# Cyclistic

Objective: How do annual members and casual riders use Cyclistic bikes differently.

Data is the last 12 months of trip data from divvy-tripdata.

Data is stored in spreadsheets from May 2022 To April 2023

License for data used provided by [Data License Agreement | Divvy Bikes]

**Cleaning log**

**Dataset:** *202205-divvy-tripdata.*

I created a copy of the data to work with saved under the name may_2022_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name may_2022_tripdata

**Data set:** *202206-divvy-tripdata*

I created a copy of the data to work with saved under the name june_2022_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name june_2022_tripdata

**Data set:** *202207-divvy-tripdata*

I created a copy of the data to work with saved under the name july_2022_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name july_2022_tripdata

**Data set:** *202208-divvy-tripdata*

I created a copy of the data to work with saved under the name august_2022_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name august_2022_tripdata

**Data set:** *202209-divvy-publictripdata*

I created a copy of the data to work with saved under the name september_2022_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name september_2022_tripdata

**Data set:** *202210-divvy-tripdata*

I created a copy of the data to work with saved under the name october_2022_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name october_2022_tripdata

**Data set:** *202211-divvy-tripdata*

I created a copy of the data to work with saved under the name november_2022_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name november_2022_tripdata

**Data set:** *202212-divvy-tripdata*

I created a copy of the data to work with saved under the name december_2022_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name december_2022_tripdata

**Data set:** *202301-divvy-tripdata*

I created a copy of the data to work with saved under the name january_20223_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name january_2023_tripdata

**Data set:** *202302-divvy-tripdata*

I created a copy of the data to work with saved under the name february_20223_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name february_20223_tripdata

**Data set:** *202303-divvy-tripdata*

I created a copy of the data to work with saved under the name march_2023_tripdata

I applied filter dropdown menus to the headers.

I hid column A to make the data easier to read and the member ids will not be used.

I deleted column E,F,G,H to make the data easier to read and station ids will not be used.

I deleted column I,J,K,L to make the data easier to read and location data can not be used per the license agreement.

After deleting all columns that will not be used in the analysis I was left with

Column B- rideable_type

Column C- strated_at

Column D- ended_at

Column E- member_casual

Using the drop filters I checked for any blank cells. None were found.

I created a new column in column F1 containing headed ride_length

in column F I subtracted the totals of column E and C giving me the total ride times.

I converted the data in column F into times in the HH:MM:SS format.

I created a new column in column G1 containing the header day_of_week

I filled column G using the weekday function using Monday as 1 through Sunday as 7

I used the drop down filter to check again for any black or errors.

I saved the cleaned data set in a cleaned data folder with the name march_2023_tripdata

All 12 months of the excel sheets were combines into a single excel workbook titled bike_tripdata

Chart for each month showing total_rides, average ride, total Monday, total Tuesday, total Wednesday, total Thursday, total Friday, total Saturday, total Sunday for each month.

A pivot chart was made for each month showing the total member and casual riders on each day of the week, for each month.

A bar graph showing the breakdown of the days of week was created for each month.

A pie chart showing the members vs casual riders was created for each month.

Only the totals for each month was transferred to a new sheet in the workbook transferring over the formulas from each month to be calculated. Page was titled totals

The total rides for each month was were added together on the totals sheet.

The member and casual riders were added together on the totals page.

Each day of the week was added together on the totals page.

The members and casual totals were used on the totals page to make a pie chart showing the difference visually.

The total riders for each month was added to a new sheet in the workbook titled by month.

The totals from the totals page were used to make a line chart showing the trend of use per month.

All the pivot charts for each month were transferred to a new sheet named pivot charts for calculations.

The pivot charts were used to make a new chart showing all the member vs casual riders for the whole year per day of the week.

The new chart showing members vs casual riders per week was used to make a bar graph showing each day of the week with a member and casual rider breakdown.

The monthly totals and all charts were transferred to a power point for a final presentation to the client.

All data and cleaned data has been preserved and is available upon request.