Growing number of afforbale housing units will have positive correlation with increasing number of people get marrige licenses in Toronto*

Jingyi Shen

January 23, 2024

With the development of economy, more and more people appeal that marriage should accompany with owning house which can provide a stable and safe living environment but also have a sense of economic security and asset accumulation. Home ownership can bring stability and a sense of belonging to both spouses and their future family. This report aims to investigate the relationship between numbers of Toronto afforable housing and marriage licenseduring 2020-2023. By exploring the data of different quarters of specific time period, I find that there is have positive correlation with increasing number of marrige licenses in Toronto and growing number of afforbale housing units.

Table of contents

1 Introduction **2**

2 Data **5**

3 Discussion **7**

*Code and data are available at:https://github.com/CSCmaster/paper1.

1 Introduction

For us, the dwelling serves as both a physical and mental environment, influencing aspects such as education, residential registration, and credit. The national marriage rate is on a continous decline, influenced by various factors. Apart from significant shifts in attitudes, there is a heightened cautiousness toward marriage, delayed onset of first marriages,and an increasing sense of independence among women. Another contributing factor is that individuals, both men and women, may have undergone wedding ceremonies without obtaining a legal license, often due to the pursuit of property ownership. Late marriage has become popular, although house prices are not the only reason, but it is a very important part of the factor. Housing prices affect every aspect of ordinary people,and the marriage rate is just another reflection of housing prices in our lives(Wrenn, Yi, and Zhang 2019). In this paper, I found data of Toronto marriage license and Toronto afforable housing during the period 2020-2023 in the open data of opendatatoronto. Then I draw a conclusion that there is positive correlation between the number of Toronto afforable housing and people get marriage license. I also find research done by statistician, they explore the correlation between housing prices, rents, and the marriage rate using data from Iranian provinces spanning the period 2002–2010. The investigation seeks evidence of a connection between elevated and persistent housing costs and the notable decline in marriage rates experienced by Iran over the last decade(Gholipour and Farzanegan 2015). # Data ##Marriage Licence Statistics In order to investigate the trend of people get marriage liscense in Toronto, I using the dataset "Marriage Licence Statistics" from from Toronto Open Data Portal(Gelfand 2020). This dataset provides details about the Toronto marriage liscense during 2020-2023. The dataset is generated to align with the Provincial legislations of the Marriage Act and the Vital Statistics Act. Additionally, it serves the operational needs and business functions of the City. This dataset contains 1095 observations of marriage liscenses with 5 variablels related to marriage liscense. This report focus on date, id, civic centre, marriages licenses, time period. By using packages"tidyverse"(Wickham et al. 2019), R(R Core Team 2020), "knitr"(Xie 2021), "janitor"(Firke 2021), "lubridate"(Grolemund and Wickham 2011), "dypler"(Wickham et al. 2021), "ggplot2"(Wickham 2016).

