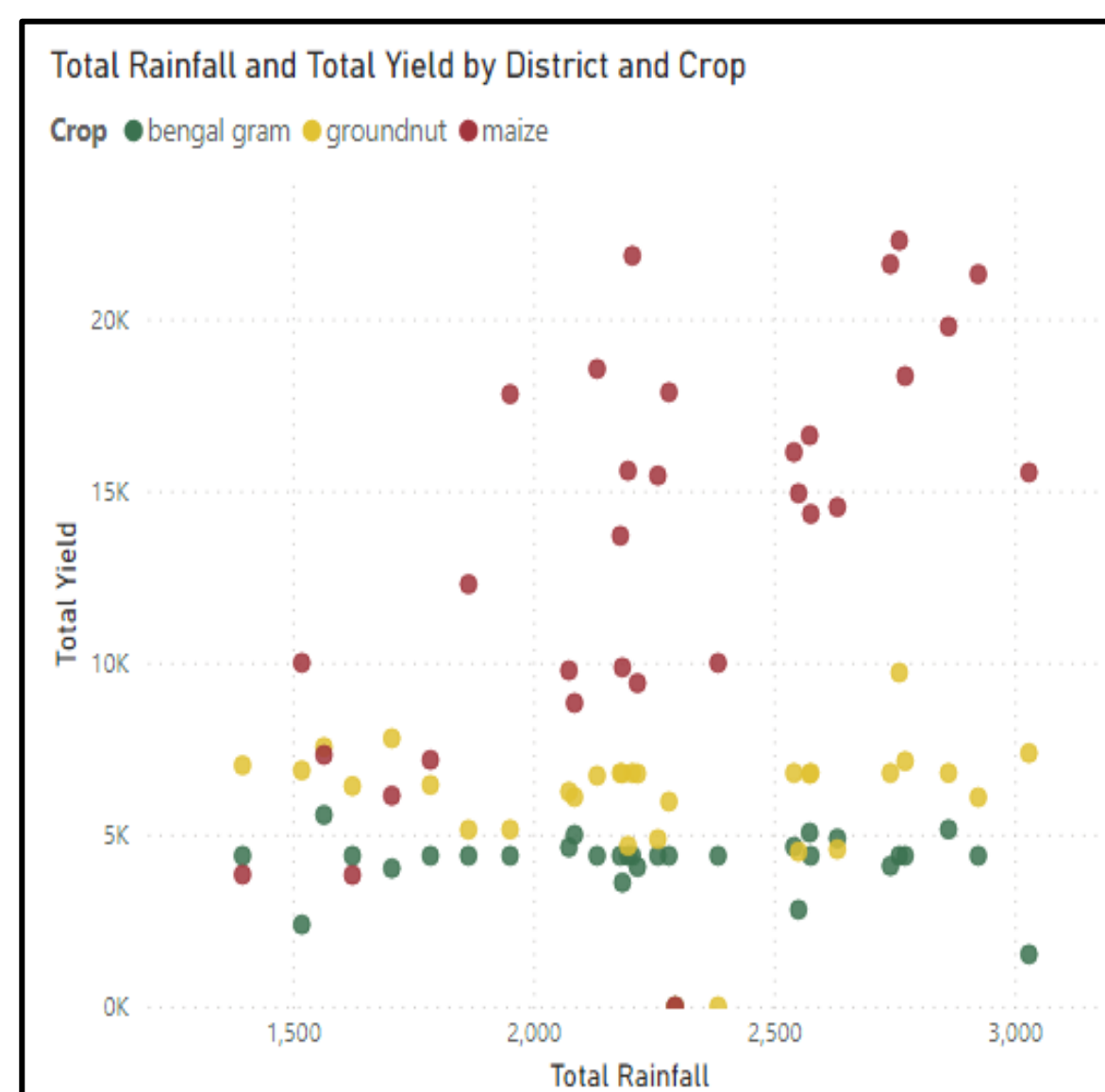
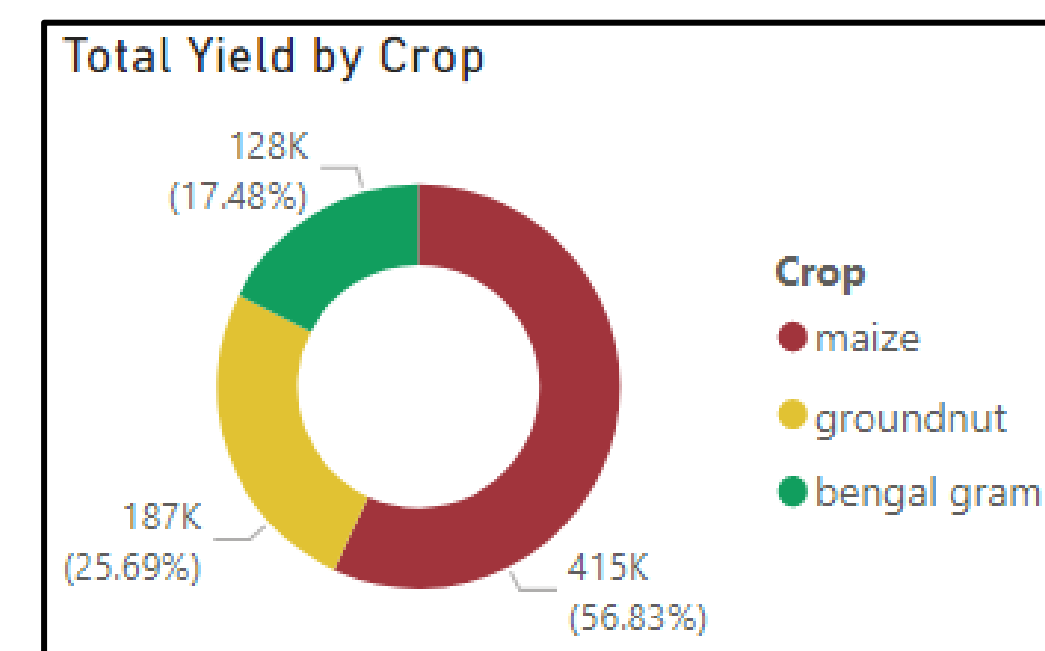
Data Analytics and Computational
Social Science Program

