Data Preparation:

1. \*\*Importing Libraries:\*\*

- Import Pandas and NumPy for data manipulation.

```python

import pandas as pd

import numpy as np

```

2. \*\*Loading the Dataset:\*\*

- Read the dataset from a CSV file into a Pandas DataFrame.

```python

df = pd.read\_csv("Weather.csv")

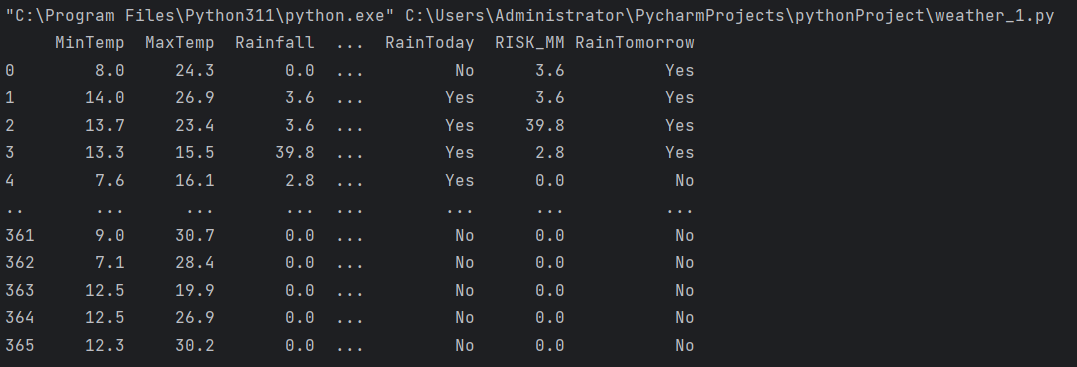
```

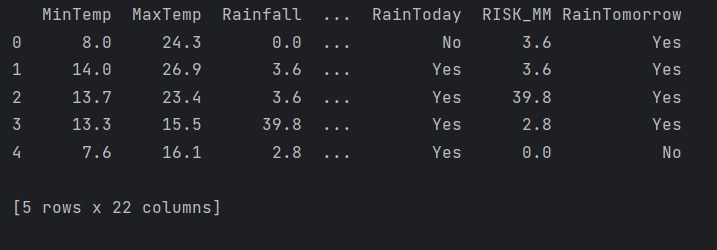
3. \*\*Checking Data Overview:\*\*

- Print the entire DataFrame and the first five rows.

```python

print(df)

 print(df.head())



4. \*\*Checking Data Structure:\*\*

- Print the shape, index, column names, data types, unique values, number of unique values, and non-null counts.

```python

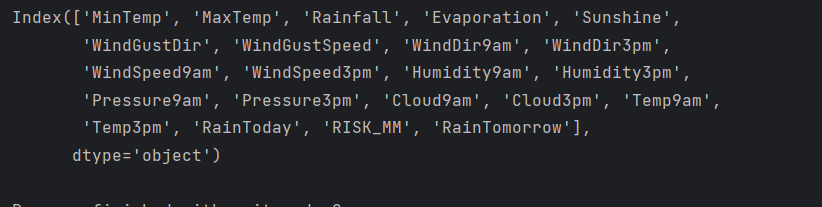
print(df.shape)



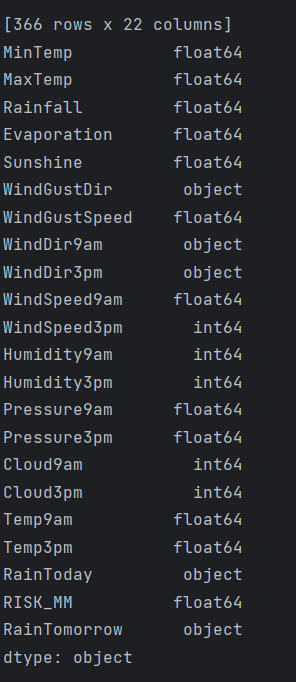
print(df.index)



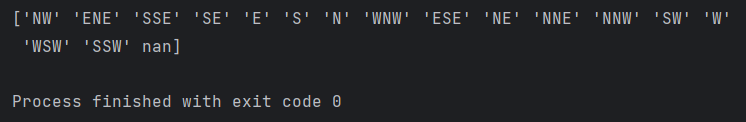
print(df.columns)



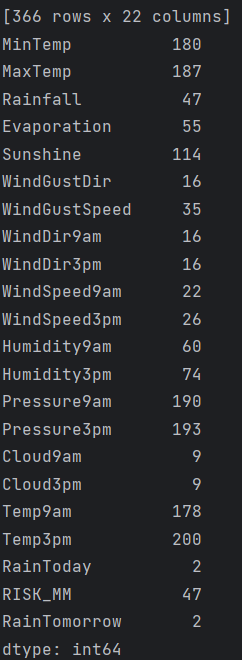
print(df.dtypes)



