**Data Collection and Preprocessing Phase**

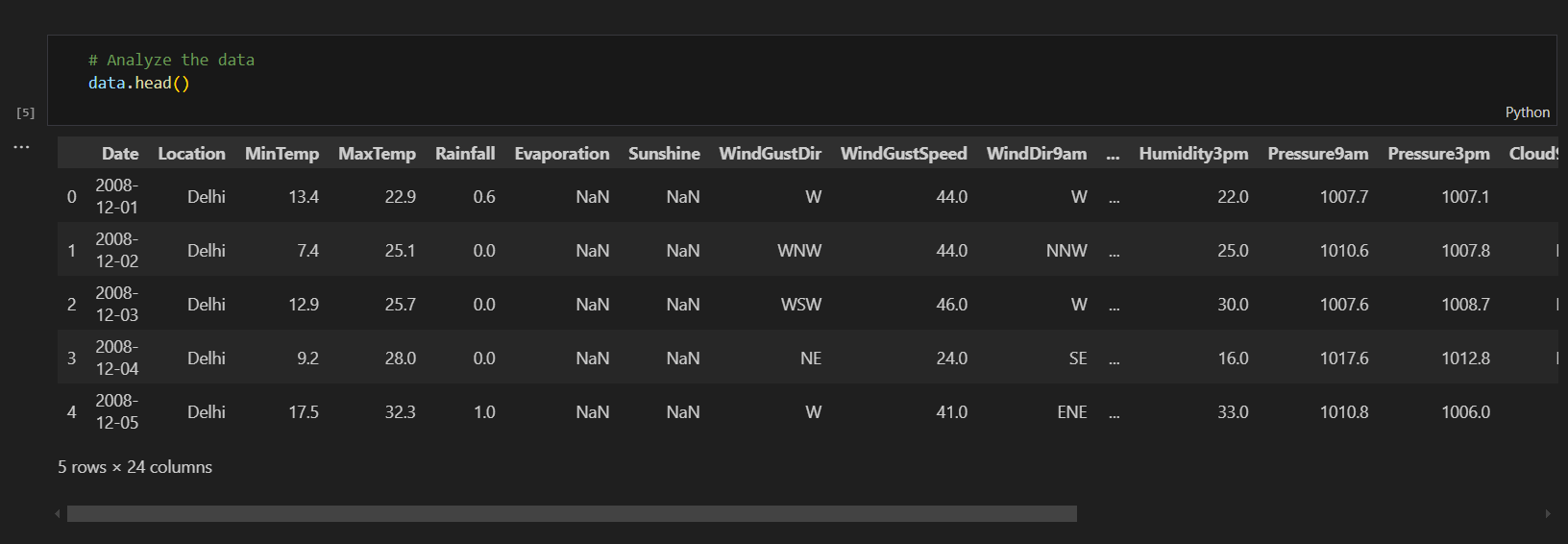
**Data Collection :**

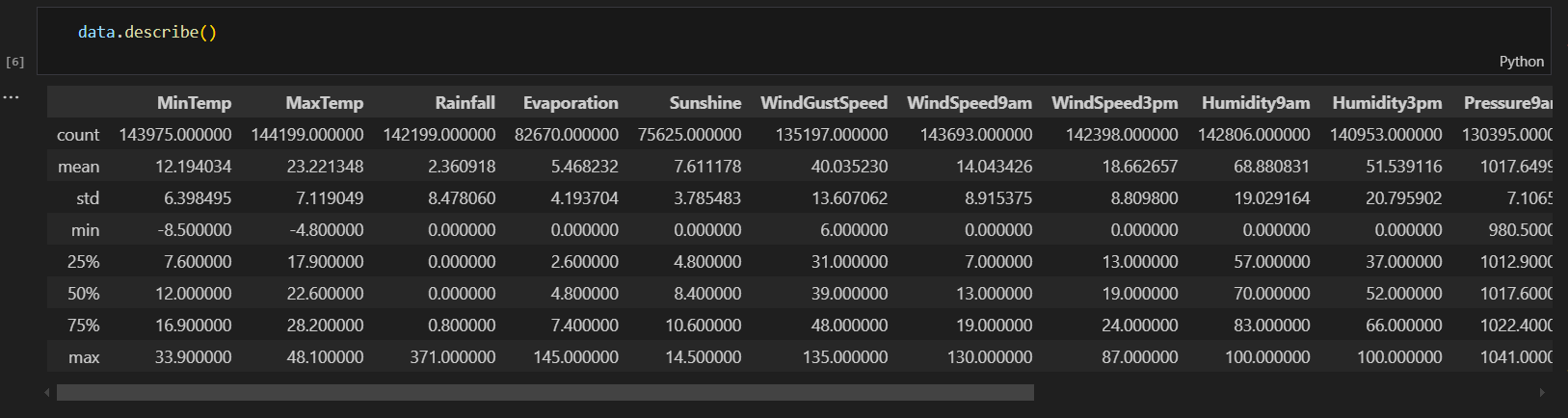
We have used the following data for the model to train on :

