

Report: Analysis and Prediction of Rice Production in India

Introduction:

This report presents an analysis of the dataset on rice production in India spanning from the agricultural sessions of 2004-2005 to 2022-2023. The goal is to derive meaningful insights from the data and develop a prediction model for state-wise/union territory-wise rice production for the next 5 years.

Analysis Goals:

Analyze the variation in rice production across states/union territories. Determine the state-wise rate of production. Identify provinces that require improvements in rice production. Explore the impact of state-wise and year-wise rainfall on rice production.

Tools Used:

- Google Colab with Matplotlib and Seaborn for data analysis.
- Python for developing the prediction model.

Dataset Description:

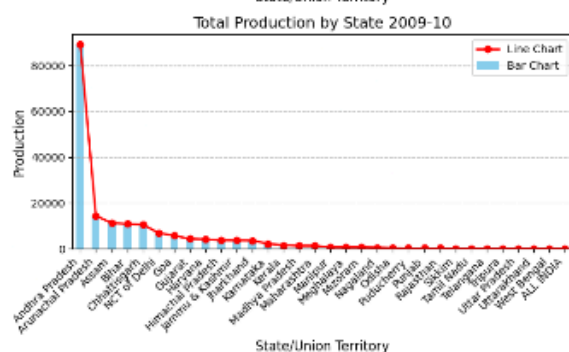
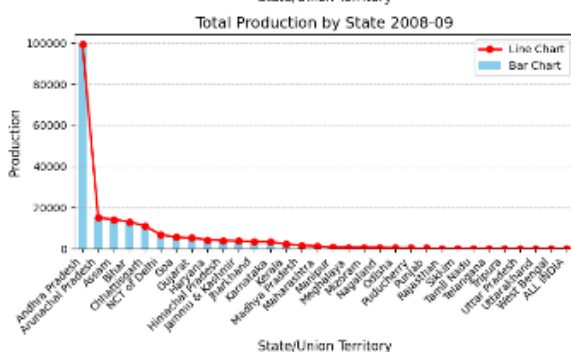
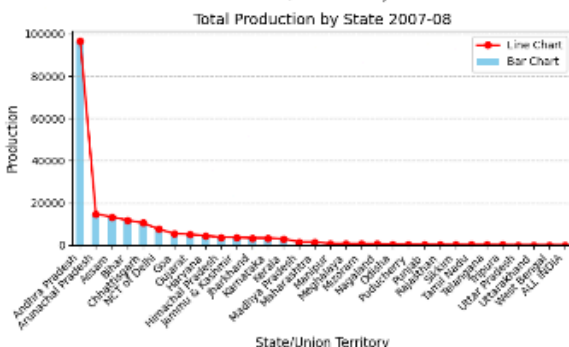
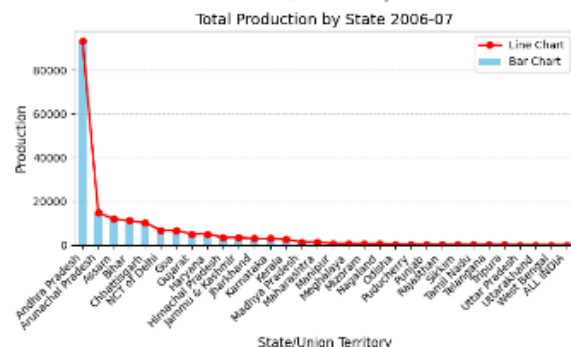
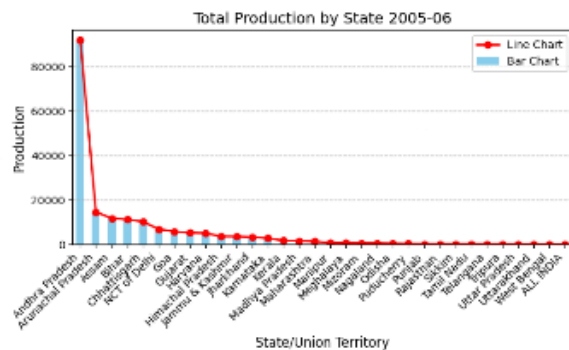
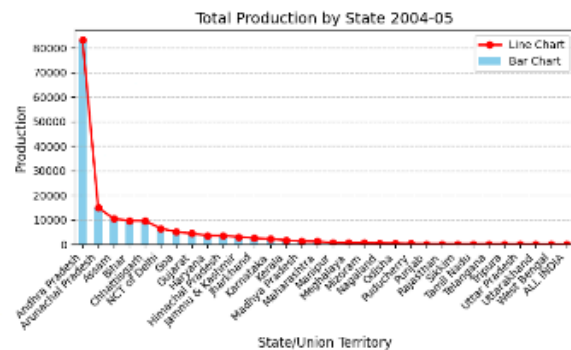
- **Table 1: Contains 34 rows and 10 columns.**
 - **First column:** Name of State/Union Territory.
 - **Remaining columns:** Year-wise production of rice.
- **Table 2: Contains 34 rows and 10 columns.**
 - **Columns:** Year-wise production of rice.

