MSBA Capstone Project

**Bi-Weekly Progress Report**

**Student Name:** Mona Ghorayeb **AUB ID:** 202152440

**Project Title:** Optimizing Growth: Sales and Market Expansion for GIFCO

**Elevator Pitch**

This project aims to optimize GIFCO’s sales and market expansion strategies by leveraging historical sales and shipment data. The focus is on creating a dashboard, forecasting sales trends, and identifying potential expansion opportunities. The project has progressed through data cleaning, feature engineering, and exploratory data analysis to prepare for predictive modeling and optimization.

**Tasks Done During the Previous Two Weeks**

First, I have extracted and manually entered some data that were given on paper:

* For the year 2016, the data was manual on paper which required manual entry on excel.

Then, I completed the data cleaning and preprocessing step where:

* Missing values were handled using mode for categorical and the median for numerical data.
* Replaced client names with anonymous client IDs for confidentiality.
* Standardized date formats and extracted Year and Month for time-based analysis.
* Removed unnecessary columns.

Then, feature engineering was performed:

* Created Profit Margin and Average Price per Unit in the sales dataset.
* Generated shipment route mapping.
* Derived Weight per Package and assigned seasons based on shipment dates.

Finally, EDA using Python was done as a first step to prepare for tableau dashboard later:

* Analyzed revenue trends, top product families, and sales growth.
* Visualized shipment volume by freight type, seasonality, and top client shipments.

**Difficulties and Challenges Encountered**

* The original dataset contained inconsistent values, missing data, and duplicate categories, requiring extensive preprocessing.
* Determining which columns were unnecessary and how to handle them without losing valuable information was a big challenge. Some columns contained redundant data, while others required transformation for better usability.
* Another challenge was deciding the best approach to structuring the data for effective analysis. This involved making decisions on categorical vs. numerical encoding, normalization, and aggregating information for better interpretability in EDA.

**Significant Changes in Proposal**

In the initial proposal, optimization of the internal operations were considered but after receiving the data, I realized that there were some limitations to this and I stuck to the sales and market expansion data analysis and optimization.

**Tasks To Be Completed**

* Continue and complete the Tableau dashboard of the sales and shipment data.
* Predictive modeling and forecasting by implementing the time-series forecasting (ARIMA, Exponential Smoothing) for sales prediction.
* Explore clustering techniques to identify potential market segments.