

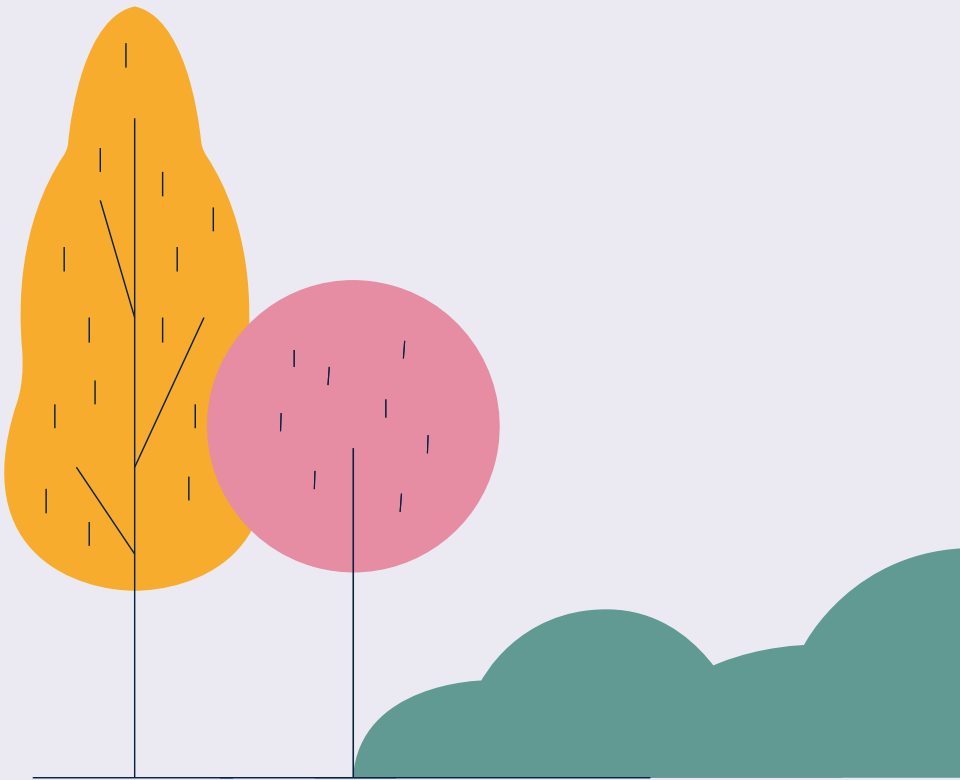
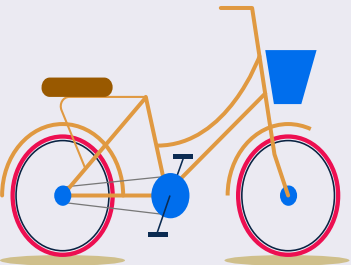
Bike sharing





Overview

Capital Bike Sharing dataset contains information related to the bike sharing program of Washington DC for the 2011 and 2012 years. To review the data set attributes, visit the [data set page on UCI](#).





data!

The dataset contains 17 attributes , 17389 instances

Problem statements



For predict

Prediction of bike rental count hourly based on the environmental and seasonal settings



For visualize

- Bike years highest season
- Top demand hours for Bikes
- Humidity and temperature effect on bike rental

