Seoul Bike Share

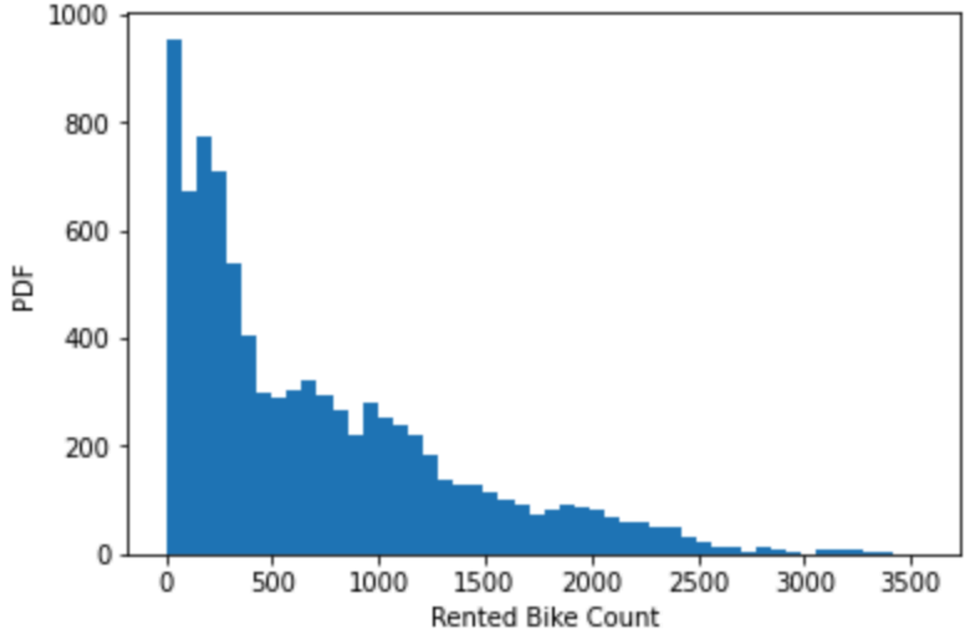
Elinor Velasquez, April 14 2021

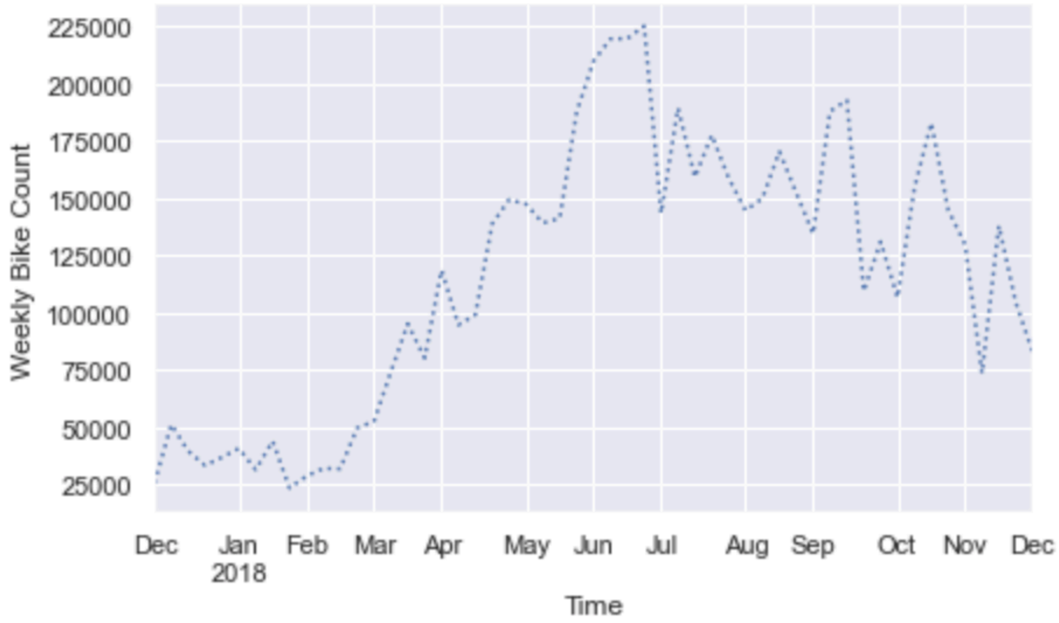
Can we predict rented bike count in Seoul, South Korea?

