**Phase 3: Data Preparation**

Gabriel Eduardo Melendez Zavala – A01638293

Mathematics and Data Science for Decision Making (Gpo 602)

In my case there is no need to merge the data because there is no other database to merge columns or data entries with the original data. Since the only client in this project is me, there are no other data entries or parameters required in this project.

In the case of subsets, I didn’t use them to discriminate a certain data group, however they can be applied for training, testing, and validation. They are used to facilitate efficient analysis, focus on relevant information, reduce computational complexity, and address specific research or analysis objectives.

Regarding adding more data, we can add more data and it can affect the analysis. However, it is not always necessary or feasible. We must consider the specific context, goals, and available resources to determine whether adding more data is appropriate for the analysis.

The data attributes for the analysis are correct and none need to be modified for a specific purpose. Each attribute is used as it was intended, and no changes were required.

In addition, the data does not need to be sorted in order to perform the analysis, no matter the order the analysis will continue to be relevant and output meaningful results.

The data was complete therefore no blank values needed to be removed, replaced or dealt with in an appropriate manner.