Advanced Excel Data Analysis: Car Dataset Insights

Overview

This project demonstrates data manipulation and analysis using Microsoft Excel. The dataset consists of various car listings with attributes such as fuel type, registration city, assembly type, transmission, and condition. Several tasks were performed to extract, transform, and analyse the data.

Tasks Performed

Task 1: Data Extraction using VLOOKUP

Used VLOOKUP to fetch the following columns from **Sheet2** to **Sheet1**:

- Fuel Type

- Registration City

- Assembly Type

- Transmission Type

- Condition

Task 2: Creating a Selection Column

A new column named **Selection** was created to classify cars based on specific conditions:

- Fuel type is either **Petrol** or **CNG**.

- The registration city **is not Karachi**.

- The car make is **not Toyota or Honda**.

Formula used: IF(AND(OR(Fuel="Petrol", Fuel="CNG"), NOT(Registration City="Karachi"), NOT(OR(Make="Toyota", Make="Honda"))), "SELECTED", "NOT SELECTED").

Task 3: Extracting Car Model

The **Model** was extracted from the "Make Model" column (e.g., extracting "Mehran VXR" from "Suzuki Mehran VXR").

Methods used:

- **Text to Columns**
- **Concatenation using CONCAT()**
- **String Functions: FIND(), RIGHT(), LEN()**

Task 4: Calculating Miles per Year

A new column **Miles_Per_Year** was created by:

- Converting kilometers to miles.
- Dividing by the number of years passed since the car's model year (assuming the current year is **2024**).
- Rounding off to three decimal places.

Formula: =ROUND((KMs * 0.621371) / (2024 - Car_Year), 3).

Analysis Questions and Answers

1. **What is the average price of CNG models in Suzuki?**
   - Used **AVERAGEIF()** function to filter Suzuki CNG models and compute the average price.

2. **How many local, automatic models of Honda are there in Lahore?**
   - Used **COUNTIFS()** function to count cars meeting all three conditions.

3. **What is the total price of imported, manual cars in Karachi?**
   - Used **SUMIFS()** function to sum the prices based on conditions.

4. **What is the difference between the average price of a local car and an imported car?**
   - Used **AVERAGEIF()** for both categories and calculated the difference.

5. **What is the percentage of CNG cars in the dataset?**
   - Used **COUNTIF()** to count CNG cars and divided it by the total number of cars to get the percentage.

Conclusion

The project showcased how Excel formulas and functions can be leveraged for data analysis. Key skills applied include **VLOOKUP, IF conditions, text extraction, mathematical operations, and data aggregation functions**. The insights obtained from the dataset can help in better understanding market trends and car pricing structures.

Dataset:

Ad ID	Car_Name	Make	Model	Variants	Year	KM_driven	miles_driven	miles_per_year	new_year(2024-year)
1019823763	Toyota Corolla Altis 1.6 cc	Toyota	Corolla	Altis	2015	130000	80778.23	8975.36	9
1032353392	Mitsubishi mini cab urgent sale	Mitsubishi	Minica		2012	115000	71457.67	5954.81	12
1035626806	Suzuki Every 2013 Model 2019 Registered Jet B	Suzuki	Every		2013	120000	74564.52	6778.59	11
1036683049	Mehran 2016 VXR original condition	Suzuki	Mehran	VXR	2016	38500	23922.78	2990.35	8
1036970933	Daihatsu Cuore 2004	Daihatsu	Cuore		2004	100000	62137.10	3106.86	20
1038612589	Mitsubishi Minica 2 door Japanese Car for sale	Mitsubishi	Minica		2007	133500	82953.03	4879.59	17
1043285959	toyota corolla 2005	Toyota	Corolla	XLI	2005	195000	121167.35	6377.23	19
1043893641	chska party call na kra	KIA	Classic		2005	129049	80187.31	4220.38	19
1045295959	Suzuki Wagon R FA Package 2019 Model 2022	Suzuki	Wagon	R	2019	18360	11408.37	2281.67	5
1045445134	Carry Daaba Minicap Mitsubishi	Mitsubishi	Minica		2017	170000	105633.07	15090.44	7
1046475308	Kia Spectra 2001 1.5 Automatic. Customised pri	KIA	Spectra		2001	50000	31068.55	1350.81	23
1046608801	suzuki every turbo 5gair Exchange posibl with H	Suzuki	Every		2016	123456	76711.98	9589.00	8
1046845167	Suzuki Bolan	Suzuki	Bolan		2017	38000	23612.10	3373.16	7
1048171759	Honda Civic VTI Automatic	Honda	Civic	Prosmetic	2005	104000	64622.58	3401.19	19
1049652233	Coure 2004 model	Daihatsu	Cuore		2004	260000	161556.46	8077.82	20
1051027975	Honda city I. D. S. I 2006	Honda	City	IDSI	2006	117790	73191.29	4066.18	18
1051075341	Suzuki Mehran VXR 2009	Suzuki	Mehran	VXR	2009	50260	31230.11	2082.01	15
1051252471	kia classic car	KIA	Classic		2002	13452	8358.68	379.94	22
1052330175	Honda city idsi vario Automatic 2007	Honda	City	Vario	2007	120000	74564.52	4386.15	17
1053425910	alto car	Suzuki	Alto		2005	200000	124274.20	6540.75	19
1054393126	SUZUKI	Cultus VXR	2001 Model	PETROL	Suzuki	Cultus	VXR	2001	20000	12427.42	540.32	23
1054535697	Changan Alsvin Lumier Silver 2022	Changan	Alsvin		2022	15000	9320.57	4660.28	2
1054733477	Suzuki Alto 2007 Like new car	Suzuki	Alto		2007	83000	51573.79	3033.75	17
1055305594	Santro Club EFI Original Condition	Hyundai	Santro		2006	153000	95069.76	5281.65	18
1055598632	cultus.	Suzuki	Cultus	VXR	2016	87000	54059.28	6757.41	8
1055660992	honda city prismatic Total guniene 03006981726	Honda	City	Aspire	2019	47000	29204.44	5840.89	5
1056841265	baleno car	Suzuki	Baleno		2005	100000	62137.10	3270.37	19
1057673672	Mitsubishi minica	Mitsubishi	Minica		2013	85000	52816.54	4801.50	11
1058158069	Suzuki wegonR VXL 2021 100% geniune low mil	Suzuki	Wagon	R	2021	21000	13048.79	4349.60	3
1058344102	coure 2006 model total geniune	Daihatsu	Cuore		2006	130000	80778.23	4487.68	18
1058709644	Daihatsu Move 2015	Daihatsu	Move		2015	86000	53437.91	5937.55	9
1059010266	Honda city auto fore sale 2005	Honda	City	IVTEC	2005	100000	62137.10	3270.37	19
1059759696	A Gift for Honda Lovers. Honda Civic 1.8 i. vtec	Honda	Civic	Prosmetic	2007	182000	113089.52	6652.32	17

