Bussiness Data Management(BDE)

BY

Shaik Irfan

sales

ETL

STEPS...!!

E

Extract

Data Cleaning

Data Duplicates,Data Mapping, Date Format,Consolidates Vlookup, Pivot Tables, Data Validation, Replace, Conditional Formatting,

T

Transfrom

Data Transfrom

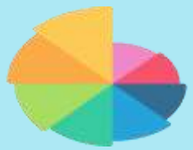
Aggregate, DatePart, Having, Format, Top clause, Ranking Functions, Joins , Set operators

L

Load

Data Visualization

Dual Axis , LOD , Show Me, Data Blending , Data Types, Filters , Hierachy , Aggregate, Calculated Fileds



carsales

E

Extract

Data Cleaning

Data Collection

- Data pertaining to various attributes such as Cus Name, Religon, Dealer Name, Gender, category, Company, price, Model, Transmission , Body Style etc in excel sheets from external sources such as Internal data, D&B Hoovers, Cars data of the Dealer.
- Initially a template in MS-Excel for the master data sheet was prepared and a few important attributes which need to be captured in it and subsequently into the database were finalized after thorough discussions with the team.
- Detailed description of the steps followed for collection of data and collating it into the master database is as follows.

Data Analysis factors

Customer Segmentation & Data Quality Assurance:

Purpose: Analyze customer demographics and preferences for segmentation, but encounter challenges in ensuring data accuracy, consistency, and completeness across columns like customer name,gender, Month Format and Currency Format.

Dealer Performance Evaluation & Data Integration Challenges:

Purpose: Assess dealer performance across regions, yet face difficulties in integrating data from various sources (e.g., different dealers' records) due to differences in formats and structures.

Market Analysis & Data Privacy/Compliance:

Purpose: Gain insights into market demand and trends, but must ensure compliance with data privacy regulations when handling customer information (e.g., customer names) and dealer data.

Price Optimization & Handling Unstructured Data:

Purpose: Analyze pricing data to optimize pricing strategies based on car features and preferences, but encounter challenges in processing unstructured data (e.g., dealer names, car company names, Month , Price) to extract meaningful insights.

Trend Analysis & Time-Based Analysis:

Purpose: Track sales trends over time to identify seasonality and market fluctuations, but require effective incorporation of the date column into analysis to ensure accurate time-series analysis techniques are employed.

I compiled data from D&B Hoovers, gathering information on 28,000 dealers, including attributes like month, gender, price, dealer name, and annual income, as depicted in the Excel sheet below

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Car_id	Month_No	Month_Name	Quarter_NO	Year	Date	Customer Name	Gender_Full	Gender	Annual_Income	Dealer_Name	Company	Model	Engine	Transmission
2	C_CND_000001	43832	43832	Q1	2020	02-01-2020	Geraldine	Male	M	13500	Buddy Storbeck's Diesel Service Inc	Ford	Expedition	Doublea Overhead Camshaft	Auto
3	C_CND_000002	43832	43832	Q1	2020	02-01-2020	Gia	Male	M	1480000	C & M Motors Inc	Dodge	Durango	Doublea Overhead Camshaft	Auto
4	C_CND_000003	43832	43832	Q1	2020	02-01-2020	Gianna	Male	M	1035000	Capitol KIA	Cadillac	Eldorado	Overhead Camshaft	Manual
5	C_CND_000004	43832	43832	Q1	2020	02-01-2020	Giselle	Male	M	13500	Chrysler of Tri-Cities	Toyota	Celica	Overhead Camshaft	Manual
6	C_CND_000005	43832	43832	Q1	2020	02-01-2020	Grace	Male	M	1465000	Chrysler Plymouth	Acura	TL	Doublea Overhead Camshaft	Auto
7	C_CND_000006	43832	43832	Q1	2020	02-01-2020	Guadalupe	Male	M	850000	Classic Chevy	Mitsubishi	Diamante	Overhead Camshaft	Manual
8	C_CND_000007	43832	43832	Q1	2020	02-01-2020	Hailey	Male	M	1600000	Clay Johnson Auto Sales	Toyota	Corolla	Overhead Camshaft	Manual
9	C_CND_000008	43832	43832	Q1	2020	02-01-2020	Graham	Male	M	13500	U-Haul CO	Mitsubishi	Galant	Doublea Overhead Camshaft	Auto
10	C_CND_000009	43832	43832	Q1	2020	02-01-2020	Naomi	Male	M	815000	Rabun Used Car Sales	Chevrolet	Malibu	Overhead Camshaft	Manual
11	C_CND_000010	43832	43832	Q1	2020	02-01-2020	Grayson	Female	F	13500	Rabun Used Car Sales	Ford	Escort	Doublea Overhead Camshaft	Auto
12	C_CND_000011	43832	43832	Q1	2020	02-01-2020	Gregory	Male	M	13500	Race Car Help	Acura	RL	Overhead Camshaft	Manual
13	C_CND_000012	43832	43832	Q1	2020	02-01-2020	Amar'E	Male	M	13500	Race Car Help	Nissan	Pathfinder	Doublea Overhead Camshaft	Auto
14	C_CND_000013	43832	43832	Q1	2020	02-01-2020	Griffin	Male	M	885000	Saab-Belle Dodge	Mercury	Grand Marquis	Doublea Overhead Camshaft	Auto
15	C_CND_000014	43832	43832	Q1	2020	02-01-2020	Harrison	Male	M	13500	Scrivener Performance Engineering	BMW	323i	Doublea Overhead Camshaft	Auto
16	C_CND_000015	43832	43832	Q1	2020	02-01-2020	Zainab	Male	M	722000	Buddy Storbeck's Diesel Service Inc	Chrysler	Sebring Coupe	Overhead Camshaft	Manual
17	C_CND_000016	43832	43832	Q1	2020	02-01-2020	Zara	Male	M	746000	C & M Motors Inc	Subaru	Forester	Overhead Camshaft	Manual
18	C_CND_000017	43832	43832	Q1	2020	02-01-2020	Zoe	Female	F	535000	Capitol KIA	Hyundai	Accent	Overhead Camshaft	Manual
19	C_CND_000018	43832	43832	Q1	2020	02-01-2020	Zoey	Female	F	570000	Chrysler of Tri-Cities	Cadillac	Eldorado	Doublea Overhead Camshaft	Auto
20	C_CND_000019	43832	43832	Q1	2020	02-01-2020	Aaliyah	Male	M	685000	Chrysler Plymouth	Toyota	Land Cruiser	Doublea Overhead Camshaft	Auto
21	C_CND_000020	43832	43832	Q1	2020	02-01-2020	Abigail	Male	M	455000	Classic Chevy	Honda	Accord	Doublea Overhead Camshaft	Auto
22	C_CND_000021	43832	43832	Q1	2020	02-01-2020	Adrianna	Male	M	13500	Clay Johnson Auto Sales	Toyota	4Runner	Overhead Camshaft	Manual
23	C_CND_000022	43832	43832	Q1	2020	02-01-2020	Joshua	Male	M	2500000	Classic Chevy	Infiniti	I30	Doublea Overhead Camshaft	Auto
24	C_CND_000023	43832	43832	Q1	2020	02-01-2020	Marcus	Male	M	585000	Diehl Motor CO Inc	Audi	A4	Overhead Camshaft	Manual
25	C_CND_000024	43832	43832	Q1	2020	02-01-2020	Arthur	Male	M	920000	Star Enterprises Inc	Porsche	Carrera Cabrio	Doublea Overhead Camshaft	Auto
26	C_CND_000025	43832	43832	Q1	2020	02-01-2020	Lizzie	Male	M	672000	Suburban Ford	Volkswagen	Jetta	Doublea Overhead Camshaft	Auto
27	C_CND_000026	43832	43832	Q1	2020	02-01-2020	Florian	Male	M	801250	Tri-State Mack Inc	Dodge	Viper	Doublea Overhead Camshaft	Auto
28	C_CND_000027	43832	43832	Q1	2020	02-01-2020	Cassandre	Female	F	820000	U-Haul CO	Buick	Regal	Doublea Overhead Camshaft	Auto
29	C_CND_000028	43832	43832	Q1	2020	02-01-2020	Syrielle	Male	M	791000	Progressive Shippers Cooperative Association No	Chrysler	LHS	Overhead Camshaft	Manual
30	C_CND_000029	43832	43832	Q1	2020	02-01-2020	Sloane	Male	M	13500	Race Car Help	Chrysler	LHS	Overhead Camshaft	Manual

