

**London College of Professional Studies**

***Assignment Cover Page***

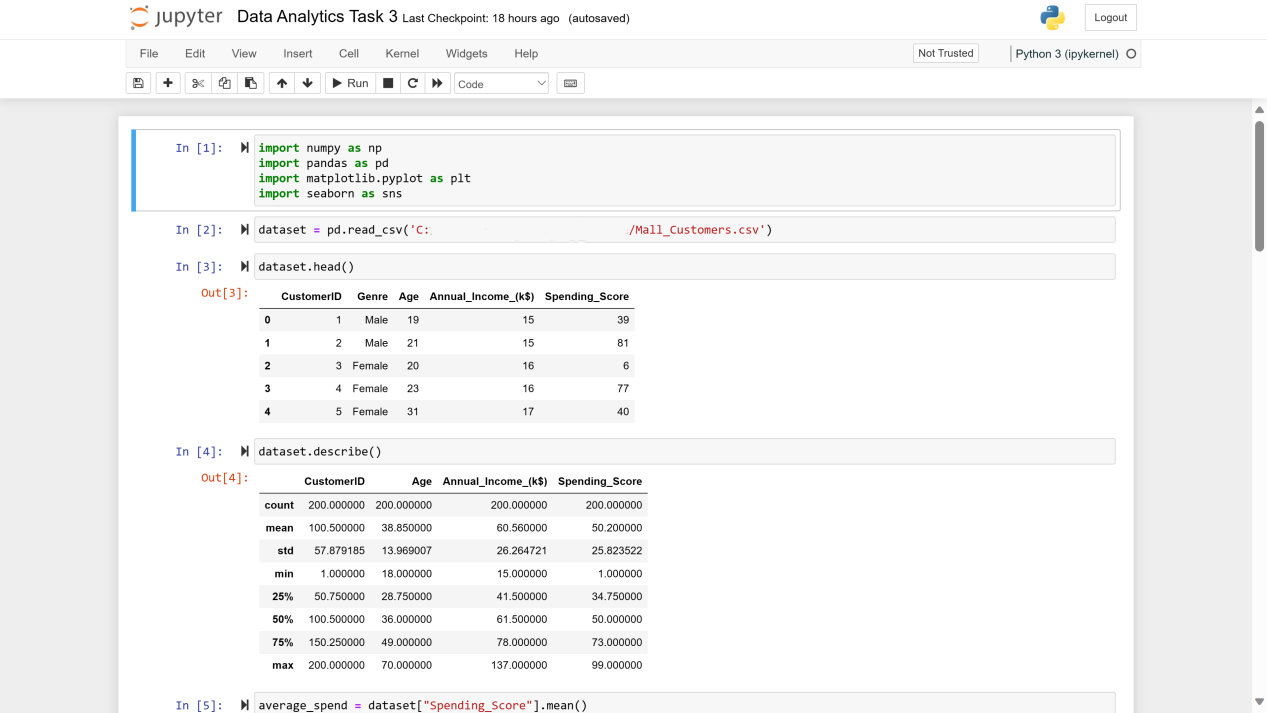
|  |  |
| --- | --- |
| **Student name** | **Christopher Moyo** |
| **Student ID** | **14034** |
| **Submission date** | **28/08/2023** |
| **Qualification title** | **OTHM Level 6 Diploma in Information and Technology** |
| **Qualification code** | **603/3789/8** |
| **Unit title** | **Advanced Data Analytics** |
| **Unit Reference Number** | **Y/617/3035** |

 I declare that the attached work is entirely my own and that all sources have been acknowledged ☒

