Data Management Plan

for

Product recommendation selling system

Version 1.3 Draft

Prepared by: Bravo Group_ Domaine expert team

05-10-2021

Project participants and affiliations:

- Retail store and Supermarket.
- Instructional Technology Center-Birzeit University.

1. Description of project aims and purpose:

This project aims to solve the problem of managing offers and suggest a proper discount based on customer behavior.

The scope of the project is developing an approach to a recommendation system based on customer purchases behavior.

The data set will be used to generate related items with each other according to the customer's behavior. This will be useful for managing commercial offers, increasing sales volume, enhancing the competitive position, increasing the market share, and Increasing customer loyalty and satisfaction.

2. Information about data:

The data set will be used for the project can be summarize as follows:

Our client	
Client_Id	Integer
Name	String
Location	String
Branch_Id	Integer
Branch_Name	String
Branch_Location	String

Customers	
Customer_Id	Integer
Gender	String
Age	Integer
Address	String
Income	Integer
Status	String

Products	
Product_Id	Integer
Product_Name	String
Branch_Id	Integer
Quantity	Integer
Cost_Price	Decima
Sales_Price	Decima
Category_ID	Integer
Category_Name	String

Transaction	
Transaction_Id	Integer
Transaction_DateTime	DateTime
Product_Id	Integer
Product_Name	String
Quantity	Integer
Sales_Price	Decimal
Discount_Value	Decimal
Vat_Value	Decimal
Net_Transaction	Decimal
Branch_Id	Integer
Branch_Name	String
Branch_Location	String
Category_ID	Integer
Category_Name	String

1. **Collecting data:**

   The source of data will be the reliable set of transactions data from Retail store/ Supermarket's point of sales or extracting from API as MS access file (.accdb).

2. **Understanding collected Dataset:**

   - **Our client:** our client represents the Retail Sales Stores and supermarkets such as:

Our client	
Client_Id	12345
Name	Bravo
Location	Ramallah
Branch_Id	12345007
Branch_Name	Blaza
Branch_Location	Nablus Street

   - **Products:** all the goods are available for sale; this data file contains the items that the customers desire to purchase for example:

Products	
Product_Id	789
Product_Name	Milk
Branch_Id	123450007
Quantity	1000
Cost_Price	10.02
Sales_Price	15
Category_ID	60
Category_Name	dairy

   - **Customers:** the target case of study, they are the power driver of competition and sustainability of our client, so they will be studied carefully according to their shopping behavior:

Customers	
Customer_Id	98756
Gender	Female
Age	26
Address	Ramalla
Income	7000
Status	married

- **Transactions:** this table represents the transactions that the customer made, this file will be reflecting the customer behavior over the time period, upon that behavior, the recommendation system will propose the probability of items related to each other, which mean if the customer purchased the product "A" what is the sold probability of product "B" with "A":

Transaction	
Transaction_Id	55555
Transaction_DateTime	05/10/2021: Tues.: 13:35
Product_Id	789
Product_Name	Milk
Quantity	4
Sales_Price	15
Discount_Value	2.01
Vat_Value	2.55
Net_Transaction	10.45
Branch_Id	12345007
Branch_Name	Blaza
Branch_Location	Nablus Street
Category_ID	60
Category_Name	dairy

3. **Preparing data for processing:** in this stage we refine the data by dropping the null values or filling the omitted data in according to a specific way, slicing the dataset, replacing the data type format for some variables…etc.

3. **Metadata format & content:**

This section is linked with an excel sheet click here to show the detailed component.

4. **Policies for access, sharing & reuse:**

1- In our system, we are dealing with sensitive data and these data are dedicated only for the retail store which will use the system, so we have to put a high restriction on access and use the data.

2- For the system customers they will have the following authorities:

3- Connect their point of sales API with our system for data integration. Import the databases for fixed period (Month, Quarter …).

4- Use the basic features to get a result from the system model like products relations, sales volume ... etc.

5- Export the result from the system and save it on any selected folder

6- Order upgrade on the database importing period and any additional features.

7- Delete their databases and the generated results from the local and cloud server

For the system Admin they will have the following authorities:

1- Monitor all customer actions during the usage of our system

2- View the system code for model tuning, auditing, features enhancements and software upgrade in general

3- Integrate the system APIs, cloud, and local servers to save the data for short and long-terms.

5. Ethics:

The databases and the generated reports are sensitive data and its customers owned

No one has the authority to use and share this data, so we have to sign an agreement between the system owner and the customer to preserve his right and go to court in case the data was leaked

6. Data storage:

We will store the data on the password protected cloud/CITRIX server and it will also be backed up on a local password protected RAID/network drive.

The data will be stored safely and securely and in full compliance with European Union data protection laws.

The imported datasets from the integrated API, the generated Metadata and results will be saved on the local sever and then we will upload it in daily basis to the cloud sever for protection and reuse.

7. Budget:

The costs are included in the “Estimating initial start-up costs table” in Vision and scope Bravo recommendation system, “Plant and Equipment” section, page 8.