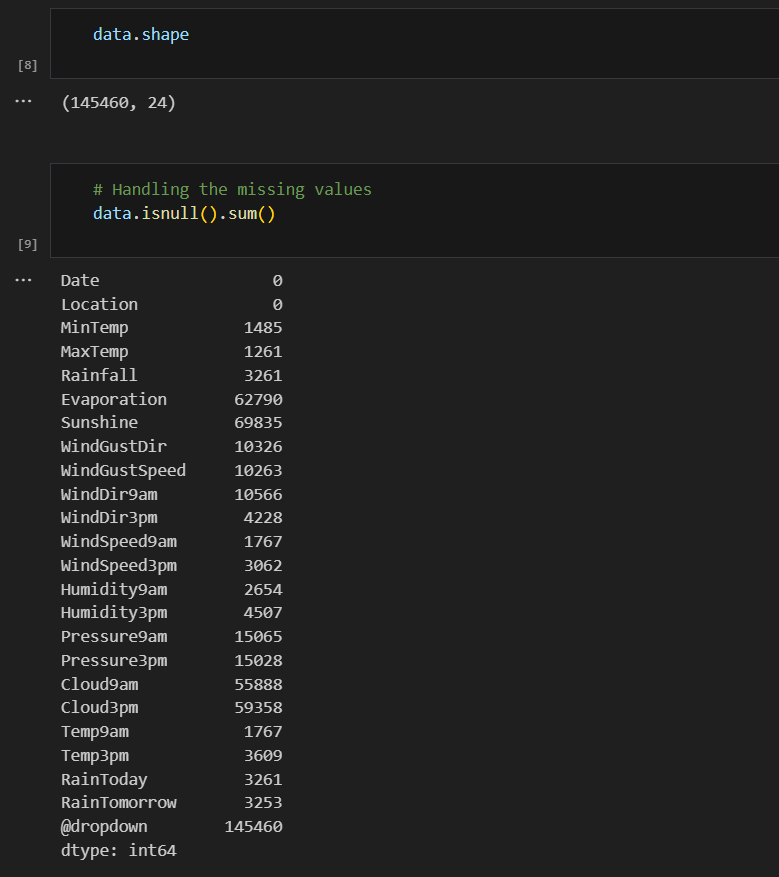
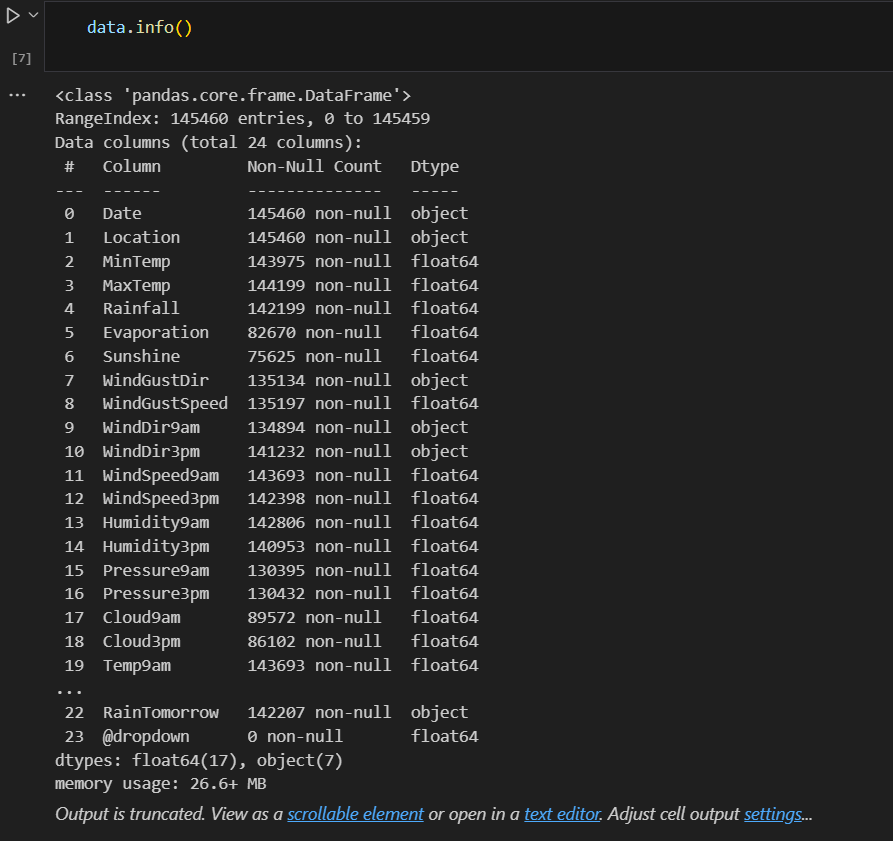
<https://docs.google.com/spreadsheets/d/1RA2OO0LZTeQykI_mvnensAjp6LM4YzWI1Tz0SUG5-Ao/edit?gid=121883362#gid=121883362>