Data Collection:
Rented Bike Share count per hour in Seoul, South Korea

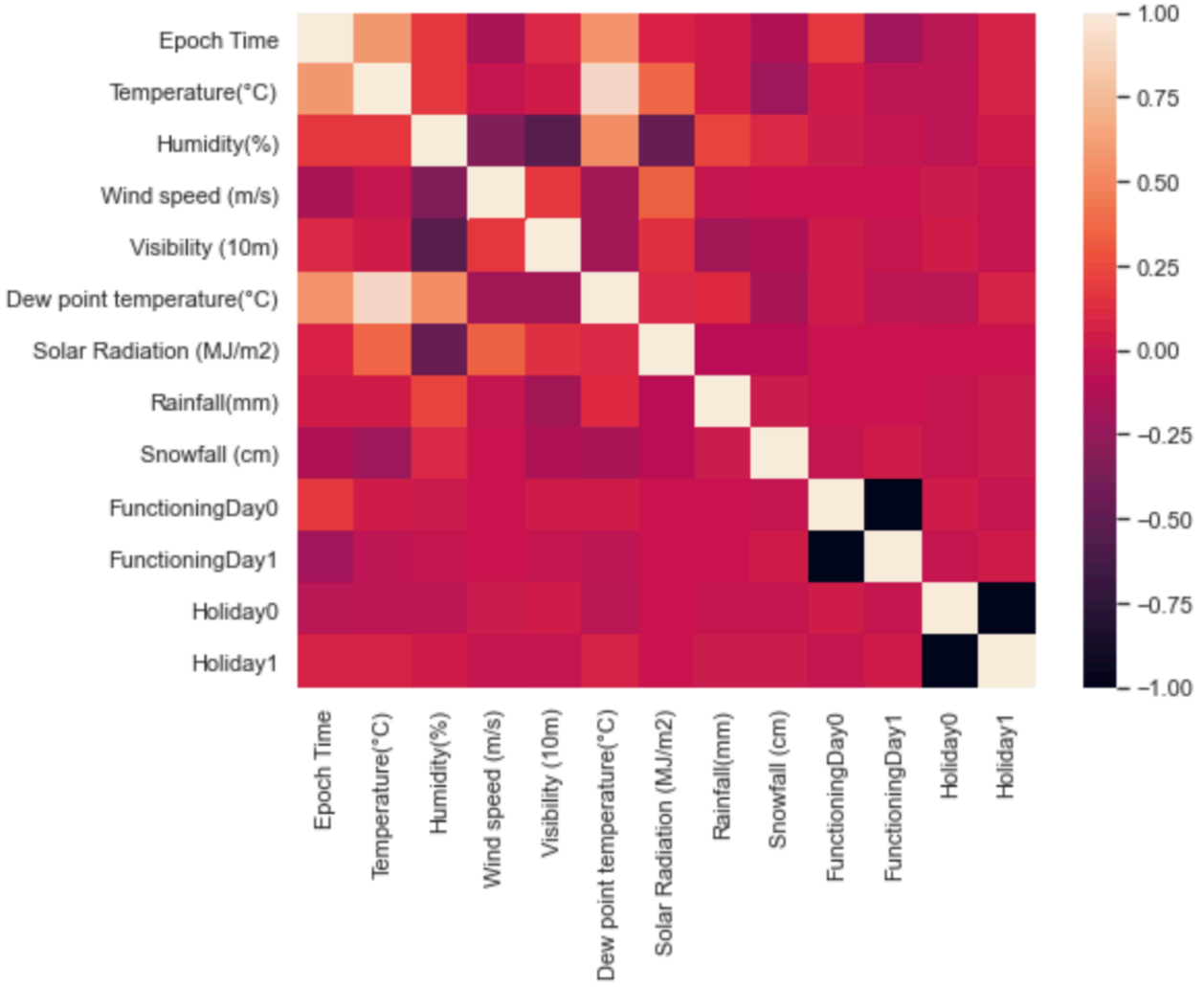
Data Retrieval:
U.C. Irvine Data Repository (Data deposited in 2020-03-01)

Can we predict weekly rented bike count?

