# Case study: How does a bike-share navigate speedy success?

## ASK

**BUSINESS TASK STATEMENT :**

The Statement of this problem gives that a Company called "Bike-Share Company" at Chicago, Wants to improve their sales and to reach the profitable outcome. To opt this, the team of the company decides to convert all the casual riders to a Annual members instead of finding for a new members. Through their analysis, it is driven as the annual members have the highest profit when compared to the casual riders.As a junior data analyst of this company, my sort of analysis would be how do annual members and casual riders use Cyclistic bikes differently? I try to solve this by understand and optimize the usage patterns of Cyclistic's bike-sharing service, we aim to analyze how annual members and casual riders use Cyclistic bikes differently. By examining key metrics such as ride duration, frequency, time of use, and routes taken, we will identify distinct behavioral trends between these two user groups. This analysis will inform strategic decisions for improving bike availability, station placement, and marketing efforts to enhance user satisfaction and operational efficiency.

## PREPARE

**DATA SOURCE DESCRIPTION :**

To Analyse and convert the casual riders to a annual members, I have chosen a Dataset from a list of datasets.

**Dataset Name:** 202307-divvy-tripdata.xsl

**Dataset Description:** The dataset have the information about the rider_id, their type, start and end timing of the bike share, stations they returned and the member's casual.

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | ride_id | rideable_type | started_at | ended_at | start_station_name | start_station_id | end_station_name | end_station_id | start_lat | start_lng | end_lat | end_lng | member_casual |
| 2 | 9340B064F0AEE130 | electric_bike | 20:06:14 | 20:22:44 | Kedzie Ave & 110th St | 20204 | Public Rack - Racine Ave & 109th Pl | 877 | 41.69241 | -87.7009 | 41.69484 | -87.653 | member |
| 3 | D1460EE3CE0D8AF8 | classic_bike | 17:05:07 | 17:18:37 | Western Ave & Walton St | KA1504000103 | Milwaukee Ave & Grand Ave | 13033 | 41.89842 | -87.6866 | 41.89158 | -87.6484 | member |
| 4 | DF41BE31B895A25E | classic_bike | 10:14:53 | 10:24:29 | Western Ave & Walton St | KA1504000103 | Damen Ave & Pierce Ave | TA1305000041 | 41.89842 | -87.6866 | 41.9094 | -87.6777 | member |
| 5 | 9624A293749EF703 | electric_bike | 08:27:44 | 08:32:40 | Racine Ave & Randolph St | 13155 | Clinton St & Madison St | TA1305000032 | 41.88411 | -87.6569 | 41.88275 | -87.6412 | member |
| 6 | 2F68A6A4CDB4C99A | classic_bike | 15:46:42 | 15:58:08 | Clark St & Leland Ave | TA1309000014 | Montrose Harbor | TA1308000012 | 41.96709 | -87.6673 | 41.96398 | -87.6382 | member |
| 7 | 9AEE973E6B941A9C | classic_bike | 08:44:47 | 08:49:41 | Racine Ave & Randolph St | 13155 | Sangamon St & Lake St | TA1306000015 | 41.88407 | -87.6569 | 41.88578 | -87.651 | member |
| 8 | E366E997FDA1582B | classic_bike | 14:30:44 | 14:37:45 | Clark St & Leland Ave | TA1309000014 | Sheridan Rd & Montrose Ave | TA1307000107 | 41.96709 | -87.6673 | 41.96167 | -87.6546 | member |
| 9 | 1BB3E73851E6C2C1 | classic_bike | 10:11:53 | 10:17:55 | Clark St & Leland Ave | TA1309000014 | Ravenswood Ave & Berteau Ave | TA1309000018 | 41.96709 | -87.6673 | 41.95792 | -87.6736 | member |
| 10 | DA1E1D0866E6566E | electric_bike | 21:57:27 | 22:08:27 | Clark St & Leland Ave | TA1309000014 | Sheffield Ave & Wellington Ave | TA1307000052 | 41.96709 | -87.6675 | 41.93625 | -87.6527 | member |

Inorder to analyse how the casual and the annual member use cyclist differently,is calculated through the rate_Length and Day_of_week. The updated Data Columns are attached Below.

| N | O |
|---|---|
| Ride_length | Day_of_Week |
| 00:16:30 | 1 |
| 00:13:30 | 1 |
| 00:09:36 | 1 |
| 00:04:56 | 6 |
| 00:11:26 | 7 |
| 00:04:54 | 2 |
| 00:07:01 | 3 |
| 00:06:02 | 6 |

# PROCESS

The Process Step involves finding the error and cleaning the data which led to analysing Phase. The tool that I have chosen to Process,Analyse the data is **EXCEL.**

## *The Reason to choose Excel ?*

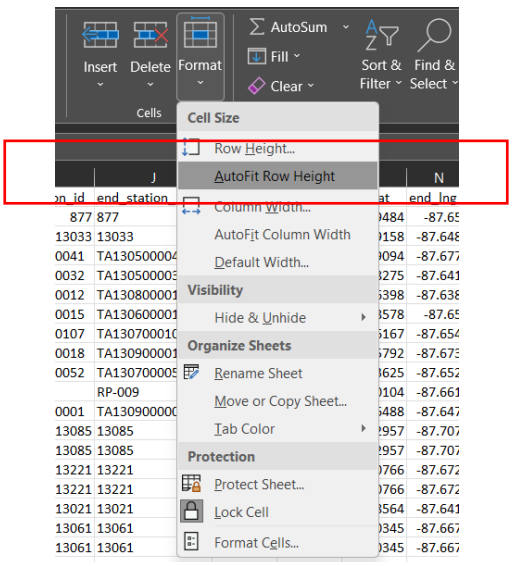
- Excel is a popular tool for analysis due to its user-friendly interface, extensive data visualization capabilities, and powerful built-in functions for statistical and mathematical calculations.
- It allows for quick data manipulation and supports a wide range of data formats, making it accessible for both beginners and advanced users.
- Additionally, Excel's pivot tables and charts facilitate efficient data summarization and reporting.

## *Steps for Data Cleaning in Excel :*

The steps that I used to clean the data is as follows,

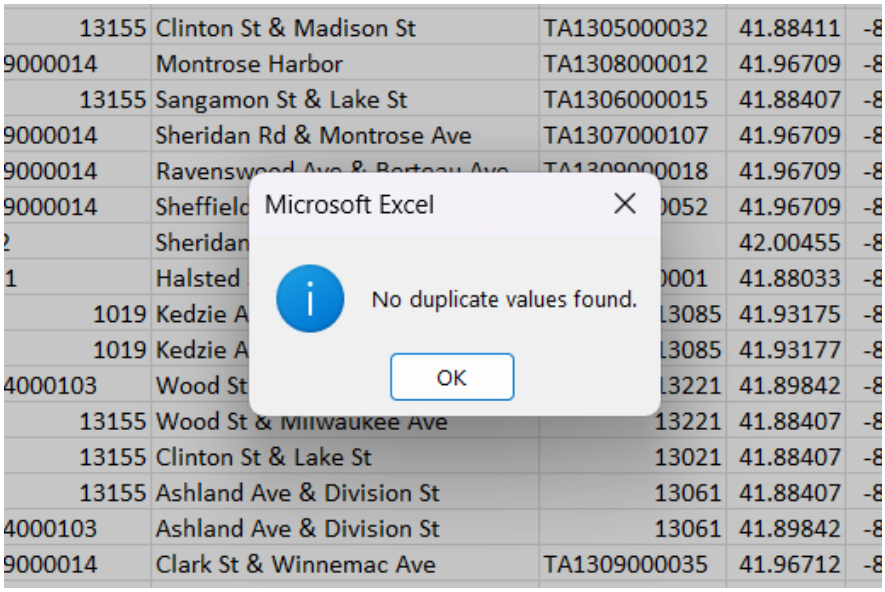
### (i) Auto fit rows and columns

Auto-fitting rows and columns in data cleaning ensures that all data is visible without manual adjustments, improving readability and accuracy. This step helps quickly identify inconsistencies or errors in the dataset.



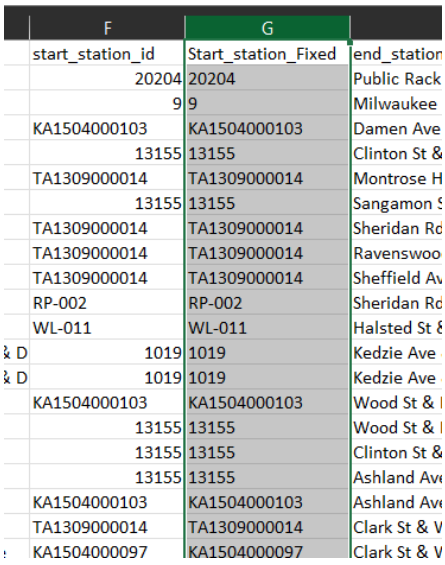
## (ii) Remove Duplicates

Identify and eliminate duplicate records to ensure each bike trip is only counted once.




## (iii) Missing Spaces

The Cells with additional spaces or trailing spaces are removed and are rearranged using the **trim()** function in the separate column beside of the original column.

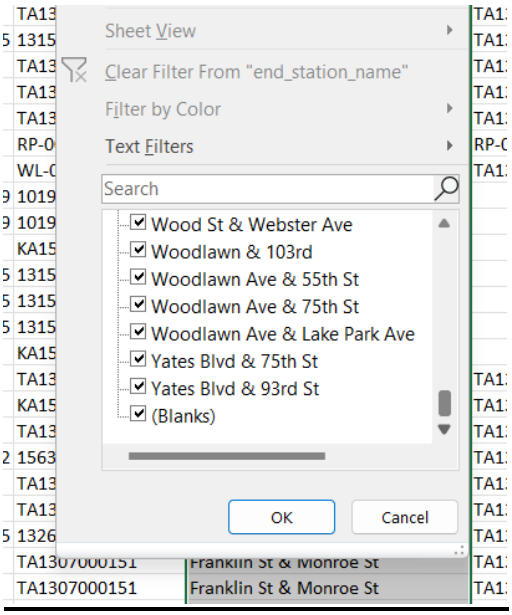



## (iv) Filling the empty cells

The dataset have the empty cells,which tend to loss the inconsistency of the data. To alter the blank spaces it is filled with the N/A to the entire table.

| | D | E | F | G | |
|---|---|---|---|---|---|
| | | | | | |
| :19:47 | 14:33:34 | NA | NA | NA | Pi |
| :39:46 | 19:46:43 | NA | NA | NA | Pi |
| :08:01 | 17:17:55 | NA | NA | NA | Br |
| :59:48 | 13:19:14 | NA | NA | NA | Pi |
| :38:53 | 22:44:57 | NA | NA | NA | Da |

Finally the data are evaluated for the consistency and accuracy using the sort&filter,data Validation and Check erros function




# ANALYSE

The data in the spread sheet are cleaned and are ready for the analysis phase.

## *Performing Data Transformation:*

| ride_id | rideable_type | started_at | ended_at | start_station_name |
|---|---|---|---|---|
| 9340B064F0AEE130 | electric_bike | 20:06:14 | 20:22:44 | Kedzie Ave & 110th St |
| D1460EE3CE0D8AF8 | classic_bike | 17:05:07 | 17:18:37 | Western Ave & Walton St |
| DF41BE31B895A25E | classic_bike | 10:14:53 | 10:24:29 | Western Ave & Walton St |
| 9624A293749EF703 | electric_bike | 08:27:44 | 08:32:40 | Racine Ave & Randolph St |
| 2F68A6A4CDB4C99A | classic_bike | 15:46:42 | 15:58:08 | Clark St & Leland Ave |
| 9AEE973E6B941A9C | classic_bike | 08:44:47 | 08:49:41 | Racine Ave & Randolph St |
| E366E997FDA1582B | classic_bike | 14:30:44 | 14:37:45 | Clark St & Leland Ave |
| 1BB3E73851E6C2C1 | classic_bike | 10:11:53 | 10:17:55 | Clark St & Leland Ave |
| DA1E1D0866E6566E | electric_bike | 21:57:27 | 22:08:27 | Clark St & Leland Ave |
| 39BF4A73A704CA85 | classic_bike | 10:51:17 | 11:03:13 | Warren Park East |
| F036C3340FF1EE0F | classic_bike | 13:27:47 | 13:46:45 | Jefferson St & Monroe St |
| 253D62F6B36B340C | electric_bike | 12:35:51 | 12:44:16 | Public Rack - Kostner Ave & D |
| 1AE843E603ED53A2 | electric_bike | 12:42:41 | 12:50:51 | Public Rack - Kostner Ave & D |
| 398A58DC66DFAD8A | electric_bike | 08:09:43 | 08:15:48 | Western Ave & Walton St |
| C87CB0A25C9682DF | classic_bike | 21:34:50 | 21:47:07 | Racine Ave & Randolph St |
| 70FFA226EBD6721F | classic_bike | 10:58:20 | 11:04:37 | Racine Ave & Randolph St |
| 00FE04677C756C6E | classic_bike | 16:14:46 | 16:27:26 | Racine Ave & Randolph St |
| B0A01513A1935927 | electric_bike | 17:00:59 | 17:06:08 | Western Ave & Walton St |
| A32770999843B987 | electric_bike | 18:29:28 | 18:32:28 | Clark St & Leland Ave |

## *Performing Descriptive Analysis:*

| | Column1 |
|---|---|
| Mean | 0.015096 |
| Standard Error | 0.000165 |
| Median | 0.007604 |
| Mode | 0.004109 |
| Standard Deviatio | 0.144918 |
| Sample Variance | 0.021001 |
| Kurtosis | 16321.48 |
| Skewness | 103.6987 |
| Range | 35.74017 |
| Minimum | -0.00309 |
| Maximum | 35.73708 |
| Sum | 11588.3 |
| Count | 767650 |
| | 0 |

## *Performing some calculations :*

| | |
|---|---|
| RATE_LENGTH (MEAN) | 0.015095812 |
| RIDE_LENGTH(MAX) | 35.73708333 |
| DAY_WEEK(MODE) | 7 |

## *Creating Pivot Table with desired column and Rows :*

The Pivot table is created with the set of functions, Here I found the relationship between the casual and annual and analysed the who owns a bike for a longer duration and on which days in the week the bike rented a most.

The table for the analysed data is attached below,

| Row Labels | Average of Ride_length |
|---|---|
| casual | 0.02245353 |
| member | 0.009507723 |
| **Grand Total** | **0.015095812** |

| Count of Ride_ | Column Labels | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Row Labels** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **Grand Total** |
| casual | 59460 | 48837 | 41237 | 27998 | 38762 | 44522 | 70542 | 331358 |
| member | 57365 | 73698 | 60460 | 53888 | 65363 | 59504 | 66014 | 436292 |
| **Grand Total** | **116825** | **122535** | **101697** | **81886** | **104125** | **104026** | **136556** | **767650** |

| Row Labels | Count of ride_id |
|---|---|
| 1 | 116825 |
| 2 | 122535 |
| 3 | 101697 |
| 4 | 81886 |
| 5 | 104125 |
| 6 | 104026 |
| 7 | 136556 |
| **Grand Total** | **767650** |

## ***Summary of the Analysis :***

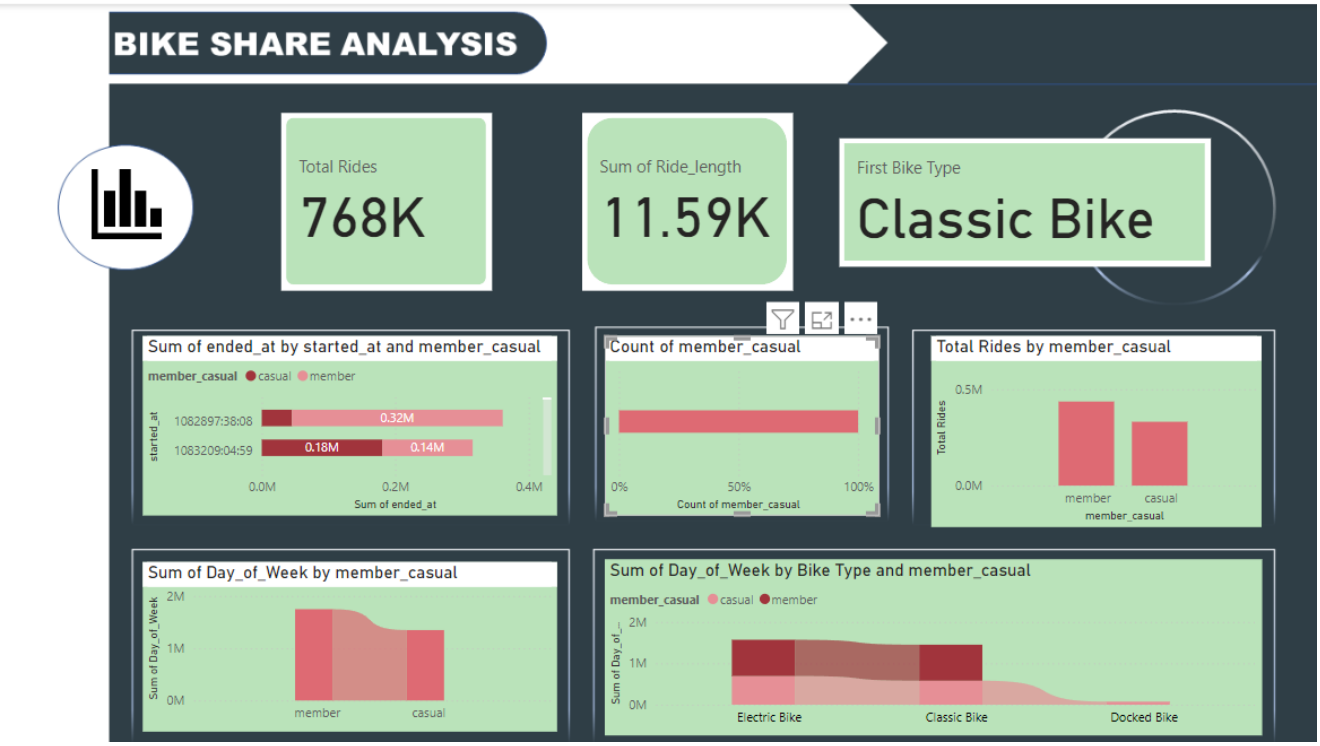
Through the “Descriptive Analysis” , I have analysed a set of functions for the given table and are listed below.

| | *Column1* |
|---|---|
| Mean | 0.015096 |
| Standard Error | 0.000165 |
| Median | 0.007604 |
| Mode | 0.004109 |
| Standard Deviatio | 0.144918 |
| Sample Variance | 0.021001 |
| Kurtosis | 16321.48 |
| Skewness | 103.6987 |
| Range | 35.74017 |
| Minimum | -0.00309 |
| Maximum | 35.73708 |
| Sum | 11588.3 |
| Count | 767650 |
| | 0 |

# SHARE

## *Visualization of the analysis :*

The Visualization is performed through the “Power BI”. Through this tool,we can analyse the count of the each member,number of days they rented,who rented a bike a lot.To the end,we got the answer of how the casual and the members differ from each other. A “Dashboard” is created with all of the visualisation elements.




## *Key considerations after Analysis:*

### Differentiation of casual and annual:

1. The “Annual riders” used the most when compared to the “Casual riders”
2. Especially,they rented a classic bike during the “Weekends”
3. The Casual riders share the bike only during the weekend or during the leisure times.

### How to Convert casual to a Annual :

1. Mostly, the consumers are attracted to the Discounts and Vouchers. So a sudden shrink of the annual members cost would merely change the casual to annual.
2. Changing the Payment options like Instalments.
3. Giving them a “Bonus Coins” for each ride and the amount of the coins collected,the sum of reduction of the amount.
4. The casual to annual converted members will be provided with the offer of “6 MONTHS free Rides.