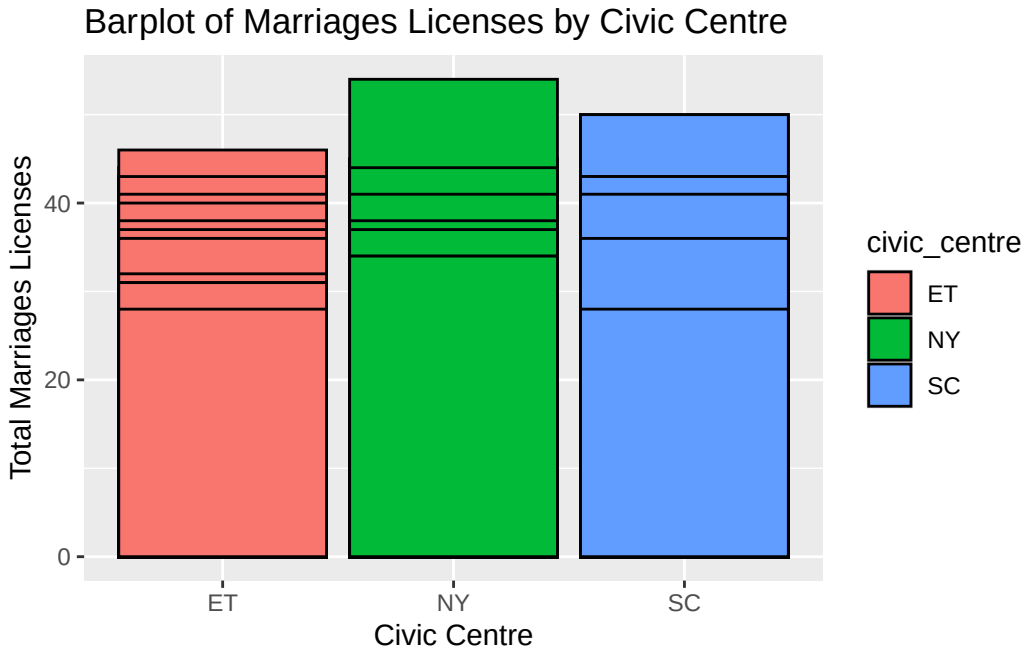
```
Rows: 1095 Columns: 5
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (2): civic_centre, time_period
dbl  (2): id, marriages_licenses
date (1): date

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
[1] 0
```

```
# A tibble: 6 x 5
  date          id civic_centre marriages_licenses time_period
  <date>     <dbl> <chr>                      <dbl> <chr>
1 2021-01-01     1 ET                            34 Q1
2 2021-01-02     1 ET                            32 Q1
3 2021-01-03     1 ET                            26 Q1
4 2021-01-04     1 ET                            24 Q1
5 2021-01-05     1 ET                            29 Q1
6 2021-01-06     1 ET                            42 Q1
```

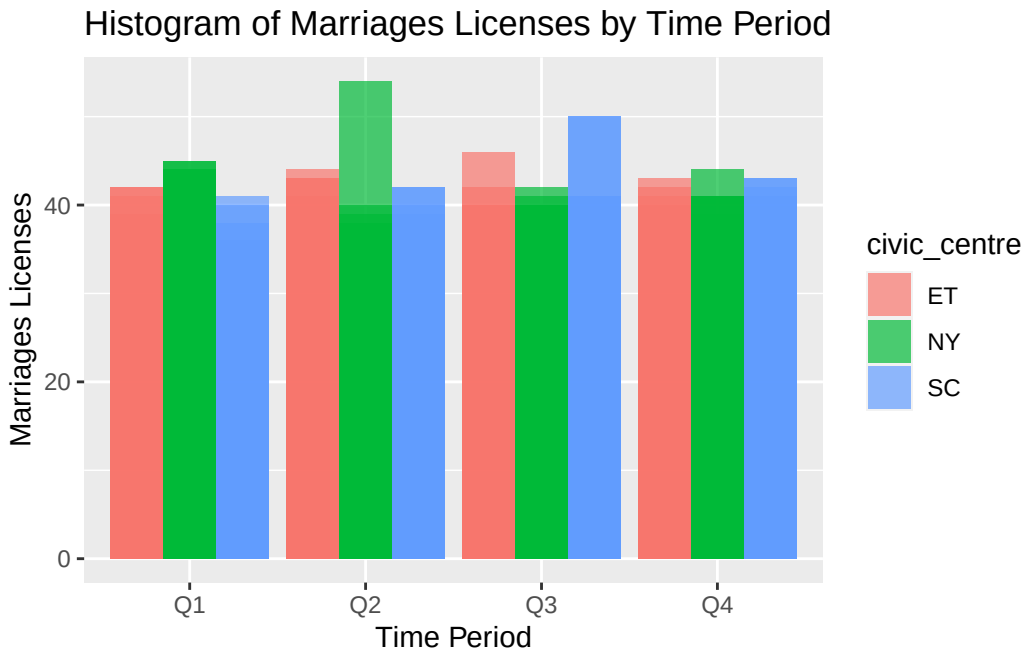
The table shows first 6 row of Marriage Licence Statistic data in Toronto. "id" in terms of a unique name for specific row in Open Data database. "civic centre" represents unique code for specific civic centre." marriage licenses" represents the number of people who get marriage licenses in Toronto. Time period is different quarters in the year that people in Toronto choose to get marriage licenses in.

```
Rows: 1095 Columns: 5
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (2): civic_centre, time_period
dbl  (2): id, marriages_licenses
date (1): date

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```



According to the barplot above, The x-axis represents civic centre, the y-axis represents marriages licenses, and the fill color is determined by civic_centre. In short, the graph is a grouped bar plot showing the total number of marriage licenses for each civic centre. The bars are grouped by civic centre, and the color distinguishes different civic centres. The plot provides a visual comparison of marriage license distribution across various civic centres. The maximum number of people get marriage licenses at the civic centre of NY which is around 58 while number of people get marriage licenses at ET and SC is 46 and 50 respectively.