Formating Date

1

Month_No	Month_Name	Quarter_N0	Year	Date	Cu
43832	43832	Q1	2020	02-01-2020	Ge
43832	43832	Q1	2020	02-01-2020	Gi
43832	43832	Q1	2020	02-01-2020	Gi
43832	43832	Q1	2020	02-01-2020	Gi
43832	43832	Q1	2020	02-01-2020	Gr
43832	43832	Q1	2020	02-01-2020	Gu
43832	43832	Q1	2020	02-01-2020	Ha
43832	43832	Q1	2020	02-01-2020	Gr
43832	43832	Q1	2020	02-01-2020	Na
43832	43832	Q1	2020	02-01-2020	Gr
43832	43832	Q1	2020	02-01-2020	Gr
43832	43832	Q1	2020	02-01-2020	An
43832	43832	Q1	2020	02-01-2020	Gr
43832	43832	Q1	2020	02-01-2020	Ha
43832	43832	Q1	2020	02-01-2020	Za
43832	43832	Q1	2020	02-01-2020	Za
43832	43832	Q1	2020	02-01-2020	Za
43832	43832	Q1	2020	02-01-2020	Zo
43832	43832	Q1	2020	02-01-2020	Zo
43832	43832	Q1	2020	02-01-2020	Aa
43832	43832	Q1	2020	02-01-2020	Ab
43832	43832	Q1	2020	02-01-2020	Ad
43832	43832	Q1	2020	02-01-2020	Jos
43832	43832	Q1	2020	02-01-2020	Ma
43832	43832	Q1	2020	02-01-2020	Ar

2

D2 ="Q" & ROUNDUP(MONTH(F3)/3,0)

	A	B	C	D	E
1	Car_id	Month_No	Month_Name	Quarter_N0	Year
2	C_CND_000001	1	Jan	Q1	2020
3	C_CND_000002	1	Jan	Q1	2020
4	C_CND_000003	1	Jan	Q1	2020
5	C_CND_000004	1	Jan	Q1	2020
6	C_CND_000005	1	Jan	Q1	2020
7	C_CND_000006	1	Jan	Q1	2020
8	C_CND_000007	1	Jan	Q1	2020
9	C_CND_000008	1	Jan	Q1	2020
10	C_CND_000009	1	Jan	Q1	2020
11	C_CND_000010	1	Jan	Q1	2020
12	C_CND_000011	1	Jan	Q1	2020
13	C_CND_000012	1	Jan	Q1	2020
14	C_CND_000013	1	Jan	Q1	2020
15	C_CND_000014	1	Jan	Q1	2020
16	C_CND_000015	1	Jan	Q1	2020
17	C_CND_000016	1	Jan	Q1	2020
18	C_CND_000017	1	Jan	Q1	2020
19	C_CND_000018	1	Jan	Q1	2020
20	C_CND_000019	1	Jan	Q1	2020
21	C_CND_000020	1	Jan	Q1	2020
22	C_CND_000021	1	Jan	Q1	2020
23	C_CND_000022	1	Jan	Q1	2020
24	C_CND_000023	1	Jan	Q1	2020
25	C_CND_000024	1	Jan	Q1	2020
26	C_CND_000025	1	Jan	Q1	2020
27	C_CND_000026	1	Jan	Q1	2020

"In this dataset, the initial format for Month_No and Month_Name is represented as '43832'. By employing steps 2 involving 'Format', 'Round', '&', and 'Month', we can accurately derive the Quarter, Month_No, and 'DDD' format to obtain concise month names like 'Jan'."

Test Tabel

\# A test table is typically employed to rewrite or replace values within an existing dataset. In our dataset, for instance, the 'Colour' column comprises entries such as 'BI', 'PW', and 'R', while the 'Gender' column contains 'M', 'Fe', and 'P'. To enhance data mapping, we can utilize the VLOOKUP function.

=VLOOKUP(Lookup_value, Table_array, Col_index, Match)
=VLOOKUP(I2, Sheet2!$A$357:$B$361, 2, 0)

\# In this example, 'I2' represents the lookup value, 'Sheet2!$A$357:$B$361' denotes the test table range, '2' indicates the column index containing the replacement values, and '0' signifies an exact match.

Test Sheet

GENDER			Color			Month_Wise	Month_No
Fe	Female		Black	Black		Jan	1
M	Male		Pale White	Pale White		Feb	2
Female	Female		P W	Pale White		Mar	3
Male	Male		BI	Black		Apr	4
F	Female		R	Red		May	5
			Red	Red		Jun	6
						Jul	7
						Aug	8
						Sep	9
						Oct	10
						Nov	11
						Dec	12

Test Tabel & Vlookup

In the dataset, the 'Gender' column contains various representations such as 'F', 'Fe', 'Female', 'M', and 'Male'. Utilizing a test table alongside Vlookup, we can efficiently replace these diverse values in a quick and effective manner, streamlining the cleaning process.

