/\* Spreadsheet formulas and steps for Cyclistic bike-share case study \*/

/\* STEP 1: Data Preparation in Google Sheets or Excel \*/

/\* Create a new sheet for 2019 Q1 data \*/

/\* Import the CSV file: Divvy\_Trips\_2019\_Q1.csv \*/

/\* Create a new sheet for 2020 Q1 data \*/

/\* Import the CSV file: Divvy\_Trips\_2020\_Q1.csv \*/

/\* STEP 2: Data Cleaning and Standardization \*/

/\* In 2019 Q1 sheet, add a new column for standardized member type \*/

/\* In cell N2 (assuming column M is the last data column), add formula: \*/

=IF(L2="Subscriber", "member", IF(L2="Customer", "casual", L2))

/\* Copy this formula down for all rows \*/

/\* Add a new column for ride\_length in minutes \*/

/\* In cell O2, add formula: \*/

=ROUND((C2-B2)\*1440, 2)

/\* Copy this formula down for all rows \*/

/\* Add a new column for day\_of\_week \*/

/\* In cell P2, add formula: \*/

=TEXT(B2, "dddd")

/\* Copy this formula down for all rows \*/

/\* In 2020 Q1 sheet, add a new column for ride\_length in minutes \*/

/\* In cell M2 (assuming column L is the last data column), add formula: \*/

=ROUND((D2-C2)\*1440, 2)

/\* Copy this formula down for all rows \*/

/\* Add a new column for day\_of\_week \*/

/\* In cell N2, add formula: \*/

=TEXT(C2, "dddd")

/\* Copy this formula down for all rows \*/

/\* STEP 3: Create a Combined Data Analysis Sheet \*/

/\* Create a new sheet called "Analysis" \*/

/\* STEP 4: Calculate Summary Statistics by Member Type \*/

/\* In cell A1, add header: "Summary Statistics by Member Type" \*/

/\* Set up column headers in row 2: Member Type, Total Rides, Average Ride Length, Median Ride Length, Min Ride Length, Max Ride Length \*/

/\* For 2019 Q1 data, in cell A3, add: "member (2019 Q1)" \*/

/\* In cell B3, add formula to count member rides: \*/

=COUNTIFS('2019 Q1'!$N:$N, "member")

/\* In cell C3, add formula to calculate average ride length for members: \*/

=AVERAGEIFS('2019 Q1'!$O:$O, '2019 Q1'!$N:$N, "member")

/\* In cell D3, add formula for median (may require array formula in Excel): \*/

=MEDIAN(IF('2019 Q1'!$N:$N="member", '2019 Q1'!$O:$O))

/\* Note: In Google Sheets, this would be: \*/

=MEDIAN(FILTER('2019 Q1'!$O:$O, '2019 Q1'!$N:$N="member"))

/\* In cell E3, add formula for minimum: \*/

=MIN(IF('2019 Q1'!$N:$N="member", '2019 Q1'!$O:$O))

/\* In Google Sheets: \*/

=MIN(FILTER('2019 Q1'!$O:$O, '2019 Q1'!$N:$N="member"))

/\* In cell F3, add formula for maximum: \*/

=MAX(IF('2019 Q1'!$N:$N="member", '2019 Q1'!$O:$O))

/\* In Google Sheets: \*/

=MAX(FILTER('2019 Q1'!$O:$O, '2019 Q1'!$N:$N="member"))

/\* Repeat for casual riders in row 4 \*/

/\* Repeat for 2020 Q1 data in rows 5-6 \*/

/\* STEP 5: Create a Pivot Table for Day of Week Analysis \*/

/\* Create a new sheet called "Day of Week Analysis" \*/

/\* Insert a pivot table using data from both 2019 Q1 and 2020 Q1 sheets \*/

/\* For the pivot table: \*/

/\* Rows: day\_of\_week \*/

/\* Columns: member\_casual \*/

/\* Values: Count of ride\_id (for number of rides) \*/

/\* Values: Average of ride\_length (for average ride duration) \*/

/\* Sort the day\_of\_week field in custom order: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday \*/

/\* STEP 6: Create a Pivot Table for Hourly Analysis \*/

/\* First, add a column for hour of day in both data sheets \*/

/\* In 2019 Q1 sheet, add column Q2 with formula: \*/

=HOUR(B2)

/\* Copy down for all rows \*/

/\* In 2020 Q1 sheet, add column O2 with formula: \*/

=HOUR(C2)

/\* Copy down for all rows \*/

/\* Create a new sheet called "Hourly Analysis" \*/

/\* Insert a pivot table using data from both sheets \*/

/\* For the pivot table: \*/

/\* Rows: hour\_of\_day \*/

/\* Columns: member\_casual \*/

/\* Values: Count of ride\_id (for number of rides) \*/

/\* Values: Average of ride\_length (for average ride duration) \*/

/\* STEP 7: Create a Pivot Table for Popular Stations \*/

/\* Create a new sheet called "Popular Stations" \*/

/\* Insert a pivot table using data from both sheets \*/

/\* For the pivot table: \*/

/\* Rows: start\_station\_name \*/

/\* Columns: member\_casual \*/

/\* Values: Count of ride\_id (for number of rides) \*/

/\* Sort by count of ride\_id in descending order \*/

/\* Filter to show only top 10 stations for each member type \*/

/\* STEP 8: Create Charts for Visualization \*/

/\* Create a column chart for rides by day of week \*/

/\* Select the data in the Day of Week Analysis sheet \*/

/\* Insert > Chart > Column Chart \*/

/\* Format the chart with appropriate titles and labels \*/

/\* Create a column chart for average ride duration by day of week \*/

/\* Select the appropriate data in the Day of Week Analysis sheet \*/

/\* Insert > Chart > Column Chart \*/

/\* Format the chart with appropriate titles and labels \*/

/\* Create a line chart for rides by hour of day \*/

/\* Select the data in the Hourly Analysis sheet \*/

/\* Insert > Chart > Line Chart \*/

/\* Format the chart with appropriate titles and labels \*/

/\* Create a bar chart for top stations for each member type \*/

/\* Select the data in the Popular Stations sheet \*/

/\* Insert > Chart > Bar Chart \*/

/\* Format the chart with appropriate titles and labels \*/

/\* STEP 9: Create a Dashboard Sheet \*/

/\* Create a new sheet called "Dashboard" \*/

/\* Copy or reference the key charts and summary statistics \*/

/\* Add text boxes with key findings and recommendations \*/

/\* STEP 10: Export Results \*/

/\* Export the Dashboard sheet as PDF \*/

/\* Export key analysis sheets as CSV files for further analysis \*/