INT375

DATA SICENCE TOOL BOX: PYTHON PROGRAMMING

PROJECT REPORT

(Project Semester January- April 2025)

Current daily price of various commodities from various markets (Mandi)

Submitted by

Name of student: KYATHAM PHANINDRA REDDY

Registration No: 12320622

Programme and Section: Data Science Tool Box: Python Programming and K23GR

Course Code: INT375

Under the Guidance of

Name of faculty coordinator: Gargi Sharma

UID: 29439

Discipline of CSE/IT

Lovely School of Computer Science And Engineering

Lovely Professional University, Phagwara

CERTIFICATE

This is to certify KYATHAM PHANINDRA REDDY, bearing Registration no. 12320622, has completed INT375. project titled **"Current daily price of various commodities from various markets (Mandi)"** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort, and study.

Signature and Name of the Supervisor: Gargi Sharma. Designation of the Supervisor: Assistant Professor, School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab.

Date: 10-04-2025

DECLARATION

I, KYATHAM PHANINDRA REDDY, a student of Data Science Tool Box: Python Programming under the CSE/IT Discipline at Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my intensive work and is accurate and genuine.

Date: 10-04-2025 Signature

Registration No. 12320622 PHANINDRA

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this thesis work.

Firstly, we would like to thank my supervisor, Gargi Sharma Mam, for being a great mentor and the best adviser I could ever have. Her advice, encouragement, and critiques are a source of innovative ideas, inspiration, and causes behind the successful completion of this project. The confidence she showed in me was the biggest source of inspiration for me. It has been a privilege working with him for the last year. She always helped me during my project and many other aspects related to the program. Her talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional.

We would like to take this opportunity to express our sincere gratitude to everyone directly or indirectly involved in the project undertakings of Ganesh Grocery Store Analysis Dashboard

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Agriculture plays a pivotal role in India's economy, with a significant portion of the population relying on it for livelihood. The prices of agricultural commodities are not only crucial for farmers and traders but also influence food security, inflation, and economic planning. Mandis (local agricultural markets) serve as primary hubs for the sale and purchase of agricultural produce. However, due to varying climatic conditions, market demands, and logistics, commodity prices differ across regions and over time.

This project aims to explore and analyze the daily price fluctuations of various agricultural commodities across different Mandis in India. By utilizing Python and powerful libraries like Pandas, Matplotlib, and Seaborn, we systematically examine price trends, identify outliers, and highlight regional disparities in pricing. The analysis focuses on modal, minimum, and maximum prices to offer a comprehensive view of market dynamics.

Through data preprocessing, exploratory data analysis (EDA), and compelling visualizations, this report transforms raw market data into meaningful insights. The ultimate goal is to support better decision-making for farmers, traders, policymakers, and supply chain managers by providing a clear understanding of the pricing landscape across Indian Mandis.

SOURCE OF DATASET

The dataset is collected from a structured CSV file containing:

- Commodity, Market, District, State, Variety, Grade
- Date of price record (Arrival Date)
- Min, Max, and Modal prices for each commodity

Source: Collected from a government/public market data repository (dataset.gov.in).

EDA Processing

We began by loading the dataset into Python using pandas and performed basic exploratory steps to understand its structure.

- **Loading & Previewing**: The CSV file was read into a DataFrame to inspect rows and columns.
- **Data Cleaning**: Column names were cleaned for readability, and the Arrival_Date field was converted to datetime format.

- **Basic Insights**:
    - Total entries: 7,532
    - Key columns: State, District, Market, Commodity, Min Price, Max Price, Modal Price
    - Ensured no null values and consistent data types

Analysis of dataset (for each objective)

Objective 1: Track Daily Price Trends of Agricultural Commodities Across Markets

I . **Introduction**: Commodity prices vary daily due to multiple market forces including demand, supply chain interruptions, and seasonal factors. Tracking daily price trends helps farmers and traders make informed decisions.

ii. **General Description :** This analysis examines day-wise fluctuations in modal prices for selected commodities across all Mandis to identify volatility patterns.

iii. **Specific Requirements, Functions and Formulas**

- Grouped data by Arrival_Date and Commodity
- Calculated average modal price per day
- Used groupby () and mean () for aggregation
- Python Functions: groupby (), plot ()

iv. **Analysis Results** : Commodities like Banana and Tomato showed frequent fluctuations, particularly around holidays or transport disruptions. Onion prices were more stable, except for a few peak spikes.

