

# Report: Citi Bike Ridership Difference between Generations

Linda Jaber<sup>1</sup>

<sup>1</sup>NYU Center for Urban Science & Progress

November 6, 2018

## Abstract

In an attempt to analyze how different generations use Citi Bike, an experiment was set to measure the difference between Citi Bike usage in terms of proportion of rides taken within a certain month between two generations —Generation X and Millennials. Recent research done at the Pew Research Center to define the age group each generation belongs to was used as a reference. The result of the analysis shows a significant difference in the usage of Citi Bike between the two aforementioned generations.

## Introduction

Citi Bike is a public bicycle sharing system serving New York City and Jersey City, New Jersey. This study investigates the difference in ridership between two generations. This analysis can help Citi Bike know the exact generation it needs to target. Figure one below shows how the distribution of the number of rides in March 2018 by birth year is negatively skewed indicating a higher usage in younger generations. Accordingly two young generations were chosen for this experiment based on the Pew Research Center ([DIMOCK, 2018](#)) definition of generational age groups, under the null hypothesis that the proportion of riders belonging to both groups is the same.

- **Generation X:** Born 1965-1980 (38-53 years old)
- **Millennials:** Born 1981-1996 (22-37 years old)

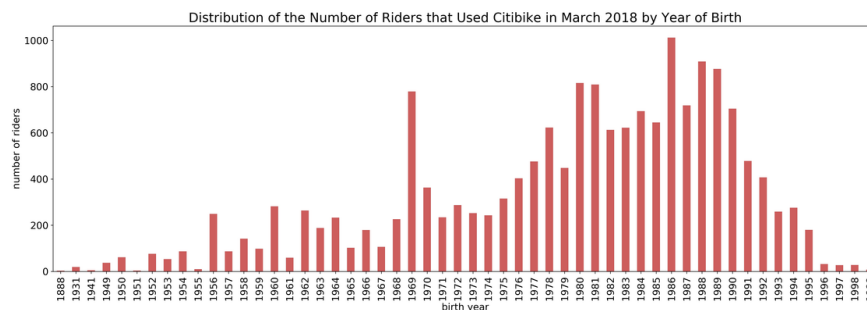


Figure 1: This is the distribution of the number of people that used Citi bike in March 2018 by year of birth. We notice that the distribution is skewed to the right. Therefore we can conclude that in that month people of smaller ages used Citi Bike more.

## Data

Citi Bike ridership data for the month of March 2018 was used as sample in this analysis. The data was grouped by birth year and only years from 1965 to 1996 were taken. The latter were divided into two groups:

- G1 > Gen X
- G2 > Millennials

I took note of my classmate review here as they pointed out the fact that the visualizations showed different generations than the ones defined.

## Methodology

To test the hypothesis and determine if there is a significant difference between the two proportions the two-tailed **two-proportion z-test** ([pro](#)) at significance level 0.05 was conducted. This approach is appropriate because the sampling method was simple random sampling, the samples were independent, and assuming each population was at least 10 times larger than its sample. My classmate review suggested that I use the Mann-Whitney U test, however the test best answers questions about rank differences where ordinal measurement scale is assumed which is not the case here.

## Conclusion

The test resulted in a large z-statistic that have a probability P-value close to zero. We concluded that the null hypothesis that the proportion of Generation X using Citi Bike is the same as that of Millennials is to be rejected. Obviously there is a significant difference in the pattern by which each generation uses Citi Bike. Millennials contribute to a higher proportion of the total rides taken in a month when compared to generation X. The analysis though doesn't account for difference in time in the year, so a better approach would be taking a larger sample that includes months from different seasons.

