SEOUL BIKE RENTAL CASE STUDY
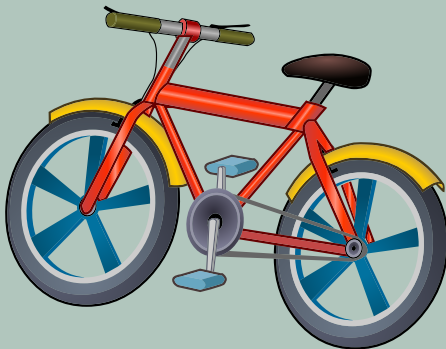
Malhar Pandya
1005893008

Tan Lin
1005102871

Anish Singla
1005801401

STEP 1. INTRODUCTION

- Brief background about rental bikes
- When do people in Seoul use bikes?
- Our process

STEP 2. DATA CLEANING

- Extracted Dates from the Date Column
- Dropped Dew Point because of Multicollinearity
- Dropped Nonfunctioning days because it wasn't relevant to our question
- Adjusted Hours to better model rental bike usage.
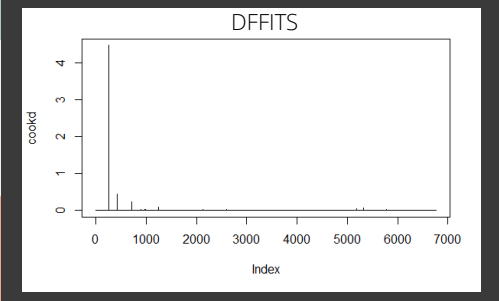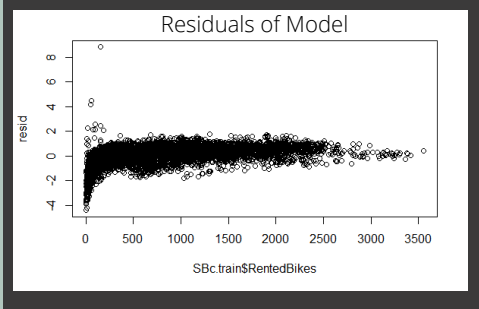- Created addional variables to model behavior.
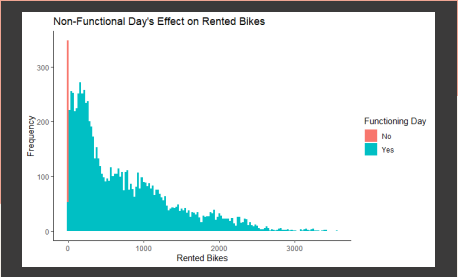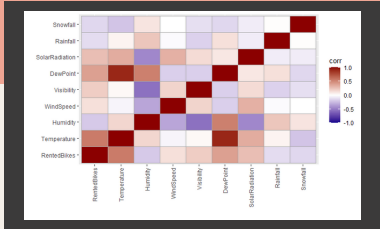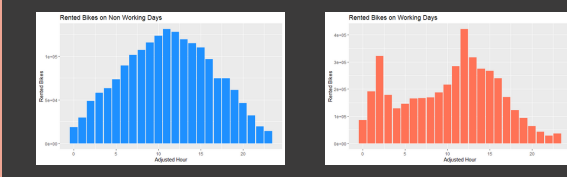
STEP 3. MODEL SELECTION

We first broke the data into train and test.
We ended up predicting **log(RentedBikes)** using:

- **Hour** + **I(Hour^2)**
- **Temperature** + **Humidity** + **Rainfall**
- **Working** + **RushHour**

We chose this model because it provided a good balance between having a low number of predictors and a high adjusted R^2 (0.63)
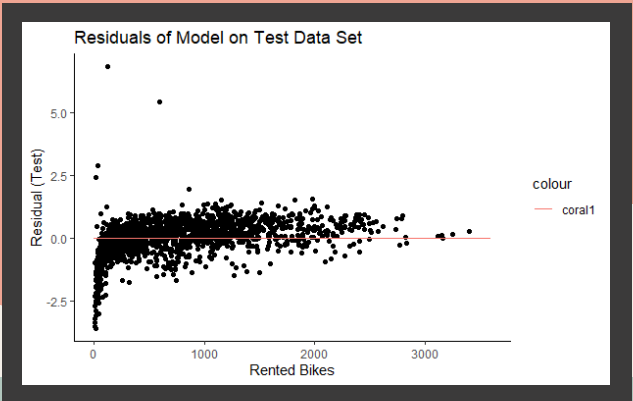
STEP 4. MODEL VALIDATION

We used the model on the validation data set and got:

- Adjusted R^2 = 0.55
- RMSE = 965.24
- MAE = 718.96

Our model has problems...

STEP 5. CONCULSION

- Model Conclusions
- Limitations