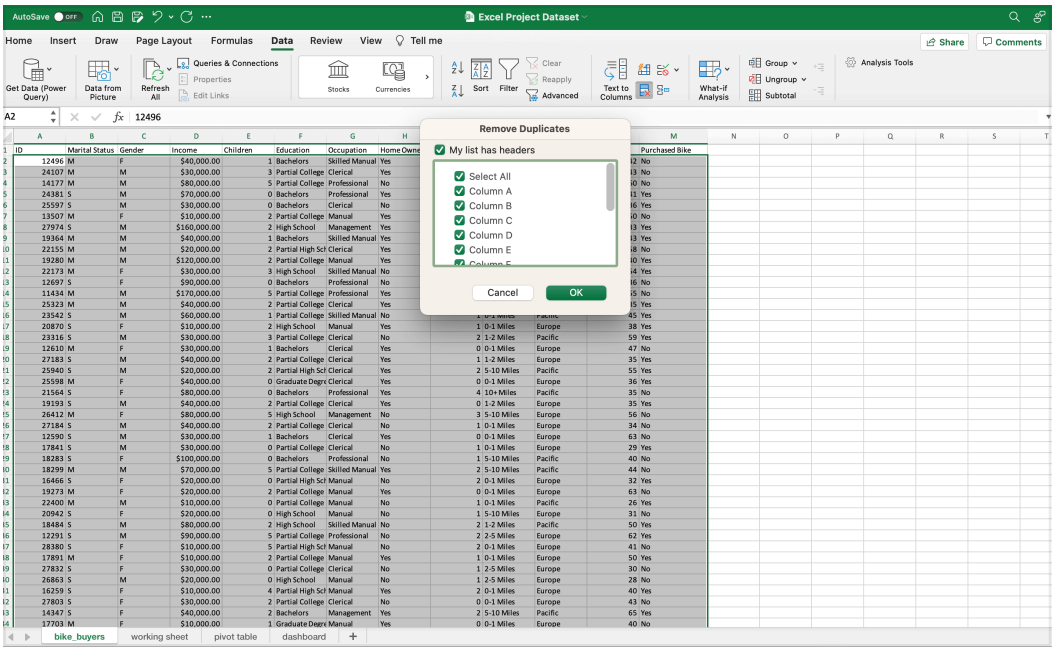
Project Log

Project: "Demographic Analysis of Bike Consumers: Understanding Buyer Profiles”

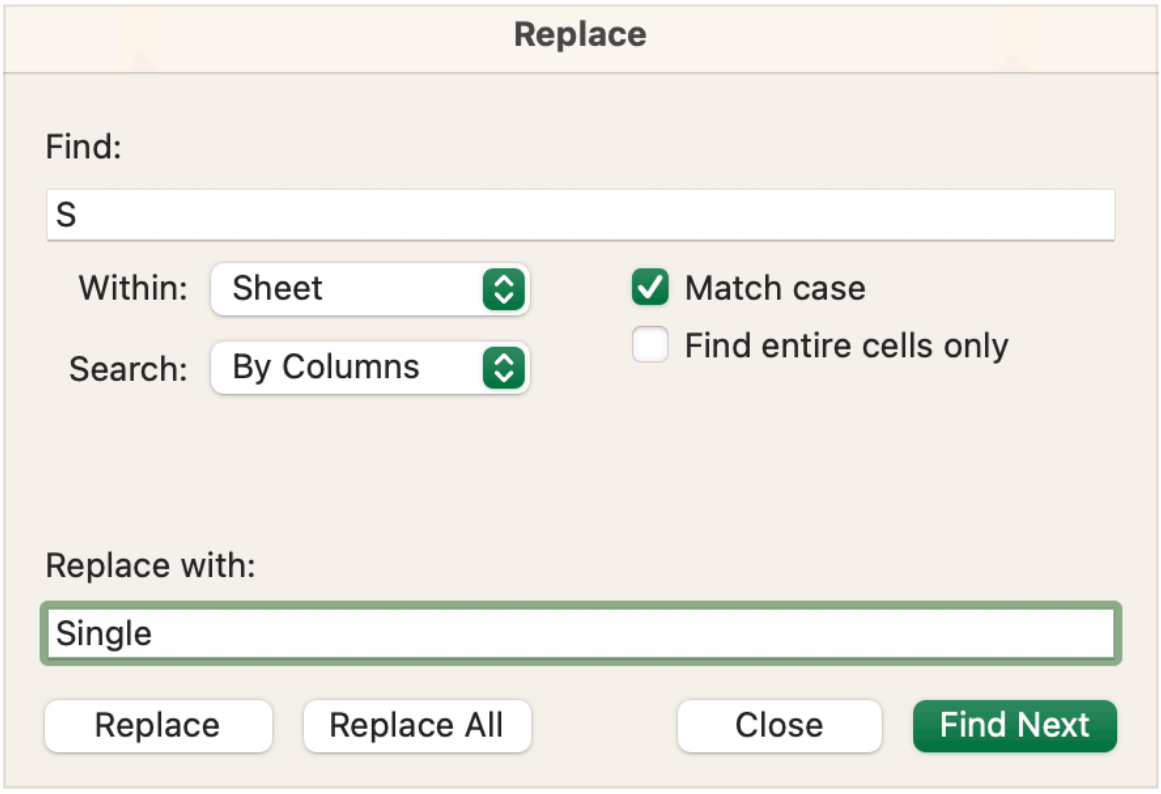
- First start of by creating a working sheet tab where I will do all the cleaning - that way I have a copy of the original raw data (I also created a pivot table and a dashboard tab).