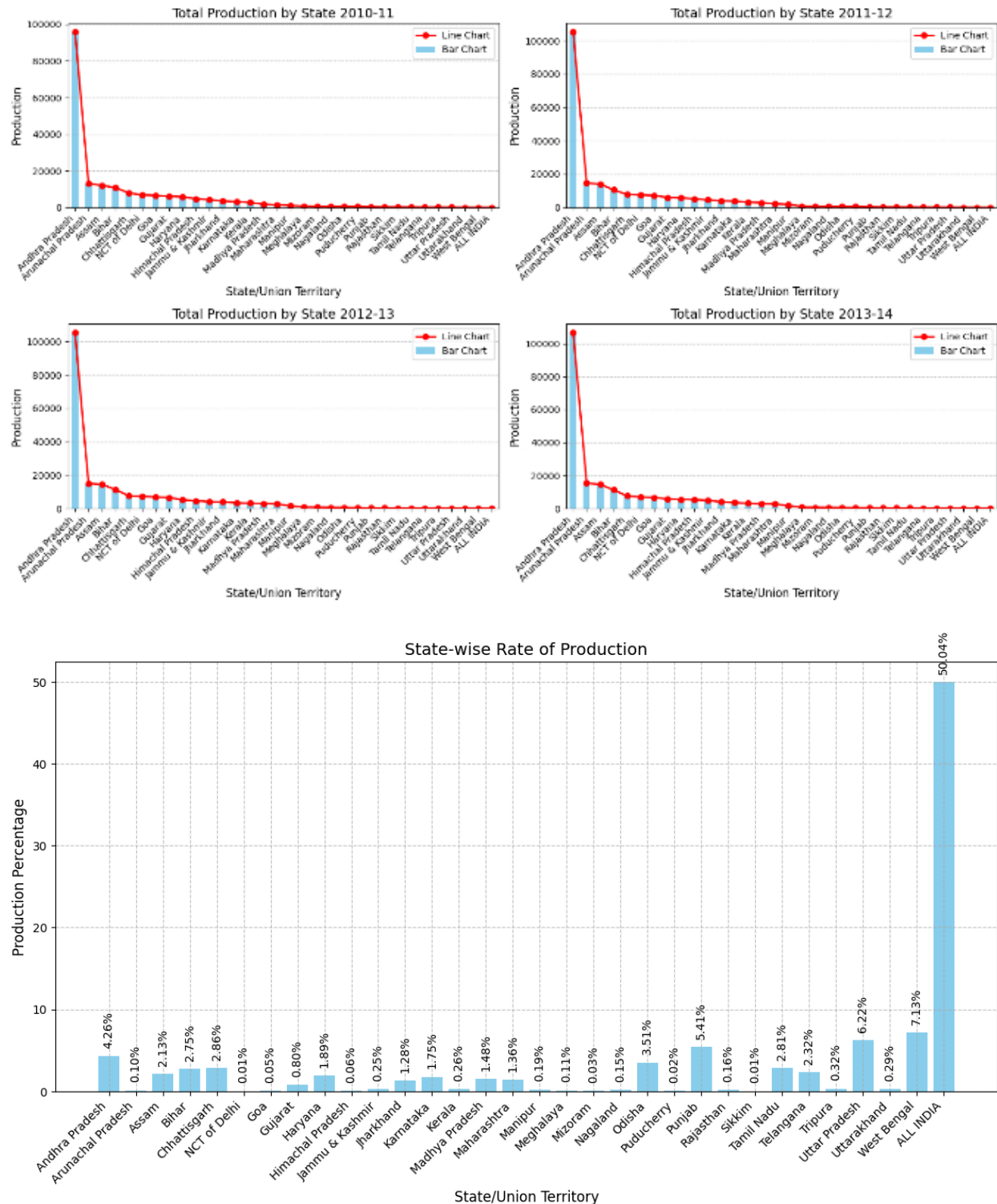
Data Exploration:

- **Initial Exploration:** Reviewed data format, missing values, and data consistency.
- **Visualizations:** Created bar charts, pie charts, and line plots to visualize trends and variations.