**Data Quality Report :**

Analyzing the data:

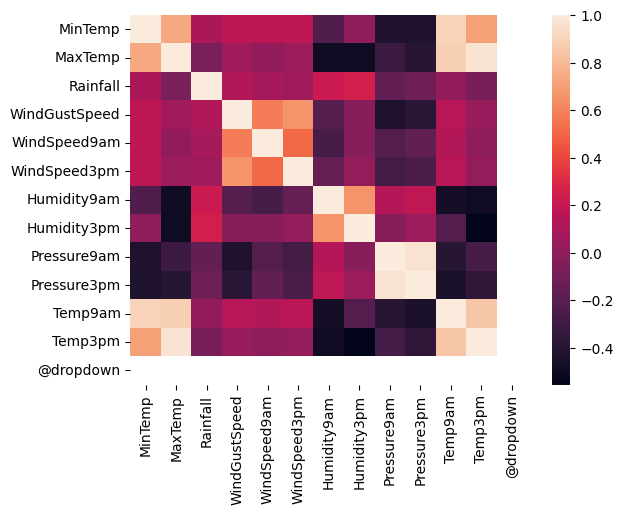
****

****

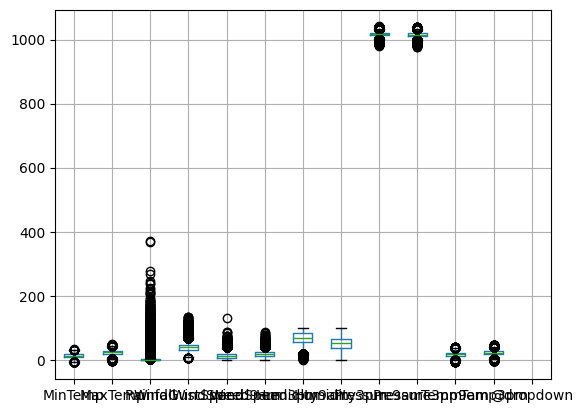
****

**Preprocessing Report : ( Visualisation Report )**

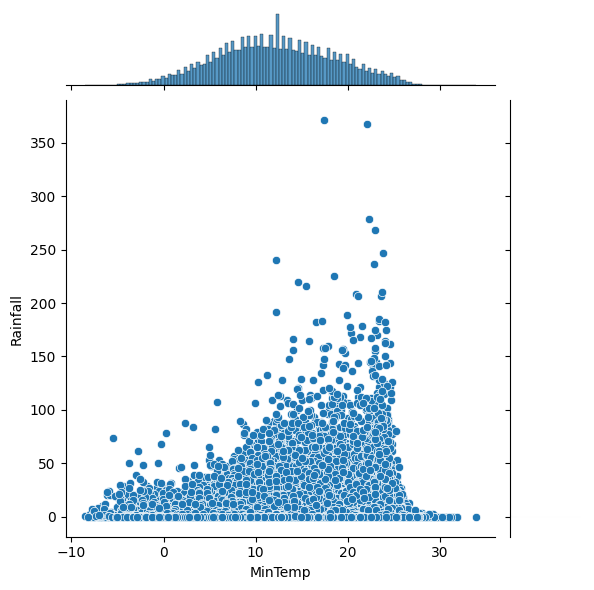
1. **The HeatMap of the data**

