**Solution Sheet**

1. Which model have you used for Covid Cases prediction? Explain your model.

Sol: - After careful data Analysis using Tableau I found that there were only 5 parameters which were really affecting the ‘Covid Cases’.

These are as follows :-

1. Population [2011]
2. Toilets Avl
3. Female Population
4. # of hospitals
5. Foreign Visitors

I used these attributes to make an Ensemble model of Linear Regression and a Random Forest.

This gave me a very reasonable RMSE and MAE.

* If we had to minimize RMSE(Root Mean Square Error) , I would have choosen ‘Linear Regression model’.
* If we had to minimize MAE(Mean Absolute Error) , I would have choosen ‘Random Forest Regression model’.

Since we don’t know the Evaluation Criteria , I chose an Ensemble of both the models.

**Final Model:**

Ensemble model of Linear Regression and Random Forest.

1. Which model have you used for Foreign Visitors Time series prediction? Explain your model.

Sol :- I used a simple Random Forest to predict the Foreign Visitors(FV) in September and Used the model from the Part 1 to Predict the Covid Cases on 1 October.

Detail:

Initially, I trained a model for predicting the FV in August using the previous four months then I used that model to predict the FV in September using the previous four months ,i.e.,

**Input** **Output**

(April, May, June, July ) 🡪 August

( May, June, July, August ) 🡪 September

Now these FV values for September was replaced in Test\_dataset with FV in August and the Ensemble model from the Part 1 was used to predict Covid Cases on 1 October.