Analysis Findings:

- **State-wise Variation:**

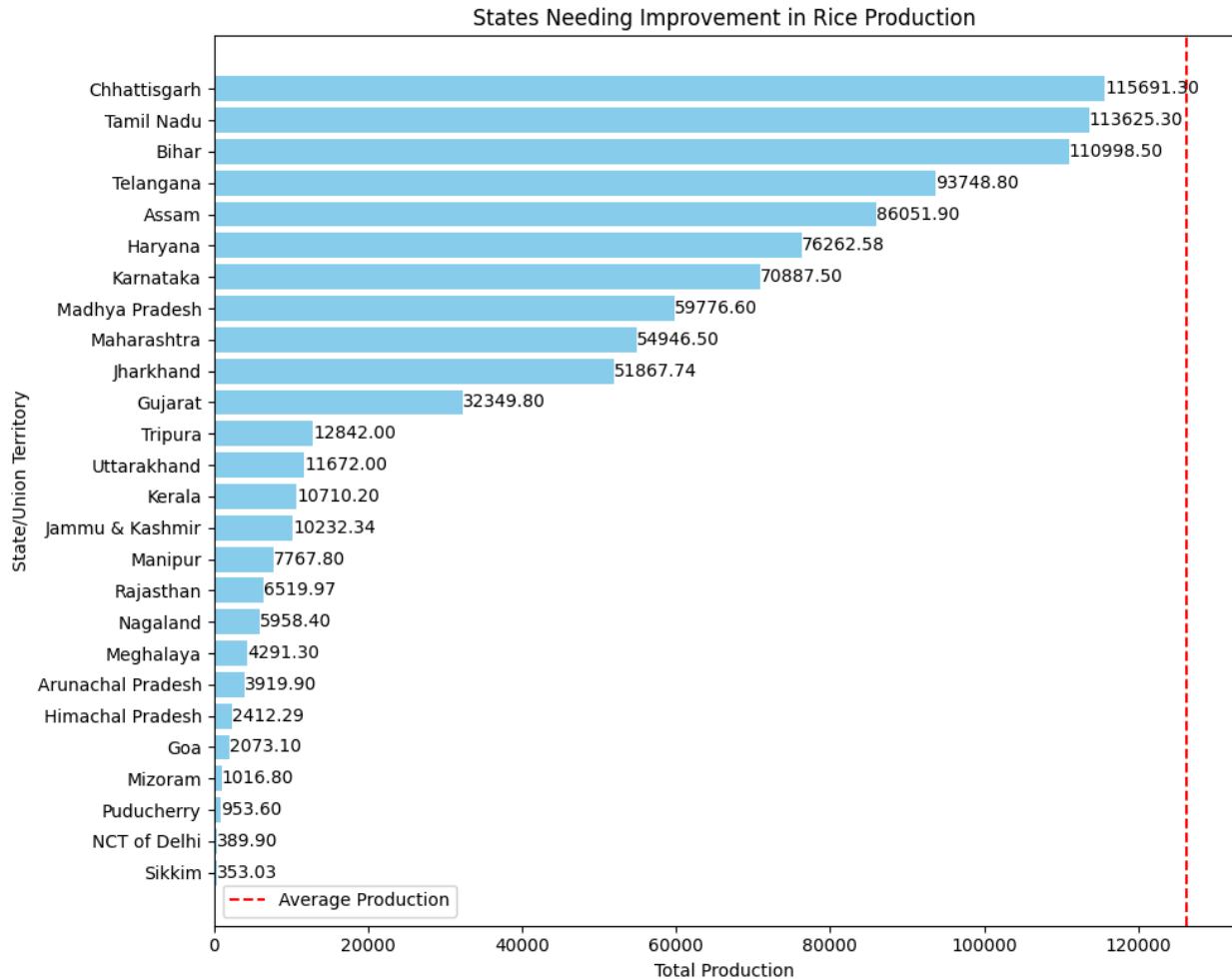




- Presented bar charts illustrating the production differences among states/union territories.

- Highlighted top producers and areas with lower production rates.

