# Dataset

The dataset I use has data for bike buyers with the below columns

ID

Marital Status Married or single

Gender Male or Female

Income

Children

Education

Occupation

Home Owner

Cars

Commute Distance

Region

Age

Purchased Bike Yes or No

# Cleaning dataset

1. Remove duplicates
2. Column **Marital Status**

Replacing M with Married and S with Single using if formula

*=IF(B2="M","Married","Single")*

1. Column **Gender**

Replacing F with Female and M with Male using if formula

*=IF(C2="F","Female","Male")*

1. Column **Income**

Convert currency to numbers without decimal digits

1. Column **Commute Distance**

Replace 10+ Miles to 10 Miles+ for the alphabetical order.

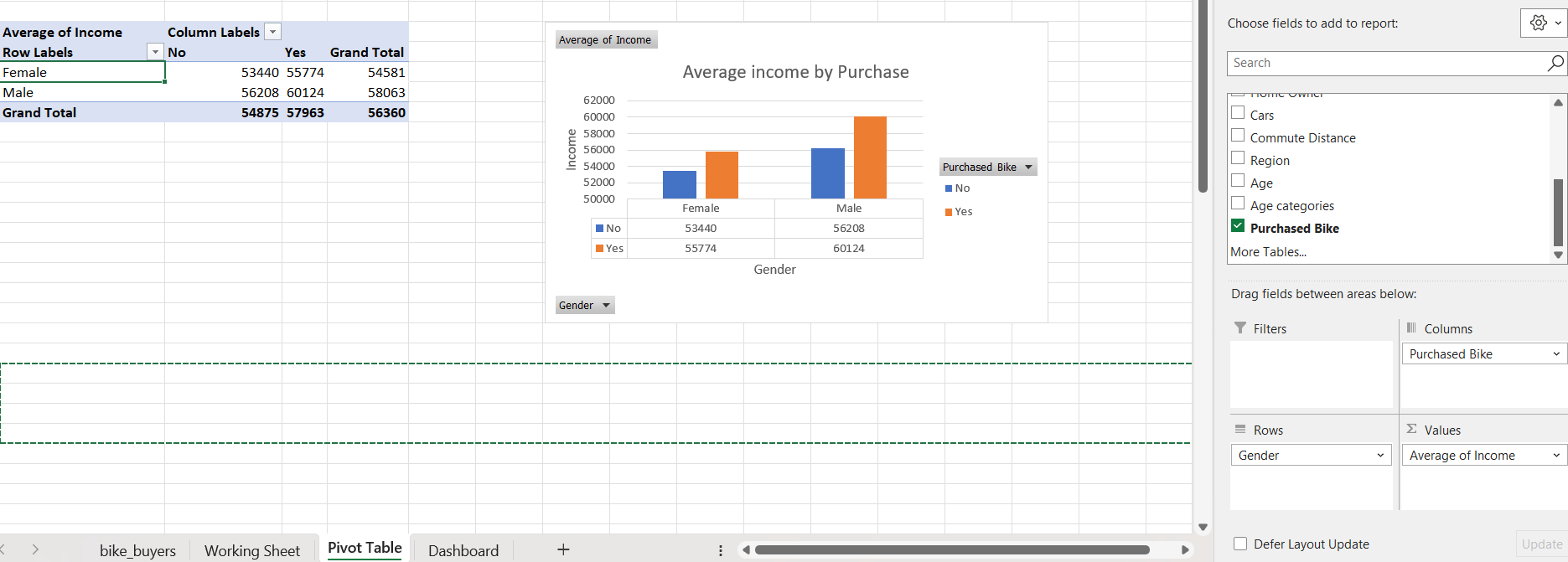
1. Add a new column (Age categories) in order to separate the buyers to three age categories (Adolescent, Middle Age, Old) using if formula