Test Table	
GENDER	
Fe	Female
M	Male
Female	Female
Male	Male
F	Female

Customer Nam	Gender_Full	Gender
	=VLOOKUP(I2,Sheet2!$A$357:$B$361,2,0)	
	VLOOKUP(lookup_value, table_array, col_ind	
Gianna	Male	M
Giselle	Male	M
Grace	Male	M
Guadalupe	Male	M
Hailey	Male	M
Graham	Male	M
Naomi	Male	M
Grayson	Female	F
Gregory	Male	M
Amar'E	Male	M
Griffin	Male	M
Harrison	Male	M
Zainab	Male	M
Zara	Male	M
Zoe	Female	F
Zoey	Female	F
Aaliyah	Male	M
Abigail	Male	M
Adrianna	Male	M
Joshua	Male	M
Marcus	Male	M
Arthur	Male	M
Lizzie	Male	M

Vlookup

Gender_Full	Year	Date
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Female	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Female	2020	02-J
Female	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J
Male	2020	02-J

Sort A to Z
Sort Z to A
Sort by Color
Sheet View
Clear Filter From "Gender"
Filter by Color
Text Filters
Search
(Select All)
F
Fe
Female
M
Male
OK

Currency Format $

"The dataset utilizes dollar ($) currency format for numerical values, ensuring alignment and consistency. This facilitates clear and visually appealing data visualization, enhancing the presentation of financial information."

R
Price ($)
$ 26,000
$ 19,000
$ 31,500
$ 14,000
$ 24,500
$ 12,000
$ 14,000
$ 42,000
$ 82,000
$ 15,000
$ 31,000
$ 46,000
$ 9,000
$ 15,000
$ 26,000
$ 17,000
$ 18,000
$ 31,000
$ 33,000
$ 21,000
$ 25,000
$ 21,000
$ 12,000
$ 18,000
$ 22,000

J
Annual_Income
$ 13,500
$ 14,80,000
$ 10,35,000
$ 13,500
$ 14,65,000
$ 8,50,000
$ 16,00,000
$ 13,500
$ 8,15,000
$ 13,500
$ 13,500
$ 13,500
$ 8,85,000
$ 13,500
$ 7,22,000
$ 7,46,000
$ 5,35,000
$ 5,70,000
$ 6,85,000
$ 4,55,000
$ 13,500
$ 25,00,000
$ 5,85,000
$ 9,20,000
$ 6,72,000
$ 8,01,250

Data Validation

In Order to restrict the kind of data that user enters we will be using data validation for Phone number of Customers. Here is an example of mobile number is shown below :

In Order to check whether mobile entered is 7 digits we need to go for data validation

S	T	U	V
Dealer_No	Body_Sty	Phone	Dealer_Regic
06457-3834	SUV	8264678	Middletown
60504-7114	SUV	6848	a
38701-8047	Passenger	7298	lle
99301-3882	SUV	6257	
53546-9427	Hatchback	7081483	Janesville
85257-3102	Hatchback	7315216	Scottsdale
78758-7841	Passenger	7727879	Austin
78758-7841	Passenger	6206512	Austin
85257-3102	Hardtop	7194857	Pasco
85257-3102	Passenger	7836892	Scottsdale
78758-7841	SUV	7995489	Austin
78758-7841	Hardtop	7288103	Pasco
60504-7114	SUV	6842408	Aurora
38701-8047	Hatchback	7558767	Greenville
06457-3834	Sedan	7677191	Middletown

Enter Your Number
Mobile no Should be 7 Digit

Data Profiling

Data profiling of our car sales dataset offers insights into the quality, patterns, and distribution of our sales data. By assessing attributes such as sales figures, regions, colors, and customer details,Car_ID , Gender,Body_Style we can uncover trends, identify target markets, and detect any anomalies or data issues. This process enables us to ensure the accuracy and reliability of our car sales data, empowering informed decision-making and strategic planning for business growth.

	Car_id	Customer Name	Gender_Full	Color_Name	Body_Style
	Valid 100% Error 0% Empty 0% 1000 distinct, 1000 unique	Valid 100% Error 0% Empty 0% 743 distinct, 555 unique	Valid 100% Error 0% Empty 0% 2 distinct, 0 unique	Valid 100% Error 0% Empty 0% 3 distinct, 0 unique	Valid 100% Error 0% Empty 0% 5 distinct, 0 unique
1	C_CND_000001	Geraldine	Male	Black	SUV
2	C_CND_000002	Gia	Male	Black	SUV
3	C_CND_000003	Gianna	Male	Red	Passenger
4	C_CND_000004	Giselle	Male	Pale White	SUV
5	C_CND_000005	Grace	Male	Red	Hatchback
6	C_CND_000006	Guadalupe	Male	Pale White	Hatchback
7	C_CND_000007	Hailey	Male	Pale White	Passenger
8	C_CND_000008	Graham	Male	Pale White	Passenger
9	C_CND_000009	Naomi	Male	Pale White	Hardtop
10	C_CND_000010	Grayson	Female	Pale White	Passenger
11	C_CND_000011	Gregory	Male	Pale White	SUV
12	C_CND_000012	Amar'E	Male	Pale White	Hardtop
13	C_CND_000013	Griffin	Male	Black	SUV
14	C_CND_000014	Harrison	Male	Pale White	Hatchback
15	C_CND_000015	Zainab	Male	Pale White	Sedan
16	C_CND_000016	Zara	Male	Pale White	Hatchback

After cleaning, the car sales dataset appears as follows :