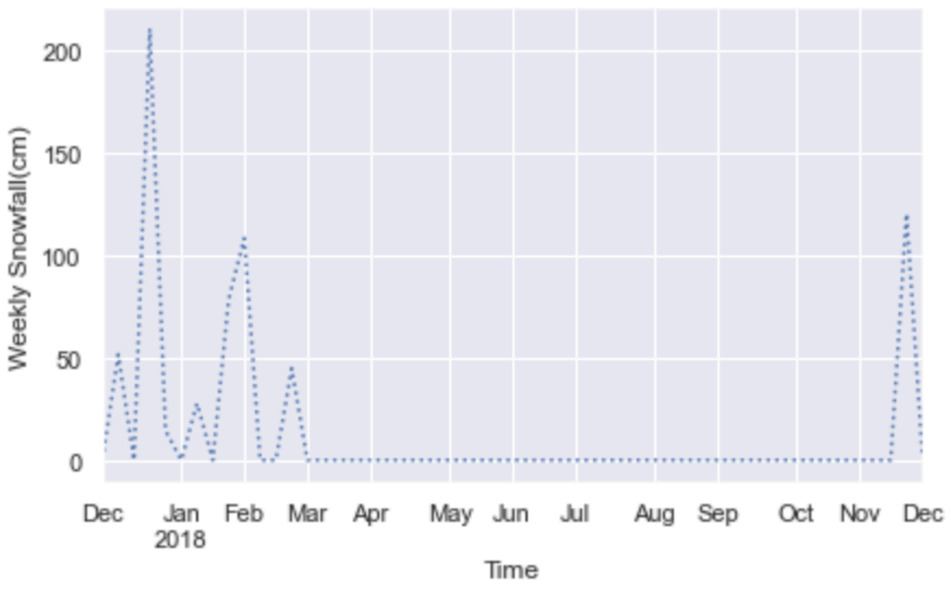


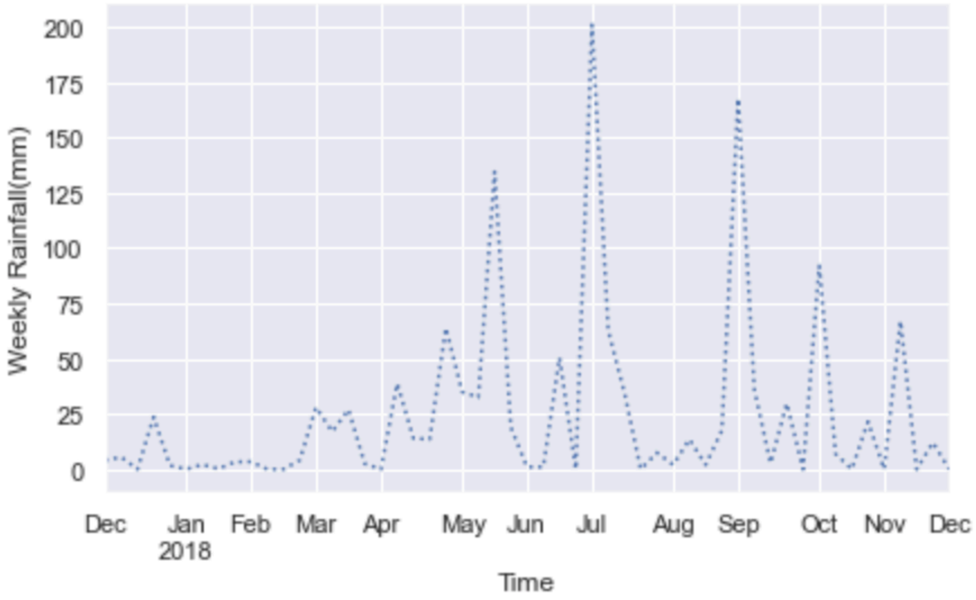


The HeatMap shows that several features are correlated

