Seasonal Variation in Marriage License Applications*

An Exploratory Analysis with Open Data Toronto

Fatimah Yunusa

September 26, 2024

IN PROGRESS What was done:Marriage License data from Toronto between 2011 and 2024 was analysed to determine whether holidays influence the number of applications. What was found: Months associated with popular holidays like February (Valentine's day) and December(Christmas), show an increase in marriage license applications. Why this matters:Understanding these patterns can help city officials anticipate changes in demand for marriage-related services and allocate resources accordingly, ensuring that civil centers are prepaered for these seasonal shifts. It also gives us an insight into how emotional and culturl influences during holidays can drive personal descisions.

Table of contents

*Code and data are available at: https://github.com/fatimahsy/Pandemic-Marriages-.git.

1 Introduction

Marriage decisions are often driven but just in terms of practicality, but in terms of deeply emotional and sometimes irrational impulses. Symbolic dates like Valentine's Day and Christmas play a key role in this. Even during unprecedented times like the COVID-19 pandemic, love and tradition lead people to marry at significant moments, despite the logistical nightmares these choices presented. This paper examines trends in marriage licenses issued in Toronto from 2011 to 2023, with a focus on seasonal periods like February and December.

Wile there have been studies showing decreases in marriage licences during global crisis, there has been little analysis on how local trends are influenced by emotionally significant dates. This paper fills that gap by examining the effect of these periods on marriage decisions in Toronto, providing information on how tradition and emotion play a part in these decisions.

We use data analysis and various visualizations to check trends from 2011-2023 with a focus on special seasons. The findings show that while marriage licence declined significantly during the pandemic, Christmas and Valentines Day continued to see small spikes, showing that symbolic dates have some sort of power in marriage decisions.

Marriage trends are a very good indicator of changes in society, reflecting broader shifts in demographics, economic circumstances, and changing cultural preferences. Understanding how external factors like seasonal holidays and different disruptions to the overall society influence marriage decisions can offer us valuable insights into human behavior and societal recovery. By identifying consistent seasonal patterns, this analysis can help wedding-retated businesses (venues, caters, DJ's, etc.)anticipate periods of higher demand and better allocate resource. Additionally, civil centers can use this data to optimise staffing in order to increase efficiency.

The Data Section 2 of this paper explores the `opendatatoronto` data and outlines tools and methods used to analyze the data and some preliminary observations observed. The Results Section 3 introduces more observations found from the data analysis. The Discussion Section 4 ties back the findings to its real world relevance. Lastly, the Conclusion Section 5 summarizes key findings and outlines future areas of study.

2 Data

The dataset for this report consist of statistics for marriage licence quantities by civic centre's in Toronto from 2011-2023. This data comes from `opendatatoronto`, a public repository that provides access to different civil datasets. These statistics were downloaded, cleaned, parsed, analysed, and visualised using R (R Core Team (2024)), a statistical programming language, along with package support from `tidyverse` (Wickham et al. (2019)), various different libraries such as:

- `ggplot2` (Wickham 2016)
- `dplyr` (Wickham et al. 2023)
- `readr` (Wickham, Hester, and Bryan 2024)
- `tibble` (Müller and Wickham 2023)

For further cleaning, the `janitor` (Firke (2023)) package was used and the `knitr` was used too.

2.1 Variables of Interest

Variable	Description	Data Type
CIVIC_CENTRE	Code representing the civic center where the marriage license was issued	Character
MARRIAGE_LICENSES	Number of marriage licenses issued in that month	Numeric
TIME_PERIOD	The year and month when the licenses were issued, formatted as YYYY-MM	Date
ID	A unique identifier for each record	Numeric

Table 1: Variable Overview

The dataset has 544 rows and 6 columns. Each row represents the number of marriage licenses issued at a civic center during a given month, along with a corresponding year and a constructed date column. ((**marriage-variables?**)) describes the variables included in the dataset. These variables include:

- `civic_centre`
- `marriage_licenses`
- `time_period`
- `id`

The data is organised by moth and aggregated for each civic center. While it would have been beneficial to have data at the daily level, the monthly data still allows for meaningful analysis of seasonal trends.

