



**UGANDA CHRISTIAN UNIVERSITY**

**Faculty of Computing and Technology**

**Object Oriented Programming**

**Master of Data Science & Analytics**

**John Tebasiima**

**S25M19/024**

**B35101**

**12<sup>th</sup> October 2025**

# **Sales Performance and Customer Insights Using Python**

## **Problem Statement**

In the competitive modern business world, organizations are becoming more dependent on the use of data-driven decision-making to enhance performance, profitability, and customer satisfaction. However, most companies in developing countries find it challenging to extract insights from their data beyond reporting performance. Organization raw sales data is stored in spread sheets or transaction systems without systematic analysis by which management can identify performance trends, customer behavior or demand in real time

The proposed project seeks to address this problem by developing a Python-based data analytics solution that extracts actionable insights from sales data and creates self-refreshing reports for seamless reporting. The project will focus on understanding sales performance across products, regions, and time periods, while also analyzing customer purchasing patterns to inform marketing and production strategies.

## **Project Relevance**

Modern business intelligence cannot do without sales analytics. Data analytics directly improve business results by defining the most sold products, regions that are not performing well, and the most valuable customers which influence key business decisions on inventory management, sales forecasting, targeted marketing among others tactical decisions.

Leveraging Python, the project will illustrate how analytics can be used to transform business strategy decisions. The suggested work is therefore applicable not just to the scholarly community but also the practical community in the industries that are aiming to enhance operational efficiency and customer awareness.

## **Project Scope**

The project will focus on analyzing a dataset containing sales transactions, including product names, customer IDs, sales quantities, prices, regions, and dates. The specific objectives are:

- To clean and preprocess the dataset using Python to handle missing, inconsistent, or duplicate records.
- To perform descriptive analytics to determine total sales, average revenue per product, and regional performance.
- To conduct customer analysis to identify key customer segments based on purchasing frequency and value.
- To develop visualizations that highlight trends in product performance, seasonal sales patterns, and customer behavior.
- To generate actionable insights to support managerial decision-making in sales planning and marketing strategies.

**Methodology**

The analysis will be conducted using the Python programming language and its associated libraries due to its versatility and extensive data science ecosystem. The workflow will involve the following key steps:

- Data Import and Cleaning: Using pandas and numpy to load and structure the dataset, handle missing values, and prepare it for analysis.
- Exploratory Data Analysis (EDA): Employing matplotlib, seaborn, and plotly to visualize trends, correlations, and anomalies.
- Customer Segmentation: Applying unsupervised learning techniques (e.g., clustering with scikit-learn) to group customers based on behavior.
- Insight Reporting: Summarizing findings through visual dashboards using streamlit or dash, allowing interactive exploration of the data.

The working data set will be adopted from Kaggle. The project will be developed in a Jupyter Notebook environment and hosted on GitHub for transparency, version control, and reproducibility.

**Tools to deploy**

Category	Tools/Libraries
Programming Language	Python 3.8
Data Analysis	pandas, numpy
Visualization	matplotlib, seaborn, plotly
Machine Learning (optional)	scikit-learn
Dashboarding	streamlit or dash
Environment	Jupyter Notebook, GitHub

**Expected Outcomes**

Upon completion, the project will produce

- A clean and structured sales dataset ready for analysis.
- Visual insights on sales performance by product, region, and customer segment.
- A summary report or dashboard highlighting key business trends.

These deliverables will collectively demonstrate how Python-based data analytics can inform business decisions and improve organizational efficiency.