bike_buyers	working sheet

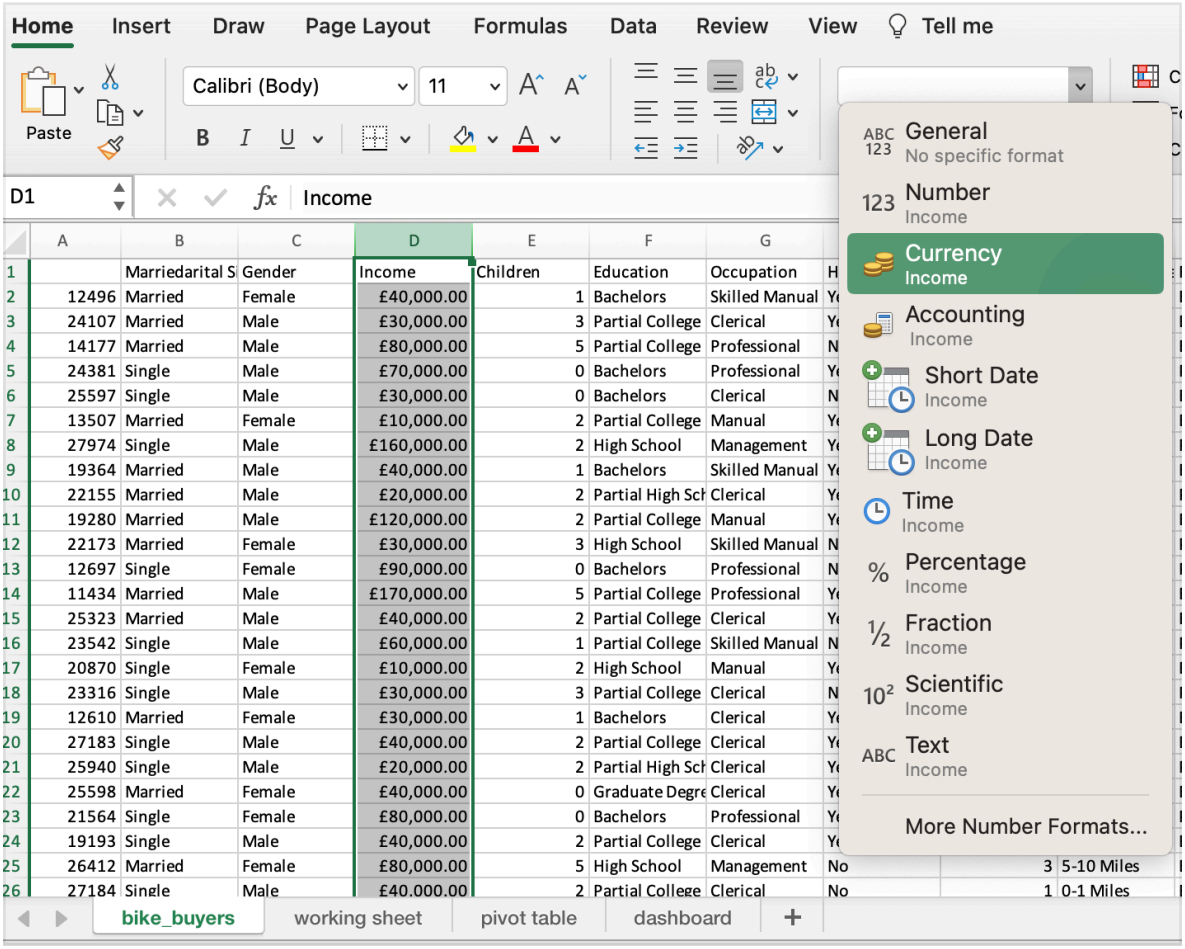
- Removing duplicates



- Using find and replace (control + H) to replace acronym’s with their actual word. This makes it easier for those using the dashboard. 2 columns changed are the gender and marital columns.



