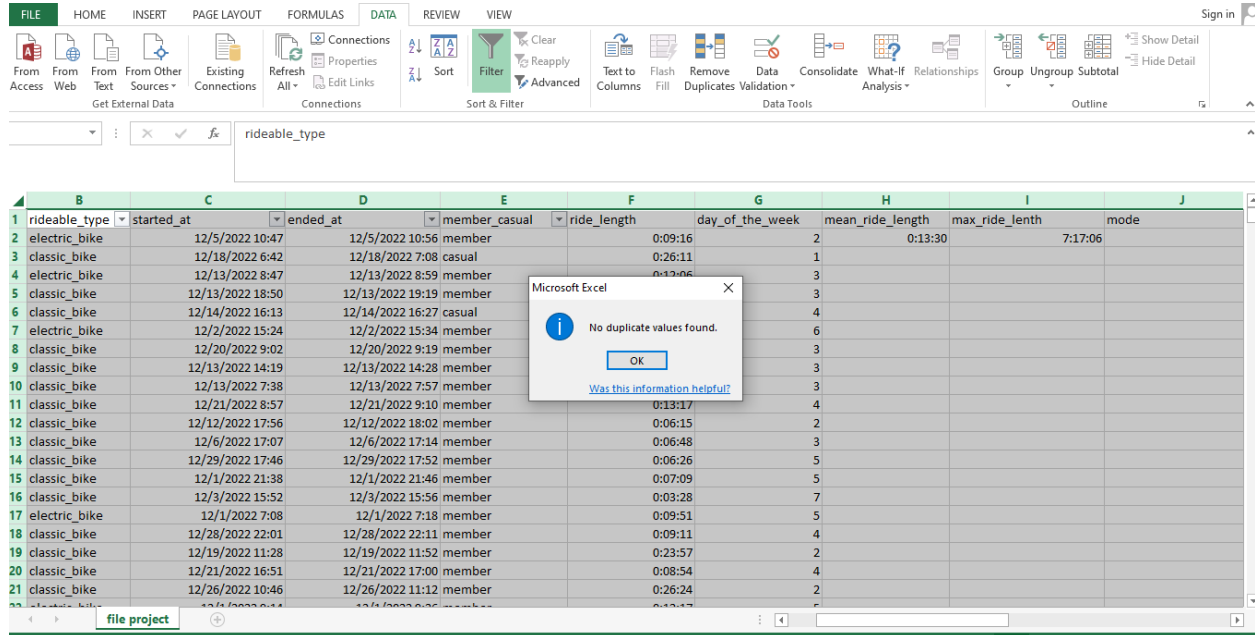# Exploratory Data Analysis in Microsoft Excel

