In [381…
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

In [382…
```
df =pd.read_csv("mandi.csv")
```

In [383…
```
df
```

Out[383…

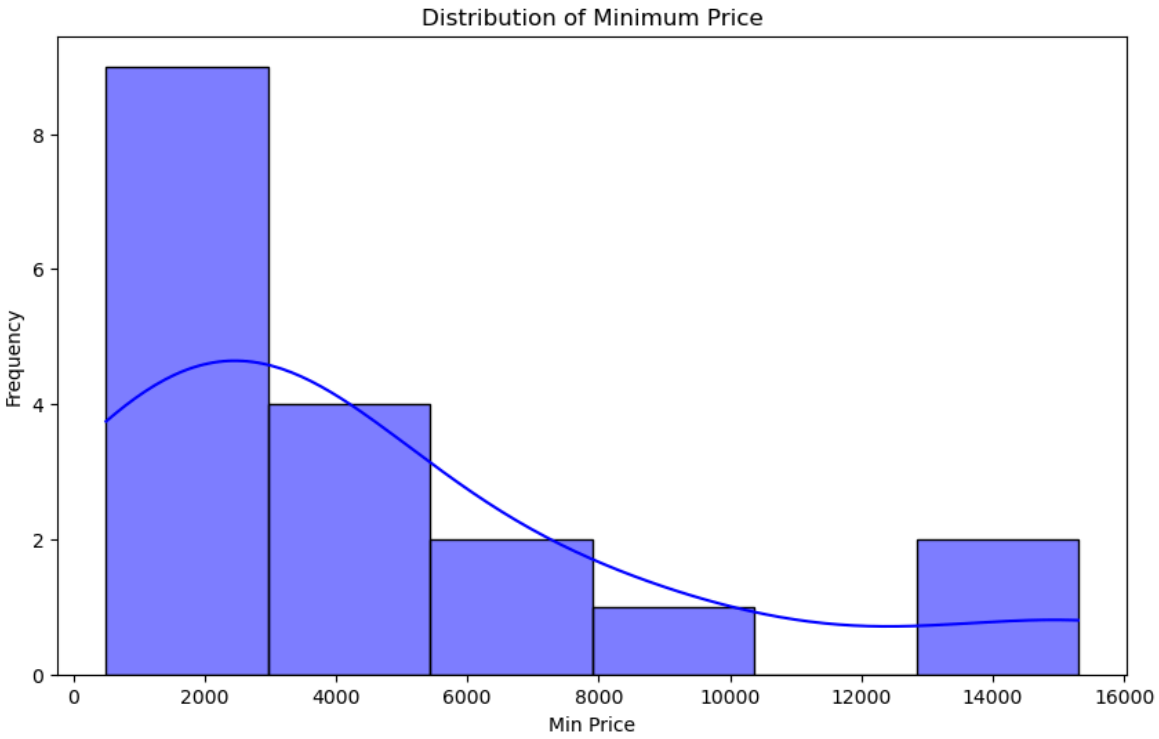
	State	District	Market	Commodity	Variety	Grade	Arrival_Da
0	Haryana	Faridabad	Faridabad	Cucumbar(Kheera)	Other	FAQ	16/03/202
1	Haryana	Faridabad	Faridabad	Grapes	Other	Large	16/03/202
2	West Bengal	Puruliya	Kasipur	Masur Dal	Masur Dal	FAQ	16/03/202
3	West Bengal	Puruliya	Kasipur	Onion	Other	FAQ	16/03/202
4	West Bengal	Puruliya	Kasipur	Rice	Other	FAQ	16/03/202
5	Haryana	Faridabad	Faridabad	Orange	Other	Large	16/03/202
6	Haryana	Faridabad	Faridabad	Water Melon	Other	Large	16/03/202
7	West Bengal	Puruliya	Kasipur	Brinjal	Other	FAQ	16/03/202
8	West Bengal	Puruliya	Kasipur	Potato	Jyoti	FAQ	16/03/202
9	Haryana	Faridabad	Faridabad	Grapes	Other	Large	16/03/202
10	Telangana	Khammam	Kothagudem	Cotton	Cotton (Unginned)	FAQ	16/03/202
11	West Bengal	Puruliya	Kasipur	Brinjal	Other	FAQ	16/03/202
12	West Bengal	Puruliya	Kasipur	Mustard Oil	Other	FAQ	16/03/202
13	Haryana	Faridabad	Faridabad	Cauliflower	Cauliflower	FAQ	16/03/202
14	Haryana	Faridabad	Faridabad	Peas Wet	Other	FAQ	16/03/202
15	Telangana	Khammam	Kothagudem	Cotton	Cotton (Unginned)	FAQ	16/03/202
16	West Bengal	Puruliya	Kasipur	Bhindi(Ladies Finger)	Bhindi	FAQ	16/03/202
17	West Bengal	Puruliya	Kasipur	Mustard Oil	Other	FAQ	16/03/202

In [384…
```
df['Arrival_Date'] = pd.to_datetime(df['Arrival_Date'], format='%d/%m/%Y')
```

Distribution of Minimum Price

In [386…

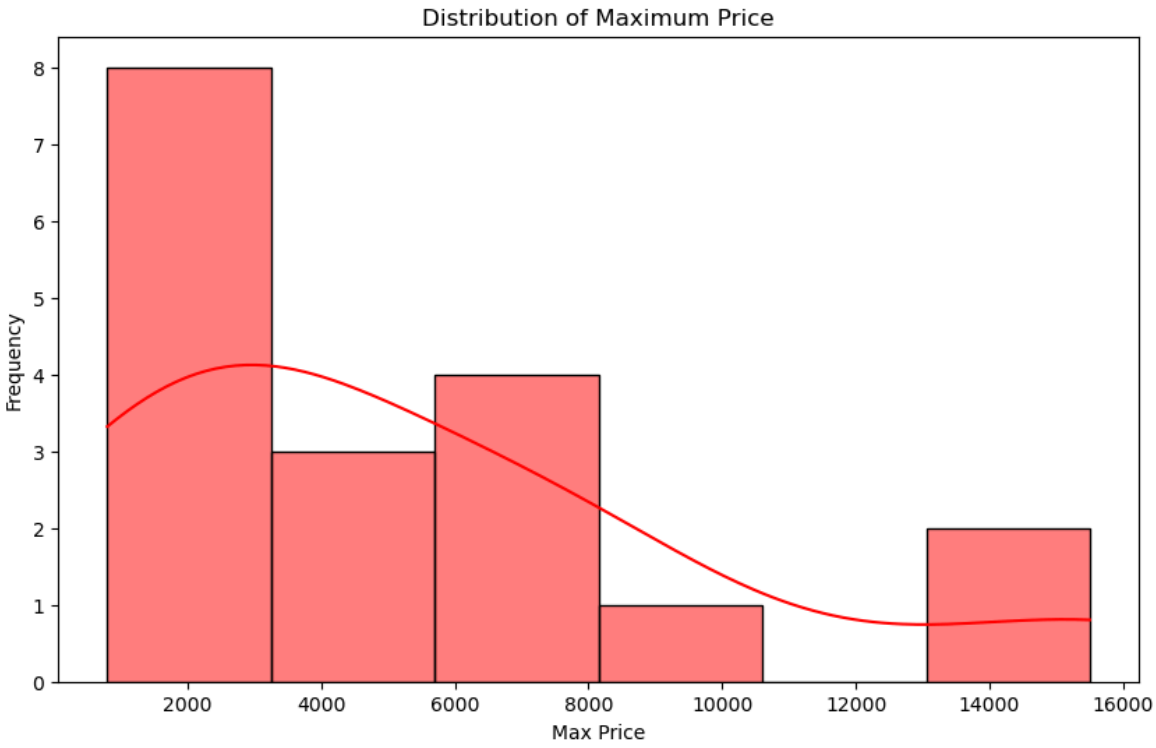
```
plt.figure(figsize=(10, 6))
sns.histplot(df['Min_x0020_Price'], kde=True, color='blue')
plt.title('Distribution of Minimum Price')
plt.xlabel('Min Price')
plt.ylabel('Frequency')
plt.show()
```



