**Data Scientist Technical Assessment**

**Overview**

Use the accompanying dataset ‘sales.csv’ to build an unsupervised learning segmentation analysis in order to effectively group the products being presented in the data. The goal is to provide reasonably discernable clusters for business to evaluate and formulate strategies around.

**Deliverables**

* Presentation of exploratory data analysis
* Data clean-up (as necessary)
* Feature reduction/addition/manipulation and rationale behind doing so
* Clustering algorithm of choice (please use from existing libraries)
* Output presentation and summaries for business

**Please use Python 3.8 or above and send back a Jupyter notebook containing commented experimental/visual work and any accompanying files**

Your Jupyter notebook must include the following:

* Comments explaining what you have done and why
* Visualization of your data/results and comments regarding the visualizations
* The version number of each Python package you use (e.g., print(sklearn.\_\_version)) and an accompanying requirements.txt file or Anaconda environment.yml file so that we can avoid version issues
* The seeds used to generate random numbers used for any purpose (if necessary)

**Data Dictionary:**

This dataset contains information about Sales Values in Dollars on American Stores between 2010 and 2011. In particular, this dataset contains:

* Stores' Area, State, Region and Size;
* Products' ID, Description, Type, Category and Sale Date;
* Accounting Info, such as Budget Margin, Profit, Total Expenses and Marking.

The file sales.csv contains a huge table of US Stores Sales info in the following columns:

* Area Code: Store's Code;
* State: Store's State;
* Market: Store's Region;
* Market Size: Store's Size;
* Profit: Profits in Dollars ($);
* Margin: Profit + Total Expenses ($) OR Sales - COGS ($);
* Sales: Values Acquired in Sales ($);
* COGS: Cost of Goods Sold ($);
* Total Expenses: Total Expenses to get the Product to Sell ($);
* Marketing: Expenses in Marketing ($);
* Inventory: Inventory Value of the Product in the Sale Moment ($);
* Budget Profit: Expected Profit ($);
* Budget COGS: Expected COGS ($);
* Budget Margin: Expected Profit + Expected Total Expenses ($) OR Expected Sales - Expected COGS ($);
* Budget Sales: Expected Value Acquired in Sales ($);
* ProductID: Product ID;
* Date: Sale Date;
* Product Type: Product Category;
* Product: Product Description;
* Type: Type;