Table 1: Sample of Toronto Marriage Licenses

ID	Civic Centre	Marriage Licenses	Time Period
15381	ET	80	2011-01
15382	NY	136	2011-01
15383	SC	159	2011-01
15384	TO	367	2011-01
15385	ET	109	2011-02
15386	NY	150	2011-02

2.2 Data cleaning

The raw data was cleaned to prepare it for analysis by removing columns, splitting the time information, and creating a date column. The cleaning process standardized the data making it easier to interpret. Here were the steps taken:

- Rename Columns:Column names were standardized using `clean_names()` for consistency.
- Split time_period into year and month: The time_period column, originally formatted as YYYY-MM, was split into separate year and month columns for easier manipulation.
- Create date column: A new date column was constructed using the `lubridate::ymd()` function.

2.3 Dataset Overview

Table 1 Shows a sample of the dataset used for the analysis of this paper. Here we can see `civic_centre` indicates the name of the civic center where the marriage licenses were issued (e.g., ET for Etobicoke, NY for North York, SC for Scarborough, and TO for Toronto).

In addition to the tabular data, Figure 1 shows the overall distribution of marriage licences by month. The boxplot illustrates the seasonal trends with the number of licences peaking in the summer months, which are particularly wedding months.

Table 2: Sample of Toronto Marriage Licenses

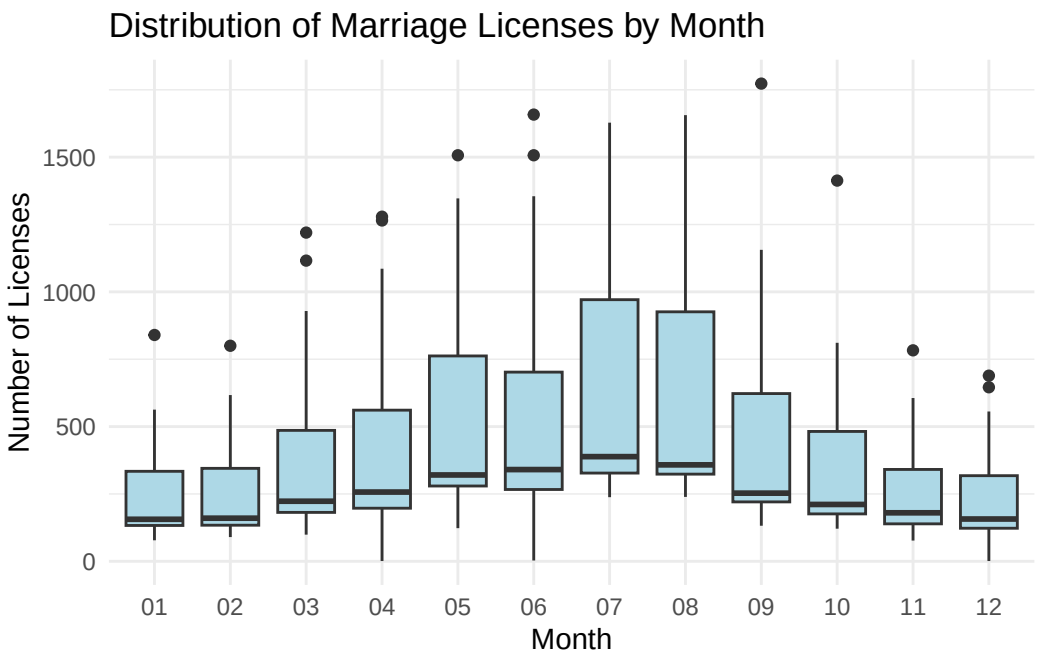

Figure 1: Figure 1: Distribution of Marriage Licenses by Month

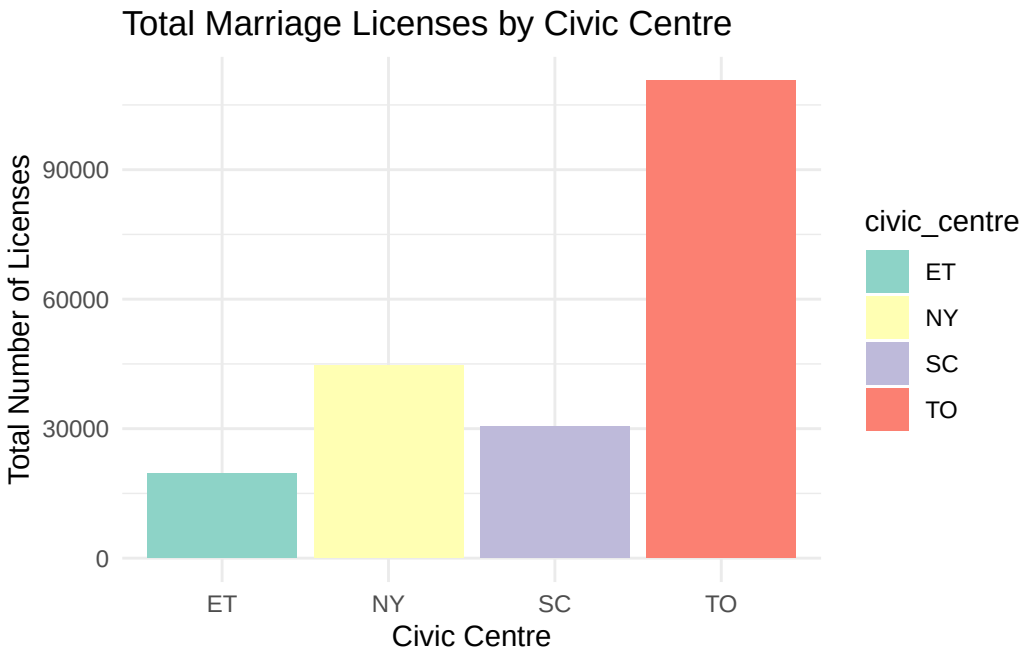

Figure 2: Figure 2: Total Marriage Licenses by Civic Centre

Figure 2 shows a breakdown of the total marriage licenses issued by each center from 2011-2023. Toronto issued the highest number of licences with over 90,000 licenses across the period, followed by North York, Scarborough and Etobicoke. The major disparity in marriage license counts across centers can mainly be attributed to differences in population density and geographical location. Toronto as the central and most densely populated area, attracts the largest number of couples applying for marriage licenses.

