# Borusan Otomotiv Data Analytics Specialist Technical Task

## Part 1 – Coding Skills

1) Please write a
   (a) SQL code
   (b) Python or R code

   which provide to convert the given dataset in the Excel document **SQL_Task.xlsx** into expected format.

## Part 2 – Data Science Skills

**Requirements:** Please submit the final results as a Jupyter notebook (with sufficient explanations) or a combination of code and documentation to explain the modeling, methodology, and results. Use whatever programming language you prefer

**Task 1: Predicting Vehicle Sale Timing at Auctions**

**Motivation:** Borusan Araç İhale (BAI) company operates a B2B platform for buying and selling used cars. The business unit wants to predict in which auction a vehicle will be sold and provide recommendations to potential sellers who are looking to sell their vehicles at auction.

**Your Task:** Your task is to build a ML model that predicts in which auction the vehicles, will be sold. Additionally, conduct exploratory data analysis (EDA) to identify the key factors that lead to earlier sales and provide actionable insights for the business unit to improve the likelihood of selling vehicles sooner.

The deliverables should include:

1. **Exploratory Data Analysis (EDA):**
   - Identify important features that influence vehicle sales timing.
   - Visualize and interpret the relationships between key features and auction outcomes.
2. **Predictive Model:**
   - Build a ML model to predict in which auction a vehicle will be sold.
   - Evaluate the model's performance using appropriate metrics.
3. **Actionable Insights:**
   - Based on the model and EDA, provide recommendations on what actions the business

unit can take to increase the chances of selling vehicles earlier.

**Data:**

**auction.sales.csv**

`vehicleid` Integer Unique identifier for each vehicle.
`Brand` String Brand of the vehicle (e.g., Renault, Peugeot).
`Model` String Model of the vehicle (e.g., Megane, Clio).
`Year` Integer Year the vehicle was manufactured.
`N_view` Integer Number of times the vehicle was viewed on the auction platform.
`N_offer` Integer Number of offers made for the vehicle.
`Seller_Type` String Type of seller (e.g., FILO).
`Ask_Price` Float Initial asking price for the vehicle (missing values are possible).
`Sale_Price` Float Final sale price for the vehicle.
`Target` Integer The auction number in which the vehicle was sold (the target variable).

**Task 2: Churn tol Revenue Forecast (BONUS)**

**Motivation:** Company 'X' follows its users that signed into its online service between Jan 23' and Apr 24'. Every group of new users is called cohorts, and the revenue the company gains from every group is sampled every month, is summarized in a table. From some reason, users that joined after Apr 24' are not counted in that table. Nevertheless, company X is still eager to develop a predictive model or a process that will estimate the monthly revenue until the end of 2024 (basing on the data it has)

**Your Task:** For every user group (cohort), you have its monthly revenue (in units of Million EUR) from the month they joined. Every group has its own unique users (a user cannot be in more than one cohort). Your task is to provide the total monthly revenue forecast from Jan 23' until Dec 24' basing on the given groups of users. It is recommended to support your model / process with visualization.

**Data:**

**Cohorts_rev.csv** – Churn DF with the following fields:

- `date` – Users joining date (month / year)
- `Cohort 1 Rev - Cohort 4 Rev` – Total monthly revenue (Million EUR) earned from the users in the specific cohort.