Bike Sales Analysis Dashboard

Summary

The agenda of this project is to create a dashboard in MS Excel to display bike sales analysis. The dataset contains data of 1000 people with key parameters such as Marital status, Gender, Income, Education, Occupation, Commute distance, Region, Age and if they purchased bike or not. Below is a snapshot of the dataset. The main focus is to analyze and visualize key metrics in the dashboard to help the stakeholders make informed business decisions.

ID	Marital Status	Gender	Income	Children	Education	Occupation	Home Owner	Cars	Commute Dist	Region	Age	Purchased Bike
12496	M	F	$40,000.00	1	Bachelors	Skilled Manual	Yes	0	0-1 Miles	Europe	42	No
24107	M	M	$30,000.00	3	Partial College	Clerical	Yes	1	0-1 Miles	Europe	43	No
14177	M	M	$80,000.00	5	Partial College	Professional	No	2	2-5 Miles	Europe	60	No
24381	S	M	$70,000.00	0	Bachelors	Professional	Yes	1	5-10 Miles	Pacific	41	Yes
25597	S	M	$30,000.00	0	Bachelors	Clerical	No	0	0-1 Miles	Europe	36	Yes
13507	M	F	$10,000.00	2	Partial College	Manual	Yes	0	1-2 Miles	Europe	50	No
27974	S	M	$160,000.00	2	High School	Management	Yes	4	0-1 Miles	Pacific	33	Yes
19364	M	M	$40,000.00	1	Bachelors	Skilled Manual	Yes	0	0-1 Miles	Europe	43	Yes
22155	M	M	$20,000.00	2	Partial High Sc	Clerical	Yes	2	5-10 Miles	Pacific	58	No
19280	M	M	$120,000.00	2	Partial College	Manual	Yes	1	0-1 Miles	Europe	40	Yes

Figure 1: Dataset snapshot

Analysis

Data Preparation

In the first step, we will check the data to find for missing values or Null values. We need to fill the missing values with some data in order to get efficient analysis results. For that we can either perform Imputation of missing values or remove the data related to the missing data altogether. If missing data is of low percentage , then you can remove data but if its on higher side, then it's not recommended as you need sufficient data to produce effective results. In that case, you can impute data i.e. provide a reasonable guess for missing data by means of taking average of the values or replacing with the most frequent value.

Check data for duplicates and remove duplicates. Excel has in-built feature to remove duplicates.

Check for data consistency by making sure to use proper terminology which can be understood by the end users. So check for any short forms or abbreviations and replace them with proper terms. For e.g. here – M for Male in Gender and M for Married in Marital status.

The Age column seems to be having a lot of variance ranging from 25 to 89 , so it's better to create Age brackets for clear and impactful data analysis. So we will use Nested IF formula to create different categories for different age brackets.

ID	Marital Sta	Gender	Income	Children	Education	Occupation	Home Own	Cars	Commute D	Region	Age	Age Bracke	Purchased Bike
27974	Single	Male	$160,000	2	High School	Management	Yes	4	0-1 Miles	Pacific	33	Middle aged	Yes
22155	Married	Male	$20,000	2	Partial High Sc	Clerical	Yes	2	5-10 Miles	Pacific	58	Old	No
25303	Single	Male	$30,000	0	High School	Manual	Yes	1	2-5 Miles	Europe	33	Middle aged	Yes
15608	Single	Female	$30,000	0	Partial College	Clerical	No	1	2-5 Miles	Europe	33	Middle aged	No
19441	Married	Male	$40,000	0	Graduate Deg	Clerical	Yes	0	0-1 Miles	Europe	25	Adolescent	Yes
22402	Married	Male	$10,000	0	Partial College	Manual	Yes	1	2-5 Miles	Pacific	25	Adolescent	Yes
15465	Married	Female	$10,000	0	Partial College	Manual	No	1	0-1 Miles	Pacific	25	Adolescent	No
23963	Married	Male	$10,000	0	Partial High Sc	Manual	No	2	0-1 Miles	Europe	33	Middle aged	No
28918	Married	Female	$130,000	4	High School	Management	No	4	More than 10	Europe	58	Old	No