	A	B	C	D	E	F	G	H	I	J	K
1	Car_id	Month_No	Month_Name	Quarter_N0	Year	Date	Customer Nam	Gender_Full	Gender	Annual_Income	Dealer_Name
2	C_CND_000001	1	Jan	Q1	2020	02-Jan-20	Geraldine	Male	M	$ 13,500	Buddy Storbeck's Diesel Service Inc
3	C_CND_000002	1	Jan	Q1	2020	02-Jan-20	Gia	Male	M	$ 14,80,000	C & M Motors Inc
4	C_CND_000003	1	Jan	Q1	2020	02-Jan-20	Gianna	Male	M	$ 10,35,000	Capitol KIA
5	C_CND_000004	1	Jan	Q1	2020	02-Jan-20	Giselle	Male	M	$ 13,500	Chrysler of Tri-Cities
6	C_CND_000005	1	Jan	Q1	2020	02-Jan-20	Grace	Male	M	$ 14,65,000	Chrysler Plymouth
7	C_CND_000006	1	Jan	Q1	2020	02-Jan-20	Guadalupe	Male	M	$ 8,50,000	Classic Chevy
8	C_CND_000007	1	Jan	Q1	2020	02-Jan-20	Hailey	Male	M	$ 16,00,000	Clay Johnson Auto Sales
9	C_CND_000008	1	Jan	Q1	2020	02-Jan-20	Graham	Male	M	$ 13,500	U-Haul CO
10	C_CND_000009	1	Jan	Q1	2020	02-Jan-20	Naomi	Male	M	$ 8,15,000	Rabun Used Car Sales
11	C_CND_000010	1	Jan	Q1	2020	02-Jan-20	Grayson	Female	F	$ 13,500	Rabun Used Car Sales
12	C_CND_000011	1	Jan	Q1	2020	02-Jan-20	Gregory	Male	M	$ 13,500	Race Car Help
13	C_CND_000012	1	Jan	Q1	2020	02-Jan-20	Amar'E	Male	M	$ 13,500	Race Car Help
14	C_CND_000013	1	Jan	Q1	2020	02-Jan-20	Griffin	Male	M	$ 8,85,000	Saab-Belle Dodge
15	C_CND_000014	1	Jan	Q1	2020	02-Jan-20	Harrison	Male	M	$ 13,500	Scrivener Performance Engineering
16	C_CND_000015	1	Jan	Q1	2020	02-Jan-20	Zainab	Male	M	$ 7,22,000	Buddy Storbeck's Diesel Service Inc
17	C_CND_000016	1	Jan	Q1	2020	02-Jan-20	Zara	Male	M	$ 7,46,000	C & M Motors Inc
18	C_CND_000017	1	Jan	Q1	2020	02-Jan-20	Zoe	Female	F	$ 5,35,000	Capitol KIA
19	C_CND_000018	1	Jan	Q1	2020	02-Jan-20	Zoey	Female	F	$ 5,70,000	Chrysler of Tri-Cities
20	C_CND_000019	1	Jan	Q1	2020	02-Jan-20	Aaliyah	Male	M	$ 6,85,000	Chrysler Plymouth
21	C_CND_000020	1	Jan	Q1	2020	02-Jan-20	Abigail	Male	M	$ 4,55,000	Classic Chevy
22	C_CND_000021	1	Jan	Q1	2020	02-Jan-20	Adrianna	Male	M	$ 13,500	Clay Johnson Auto Sales
23	C_CND_000022	1	Jan	Q1	2020	02-Jan-20	Joshua	Male	M	$ 25,00,000	Classic Chevy
24	C_CND_000023	1	Jan	Q1	2020	02-Jan-20	Marcus	Male	M	$ 5,85,000	Diehl Motor CO Inc
25	C_CND_000024	1	Jan	Q1	2020	02-Jan-20	Arthur	Male	M	$ 9,20,000	Star Enterprises Inc
26	C_CND_000025	1	Jan	Q1	2020	02-Jan-20	Lizzie	Male	M	$ 6,72,000	Suburban Ford
27	C_CND_000026	1	Jan	Q1	2020	02-Jan-20	Florian	Male	M	$ 8,01,250	Tri-State Mack Inc

