

Excel Project



Bike Sales Analysis

The Dataset shows Bike Sales Data which includes the following columns; ID, Marital status, Gender, Income, Children, Education, Occupation, Home Owner, Cars, Commute Distance, Region, Age and Purchased Bike.

The objective of this analysis is to explore the relationship between various factors such as Marital status, Gender, Income, Education, Occupation, and Home Owner status with bike sales. This analysis aims to identify patterns and trends that may influence sales.

performance and provide insights into target markets to help tailor marketing strategies accordingly and help make informed decision on the following basis;

1. **Market Research:** To provide valuable insights into market trends and consumer preferences. Here we examined the relationship between the average Income of buyers and Purchased Bike, to understand how price sensitivity affects purchasing decisions.
2. **Demand Forecasting:** To predict the likelihood of a customer purchasing a bike based on their demographic and socioeconomic attributes.
3. **Customer Segmentation:** By examining demographic factors such as Age, Gender, Marital status, and Region, we can segment customers into different groups based on their characteristics. This segmentation can provide insights into target markets and help tailor duly marketing strategies
4. **Sales Performance Analysis:** The relationship between various Marital status, Gender, Education, Occupation, and Home Owner status with bike sales. This analysis can help identify patterns and trends that may influence sales performance.

What our dataset looks like:

1	ID	Marital Status	Gender	Income	Children	Education	Occupation	Home Owner	Cars	Commute Distance	Region	Age	Purchased Bike
2	12496	M	F	40000	1	Bachelors	Skilled Manual	Yes	0	0-1 Miles	Europe	42	No
3	24107	M	M	30000	3	Partial College	Clerical	Yes	1	0-1 Miles	Europe	43	No
4	14177	M	M	80000	5	Partial College	Professional	No	2	2-5 Miles	Europe	60	No
5	24381	S	M	70000	0	Bachelors	Professional	Yes	1	5-10 Miles	Pacific	41	Yes
6	25597	S	M	30000	0	Bachelors	Clerical	No	0	0-1 Miles	Europe	36	Yes
7	13507	M	F	10000	2	Partial College	Manual	Yes	0	1-2 Miles	Europe	50	No
8	27974	S	M	160000	2	High School	Management	Yes	4	0-1 Miles	Pacific	33	Yes
9	19364	M	M	40000	1	Bachelors	Skilled Manual	Yes	0	0-1 Miles	Europe	43	Yes
10	22155	M	M	20000	2	Partial High School	Clerical	Yes	2	5-10 Miles	Pacific	58	No
11	19280	M	M	120000	2	Partial College	Manual	Yes	1	0-1 Miles	Europe	40	Yes
12	22173	M	F	30000	3	High School	Skilled Manual	No	2	1-2 Miles	Pacific	54	Yes
13	12697	S	F	90000	0	Bachelors	Professional	No	4	10+ Miles	Pacific	36	No
14	11434	M	M	170000	5	Partial College	Professional	Yes	0	0-1 Miles	Europe	55	No
15	25323	M	M	40000	2	Partial College	Clerical	Yes	1	1-2 Miles	Europe	35	Yes
16	23542	S	M	60000	1	Partial College	Skilled Manual	No	1	0-1 Miles	Pacific	45	Yes
17	20870	S	F	10000	2	High School	Manual	Yes	1	0-1 Miles	Europe	38	Yes
18	23316	S	M	30000	3	Partial College	Clerical	No	2	1-2 Miles	Pacific	59	Yes
19	12610	M	F	30000	1	Bachelors	Clerical	Yes	0	0-1 Miles	Europe	47	No
20	27183	S	M	40000	2	Partial College	Clerical	Yes	1	1-2 Miles	Europe	35	Yes
21	25940	S	M	20000	2	Partial High School	Clerical	Yes	2	5-10 Miles	Pacific	55	Yes
22	25598	M	F	40000	0	Graduate Degree	Clerical	Yes	0	0-1 Miles	Europe	36	Yes
23	21564	S	F	80000	0	Bachelors	Professional	Yes	4	10+ Miles	Pacific	35	No
24	19193	S	M	40000	2	Partial College	Clerical	Yes	0	1-2 Miles	Europe	35	Yes
25	26412	M	F	80000	5	High School	Management	No	3	5-10 Miles	Europe	56	No
26	27184	S	M	40000	2	Partial College	Clerical	No	1	0-1 Miles	Europe	34	No
27	12590	S	M	30000	1	Bachelors	Clerical	Yes	0	0-1 Miles	Europe	63	No
28	17841	S	M	30000	0	Partial College	Clerical	No	1	0-1 Miles	Europe	29	Yes
29	18283	S	F	100000	0	Bachelors	Professional	No	1	5-10 Miles	Pacific	40	No
30	18299	M	M	70000	5	Partial College	Skilled Manual	Yes	2	5-10 Miles	Pacific	44	No

