**Description**

**If you have any questions, please reach out to your ambassador Emmanuel Kouphoh -**[**eaedk**](https://zindi.africa/users/eaedk)**.**

Income inequality - when income is distributed in an uneven manner among a population - is a growing problem in developing nations across the world. With the rapid rise of AI and worker automation, this problem could continue to grow if steps are not taken to address the issue.

The objective of this challenge is to create a machine learning model to predict whether an individual earns above or below a certain amount.

This solution can potentially reduce the cost and improve the accuracy of monitoring key population indicators such as income level in between census years. This information will help policymakers to better manage and avoid income inequality globally.

**Evaluation**

The error metric for this competition is the F1 score, which ranges from 0 (total failure) to 1 (perfect score). Hence, the closer your score is to 1, the better your model.

**F1 Score:**A performance score that combines both precision and recall. It is a harmonic mean of these two variables. Formula is given as: 2\*Precision\*Recall/(Precision + Recall)

**Precision:**This is an indicator of the number of items correctly identified as positive out of total items identified as positive. Formula is given as: TP/(TP+FP)

**Recall / Sensitivity / True Positive Rate (TPR):**This is an indicator of the number of items correctly identified as positive out of total actual positives. Formula is given as: TP/(TP+FN)

Where:

TP=True Positive

FP=False Positive

TN=True Negative

FN=False Negative

The submission file is large and takes ~3 min to process, this time will increase when everyone makes a submission on the final day. Don't worry, as long as your file is submitted before the deadline it will be scored and taken into consideration.

This is an example of what your submission file should look like.

Where 1 indicates that the individual is above the limit and 0 indicates the individual is below the limit.

ID income\_above\_limit

ID\_TZ209502 0

ID\_TZ209511 1