# Comprehensive Customer Value and Management Use Casegrafik, yazı tipi, logo, meneviş mavisi içeren bir resim Açıklama otomatik olarak oluşturuldu

The purpose of this document is to outline a comprehensive use case that encompasses key aspects of customer data analysis for a Data Analytics role. This use case aims to leverage customer data to gain insights into customer behavior, predict customer churn, estimate the Customer Lifetime Value (CLV), and perform customer segmentation. These tasks are critical for developing effective marketing strategies, improving customer retention, and ultimately driving business growth. The document will guide candidates through the process of analyzing the data sets provided, developing predictive models, and formulating recommendations for strategic business decisions.

## Data Sets Description

### `cust\_best\_sample.csv`:

This data set includes demographic information of customers. Each customer is provided with a unique identifier. However, there are some missing and irregular entries in this data set. Notably, the 'date\_of\_birth' column contains missing values, which should be considered during the data cleaning and preprocessing stage.  
- `unique\_customer\_id`: The unique identifier of the customer.  
- `gender`: The gender of the customer.  
- `date\_of\_birth`: The date of birth of the customer (Some values may be missing).

### `cust\_sample.csv`:

This data set contains unique identifiers for customers across different systems and can be used to link with another data set.  
- `cb\_customer\_id`: The identifier of the customer in another system.  
- `unique\_customer\_id`: The unique identifier of the customer.

### `trx\_sample.csv`:

Customer transaction data are located in this data set. Various financial details about the transactions are provided, but some of these may have been scaled or transformed. Candidates are expected to interpret and understand these attributes accurately.  
- `cb\_customer\_id`: The customer identifier that can be linked with `cust\_sample.csv`.  
- `transaction\_date`: The date of the transaction.  
- `amount\_after\_discount`: The transaction amount after discount (Values may have been transformed).  
- `cb\_branch\_id`: The identifier of the branch where the transaction took place.  
- `amount\_before\_discount`: The transaction amount before discount (Values may have been transformed).  
- `amount\_discount`: The discount amount applied (Values may have been transformed).

## Expectations from the Candidate

Candidates are expected to perform the following tasks using the datasets:

**1. Data Exploration and Cleaning** - Inspect the data to identify missing and inconsistent data points.  
 - Determine appropriate methods for filling in missing data points or cleaning the data.

**2. Feature Engineering**  
 - Derive new features from the existing data.  
 - Extract features reflecting customer behavior from transaction data.

**3. Data Association:**   
 - Link two sets of unique customer identifiers to create an integrated customer profile.  
 - Combine transaction data with customer demographic information to create a comprehensive customer database.

4. **Modeling and Evaluation**:

Develop at least one model:

- Churn prediction model.  
 - CLV prediction model.  
 - Customer segmentation model.

Select and apply appropriate metrics to assess the performance of the models.

**5. Dockerization**

Containerize the entire project, including:

* Data preprocessing scripts.
* Feature engineering and modeling code.
* Any model files and dependencies.

Write a **Dockerfile** that ensures:

* Reproducibility of the environment.
* Minimal image size for efficiency.

Test the Dockerized solution locally to ensure it runs seamlessly.

**6. Publishing and Sharing**

Push the Docker image to **Docker Hub**.

Create a **GitHub repository** containing:

* All project files, excluding sensitive data (e.g., raw datasets).
* A README.md file with:

Project overview.

Setup and run instructions using Docker.

Summary of findings and recommendations.

* Docker-related files, including the Dockerfile and instructions for pulling and running the Docker image.

Please provide access to both the Docker Hub repository and the GitHub repository.