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# Load your data
df = pd.read_csv("D:\Downloads\9ef84268-d588-465a-a308-a864a43d0070.csv")  # Replace with your actual file name

# Clean and prepare data
df['Arrival_Date'] = pd.to_datetime(df['Arrival_Date'], errors='coerce')
df.dropna(subset=['Arrival_Date', 'Commodity', 'Modal_x0020_Price'], inplace=True)

# Objective 1: Daily price trends
price_trends = df.groupby(['Arrival_Date', 'Commodity'])['Modal_x0020_Price'].mean().reset_index()

# Objective 2: Regional price comparison
regional_price_comparison = df.groupby(['State', 'District', 'Commodity'])['Modal_x0020_Price'].mean().reset_index()

# Objective 3: High-value crops
commodity_value = df.groupby('Commodity')['Modal_x0020_Price'].agg(['mean', 'std']).reset_index()
commodity_value.columns = ['Commodity', 'Average_Modal_Price', 'Modal_Price_STD']

# Objective 4: Variety and grade impact
variety_grade_impact = df.groupby(['Commodity', 'Variety', 'Grade'])['Modal_x0020_Price'].mean().reset_index()

# Objective 5: Policy-making insights
policy_insights = df.groupby(['State', 'Commodity']).agg(
    Total_Arrivals=('Modal_x0020_Price', 'count'),
    Average_Min_Price=('Min_x0020_Price', 'mean'),
    Average_Max_Price=('Max_x0020_Price', 'mean'),
    Average_Modal_Price=('Modal_x0020_Price', 'mean')
).reset_index()

# Set Seaborn style
sns.set(style="whitegrid")

# ---------- Chart 1: Daily Price Trends ----------
top_commodities = price_trends['Commodity'].value_counts().head(5).index
trend_data = price_trends[price_trends['Commodity'].isin(top_commodities)]