INTRODUCTION

Agriculture is a key activity of human being since it provides basic needs such as food, clothing and shelter. It has been demonstrated that every 1% increase in agricultural yield translates into a 0.6–1.2% decrease in the numbers of absolute poor households in the world.

In the state of Telangana, India throughout the year there 27 important crops that are grown, out of which our project focuses on Maize, Groundnut and Bengal Gram. It has been observed over the years that out of these three, Maize has the highest yield.

There are a variety of factors associated with crop yield and the risks involved with farming. The four most important factors that influence crop yield are soil fertility, availability of water, climate, and diseases or pests. When we compare the total rainfall with the total yield for each of the three crops, it is evident that there is correlation between rainfall and yield for Maize, whereas for Bengal Gram and Groundnut a significant relationship isn't visible. But correlation doesn't imply causation, so we wanted to see if could predict the yield accurately based on temperature and rainfall of the area.

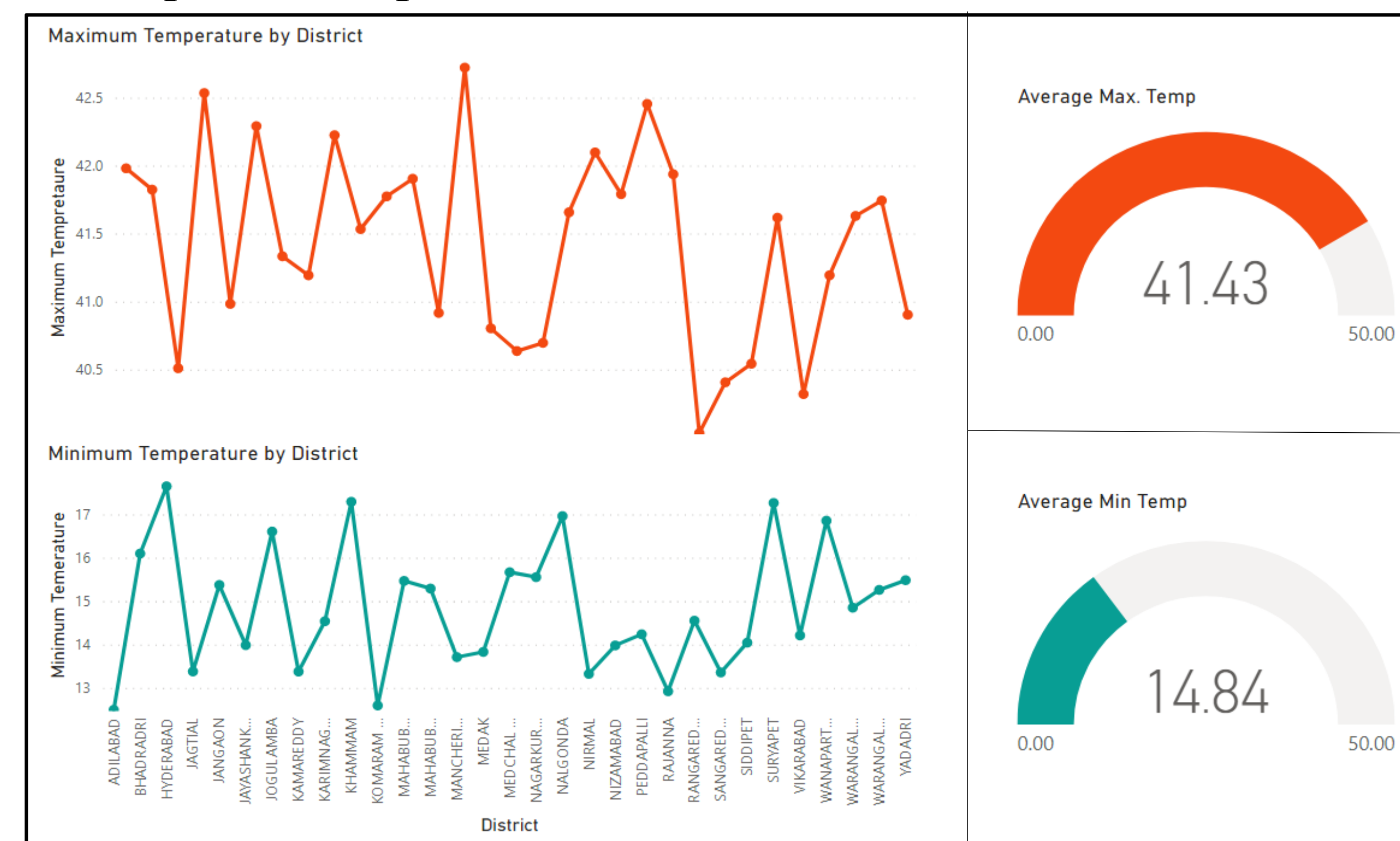


METHODOLOGY

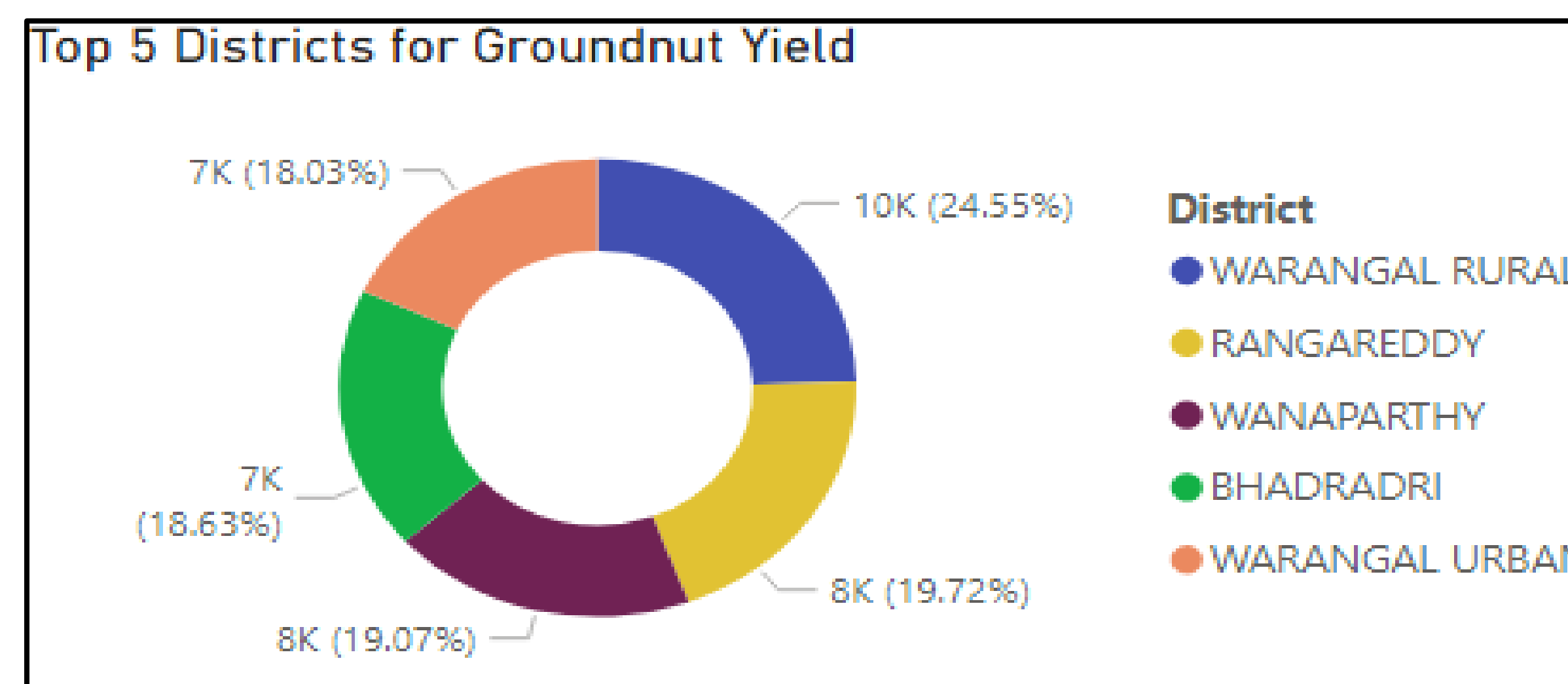
- STEP - 1 • Collect Data from various sources based on our problem statement.
- STEP - 2 • Combine and clean the datasets.
- STEP - 3 • Explore the dataset using visualizations tools such as Excel, Tableau and Power BI.
- STEP - 4 • Perform Principal Component Analysis (PCA) to identify the factor that has most impact on the yield.
- STEP - 5 • Perform Linear Regression on the dataset to predict total yield using Python.
- STEP - 6 • Optimize the model using XGBoost and Cross Validation to get better predictions.

