The Number of Marriages in Toronto Has Seasonality*

Zihao Liu

07 February 2022

Abstract

Statistics on Marriages has a long history, and the number of marriages tends to be more in a certain period of a year. Toronto, a city with a large population, can be a good object to study the marriages characteristics, this report uses the marriage licenses data in Toronto to do the analysis, and found that marriages tend to be more in summer, and less in winter. Also, the number of marriages is much lower after COVID-19. In addition, the findings could be used for economical and social studies, and forecasting the trending of marriages for related industries.

1 Introduction

The distribution of marriages is not uniform, people tend to get married in certain periods of time in a year(Thomas 1924). As early as 1838, the research and statistics on marriage began, throughout the research from 1838 to 1890, it was found that the number of marriages always reached its peak in the fourth quarter of a year and bottomed out in the first quarter of a year(Thomas 1924). Marriage is important to individuals, and number of marriages is important to the society. This report aims to study the association between marriages and time in Toronto.

This report analyzes the number of marriage licenses issued by civic centers in Toronto, and with the information, I can study how marriages change as time and season change. And I found that the number of marriage licenses issued in Toronto is seasonal. The number of marriage licenses issued in summer is much higher than in winter. In addition, the number of marriages dropped a lot after the start of COVID-19. This report reveals some of the relationship between time and marriages in Toronto, and the analysis can be used for social studies, planning of human resources at civic centers, and predicting the trending for wedding business.

The remainder of the report will set up and analyze the data using R programming language(R Core Team 2020). And R packages `tidyverse` (Wickham et al. 2019), `knitr`(Xie 2021b), `dplyr`(Wickham et al. 2021),`ggplot2`(Wickham 2016), `here`(Müller 2020),and `bookdown`(Xie 2021a) are used in this report.

2 Data

The R package `opendatatoronto` (Gelfand 2020) is used to import the data from Open Data Toronto Portal. The data is formed by the monthly marriage licenses issued at the civic centers in Toronto, it includes the record from January 2011 to January 2022. The data might not to be a complete representation of the marriages situations in Toronto, because it only keeps record of the number of marriage licenses, which are legally registered marriages. However, due to fact that every registered marriage is associated with a license, the preciseness of the data is high and the data is highly reliable. The data contains 485 observations and 4 variables, which are id, Civic center, Marriage licenses, Time period. In the process of cleaning data, id was removed, and two more variables were added, which are the month and the year of time period.

Table 1 is a sample of the cleaned dataset. Civic centers are the locations where marriage licenses are issued. Marriage license is the number of marriage licenses issued at the civic center in that particular time period.

*Code and data are available at: https://github.com/ZihaoLiu2/Toronto_Marriage_Statistics