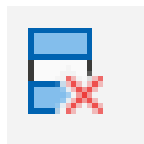
The following steps were taken for this analysis;

Data Cleaning

The dataset comprises of 14 columns (listed above) and 1027 rows.

- Remove Duplicate - Duplicate data can cause inconsistencies in the dataset, making it difficult to maintain data integrity. Removing duplicates allows one to focus on unique data points, enabling more accurate and meaningful analysis, which leads to better insights and decision-making.

This was achieved by Highlighting the data table, Select the “Data” tab, click on Remove Duplicates



This automatically checks the table for duplicate information and get rid of them. Here we had 26 duplicates which were successfully removed leaving us with 1001 rows of Data.

- Filter - Filtering allows focusing on specific information in a large dataset or table. By filtering data, one can quickly find values and control what to see and what not to see, excluding irrelevant (at the time) entries from view. Filtering can help identify and remove duplicate or redundant data entries such as blanks.

Here we used Filter to remove Blanks; Highlight the Table, Select the “Data” tab, click on “Filter”



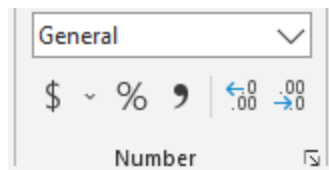
The above image appeared on all the column headers, click on “Marital status” column (any column header can be selected but it is advised to choose one with few unique values, makes it easier because blanks are usually the last option.), deselect all and select only blanks. Only blank cells will be viewed, select and delete all. Go back to the filter arrow, select all. We had one blank row and one blank column, which were successfully deleted, leaving us with 13 columns and 1000 rows of data.

- Find and Replace - This helps to find whatever value you're looking for and replace it with your chosen new value.

The Marital Status and Gender Columns contain two unique single values each, which the both column has one in common, “M”, “S”, and “M”, “F”, respectively. These letters were found and replaced with their full form for easier analysis of the data and for better understanding after the analysis. Select the “Filter & Select” command which is found on the “Home” tab, but first you have to select the column you wish to find the value in, in this case “Marital Status”- click on “Replace”, add the value you wish to find on the “Find” space and the new value on the “Replace” space. Just below them are three options; “Within” - select “Sheet”, “Search” - select “By column”, ignore the “Look in” option, click on “Replace All”. “M” was replaced with “Married”, “S” with “Single”.

Same steps were applied to the “Gender” column but “M” was replaced with “Male”, “F” with “Female”.

- Format Change - The Income column was changed to currency and decimals on the income values as a result of the format were also removed.



Click on “General” on the “Home” tab, select currency, click on the “.00 with arrow pointing to your right” to remove the decimals.

- Creating another Column named “Age Brackets” - The “Age” column contains lots of variables which will make it difficult for data exploration and visualization. This was done by grouping the ages into ranges using “Nested IFS” formula.

```
=IF(L2>54,"Old",IF(L2>=31,"Middle age",IF(L2<31,"Adolescent")))
```

“L2” is the cell name of the first row in column “Age”.

First off Select the column after Age, right click on it, select “insert”. This creates a new column, the column was named “Age Brackets”. The above formula was built on the first row of the new column, then flash fill for the rest of the column.

Flash fill was done by clicking on the cell, take your cursor to the button right of the cell, a black + appears, double click on it.