RESULT

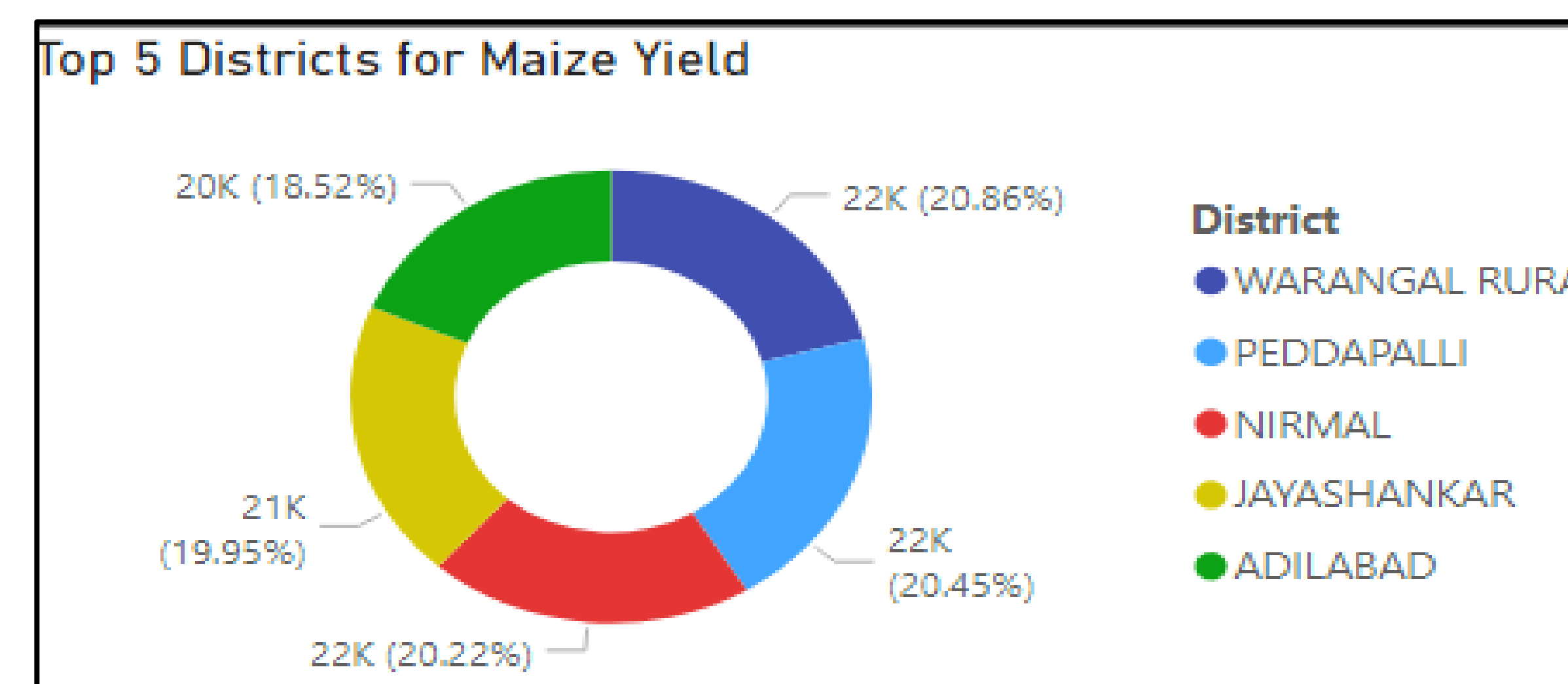
There are 31 districts in the state of Telangana for which we conducted an analysis based on the temperature range, rainfall and the 3 crops – Maize, Groundnut and Bengal Gram. The relation between the districts and temperature and the crop yield for each of the crop was as depicted below.