Figure 2: After data preparation

Creating Pivot tables and Pivot charts

Once we have data prepared and validated, we can move forward for our analysis. We will select all the data and insert pivot table. Now we will select the necessary parameters to display various metrics for analysis as below:
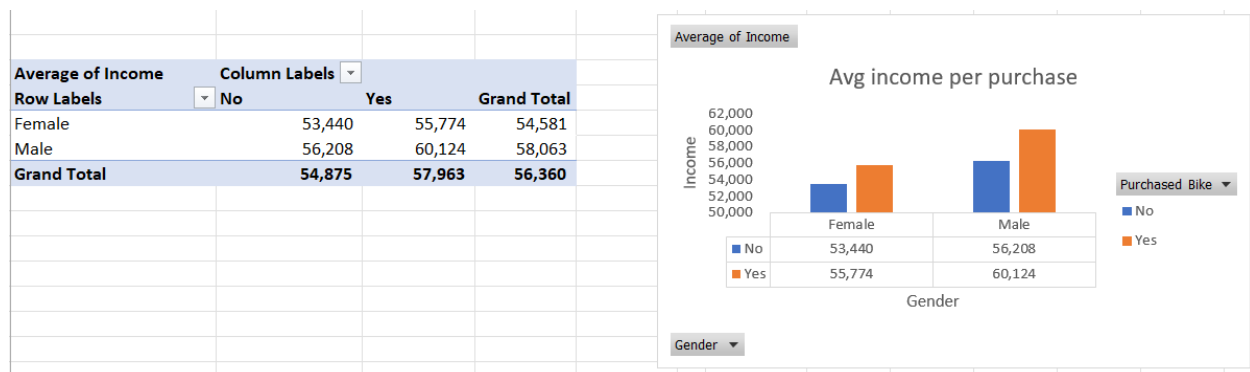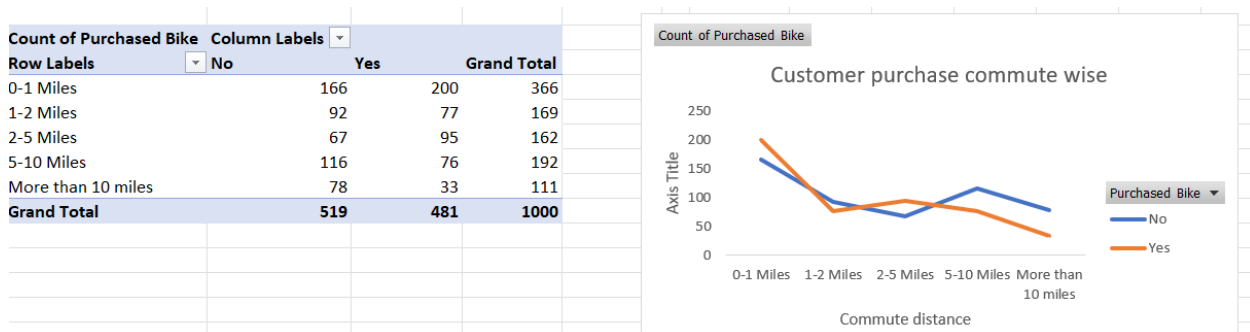

Average of Income	Column Labels		
Row Labels	No	Yes	Grand Total
Female	53,440	55,774	54,581
Male	56,208	60,124	58,063
Grand Total	54,875	57,963	56,360

Figure 3: Avg income per purchase

Count of Purchased Bike	Column Labels		
Row Labels	No	Yes	Grand Total
0-1 Miles	166	200	366
1-2 Miles	92	77	169
2-5 Miles	67	95	162
5-10 Miles	116	76	192
More than 10 miles	78	33	111
Grand Total	519	481	1000

Figure 4: Customer purchase commute wise

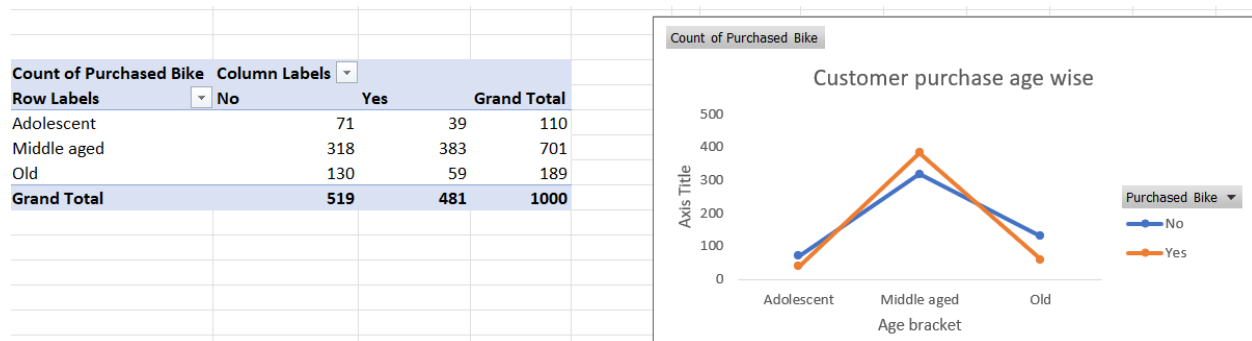| Count of Purchased Bike	Column Labels		
Row Labels	No	Yes	Grand Total
Adolescent	71	39	110
Middle aged	318	383	701
Old	130	59	189
Grand Total	519	481	1000

