# **Homework 1 - Business Intelligence**

**Group S**

* Ron Itzhaki
* Ron Shavit
* Shaked Misrael
* Asaf Hershko 312526841

## **Part 1 – Dataset**

The "Online Retail Transaction" dataset contains information on customer transactions through an online retail platform. The dataset includes data on the products purchased, the quantity of each product, the date and time of each transaction, the price of each product, the unique identifier for each customer who made a purchase, and the country where each customer is located. This dataset can be used to analyze customer behavior and preferences, identify popular products, and optimize pricing and marketing strategies. The dataset is well-suited for data analysis and machine learning applications. It contains a large volume of transactional data that can be used to train predictive models and make data-driven decisions.

[Dataset link](https://www.kaggle.com/datasets/abhishekrp1517/online-retail-transactions-dataset?select=Online+Retail.csv)

The data set contains one table, so we normalized it into two different tables:

**First table – product\_data**

This table contains the data of the products sold.

**Table #1 columns are:**

* StockCode: A code used to identify the product that was purchased
* Description: A brief description of the product that was purchased
* UnitPrice: The price of one unit of the product that was purchased

**Second table – sales\_data**

This table contains the sales data and transactions made by customers.

**Table #2 columns are:**

* StockCode: A code used to identify the product that was purchased
* Quantity: The quantity of the product that was purchased
* InvoiceDate: The date and time that the purchase was made
* CustomerID: The unique identifier for the customer who made the purchase
* Country: The country where the customer who made the asset is located

## **Part 2 – Business Questions**

**Question 1** - What is the most popular product category among customers, and how can the company optimize pricing strategies for those products?

KPIs for the first business question:

1. **Sales revenue generated by each product category over time**

This KPI meets the SMART criteria because it is specific (product category), measurable (sales revenue), achievable (can be tracked over time), relevant (to pricing strategy optimization), and time-bound (over time).

1. **Average unit price and quantity sold per product category**

This KPI meets the SMART criteria because it is specific (product category), measurable (unit price and quantity sold), achievable (can be tracked over time), relevant (to pricing strategy optimization), and time-bound (over time).

1. **Conversion rate for each product category**

This KPI meets the SMART criteria because it is specific (product category), measurable (conversion rate), achievable (can be tracked over time), relevant (to pricing strategy optimization), and time-bound (over time).

**Question 2** - How does customer location affect their purchasing behavior, and what can the company do to increase sales in regions with lower customer engagement?

KPIs for the second business question:

1. **Sales revenue generated by each region**

This KPI meets the SMART criteria because it is specific (region), measurable (sales revenue), achievable (can be tracked over time), relevant (to increasing sales in regions with lower customer engagement), and time-bound (over time).

1. **Customer acquisition rate by region**

This KPI meets the SMART criteria because it is specific (region), measurable (customer acquisition rate), achievable (can be tracked over time), relevant (to increasing sales in regions with lower customer engagement), and time-bound (over time).

1. **Customer retention rate by region**

This KPI meets the SMART criteria because it is specific (region), measurable (customer retention rate), achievable (can be tracked over time), relevant (to increasing sales in regions with lower customer engagement), and time-bound (over time).