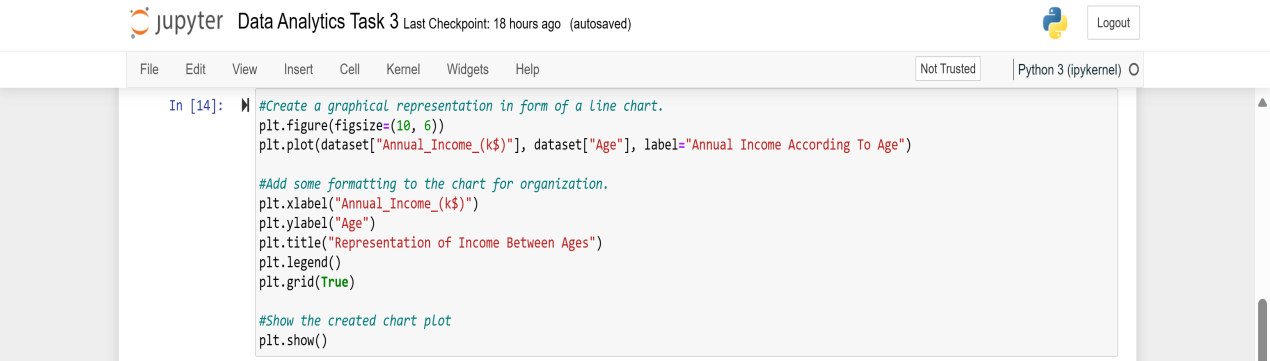
Our first predictive analysis was performed on a Small and medium-sized enterprise (SME). The focus of this analysis was towards the clients that the enterprise received in relation to their age, annual income, and spending score whenever they visited the enterprise premises. The tools used were Jupyter Notebook and Python programming language. Below is the demonstration.

The used dataset is borrowed from Shruti Iyyer on the Kaggle data analytics platform.

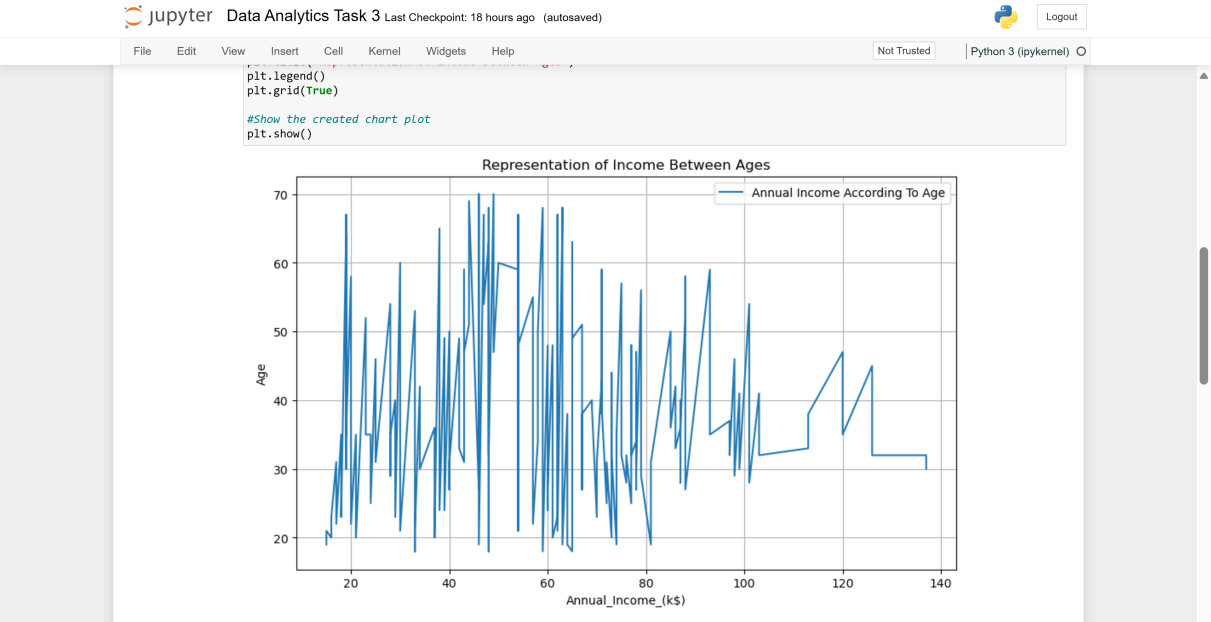
1. Adding required libraries and getting the relevant dataset file read into the notebook for analysing.



