**Data Science Take-Home Exercise**

Sept. 2022

**Analysis & Modeling**

**Setup**

A client is looking for help to improve their marketing, and they have provided historical response data for us to work with. This data can be found in the attached file named DataScience\_exercise\_data.csv.

The data represents 4,000 individuals who were recently marketed to along with a flag indicating whether they subsequently responded (target = 1) or not (target = 0). The data also includes a unique identifier for each record along with several demographic attributes.

Please use Python for tasks 1-3. In addition to the output associated with each part, please provide a copy of the code used in this exercise. A Jupyter Notebook often works well, but you may provide your answers and code in any format you wish.

**Part 1: Hypothesis Testing**

Use a Hypothesis Test to determine if responders and non-responders tend to have the same age.

Output: A written summary of the steps of your hypothesis test and the result of your test.

**Part 2: Predictive Modeling**

Build a Binary Classifier to help this client better predict responders in future campaigns.

Output: A brief writeup of your approach, including the algorithm you used, any data analysis or preparation you pursued, and the performance metrics you used to evaluate your final model.

**Part 3: Data Visualization**

Create a visualization of the Cumulative Gains Chart for your model. If you are unfamiliar, see here for more information: <https://www.ibm.com/docs/en/spss-statistics/24.0.0?topic=overtraining-cumulative-gains-lift-charts>

Output: The image of your Cumulative Gains Chart.

**Part 4: Brief Writeup**

Provide an executive summary of your findings from tasks 1-3. Treat this as if you were sending an email to a teammate or your supervisor bringing them up to speed on your analysis and findings.

Output: A written executive summary.

**Part 5: Secondary Visualization (Optional)**

Create a second visualization of your choice to help illustrate your model or the results it can yield to the client.

Output: The image of your visualization/chart.