- Confirm all of it IS currency int he income column



	A	B	C	D	E	F	G
1		Marriedarital S	Gender	Income	Children	Education	Occupation
2	12496	Married	Female	£40,000.00	1	Bachelors	Skilled Manual
3	24107	Married	Male	£30,000.00	3	Partial College	Clerical
4	14177	Married	Male	£80,000.00	5	Partial College	Professional
5	24381	Single	Male	£70,000.00	0	Bachelors	Professional
6	25597	Single	Male	£30,000.00	0	Bachelors	Clerical
7	13507	Married	Female	£10,000.00	2	Partial College	Manual
8	27974	Single	Male	£160,000.00	2	High School	Management
9	19364	Married	Male	£40,000.00	1	Bachelors	Skilled Manual
10	22155	Married	Male	£20,000.00	2	Partial High Sch	Clerical
11	19280	Married	Male	£120,000.00	2	Partial College	Manual
12	22173	Married	Female	£30,000.00	3	High School	Skilled Manual
13	12697	Single	Female	£90,000.00	0	Bachelors	Professional
14	11434	Married	Male	£170,000.00	5	Partial College	Professional
15	25323	Married	Male	£40,000.00	2	Partial College	Clerical
16	23542	Single	Male	£60,000.00	1	Partial College	Skilled Manual
17	20870	Single	Female	£10,000.00	2	High School	Manual
18	23316	Single	Male	£30,000.00	3	Partial College	Clerical
19	12610	Married	Female	£30,000.00	1	Bachelors	Clerical
20	27183	Single	Male	£40,000.00	2	Partial College	Clerical
21	25940	Single	Male	£20,000.00	2	Partial High Sch	Clerical
22	25598	Married	Female	£40,000.00	0	Graduate Degre	Clerical
23	21564	Single	Female	£80,000.00	0	Bachelors	Professional
24	19193	Single	Male	£40,000.00	2	Partial College	Clerical
25	26412	Married	Female	£80,000.00	5	High School	Management
26	27184	Single	Male	£40,000.00	2	Partial College	Clerical

- Commute distance is giving us ranges. I will keep it for now but may

change it later for better visualisations.

J
Commute Dista
0-1 Miles
0-1 Miles
2-5 Miles
5-10 Miles
0-1 Miles
1-2 Miles

- Currently we have ages as individuals but I will put them in brackets or ranges in a new column because it will later help me keep the visualisation looking cleaner and sharper and easier to digest.

- Use IF statement to create the age ranges - our query will be, if they are less than the age of 31, then they are adolescent, and if not, I will brand it as invalid.

```
=IF(L2<31, "adolescent", "invalid")
```

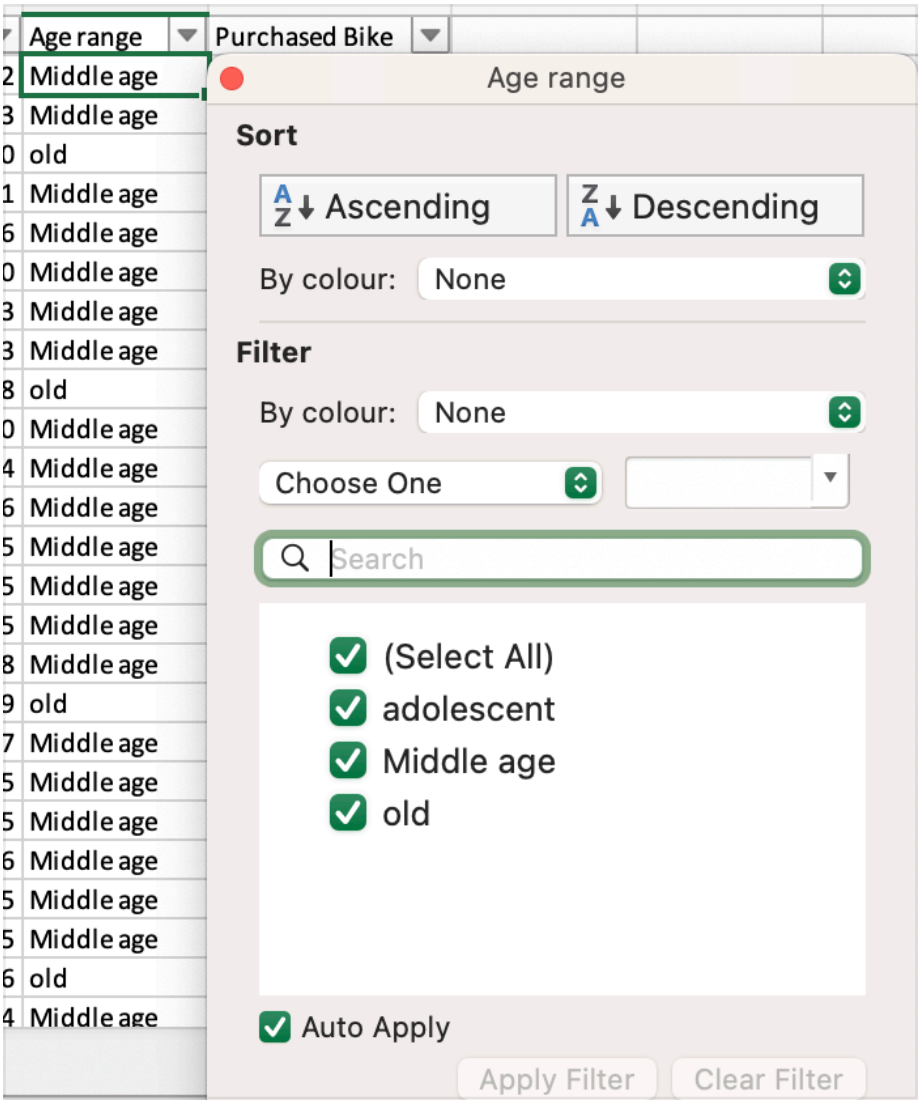
L	M
Age	Age range
42	invalid
43	invalid
60	invalid
41	invalid
36	invalid
50	invalid
33	invalid
43	invalid
58	invalid
40	invalid
54	invalid
36	invalid
55	invalid
35	invalid
45	invalid
38	invalid