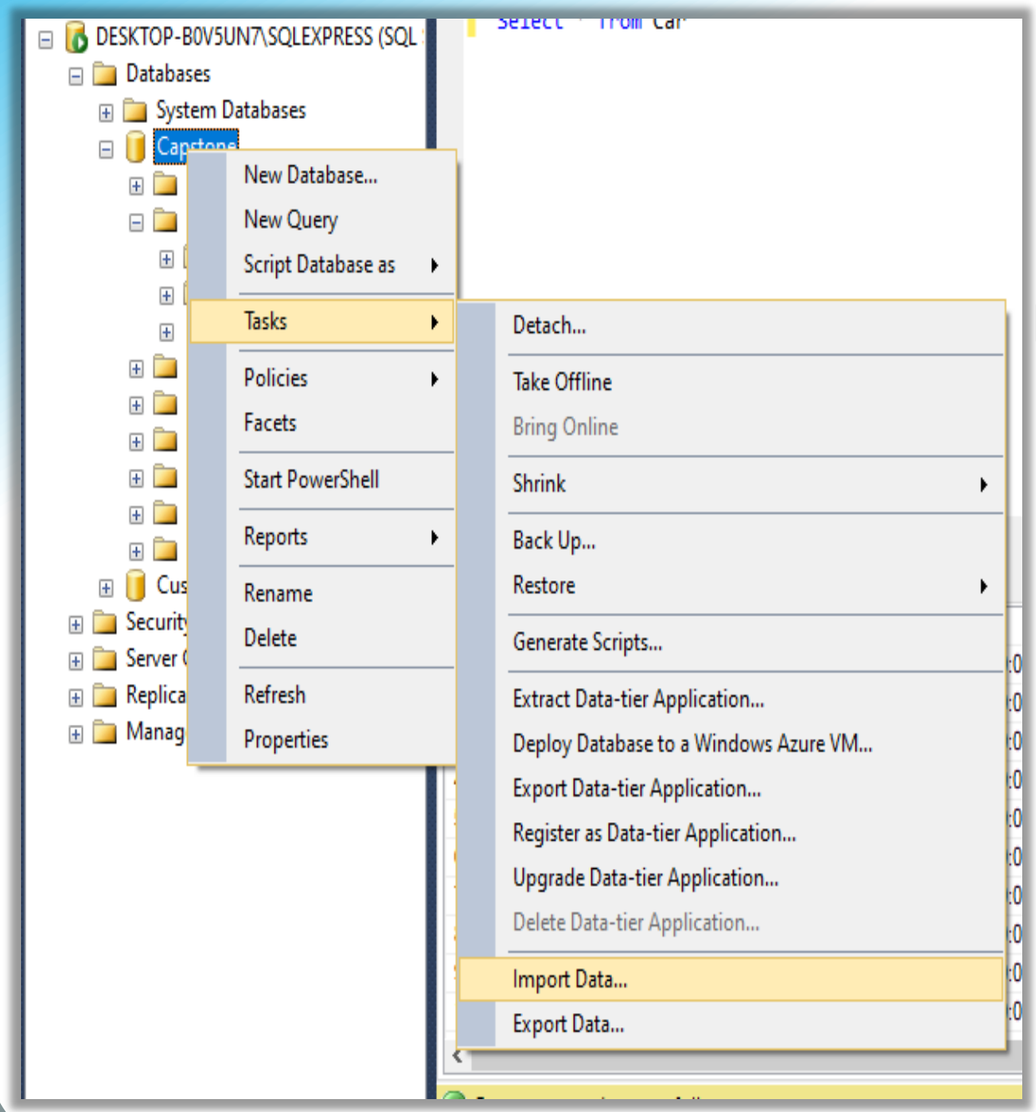
T

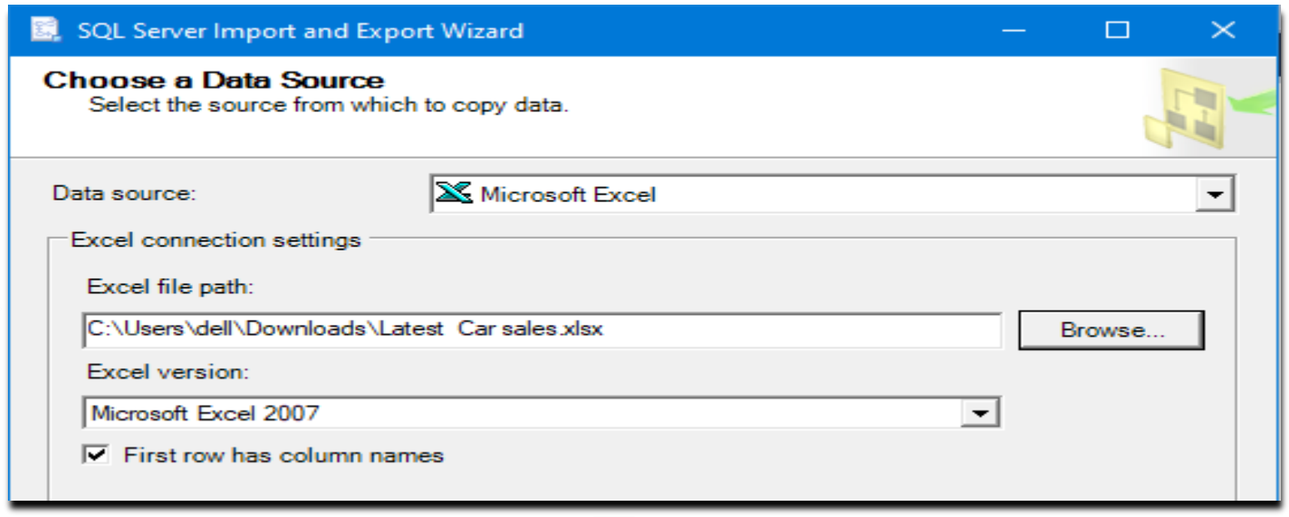
Transfrom

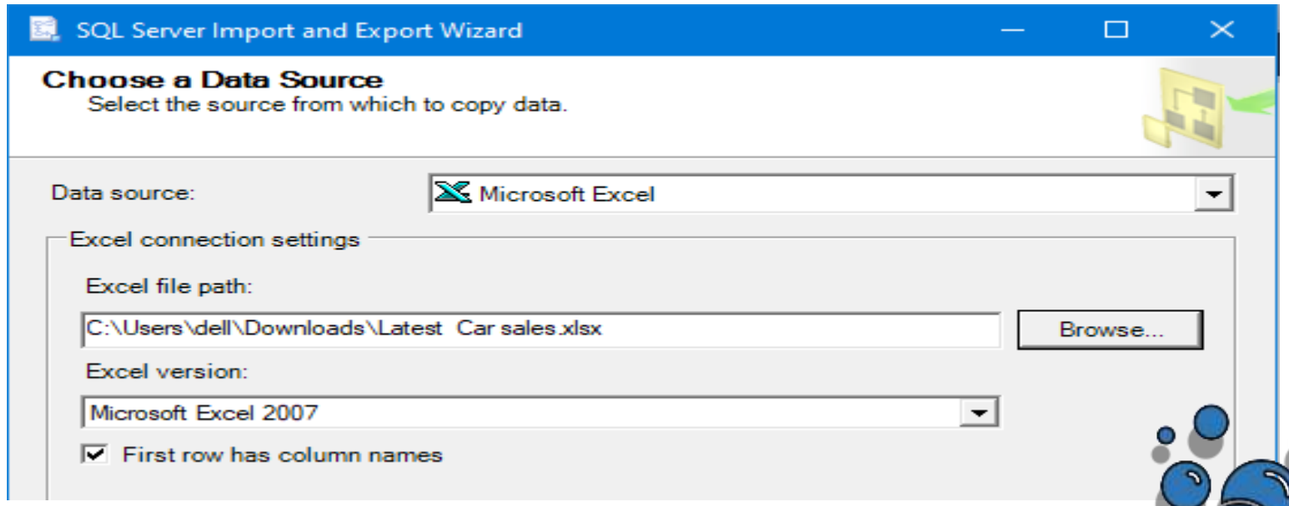
Why we want use SQL instead os Excel for transformation :-

- *In SQL Database it is easier to extract data as per our requirement.*
- *In any organisation there may be a larger number of master data files and as result maintaining a database can help.*
- *MS – Excel has a limited capacity to store upto 10 lakhs data.*
- *Hence under those circumstances where we need to deal with much larger volumes of data importing into SQL is useful.*

After cleaning or extracting the data, we have the option to insert it into an existing database or create a new one. Utilizing the Import Wizard, we can seamlessly insert our data into SQL. Here are the steps involved.