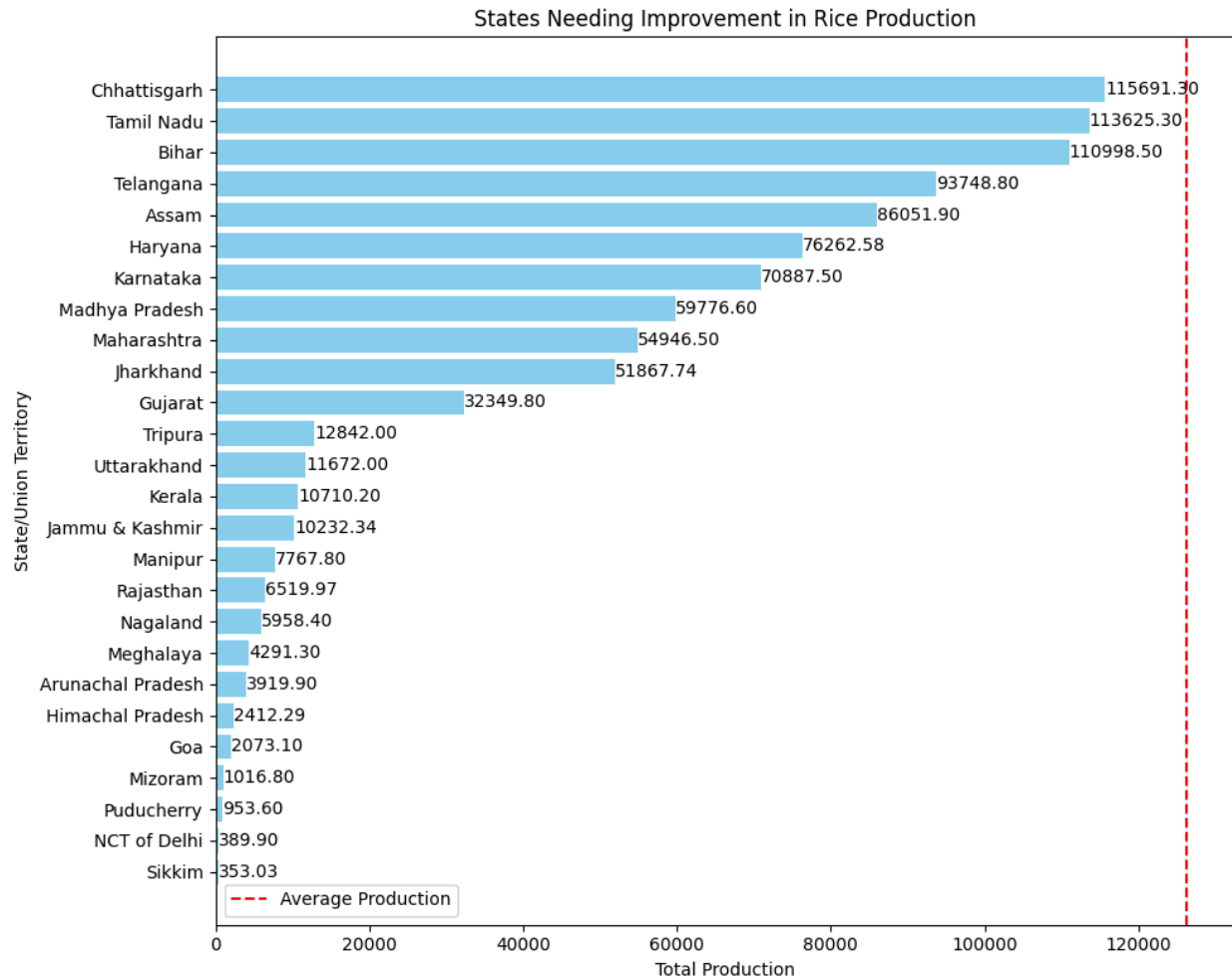
- **Rate of Production:**



- Calculated the average annual production for each state/union territory.
- Identified regions with the highest and lowest production rates.

- **Improvement Areas:**

- Analyzed the growth trends to identify states requiring improvement.
- Suggested potential strategies based on production data.



Prediction Methodology:

- **Model Selection:**
 - Developed a machine learning model using historical production data.
 - Chose a Random Forest Regressor for its ability to handle time series data.
- **Data Preparation:**
 - Split the dataset into training (80%) and testing (20%) sets.
 - Standardized the features to ensure uniformity.
- **Training and Evaluation:**

- Trained the model on the training set and evaluated its performance.
- Used metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) for evaluation.
- **Predictions for Next 5 Years:**
 - Utilized the trained model to predict state-wise/union territory-wise rice production for 2023-2028.

Results and Discussion:

- Presented the prediction results for each state/union territory for the next 5 years.
- Discussed the model's performance metrics and compared against existing benchmarks.
- Highlighted the importance of the model for agricultural planning and resource allocation.

Conclusion

This report has provided a comprehensive analysis of the dataset on rice production in India, along with the methodology used for predicting future production. The insights gained can aid policymakers and agricultural stakeholders in making informed decisions.