Seasonal Variations in Marriage License Applications*

An Exploratory Analysis with Open Data Toronto

Fatimah Yunusa

September 27, 2024

Marriage License data from Toronto between 2011 and 2024 was analysed to explore trends in marriage licience number and the influence of emotionally significant dates like valentines day & christmas on these numbers. Couples are more practical in choosing wedding dates, with significantl higher numbers of marriage liciences issued during the summer months.These results hilighthow societal traditions around marriage are shaped more by practical considerations than romantic ideals. 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.

Table of contents

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

1 Introduction

Marriage decisions are often thought to be driven by deeply emotional and symbolic impulses, such as getting married on Valentine's Day or during the Christmas holidays. However, this paper finds that practical considerations like favorable weather and scheduling play a more significant role in choosing wedding dates. By examining marriage license data in Toronto from 2011 to 2023, this paper looks at trends in key periods such as February, the summer months, and December showing how couples tend to prioritize convenience over sentimentality, particularly favoring summer months for weddings.

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 or practicality. This paper fills that gap by examining the effect of these periods on marriage decisions in Toronto, providing information on how tradition, emotion, and practicality play a part in these decisions.

We use data analysis and different visualizations to examine marriage license trends from 2011 to 2023, with a focus on key seasons such as summer, Christmas, and Valentine's Day. While marriage license issuance declined significantly during the pandemic, these symbolic dates continued to show small spikes, indicating that emotional or symbolic significance still plays a role in some marriage decisions. However, the data leans more towards suggesting that practical considerations, such as weather and scheduling, are the primary drivers of marriage planning, with the summer months consistently seeing the highest activity.

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. By identifying consistent seasonal patterns, this analysis can help wedding-related 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. 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 and summarizes key findings and outlines future areas of study.

Table 1: Variable Overview

[!h]

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

2 Data

The dataset for this report consist of statistics for marriage licence quantities by civic center'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

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.(Table 1)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 2: Sample of Toronto Marriage Licenses

Table 3: 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 2)Shows a sample of the dataset used for the analysis of this paper.The dataset has 544 rows and 6 columns. 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.

(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.