Data cleaning

Check any outlier and handle it

Check if have any inconsistent

1

2

3

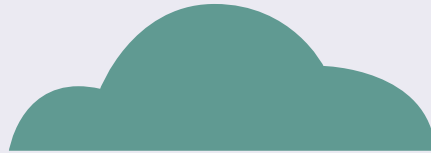
4

5

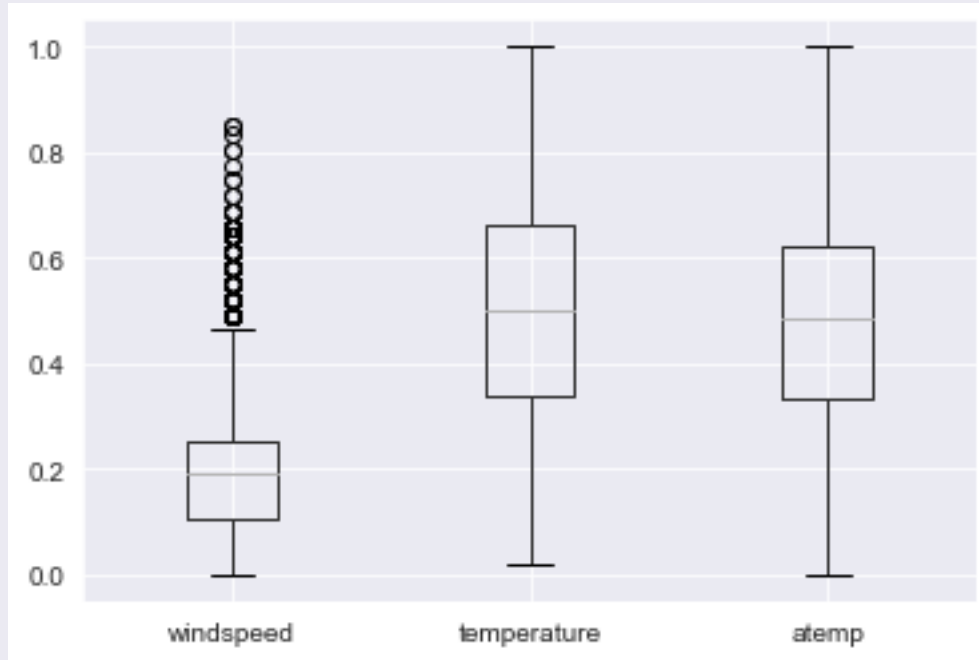
