Name: Kenny Mar N. Gutierrez Subject: MIS612

Date:08/31/2022

**OVERVIEW**

The integration of data sets in this documentation that can be valuable in the operational component of any organization where decision makers can oversee the probable outcome and impact of its features and operations. Data can be turned into usable information, where consolidated information will generate solutions.

Manual data interpretation will take time and will be limited. For example, a hundred pieces of data cannot be comprehended at a single glance. Thus, data science has made a significant contribution to technology by facilitating straightforward interpretation.

For example, this analysis is focused on App Usage on as basis and is designed to leverage this Clothes Retailer's Database as the datasets to be interpreted.

# About This Data

This is a list of 500 customers dataset includes Email, Address, Avatar, Avg. Session Length, Time on App, Time on Website, Length of Membership, Yearly Amount Spent.

**OBJECTIVES**

A clothes retailer that offers both in-person style consultations and online clothing sales. Customers visit the store, meet with a personal stylist during sessions or meetings, and then leave the store to get the garments they want online or through a mobile app.

The business is debating whether to concentrate its efforts on the mobile app experience or its website.

* Average length of a session for in-store style consultations.
* App usage duration, measured in minutes.
* Time on Website: The typical number of minutes spent on a website.
* Membership Duration: How long has the customer been a member?
* Annual Amount Spent: How much money did the consumer spend during the year?

**PROCESS AND METHOD**

The following procedure and methods were used:

1. Importing the required libraries
2. Reading the dataset
3. Exploratory Data Analysis (EDA)
4. Visualizing Data using Seaborn library
5. Data Modeling
   1. Separating the data into features and target variable.
   2. Splitting the data into training and test sets.
   3. Implementing the Linear Regression Model.
   4. Evaluating the prediction scores.
   5. Visualizing the results
6. Evaluating the model accuracy

* **Numpy**: offers comprehensive mathematical functions, random number generators, linear algebra routines, Fourier transforms, and moret
* **Panda**: is used to analyze data
* **Seaborn**: data visualization library based on matplotlib
* **Matplotlib**: is a comprehensive library for creating static, animated, and interactive visualizations

**RESULTS AND DISCUSSION**

**Chart

Description automatically generated with medium confidence**

**Figure 1.**

This is an excellent approach to depict data since it highlights the relationship between variables. In this instance, Avg. Session Length, Time on App, Time on Website, Length of Membership, Yearly Amount Spent.

**Scatter chart

Description automatically generated**

**Figure 2.**

The measures and classes in this dataset make it an ideal choice for machine learning since they make it easy to distinguish between different classes. The grid below displays a map of the data. There are 5 measurements, hence a 5x5 plot is produced.

**Chart, scatter chart

Description automatically generated**

**Figure 3.**