****

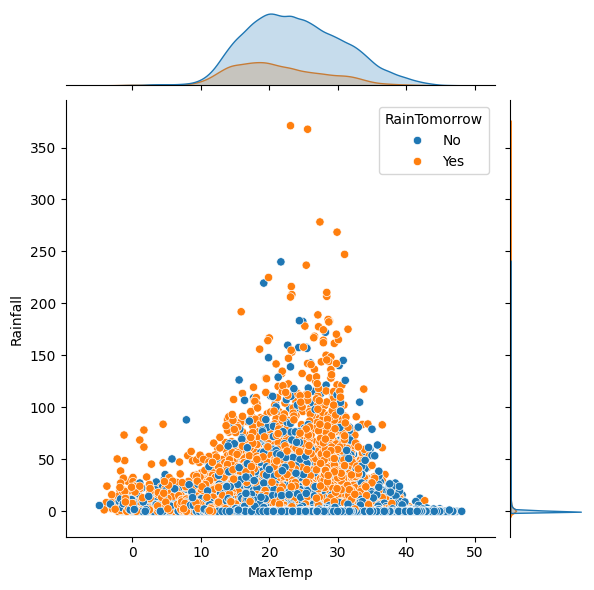
1. **The Boxplot**

****

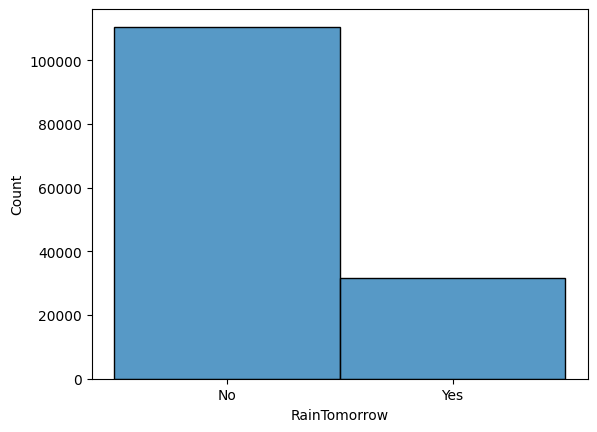
1. **The Jointplot - 1**

****

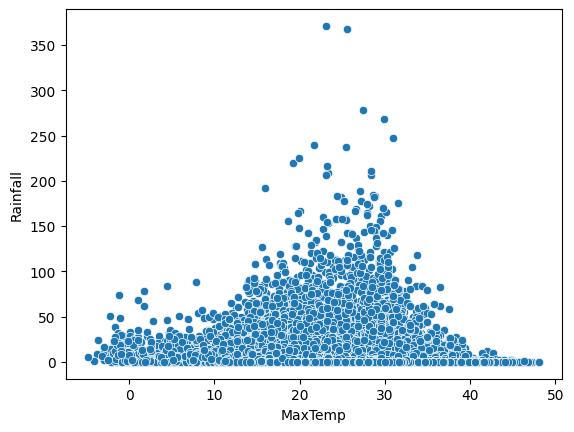
1. **The Jointplot - 2**

****

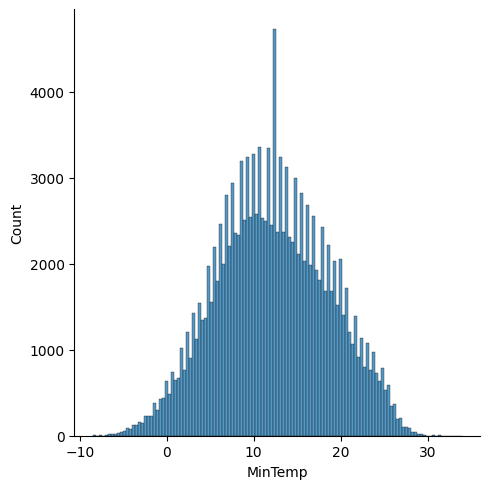
1. **The Histogram**