Check Missing & duplicate data

Rename column

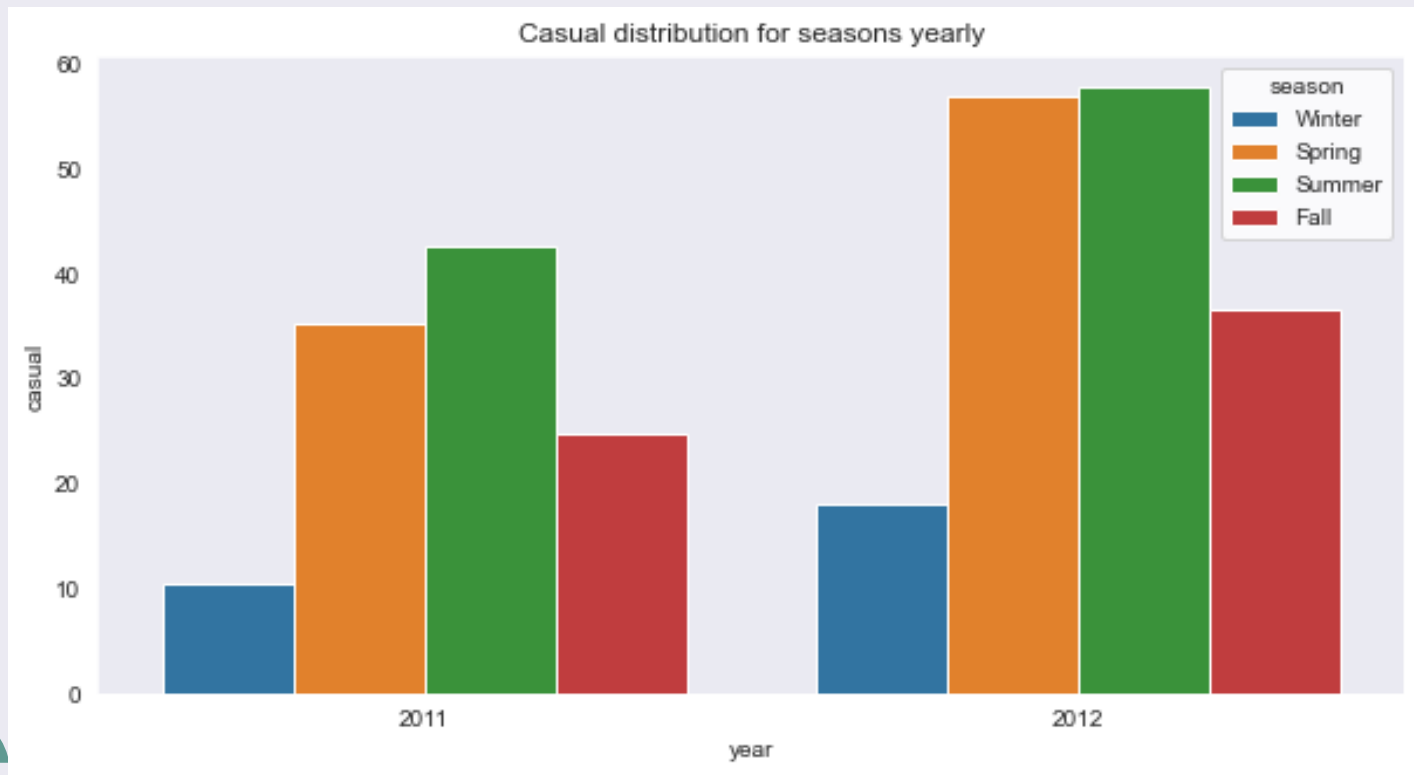
Map data nice to visualize



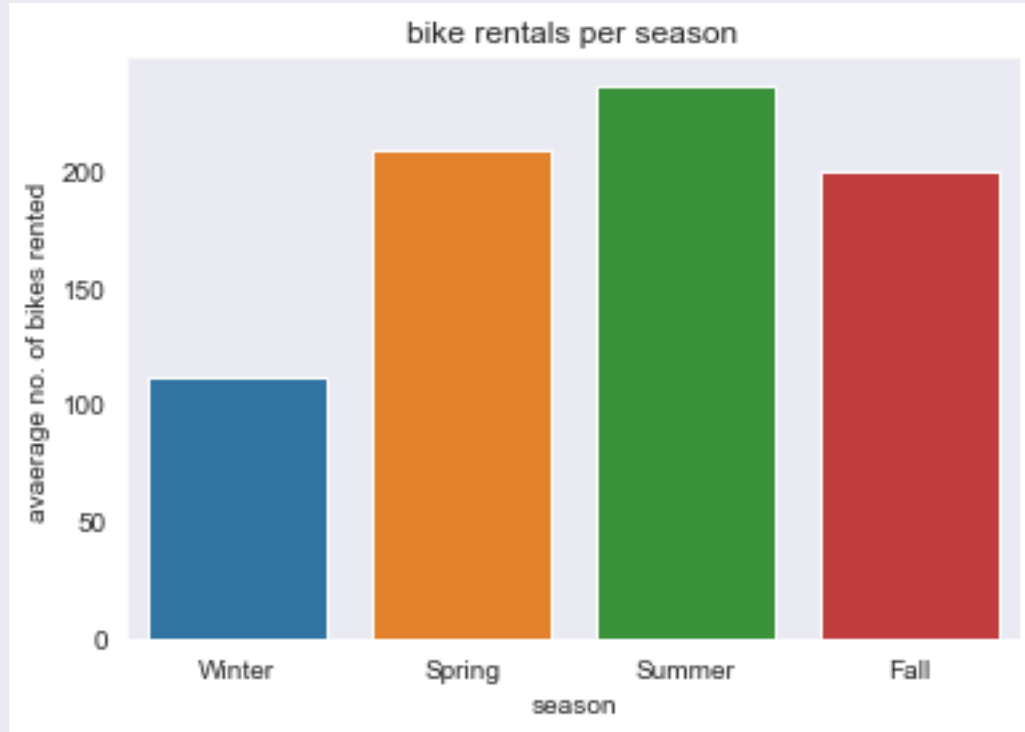