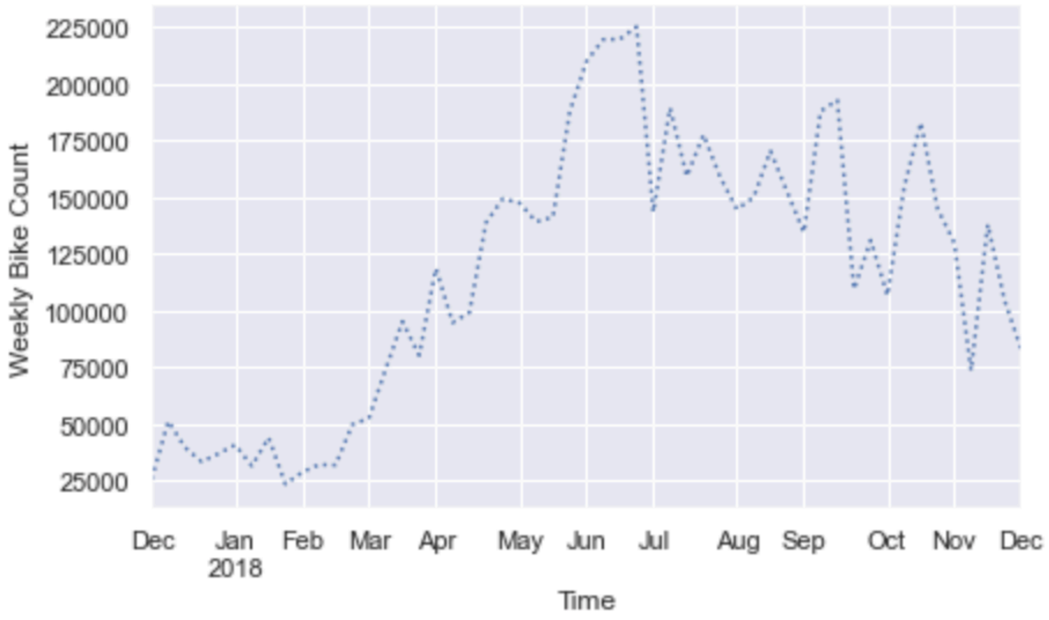


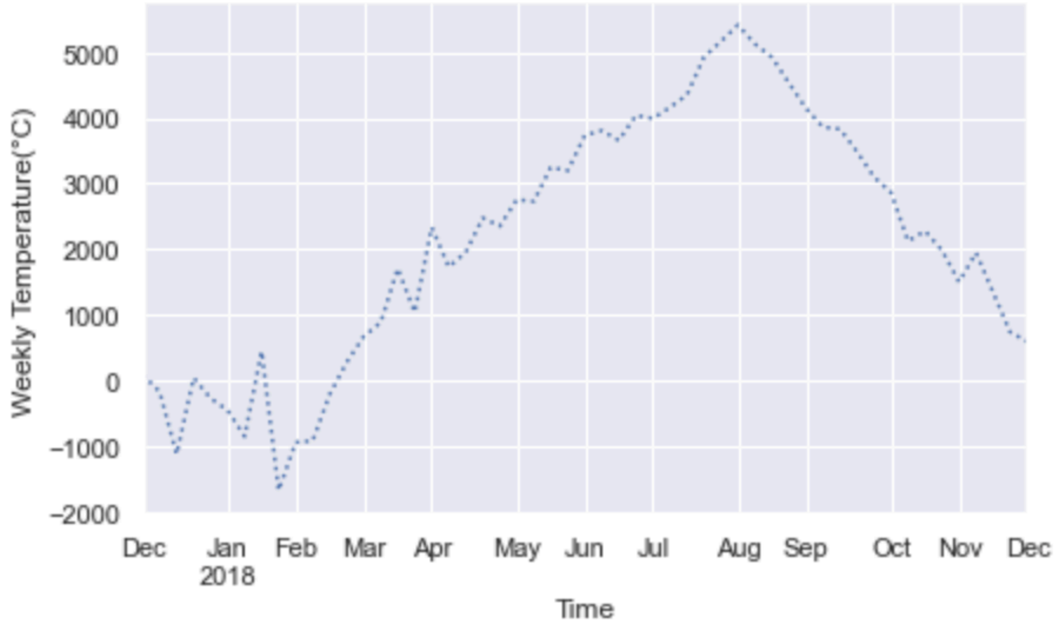
Rainfall and Snowfall in Seoul is a constant