Distribution of Maximum Price

In [388…

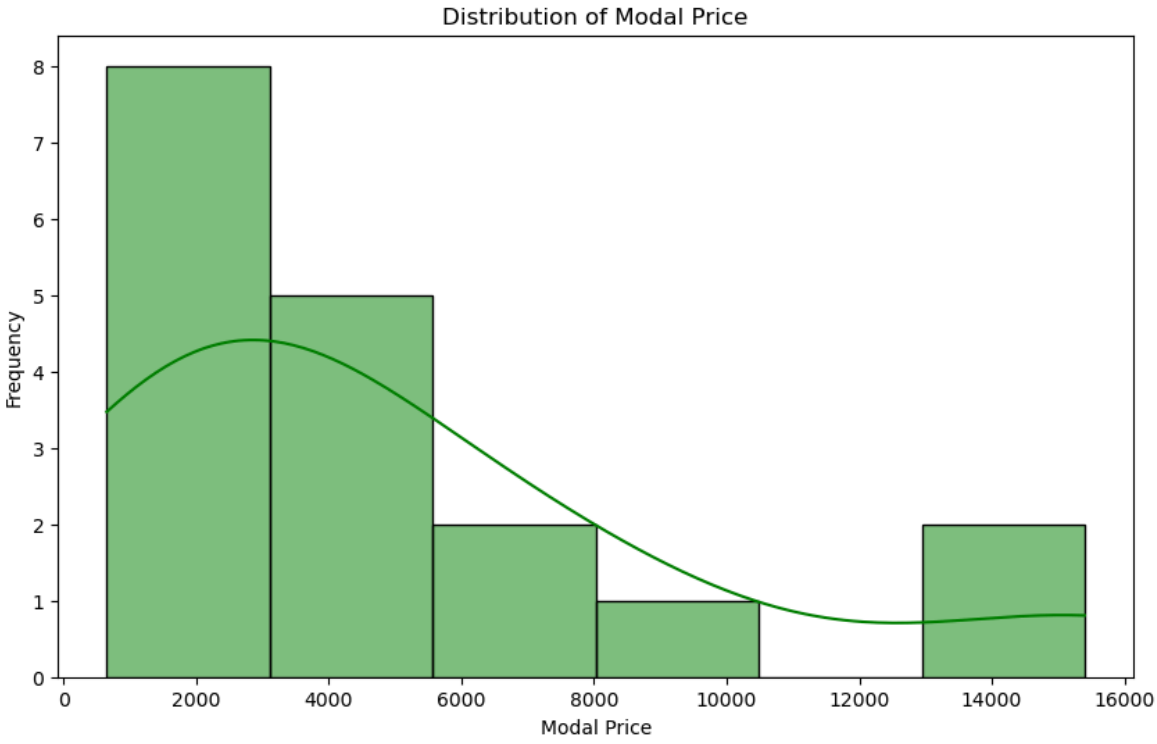
```
plt.figure(figsize=(10, 6))
sns.histplot(df['Max_x0020_Price'], kde=True, color='red')
plt.title('Distribution of Maximum Price')
plt.xlabel('Max Price')
plt.ylabel('Frequency')
plt.show()
```



Distribution of Modal Price

In [390…

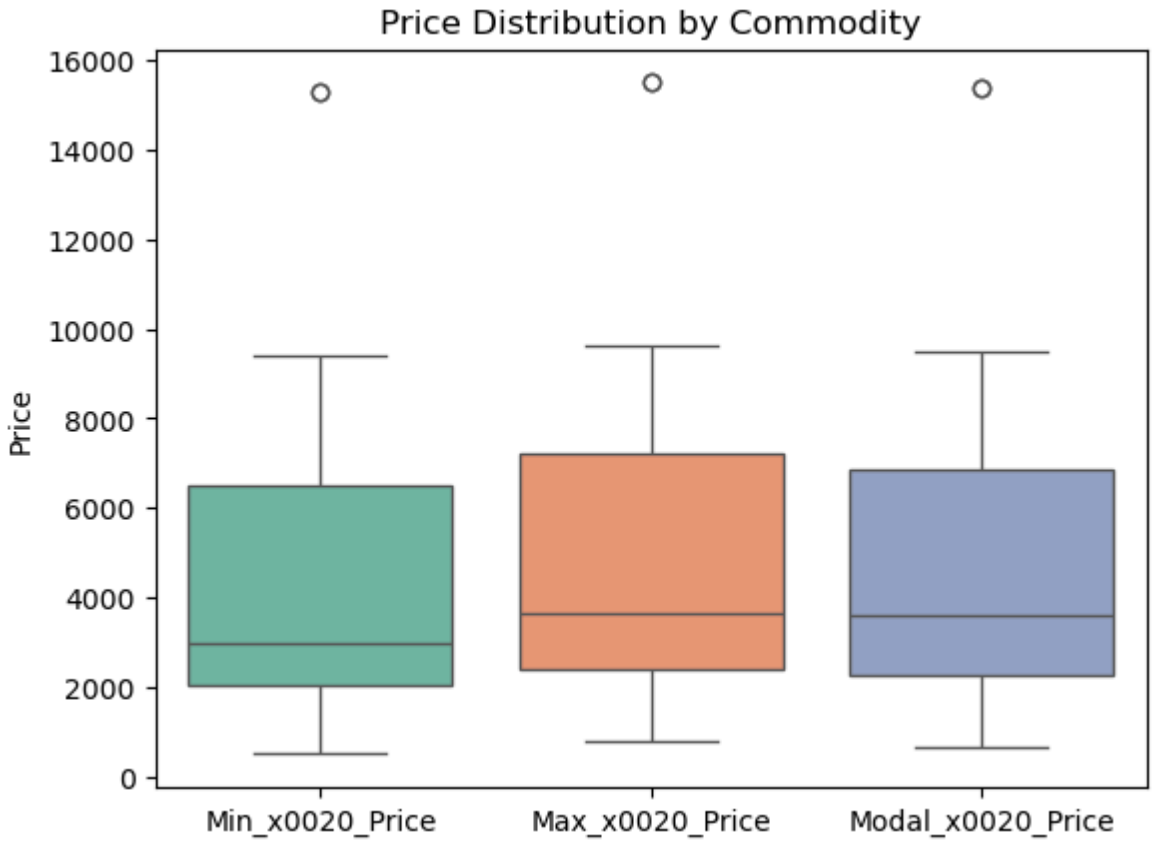
```
plt.figure(figsize=(10, 6))
sns.histplot(df['Modal_x0020_Price'], kde=True, color='green')
plt.title('Distribution of Modal Price')
plt.xlabel('Modal Price')
plt.ylabel('Frequency')
plt.show()
```