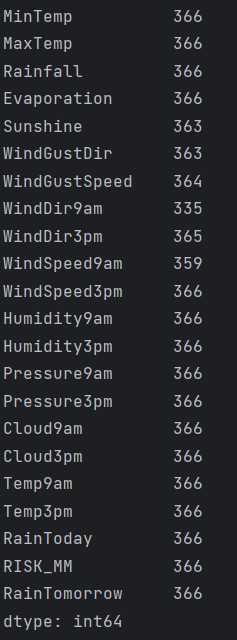
print(df['WindGustDir'].unique())



print(df.nunique())



print(df.count())



**Data Analysis:**

5. \*\*Exploring Specific Columns:\*\*

- Print unique values and their counts for the 'WindGustDir' column.

```python

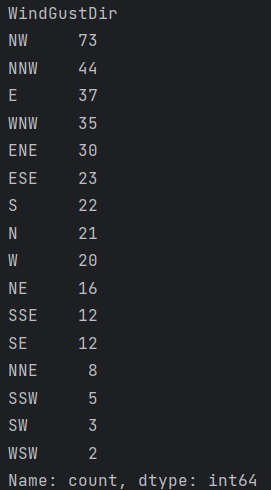
print(df['WindGustDir'].value\_counts())

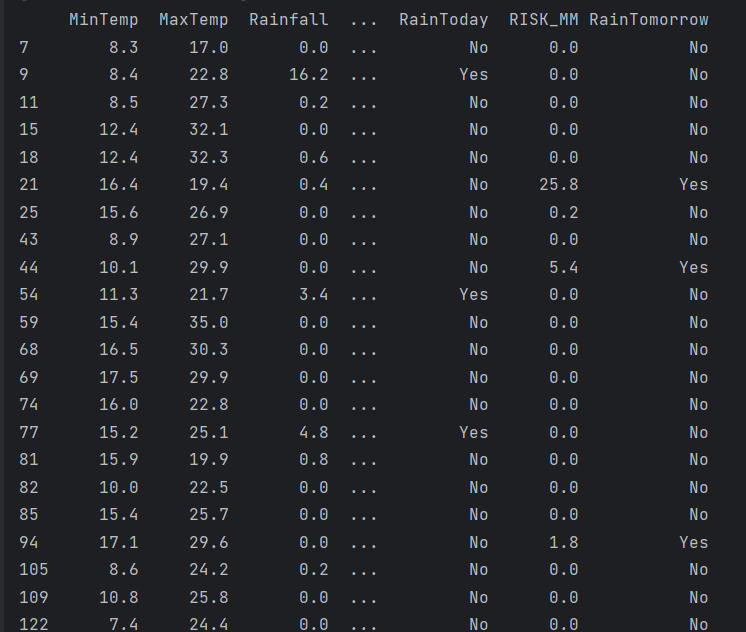
```

6. Grouping Data:

Use groupby() to group data by 'WindGustDir' and get the group where it's 'E'.

print(df[df.WindGustDir == 'E'])

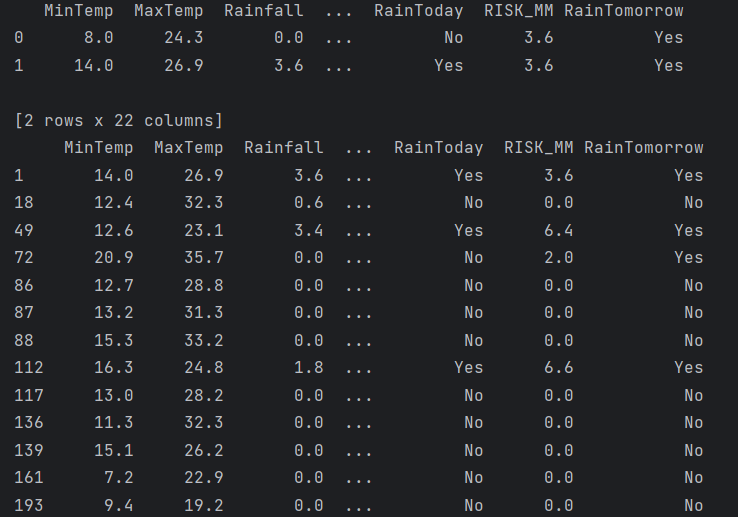




**Filtering Data Based on Numeric Condition ('WindSpeed9am'):**

Filter data where 'WindSpeed9am' is exactly 4 km/hr.