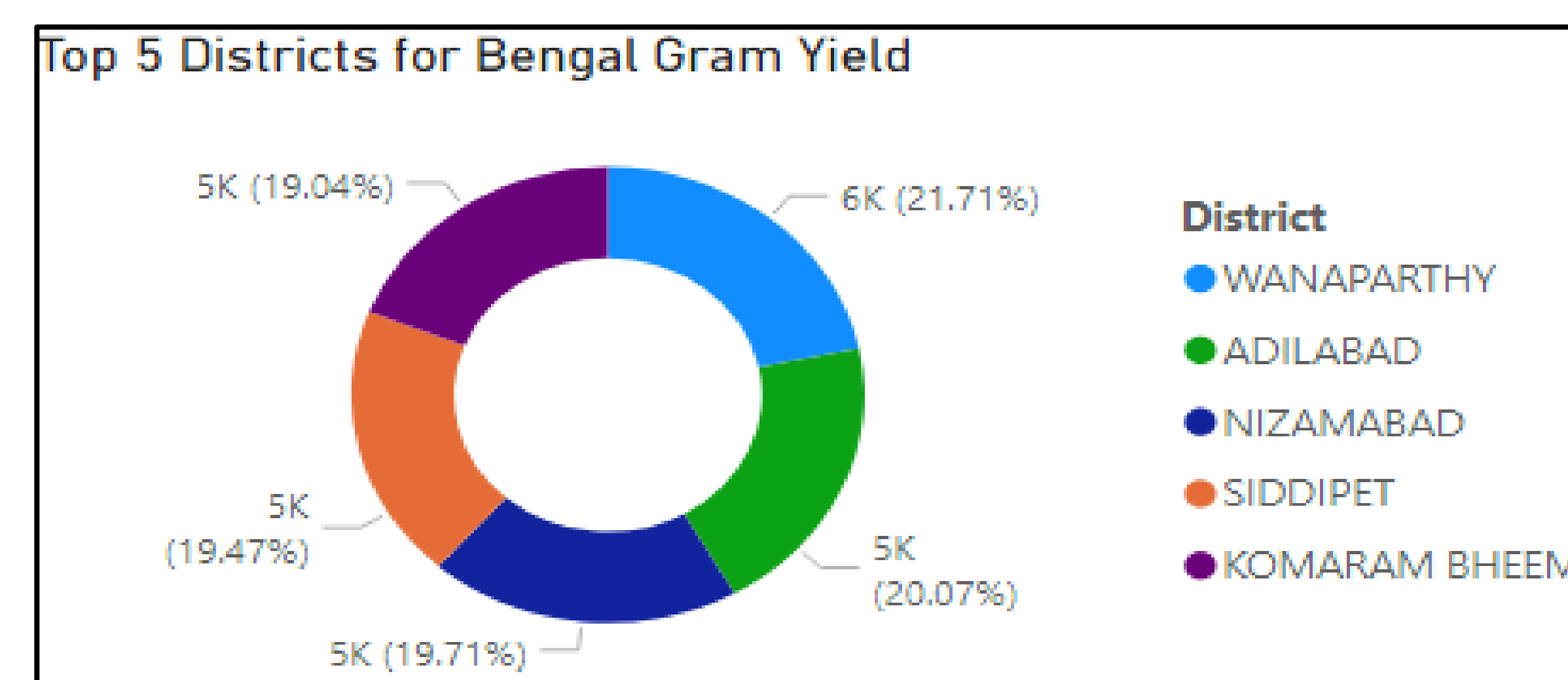
Graph 1: Average Maximum and Minimum Temperature of each district and Telangana state as a whole