1	ID	Marital status	Gender	Income	Children	Education	Occupation	Home Owner	Cars	Commute Distance	Region	Age	Age Brackets	Purchased Bike
2	12496	Married	Female	\$40,000	1	Bachelors	Skilled Manual	Yes	0	0-1 Miles	Europe	42	Middle age	No
3	24107	Married	Male	\$30,000	3	Partial College	Clerical	Yes	1	0-1 Miles	Europe	43	Middle age	No
4	14177	Married	Male	\$80,000	5	Partial College	Professional	No	2	2-5 Miles	Europe	60	Old	No
5	24381	Single	Male	\$70,000	0	Bachelors	Professional	Yes	1	5-10 Miles	Pacific	41	Middle age	Yes
6	25597	Single	Male	\$30,000	0	Bachelors	Clerical	No	0	0-1 Miles	Europe	36	Middle age	Yes
7	13507	Married	Female	\$10,000	2	Partial College	Manual	Yes	0	1-2 Miles	Europe	50	Middle age	No
8	27974	Single	Male	\$160,000	2	High School	Management	Yes	4	0-1 Miles	Pacific	33	Middle age	Yes
9	19364	Married	Male	\$40,000	1	Bachelors	Skilled Manual	Yes	0	0-1 Miles	Europe	43	Middle age	Yes
10	22155	Married	Male	\$20,000	2	Partial High School	Clerical	Yes	2	5-10 Miles	Pacific	58	Old	No
11	19280	Married	Male	\$120,000	2	Partial College	Manual	Yes	1	0-1 Miles	Europe	40	Middle age	Yes
12	22173	Married	Female	\$30,000	3	High School	Skilled Manual	No	2	1-2 Miles	Pacific	54	Middle age	Yes
13	12697	Single	Female	\$90,000	0	Bachelors	Professional	No	4	10+ Miles	Pacific	36	Middle age	No
14	11434	Married	Male	\$170,000	5	Partial College	Professional	Yes	0	0-1 Miles	Europe	55	Old	No
15	25323	Married	Male	\$40,000	2	Partial College	Clerical	Yes	1	1-2 Miles	Europe	35	Middle age	Yes
16	23542	Single	Male	\$60,000	1	Partial College	Skilled Manual	No	1	0-1 Miles	Pacific	45	Middle age	Yes
17	20870	Single	Female	\$10,000	2	High School	Manual	Yes	1	0-1 Miles	Europe	38	Middle age	Yes
18	23316	Single	Male	\$30,000	3	Partial College	Clerical	No	2	1-2 Miles	Pacific	59	Old	Yes
19	12610	Married	Female	\$30,000	1	Bachelors	Clerical	Yes	0	0-1 Miles	Europe	47	Middle age	No
20	27183	Single	Male	\$40,000	2	Partial College	Clerical	Yes	1	1-2 Miles	Europe	35	Middle age	Yes
21	25940	Single	Male	\$20,000	2	Partial High School	Clerical	Yes	2	5-10 Miles	Pacific	55	Old	Yes
22	25598	Married	Female	\$40,000	0	Graduate Degree	Clerical	Yes	0	0-1 Miles	Europe	36	Middle age	Yes
23	21564	Single	Female	\$80,000	0	Bachelors	Professional	Yes	4	10+ Miles	Pacific	35	Middle age	No
24	19193	Single	Male	\$40,000	2	Partial College	Clerical	Yes	0	1-2 Miles	Europe	35	Middle age	Yes
25	26412	Married	Female	\$80,000	5	High School	Management	No	3	5-10 Miles	Europe	56	Old	No
26	27184	Single	Male	\$40,000	2	Partial College	Clerical	No	1	0-1 Miles	Europe	34	Middle age	No
27	12590	Single	Male	\$30,000	1	Bachelors	Clerical	Yes	0	0-1 Miles	Europe	63	Old	No
28	17841	Single	Male	\$30,000	0	Partial College	Clerical	No	1	0-1 Miles	Europe	29	Adolescent	Yes
29	18283	Single	Female	\$100,000	0	Bachelors	Professional	No	1	5-10 Miles	Pacific	40	Middle age	No
30	18299	Married	Male	\$70,000	5	Partial College	Skilled Manual	Yes	2	5-10 Miles	Pacific	44	Middle age	No

Our Data after cleaning.