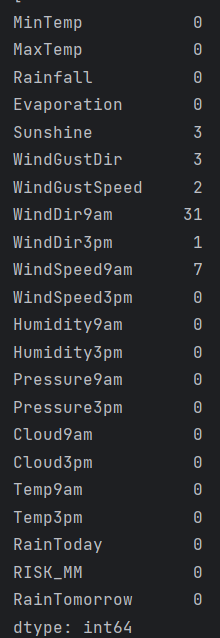
print(df[df['WindSpeed9am'] == 4])



Handling Null Values:

Print the count of null values in each column.

print(df.isnull().sum())

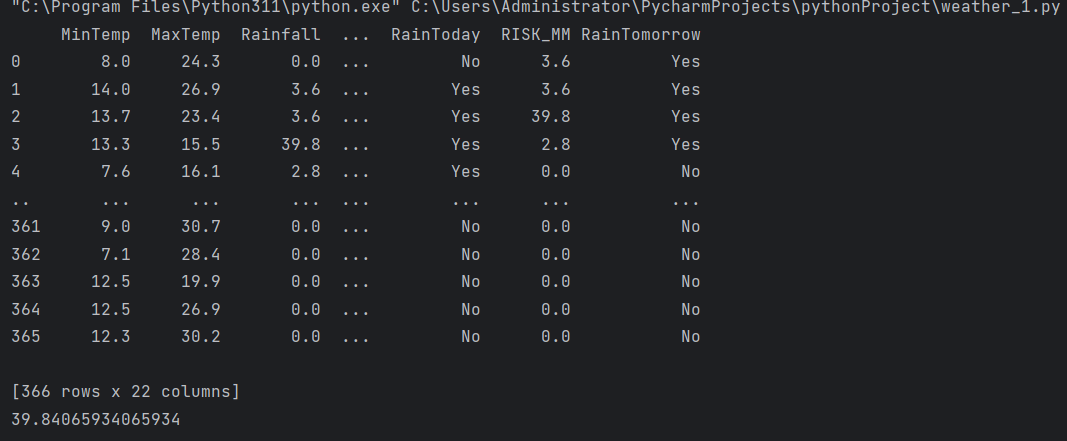


Additional Analysis:

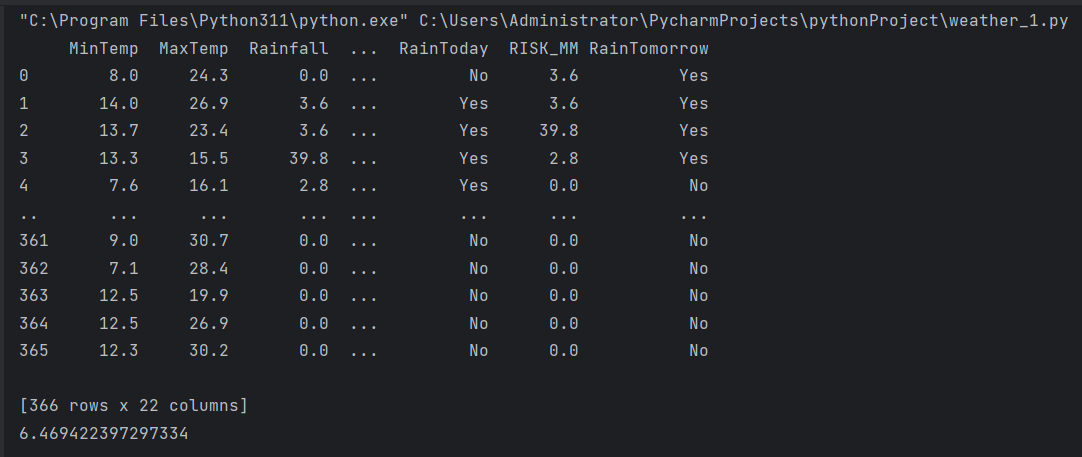
Calculating Mean, Standard Deviation, and Variance:

Print mean 'WindGustSpeed', standard deviation of 'Pressure3pm', and variance of 'Humidity9am'.