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

# Making pivot tables to answer questions

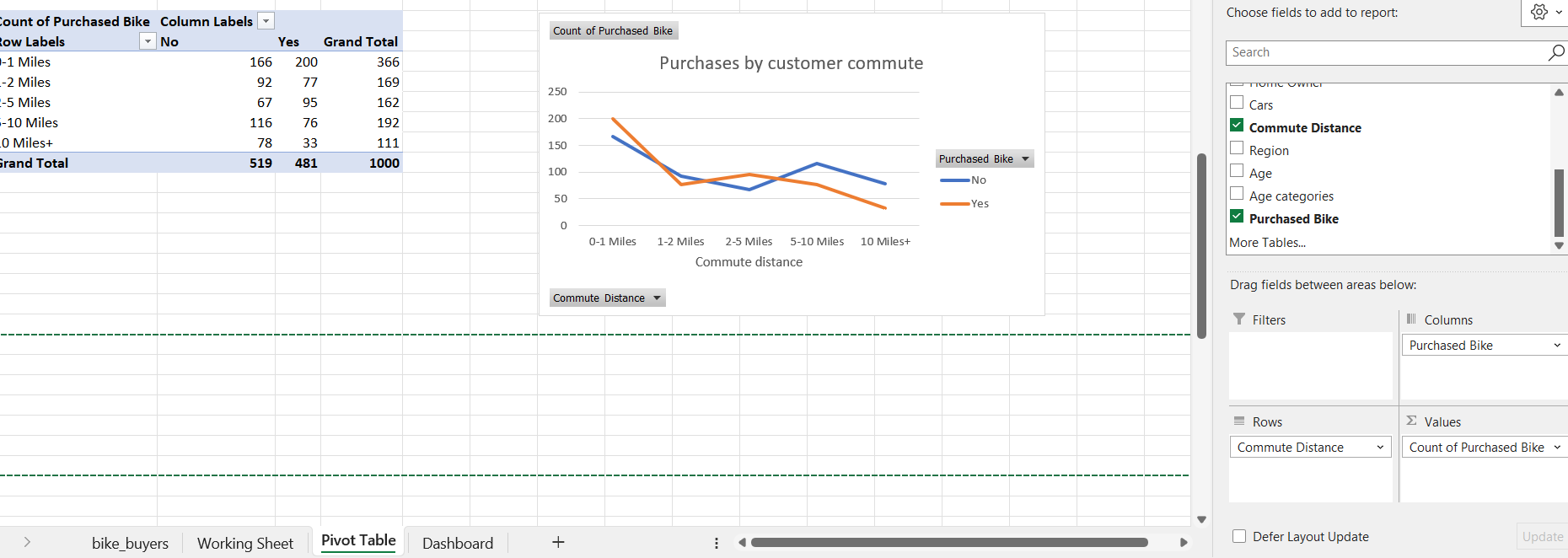
1. What is the average income of the customers who bought a bicycle? How many of them are men and how many are women?

I use a pivot table with gender for rows, purchased bike as columns and average income as values.



1. Is the number of the purchased bicycles dependent on the commute distance of each person?

I use a pivot table with commute distance for rows, purchased bike as columns and count of purchased bike as values.

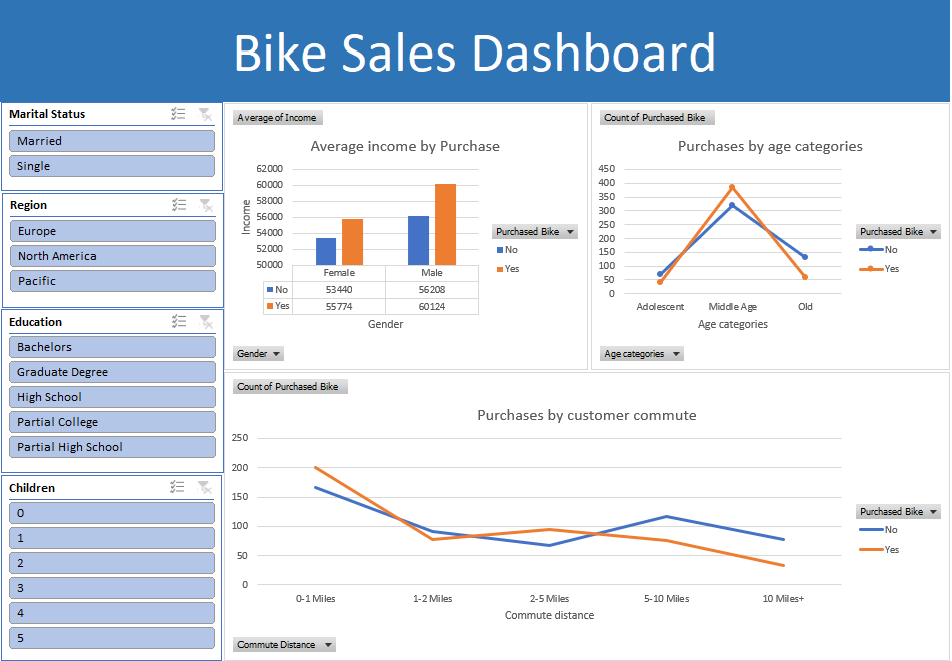


1. Is the number of the purchased bicycles dependent on the age of each person?

I use a pivot table with age categories for rows, purchased bike as columns and count of purchased bike as values.



# Dashboard



I use the charts that I made from pivot tables to create the dashboard. Along with the charts I also added some filters for marital status, region, education and number of children.