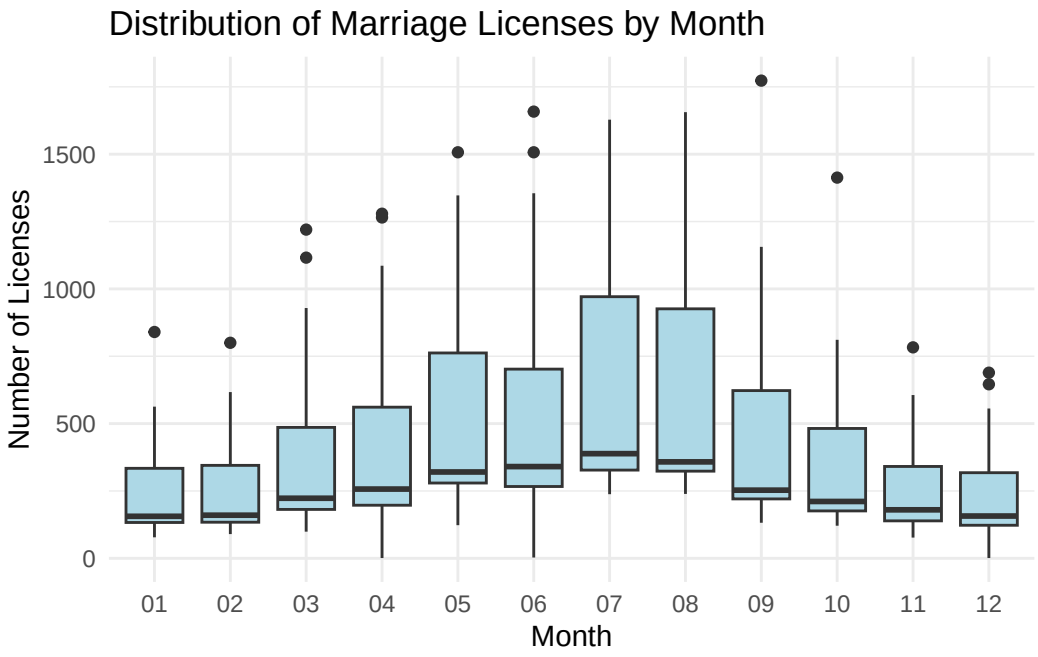

Figure 1: Distribution of Marriage Licenses by Month

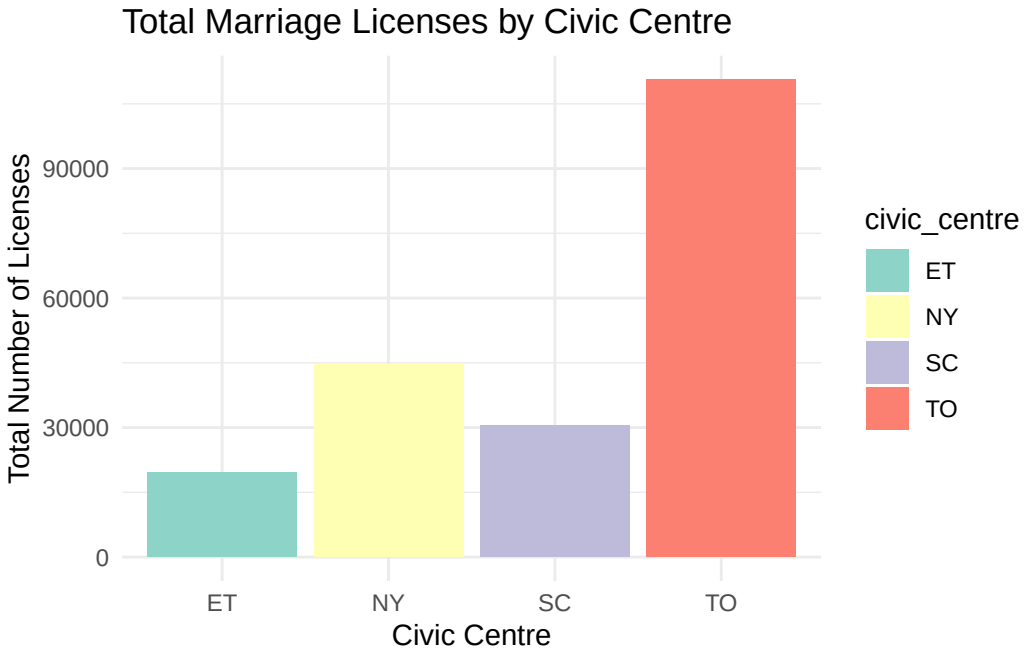

Figure 2: Total Marriage Licenses by Civic Centre

3 Results

3.1 Licenses by Year and Month

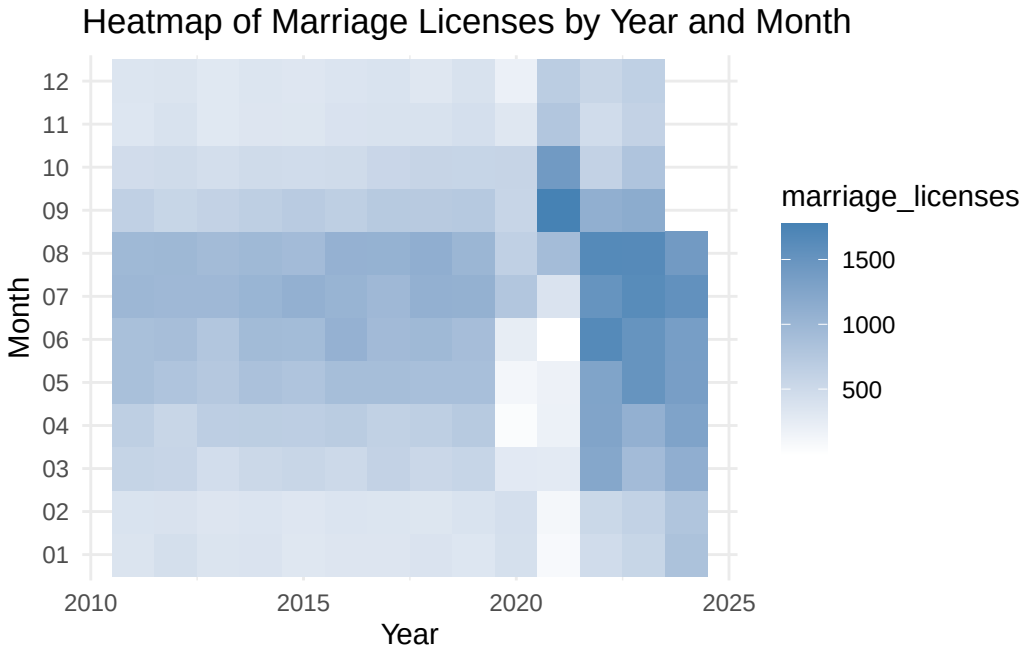

Figure 3: Heatmap of Marriage Licenses by Year and Month

(1) 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 4)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 4: Total Marriage Licenses Issued in February for Each Year