****

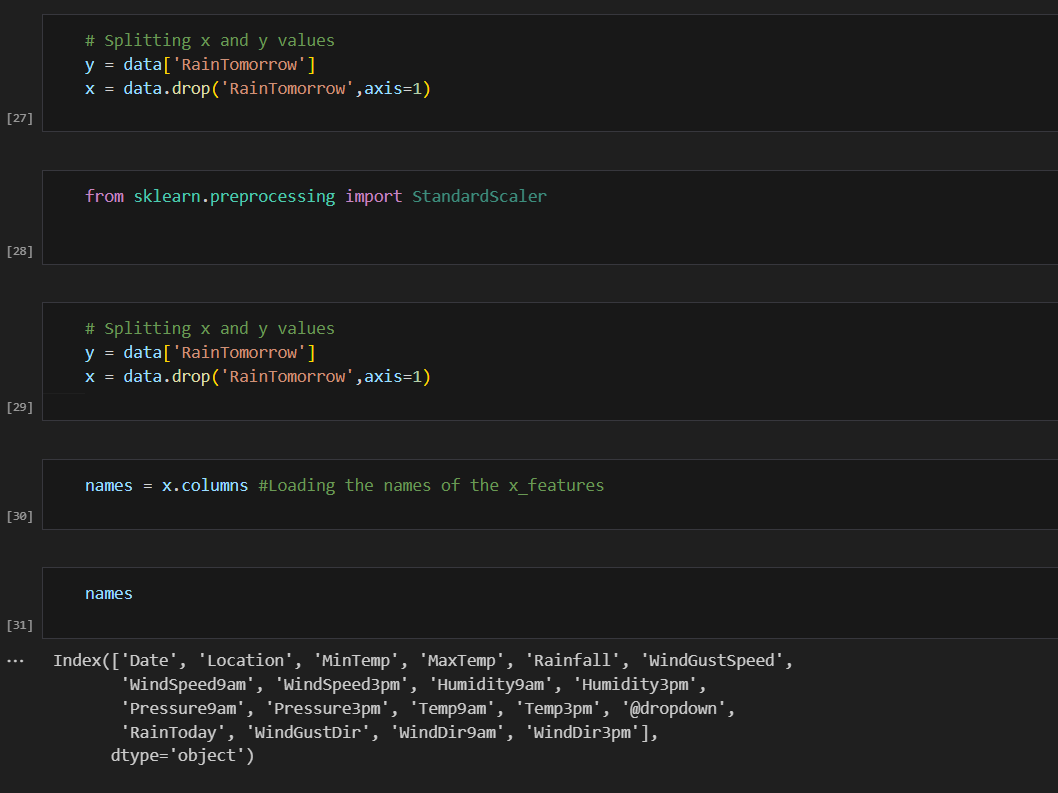
1. **The Scatter Plot**

****

1. **The Distribution Plot**

****

**Splitting the data:**

****

**# Splitting the data into train and test**

**x\_train,x\_test,y\_train,y\_test = model\_selection.train\_test\_split(x,y,test\_size=0.2, random\_state=0)**