Create a Pivot Table

Pivot Table is a great tool for summarizing and analysing data in Excel. Can be used to perform calculations on our data based on certain criteria.

- Open a new spreadsheet, rename it Pivot Table
- Click on the “Insert” tab and select Pivot Table
- To select the Table/Range, open the spreadsheet that contains your data and select the Table
- After creating a pivot table, all the column headers appears on the pivot table fields. Here we can drag the different fields to either of the following areas; Column, Row, Value or Filter depending on your findings.

Findings;

1. Average Income of Males and Females who purchased and did not purchase bikes:

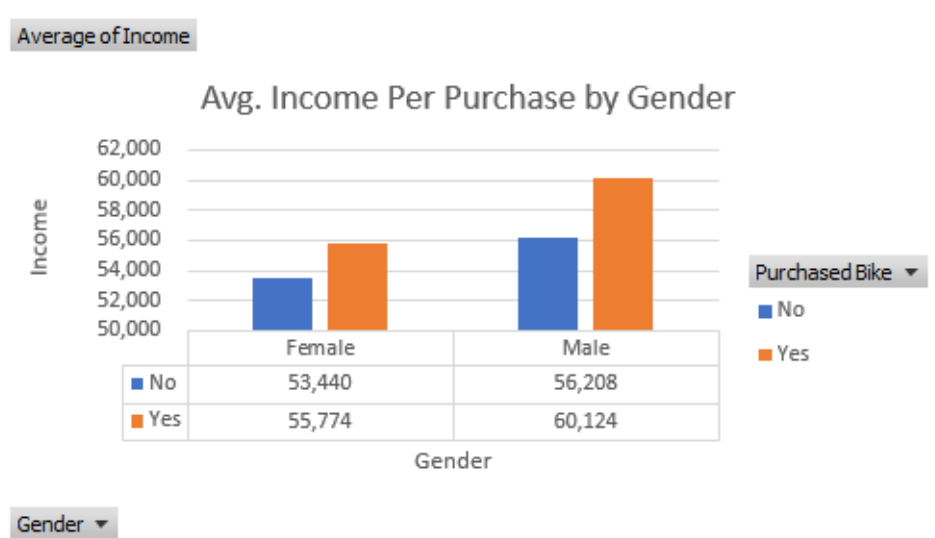
To check if individuals income has any effect on their capability to purchase a bike, if there's any income change whether or not they purchase a bike and also to compare income earned by the two gender.

- The Income was dragged to the Values field, Gender to the Rows and Purchased Bike to the column, Here's what our Pivot Table looks like:

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

- The Default calculation for Values is sum, to get the Average, click on the value field which is "Sum of Income", select "Value field setting", Select Average. Remove the decimals as we explained above.
- Create a visualization of the pivot table using a column chart; Click on Insert tab, Click on Recommended charts or PivotChart, Select Clustered column chart.

Arrange the chart, add necessary information (such as the Axis title, chart title) that will make it more understandable using the Chart Element (appears like a cross in a box when you click on the chart).



2. . Commute Distance vs Bike purchased: The effect of commute distance on bike purchase;

- Go through the same pivot table creation as above but this time, you're creating inside an existing table. In the same spreadsheet, click on an empty cell, click on

the Insert tab then Pivot Table, Select the table with your data and click Ok.

- The Commute Distance was dragged to the Rows, Purchased Bike to the Column, Purchased Bike also to the Value field. The Purchased Bike in the Value Field becomes the Count of Purchased Bike because Values placed in the “Value Field” comes out as numbers.
- Find and Replace 10+ Miles to More than 10 Miles. This is because the + sign on the number placed it directly after 0-1 Mile whereas it's supposed to be the last value in the Row.

This was done by going back to the Data Source sheet, click on the column, CTRL + H to bring out the Find & Replace Command, enter the values as it should be, click on Replace All. Go back to the Pivot Table sheet, go to Data Tab, Click on Refresh All.