Table 4: 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

3.3 Summer Licenses

(2) presents the number of marriage licenses issued during the summer months(June, July, and August) for each year from 2011 to 2023. Each year is divided into 3 bars for each summer month, showing the distribution licenses within the summer sessions. July consistently has the highest number of marriage licenses issued compared to June and August, particularly in 2021 and 2022 where it reached over 2,500 licenses. 2020 shows a drop in marriage licenses with all three months experiencing reduced activity. July still led in numbers that year. A recovery period can be seen in 2021 and 2022 where the number of licenses exceeded pre pandemic levels. Generally, these months have higher numbers that other months in the calender.

3.4 December Licenses

(3) shows the total number of marriage licenses issued in December. The data shows that similar to other winter months, December typically experiences lower marriage license activity compared to the summer months. However, the numbers vary year by year

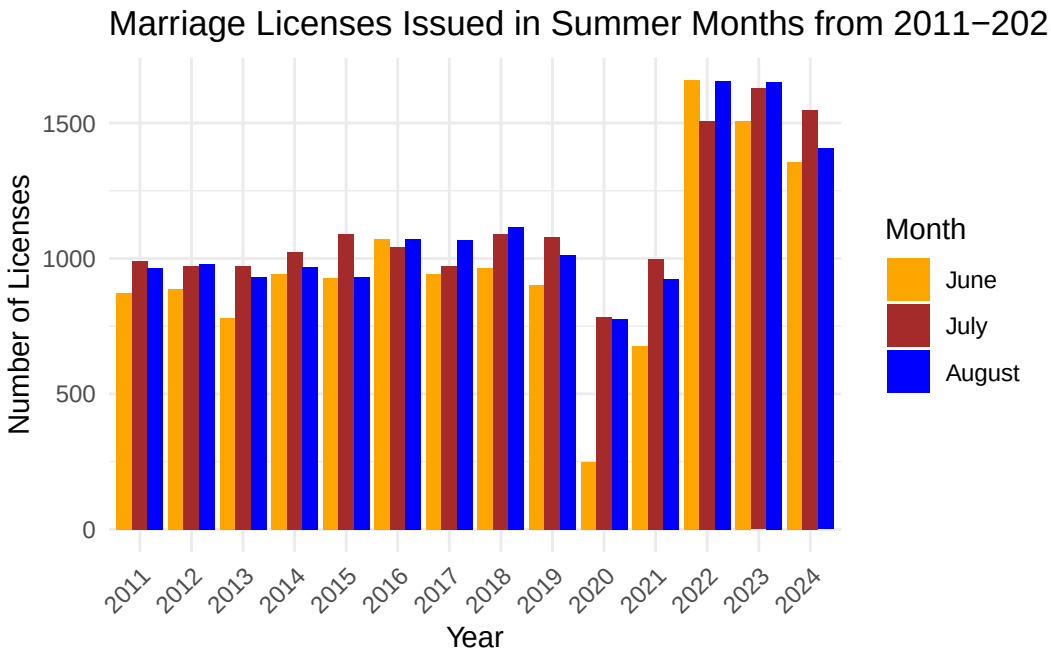

Figure 4: Total Marriage Licenses Issued in Summer Months from 2011-2023

with notable fluctuations.The graph shows relatively stable activity for most years with licenses ranging between 680-785.The highest number of licenses issued is December occurred in 2019, with a total of 855 licenses. The lowest number was recorded in 2020, during the height of the pandemic, when only 368 licenses were issued.

```
NULL
```

4 Discussion

Using Data about marriage licenses in Toronto, this paper tries to to see if there is a visible pattern that can suggest that people get married based on emotions or practicality. It focuses on emotionally significant dates like Valentines day, Christmas, and summer. The findings reveal that practical considerations, like weather and guest availability tend to outweigh emotional or sentimental motivations as seen in Section 3. There are consistent increases during the summer months compared to the relatively lower numbers on Valentine's Day or during the Winter.