1.0   CLEANING PROCESS

1. Removing duplicates

2. Using the function "find and replace", I located all blanks in spreadsheet and deleted them.

3. Made sure that all the tables are consistent (column names etc.)

4. Removed all the extra spaces using TRIM function. (Where applicable)

5. Made sure that all dates are in the same format.

6. Used FILTER function to be sure that there is not any error or unusual values.

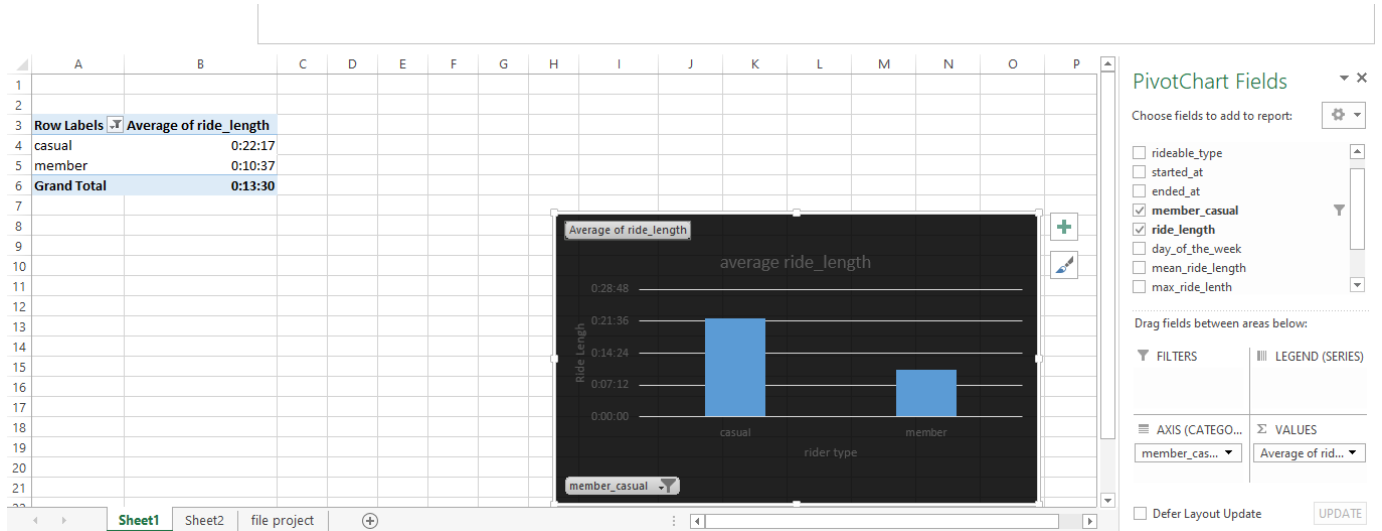7. Used Filter function to check for blanks

2.0   DATA EXPLORATION

1. I created new column called "ride length" and set up values in "time" HH: MM: SS 37:30:55

2. I created new column called "day of week" Where Sunday is 1 and Saturday is 7

3. I deleted columns which I will not use for the analysis, making tables easy to read and   understand. (start_lat, start_lng, end_lat,end_lng)

4. Calculated mean (average) and max values in column "ride length"

5. Calculated mode in column "day of week", to discover the most frequently occurring value that appears in this column.
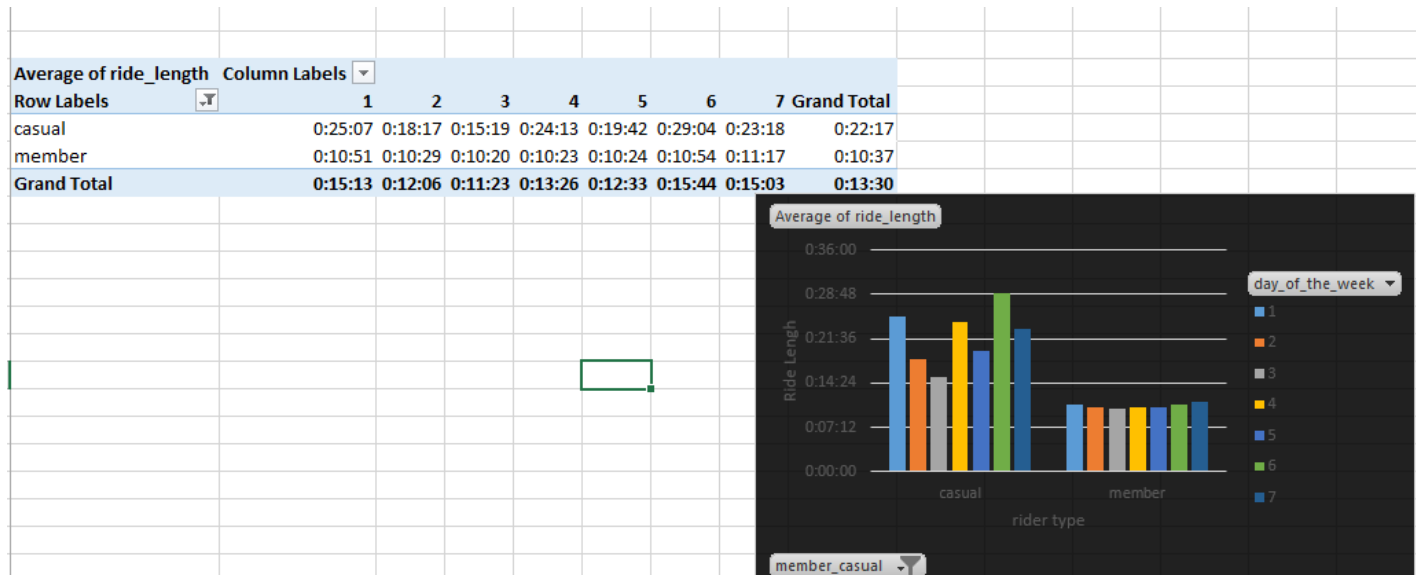