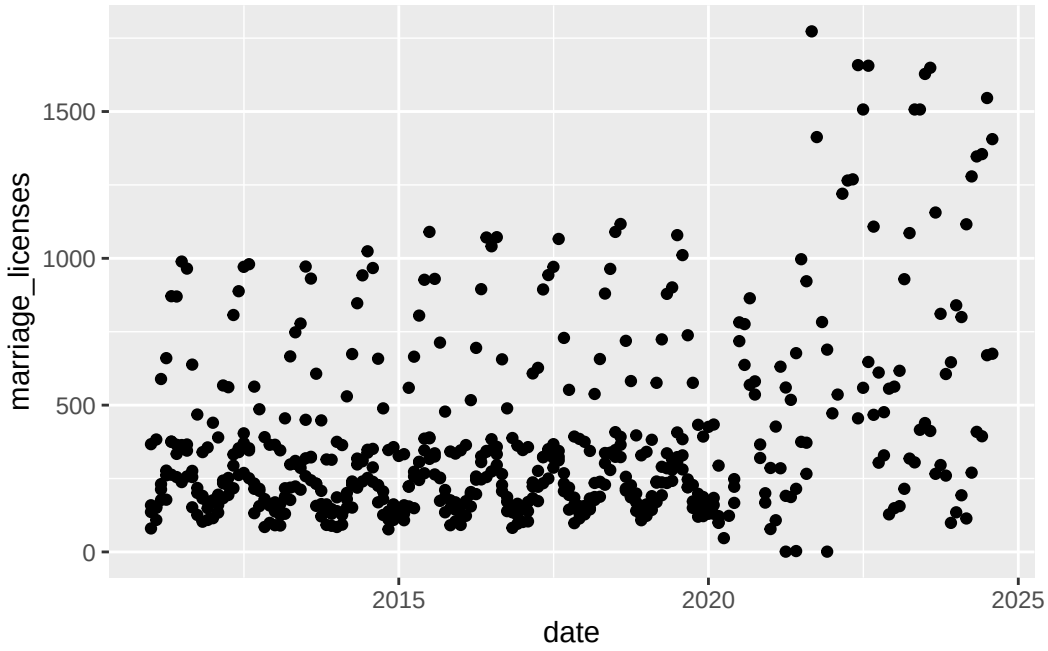

Figure 3: Marriage Licinces By Date

The scatter plot (Figure 3) displays the number of marriage licenses issued in Toronto over time, spanning from before 2011 to approximately 2025.Each point represents the number of liciences issued in a specific month. the scatter plot illustrates both the overall distribution of licenses over time and any significant deviations or disruptions to typical patterns.

3 Results

3.1 Licenses by Year and Month

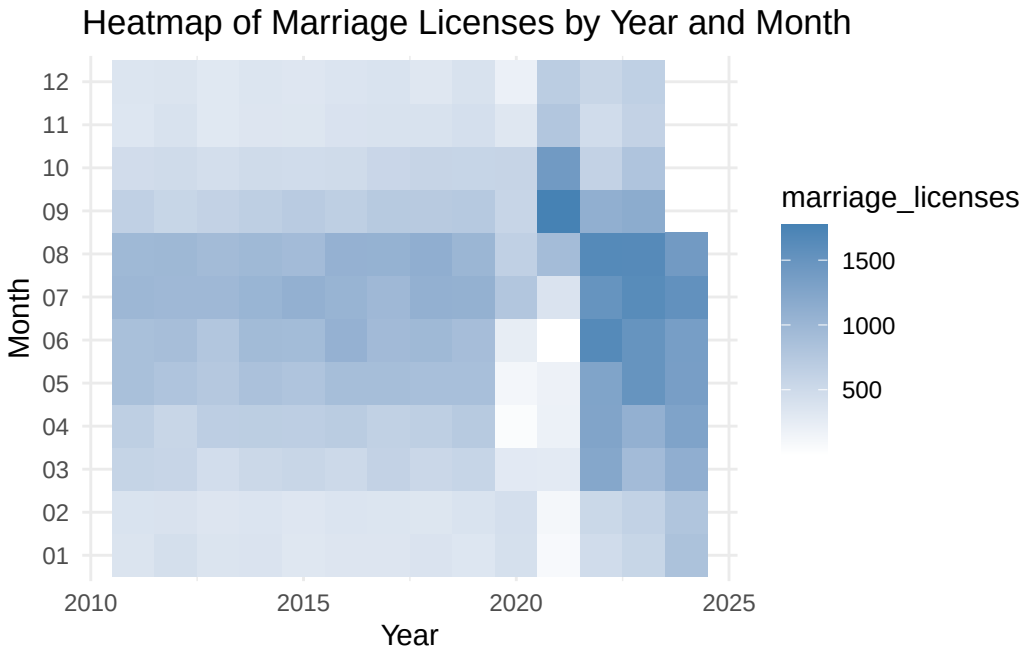

Figure 4: Figure 4: Heatmap of Marriage Licenses by Year and Month

Figure 4 shows a heatmap of the distribution of marriage licenses issued by month and year from 2022 to 2024. The x-axis represents years and the y-axis represents the months. The colour gradient illustrates the number of marriage licenses issued, with the darker shades representing more licenses issued. Higher volumes of marriage licenses are typically observed during the summer months(June-August) across all years, as shown by the darker shades on the heatmap. Winter months tend to have fewer marriage licenses issued. In 2020, there was a clear drop in the number of licences, particularly during the months that typically see the most activity.