check outlier



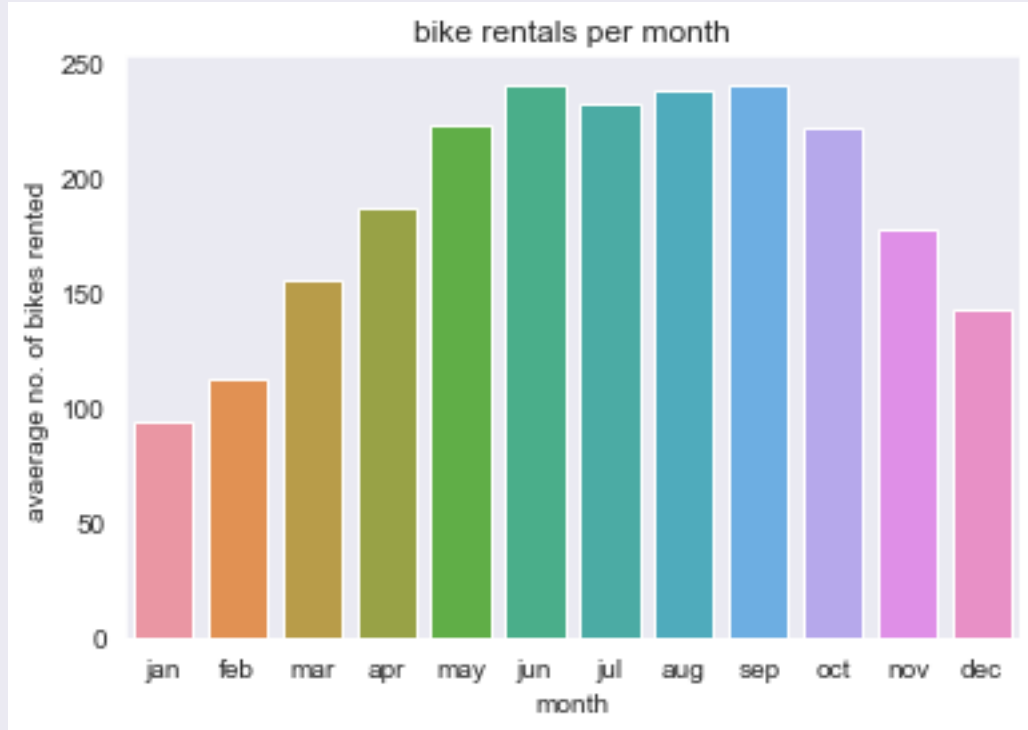
Year



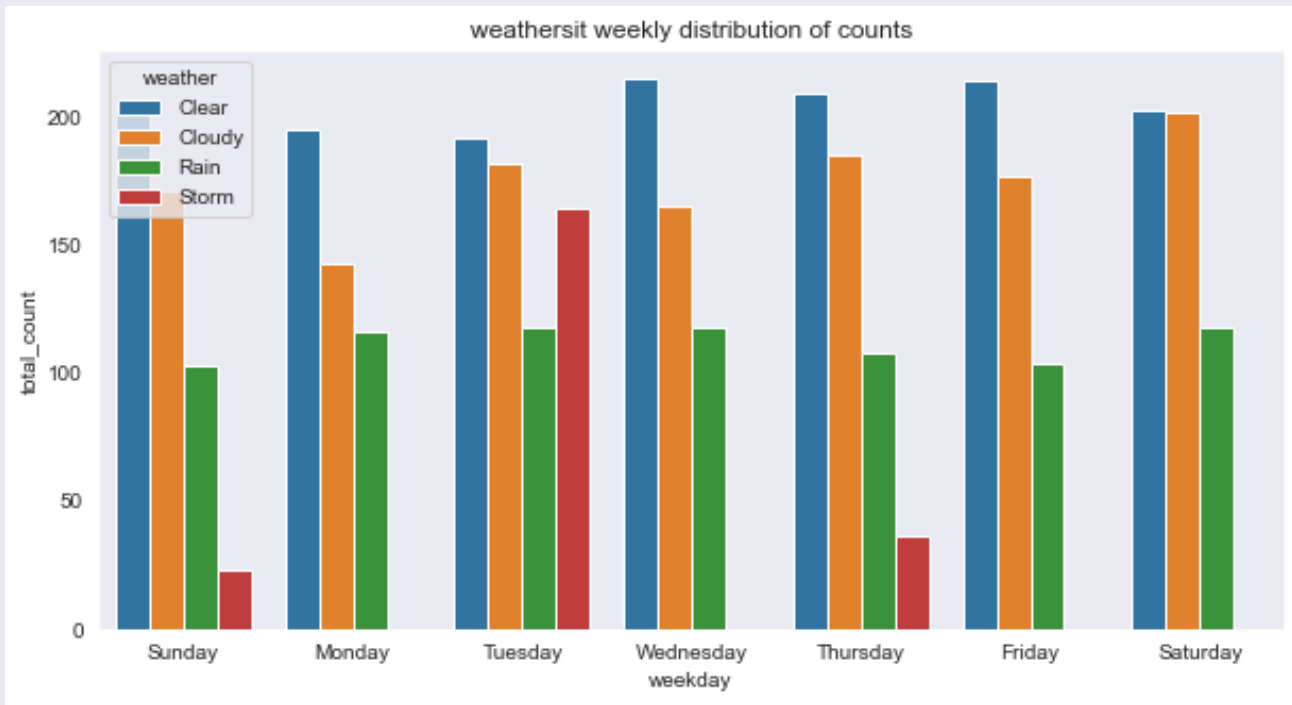
seasons



months

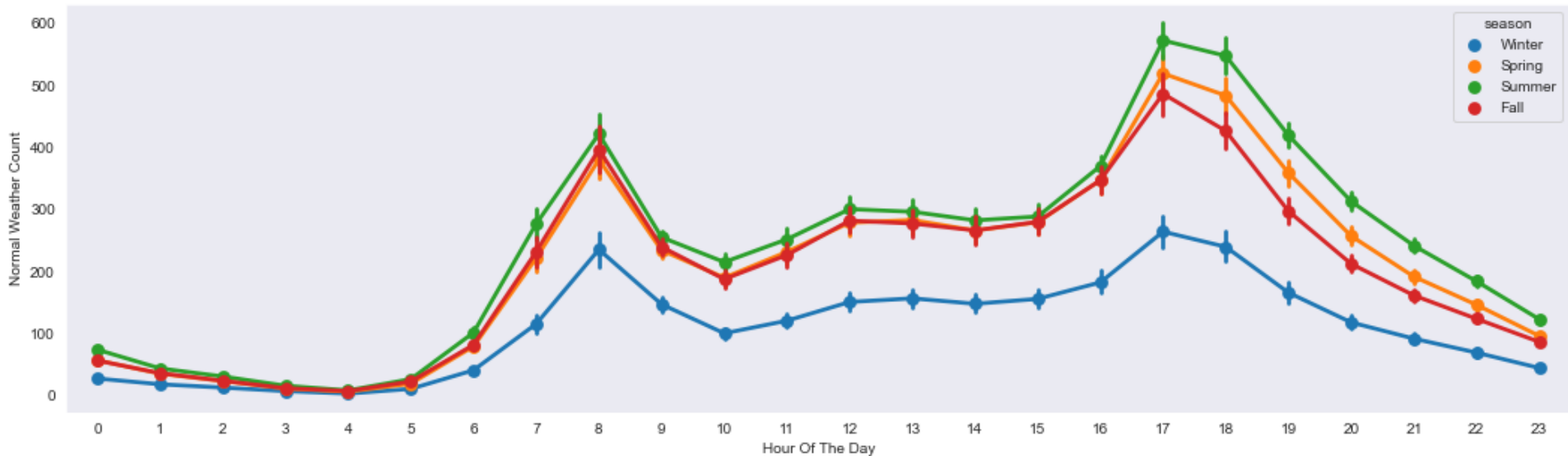


