

Anna tutorial week 3

Anna Li

1/22/2022

```
### Preamble ###  
#Purpose:Provide a glimpse at Toronto's marriage license registration situation.  
#Author: Anna Li  
#Email:annadl.li@mail.utoronto.ca  
#Date: Jan 22, 2021
```

Github

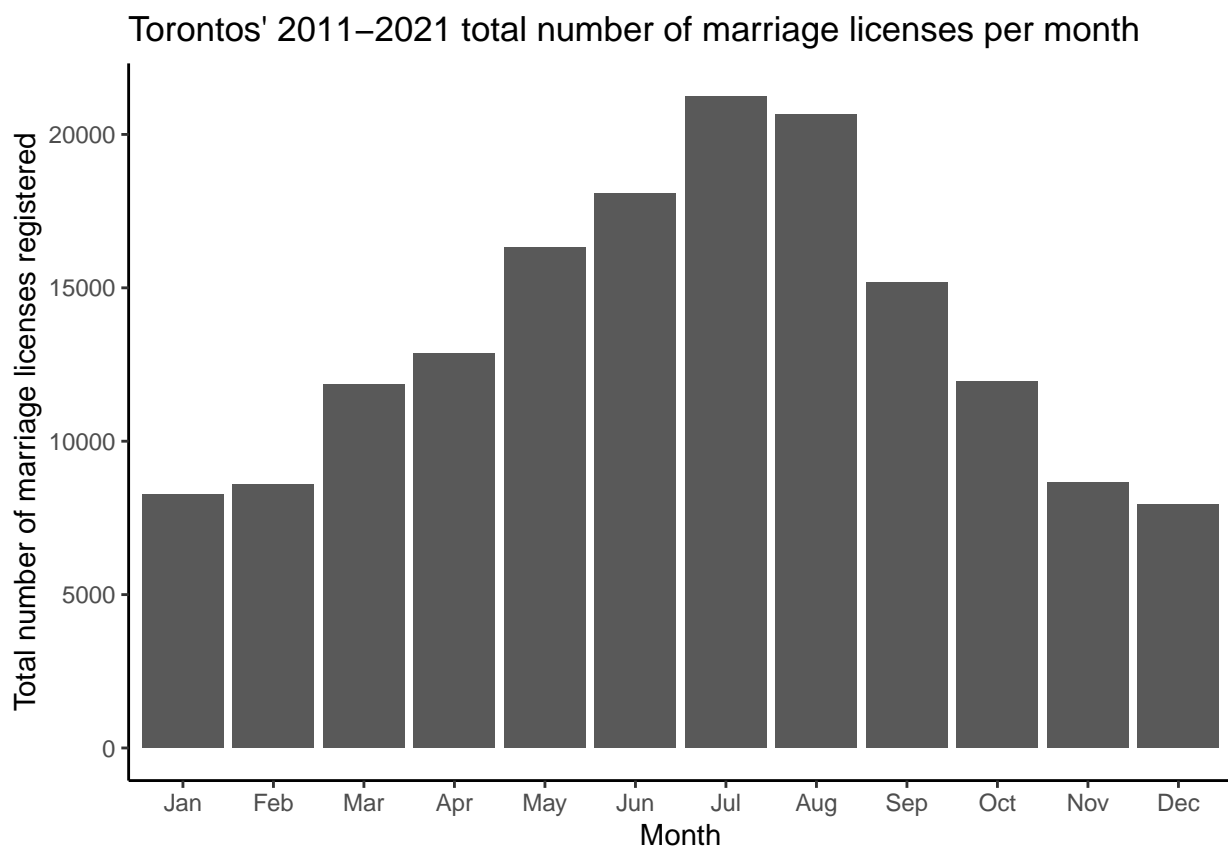
github repository link is <https://github.com/annadlli/STA304-week3tutorial.git>

Context about dataset

The dataset is released by the City Clerk's Office for the City of Toronto. It compiles information from 2011-01 to 2021-12 regarding marriage license statistics for Toronto. It compiles data from the 4 Civic Centres located in Scarborough, Etobicoke, North York, and Toronto.

```
#get and clean dataset  
dataset <- list_package_resources("e28bc818-43d5-43f7-b5d9-bdfb4eda5feb") |> get_resource()  
cleaned_dataset<- dataset %>% clean_names() |> select(marriage_licenses,time_period) |>  
  mutate(time_period = ym(time_period)) |>  
  mutate(time_period = as.integer(month(time_period)))
```

```
#Create graph  
cleaned_dataset |> group_by(time_period) |>  
  summarize(sum = sum(marriage_licenses)) |>  
  mutate(time_period = month.abb[time_period]) |>  
ggplot(aes(x=time_period, y=sum)) +  
  geom_col() +  
  labs(x="Month",y= "Total number of marriage licenses registered",  
       title ="Torontos' 2011-2021 total number of marriage licenses per month")+  
  scale_x_discrete(limits = month.abb)+  
  theme_classic()
```



Citations R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Sharla Gelfand (2020). opendatatoronto: Access the City of Toronto Open Data Portal. R package version 0.1.4. <https://CRAN.R-project.org/package=opendatatoronto>