Egypt University of Informatics

Faculty of Computer and Information Systems

Data Analysis Course – Project Report

Used Cars Market Analysis in Egypt

Students Name: Al Moataz bellah Mohamed, Ali Amin, Ahmed El Khateeb, Ahmed Sherif

Submission Date: 24/5/2024

# 1. Introduction

The automotive sector in Egypt, particularly the used car market, has undergone notable changes. Economic pressures and consumer demand have shifted interest toward used vehicles. This report leverages data visualizations to explore key trends such as brand distribution, pricing behavior based on mileage, model year, and transmission type.

# 2. Research Objectives

This report addresses the following questions:

1. How does mileage (Kilometers) affect car price?

2. Does color effect the pricing of a car?

3. Do newer models with high mileage sell for less than older low-mileage models?

4. Is there a statistically significant difference in the average price of used cars with automatic transmission versus manual transmission in Egypt?

# 3. Hypothesis Questions

H1: Cars with higher mileage are priced lower.

H2: "Does the color of a Hyundai used car significantly affect its market price in Egypt?"

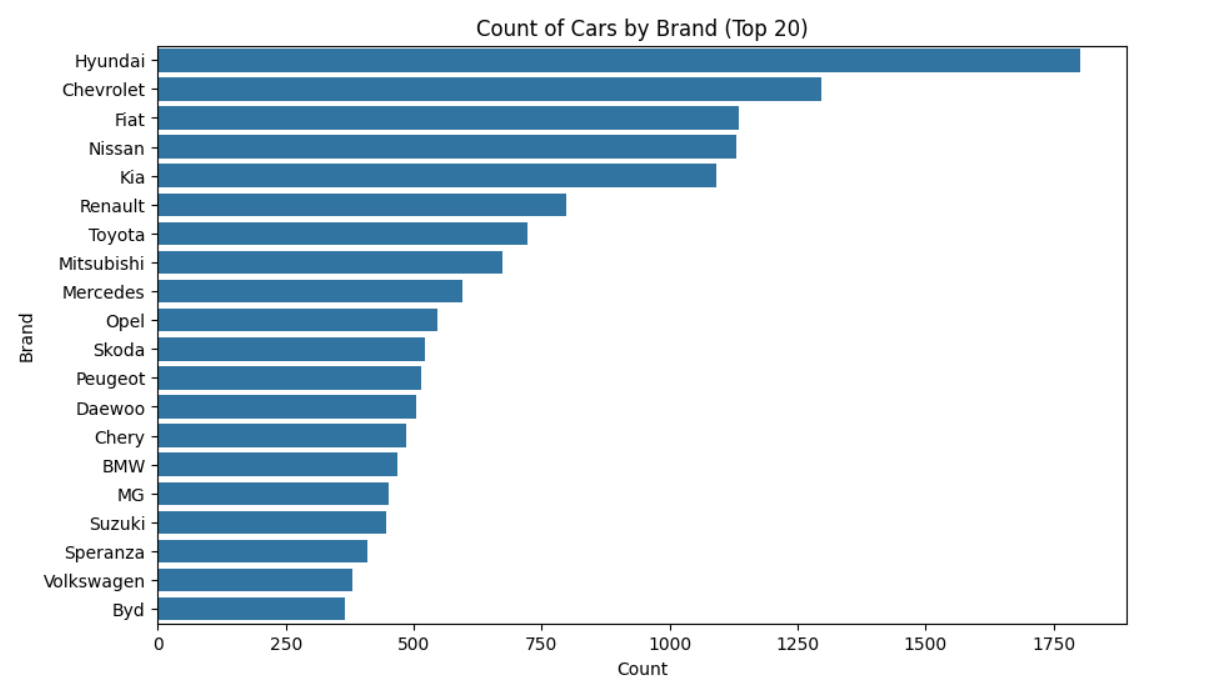
H3: Newer models retain value even when driven more.

H4: There is a statistically significant difference in average prices between automatic and manual transmission used cars in Egypt.

