Understanding Factors Affecting Bike Sharing Demand in Seoul

Ahmad Sadeed (asadeed2), Deepa Nemmili Veeravalli (deepan2), Rui Zou (ruizou4)

2022-07-16

Description of the data file

This data file contains count of public bikes rented at each hour in Seoul Bike Sharing System with the corresponding weather data and holidays information. It has 14 variables and 8760 observations. We are interested in using Rented.Bike.Count (a numeric variable) as our response variable and explore how other factors (3 categorical variables and several continuous numeric variables) affect the count of bikes rented at each hour. Among the other 13 variables which we plan to use as potential predictors, we know from intuition that some may have more importance than others, like temperature, humidity, wind speed, visibility, seasons, and holiday, etc.

Background information on the data set

The original data comes from http://data.seoul.go.kr. The holiday information comes from SOUTH KOREA PUBLIC HOLIDAYS. A clean version can be found at UCI Machine Learning Repository.

Attribute Information:

- Date : month/day/year
- Rented Bike count - Count of bikes rented at each hour
- Hour - Hour of the day
- Temperature - Temperature in Celsius
- Humidity - %
- Windspeed - m/s
- Visibility - 10m
- Dew point temperature - Celsius
- Solar radiation - MJ/m2
- Rainfall - mm
- Snowfall - cm
- Seasons - Winter, Spring, Summer, Autumn
- Holiday - Holiday, No holiday
- Functional Day - Functional or Non-functional days of rental bike system

Our Interest

This data set is interesting to us both personally and business-wise. Recently we have seen a rise in the delivery, accessibility, and usage of regular and electric rental bikes. There are clear environmental, health, and economical benefits associated with the usage of bikes as a mode of transportation. We would like to find out what factors lead to an increase in number of bikes rented and what factors have inverse effect on using

rental bikes. Learning about such factors can help a bike rental business manage its inventory and supply without any hindrance. It can also help cities plan accordingly due to an increase of bikers, e.g. opening up more bike lanes during certain days or seasons. Environmentally, we will have a better understanding of the feasibility of turning a city into a "bike city" or looking at alternative options if a city is not friendly to bikers due to harsh weather conditions.

Data in R

The data file can be successfully loaded into R. We have printed out the structure and first few rows of the data file below.

The column names in the `csv` file contains measurement units (like `Wind speed (m/s)`, `Solar Radiation (MJ/m2)`) and characters such as ° and %. We load the data using cleaned up column names.

```
columns = c("Date","Rented.Bike.Count","Hour","Temperature","Humidity",
            "Wind.Speed","Visibility","Dew.point.temperature",
            "Solar.Radiation","Rainfall","Snowfall","Seasons","Holiday",
            "Functioning.Day")
bike = read.csv("../data/SeoulBikeData.csv", col.names = columns)
str(bike)
```

```
## 'data.frame':    8760 obs. of  14 variables:
##  $ Date                 : chr  "01/12/2017" "01/12/2017" "01/12/2017" "01/12/2017" ...
##  $ Rented.Bike.Count    : int  254 204 173 107 78 100 181 460 930 490 ...
##  $ Hour                 : int  0 1 2 3 4 5 6 7 8 9 ...
##  $ Temperature          : num  -5.2 -5.5 -6 -6.2 -6 -6.4 -6.6 -7.4 -7.6 -6.5 ...
##  $ Humidity             : int  37 38 39 40 36 37 35 38 37 27 ...
##  $ Wind.Speed           : num  2.2 0.8 1 0.9 2.3 1.5 1.3 0.9 1.1 0.5 ...
##  $ Visibility           : int  2000 2000 2000 2000 2000 2000 2000 2000 2000 1928 ...
##  $ Dew.point.temperature: num  -17.6 -17.6 -17.7 -17.6 -18.6 -18.7 -19.5 -19.3 -19.8 -22.4 ...
##  $ Solar.Radiation      : num  0 0 0 0 0 0 0 0 0.01 0.23 ...
##  $ Rainfall             : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Snowfall             : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ Seasons              : chr  "Winter" "Winter" "Winter" "Winter" ...
##  $ Holiday              : chr  "No Holiday" "No Holiday" "No Holiday" "No Holiday" ...
##  $ Functioning.Day      : chr  "Yes" "Yes" "Yes" "Yes" ...
```

```
head(bike)
```

```
##         Date Rented.Bike.Count Hour Temperature Humidity Wind.Speed Visibility
## 1 01/12/2017               254    0        -5.2       37        2.2       2000
## 2 01/12/2017               204    1        -5.5       38        0.8       2000
## 3 01/12/2017               173    2        -6.0       39        1.0       2000
## 4 01/12/2017               107    3        -6.2       40        0.9       2000
## 5 01/12/2017                78    4        -6.0       36        2.3       2000
## 6 01/12/2017               100    5        -6.4       37        1.5       2000
##   Dew.point.temperature Solar.Radiation Rainfall Snowfall Seasons    Holiday
## 1                 -17.6              0        0        0  Winter No Holiday
## 2                 -17.6              0        0        0  Winter No Holiday
## 3                 -17.7              0        0        0  Winter No Holiday
## 4                 -17.6              0        0        0  Winter No Holiday
## 5                 -18.6              0        0        0  Winter No Holiday
## 6                 -18.7              0        0        0  Winter No Holiday
```

```
##   Functioning.Day
## 1             Yes
## 2             Yes
## 3             Yes
## 4             Yes
## 5             Yes
## 6             Yes
```