Boxplot: Min, Max, Modal Prices by Commodity

In [392…

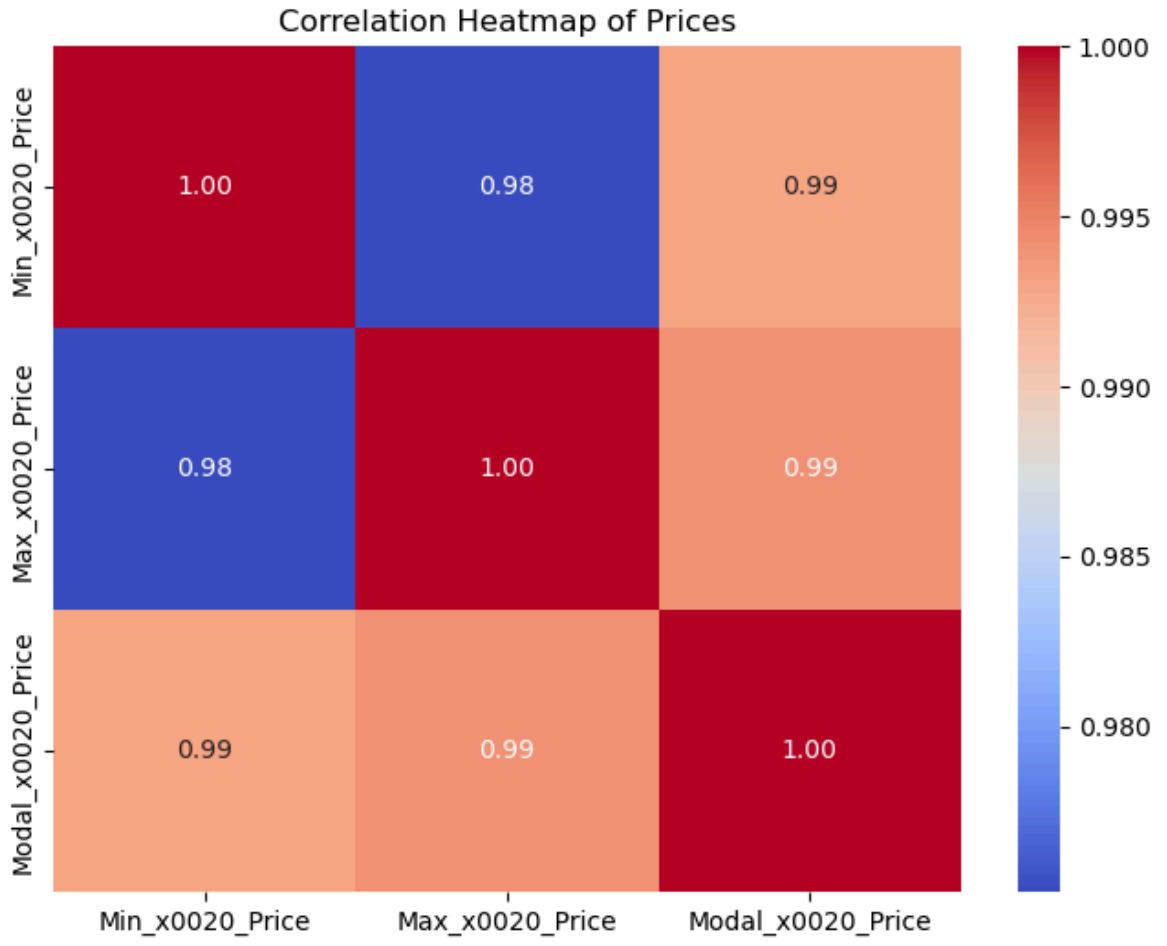
```
sns.boxplot(data=df[['Min_x0020_Price', 'Max_x0020_Price', 'Modal_x0020_Price']]
plt.title('Price Distribution by Commodity')
plt.ylabel('Price')
plt.show()
```



Correlation Heatmap

In [394…

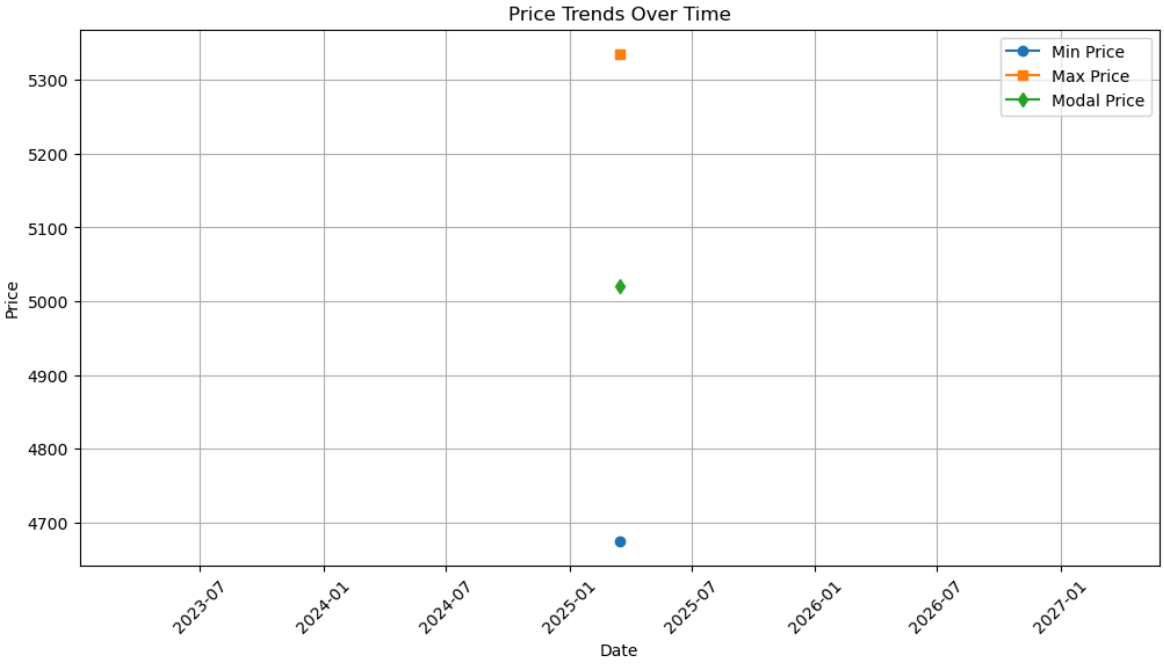
```
corr = df[['Min_x0020_Price', 'Max_x0020_Price', 'Modal_x0020_Price']].corr()
plt.figure(figsize=(8, 6))
sns.heatmap(corr, annot=True, cmap='coolwarm', fmt='.2f')
plt.title('Correlation Heatmap of Prices')
plt.show()
```



Price Trends Over Time (Line Plot)

In [396…

```
plt.figure(figsize=(12, 6))
plt.plot(df_grouped_by_date['Arrival_Date'], df_grouped_by_date['Min Price']
plt.plot(df_grouped_by_date['Arrival_Date'], df_grouped_by_date['Max Price']
plt.plot(df_grouped_by_date['Arrival_Date'], df_grouped_by_date['Modal Price
plt.title('Price Trends Over Time')
plt.xlabel('Date')
plt.ylabel('Price')
plt.legend()
plt.xticks(rotation=45)
plt.grid(True)
plt.show()
```

