Questions

Intro Question Total sales in 2015 across all sellers
0 Show amount of cars sold per month
1 Which car brand sold the most vehicles in 2015
2 Which vehicle model is the most sold overall models across all brands
3 What category of vehicle is the most sold
4 What percentage of the sales do all types of vehicle encompass
5 Most popular colour for toyota cars
6 How many manuals and how many automatics are in the database sorted by each year of production (not sold)?
7 Average odometer of a car when sold
8 Most popular interiors?
9 Which seller has the highest total sales in $?
10 Rank sellers by the amount of cars they sold in total descending

Exploratory Data Analysis Report

Introduction

This dataset provides a comprehensive collection of information pertaining to the sales transactions of various vehicles, new and used from 19 to 2015.

Available details are the following :

Year	Condition Rating 1 To 49
Brand	Odometer Reading 1 To 999’999
Model	Exterior Colour
Trim	Interior Colour
Category	Seller Name
Transmission Type	Manheim Market Report (Mmr) Value
Vin (Vehicle Identification Number)	Selling Price
State Of Registration	Sale Date

Data Cleaning

Deleted the column MMR as it is not relevant for our analysis.

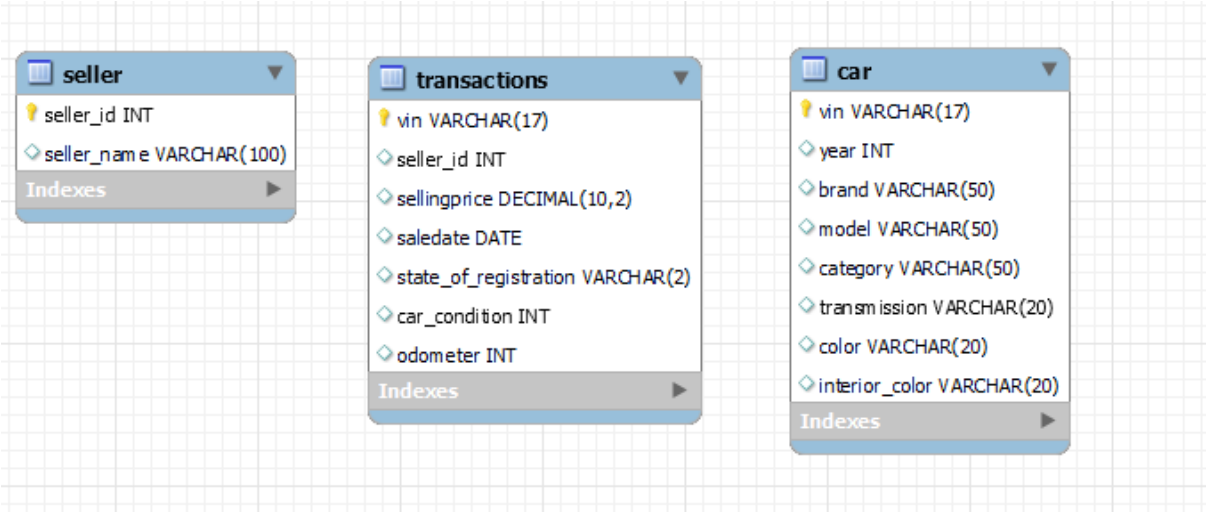
Deleted the trim column as it is not relevant since we are conducting a general analysis on the cars market.

Cleaned the sale date column to obtain only the day, month and year in SQL preferred format.

Eliminated every row where at least one column has a null value, that way we are working with a dataset without null values :) .

mySQL Upload

I exported my python cleaned dataset into an excel file, sent it to google sheets to divide it into 3 tables linked by a common id/key.



Then downloaded each table as a csv to import in mySQL.

The schema and tables creation was done by SQL queries.

I also had to change the permissions inside the my.ini file from SQL server to allow me to query data import from a local csv file.

I then queried the import from the csv files and checked if all matched correctly.