**Bhakti Patil, Shashank Rewankar, Prithvi Bhatt, Arnav Laturkar**

**Project - Phase I: Planning**

**IFT 533- Data Visualization and Reporting for IT**

**Prof. Asmaa Elbadrawy**

**November 17, 2023**

**Section I: Description of the Dataset:**

**Online Automotive Sales Statistics'23 (Volkswagen) :**

There are a variety of publicly available datasets that provide statistics on Automotive Sales (on Kaggle and related websites), however there aren't many Volkswagen-specific datasets available online. The data for this dataset came from Turkey's leading online real estate and vehicle sales marketplace. The data solely applies to Volkswagen automobiles and covers the three months from January 1 to March 31, 2023.

This dataset contains data scraped from 7216 customers in order to examine the automotive market. A row in the dataset correlates to other columns for various car types, costs, color, and locations. Numerous null entries are present in the Date and Fuel columns and can be pre-processed before real visualizations are made. The dataset provides useful information about Volkswagen automotive sales and advertising tendencies in Turkey during the first quarter of 2023. The data can be utilized to identify patterns and trends in customer behavior, such as the most popular models, transmission styles, and fuel kinds. The data might also be used to discover which cities have the highest demand for Volkswagen vehicles and to evaluate the effectiveness of advertising campaigns.

The Online Automotive Sales Statistics'23 (Volkswagen) dataset includes 7216 records (rows) and 13 attributes (columns).

Attributes with what they mean, their Data types and Domain values:

1. Customer ID - Ordinal, Domain - Integer [0 - 7216]

A unique identifying number assigned to the person who is advertising.

1. Advertisement number - Categorical, Domain - Integer [195460036 – 777835401]

Unique identification for the AD

1. Brand - Categorical, Domain - String

Brand name of the car.

1. Model - Categorical, Domain - String

Model name of the car.

1. Variant - Categorical, Domain - String

Version of the car.

1. Year - Interval, Domain - Date type

Year when the car was developed and ready for use.

1. Kilometer - Categorical, Domain - Integer

Gives information about the distance the car has traveled.

1. Color - Categorical, Domain - String

Specifies the car color.

1. Transmission - Categorical, Domain - String [Automatic, Manual]

Takes two values and provides the transmission type of the car.

1. Fuel - Categorical, Domain - String [Diesel, Gas, Gasoline]

Specifies the type of fuel the car uses to run on.

1. City - Categorical, Domain - String

Specifies the city where the car is advertised.

1. AD Date - Interval, Domain - Date Type

Date information of the ad release.

1. Price - Ordinal, Domain - Integer

Price of the car.

Based on the above information regarding this dataset, the prospective visualizations can be as follows:

1. Many apps for the business group.
2. The quantity of automobiles in each class
3. The number of kilometers depending on the fuel type of the car.
4. Sales as a result of the advertising.
5. Ten most popular car models in a given segment.
6. Ten most expensive cars in the dataset.

**Section II: Prospective Dashboard Users**

Following is the list of prospective users for our planned dashboard:

* **Sales Managers:** They can use the dashboard to monitor the performance of sales, segment customers, predict and set sales targets, keep track of inventory levels and also understand geographical effects on the sales of cars to make decisions accordingly.
* **Marketing Managers:** They can use the dashboard to swiftly assess the overall situation of the marketing efforts, evaluate the effect of the marketing campaigns, optimize lead generation strategies and understand the amount of budget allocation requirements.
* **Operations Managers:** They can use the dashboard to monitor the various inventory levels, identify popular vehicles, anticipate demands of various vehicles, streamline the process of order fulfillment and optimize the entire supply chain process by identifying bottlenecks.
* **Business Owners:** Business Owners can monitor the key performance indicators (KPIs) such as total sales, revenue and profit margins, financial metrics such as expenses and profits, forecast sales trends, monitor customer feedback to improve upon, efficiency of the overall business and make key strategic decisions to expand the business.
* **Data Analysts:** The analysis team working for the vehicle company can use the dashboard to identify specific patterns or hidden relationships in the data. They can also further segment the customer base based on various attributes and their combinations and understand their behaviors. They can also further refine the dashboard and its visualizations for better clarity and even create new visualizations.

**Section III: List of User Requirements & Potential Questions**

**User Requirements:**

* Data Analysts Requirements: By analyzing the sales trends of Volkswagen cars in Turkey during the first quarter of 2023, analysts in the automotive industry may better understand the market demand for different car models and variants.
* Data Scientists Requirements: Using the dataset, data scientists may create prediction models that can be used to forecast demand for different Volkswagen car models and variants based on a variety of variables, including price, year, kilometers driven, city, and transmission type.
* Marketing Decision Making: Marketing and advertising teams can use this dataset to evaluate the effectiveness of their campaigns to promote Volkswagen cars in different cities and to modify their ads as needed to reach a wider audience.
* Car Production Decision Making: The dataset can assist automakers in better understanding consumer preferences and needs, which will enable the development of new Volkswagen car models and variants that better satisfy those needs.
* Rental Companies Analysis: Car rental companies can use the dataset to ascertain the demand for different Volkswagen car models and variants in different cities. Based on this information, the companies can modify the size of their fleets and their prices.
* Financial Analysts: Financial analysts can use the dataset to project Volkswagen car sales trends in Turkey going forward and to calculate the revenue earned by car sales.
* Learning purpose and Classwork: The dataset can be used by academics and students to explore various statistical and machine learning techniques for data analysis and to obtain hands-on experience in real-world applications.

**Potential Questions:**

* Which are the most popular car models? Which car models are the most popular in each city?
* Which car colour is the most popular among customers?
* Which year of production of cars is the most popular among customers?
* What is the average distance travelled by cars that use different fuel types?
* What are the most popular variants of the different car models?
* What are the price differences between different variants of the same car models?
* How does the average price of gasoline cars compare to diesel and hybrid cars?
* What is the most popular transmission type in cars purchased? What are the counts of fuel used by the car by different transmission types?
* Which cities/regions have the most customers that purchase cars? What are the most popular car models there? Which cities have the fewest customers?
* Which month of advertisement dates had the highest effect on the sales of cars?
* Which advertisement number brought in the highest number of car sales?
* What are the averages prices of cars based on types of transmission mode used by the car?

**Section IV: References**

**Dataset Link –**

<https://www.kaggle.com/datasets/bimervos/online-automotive-sales-statistics-volkswagen>

**Mural Link –**

<https://app.mural.co/t/dvassignment0284/m/dvassignment0284/1699563981327/42f506657e2cd9915e45641bf72383a216df88c2?sender=ua05eef829bc2a8f507a20491>