days

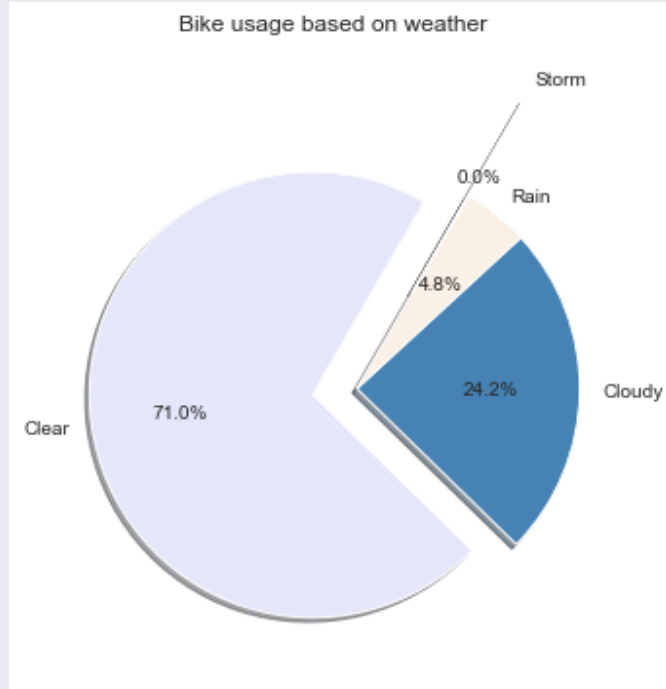


hours

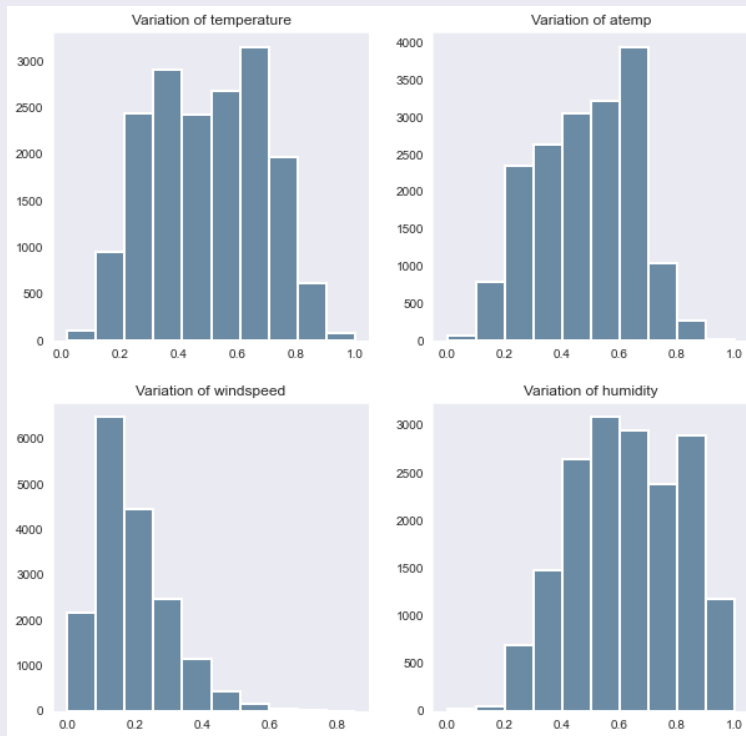
Normal Weather By Hour Of The Day Across Season



