**Data Description**

The data contains 80 variables that characterise the demographic and socio-economic situation of 181

galaxies over a period of at most 26 years. A composite index is given that measures their well-being.

However, the demographic and socio-economic variables that influence this index is not known. We

seek to determine, **what makes the galaxies better off**?

We would like you to use the data and:

1. Tell us which variables best explain the variance of the well-being index

2. Determine the future well-being values of the galaxies

**Submission Instructions and Format**

We have provided you data with observed values of the well-being index of each galaxy and a

validation dataset that requires the prediction of the future well-being index.

**Kindly submit:**

1. A report that discusses the demographic and socio-economic determinants of the galaxies'

wellbeing.

**Submission Format**: The report should be a **pdf of a slide presentation** of not more than 5 slides

2. The predicted future well-being index values with the highest possible level of certainty using

data in the validation dataset.

**Submission Format**: A csv file Saved as "firstname\_lastname\_DSA.csv" containing:

**Variable Description**

ID Unique identifier of the observations in the validation dataset

Predicted Well-Being Index Prediction for the Well-Being Index

3. Analysis file

Submission Format: python/R notebook with detailed comments and organised analysis of EDA with

visualisations, well documented analytical process and test results.