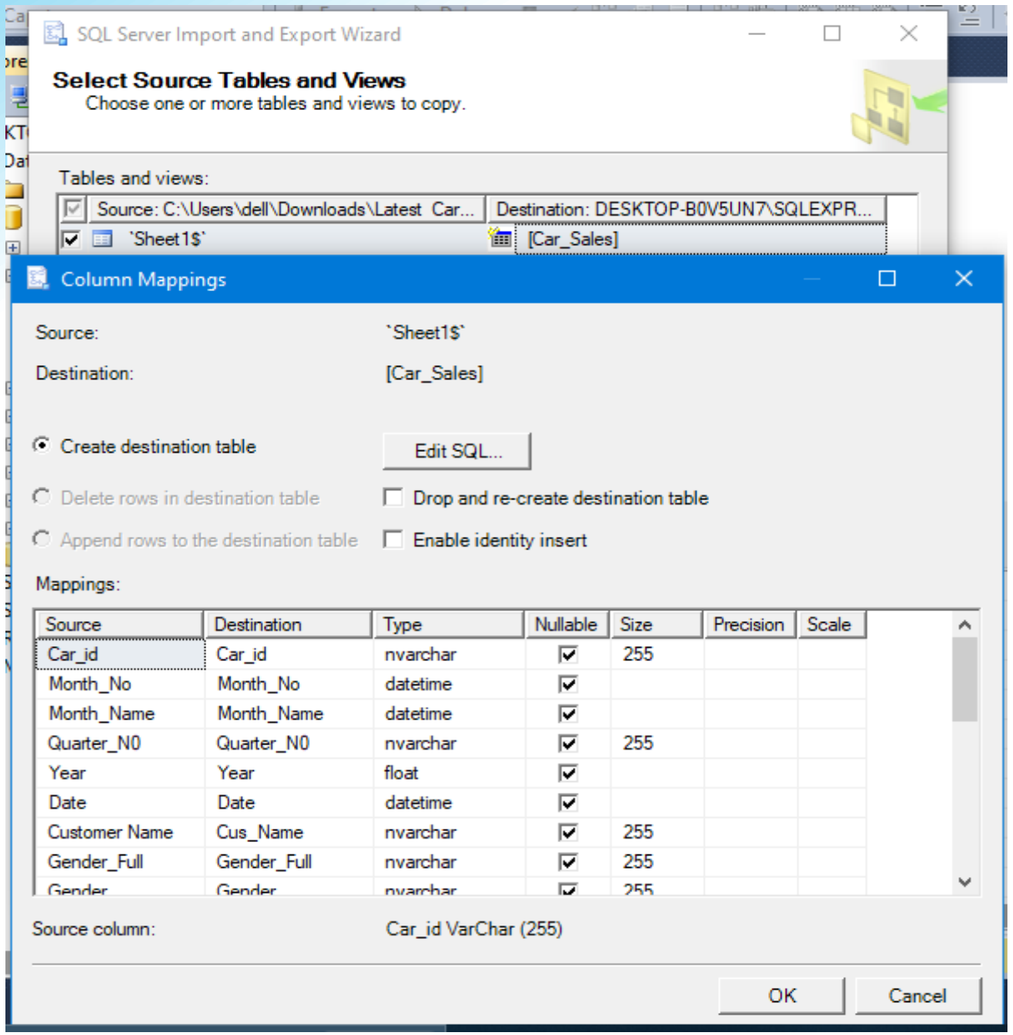


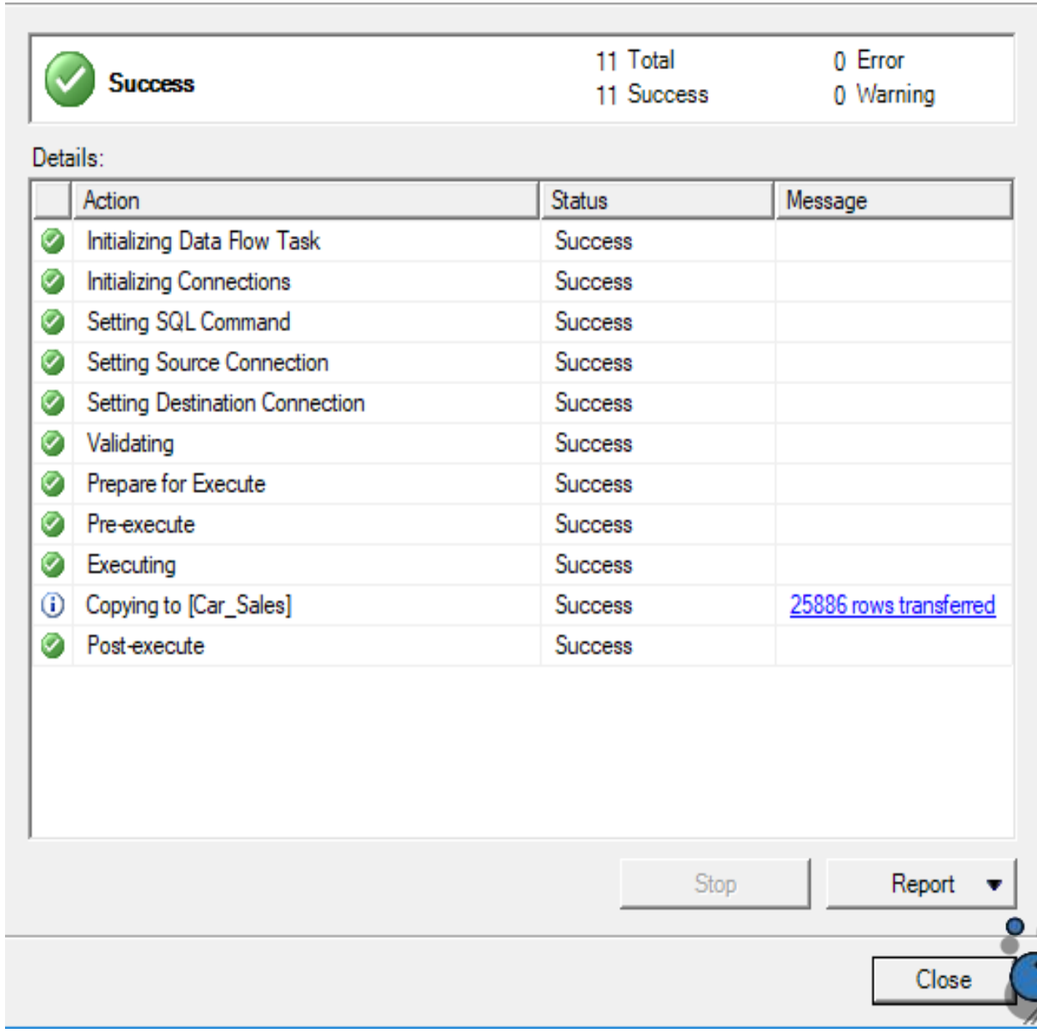




Data Edit Mapping

After selecting the SQL Client 1.0 option, we can utilize mapping to tailor our data, adjusting column names, data types, and sizes as needed. For instance, if our table is named 'Car' and the 'Customer Name' column comprises two words, SQL may display it as '[customer Name]'. To resolve this issue, we can rename the column to 'Cus_Name', ensuring seamless integration and accurate representation of the data. We can see 26000 rows are there.





To extract the total sales of cars using the aggregate function 'SUM', we obtained a total sales figure of '19862067856.00'. However, upon applying the 'FORMAT' function for improved readability, the result appeared as '19862.1M'.

```sql
-- Total sales

Select Sum(Annual_income)  As Total_sales From Car

-- Formatting Sum sales

SELECT CASE WHEN SUM(Annual_income) >= 1000000 THEN FORMAT(SUM(Annual_income) / 1000000.0, '0.#') + 'M'
        ELSE FORMAT(SUM(Annual_income) / 1000.0, '0.#') + 'K' END AS Annual_Income_Millions
From Car;
```

100 %

Results | Messages

	Total_sales
1	19862067856.00

	Annual_Income_Millions
1	19862.1M

```sql
Select CASE WHEN SUM(Annual_income) >= 1000000 THEN FORMAT(SUM(Annual_income) / 1000000.0, '0.#') + 'M'
        ELSE FORMAT(SUM(Annual_income) / 1000.0, '0.#') + 'K' END AS Annual_Income_Millions ,
    SUM(Annual_Income) over(Order by date) as Running_Total,
    DATEPART(QUARTER, Date) AS Quarters,
    DATEPART(MONTH, Date) AS Months,
    YEAR(Date) AS Year
From Car
Where YEAR(Date) = 2020
Group by YEAR(Date),DATEPART(QUARTER, Date),DATEPART(MONTH, Date),Annual_Income ,Date
Order by DATEPART(QUARTER, Date),DATEPART(MONTH, Date);
```

100 %

Results | Messages

	Annual_Income_Millions	Running_Total	Quarters	Months	Year
1	121.5K	23115750.00	1	1	2020
2	210K	23115750.00	1	1	2020
3	455K	23115750.00	1	1	2020
4	535K	23115750.00	1	1	2020
5	570K	23115750.00	1	1	2020
6	585K	23115750.00	1	1	2020
7	635K	23115750.00	1	1	2020
8	672K	23115750.00	1	1	2020
9	675K	23115750.00	1	1	2020
10	685K	23115750.00	1	1	2020
11	722K	23115750.00	1	1	2020
12	746K	23115750.00	1	1	2020
13	750K	23115750.00	1	1	2020
14	791K	23115750.00	1	1	2020

Employing DatePart with Year, Month, and Quarter enables a granular understanding of sales performance throughout various timeframes.
Running_Sales tracks cumulative sales over time, offering insights into evolving trends, while Total_Sales presents the aggregate sum of all car sales, reflecting the overall revenue.

