# Cars Market Analysis in Egypt

**Team Members:** Mahmoud Ahmed Abdelazez Shimy (Leader), Selim Sayed Hussein Mekawy  
**Supervisor:** Merna Saeed  
**Organization:** AST

## 1. Project Planning

**Project Title:** Cars Market Analysis in Egypt  
**Project Type:** Data Analysis / Data Science (Graduation Project)

**Project Description (brief):**  
This project analyses used-car market listings in Egypt by combining CSV datasets (scraped from Hatla2ee and a Kaggle dataset). The analysis explores pricing trends across brands, models, manufacturing year, transmission, fuel type, engine capacity, kilometers, governorates and color to provide actionable insights for dealerships, buyers and financial institutions.

**Main Objectives:** - Identify factors that significantly affect used-car prices in Egypt. - Produce location- and brand-level pricing insights and recommendations. - Create a reproducible data-processing pipeline starting from CSV files to cleaned, analyzed data and visualizations.

**Tools & Technologies:** Excel, Power BI, SQL, Git for version control, CSV files as the data source.

**Project Phases & Timeline:** - Research & Data Gathering — 1 week - Data Cleaning & Preparation — 2 weeks - Exploratory Data Analysis & Feature Engineering — 2 weeks - Modeling & Statistical Tests (if needed) — 1 week - Reporting & Documentation — 1 week.

**Team Roles (assigned):**

* **Mahmoud Ahmed Abdelazez Shimy** — Project lead, data cleaning lead, final reporting.
* **Selim Sayed Hussein Mekawy** — Data analysis & visualizations, exploratory analysis, assisting with feature engineering.

## 2. Stakeholder Analysis

| **Stakeholder** | **Role** | **Responsibilities** |
| --- | --- | --- |
| Project Team (Mahmoud, Selim) | Project execution, deliverables | Data collection, cleaning, analysis, reporting |
| Supervisor — Merna Saeed | Academic supervision | Project guidance, evaluation, feedback |
| Organization — AST | Project host / sponsor | Provide institutional support, possible dataset access or deployment support |
| Data Sources (Hatla2ee, Kaggle) — providers | Provide raw listing data | Allow lawful use of dataset; data quality responsibility |
| End Users | Dealerships, buyers, financial institutions, insurers | Use insights to price, buy/sell and underwrite loans/policies |
|  |  |  |

## 3. Database Design

Source data: CSV files combined from Hatla2ee (scraped) and a Kaggle dataset. For reproducibility and future extension, we propose a relational database schema (ERD) designed to store normalized listing data.

**Proposed Tables (logical):**

1. **brands** (brand\_id PK, name)  
2. **models** (model\_id PK, brand\_id FK → brands.brand\_id, name)  
3. **locations** (location\_id PK, governorate, city, raw\_location\_text)  
4. **listings** (listing\_id PK, model\_id FK → models.model\_id, location\_id FK → locations.location\_id, manufacture\_year, engine\_cc, transmission, fuel\_type, kilometers, color, body\_type, price\_egp, price\_currency, price\_numeric)  
5. **sources** (source\_id PK, name, url, notes) — to record Hatla2ee and Kaggle provenance  
6. **listing\_source** (listing\_id FK, source\_id FK, original\_url) — many-to-many if needed

**Primary Keys / Foreign Keys:** - brands.brand\_id (PK) - models.model\_id (PK), models.brand\_id (FK → brands.brand\_id) - locations.location\_id (PK) - listings.listing\_id (PK), listings.model\_id (FK → models.model\_id), listings.location\_id (FK → locations.location\_id)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_