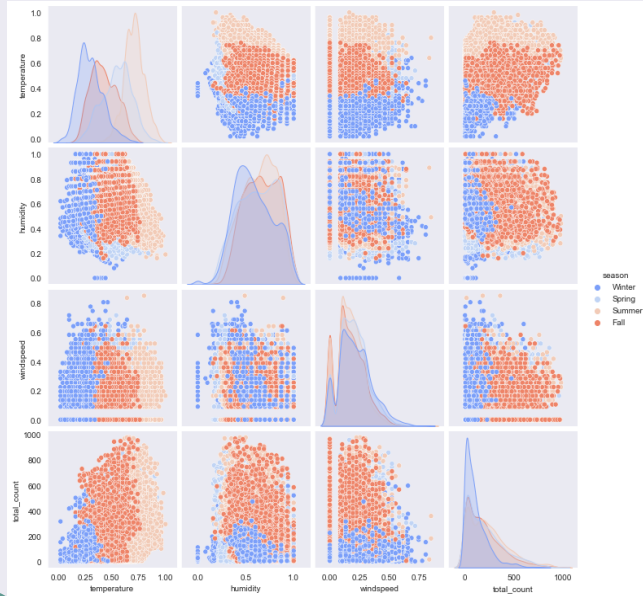
Bike usage on weather



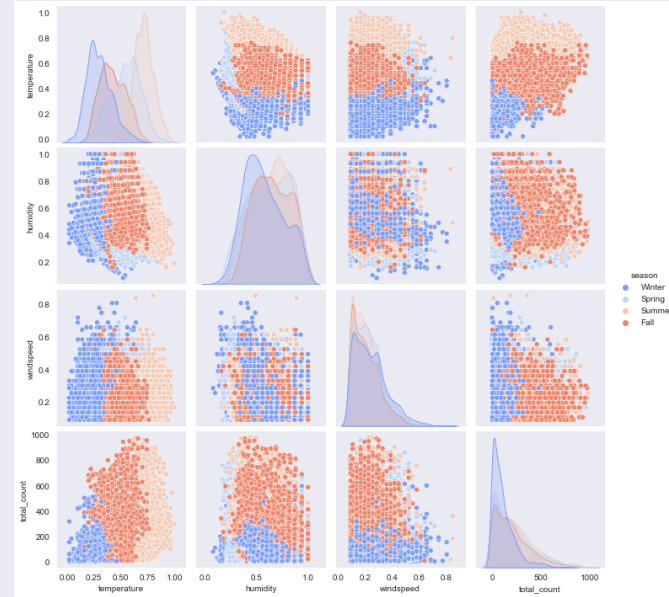
Humidity and temperature effect



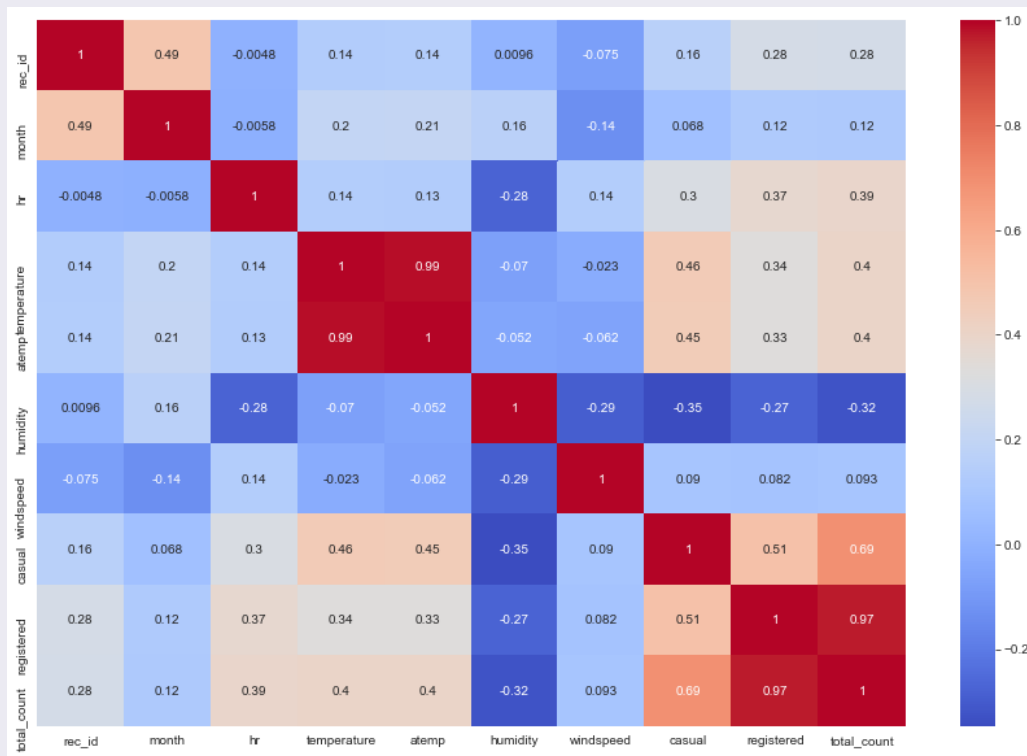
Winds speed after fill by KNN



Winds speed after fill by KNN



Visualize correlation



Drop unnecessary column

Record Id

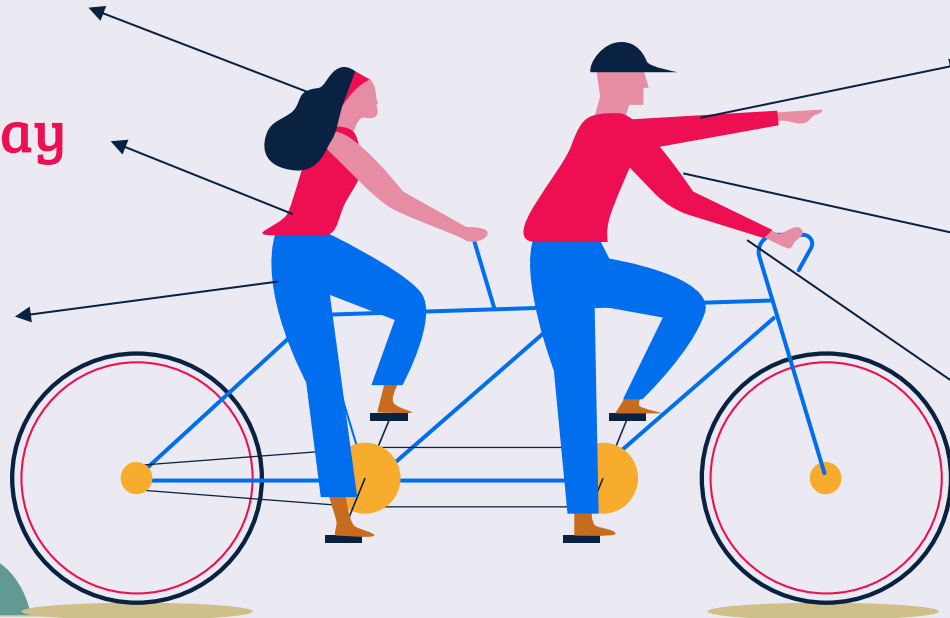
Date day

