BIKES CHAT GPT PROJECT

Environment: source maruenv/bin/activate ???

Para activarlo en powershell : ./maruenv/Scripts/Activate.ps1

A company that rents out bikes wants to use a model to predict the number of bikes that will be rented out on a given day. The model will take into account various factors such as the date (month, day, whether it is a holiday or weekend), as well as weather data (temperature, precipitation, etc.) to make accurate predictions about bike rental demand. This will help the company to better plan for staffing and bike inventory, and optimize revenue. My dataset is located in "../../data/raw/daily-bike-share.csv". Can you write python for me that i can read this csv into a pandas df?

Can you explain this dataset and how we can create a prediction model for the rental? These are the columns in the dataset: 'instant', 'dteday', 'season', 'yr', 'mnth', 'holiday', 'weekday', 'workingday', 'weathersit', 'temp', 'atemp', 'hum', 'windspeed', 'rentals'

a brief description of each column in the dataset:

* 'instant': A unique identifier for each row
* 'dteday': The date of the observation
* 'season': The season in which the observation was taken (1=spring, 2=summer, 3=fall, 4=winter)
* 'yr': The year in which the observation was taken (0=2011, 1=2012)
* 'mnth': The month in which the observation was taken (1 to 12)
* 'holiday': Whether or not the day was a holiday (0=no, 1=yes)
* 'weekday': The day of the week (0=Sunday, 1=Monday, ..., 6=Saturday)
* 'workingday': Whether or not the day was a working day (0=no, 1=yes)
* 'weathersit': The weather situation on the day of the observation (1=clear, 2=mist/cloudy, 3=light rain/snow, 4=heavy rain/snow)
* 'temp': The normalized temperature in Celsius
* 'atemp': The normalized feeling temperature in Celsius
* 'hum': The normalized humidity
* 'windspeed': The normalized wind speed
* 'rentals': The number of bike rentals for that day

To predict the daily number of bike rentals based on the data from the dataset we can use various machine-learning algorithms, such as regression or decision trees.

If I had to use the dteday feature, its an object and must convert it to numeric as to normalize it or use like that in the ml model, now I will delete it because there are already the variables year and month of the observation

## Convert date column to datetime type ()

X['dteday'] = pd.to\_datetime(X['dteday'])

## Extract year, month, and day features from the datetime column

X['year'] = X['dteday'].dt.year

X['month'] = X['dteday'].dt.month

X['day'] = X['dteday'].dt.day

## Convert the year, month, and day columns to numeric type

X['year'] = pd.to\_numeric(X['year'])

X['month'] = pd.to\_numeric(X['month'])

X['day'] = pd.to\_numeric(X['day'])