- I will then build on the above IF statement and make it into a nested IF statement.- so I will be asking, if the age is bigger than or equivalent to 31, then they are middle aged,

```
=IF(L2>=31, "Middle age",IF(L2<31, "adolescent", "invalid"))
```

- Add one more IF statement so we have 3 age categories by the end (those less than 31 years old, those over/or 31 years old and those over the age of 54) - so that is:
  ~ <31 adolescent
  ~ >=31 middle aged
  ~ >54 old

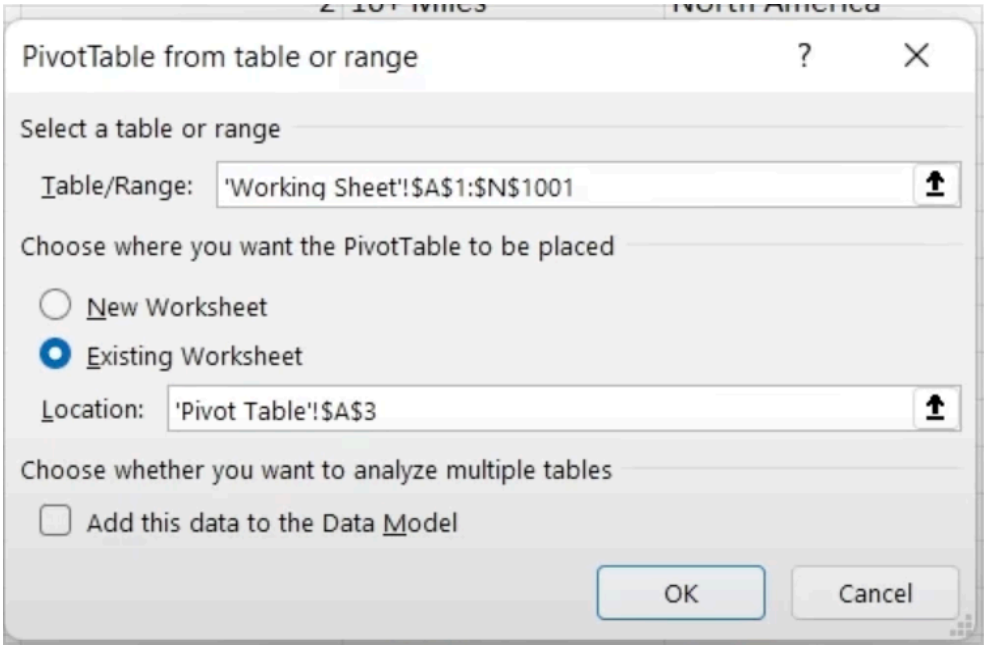
```
=IF(L2>55, "old",IF(L2>=31, "Middle age",IF(L2<31, "adolescent", "invalid")))
```



Now we have these categories for age which is much more usable than individual ages.