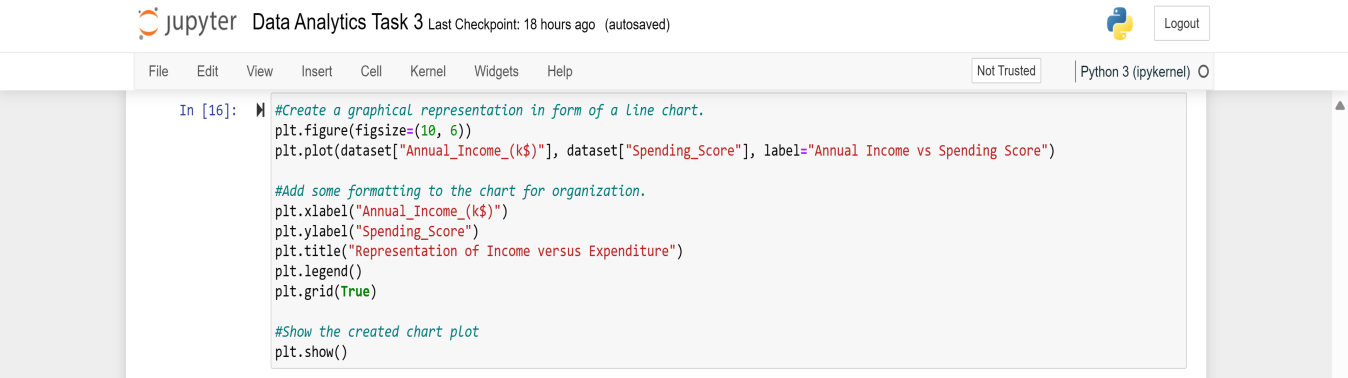
* We begin by importing the relevant libraries we are to use for analysis in line 1, and then prompt the pandas library to read our dataset in line 2.
* Lines 3 and 4 assist us in knowing the first 5 members of the dataset and the description of the dataset we are using.
* Line 5 is created in order to create an average on the spending score of each client as we analyse the collected data.

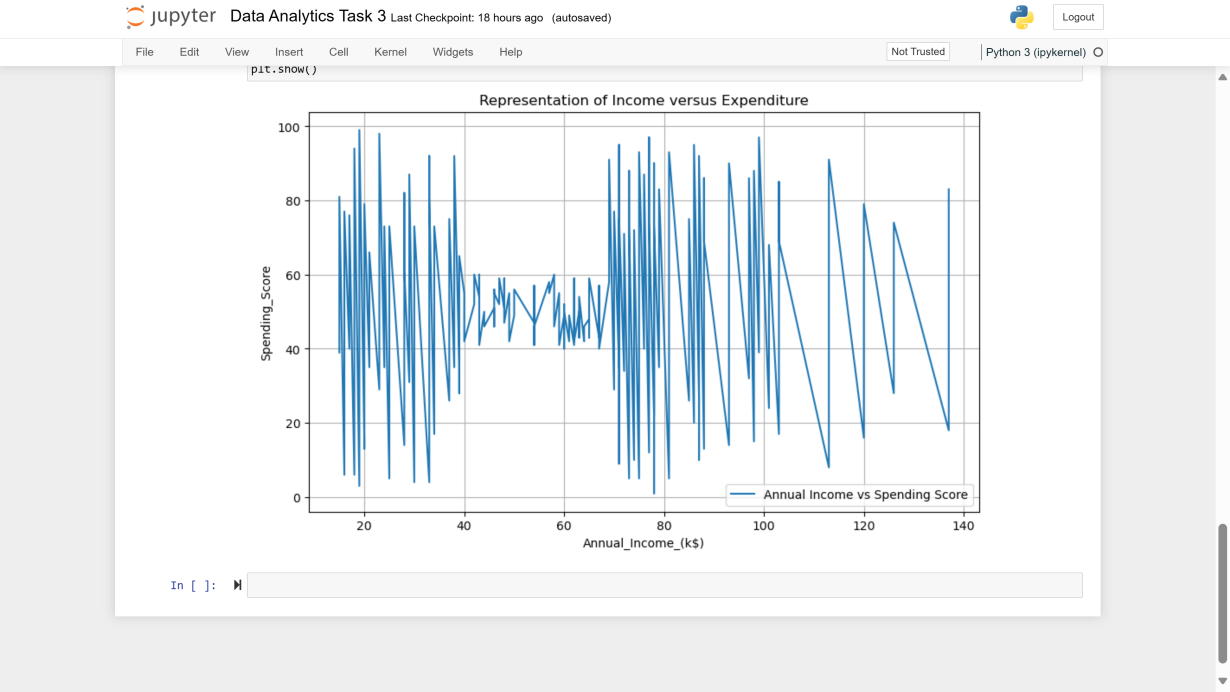
1. We then follow up by showing the analysis in form of line graphs and bar-charts as shown below.