Count of Purchased 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

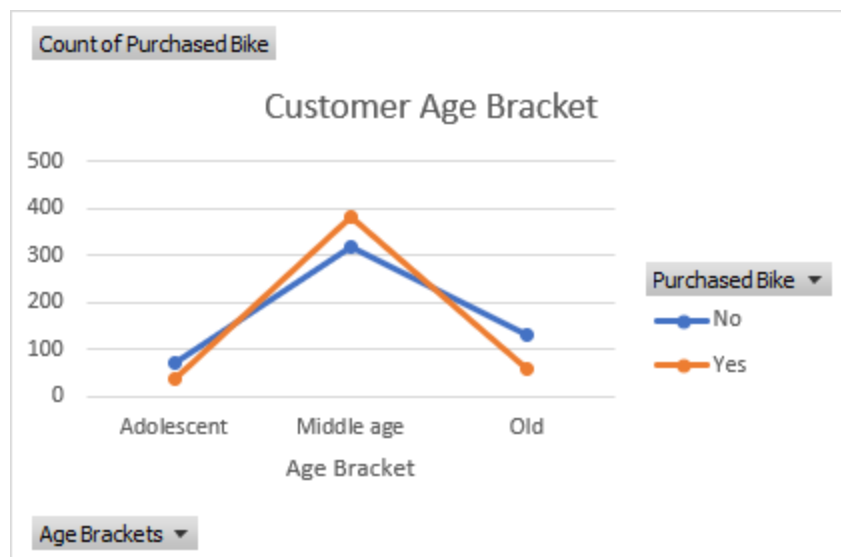
- Create a visual representation using a Line Chart



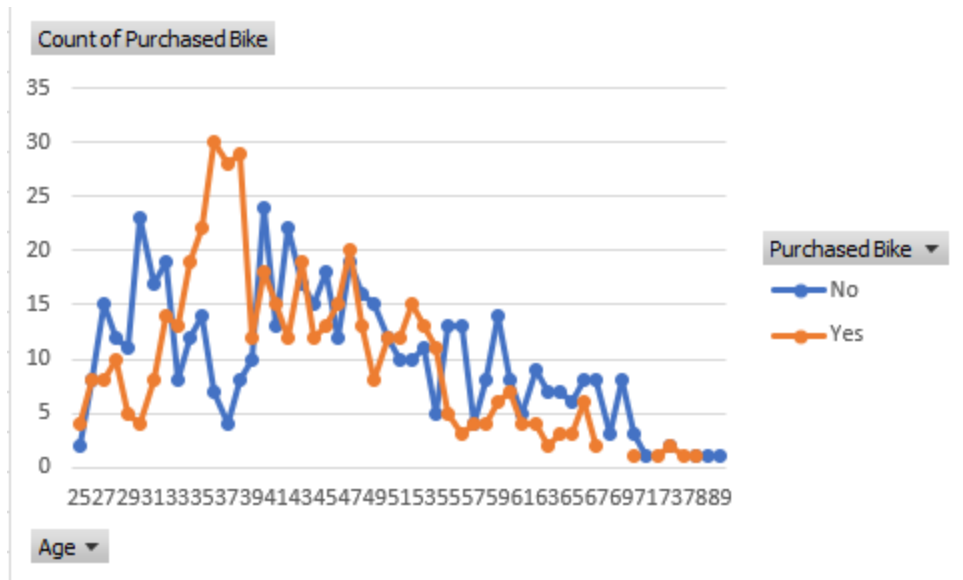
- Modify your Chart using the Chart Elements.
3. Bike purchased per age bracket: To analyse which Age Bracket bought the most bikes.
- A Pivot Table was created in the existing sheet using the same Data Source.
 - Age Brackets was dragged to the Rows Field, Purchased Bike to the Column and Value Fields.

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

- Visualize the table using Line with Makers.



Creating a Visual Representation of the above using the Age Column instead of Age Brackets Column;