# 4. Visual Analysis & Hypothesis Testing

## 4.1 Count of Cars by Brand (Top 20)

This chart highlights the 20 most common car brands listed. Hyundai, Chevrolet, and Fiat lead in listing frequency.



## 

## Step 1: Define null and alternative hypotheses

## 4.2 Relationship between Price and Kilometers Driven

A scatter plot shows that, generally, higher mileage cars tend to be cheaper, though exceptions exist.

A graph of a car with blue dots

AI-generated content may be incorrect.

**. H₀**: There is no significant correlation between mileage and price of used cars.

**. H₁**: There is a significant correlation between mileage and price of used cars.

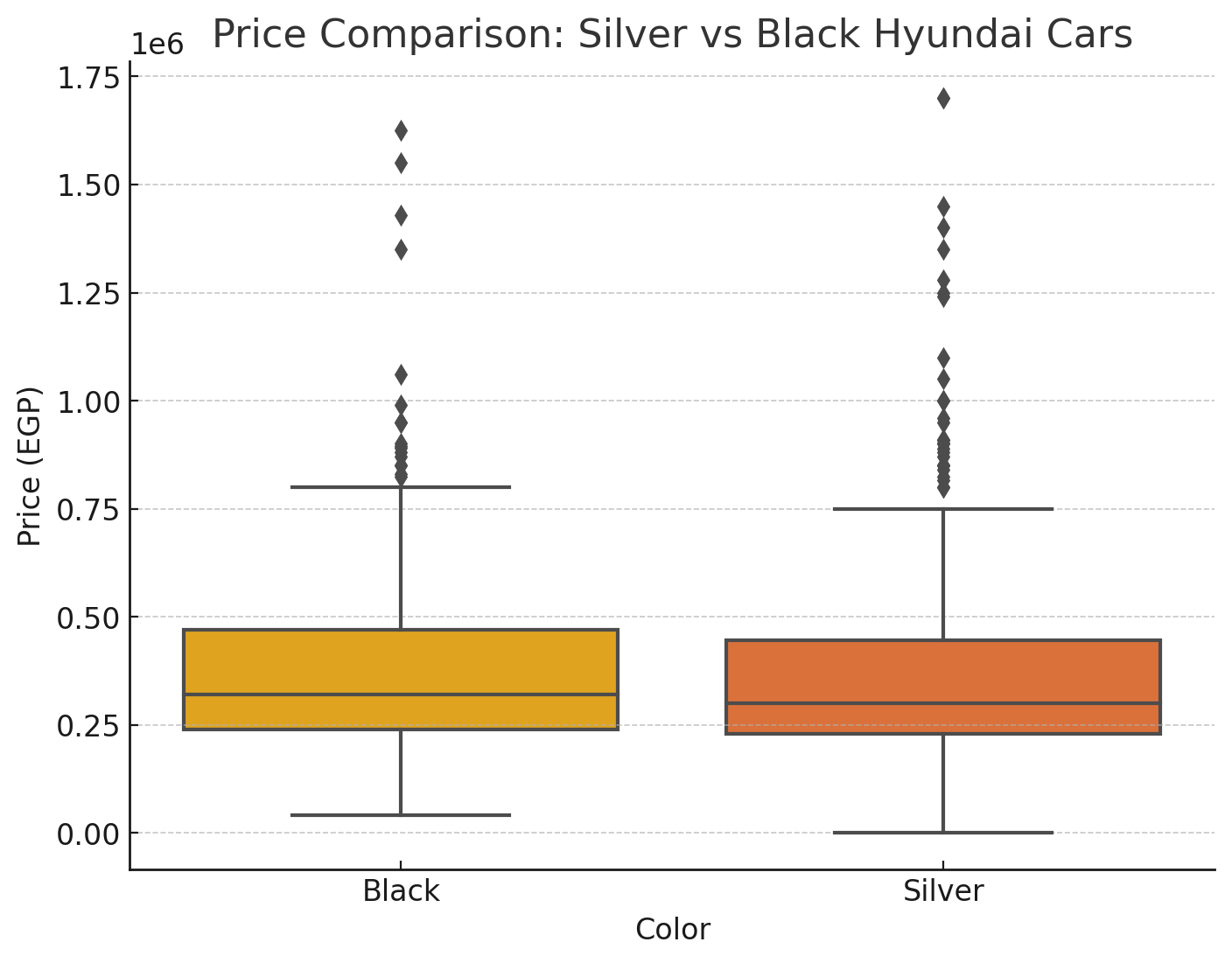
**Step 2: Choose the appropriate test**  
We used the **Pearson correlation test**, which is appropriate for examining the linear relationship between two continuous variables: mileage and price.