3.2 Valentines Day Licenses

Table 3 shows the total number of marriage licenses issued in February for each year between 2011-2024.Contrary to my earlier belief, February is typically a low-activity month for marriage license, with most years seeing less than 900 licenses issued. However, in 2020, 908 licenses were issued in February, the ost in the whole period. They year 2021 shows a significant

drop in licenses, with only 535 licenses issued, reflecting the ongoing impact of the COVID-19 pandemic. By 2024, the number of february licenses reached 993, the highest for any February in the dataset, indicating a recovery and potential resurgence in wedding activity during this month.

Table 2: Table: Total Marriage Licenses Issued in February for Each Year

Year	February Licenses
2011	796
2012	879
2013	725
2014	717
2015	730
2016	830
2017	770
2018	820
2019	873
2020	908
2021	535
2022	536
2023	773
2024	993

Total Marriage Licenses Issued in February for Each Year

3.3 Pre,During & Post Pandemic Licenses

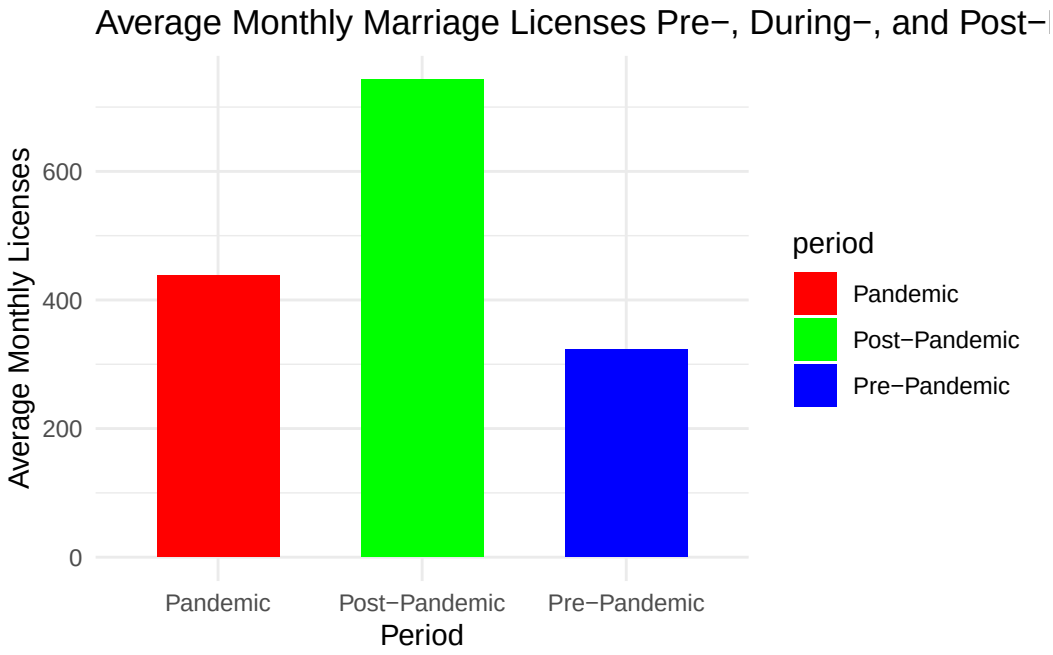

Figure 5: Figure: Average Monthly Marriage Licenses Pre-, During-, and Post-Pandemic

Figure 5 shows the average number of licenses issued per month during three distinct periods: pre-pandemic, pandemic & post-pandemic.Before the pandemic(2011-2019),there was an average of approximately 300 licenses issued per mont. During the pandemic, the red bar on the graph shows a significant drop in the average monthly licenses with approximately 450 licenses issued monthly. This reflects the overall decline in marriage due pandemic-related restrictions and disruptions. The green bar represents the post-pandemic period, showing an increase in marriage licenses issued with an average of 700 licenses per month. This suggest a rebound in normal activities. This hart clearly shows the disruption caused by the pandemic as well as later recovery in the following years, with post-pandemic averages exceeding pre-pandemic averages.