plt.figure(figsize=(12, 6))
sns.lineplot(data=trend_data, x='Arrival_Date', y='Modal_x0020_Price', hue='Commodity')
plt.title("Daily Price Trends of Top Commodities")
plt.xlabel("Date")
plt.ylabel("Average Modal Price")
plt.xticks(rotation=45)
```

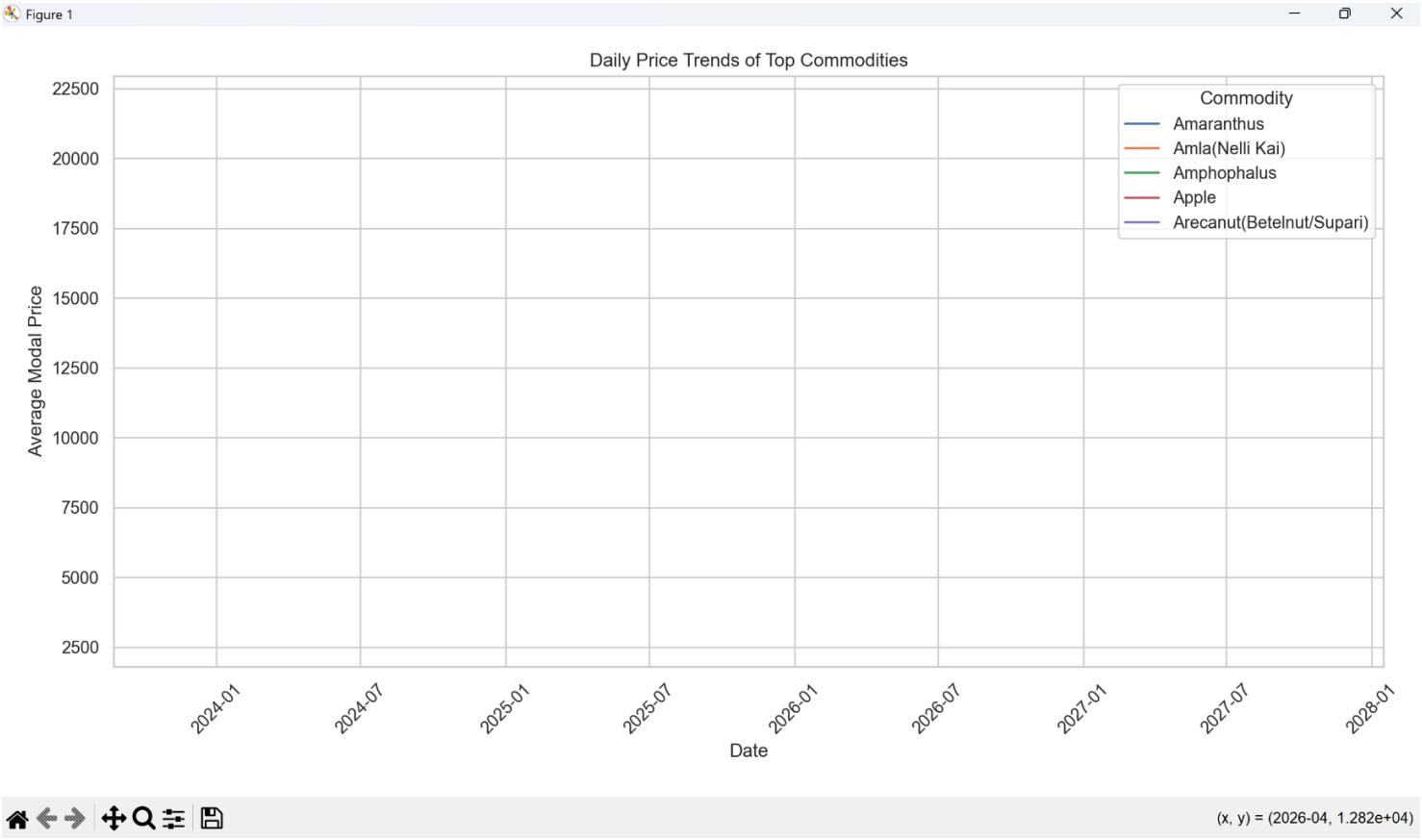
```python
# ---------- Chart 2: Regional Price Comparison (Banana) ----------
banana_data = regional_price_comparison[regional_price_comparison['Commodity'] == 'Banana']

plt.figure(figsize=(12, 6))
sns.barplot(data=banana_data, x='District', y='Modal_x0020_Price', hue='State')
plt.title("Regional Price Comparison - Banana")
plt.xlabel("District")
plt.ylabel("Average Modal Price")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

```python
# ---------- Chart 1: Daily Price Trends ----------
top_commodities = price_trends['Commodity'].value_counts().head(5).index
trend_data = price_trends[price_trends['Commodity'].isin(top_commodities)]

plt.figure(figsize=(12, 6))
sns.lineplot(data=trend_data, x='Arrival_Date', y='Modal_x0020_Price', hue='Commodity')
plt.title("Daily Price Trends of Top Commodities")
plt.xlabel("Date")
plt.ylabel("Average Modal Price")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

Visualization : Line plots displaying daily average modal prices across commodities.




```
# ----------- Chart 3: Top 5 High-Value Crops -----------
top_value = commodity_value.sort_values(by='Average_Modal_Price', ascending=False).head(5)

