You are working with the government to transform your city into a smart city. The vision is to convert it into a digital and intelligent city to improve the efficiency of services for the citizens. One of the problems faced by the government is traffic. You are a data scientist working to manage the traffic of the city better and to provide input on infrastructure planning for the future.

The government wants to implement a robust traffic system for the city by being prepared for traffic peaks. They want to understand the traffic patterns of the four junctions of the city. Traffic patterns on holidays, as well as on various other occasions during the year, differ from normal working days. This is important to take into account for your forecasting.

**Your task**

To predict traffic patterns in each of these four junctions for the next 4 months.

The sensors on each of these junctions were collecting data at different times, hence you will see traffic data from different time periods. To add to the complexity, some of the junctions have provided limited or sparse data requiring thoughtfulness when creating future projections. Depending upon the historical data of 20 months, the government is looking to you to deliver accurate traffic projections for the coming four months. Your algorithm will become the foundation of a larger transformation to make your city smart and intelligent.

**Data Dictionary**

|  |  |
| --- | --- |
| **Variable** | **Description** |
| ID | Unique ID |
| DateTime | Hourly Datetime Variable |
| Junction | Junction Type |
| Vehicles | Number of Vehicles (Target) |

**sample\_submission.csv**

|  |  |
| --- | --- |
| **Column Name** | **Description** |
| ID | Unique ID |
| Vehicles | Number of Vehicles (Target) |

## ****Evaluation Metric****

The evaluation metric for this competition is Root Mean Squared Error (RMSE)

## ****Public and Private split****

Note that the test data is further randomly divided into Public (25%) and Private (75%) data. Your initial responses will be checked and scored on the Public data.

The final rankings would be based on your private score which will be published once the competition is over.

## ****Guidelines for Final Submission****

Please ensure that your final submission includes the following:

1. Solution file containing the predicted number of vehicles for next 4 months (format is given in sample submission csv)
2. Code file for reproducing the submission, note that it is mandatory to submit your code for a valid final submission