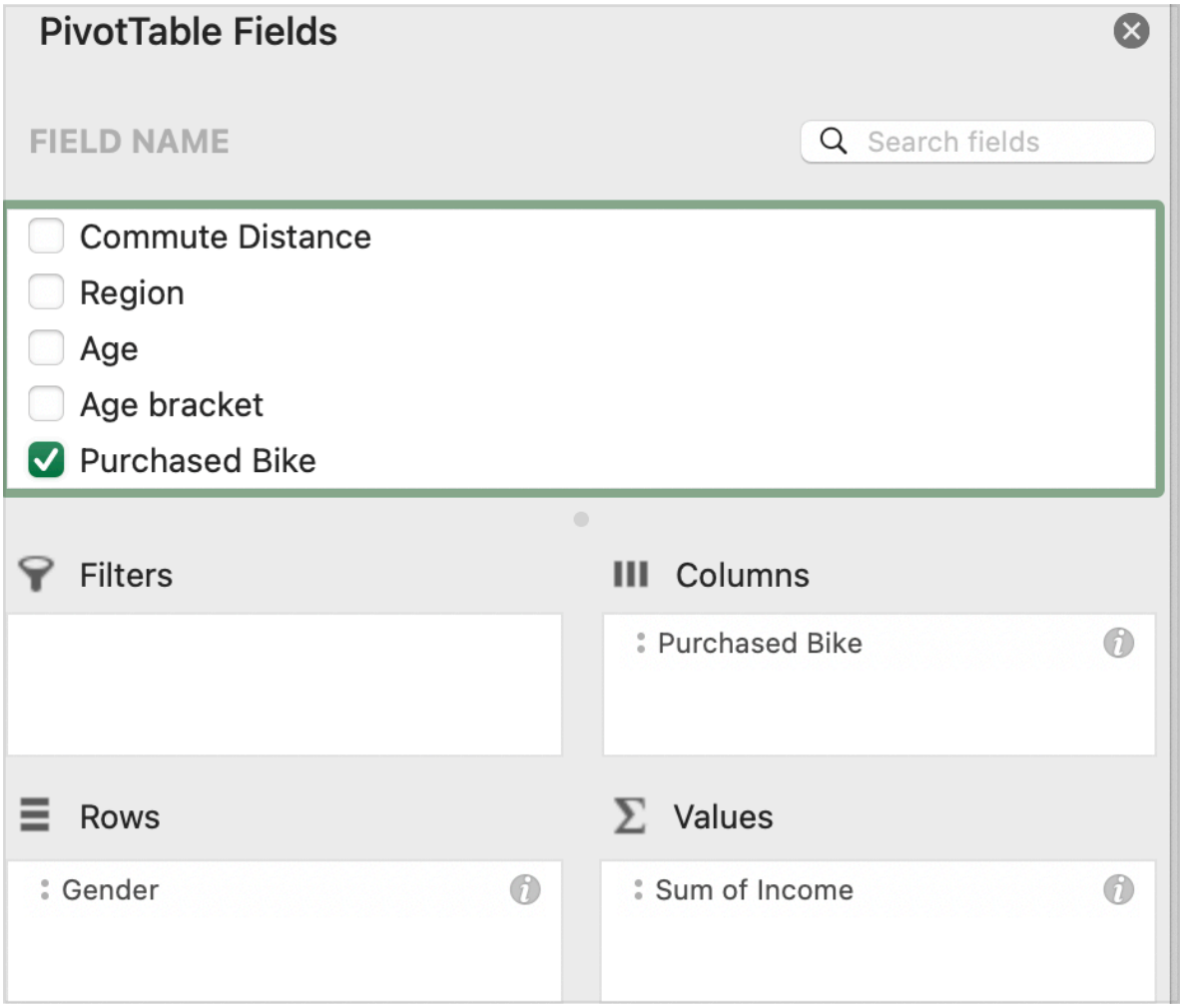
Pivot table

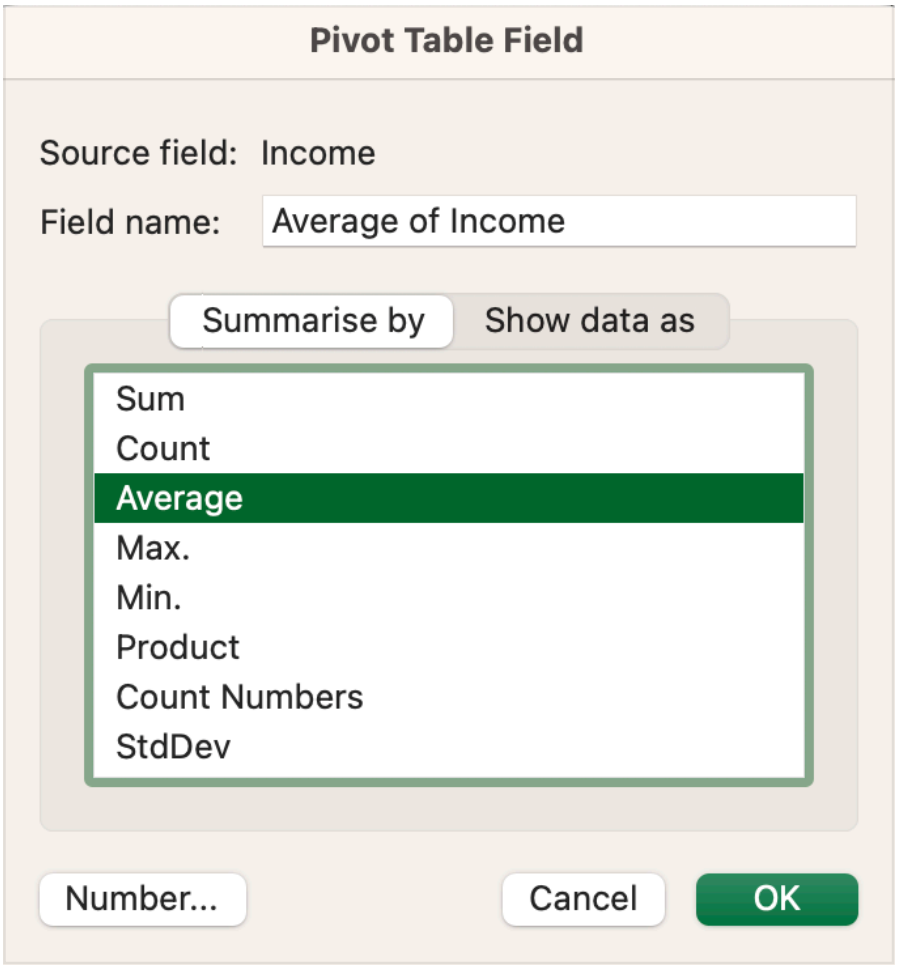
Now we are going to build our pivot table with our worksheet.



- Now I can select what columns/data I want to work with

- The first thing I will look at is the average income of someone who has either bought a bike or did NOT buy a bike (a value we need is income, break them into gender, and find the average for salary/income).