3.4 December Licenses

Table 3: Table: Total Marriage Licenses Issued in December for Each Year

Year	December Licenses
2011	785

Year	December Licenses
2012	785
2013	679
2014	785
2015	752
2016	747
2017	787
2018	717
2019	855
2020	368
2021	690
2022	684
2023	745

Table 3: Total Marriage Licenses Issued in December for Each Year

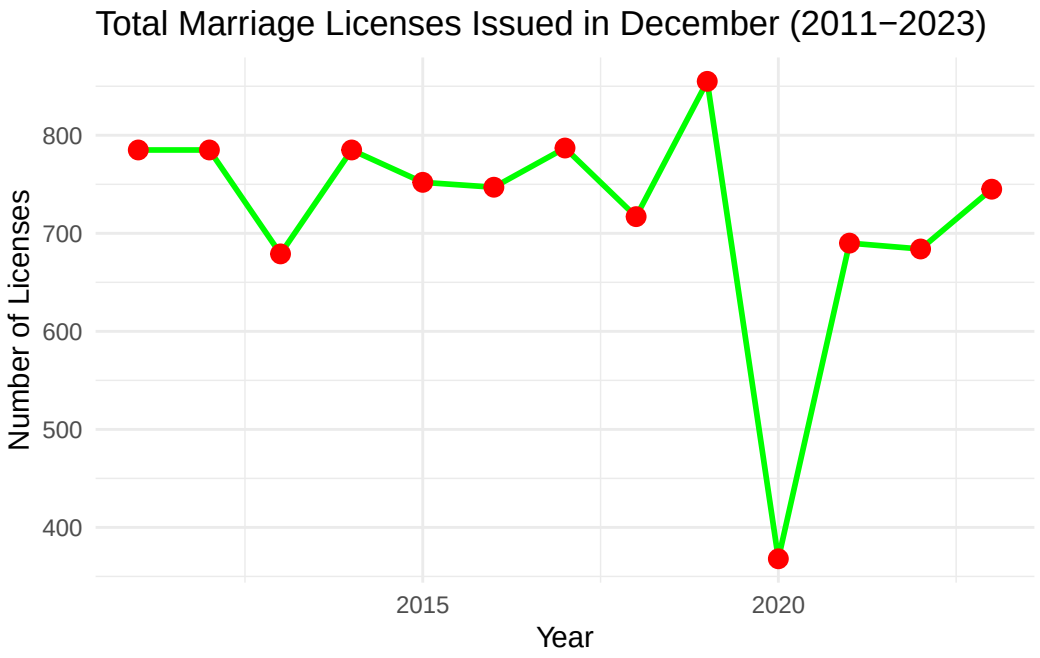

Figure 6: Figure 2: Total Marriage Licenses Issued in December (2011-2023)

3.5 Patterns

4 Discussion

5 Conclusion

‘‘

Appendix

Table 4: Table: Total Monthly Marriage Licenses by Year

Month	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
01	742	902	763	785	739	725	742	815	777	895	364	472	712	975
02	796	879	725	717	730	830	770	820	873	908	535	536	773	993
03	1210	1227	990	1062	1160	1073	1251	1137	1222	616	916	1220	1144	1230
04	1376	1232	1360	1294	1344	1395	1298	1322	1444	47	752	1265	1404	1549
05	1885	1650	1581	1682	1663	1774	1774	1748	1735	123	705	1269	1812	1756
06	1824	1843	1579	1806	1927	2011	1889	1918	1738	637	895	2113	1923	1749
07	1943	2015	1999	1962	2184	2047	1945	2169	2117	1500	1372	2066	2067	2216
08	1933	1930	1821	1845	1855	2059	2046	2197	2005	1413	1560	2303	2061	2081
09	1321	1143	1229	1280	1391	1356	1437	1400	1426	1433	1773	1575	1422	NA
10	1013	1065	940	1001	1010	986	1109	1182	1126	1117	1413	915	1107	NA
11	816	826	709	673	758	753	836	905	907	686	783	805	866	NA
12	785	785	679	785	752	747	787	717	855	368	690	684	745	NA

Total Monthly Marriage Licenses by Year

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

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, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2024. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.