Graph 2: Top 5 Districts with highest yield for Groundnut



Graph 3: Top 5 Districts with highest yield for Maize

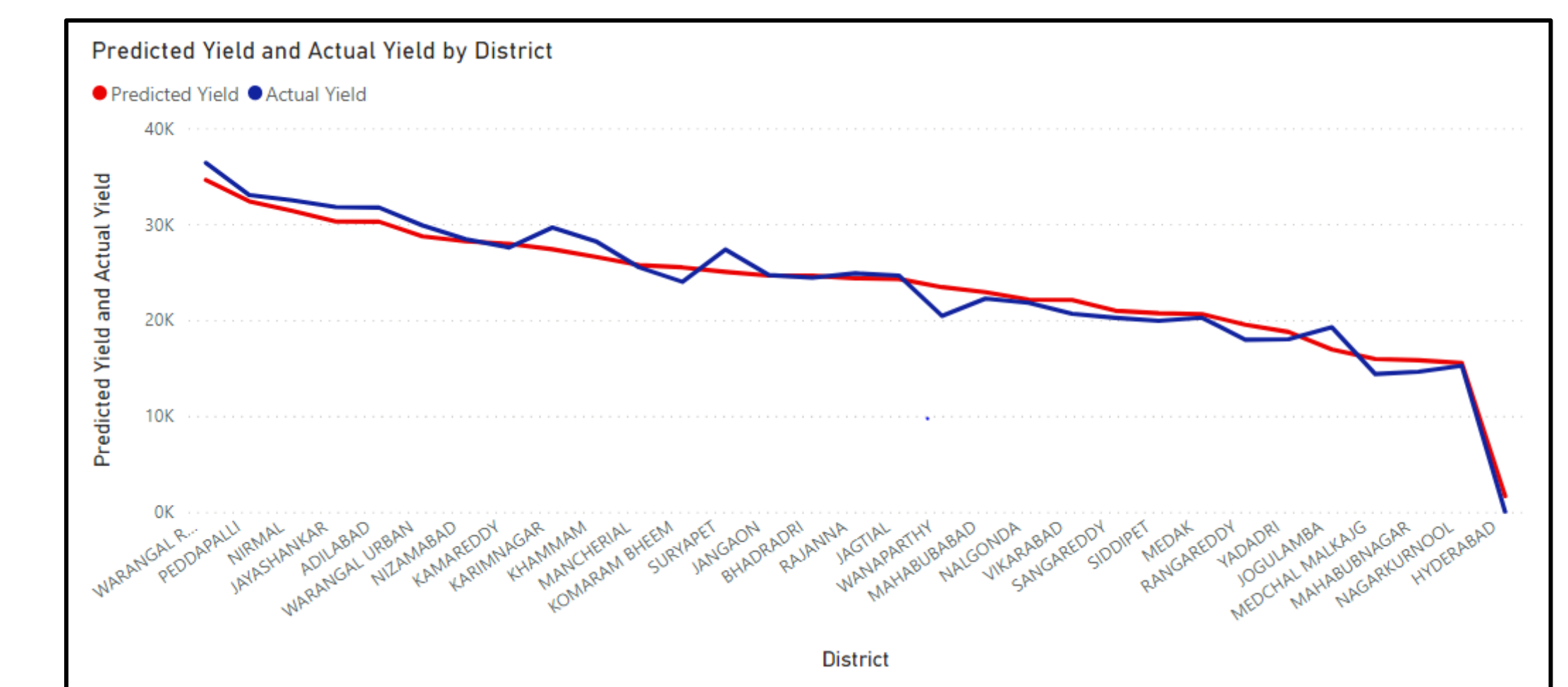


Graph 4: Top 5 Districts with highest yield for Bengal Gram

CONCLUSION

	0	1	2	3	Variability (%)	Cummulative Variability (%)
PC1	Total Rainfall	Max. Temp	Min Temp	Crop	0.468911	0.468911
PC2	Min Temp	Max. Temp	Total Rainfall	Crop	0.250953	0.719863
PC3	Crop	Total Rainfall	Max. Temp	Min Temp	0.181818	0.901682

Table 1 : Dimensionality Reduction using PCA to capture 85% or more variability



Graph 5 : Comparison of the Actual Yield and Predicted Yield using the model

POLICY RECOMMENDATION

Since, Telangana is not surrounded by any of the major water bodies, the major sources of water is the rainfall and the rivers flowing through the state. The first strategy and the actions relate to 'Rain Water Harvesting and Ground-water Conservation/Recharge'.

URBAN AREAS:

The ground water conservation in urban areas emphasizes on roof-top rainwater harvesting systems on buildings along with combined recharge systems to be implemented.

RURAL AREAS:

In respect of ground-water conservation in rural areas, the emphasis is more on rehabilitation of village ponds and reservoirs for water storage and conservation.

INDUSTRIAL AREAS:

In industrial areas, we should ensure careful implementation of the rain water harvesting and promote use of the ground water recharge techniques. The polluted effluent of industries be treated for its maximum reuse and recycle. Due to risk of ground-water pollution, 'recharge well method' should not be encouraged.

FUTURE SCOPE

- Here we had taken into consideration only two factors i.e., the temperature and rainfall of the districts. There are other factors such as the fertility and type of soil present in the area which would affect the crop yield. On proper analysis we can help find the soil type and fertilizers which can be helpful in maximizing and predicting the crop yield.
- P.C.A can be used on images of the field and crop to detect if they are infected by any disease or the presence of weed, which also affects the quality of the crop.



QR code for video link