```
Warning in geom_histogram(stat = "identity", position = "dodge", binwidth = 5,
: Ignoring unknown parameters: `binwidth`, `bins`, and `pad`
```



In conclusion, the graph is a histogram that visualizes the distribution of marriages licenses over different time periods, with bars grouped by civic centre. The x-axis represents quarters of a year , the y-axis represents numbers of marriages licenses that people got in each quarter in Toronto. The fill color is determined by the civic enter variable, ET represents in color red, NY represents color green, SC represents color blue. The shape of this histogram is slightly right_skewed with the maximum number of number of people (55) get marriage licenses is at Q2 which is around April to June in the NY civic centre. During the first and fourth quarter of the year, there are around 43 people get their marriage licenses in three different civic centre. There are around 45 people get marriage licenses in the third quarter.

2 Data

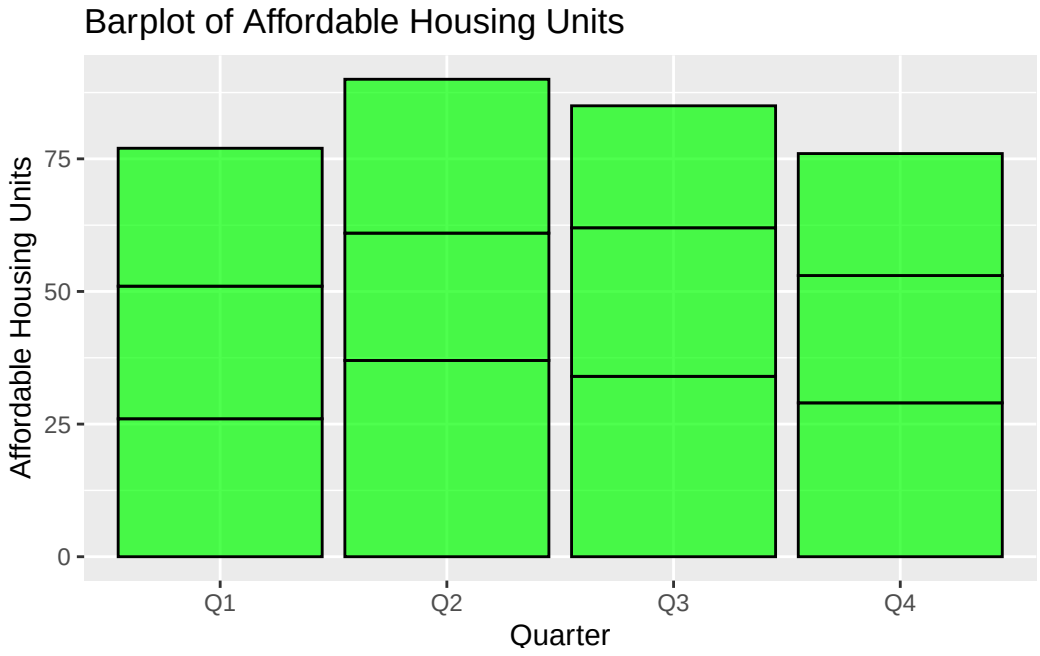
##Active Affordable and Social Housing Units This dataset of "Active Affordable and Social Housing Units" provides details about the currently managed social housing and affordable housing units within the Toronto's jurisdiction during the year 2020-2023 from Toronto Open Data Portal(Gelfand 2020). The City oversees social housing to ensure affordability for households with low to moderate incomes. Additionally, affordable housing in this context refers to rental residences under the City's management that align with the provincial criteria for affordability. This dataset contains 12 observational of affordable social housing units and 3 variables which are quarter, subsidized housing units, affordable housing units. By using packages"tidyverse"(Wickham et al. 2019), R(R Core Team 2020), "knitr"(Xie 2021),"janitor"(Firke 2021),"lubridate"(Grolemund and Wickham 2011),"ggplot2"(Wickham 2016).