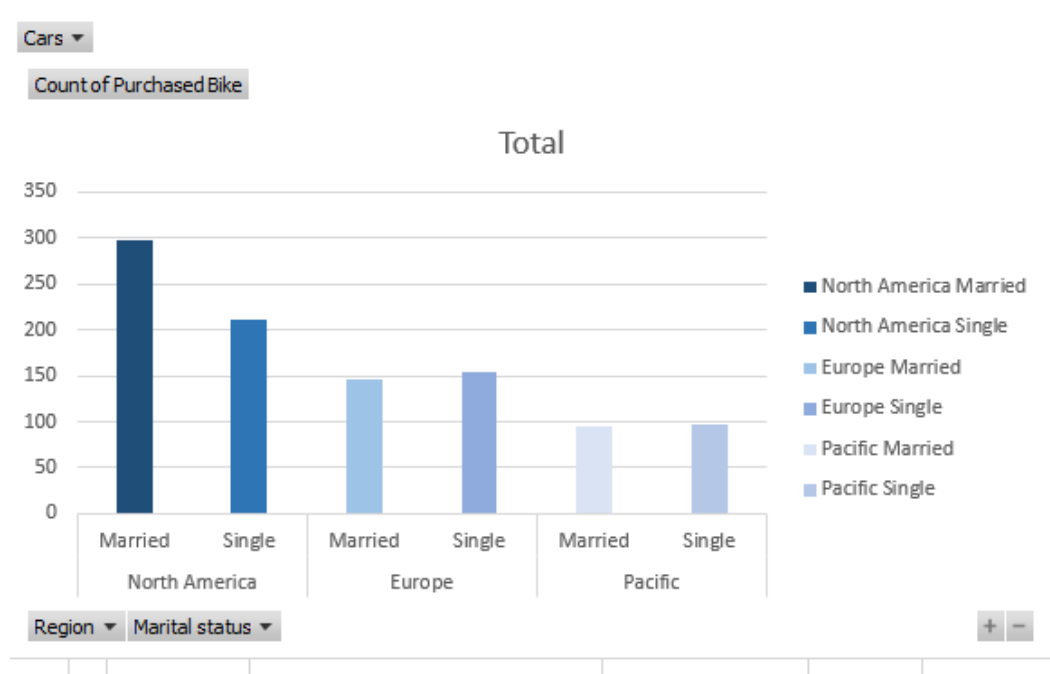


Oops that didn't go well.

- Count of Purchased Bikes by Region and Marital Status, The data is further filtered by Cars;

Filters	Columns
Cars	
Rows	Values
Region	Count of Purchased Bi...
Marital status	

Cars	(All)
Row Labels	Count of Purchased Bike
North America	508
Married	297
Single	211
Europe	300
Married	146
Single	154
Pacific	192
Married	95
Single	97
Grand Total	1000



Creating a Dashboard

The main objective of a data analysis project is to develop dashboards that are visually attractive, easy to comprehend, and interactive. To enhance the interactivity and dynamism of the dashboards, we can incorporate slicers that allow us to filter the results according to our requirements.

Steps;

- Open a New Spreadsheet
- Select “View” tab
- Untick the “Gridlines”
- Make a Dashboard header either by selecting multiple cells and click “Merge & centre” in the “Home” tab or click on the “Insert” tab, select “Text box” from the “Text” command.
- Use your cursor to create the amount of space you need for the header
- To colour the header, go to “Home” tab, select “colour fill”, and choose any colour of choice.
- Name the header, change the colour of the text by selecting any colour of choice from “Font colour” on the “Home” tab.

- Copy and Paste your Pivot Charts
- Align them with your cursor or click on the chart for the “Format” tab to appear, select any of the “Align” option depending on which side you want your chart to align to.