Figure 5: Customer purchase age wise

Applying filters using Slicer

Slicers are used as one-click filters to filter out any required information within fraction of time. We use slicers to filter out data via different regions , marital status or occupation. This helps us to gain insights on various parameters and how it affects the outcome.
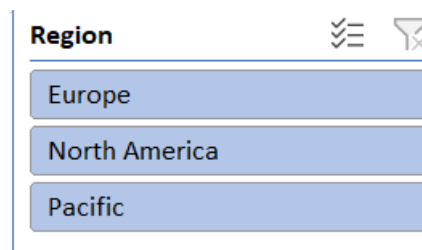

Figure 6: Region filter

Creating Final Dashboard

For the Dashboard, I will select another sheet to display all the pivot charts in one place and add slicers as well. I will also add a title for the dashboard for a better look and feel.
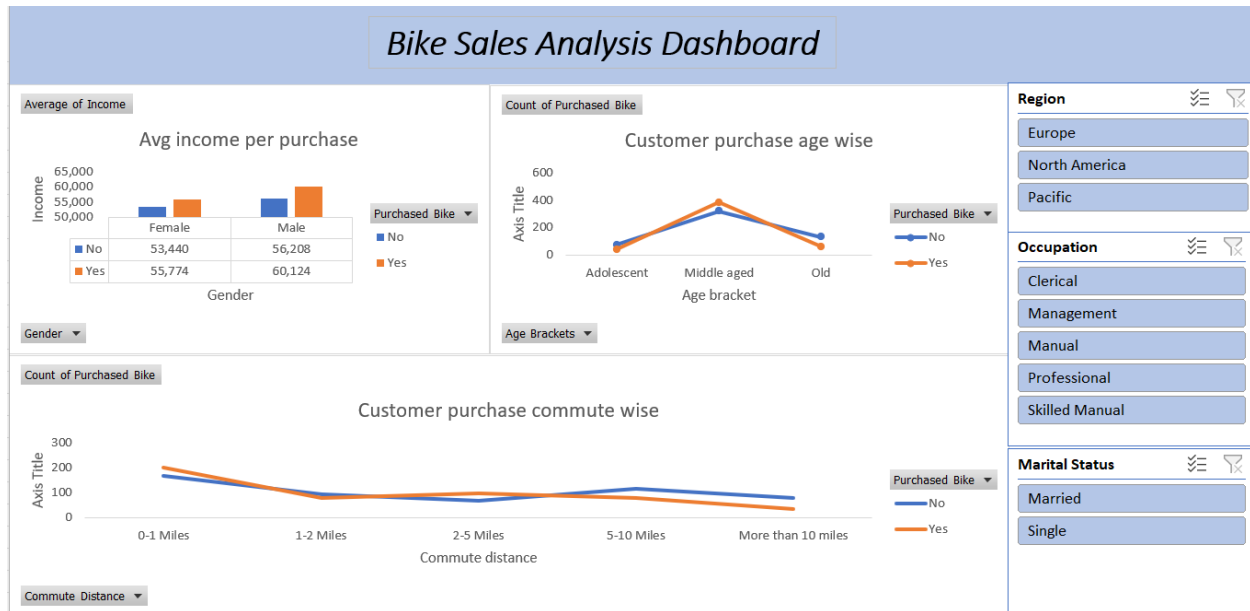

Figure 7: Final Dashboard

Conclusion

This dashboard project in Excel helped me to understand data preparation tasks such as Data Cleaning and Data Validation. Further I got good hands-on on the pivot functionality to create pivot charts for the dashboard. Using Slicers to filter data is another key learning I was able to showcase in this project. This type of dashboard can help the stakeholders with key metrics and statistical analysis for making better informed business decisions.