**CCT College Dublin**

**Assessment Cover Page**

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

|  |  |
| --- | --- |
| **Module Title:** | **Programming for DA**  **Statistics for Data Analytics**  **Machine Learning for Data Analysis**  **Data Preparation & Visualisation** |
| **Assessment Title:** | **MSC\_DA\_CA1** |
| **Lecturer Name:** | **David McQuaid**  ***Sam Weiss/* *David Gonzalez***  ***Bharathi Chakravarthi/* *Marina Iantorno***  ***Muhammad Iqbal*** |
| **Student Full Name:** | **Syed Asad Ailia** |
| **Student Number:** | **2023408** |
| **Assessment Due Date:** | **12th November,2023** |
| **Date of Submission:** | **12th November,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**

*In today's data-driven world, data analytics is becoming more and more significant, having a major impact on many facets of business, science, and society. In this report, we used data analytics techniques to clean insightful information from the mean and median hourly earnings across various economic sectors and employment statuses in Ireland. We carefully imported and checked the information to respond to specific questions. The development of a machine learning model for extracting output parameters from the validation dataset was the final step in this extensive process, which began with data preparation and continued with graphical representation using statistical techniques to identify trends. The Python framework was used to fulfil the programming requirements, and the entire project was recorded in a Jupyter Notebook.*

**Introduction**

Data analytics is a vital tool in today's data-driven world, especially when examining hourly wages across a range of economic sectors and employment status. With this analytical method, data is systematically examined and interpreted to find trends, patterns, and insightful information. Data analytics provides organizations, policymakers, and researchers with a thorough understanding of the economic landscape when it comes to hourly earnings. It also provides data-driven insights into labour markets and income disparities. An important economic measure, hourly earnings, differ greatly between industries and job types, which reflects the complexity of the modern workforce. Data analytics makes it easier to explore this complex terrain by highlighting differences, pointing out causes, and facilitating data-driven decision-making. Our ability to analyze the data is further improved by statistical methods and machine learning models, which help us forecast future trends and the effects of policies in addition to describing the current state of affairs. When it comes to addressing economic disparities in a world where data-driven insights are essential, data analytics becomes a crucial ally that leads the way toward a more just and prosperous future. Here, we'll examine a few crucial data analytics pipelines to gain deeper understanding.