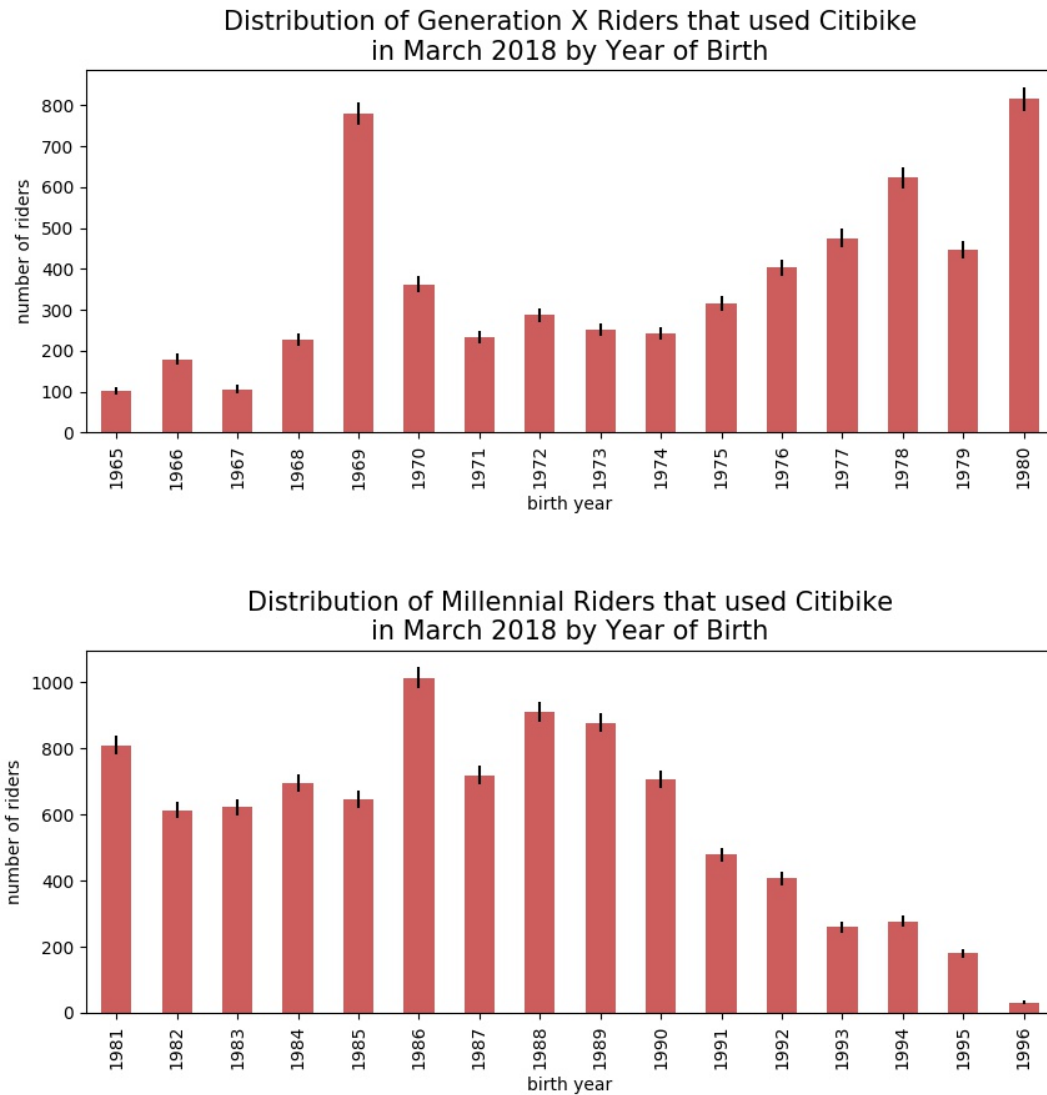


Figure 2: These are the absolute counts distributions of Gen X and Millennial riders respectively that used Citi Bike in March 2018 by year of birth. We notice that the number of riders differs between the two age groups, where it looks higher for Millennials that show a mean of 577 rides compared to a mean of 365 rides for Gen X. From here we want to do more analysis to see if this difference in numbers is significant.

