ISSU0053 Data Science and Big Data Analytics **Report ID:** 1043364

UCL International Summer School for Undergraduates 2019

**Assessment I: Computer Practical Work and Write-up (50%)**

|  |  |  |
| --- | --- | --- |
| **Task** | **Mark** | Comments |
| T1: Explore the dataset | A+ |  |
| T2: Linear **/** logistic regression **/** LDA | B |  |
| T3: Decision Tree methods | B |  |
| **Overall task:** | A-B |  |
| Use of Figures | A-B |  |
| Introduction, Commentary, Summary | C |  |
| Coding Technique | B |  |
| Writing Standard | C |  |
| Structure & Presentation | C |  |
| **Overall:** | B  (60%) |  |

**Strengths:**

- very good EDA with good comments and appropriate figures and correct conclusions

**Focus on improvement:**

- you have clearly not understood the difference between regression and classification. You are trying to predict a continuous variable (number of bicycles). Why do you use logistic regression and KNN for that?

- there are no comments and no summary after EDA, which again leads to the conclusion that you only followed the instructions but did actually not know how to interpret the output

- after EDA report becomes hard to read due to all the warning messages (which should have alarmed you that something is going wrong by the way)