Price	FueL	Registrartion City	Assembly	Transmission	Condition	selection
4000000	Petrol	Karachi	Local	Automatic	Used	NOT SELECTED
1150000	Petrol	Lahore	Imported	Automatic	Used	SELECTED
1900000	Petrol	Lahore	Imported	Automatic	Used	SELECTED
1180000	Petrol	Hyderabad	Local	Manual	Used	SELECTED
700000	Petrol	Karachi	Local	Manual	Used	NOT SELECTED
1300000	Petrol	Lahore	Imported	Manual	Used	SELECTED
1750000	Petrol	Islamabad	Local	Manual	Used	SELECTED
550000	Petrol	Islamabad	Local	Manual	Used	SELECTED
3500000	Petrol	Unregistered	Imported	Automatic	Used	SELECTED
1500000	Petrol	Lahore	Imported	Automatic	Used	SELECTED
1200000	Petrol	Lahore	Imported	Automatic	Used	SELECTED
1750000	Petrol	Lahore	Imported	Automatic	Used	SELECTED
1300000	Petrol	Karachi	Local	Manual	Used	NOT SELECTED
1515000	Petrol	Karachi	Local	Automatic	Used	NOT SELECTED
650000	Petrol	Faisalabad	Imported	Manual	Used	SELECTED
1625000	Petrol	Karachi	Local	Manual	Used	NOT SELECTED
540000	Petrol	Karachi	Local	Manual	Used	NOT SELECTED
595000	CNG	Haripur	Imported	Manual	Used	SELECTED
1800000	Petrol	Karachi	Local	Automatic	Used	NOT SELECTED
750000	Petrol	Lahore	Imported	Manual	Used	SELECTED
500000	Petrol	Karachi	Local	Manual	Used	NOT SELECTED
4250000	Petrol	Islamabad	Local	Automatic	Used	SELECTED
890000	Petrol	Lahore	Local	Manual	Used	SELECTED
950000	Petrol	Lahore	Imported	Manual	Used	SELECTED
1825000	Petrol	Karachi	Local	Manual	Used	NOT SELECTED
3995000	Petrol	Karachi	Local	Automatic	Used	NOT SELECTED
850000	Petrol	Islamabad	Local	Manual	Used	SELECTED
1345000	Petrol	Islamabad	Imported	Automatic	Used	SELECTED
2729000	Petrol	Punjab	Local	Manual	Used	SELECTED
920000	Petrol	Attock	Imported	Manual	Used	SELECTED
2325000	Petrol	Lahore	Imported	Automatic	Used	SELECTED
1450000	Petrol	Karachi	Imported	Automatic	Used	NOT SELECTED
1800000	Petrol	Karachi	Local	Automatic	Used	NOT SELECTED
2025000	Petrol	Karachi	Local	Manual	Used	NOT SELECTED

Questions:

	TASKS	[5 marks each]
	Task1	Use vlookup to bring the following columns from sheet2 to sheet1: 1) fuel 2) registration city 3) assembly 4) transmission 5) condition
	Task2	Create a new column called **selection** that will show SELECTED for the cars that meet ALL the following conditions: 1) Fuel is either Petrol or CNG 2) City is not Karachi 3) Make could be any except Toyota and Honda Column should show NOT SELECTED for all the rest You will need to combine AND(), OR() and NOT() for the condition for IF()
	Task3	Extract Model from the Make Model column into a separate column (for example extract Mehran VXR from Suzuki Mehran VXR) hint: you can do this in different ways. You can use text to column and then CONCAT(). Look into how you can also do this in one go using the FIND(), RIGHT() and LEN() functions
	Task4	make a miles_per_year column (change kms to miles and then divide by the number of years that have passed between car year and current year 2024) Round it off to 3 decimal places

QUESTIONS	[2 marks each]			
1	what is the avg price of CNG models in suzuki?	816292.9293		
2	how many local, automatic models in Honda in Lahore?	304		
3	what is the total of all the prices of imported, manual cars in Karachi?	342497996		
4	what is the difference between avg price of a local car and avg price of an imported car?	2146813.503	1820826	325987.8
5	what is the percentage of CNG cars in the dataset?	2.14		