COVID-19 Data Science Challenge August 2020

Guidelines for Reading your Dashboard for Phase 1

In your dashboard, we report a selected set of metrics related to your model's **prediction accuracy** related to sum_visit_counts of 304 unique uuids for week_18 as well as your **effort** in achieving this accuracy.

Your dashboard has two components: Your scores and Your peers.

- Your scores tell you about individual metrics.
- Your peers provide you some more insight into your current standing in the hack with 47 participants and 7 submissions in phase 1.

In both components, we present the following metrics:

Metric	Brief Description/Formula	Data type
`MSPE`	The Mean Squared Prediction Error for the `sum_visit_counts` for `week 18` is calculated as following: $MSPE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$; with n being the 304 unique `uuids` in week 18; if your cells where "empty" for certain `uuids` we assume that you submitted a `0`; the lower your MSPE, the better your model; please note that the MSPE is sensitive to outliers. The MSPE is commonly used in other Data Science Challenges; thus, in this challenge, we will be using the MSPE	FLOAT (rounded to two digits)
`MAPE`	The Mean Absolute Prediction Error for the `sum_visit_counts` for `week 18` is calculated as following: $MAPE = \frac{1}{n}\sum_{i=1}^{n}(\lvert Y_i - \hat{Y}_i \rvert)$; with n being the 304 unique `uuids` in week 18; if your cells where "empty" for certain `uuids` we assume that you submitted a `0`; the lower your MSPE, the better your model	FLOAT (rounded to two digits)
`Notebook Effort`	The `Notebook Effort` score describes your effort in using your notebook to build your model; it considers the number of cells as well as your activity in modifying and running them.	INTEGER
`BigQuery Effort`	The `BigQuery Effort` describes your effort in understanding the data we provided you; it considers the number and the nature of the jobs you launched via your notebook.	INTEGER

> *How to read your scores and your peers? Here some tips:*
>
> - *The lower your MSPE the better; however, you do not necessarily need to have a value close to 0 to win the hack. Make sure you also compare the mean and the standard deviation in the `your`*

> *`peers` tab;*
> - *Your standing relative to the mean tells you how you are performing compared to the mean! Also, note that a large distance between you and the mean does not mean that you cannot catch up (or the other way around); we will eventually rank-order all participants to select the Best Solutions.*
> - *Your Big Query Effort and your Notebook Effort count towards your overall effort during the hack. To allocate improvement spirit incentives we will consider both - your MSPE as well as your effort*
>
> *And a final note: Only the TWO best submissions will be considered for the final MSPE and MAPE calculation! And do not forget: According to the rules of the challenge, only 2 submissions are needed to qualify for an award.*