```
Rows: 12 Columns: 3
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (1): quarter
dbl (2): subsidized_housing_units, affordable_housing_units

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# A tibble: 6 x 3
  quarter subsidized_housing_units affordable_housing_units
  <chr>                      <dbl>                     <dbl>
1 Q1                            17                        26
2 Q2                            25                        37
3 Q3                            12                        34
4 Q4                            20                        29
5 Q1                            27                        25
6 Q2                            22                        24
```

The table shows first 6 rows of affordable housing Units in Toronto. The predictor "Quarter" represents the four different time period of the year. Predictor "subsidized housing unit" in terms of residential properties constructed with government support which aimed at ensuring more economical rental options for low-income households, are overseen by the City. Social housing encompasses residences categorized as Rent-Geared-to-Income (RGI) housing with the typical RGI rent amounting to 30 percent of a household's monthly Adjusted Family Net Income. "Affordable housing unit" represents Residential properties managed by the City, where monthly rents covered utilities are equal to or less than 100 percent of the Average Market Rent (AMR). The Canadian Mortgage and Housing Corporation (CMHC) releases the AMR figures on an annual basis.

```
Rows: 12 Columns: 3
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (1): quarter
dbl (2): subsidized_housing_units, affordable_housing_units

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```



In a nutshell, the graph is a bar plot that illustrates the distribution of affordable housing units over different quarters. Each bar represents the count of affordable housing units for a specific quarter. The bars are filled with green color, outlined in black, and slightly transparent for better visualization. The plot provides a visual overview of how the number of affordable housing units varies across quarters. According to the graph, the maximum number of affordable housing units is 90 which is occuried at Quarter 2,the lowest number of affordable housing units is 75 which is occuried at Quarter1 and Quarter4.

3 Discussion

With the analysis of two dataset, I found that there is positive correlations between growing number of affordable housing units and increasing number of people get marriage licenses in Toronto. With comparison of figure 2 and 3, Quater 2 meets the maximum point of both number of affordable housing units and number of people get marriage licenses in Toronto while quarter 1 and 4 meet the lowest ponit of both number of affordable housing units and number of people get marriage licenses in Toronto, it shows that there is strong connections between housing units and marriage.

From my perspective, Owning a house gives me a sense of security. Having your own house can provide a stable living environment without facing housing problems caused by increasing rent or the expulsion of the landlord which can provide a stable sense of security. Owning a house is good for the stability of marriage and the growth of children. Having your own home can provide a stable environment for married life and the growth of children, which is conducive to the growth of relationships and children. If the long-term living conditions are unstable, it will

have a certain negative impact on these two aspects.However, in specific stage under personal and family conditions, short-term renting is also acceptable. The choice between owning and renting also needs to be made according to each person's unique situation.

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://sharlagelfand.github.io/opendatatoronto/, https://github.com/sharlagelfand/opendatatoronto/.

Gholipour, Hassan F, and Mohammad Reza Farzanegan. 2015. "Marriage Crisis and Housing Costs: Empirical Evidence from Provinces of Iran." *Journal of Policy Modeling* 37 (1): 107–23.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

R Core Team. 2020. *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*. https://CRAN.R-project.org/package=ggplot2.

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*. https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr.

Wrenn, Douglas H, Junjian Yi, and Bo Zhang. 2019. "House Prices and Marriage Entry in China." *Regional Science and Urban Economics* 74: 118–30.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.