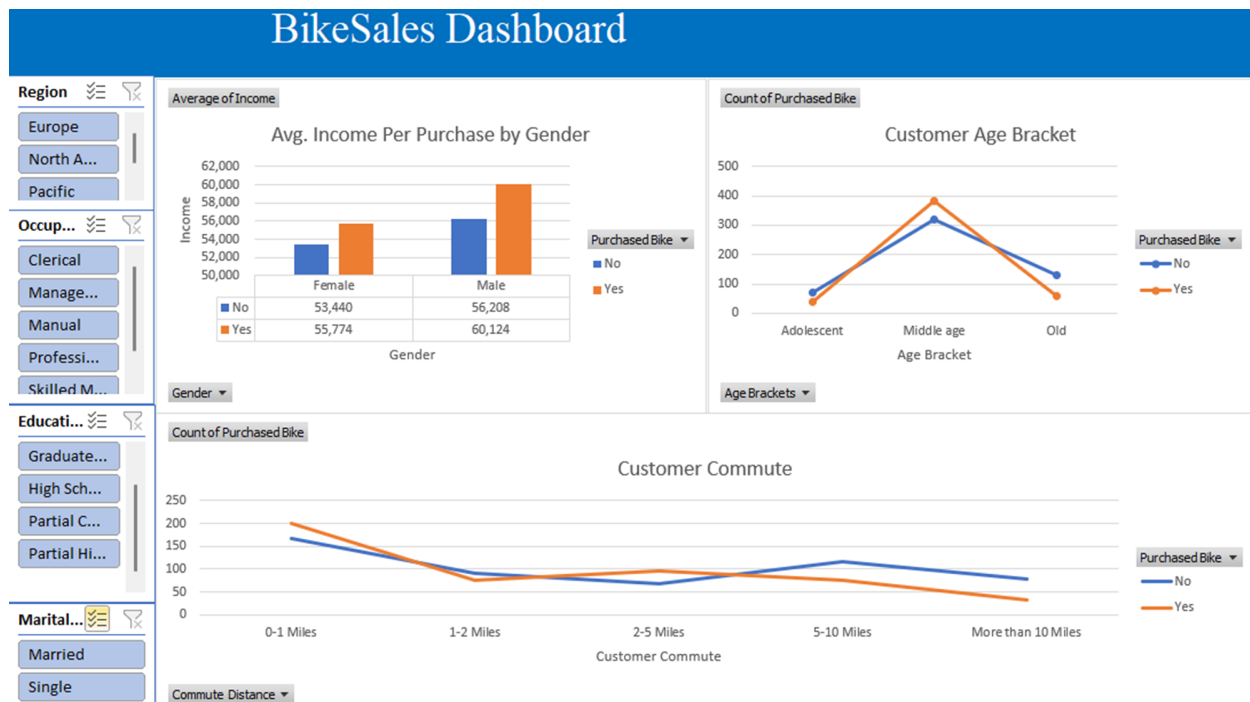
Inserting Slicers

Slicers are a visual filter that can be added to pivot tables and pivot charts in Excel. They allow you to filter data quickly and easily by clicking on a button, making it simple to understand what data is currently displayed. Slicers are especially useful for creating interactive dashboards that allow users to filter data according to their needs. By adding slicers to pivot charts, you can make your dashboards more dynamic and interactive, allowing users to filter data based on their requirements.

- Click on the “PivotChart Analyze” Tab, Click on “Insert Slicers” and select any of the column heads you wish to filter based on.
- Select as much as you want but be careful not to cluster your Dashboard
- Using your cursor, arrange the slicers by the side of your charts.

To connect each slicer to all the Pivot Charts;

- Click on the slicer which causes a “Slicer” tab to appear, click on it then on “Report connections”. Tick all the Pivot Table listed.



Conclusion

From the above analysed Data;

The **Table 1** shows that the Average Income of Individuals in both gender who purchased bikes is higher than those who did not. Male has higher Average Income than the Females and they bought more bikes than the Female.

In both Gender, the difference in the Average income of the “Yes” and “No” category is less than 10% of the Grand Total.

Table 2 - Commute Distance vs Bike Purchased:

Individuals that commute within 0-1 Miles has the highest purchase followed by 2-5Miles while those that commuted more than 10miles away has the least purchase. 0-1Miles category also has the highest number of no purchase followed by 5-10Miles.

Table 3 - Bike Purchased Per Age Bracket:

The Middle age group (31 -54years) bought more bikes followed by the Old group (55 years and above) with a wide range of over 200+ purchases. The Old and Adolescent (30 and below) fall in the “No” category than they do in the “Yes”.

Table 4 - Count of Purchase by Region and Marital Status, filtered by Cars:

More purchases were made in North America, followed by Europe then Pacific. Married folks made more purchase than the Single overall, though in Europe and Pacific, Single Folks made more purchases by a few than the Married while in North America, the reverse was the case.

Individuals with 2 cars made the highest purchase followed by 1 car then No car on a close range. People with 4 cars made the least purchase.

The Males, People that commune within 0-1Mile, the Middle age, the Married and North America made more purchases than their cohorts, though the differences in some wasn't much.

Untitled