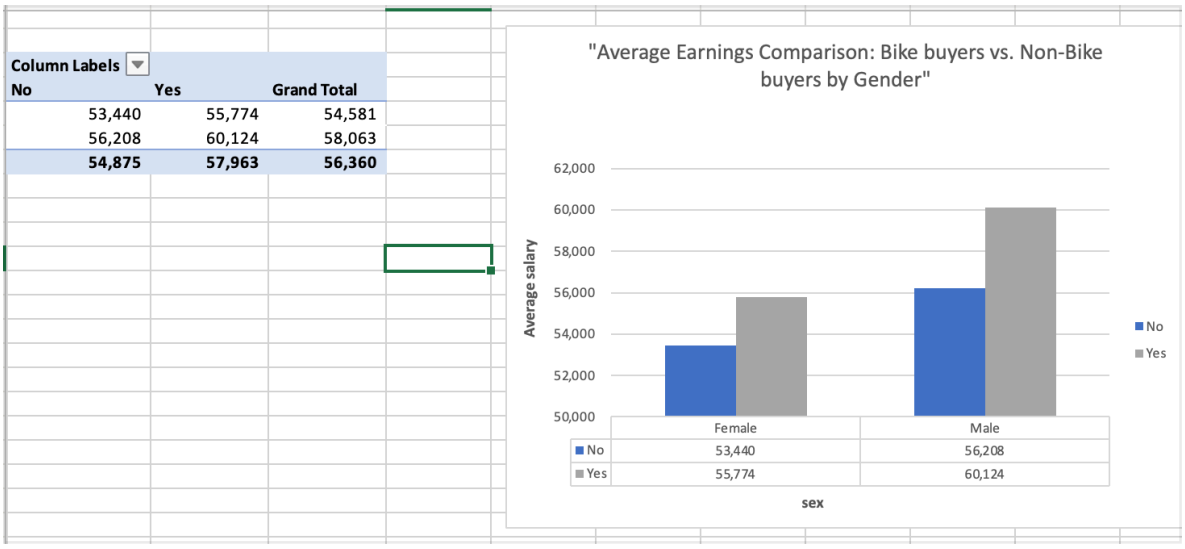

PivotTable Fields
FIELD NAME
Search fields
Commute Distance
Region
Age
Age bracket
Purchased Bike
Filters
Columns
Purchased Bike
Rows
Gender
Values
Sum of Income



Average of Income	Column Labels		
Row Labels	No	Yes	Grand Total
Female	53440	55774.05858	54580.7771
Male	56208.17844	60123.96694	58062.62231
Grand Total	**54874.75915**	**57962.57796**	**56360**

Findings from the above data?: those with a bit more money are buying bikes, and the men are earning more than women on average.

- Now I select a 2d column graph to represent my data, make visual customisations and add data table
- Decrease the amount of decimals and add comas to make it look more clean and digestible



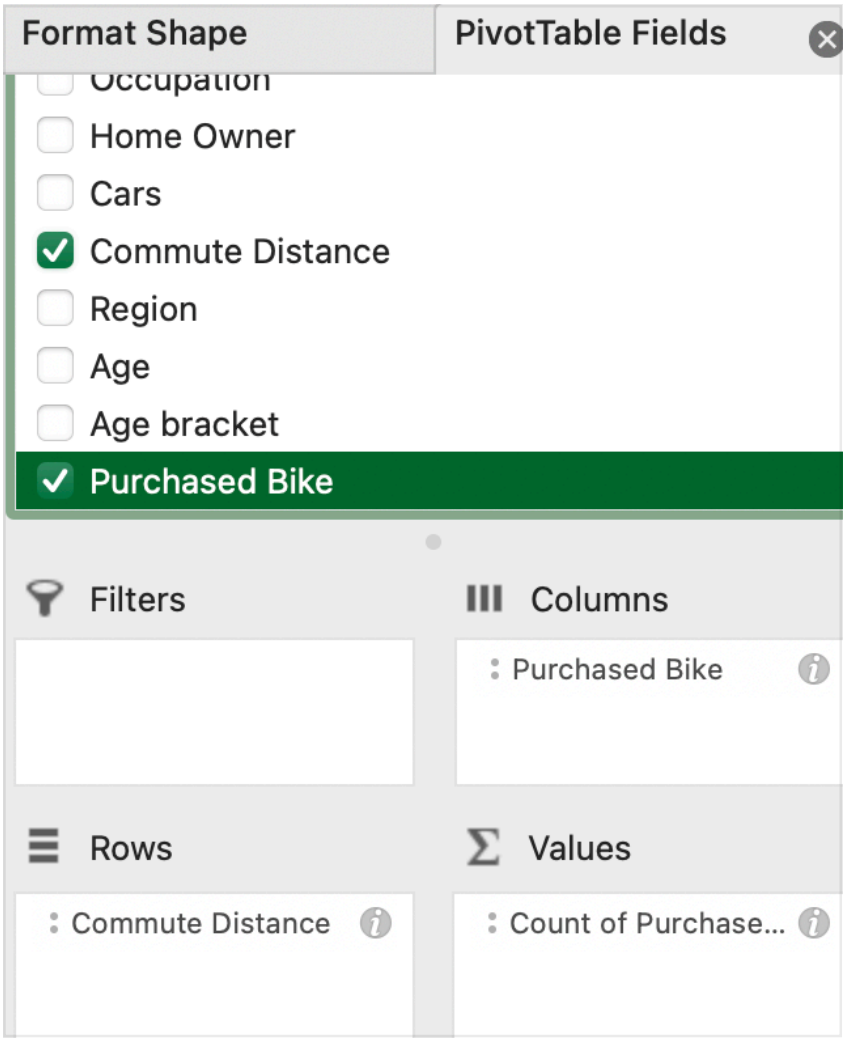
What was I looking at in the above table?