**Step 3: Calculate the p-value**

* **Pearson r** = -0.17
* **p-value** = 0.000

**Step 4: Determine statistical significance**  
Since **p < 0.05**, we **reject the null hypothesis**.  
✅ There is a statistically significant negative correlation between mileage and price.

**4.3 Relation between Color and Price**

****

**Step 1: Define Hypotheses**

* **H₀ (Null Hypothesis):** There is no significant difference in the average price of Silver and Black Hyundai cars.
* **H₁ (Alternative Hypothesis):** There is a significant difference in the average price of Silver and Black Hyundai cars.

## Step 2: Test Chosen

We used Welch’s t-test, which compares the mean prices of two independent groups that may have unequal variances.

## Step 3: Calculate the p-value

## p-value = 0.591635

**Step 4: Decision**

Since p > 0.05, we fail to reject the null hypothesis.  
⚠️ There is no statistically significant difference in price between Silver and Black Hyundai used cars.

## 4.4 Price by Model Year and Mileage Group

Vehicles that are newer and have low mileage are priced the highest, confirming expectations.

A graph of different colored bars

AI-generated content may be incorrect.  
**3. Model Year + Mileage Combined**

**Step 1: Define null and alternative hypotheses**

* **H₀**: There is no significant price difference between newer cars with high mileage and older cars with low mileage.
* **H₁**: There is a significant price difference between newer cars with high mileage and older cars with low mileage.

**Step 2: Choose the appropriate test**  
We used **Welch’s t-test** to compare prices between two independent groups:

* Group 1: Newer cars (after 2010) with high mileage
* Group 2: Older cars (before 2010) with low mileage

**Step 3: Calculate the p-value**

* **p-value** = 0.00
* **Step 4: Determine statistical significance**  
  Since **p < 0.05**, we **reject the null hypothesis**.  
  ✅ Newer high-mileage cars tend to retain more value than older low-mileage ones.

## 4.5 Price Distribution by Transmission Type

Automatic cars show a broader and generally higher price range compared to manual ones.

A graph of a car price

AI-generated content may be incorrect.

**2. Transmission Type vs. Price**

**Step 1: Define null and alternative hypotheses**

* **H₀**: There is no significant difference in the prices of used cars between automatic and manual transmission types.
* **H₁**: There is a significant difference in the prices of used cars between automatic and manual transmission types.

**Step 2: Choose the appropriate test**  
We used **Welch’s t-test** (a variation of the t-test that does not assume equal variances) to compare the mean prices of two independent groups.

**Step 3: Calculate the p-value**

* **p-value** = 1.26e-13

**Step 4: Determine statistical significance**  
Since **p < 0.05**, we **reject the null hypothesis**.  
✅ There is a statistically significant difference in average prices between automatic and manual cars.

# 5. Statistical Testing Framework

1. Define null and alternative hypotheses.  
2. Set significance level at α = 0.05.  
3. Apply correlation, group averages, and t-tests for hypothesis validation.  
4. For transmission comparison, a Welch’s t-test was conducted to evaluate price differences.  
5. Conclusions are based on p-values and confidence intervals.

# 6. Discussion & Conclusion

The analysis confirmed that newer models and luxury brands hold greater value. Mileage has a minimal standalone impact. A statistically significant difference was observed between average prices of automatic and manual cars, supporting the hypothesis that transmission type influences market price.