rideable_type	started_at	ended_at	member_casual	ride_length	day_of_the_week	mean_ride_length	max_ride_lenth	mode
electric_bike	12/5/2022 10:47	12/5/2022 10:56 member		0:09:16	2	0:13:30	7:17:06	
classic_bike	12/18/2022 6:42	12/18/2022 7:08 casual		0:26:11	1			
electric_bike	12/13/2022 8:47	12/13/2022 8:59 member		0:12:06	3			
classic_bike	12/13/2022 18:50	12/13/2022 19:19 member		0:29:01	3			
classic_bike	12/14/2022 16:13	12/14/2022 16:27 casual		0:14:11	4			
electric_bike	12/2/2022 15:24	12/2/2022 15:34 member		0:09:27	6			
classic_bike	12/20/2022 9:02	12/20/2022 9:19 member		0:17:51	3			
classic_bike	12/13/2022 14:19	12/13/2022 14:28 member		0:08:39	3			
classic_bike	12/13/2022 7:38	12/13/2022 7:57 member		0:18:54	3			
classic_bike	12/21/2022 8:57	12/21/2022 9:10 member		0:13:17	4			
classic_bike	12/12/2022 17:56	12/12/2022 18:02 member		0:06:15	2			
classic_bike	12/6/2022 17:07	12/6/2022 17:14 member		0:06:48	3			
classic_bike	12/29/2022 17:46	12/29/2022 17:52 member		0:06:26	5			
classic_bike	12/1/2022 21:38	12/1/2022 21:46 member		0:07:09	5			
classic_bike	12/3/2022 15:52	12/3/2022 15:56 member		0:03:28	7			
electric_bike	12/1/2022 7:08	12/1/2022 7:18 member		0:09:51	5			
classic_bike	12/28/2022 22:01	12/28/2022 22:11 member		0:09:11	4			
classic_bike	12/19/2022 11:28	12/19/2022 11:52 member		0:23:57	2			
classic_bike	12/21/2022 16:51	12/21/2022 17:00 member		0:08:54	4			
classic_bike	12/26/2022 10:46	12/26/2022 11:12 member		0:26:24	2			

6. Created pivot table to get initial insight of data set

7. Using pivot table I calculated average ride length, average ride length by day, and number of rides per day, and displayed results through charts to make it easier for stakeholders to understand

2.1 Average Ride_Length

Row Labels	Average of ride_length
casual	0:22:17
member	0:10:37
Grand Total	0:13:30

2.2 Average ride length by day (Where #1 is Sunday, #7 is Saturday)

Average of ride_length	Column Labels							
Row Labels	1	2	3	4	5	6	7	Grand Total
casual	0:25:07	0:18:17	0:15:19	0:24:13	0:19:42	0:29:04	0:23:18	0:22:17
member	0:10:51	0:10:29	0:10:20	0:10:23	0:10:24	0:10:54	0:11:17	0:10:37
Grand Total	0:15:13	0:12:06	0:11:23	0:13:26	0:12:33	0:15:44	0:15:03	0:13:30

2.3   Number of rides per day (Where #1 is Sunday, #7 is Saturday)

Count of member_casual	Column Labels							
Row Labels	1	2	3	4	5	6	7	Grand Total
casual	5343.00	4936.00	5917.00	5678.00	8177.00	7156.00	7687.00	44894.00
member	12117.00	18889.00	22108.00	20162.00	27095.00	19662.00	16879.00	136912.00
Grand Total	17460.00	23825.00	28025.00	25840.00	35272.00	26818.00	24566.00	181806.00