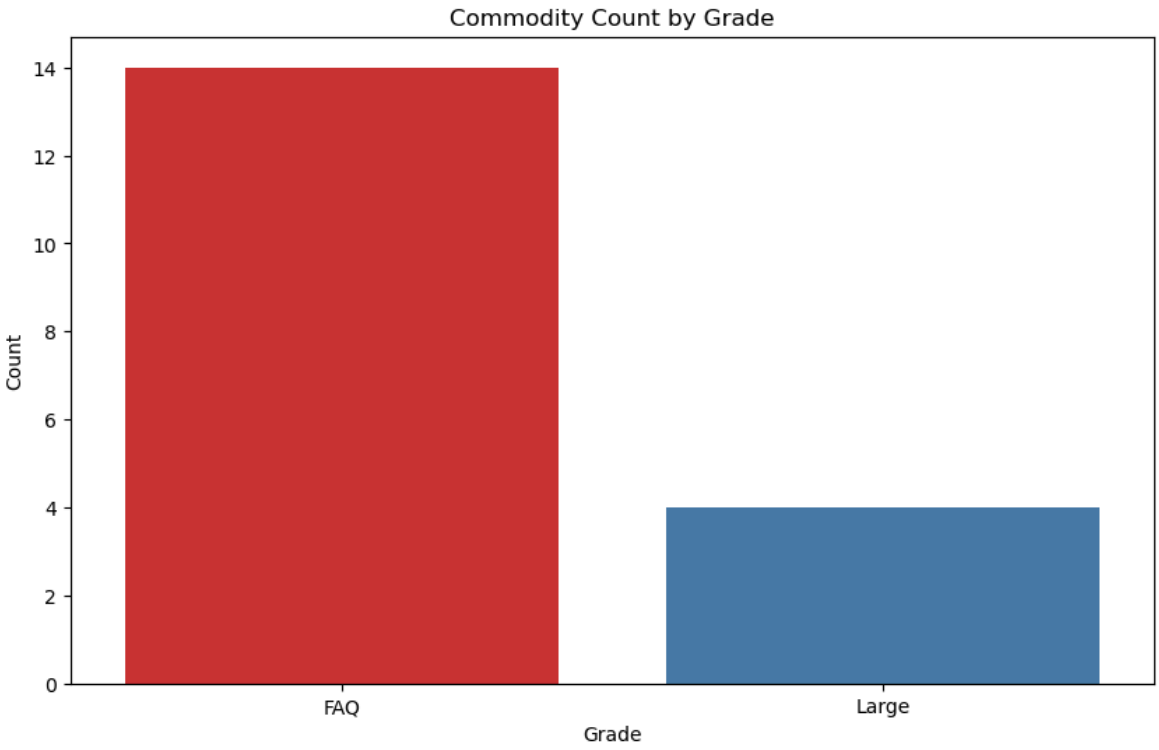


Count of Commodities by Grade

In [398…

```
plt.figure(figsize=(10, 6))
sns.countplot(x='Grade', data=df, palette='Set1')
plt.title('Commodity Count by Grade')
plt.ylabel('Count')
plt.show()
```

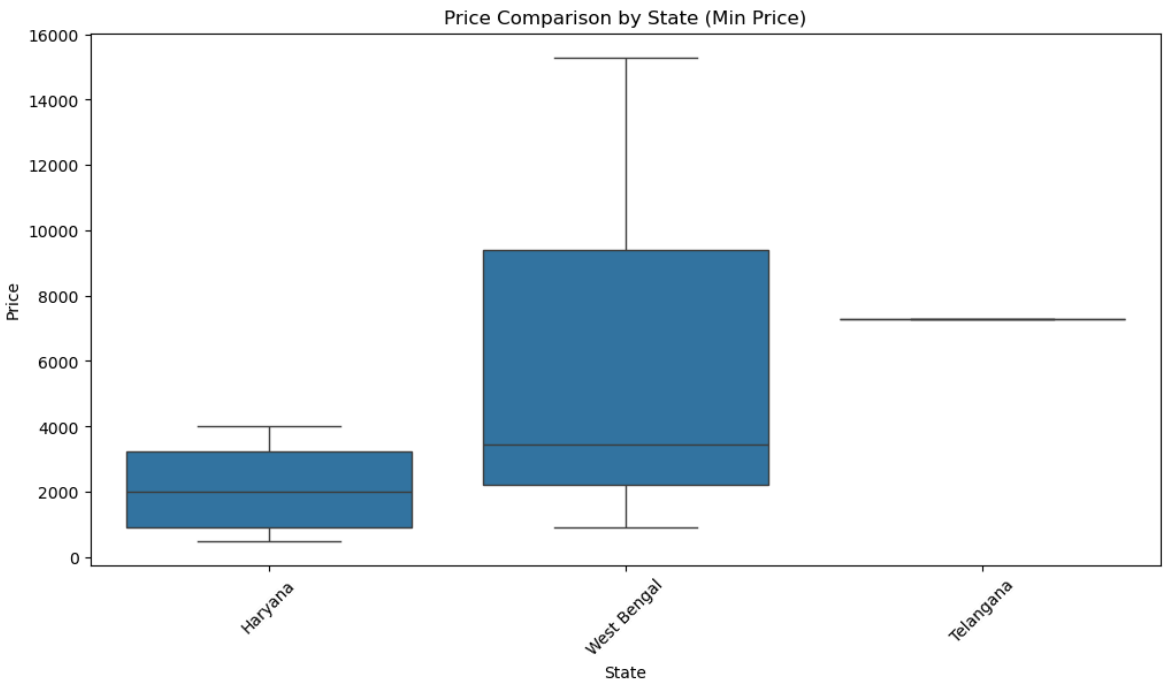
```
/var/folders/ky/v7pnydl97tvccw05b55n4xx00000gn/T/ipykernel_26112/2130381866.py:2:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.countplot(x='Grade', data=df, palette='Set1')
```



Price Comparison by State

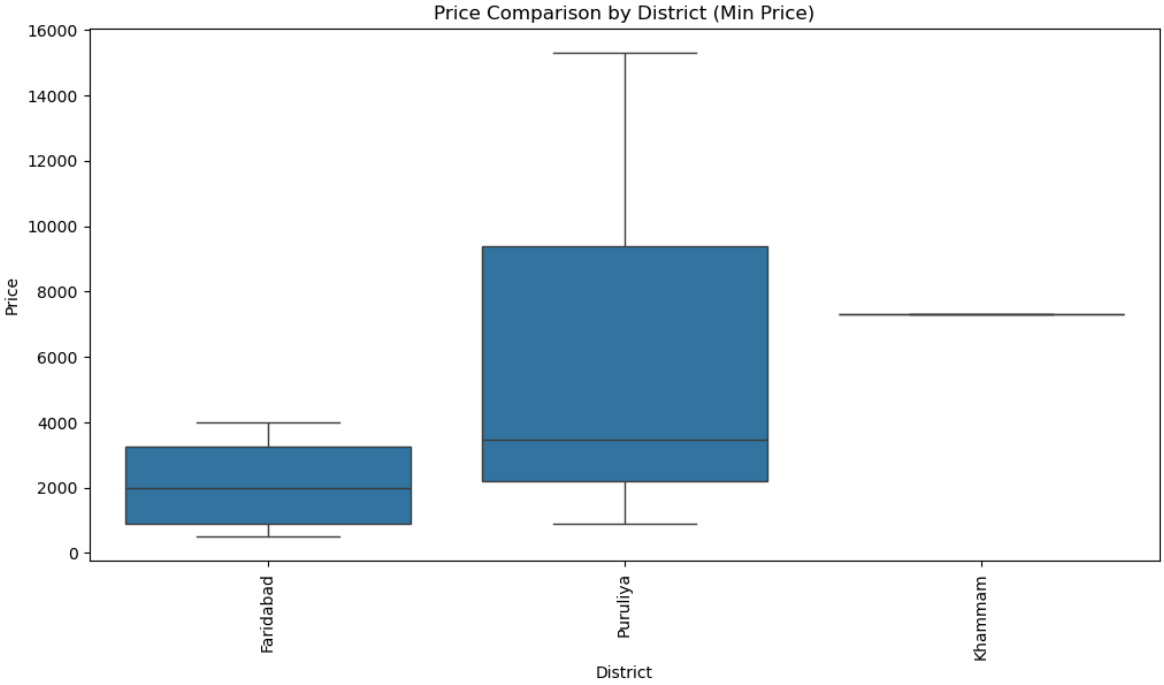
In [400…

```
plt.figure(figsize=(12, 6))
sns.boxplot(x='State', y='Min_x0020_Price', data=df)
plt.title('Price Comparison by State (Min Price)')
plt.ylabel('Price')
plt.xlabel('State')
plt.xticks(rotation=45)
plt.show()
```



