Python Project

The **weatherAUS.csv** dataset contains $n = 36,881$ daily observations from 45 Australian weather stations. The weather data were obtained from the Australian Commonwealth Bureau of Meteorology. The data has been processed to provide a binary target variable RainTomorrow (whether there is rain during the next day; No/Yes) and a continuous target (risk) variable RISK_MM (the amount of rain recorded during the next day). The data set includes the following variables:

Date: The date of observation (a date object).
Location: The common name of the location of the weather station.
MinTemp: The minimum temperature in degrees centigrade.
MaxTemp: The maximum temperature in degrees centigrade.
Rainfall: The amount of rainfall recorded for the day in millimeters.
Evaporation: Class A pan evaporation (in millimeters) during 24 h (until 9 AM).
Sunshine: The number of hours of bright sunshine in the day.
WindGustDir: The direction of the strongest wind gust in the 24 h to midnight.
WindGustSpeed: The speed (in kilometers per hour) of the strongest wind gust in the 24 h to midnight.
WindDir9am: The direction of the wind gust at 9 AM.
WindDir3pm: The direction of the wind gust at 3 PM.
WindSpeed9am: Wind speed (in kilometers per hour) averaged over 10 min before 9 AM.
WindSpeed3pm: Wind speed (in kilometers per hour) averaged over 10 min before 3 PM.
RelHumid9am: Relative humidity (in percent) at 9 AM.
RelHumid3pm: Relative humidity (in percent) at 3 PM.
Pressure9am: Atmospheric pressure (hpa) reduced to mean sea level at 9 AM
Pressure3pm: Atmospheric pressure (hpa) reduced to mean sea level at 3 PM.
Cloud9am: Fraction of sky obscured by cloud at 9 AM. This is measured in "oktas," which are a unit of eighths. It records how many eighths of the sky are obscured by cloud. A 0 measure indicates completely clear sky, while an 8 indicates that it is completely overcast.
Cloud3pm: Fraction of sky obscured by cloud at 3 PM; see Cloud9am for a description of the values.
Temp9am: Temperature (degrees C) at 9 AM
Temp3pm: Temperature (degrees C) at 3 PM.
RainToday: Integer 1 if precipitation (in millimeters) in the 24 h to 9 AM exceeds 1 mm, otherwise 0.
RISK_MM: The continuous target variable; the amount of rain recorded during the next day.

Tasks
Make an exploratory data analysis of the given data considering the fact to design regression models. Design various predictive models to predict tomorrow's rain amount (a continuous target variable). Use as many regression techniques as per my knowing. Create train-test split and compare the performance of the models on the out-of-sample (i.e., the test data). Also compare the training and test performances of the models.