ISSU0053 Data Science and Big Data Analytics **Report ID:** **1043376**

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 | C |  |
| T3: Decision Tree methods | F | N/A |
| **Overall task:** | C-D |  |
| Use of Figures | B |  |
| Introduction, Commentary, Summary | C |  |
| Coding Technique | B |  |
| Writing Standard | B |  |
| Structure & Presentation | B |  |
| **Overall:** | C  (55%) |  |

**Strengths:**

**Focus on improvement:**

- you have no missing values (apart from the new price, which you delete) – why do you explore missing values? I don’t understand

- you clearly have not understood the difference between regression and classification. You want to predict a continuous variable: price. Why are you using logistic regression and KNN instead of Linear Regression and Decision Trees?

- No tree based methods implemented

- no summary of your findings/conclusion of your report