FINAL PROJECT

**Note:**

This would be a light introduction to a real world scenario where your job as a data scientist is to develop a good supervised regression model to predict dengue fever cases so Medics , Govt and NGOs can take precautionary steps and policies in advance.

You **ALL** would be competing against each other for the best model

If two competing parties have **same MAE score** they both get a **zero in the project**

**Sign up with your own FULL NAME or else you get a zero**

To be submitted individually no teams allowed

Incase presentation is not held due to shortage of time [Task 1 would be marked out of 10 marks]

**Task:**

Participate in the **DengAI: Predicting Disease Spread** challenge at Drivendata.org

Go to the site <https://www.drivendata.org/competitions/44/dengai-predicting-disease-spread/>

Sign Up - **with your own FULL NAME**

Then

Join the competition <https://www.drivendata.org/competitions/44/dengai-predicting-disease-spread/>

Following the rules of datadriven

1. Go through the whole DS process applying everything you know to the challenge and make a final predictive sample. (Make a submission with your own full name on Datadriven platform(predictions only) and code on your github account) – [6 marks]

2. Prepare a short ppt on how you approached the problem , based on which there would be a short oral Q/A session – [4 marks]

**Keep these in your ppt (Atleast and add other important findings /insights/ approaches you took )**

Why you did what you did and your final best score in the challenge on their test set with a screenshot of the leader board with your name