**#DataMadness Project Proposal Form**

|  |  |
| --- | --- |
| **i-number** | **Student name** |
| 6213546 | Tran, Truc |
| 6194619 | Ha, Viet |
| 6219876 | Hermans, Teun |

(Provisional) Title:

*Can change during the project, but make sure that you pick a title that attracts the audience*

Relationship between wearing mask and percepting languages

Questions and Objectives:

*List at least 3 questions (high-level) of what you want to address in your analysis. More are welcome, but think of the workload*

1. Does wearing masks more often facilitate masked speech perception?

2. Which variable(s) affect the most to masked speech perception?

3. Is there a significant difference between people who have positive and negative attitude towards covid on masked speech understanding?

The Dataset

*Attach the dataset or point to a web link that is available (if very big or already existing). Highlight why it can answer your questions.If the dataset does not come with datasheets, you are going to create them to best of your knowledge (see extra form)*

Click [here](https://drive.google.com/file/d/1Uw8B63gG4mRxPqCvO3dXpERQFqZzZ6yc/view?usp=sharing).

\*\*Rough\*\* Outline of the Analysis

*This is a rough sketch of the analysis you want to do on the dataset and which techniques you should use to answer your questions. Not detailed at the proposal, but should act as a guide as you work on the project.*

+ Clean, visualize and summarize statistics of dataset

+ Use Multi-variables Linear Regression model. Train and choose the best *k* variables using stepwise selection. (Q1+2)

+ Hypothesis testing for finding difference between people who wear and do not wear mask on masked speech understanding. (Q3)