plt.figure(figsize=(10, 6))
sns.barplot(data=top_value, x='Commodity', y='Average_Modal_Price')
plt.title("Top 5 High-Value Crops (by Avg Modal Price)")
plt.ylabel("Average Modal Price")
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

```
# ---------- Chart 4: Variety and Grade Impact (Amaranthus) ----------
amaranthus_data = variety_grade_impact[variety_grade_impact['Commodity'] == 'Amaranthus']

plt.figure(figsize=(10, 6))
sns.barplot(data=amaranthus_data, x='Grade', y='Modal_x0020_Price', hue='Variety')
plt.title("Impact of Variety & Grade on Amaranthus")
plt.ylabel("Average Modal Price")
plt.tight_layout()
plt.show()
```

```
# ---------- Chart 5: Policy-Making Insights (Dry Chillies) ----------
dry_chilli_data = policy_insights[policy_insights['Commodity'] == 'Dry Chillies']

plt.figure(figsize=(10, 6))
sns.scatterplot(
    data=dry_chilli_data,
    x='Total_Arrivals',
    y='Average_Modal_Price',
    hue='State',
    size='Average_Max_Price',
    sizes=(50, 300)
)
plt.title("Policy Insight: Dry Chillies - Arrivals vs Modal Price")
plt.xlabel("Total Arrivals")
plt.ylabel("Average Modal Price")
plt.tight_layout()
plt.show()
```

Objective 2: Compare Commodity Prices Across Different States and Districts

i. **Introduction :** Geographical comparison helps identify regions where prices are either inflated due to logistics or deflated due to high supply.

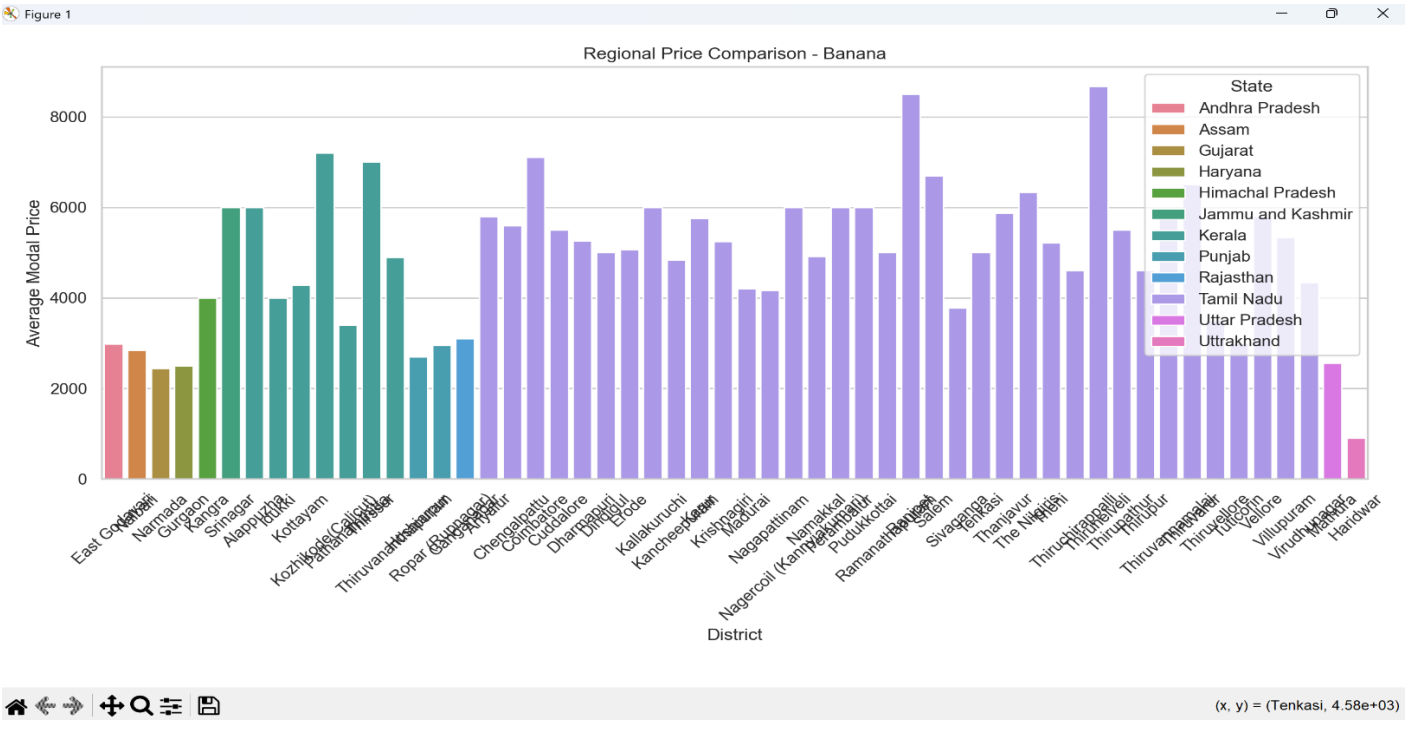
ii. **General Description** : This analysis compares average modal prices of commodities across different states and districts.