Price Comparison by District

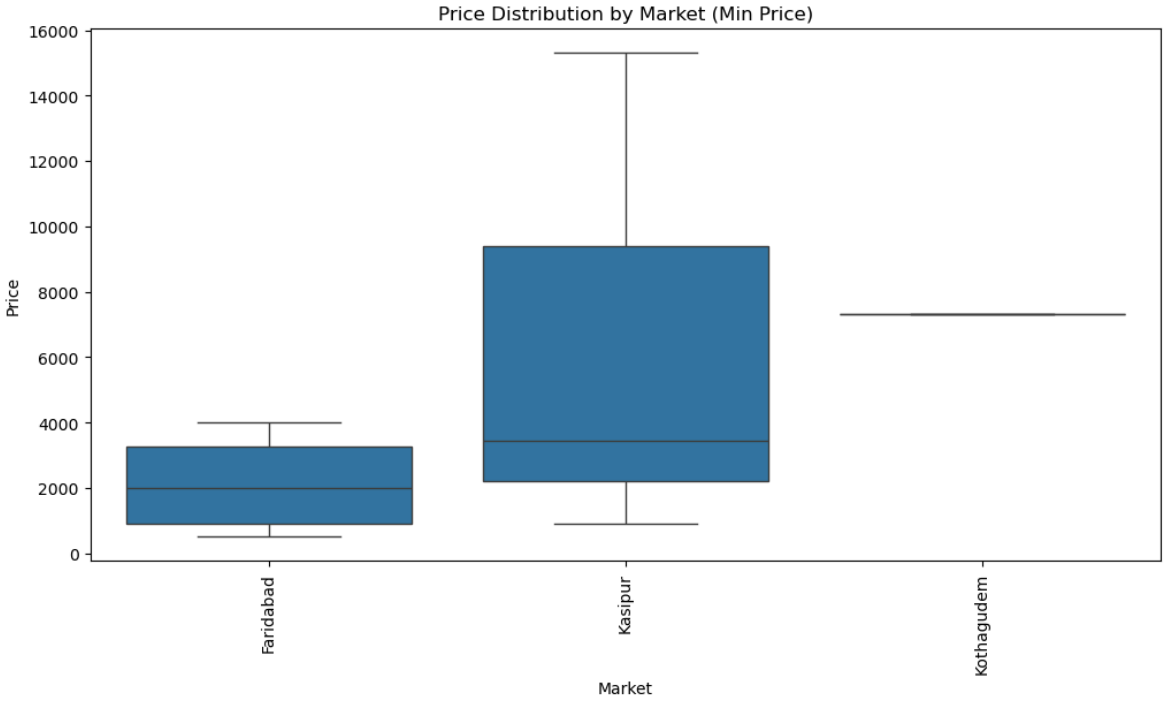
In [402…
```
plt.figure(figsize=(12, 6))
sns.boxplot(x='District', y='Min_x0020_Price', data=df)
plt.title('Price Comparison by District (Min Price)')
plt.xlabel('District')
plt.ylabel('Price')
plt.xticks(rotation=90)
plt.show()
```



Distribution of Prices by Market

In [404…
```
df.rename(columns=lambda x: x.replace('_x0020_', ' '), inplace=True)
df = df.dropna(subset=['Min Price'])
df['Market'] = df['Market'].astype(str)
plt.figure(figsize=(12, 6))
sns.boxplot(x='Market', y='Min Price', data=df)
plt.title('Price Distribution by Market (Min Price)')
plt.xlabel('Market')
plt.ylabel('Price')
plt.xticks(rotation=90)
```

Out[404…
```
([0, 1, 2],
 [Text(0, 0, 'Faridabad'), Text(1, 0, 'Kasipur'), Text(2, 0, 'Kothagudem')])
```

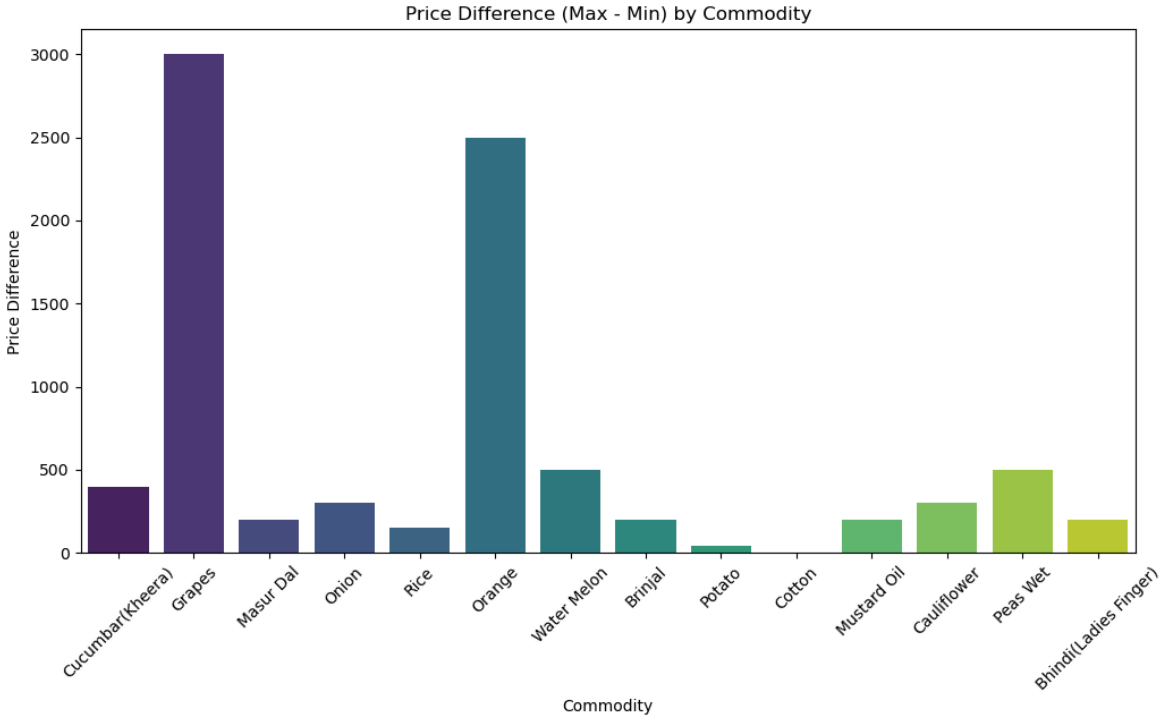


Price Difference (Max - Min) by Commodity

In [406…

```
df['Price_Diff'] = df['Max Price'] - df['Min Price']
plt.figure(figsize=(12, 6))
sns.barplot(x='Commodity', y='Price_Diff', data=df, palette='viridis')
plt.title('Price Difference (Max - Min) by Commodity')
plt.ylabel('Price Difference')
plt.xticks(rotation=45)
plt.show()
```

```
/var/folders/ky/v7pnydl97tvccw05b55n4xx00000gn/T/ipykernel_26112/374020366.py:3:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.barplot(x='Commodity', y='Price_Diff', data=df, palette='viridis')
```