These results challenge the assumption that romantic or symbolic dates drive marriage decisions, suggesting instead that couples tend to prioritize practicality when planning their weddings.

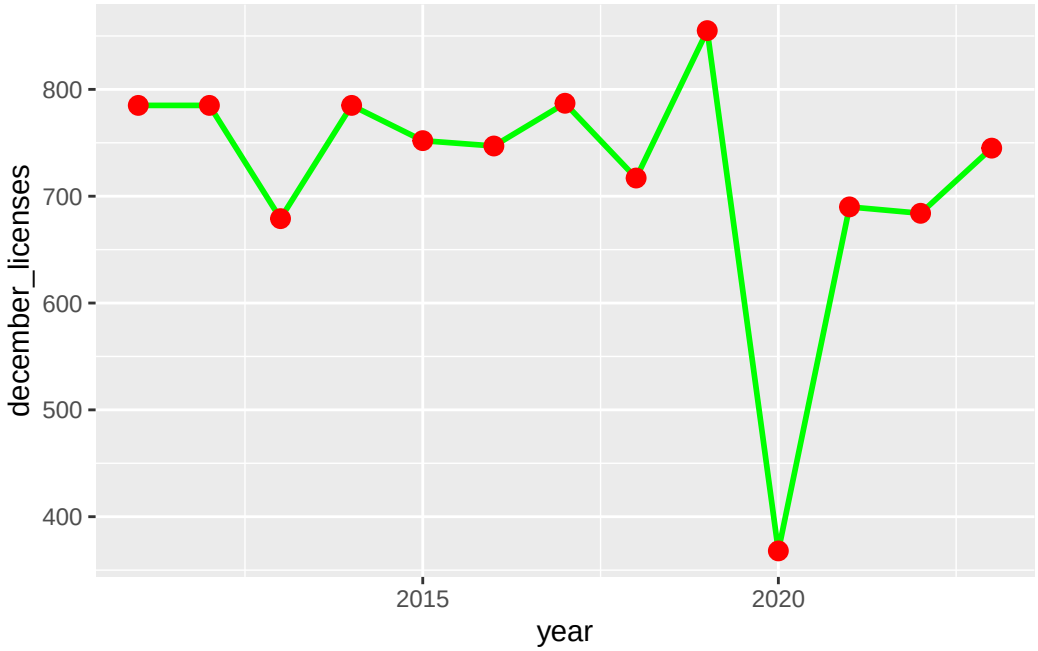

Figure 5: Total Marriage Licenses Issued in December (2011-2023)

One of the clearest patters shown from the data is the overwhelming preference for summer months when it comes t getting married. Across all years, these moths far exceed other months. This is largely attributed to the fact that during summer, logistics are much easier to work out. People are able to have weddings outside which is usually preferred. People also do not have to worry about bad weather like snow storms or heavy wind. Having a wedding during the summer also makes pictures look better since the sun is shining pretty well. In terms of guest availability, summer is usually optimal because there is no school, and many people take time off work. Even thought dates like valentines day hold cultural and romantic significance, they do not show a corresponding increase in marriage activity. This indicates that while couples may enjoy celebrating love or symbolic days, their wedding planning is more practically driven by factors like convenience, weather, and timing.

While the findings offer important insights into marriage license trends, there are a few limitations. First, the analysis is restricted to marriage license data which does not account for actual ceremonies or couples' personal motivations for choosing specific dates. While the data clearly shows summer dates are more popular, it cannot directly explain why couples make these decisions or weather other factors like venue availability or budget considerations play a role. Additionally, the dataset does not include demographic information like age, income, race or cultural background which could help provide a better understanding of why certain groups might favour particular wedding dates.

Future studies could have more of a focus on the part economic factors play in marriage licence

number.For example how does the cost of weddings or affordability of venues influence marriage licence data? Cross-regional comparisons could also show some valuable trends. Researchers could look into whether these practical considerations are consistent across different regions of if local traditions, economic conditions, or cultural preferences influence wedding planning.

‘‘

Appendix

Table 5: 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.