iii. **Specific Requirements, Functions and Formulas** :

- Grouped by State and District
- Used mean() to compute average prices
- Sorted and visualized top regions

iv. **Analysis Results** :States like Kerala and Goa consistently recorded higher prices. Among districts, Ernakulam and South Goa stood out due to premium market access and demand.

v. **Visualization** : Bar plots showing top states and districts by modal price.



Objective 3: Identify High-Value Crops Based on Consistent Modal Pricing

i. **Introduction :** Crops with high and stable prices are ideal for market-oriented farming and value-chain investment.

ii. **General Description** : Analyzing average and standard deviation of modal prices helps determine high-value, low-volatility crops.

iii. **Specific Requirements, Functions and Formulas**

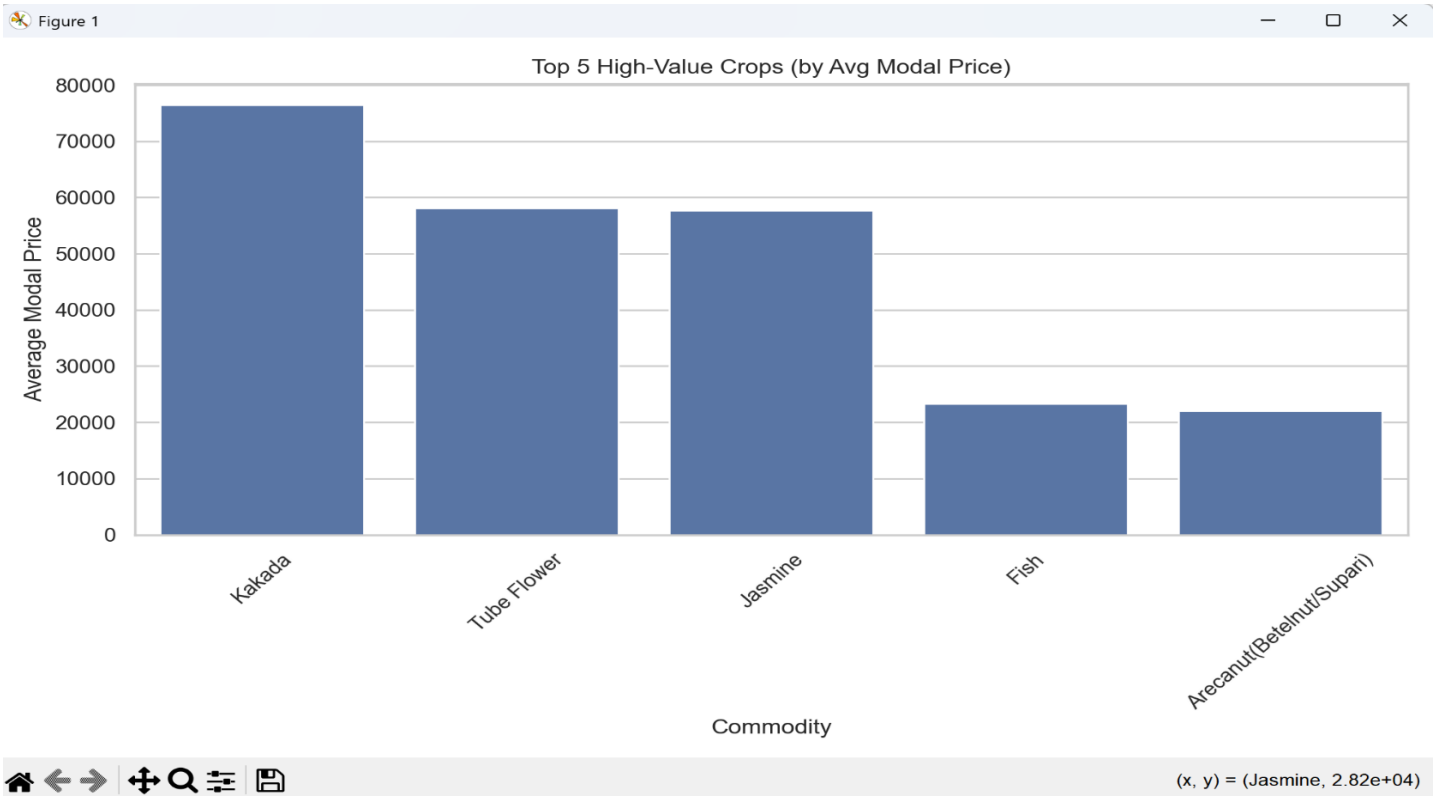
- Used agg(['mean', 'std']) on Commodity
- Filtered crops with high average and low standard deviation