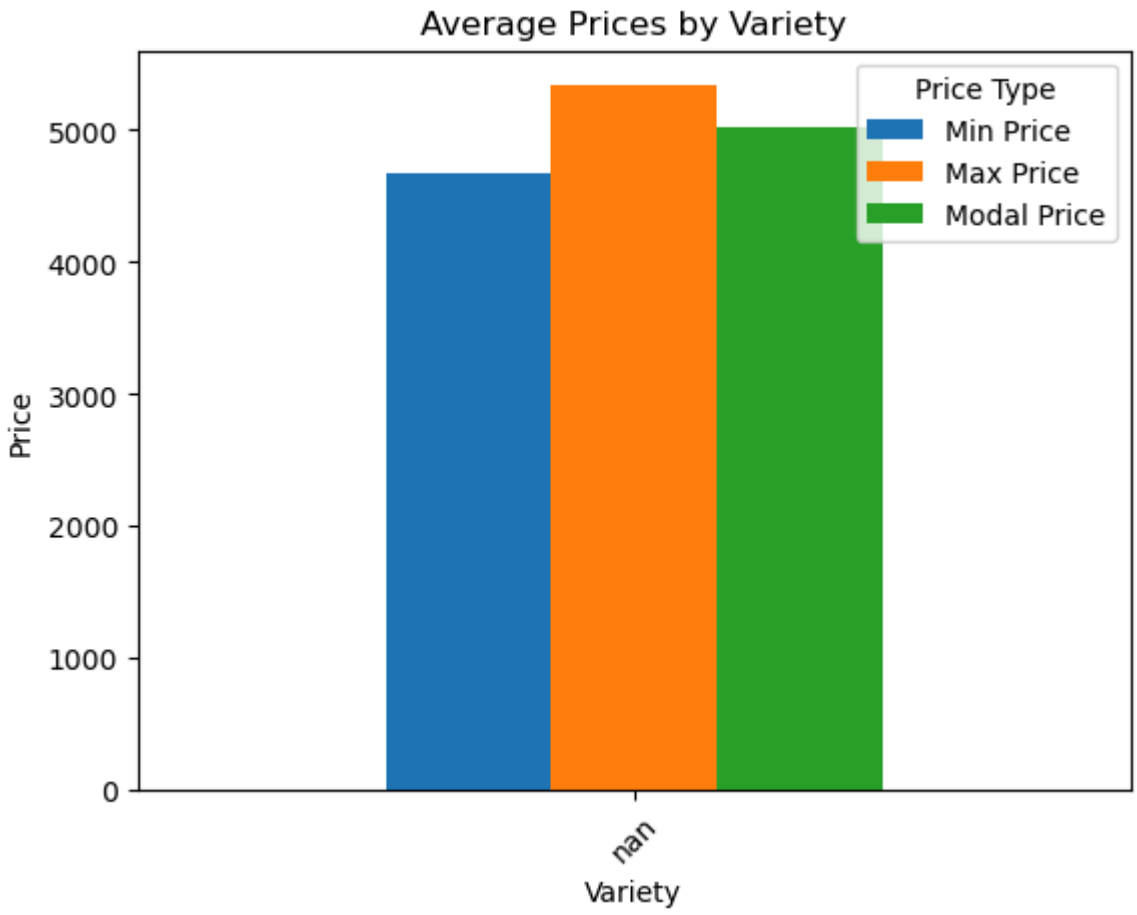


Average Prices by Variety

In [408…

```
plt.figure(figsize=(12, 6))
df_avg_prices.set_index('Variety').plot(kind='bar', stacked=False)  # Change to

plt.title('Average Prices by Variety')
plt.ylabel('Price')
plt.xlabel('Variety')
plt.xticks(rotation=45)
plt.legend(title="Price Type")
plt.show()
```

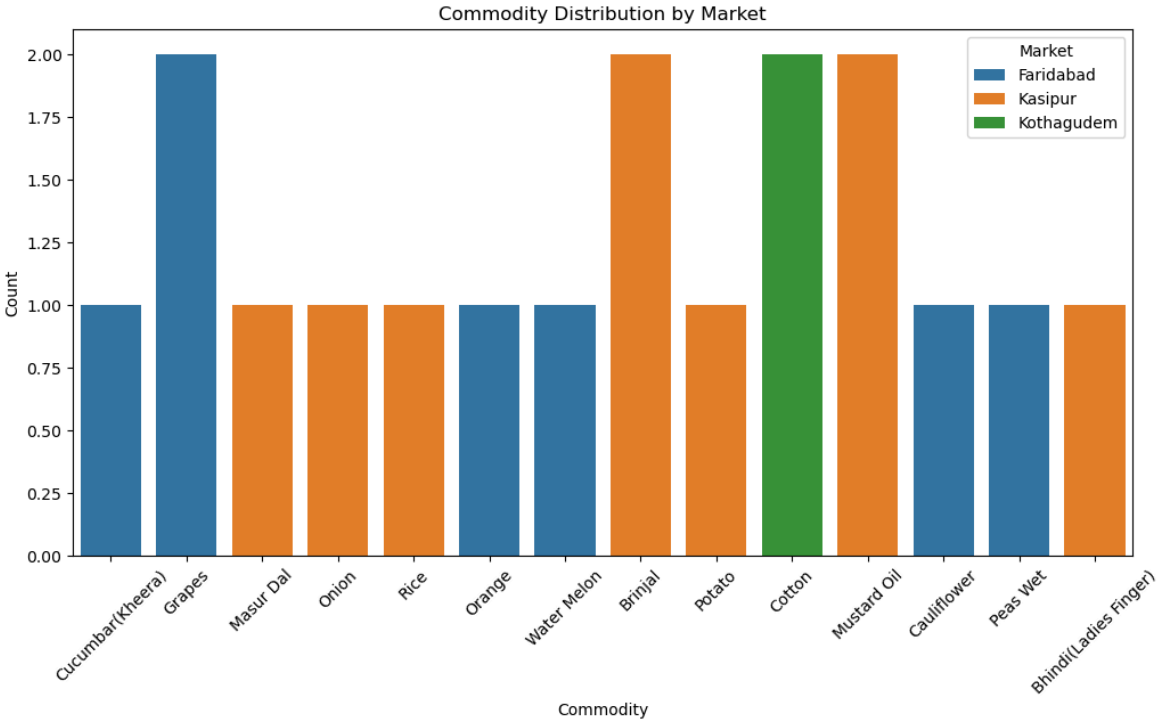
```
<Figure size 1200x600 with 0 Axes>
```



Commodity Distribution by Market

In [410…

```
plt.figure(figsize=(12, 6))
sns.countplot(x='Commodity', hue='Market', data=df)
plt.title('Commodity Distribution by Market')
plt.ylabel('Count')
plt.xticks(rotation=45)
plt.show()
```

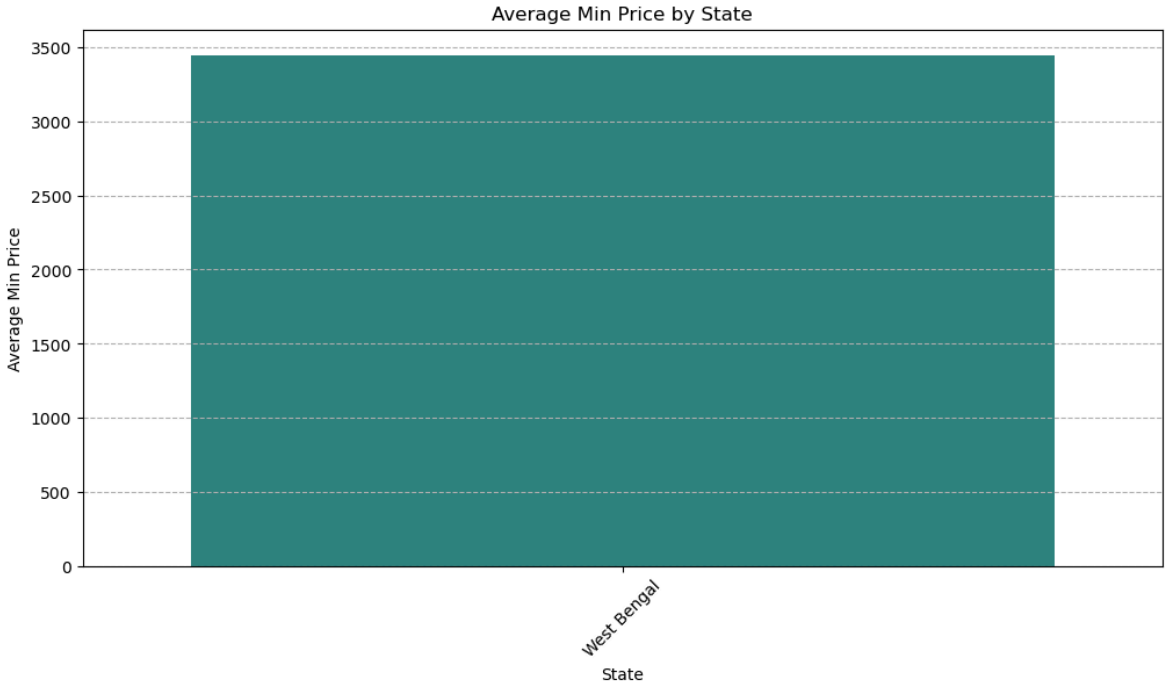


State Wise Price Trends

In [412…

```
plt.figure(figsize=(12, 6))
df_avg_price = df_state.groupby('State')['Min Price'].mean().reset_index()
sns.barplot(x='State', y='Min Price', data=df_avg_price, palette='viridis')
plt.title('Average Min Price by State')
plt.xlabel('State')
plt.ylabel('Average Min Price')
plt.xticks(rotation=45)
plt.grid(axis='y', linestyle='--')
plt.show()
```