Utilizing the TOP clause, we limit the results to the top 10 sales across all customer names, along with corresponding engine and transmission details. Additionally, the ROW_NUMBER ranking function assigns rankings based on annual income.

```sql
SELECT TOP 10 [Cus_Name], Transmission, Engine, Annual_Income, ROW_NUMBER() OVER(ORDER BY Annual_Income DESC) AS RowNumber
FROM Car
ORDER BY Annual_Income DESC;
```

100 %

Results | Messages

	Cus_Name	Transmission	Engine	Annual_Income	RowNumber
1	Assia	Auto	Doublea Overhead Camshaft	11200000.00	1
2	Sofia	Manual	Overhead Camshaft	8000000.00	2
3	Makayla	Manual	Overhead Camshaft	7650000.00	3
4	Peyton	Auto	Doublea Overhead Camshaft	6800000.00	4
5	Josh	Auto	Doublea Overhead Camshaft	6600000.00	5
6	Eileen	Manual	Overhead Camshaft	6500000.00	6
7	Kayden	Manual	Overhead Camshaft	6460000.00	7
8	Avery	Manual	Overhead Camshaft	6400000.00	8
9	Clara	Auto	Doublea Overhead Camshaft	6400000.00	9
10	Yaakov	Auto	Doublea Overhead Camshaft	6240000.00	10

The query results are displayed based on Gender, Body_Style, and Color_Name, utilizing Aggregate functions, the WHERE clause, and the GROUP BY clause for data aggregation and filtering. This approach ensures accurate and organized data presentation.

```sql
select Gender_Full , Sum(Annual_Income) as Total_sales From Car_sales
Where Gender_Full is NOt NULL
Group by Gender_full

Select Body_Style, Sum(Annual_Income) as Total_sales  from Car
Where Body_Style is Not NULL
Group by Body_Style

Select Color_Name , Sum(Annual_Income) as Total_sales  from Car
Where Color_Name is Not Null
Group by Color_Name
```

100 %

Results | Messages

	Gender_Full	Total_sales
1	Male	16000557663.00
2	Female	3861510193.00

	Body_Style	Total_sales
1	Passenger	3236302457.00
2	Hardtop	2440182530.00
3	Hatchback	5215994590.00
4	Sedan	3721443504.00
5	SUV	5248144775.00

	Color_Name	Total_sales
1	Red	3960364265.00
2	Pale White	9359040330.00
3	Black	6542663261.00

L

Load

Why Tableau ?

To effectively visualize data and extract critical insights, we aim to uncover valuable business opportunities and pinpoint target areas, facilitating informed decision-making on crucial matters.

- In our car sales dataset, focusing on dealer names and regions, we aim to pinpoint optimal locations for establishing centers, expanding our product and sales portfolio, and identifying key areas for strategic investment.
- By analyzing this data, we can develop targeted marketing strategies to maximize business growth and capitalize on emerging opportunities.

Tableau offers versatile connectivity, enabling seamless integration with various file types, including Excel files. In this scenario, we'll leverage Tableau to analyze historical data extracted from Excel, providing valuable insights based on past trends rather than real-time data.

KPI

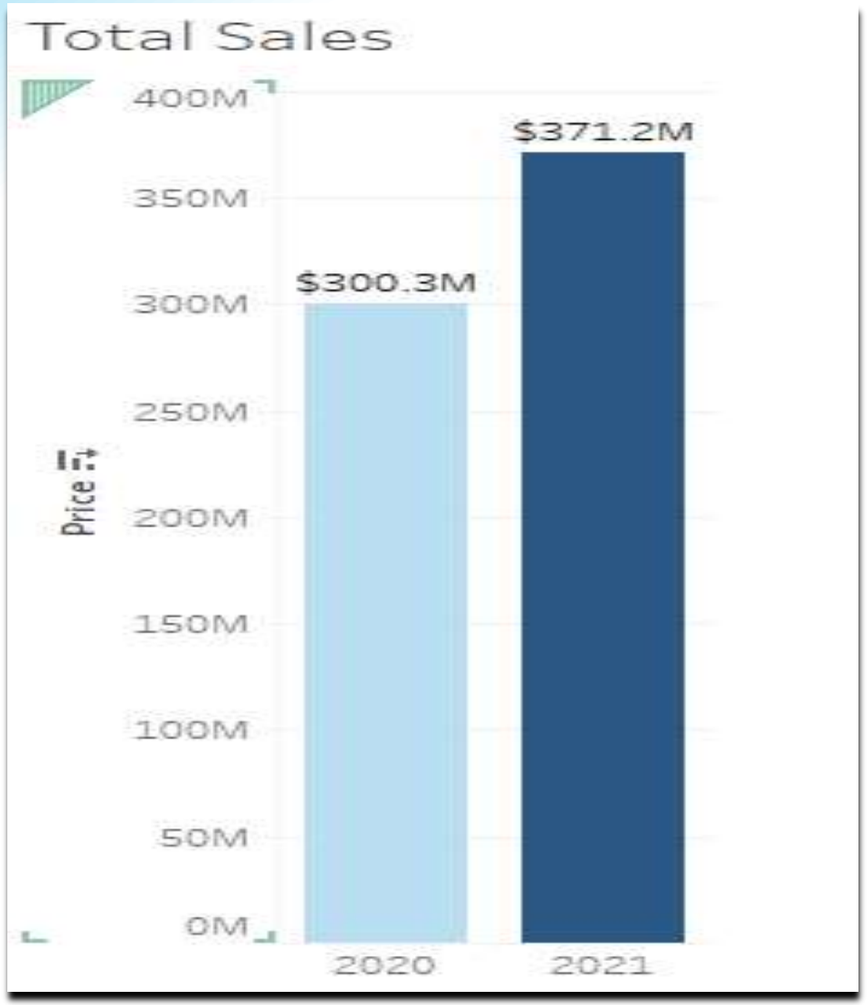
Here, we've highlighted key performance indicators (KPIs) for the entire sales period, including Total Sales, Average Car Sales, and Cars Sold.

