DOCUMENT OBSERVATIONS ABOUT DATA QUALITY AND COMPLETENESS

THE OBJECTIVE OF THIS CASE STUDY IS THE PREDICTION OF BIKE,CAR,BICYCLE RENTAL COUNT ON DAILY BASED ON THE

ENVIRONMENTAL AND SEASONAL SETTINGS. THE DATASET(1,2,3) CONTAINS 1000 OBSERVATIONS, 16 PREDICTORS AND 1 TARGET VARIABLE. THE PREDICTORS ARE DESCRIBING VARIOUS ENVIRONMENT FACTORS AND SETTINGS LIKE SEASON, HUMIDITY ETC. WE NEED TO BUILD A PREDICTION MODEL TO PREDICT ESTIMATED COUNT OR DEMAND OF BIKES ON A PARTICULAR DAY BASED ON THE ENVIRONMENTAL FACTORS

DATASET

THE DATA SET CONSIST OF 1000 OBSERVATION RECORDED OVER A PERIOD YEARS IT HAS 16 PREDICTORS OR VARIABLES AND 1 TARGET VARIABLE. ALL THE VARIABLES ARE DESCRIBED IN TABLE 1.

Variable names	Description
instant	Record index
dteday	Date
season	Season (1:springer, 2:summer, 3:fall, 4:winter)
yr	Year
mnth	Month (1 to 12)
hr	Hour (0 to 23)
holiday	Weather day is holiday or not (extracted from Holiday Schedule)
weekday	Day of the week
weathersit	1: Clear, Few clouds, Partly cloudy, Partly cloudy 2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist 3: Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds 4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog
temp	Normalized temperature in Celsius.
atemp	Normalized feeling temperature in Celsius.
hum	Normalized humidity
windspeed	Normalized humidity
casual	Count of casual users
registered	Count of registered users
cnt	Count of total rental bikes including both casual and registered