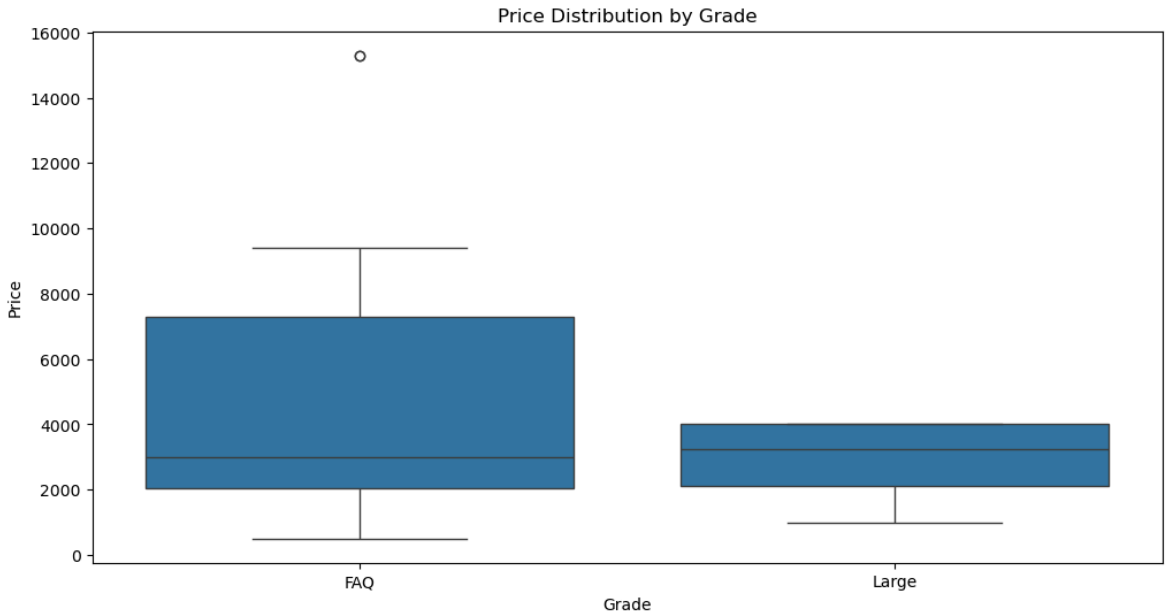
```
/var/folders/ky/v7pnydl97tvccw05b55n4xx00000gn/T/ipykernel_26112/3000224058.py:3:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v
0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effe
ct.

  sns.barplot(x='State', y='Min Price', data=df_avg_price, palette='viridis')
```



Price Distribution by Grade

In [414…

```
plt.figure(figsize=(12, 6))
sns.boxplot(x='Grade', y='Min Price', data=df)
plt.title('Price Distribution by Grade')
plt.ylabel('Price')
plt.show()
```

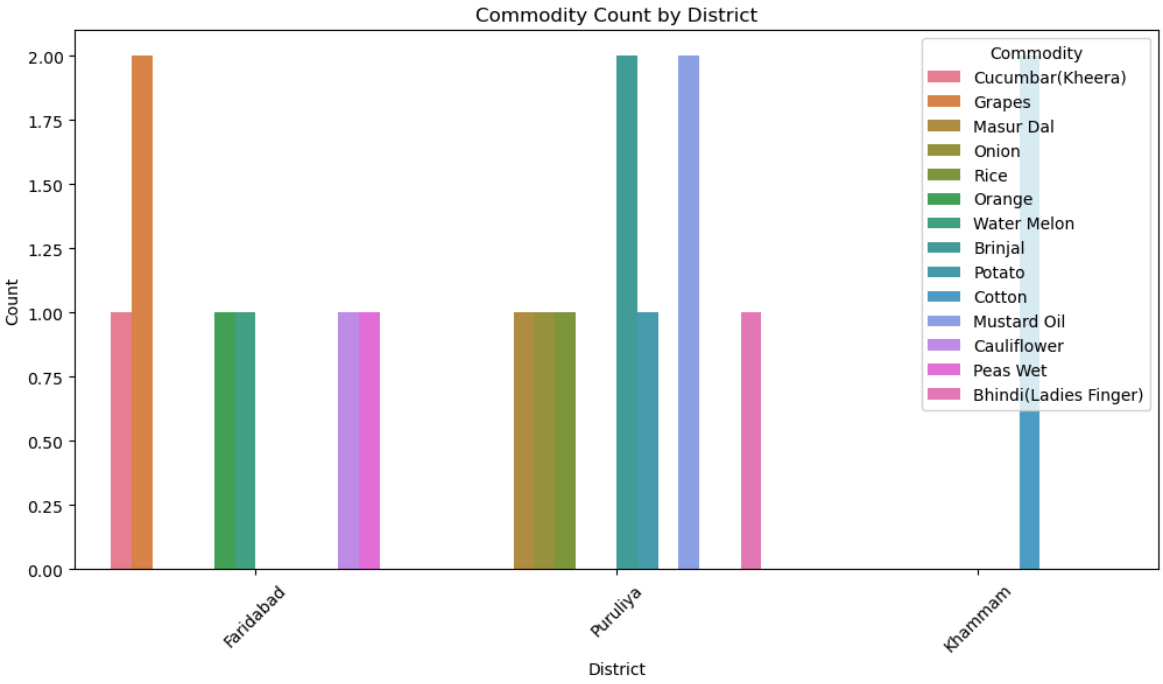


Commodity Count by District

In [416…

```
plt.figure(figsize=(12, 6))
sns.countplot(x='District', hue='Commodity', data=df)
plt.title('Commodity Count by District')
plt.ylabel('Count')
```

```
plt.xticks(rotation=45)
plt.show()
```



Price Distribution for Specific Commodity

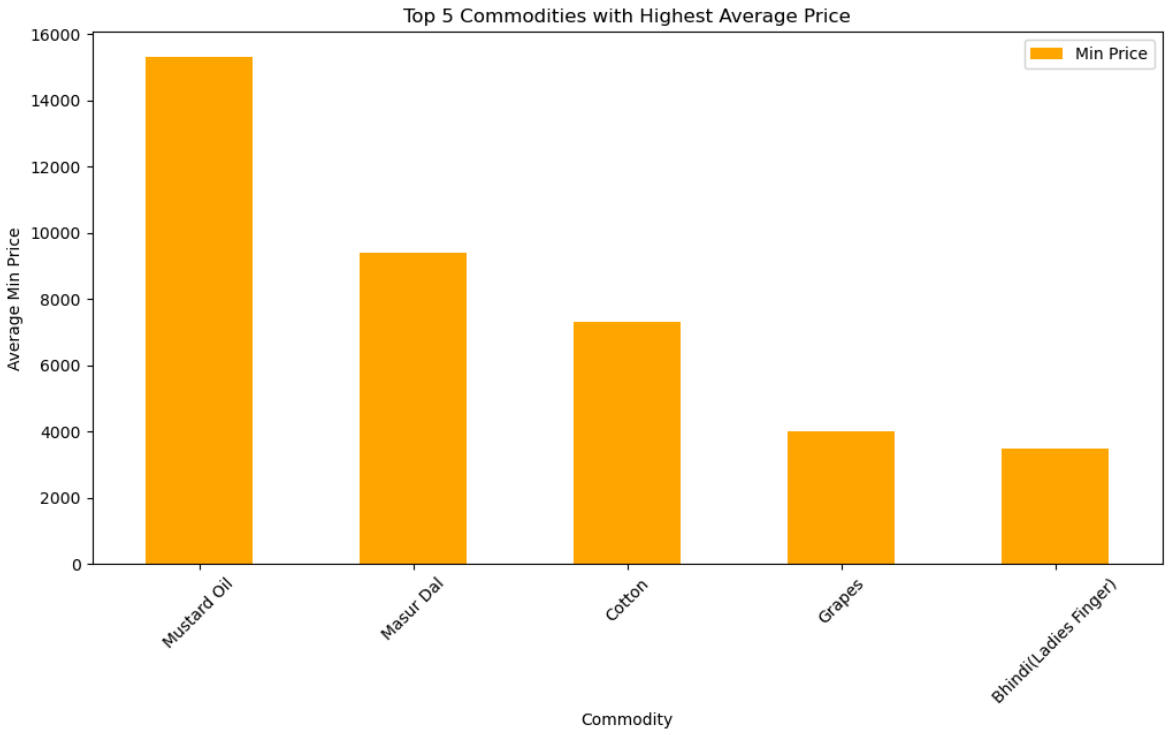
In [418…

```
masur_dal_df = df[df['Commodity'] == 'Masur Dal']
plt.figure(figsize=(10, 6))
sns.histplot(masur_dal_df['Min Price'], kde=True, color='purple')
plt.title('Price Distribution for Masur Dal')
plt.xlabel('Min Price')
plt.ylabel('Frequency')
plt.show()
```