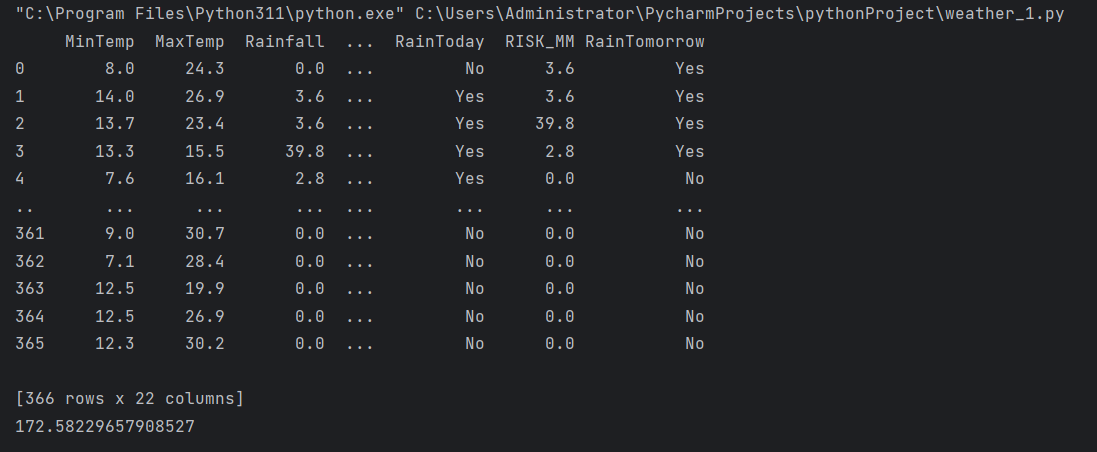
print(df.WindGustSpeed.mean())



print(df.Pressure3pm.std())



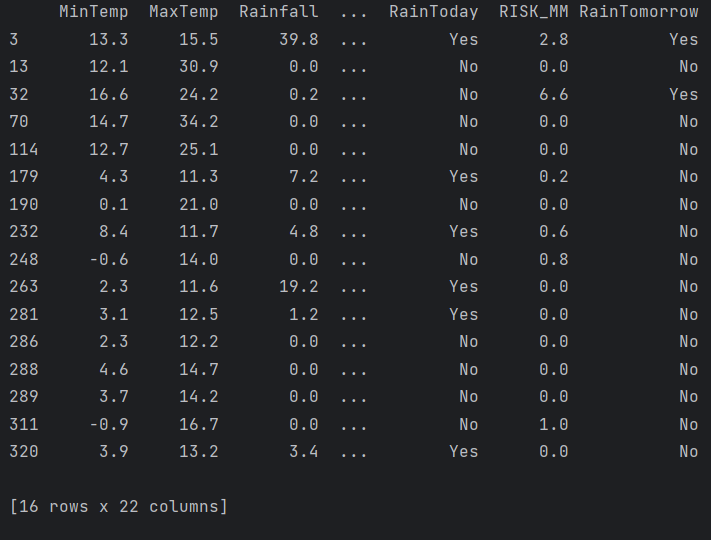
print(df['Humidity9am'].var())



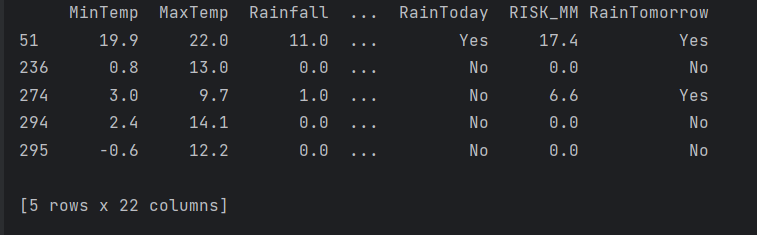
Filtering Data Based on Conditions ('WindDir9am', 'WindSpeed3pm', 'Cloud9am'):

Filter data where 'WindDir9am' is 'WNW', 'WindSpeed3pm' is above 24, and 'Cloud9am' is 8.

print(df[df['WindDir9am'] == 'WNW'])



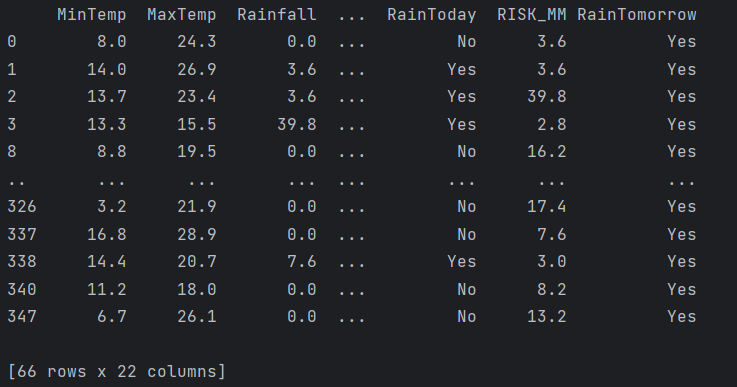
print(df[(df['WindSpeed3pm'] > 24) & (df['Cloud9am'] == 8)])



Filtering Data for Rain Prediction ('RainTomorrow' and 'RainToday'):

Filter data where 'RainTomorrow' is 'Yes' and where 'RainToday' is 'No' or 'RISK\_MM' is above 22.6.

print(df[df['RainTomorrow'] == 'Yes'])



print(df[(df['RainToday'] == 'No') | (df['RISK\_MM'] > 22.6)])