#13.3K | YOY 24.6%

$371.2M | YOY 23.6%

28.0K | YOY -0.8%

Visualizations

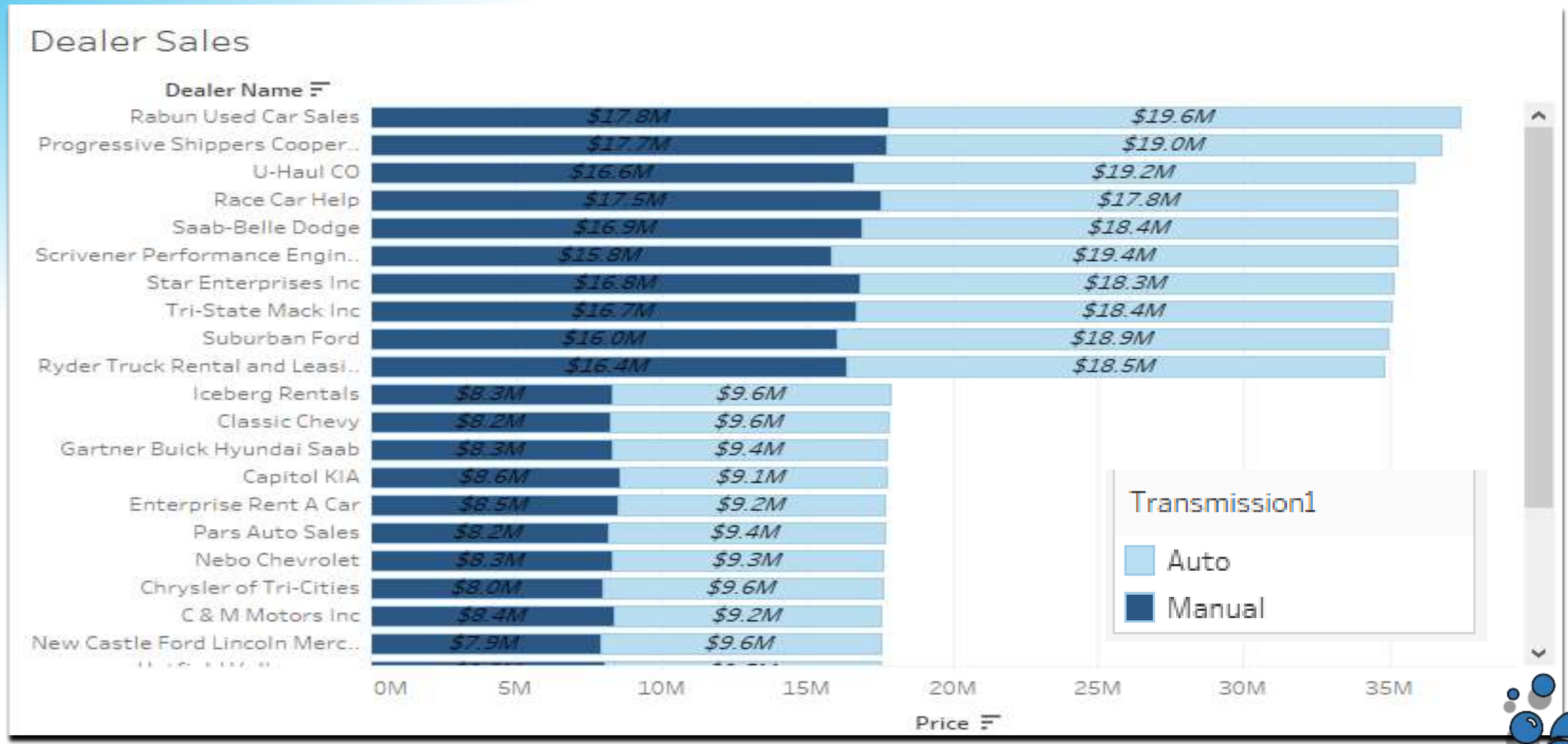


"Our dataset spans two years, encompassing data from both 2020 and 2021. Specifically, we observed sales of $300.3 million in 2020, followed by a growth to $371.2 million in 2021."

In our dataset, we observe four distinct body styles: SUV, Hardtop, Hatchback, Passenger, and Sedan. Notably, SUV recorded the highest sales of $170.6 million, followed by Hardtop with sales reaching $86.6 million over the two-year period."

In our dataset, we have data for three colors: Black, Pale White, and Red. Notably, Pale White recorded the highest sales of $309.4 million, while Red had the lowest sales at $137.7 million.

The dealer 'Rabun Used Car Sales' achieved the highest sales in both transmission types. Specifically, sales for 'Manual' transmission totaled $17.8 million, while 'Auto' transmission sales reached $19.6 million.



In Tableau, I created five Level of Detail (LOD) calculations to analyze various aspects of sales data:

1. YTD Sales: Represents total sales year-to-date.
2. PY YTD Sales: Reflects previous year's total sales.
3. YTD Avg Price: Calculates the average price of cars sold by each car company year-to-date.
4. YTD Car Sold: Indicates the total number of cars sold by each car company year-to-date.
5. YOY Growth Car Sales: percentage growth of car sales year-over-year, factoring in average price.

Company Growth

Company	YTD Sales	PY YTD Sales	YTD Avg price	YTD Car Sold	YOY Growth Car sales
Chevrolet	$27.1M	$20.5M	26.0K	1,043	34.4%
Ford	$25.4M	$21.8M	28.7K	886	21.7%
Dodge	$25.0M	$19.1M	26.4K	949	31.4%
Oldsmobile	$19.6M	$15.8M	31.6K	622	27.2%
Mercedes-B	$19.0M	$15.6M	26.7K	714	25.0%
Volkswagen	$18.2M	$15.8M	25.4K	718	16.7%
Mitsubishi	$18.7M	$15.3M	26.6K	705	23.3%
Toyota	$17.5M	$15.3M	29.5K	593	14.7%
Chrysler	$16.0M	$13.1M	25.9K	618	23.1%
Lexus	$15.2M	$12.1M	34.1K	445	24.6%
Cadillac	$15.3M	$11.4M	42.2K	363	25.6%
Mercury	$14.1M	$10.8M	29.0K	487	25.8%
Nissan	$12.7M	$11.2M	26.9K	473	14.5%
Pontiac	$13.2M	$10.1M	29.5K	448	28.7%
Volvo	$12.8M	$9.1M	27.9K	458	38.4%
Honda	$10.8M	$9.1M	27.6K	392	24.1%
BMW	$11.4M	$8.4M	25.7K	445	29.0%
Saturn	$9.9M	$8.3M	30.4K	326	25.4%
Plymouth	$9.3M	$8.8M	28.6K	327	12.8%
Acura	$9.3M	$7.7M	24.9K	376	20.1%
Lincoln	$8.6M	$6.8M	32.1K	269	20.6%
Buick	$7.8M	$6.9M	32.2K	243	24.0%
Subaru	$6.4M	$4.9M	27.7K	230	31.4%
Audi	$5.8M	$4.9M	22.2K	259	23.9%
Porsche	$4.6M	$3.6M	22.6K	203	28.5%
Saab	$4.2M	$3.5M	37.4K	111	12.1%
Jeep	$4.2M	$3.5M	20.8K	200	22.7%
Infiniti	$3.7M	$2.0M	30.5K	122	67.[illegible]%
Hyundai	$2.6M	$2.6M	19.1K	134	3.1%
Jaguar	$2.5M	$2.0M	24.5K	102	[illegible]

Conclusion

- Data analytic tools serve as invaluable assets, empowering companies to make data-driven decisions with greater precision and confidence, leveraging insights from car sales, regions, colors, customer names, and more.
- Visualizations derived from our dataset have been instrumental in pinpointing crucial target areas for our company's expansion into the market.
- Strategic investment in these identified target areas, coupled with the acquisition of appropriate resources, is imperative for the company's growth and success in the competitive market landscape.
- Utilizing data visualization to identify peak months, the company can strategically offer attractive discounts and promotions to entice potential customers, thereby boosting revenue and expanding market reach effectively.

THANK YOU