- Did their income change whether they bought a bike or did not buy a bike?
- For those that did buy bikes, what was the reason?
- Should we cater more to customers who make more money?

Second table

Next I looked at commute - it is very important because it could tell us whether distance from their house to work makes a difference with regards to who buys a bike and who doesn't.

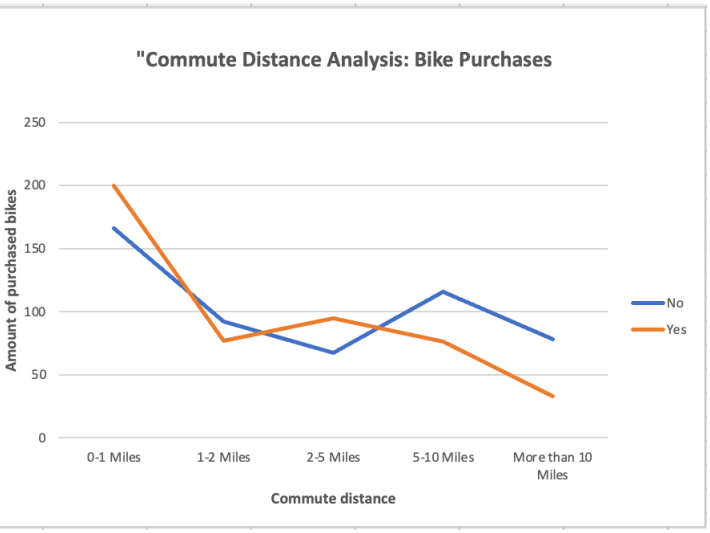
What our table will show us is the commute distance, whether they bought a bike or not, and how many of them bought a bike or not (a count).



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

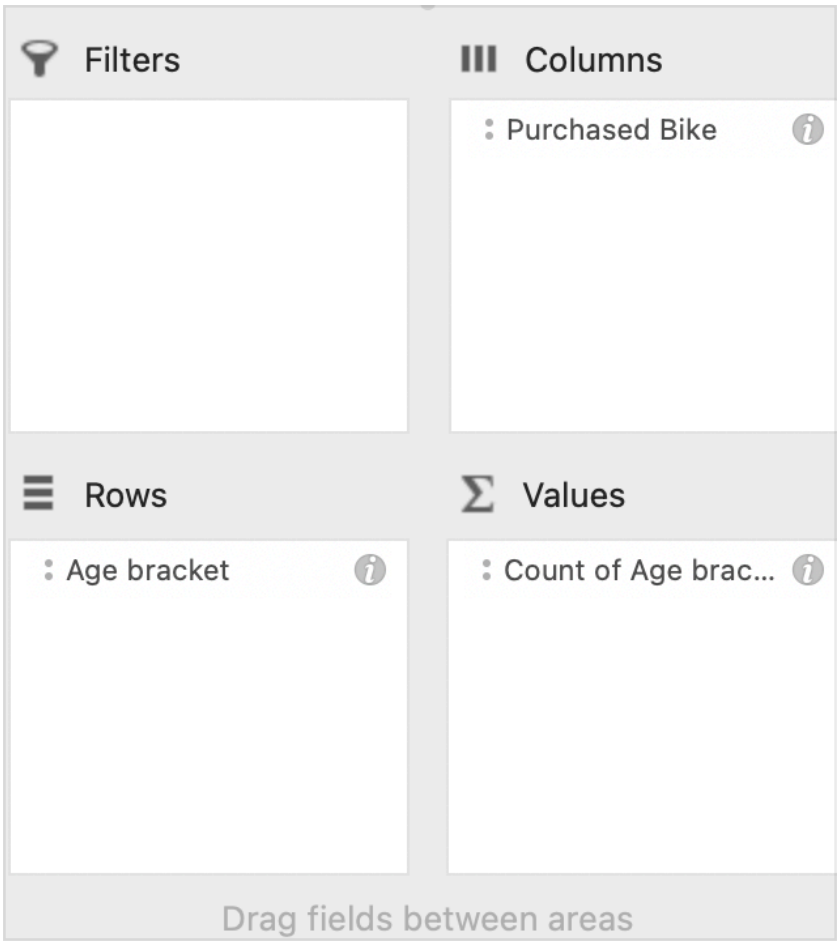
- Pick a graph
- Change the 10+ miles to More than 10 miles so it is moved to the bottom of the table - more visually appealing.