year

Rigistered &
casual

Holiday &
weekday

atemp



Feature engineering

Encode categorical column

Hour & season
Drop first column

Fill zero value in winds speed

by KNN neighbors value

Scaler numerical column

By standard scaler

Cross validation

By 5 Kfolds




Models

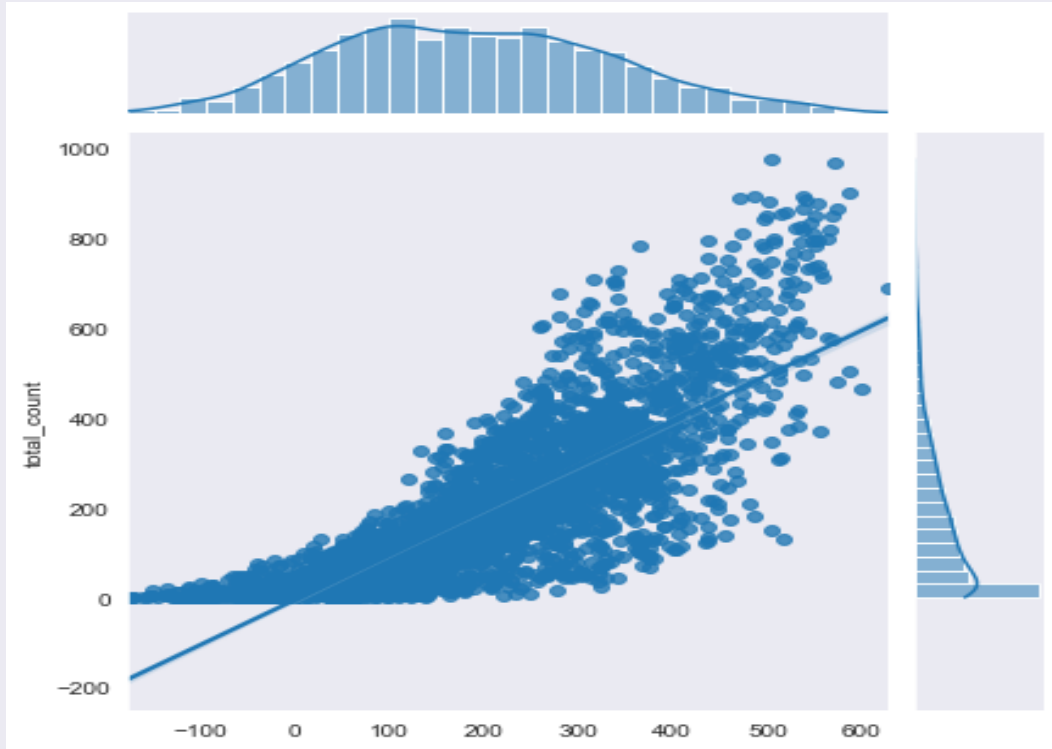


R square

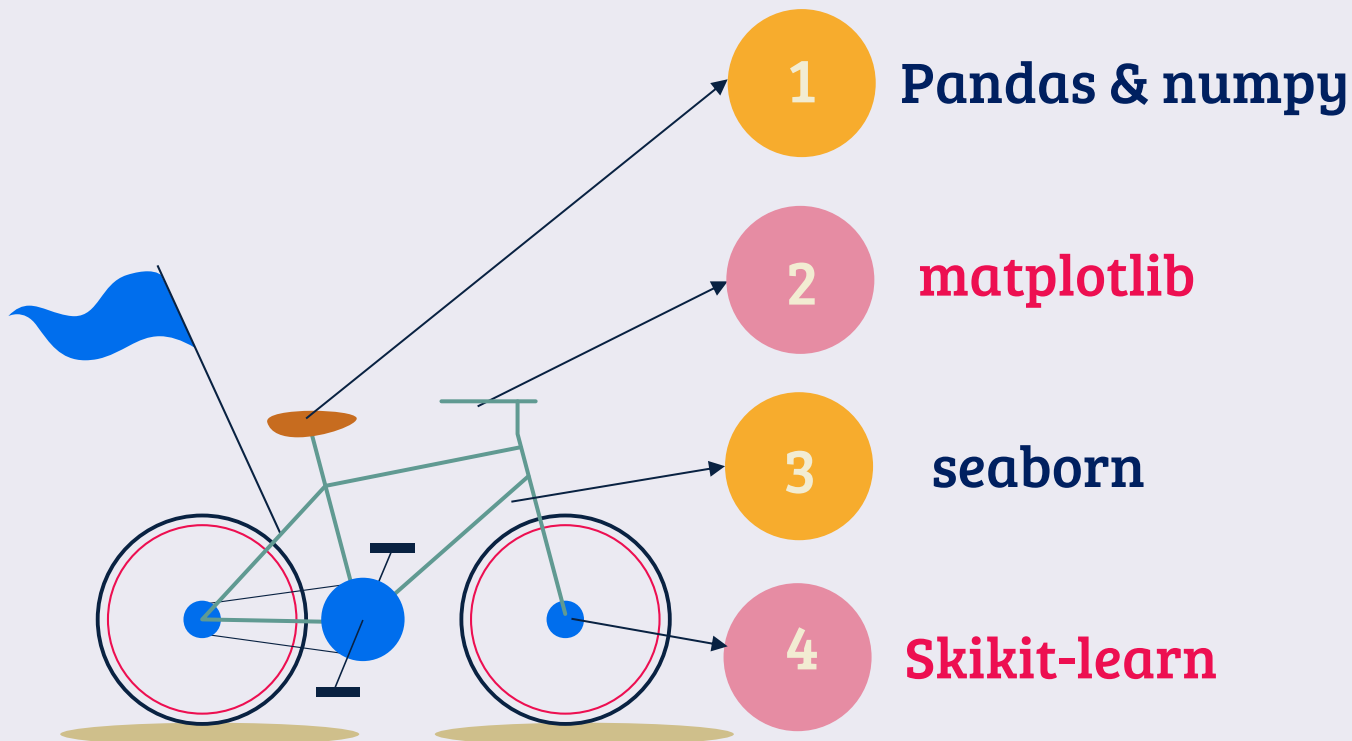
Linear Regression	0.69
Ridge Regression	0.69
Lasso Regression	0.69
Polynomial degree (2)	0.90



Result



Tools





Our best experiences


66 Route

Mercury is the closest planet to the Sun

East Coast

Venus has a beautiful name, but it's very hot





“The best way to predict your
future is to create it.”

Thanks

Do you have any questions?