Top 5 Commodities with Highest Average Price

In [420…

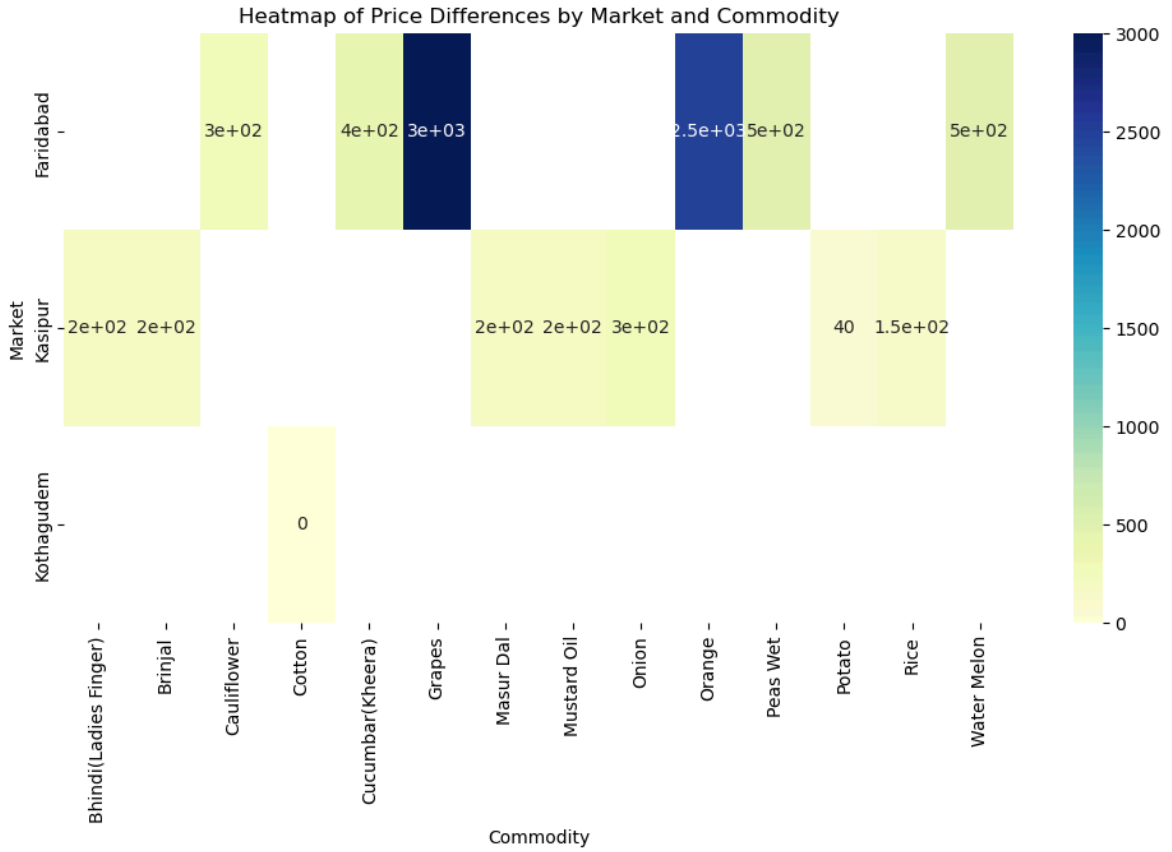
```
top_commodities = df.groupby('Commodity')[['Min Price']].mean().sort_values(by='
top_commodities.plot(kind='bar', figsize=(12, 6), color='orange')
plt.title('Top 5 Commodities with Highest Average Price')
plt.ylabel('Average Min Price')
plt.xticks(rotation=45)
plt.show()
```



Heatmap of Price Differences by Market and Commodity

In [422…

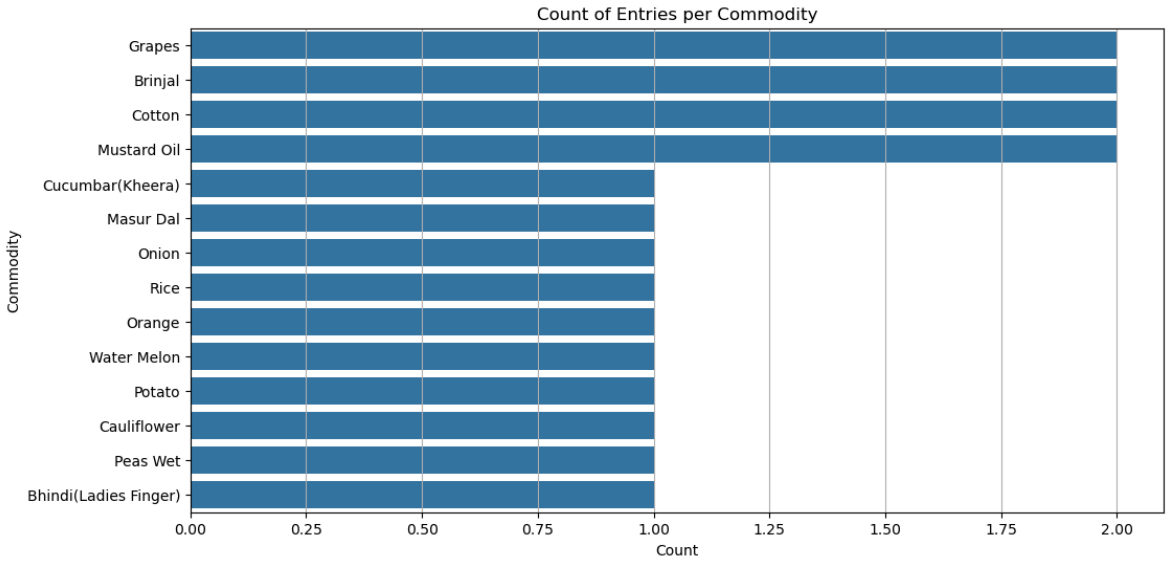
```
price_diff_by_market = df.pivot_table(index='Market', columns='Commodity', value
plt.figure(figsize=(12, 6))
sns.heatmap(price_diff_by_market, annot=True, cmap='YlGnBu')
plt.title('Heatmap of Price Differences by Market and Commodity')
plt.show()
```



Market Price Trends for Specific Commodities

In [424…

```
plt.figure(figsize=(12, 6))
sns.countplot(y='Commodity', data=df, order=df['Commodity'].value_counts().index
plt.title('Count of Entries per Commodity')
plt.xlabel('Count')
plt.ylabel('Commodity')
plt.grid(True, axis='x')
plt.show()
```



In [ ]:

In [ ]:

In [ ]: