# Data Analysis – Trends & Relationships

This is a documentation of trends & relationships found in the analysis step, a long with insights derived from them.

## Trend 1

* By viewing relationship 1 in a season-wise fashion, we can conclude that seasonal factors affect usage a lot and they affect casual users much more than member users, in cold seasons, casual users ride much less compared to warmer seasons.
* **Test :** To test this hypothesis, I obtained Chicago’s daily average temperature data from Internet(<https://www.visualcrossing.com/weather-history/Chicago,United%20States/us/2022-07-01/2023-06-30>) and did a correlation test with daily usage and daily average temperature.
  + The correlation coefficient between temperature and daily usage was 0.871 and 0.852 for member and casual users respectively. Thus, we can see that there is **a high correlation between temperature and daily usage** of all users.
  + Taking regression lines that use temperature to predict the total usage of the day, we get a slope of 198.4717 for member users and 219.0595 for casual users. Combining this with the information of correlation coefficients, we can say that casual users are affected more by the temperature(greater slope).

## Trend 2

* Casual users & member users have varying average usage distribution in a day.
* Both types of users max out their usage on around 5 pm.
* Member users have a lower peak ride usage at around 8 am as well as casual users, though casual users have a much smaller peak.
* Casual members have increasing ride usage from 4am to 5pm (the highest peak), while member users usage decrease after the first peak at around 8 a.m.
* The 2 patterns in daily riding clock-time does not vary much across seasons.
* Across all seasons, member users have a small peak ride usage at around 8 am while casual users don't. Both member and casual users have a highest peak ride usage at around 5pm (17:00). Casual members have increasing ride usage from 4am to 5pm (the peak).

## Trend 3

All seasons - huge discrepancy on rides distribution throughout the week

- member users take rides more often on weekdays, and less on weekends

2022-7 8 9 - member users peak at Wednesday and has a decreasing till Sun (usage frquency on Tue Wed Thu Fri) are very similar and close to each others

HOT casual users take rides more often on weekends, and less on weekdays

casual users peak at Saturday, then has a decreasing trend until Mon (huge peak on Sat)

2022-10 11 12 - member users peak at Thursday and has a decreasing trend until Monday

COLD casual users take rides more often on weekends, and less on weekdays

casual users peak at Saturday, then has a decreasing trend until Thursday

2023-1 2 3 - member users peak at Tues and has a decreasing trend until Friday

COLD casual users take rides more evenly throughout the week

casual users have 2 peaks(Tue,Sat)

the main discrepancy in this period occurs at Thu, Fri, Sat

2023-4 5 6 - member users peak at Thursday and has a decreasing trend until Monday

HOT casual users take rides more often on weekends, and less on weekdays

casual users peak at Saturday, then has a decreasing trend until Tue (huge peak Sat)

## Trend 4

* casual users switched toward classic bikes in 2023, while used a lot electric bike in 2022
  + maybe due to changes in supply of different types of bikes? -- more data needed

## Relationship 1

* There are more member users usage than casual users usage

|  |  |  |
| --- | --- | --- |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_| Usage Type  Season | | Member | Casual |
| 2022 Jul, Aug, Sep | 54.10% | 45.90% |
| 2022 Oct, Nov, Dec | 67.15% | 32.86% |
| 2023 Jan, Feb, Mar | 77.33% | 22.67% |
| 2023 May, Jun, Jul | 61.10% | 38.93% |

## Relationship 2

* member has slightly shorter & less variable ride duration on average compared to casual usage

|  |  |  |  |
| --- | --- | --- | --- |
| Season | Type | Mean (hour) | Stdv (hour) |
| 2022-Jul, Aug, Sep | Casual | 0.354097417 | 0.716872474 |
| Member | 0.201238386 | 0.339687871 |
| 2022-Oct, Nov, Dec | Casual | 0.248520837 | 0.67382876 |
| Member | 0.170573323 | 0.331441302 |
| 2023-Jan, Feb, Mar | Casual | 0.281862425 | 0.543842988 |
| Member | 0.184871561 | 0.317909974 |
| 2023-Apr, May, Jun | Casual | 0.352994105 | 0.595278897 |
| member | 0.216151471 | 0.306361379 |

## Insight 1

* From trend 1, relationship 1 & 2, I conclude that member users ride shared bikes for more stable & consistent activities, while casual users ride shared bikes for more versatile activities

## Insight 2

* From insight 1, trend 2, I conclude that member users ride shared bikes mainly for commutation to and from workplaces, whilst casual users ride shared bikes mainly for other casual activities.

## Conclusion

* By performing exploratory analysis in the data, I identified several trends & relationships in it. The main difference in member and casual user is that usage of our shared bikes for commutation weighs heavily among member users, whilst this proportion is relatively small in casual users.