* Cell 14 is used to create a graphical representation using a line chart of the Annual income of each client in relation to their age in the enterprise. The result is the screenshot below:



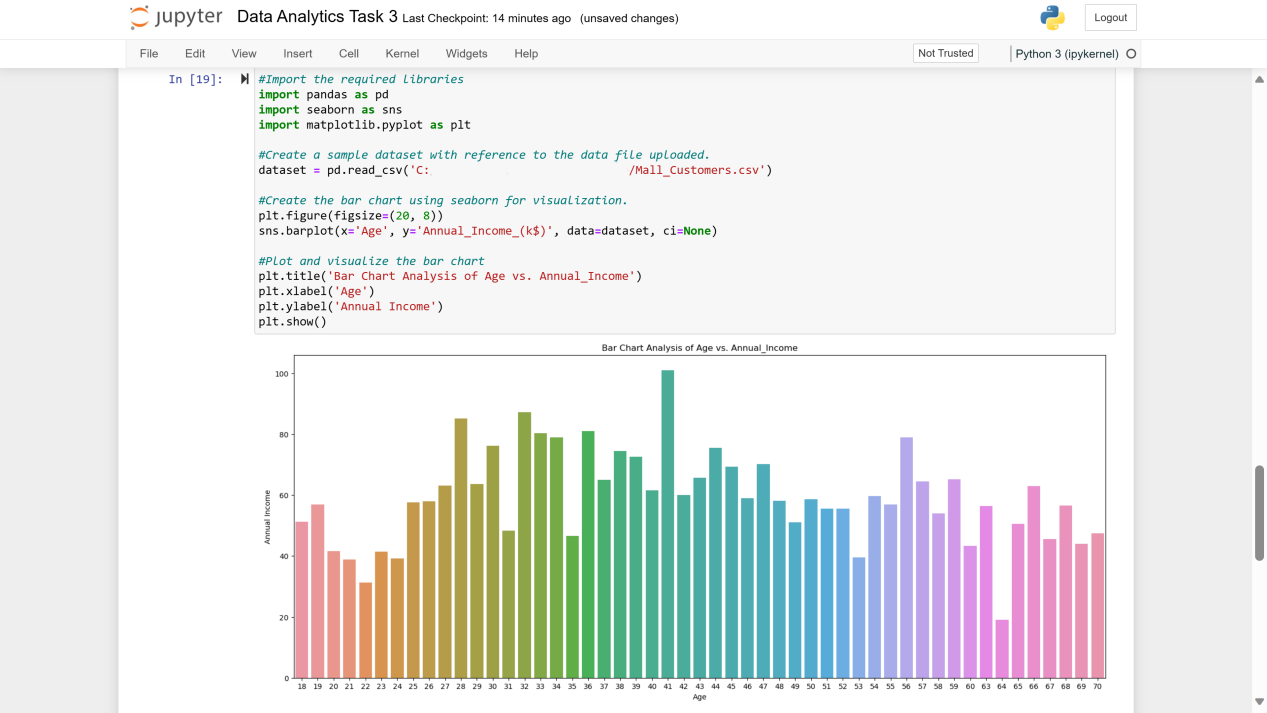
(Age vs. Annual Income)



* We can also do the same for the representation of the annual income versus the spending score of the clients in the enterprise. The result is the screenshot below:

However, using line charts is not always the best way to represent data that is in large groups, because the result is not always read as accurate as it should be. To fix this, we used bar charts to clearly visualize the analysis and predict the turnout of the age, annual income, and spending score of the clients in the enterprise. This is demonstrated below.