Count of Purchased	Colum		
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**

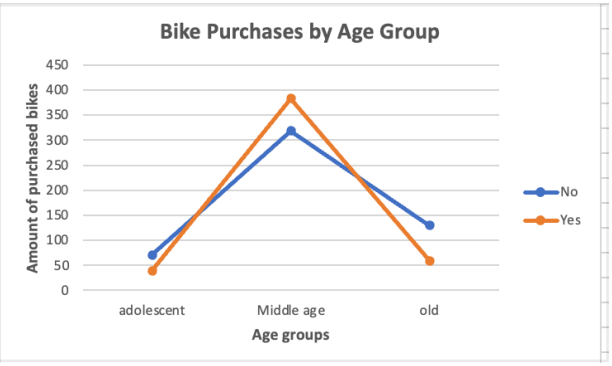


Third/Final table

- Focus on our age brackets - who purchased a bike and how many bikes were purchased

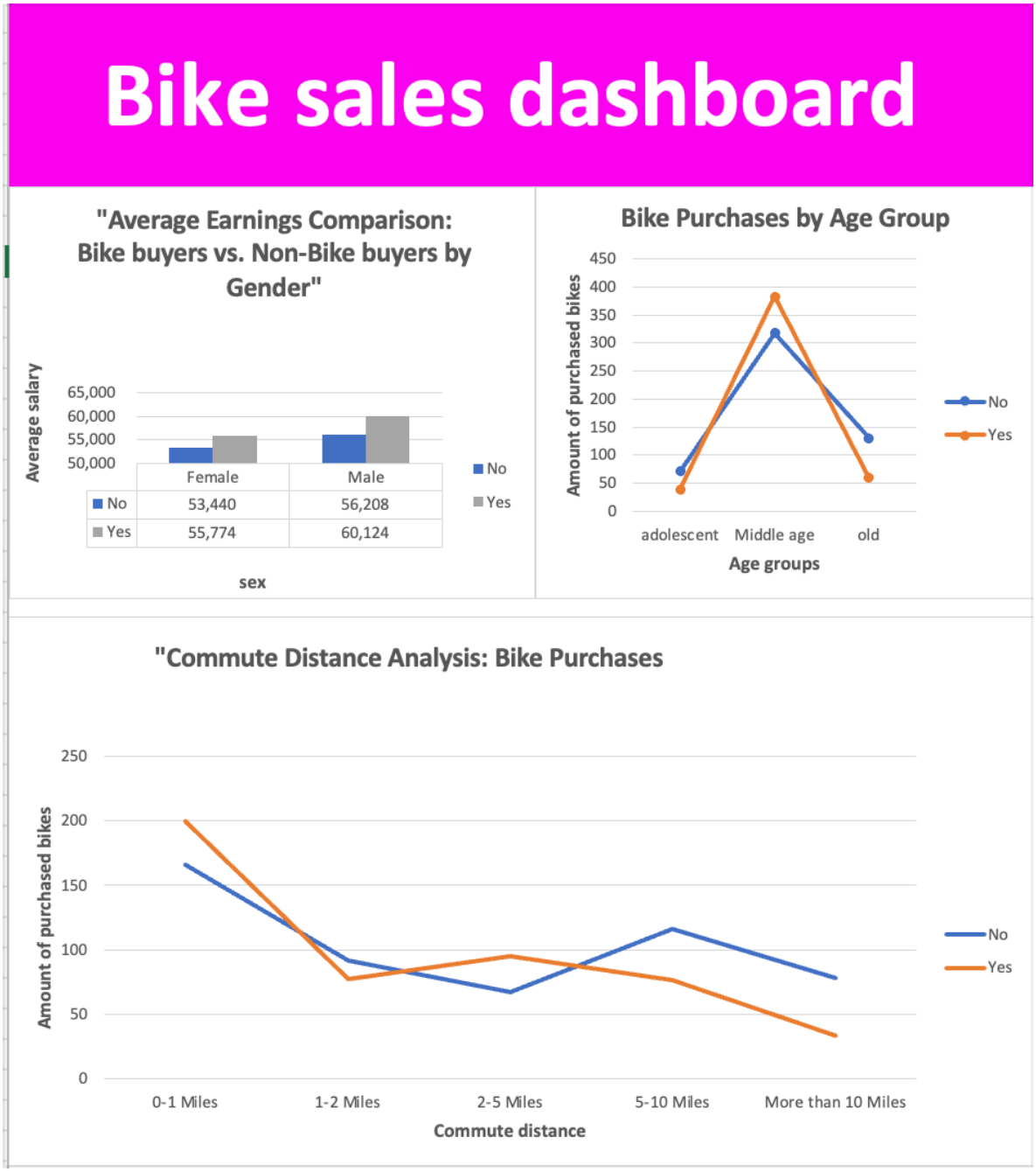


Count of Age bracket	Column Labels		
Row Labels	No	Yes	Grand Total
adolescent	71	39	110
Middle age	318	383	701
old	130	59	189
Grand Total	**519**	**481**	**1000**



Creating the dashboard

- Copy over the pivot tables to the dashboard tab
- Make dashboard visually appealing (remove gridlines etc)



We can ask questions like:

- Are married people making more purchases?
- Is there a relationship between education level and bike purchases? If so why? Could it be that those with higher level of education are likely to have higher paying jobs and therefore have more money to spend on a bike?
- Many more questions we can ask.

I started off by applying the marital status as a slicer connected to the "Average Earnings Comparison: Bike buyers vs. Non-Bike buyers by Gender" but then I connected it to all pivot tables.

I then added region and education as slicers.

This means I can make deeper analysis, take in various factors when making analysis.

For a business, it could inform them on marketing, target audience, where to open physical shops, what kind of bikes to sell more, ascetic on bikes, bike sizes to stock up on, and answer many more questions.

To add to this, we can start digging on the 'Why'? Why are more educated people buying bikes? Why are young people not buying as much bikes

The final Dashboard