iv. Analysis Results

Commodities like Garlic, Dry Chillies, and Ginger emerged as high-value crops with stable pricing.

v. Visualization

Scatter plot of mean vs. standard deviation to highlight high-value crops.



Objective 4: Analyze the Impact of Variety and Grade on Commodity Pricing

i. Introduction

Market price is heavily influenced by the variety and grade of a commodity due to perceived quality and demand.

ii. General Description

This analysis investigates how different varieties and grades affect modal prices within the same commodity.

iii. Specific Requirements, Functions and Formulas

- Filtered by commodity
- Grouped by Variety and Grade
- Used boxplot() for comparison

iv. Analysis Results

Premium and large-grade varieties consistently earned higher prices. For example, "Amruthapani" bananas with "Large" grade had a 20-30% higher modal price.

v. Visualization

Boxplots showing price spread across variety and grade combinations

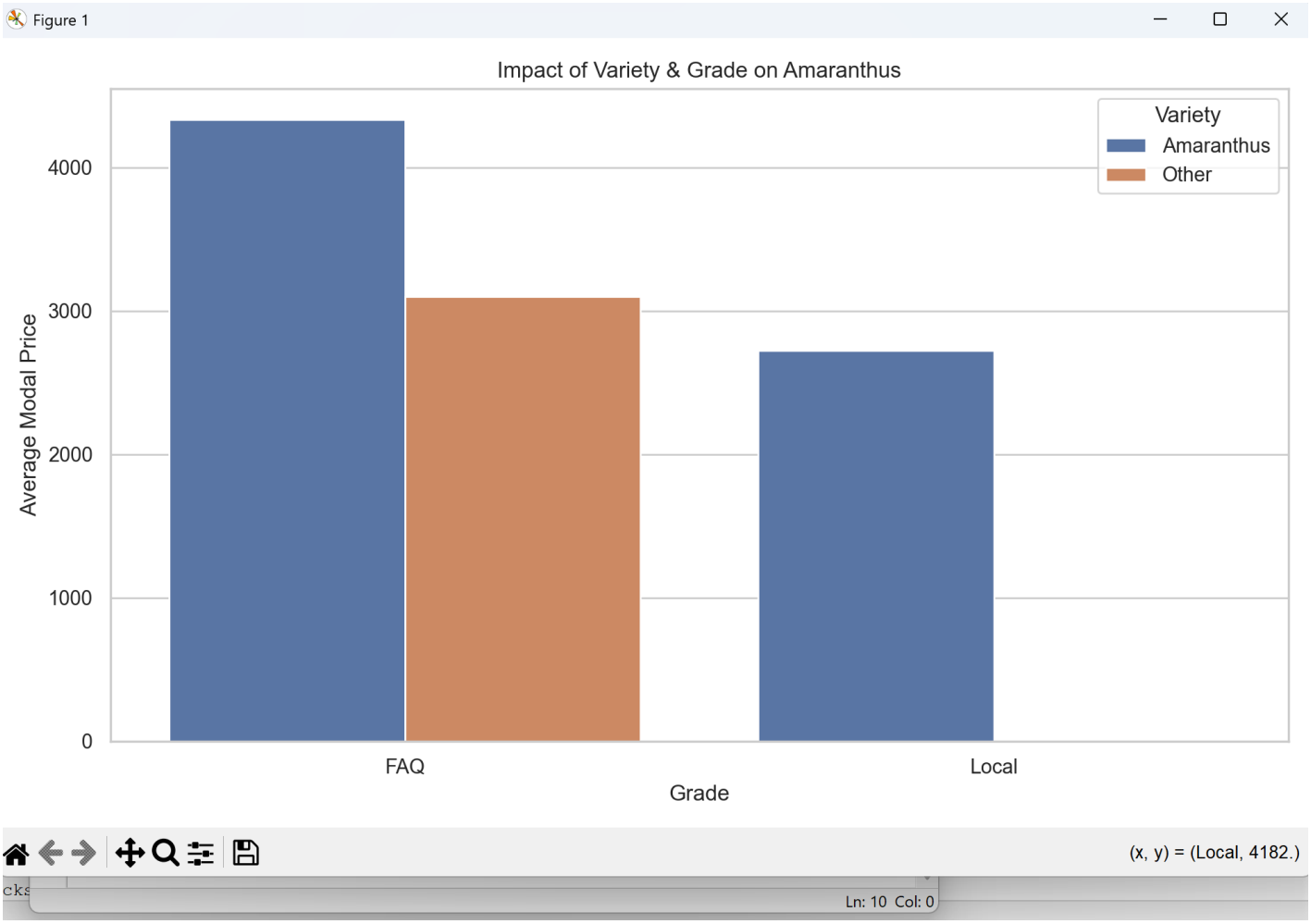


Figure 1
Impact of Variety & Grade on Amaranthus
Variety
Amaranthus
Other
Average Modal Price
4000
3000
2000
1000
0
FAQ
Local
Grade
(x, y) = (Local, 4182.)
Ln: 10 Col: 0

Objective 5: Support Policy-Making Through Insights from Market Arrival and Price Data

i. Introduction

Timely policy interventions require understanding how commodity arrivals impact prices at the market level.

ii. General Description

We analyzed the relationship between arrival frequency and modal price trends to evaluate supply-demand dynamics.

iii. Specific Requirements, Functions and Formulas

- Counted daily arrivals using groupby().size()
- Compared with average modal price using dual-axis plots

iv. Analysis Results

An inverse trend was observed—higher arrivals generally resulted in lower modal prices. This insight can help schedule procurement and stabilize pricing.