Table 1: First 10 rows of the dataset of marriage licenses in Toronto

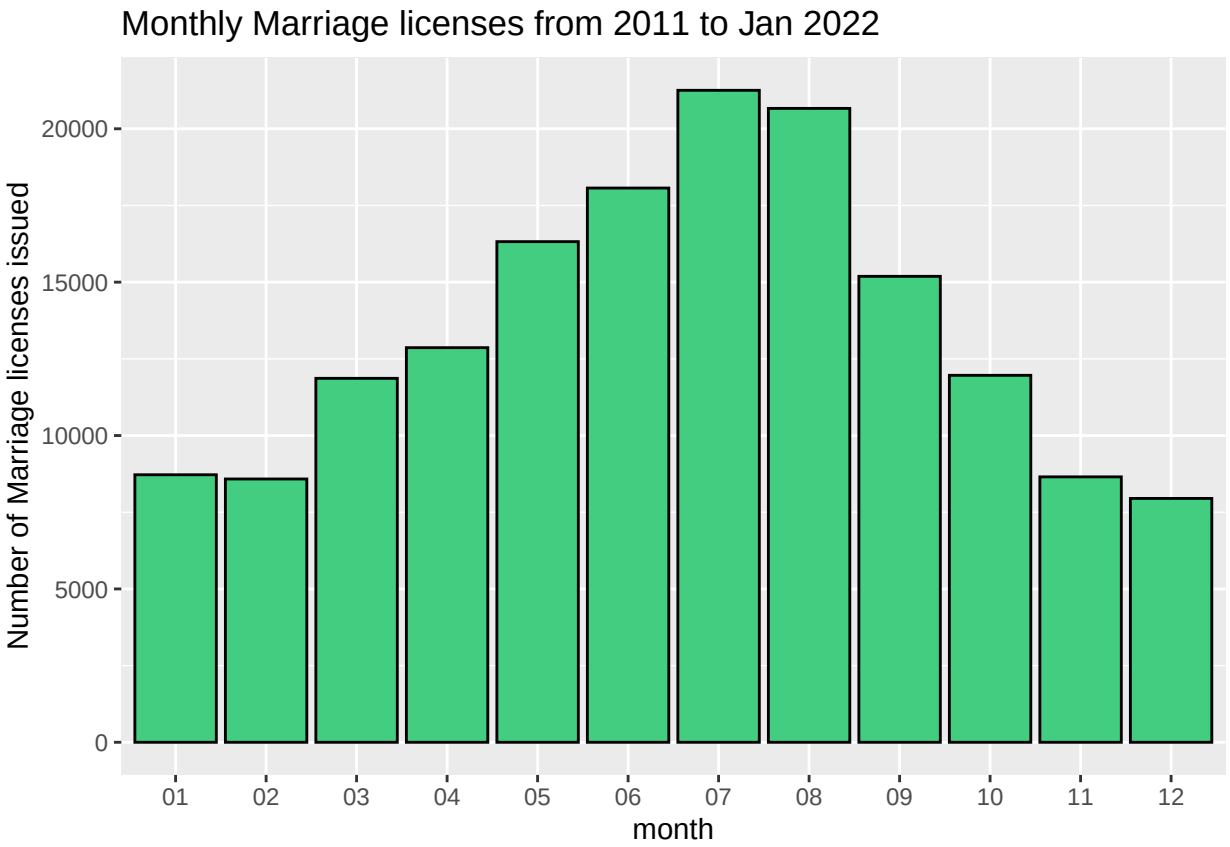
Civic center	Marriage licenses	Time period	Month	Year
ET	80	2011-01	01	2011
NY	136	2011-01	01	2011
SC	159	2011-01	01	2011
TO	367	2011-01	01	2011
ET	109	2011-02	02	2011
NY	150	2011-02	02	2011
SC	154	2011-02	02	2011
TO	383	2011-02	02	2011
ET	177	2011-03	03	2011
NY	231	2011-03	03	2011

Table 2: Yearly marriage licenses from 2011 to 2022 Jan

Year	Marriage licenses
2011	15644
2012	15497
2013	14375
2014	14892
2015	15513
2016	15756
2017	15884
2018	16330
2019	16225
2020	9743
2021	11758
2022	472

Each time period is one month long. Month and Year are just separated from the time period, so monthly total and annual total issued marriage licenses can be calculated more easily. The data contains the record from 2011 to 2022, so the first ten rows in the table above are the record in January, February and March of 2011.

Table 2 shows the total number of marriage licenses issued in a year from 2011 to 2022. From 2011 to 2019, the number is around 15000 for each year, and in 2018 and 2019, it has an increasing trend. From 2020 to 2021, the number of marriage licenses issued dropped a lot, the number in 2020 is only 60% of that in 2019. After the start of COVID-19, the number of marriage licenses decreased for nearly 40%, and the number recovers a bit in 2021. Since the data is updated to January 2022, there are only 472 marriage licenses issued in 2022 so far.



The plot above exhibits the number of marriage licenses in different months in Toronto. The number peaks in July and August, over 40000 marriage licenses were issued in July and August in Toronto in the past 11 years. And as time goes from September to December, the number gradually decreases and hits the bottom in December.Then, from January to February, the number stays at a fairly low level, and from March to June, the number increases gradually and peaks in July. It seems like the number of marriage licenses is influenced by the month changes, or in other words, the seasons. In the summer months, the number is much higher than in winter months. The number in July is more than the number in December doubled. For patterns of the seasonality, the number of marriage licenses is highest in summer, gradually decreasing in fall, hitting the bottom in winter, and gradully recovering in spring.

References

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Thomas, Dorothy S. 1924. "Changes in Marriage Seasons." *Economica*, no. 10: 97–106. http://www.jstor.org/stable/2547876.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*.

Xie, Yihui. 2021a. *Bookdown: Authoring Books and Technical Documents with R Markdown*. https://github.com/rstudio/bookdown.

———. 2021b. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. https://yihui.org/knitr/.