Data Wrangling

Tzu-Ying Chen

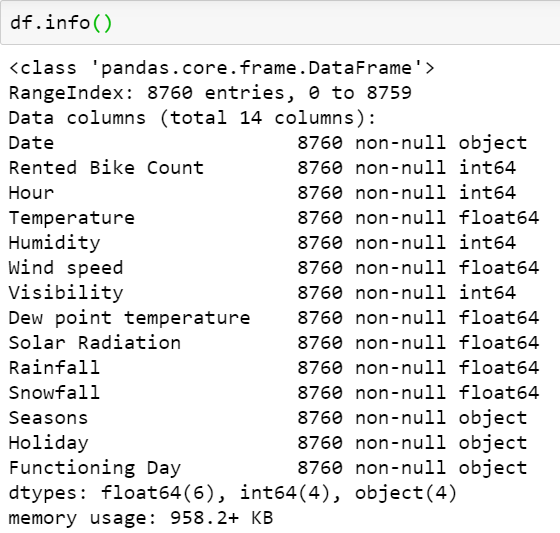
# Data Collection

The data for this capstone project contains public bikes rental information in Seoul Bike Sharing System and is acquired from UCI Machine Learning Repository. Link to the data source is below:

<https://archive.ics.uci.edu/ml/datasets/Seoul+Bike+Sharing+Demand#>

The data set is in csv format, which could be directly read into python console using python pandas library.

# Data Cleansing and Data Wrangling

The only thing that I need to do when reading dataset was to set the correct encoding when parsing the data using pd.read\_csv. The data set comes fairly clean, with a total of 14 columns and 8,760 data points without any null values in the data set. There were not outliers, and do not need to drop and/or fill any missing values.