My Walkthrough steps following Alex The Analyst, Excel Final project: Made by Luke Yeo

Note:

* Dataset:<https://github.com/AlexTheAnalyst/ExcelTutorial/blob/main/Excel%20Project%20Dataset.xlsx>
* Note, the styling and designing of a dataset is not the main focus here.

Screening from the Data In Excel for a few seconds:

1. There might be duplicates
2. Income has the ‘$’ and ‘.00’
3. The header isn’t formatted properly

Doesn’t seem like there is much cleaning needed at first glance

Data Cleaning and Formatting:

Create a dashboard sheet, pivot table sheet, working sheet

1. Copy the raw data into working sheet
2. Remove Duplicates by going data -> remove duplicates: there were 26 duplicates found.
3. For Gender column, there are ‘M’ and ‘F’. However just in case it isn’t clear enough, we will put it as ‘Male’ and ‘Female’. -> -> Using ‘Find and Replace’, which is CRTL + H , change the setting to search by columns (default will be rows)
4. For Marital Status column, there are ‘M’ and ‘F’. However just in case it isn’t clear enough, we will put it as ‘Male’ and ‘Female’. -> Using ‘Find and Replace’, which is CRTL + H, change the setting to search by columns (default will be rows)
5. For Income Column, it has a dollar sign and is general -> Make it to ‘currency’ or ‘Numerical’ (though numerical, you will have to take out the $ sign and .00)
6. For Age column, we do want a age bracket like Young or Old -> to do so, create a new column, use the IF function:

**=IF(M2>50,"Old",IF(M2>=31,"Middle Age",IF(M2<31,"Adolescent","Invalid")))**

Data Visualization (Pivot Table):

* Go to Insert -> Table -> CRTL A the entire working sheet (FOR EACH TABLE)

1. Table 1: Income base off gender and if they purchase a bike. Values: Sum of income (right click to ‘Value Field Setting’ then click ‘Average’, Rows: Gender, Column: Purchase Bike.
2. Chart 1: go to INSERT, then click on the first column -> right click to add Axis Title and Chart Title
3. Table 2: To see if the diff commuting means more people buy bikes or not. Row: Commute Distance, Column: Purchase Bike, Values: Count of Purchase Bike

**3.5)** The original Commute Distance column has an issue in regards to going biggest value to smallest value -> Using ‘Find and Replace’, which is CRTL + H, Find ‘10+ Miles’ to ‘More than 10 Miles’

1. Chart 2: go to INSERT, then click on the first Line -> right click to add Axis Title and Chart Title (If did chart already before 3.5), go to Data at the top and click ‘Refresh All’)
2. Table 3: Age base on bike purchase -> Row: Age Bracket, Column: Purchase Bike, Value: Count of Purchase Bike
3. Chart 3: go to INSERT, then click on the first Line, this time we pick the ones with Dots -> right click to add Axis Title and Chart Title, add data labels too

Data Visualization (Dashboard):

1. Copy all Charts in Pivot Table Sheet to Dashboard Sheet
2. Get Rid of all the gridlines, click on the left Beside Column A + Top of Row 1 then click View at the top, then untick the gridlines
3. Create a box, puck a random colour and make the Header of the dashboard -> then click on ‘Merge & Center’
4. Click on any graph, go to ‘Pivot Chart Analyze’ at the top then click ‘Insert Slicer’. (See what else we want to add)
5. To connect to other graphs, click ‘Report Connection’ then tick all the pivot tables
6. Repeat step 4-6 till we are satisfied