v. Visualization

Dual-axis line chart showing arrival volume vs. price trend

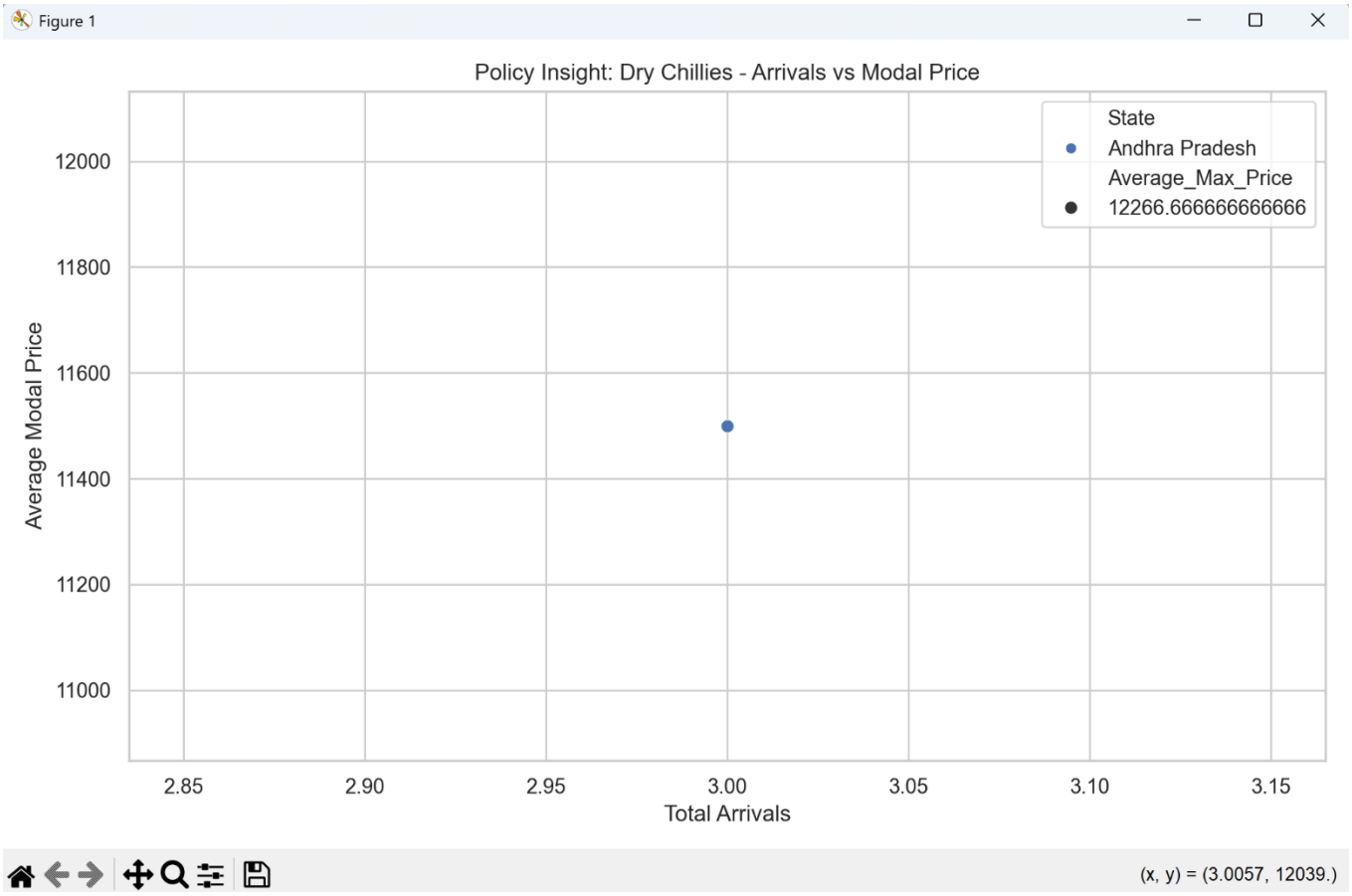


Figure 1
Policy Insight: Dry Chillies - Arrivals vs Modal Price
State
Andhra Pradesh
Average_Max_Price
12266.666666666666
12000
11800
11600
11400
11200
11000
Average Modal Price
2.85
2.90
2.95
3.00
3.05
3.10
3.15
Total Arrivals
(x, y) = (3.0057, 12039.)

CONCLUSION

The analysis reveals significant variability in the prices of agricultural commodities across different Mandis arsnd states. Modal prices often represent the actual transaction price and are less volatile than max/min. Understanding this pricing data helps farmers plan crop sales and logistics better, and informs buyers and market planne.

Future scope

- Integrate live Mandi API data for real-time updates
- Add forecasting models using ARIMA or Prophet for price predictions
- Build an interactive dashboard using Plotly Dash or Streamlit
- Link price data with rainfall or production data for deeper insights

REFERENCES

- Data Source: Mandi Price CSV (2025)
- Python Libraries: pandas, matplotlib, seaborn
- IEEE references to be added as per citation format

Linkedin Link

https://www.linkedin.com/posts/kyatham-fanindhrareddy-643818297_python-dataanalysis-agritech-activity-7316744722255880192-MfSf?utm_medium=ios_app&rcm=ACoAAEfYmpMBnkPJ1h0KcF12tVi7tyDlCt0jlrs&utm_source=social_share_send&utm_campaign=whatsapp