

### Bike sharing in Seoul

- Currently Rental bikes are introduced in many urban cities for the enhancement of mobility comfort. It is important to make the rental bike available and accessible to the public at the right time as it lessens the waiting time. Eventually, providing the city with a stable supply of rental bikes becomes a major concern. **The crucial part is the prediction of bike count required at each hour for the stable supply of rental bikes.**
- The dataset contains weather information (Temperature, Humidity, Windspeed, Visibility, Dewpoint, Solar radiation, Snowfall, Rainfall), the number of bikes rented per hour and date information.

My mission, should I choose to accept it, is to analyse and exploit the dataset to predict to the best the number of bikes that have to be available in real time in the store to respond to the consumer request.

The work was carried out in 3 stages :

- Data Preparation
- Data Visualisation
- Modeling

The first stage was Data Preparation : wich means I had to sort all the information available in the dataset. Checking and delete potential mistakes that could make our work wrong. So we check null values and NaN values to clean the dataset and make it more exploitable.

The second stage is Data Visualisation. It consists of creating graphes and curves in order to visualise the best way all the information contained in the csv. Thus, we can see all the parameters correlated to the number of bikes rented. This means that these are the parameters that influence the number of rented bikes in Seoul.

Finally, the third stage is Modeling. In order to do that we have to :

- Scale data
- Prepare date for future Regression
- Perform a grid search to search for the best model and best hyper parameters

Then we establish the best prediction model to guess at the best the number of bikes that have to be available.

Thanks to this work, i can conclude that the best prediction model is Decision Tree, is allows us to have the minium error during the cross validation. It is the most efficient way to predict the number of bikes the store will have to rent.

Using a new parameter, we can obtain better results. We add a column with the number of bikes rented during the last hour. The company might not always have this information, but this could lead to a strategy of live gestion of bikes in the city.