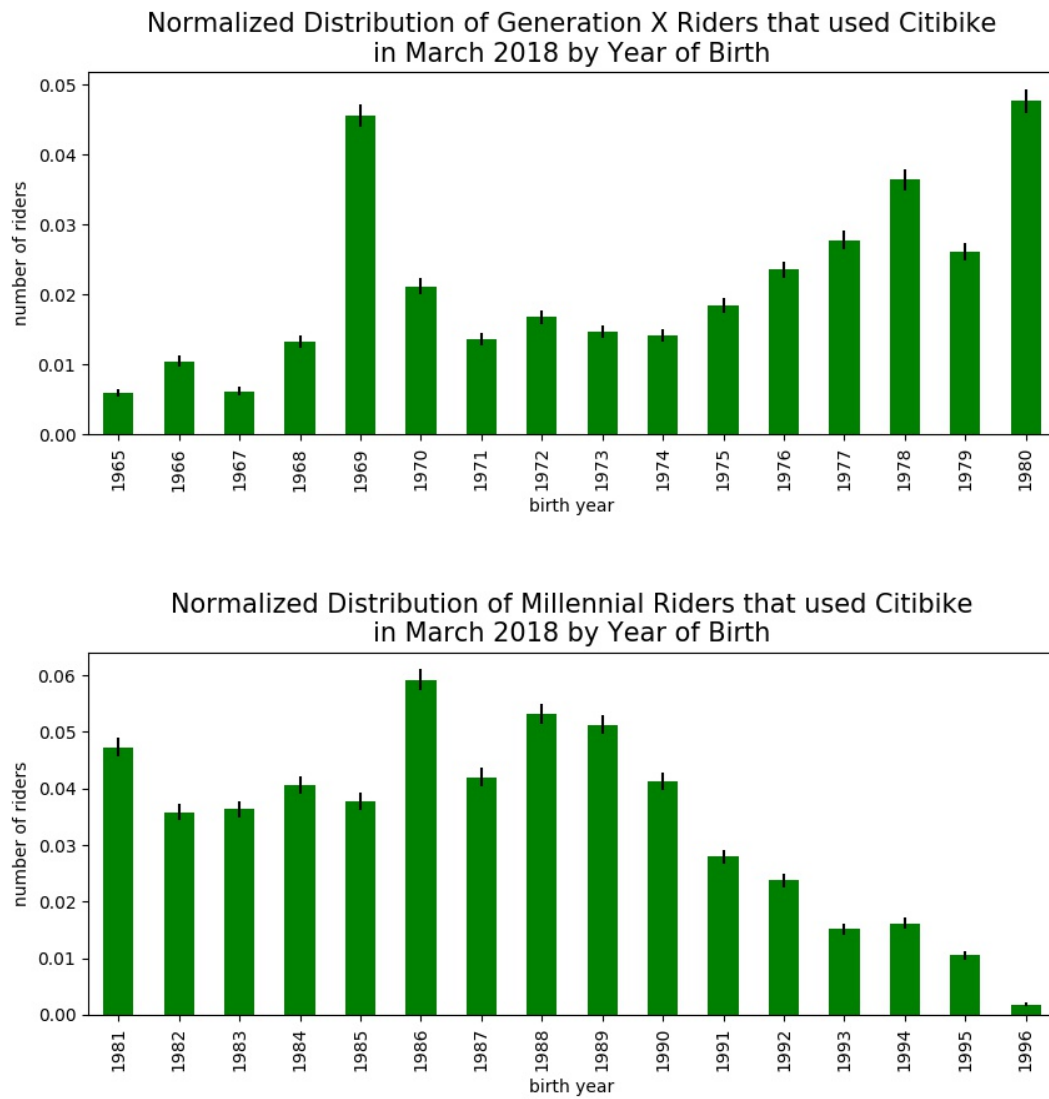


Figure 3: These are the normalized distributions of Gen X and Millennial riders respectively that used Citi Bike in March 2018 by year of birth.



Figure 4: Proportion of Citi Bike rides for two generations GenX and Millennials in March 2016. We notice there is a difference in the proportions and conclude the need for further testing to check if the difference is significant before we reject the null.

## References

Hypothesis Test: Difference Between Proportions.

MICHAEL DIMOCK. Defining generations: Where Millennials end and post-Millennials begin. 2018.