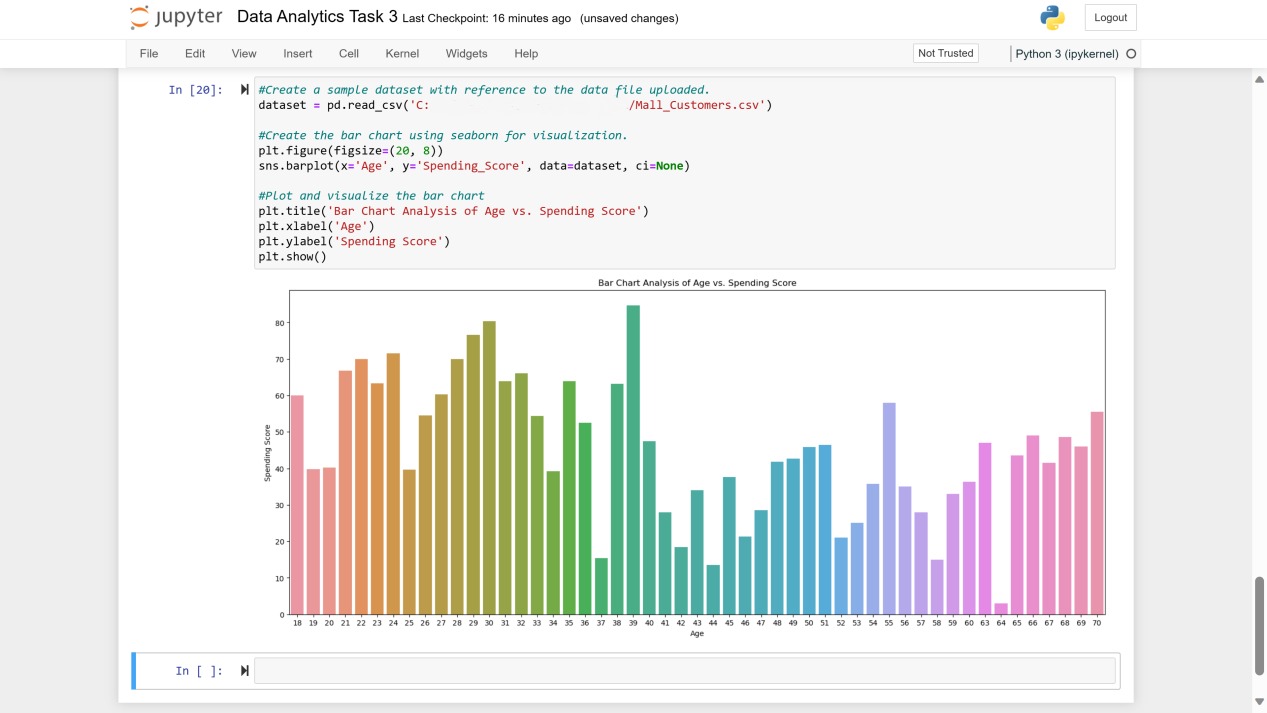
1. **Age vs. Annual Income.**



In the cell we:

* First import the required libraries.
* Then create a sample dataset with reference to the dataset we read from the csv file.
* And create the bar chart using the seaborn library.
* Finally plot and visualize the bar chart.

1. **Age vs. Spending Score.**



In this cell we follow every step from the previous representation, except importing the required libraries as they are already active.

**References.**

1. Iyyer, S. (2019). *Customer data*. Available at: <https://www.kaggle.com/datasets/shrutimechlearn/customer-data?resource=download.> [Accessed 26 August 2023].