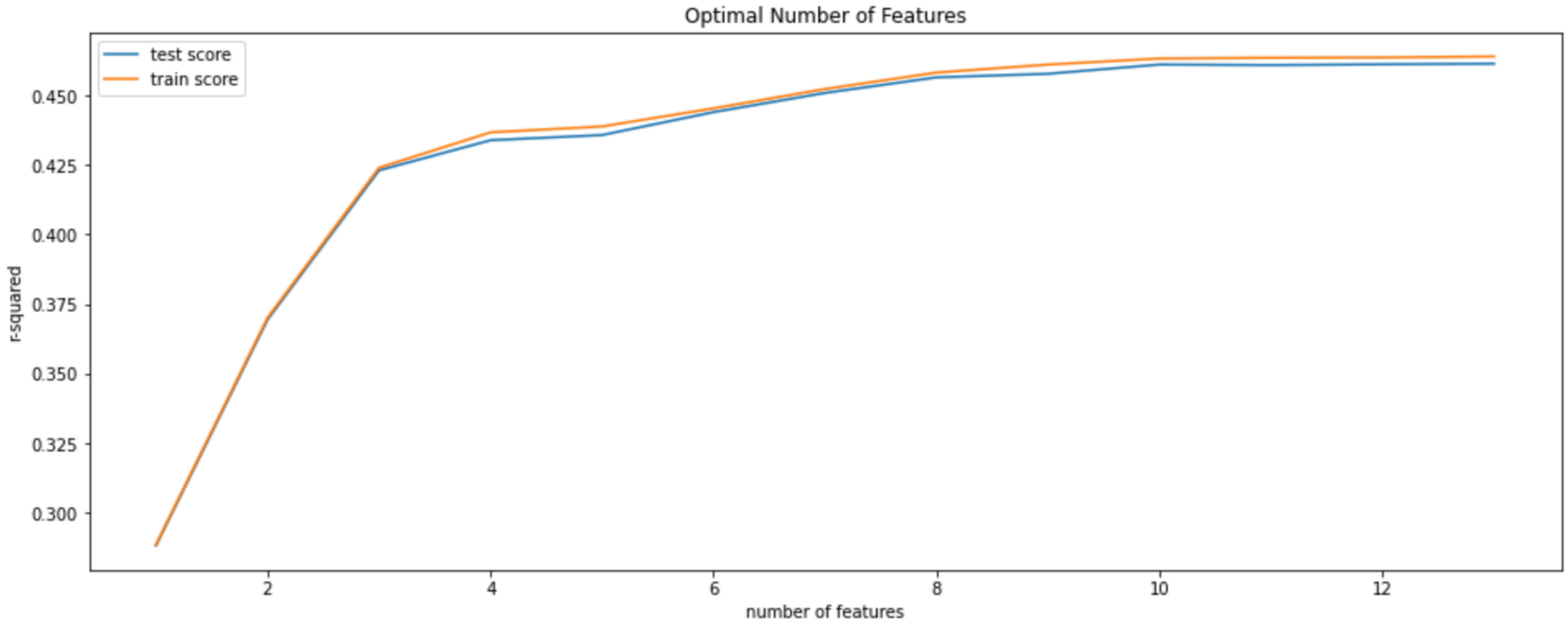


Weekly Bike Rental corresponds to Weekly Temperature

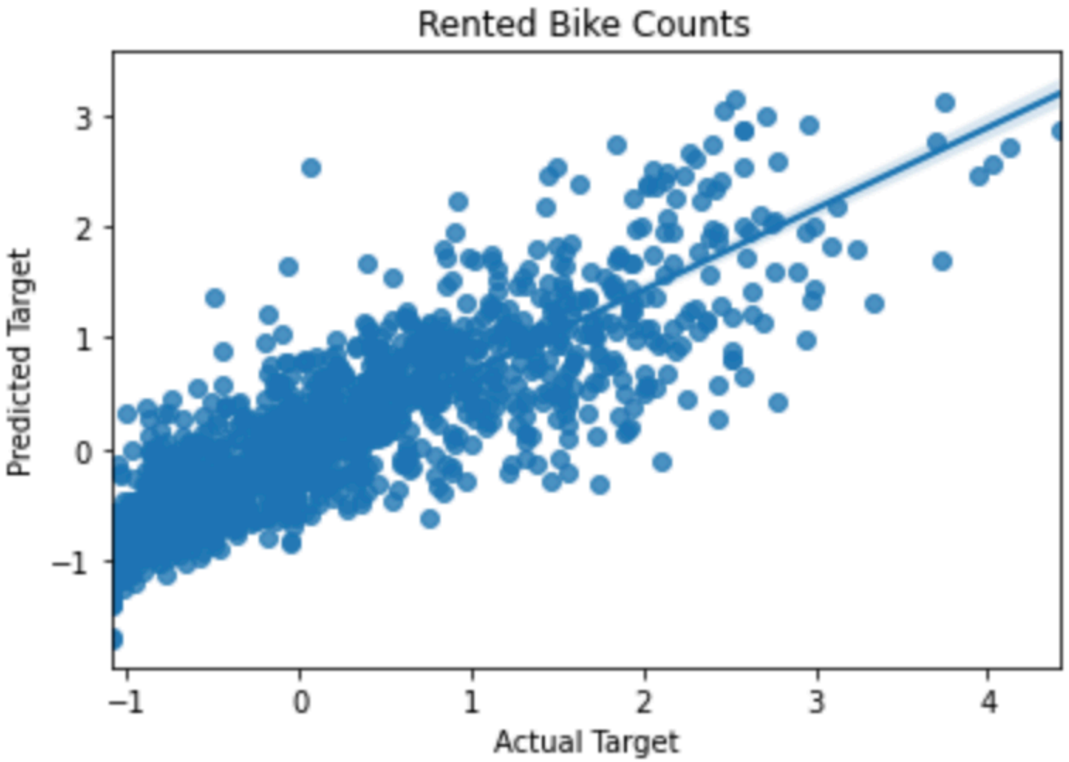




Hyperparameter Tuning Optimizes the Model of the Data



Predicted Target corresponds to Actual Target



Best Model:
CatBoost Regressor

R-squared = 0.75

Workflow

Seoul Bike Share

- *Problem Identification & Formulation: Predict bike rental count*
- *Data Cleaning & Understanding and EDA: Pandas Profiling used to explore the dataset, Missing values confirmation, Categorical variables confirmed/transformed, Normalization of data*
- *Modeling & Evaluation: Linear Regression, Multiple Linear Regression, Support Vector Machine Regression, Multi-layer Perceptron Classifier, CatBoost Regressor (Hyper-parameters: Depth, Bagging Temperature)*
- *Findings & Recommendations: CatBoost Regressor (0.75 Accuracy)*

Future Work

Seoul Bike Share

- Advanced time series modeling
- Step 1: Visualize the data via time-series decomposition: Decompose the data into three components (trend, seasonality, noise)
- Step 2: Forecasting
- Time series forecasting with ARIMA (statsmodels)
- Time series forecasting with Facebook's Prophet