**CCT College Dublin**

**Assessment Cover Page**

*To be provided separately as a word doc for students to include with every submission*

|  |  |
| --- | --- |
| **Module Title:** | *Advanced Data Analytics*  *Big Data Storage and Processing* |
| **Assessment Title:** | *MSC\_DA\_BD\_ADAv5* |
| **Lecturer Name:** | *David McQuaid*  *Muhammad Iqbal* |
| **Student Full Name:** | Jose Luis Alvarado Carranza |
| **Student Number:** | 2020092 |
| **Assessment Due Date:** | 08/11/2023 |
| **Date of Submission:** | 08/11/2023 |

**Declaration**

|  |
| --- |
| By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution. |

ABSTRACT

INTRODUCTION

I have uploaded the Project Tweets .csv into both of the technologies used in the course. Vis a vie Hbase, SparkSQL. And I have made some comparisons in my head on the performance of each off them. I think that for the task is it more convenient to use Spark SQL than Hbase because it is visually easier to comprehend and one doesn not have to be switching between terminals to eeecute the commands.