# Data Science Assignment 01

# 01 Page Summary

Group 09 (DS_2019_007)

MSc in Computer Science 2019
Department of Computer Science and Engineering
University of Moratuwa

# Summary

## Group Members and Responsibilities

**199329V, H.E.A.K.P. Hewawasam**
Developing Analytical Insights, Other Analytics, Executive Summary, Detailed Report

**199355V, W.L.D.W.P. Perera**
Developing Analytical Insights, Visualizations, Hypothesis Testing, Detailed Report

**199342E, C.G. Madage**
Developing Analytical Insights, Hypothesis Testing, Executive Summary, Detailed Report

**199361K, M.D.R.N. Senanayake**
Developing Analytical Insights, Sampling, bias & relevant matters, Detailed Report

## Github Repository

https://github.com/crishalm/dataScience-assignment1

## Hypothesis Testing

**People with high education level have high earnings than people with low education levels.**

Ph : Average earnings of people with education level higher than education median
Pl : Average earnings of people with education level lower than education median
H0 : Ph = Pl
Ha : Ph > Pl

## Assumptions

- Entries above education level 20 have minimum effect on overall statistics.
- Entries above number of kids 20 have minimum effect on overall statistics.
- Entries above earnings of USD 150,000 does not affect the overall statistics.
- Even the persnum is same, we have considered them as separate individuals

## References

[1]  https://psidonline.isr.umich.edu/Guide/2015DataHighlights.pdf
[2]  https://psidonline.isr.umich.edu/
[3]  https://psidonline.isr.umich.edu/data/Documentation/UserGuide2013.pdf
[4]  https://psidonline.isr.umich.edu/data/Documentation/UserGuide2015.pdf