Data analysis of dangerous dog order to find out which breed of dog and what area in the city is more dangerous by data collected from 2017-2023 from opendatatoronto*

Analyzing the circumstance of dog attack and compare data to find out the most dangerous breed and area from 2017-2023

Yuanchen Miao

September 24, 2024

In this paper, Registry of Dogs Subject to and Issued a Dangerous Dog Order is being analyzed and tested.After pandemic, more and more people choose to have a pet, most of them are dogs, dog attacks are becoming more often than before. The following paper discusses about the frequency of dangerous dog act along time and what breed of dogs are causing more dog attack issue in Toronto between 2017 to 2024. With the test to the dataset from open data toronto, dogs that are assigned as dangerous dog increased a lot after pandemic and bulldog and shepherd breed are showed to be more aggressive.

Table of contents

*Code and data are available at: https://github.com/NevaeH-9/dangerous-dog-attack.git.

1 Introduction

Dog attack is the most normal and minor harm to people in the neighborhood or society, but it is still causing a serious circumstances to individual safety and mental health. There are a lot of reasons for dog to attack people, under pressure, defending themselves, barriers, etc. Dog training is a must for dog owners to do, but some of the breed are showing more aggressive desire and some of the owners do not pay enough attention on dog training. Avoiding such breed or area is very important for people who wants to own a dog or government to reduce the number of dangerous dog act report and protect themselves.

There are a lot of people reporting they are getting attack by dogs every year, 2726 service requests are reported for potentially dog attack in 2023 (Draaisma 2023). Dog attack can cause a very bad and terrifying circumstance, if the dog is not vaccinated, it could carry lyssavirus and other serious virus with it, individuals have chance to get infected and have life threatened after getting bites by dogs. Getting attack by dogs will also leave a terrifying experience to individuals that causing these individuals panics in a period of time. On June 28, 2023, global news have reported that a 6-year-old girl bitten by two dogs in Toronto and it leaves her serious injured, the girl got a cut on her leg bleeding and the bite was really bad. (Global News 2023). There is also another report on February 9, 2024, A woman was seriously bitten by dog when she was waiting for bus and she was seriously injured on her left arm and face. She described the dog attack as "Scared of my life" (Ahmar Khan 2023). Dog attacks are very dangerous when it happens, individuals should take the threat more seriously. Thus, avoid breed with more aggressive desire and be more careful in higher dog attack frequency area should be considered.

In this paper, The data set **"About registry of Dogs Subject to and Issued a Dangerous Dog Order"** (Toronto 2023) that records the dogs that have been issued as dangerous dog and the reports of these dogs from 2017 to 2024 has been used to test which area in Toronto has more dog attacks and what breed of dogs are tend to have more aggressive desire, in other words, more dangerous to be close to these breed. This paper finds evidence to suggest that dogs with bulldog and shepherd gene are more aggressive and dog with bulldog gene are having worse bite circumstances. Owners of these breed of dogs should spend more time on training these dog and people should avoid getting too close to these dogs. The area with postal code "M4L" are having more dangerous dogs than other areas, people needs to be more careful when they are in this area.

The structure of the paper are as follows: Section 2.1 introduced the data, programming language and package used in the paper. Section 2.2 introduces that what has been done to clean the original data and variables we are using to do the comparison. Also, what has been done to reduced the variable in the original data. Section 3 are having tests and compare the variables from the data. Brief explain about the result we get after the tests. Section 4 discusses the tests in details and find results of dangerous breed and area. Section 5 discusses the limitation of the data and analyze. Introduce the next step to make the paper more reliable.

2 Data

2.1 Overview

The dataset **"About registry of Dogs Subject to and Issued a Dangerous Dog Order"** (Toronto 2023)has been used in this paper to analyze what breed of dogs are tend to be more aggressive and causing worse attack circumstances, and what areas in toronto are having more frequent dog attacks. Programming language R (R Core Team 2023) is been used in this paper to clean and analyze the dataset. Packages used to create graph and make the whole paper tidy are: tidyverse(Wickham et al. 2019), ggplot2 (Wickham 2016), dplyr(Wickham et al. 2023), here(Müller 2020) and knitr (Xie 2023). Package opendatatoronto (Gelfand 2022)are used to download the dataset to R. Codes are adapted from Tellingstorieswithdata (Alexander 2023)

2.2 Clean data

The raw data downloaded from opendatatoronto (Gelfand 2022) have no variables showing a "NA" and none of these data showing 0. Since the paper only cares about the bite circumstances after attack, dog breed and forward sortation area, so other columns are removed from the raw data. There are a lot mix breed contained in the original data, these breeds are being consiered as the same main breed, for example, "LABRADOR RETR / BORDER COLLIE" are considered as "LABRADOR RETR". Only the first 3 digits of the postal code are being recorded in the data, thus, the range of the area being analyzed will be as same as the range of 3 digits of the postal code. Table 1 is giving first 5 rows for a brief view of cleaned data.

Table 1: Cleaned data

Breed	Bite_Circumstance	Date_of_Dangerous_Act	Forward_Sortation_Area
BEAGLE	VERY SEVERE	2017-03-09	M6R
AKITA	SEVERE	2017-03-11	M3M

Table 1: Cleaned data

Breed	Bite_Circumstance	Date_of_Dangerous_Act	Forward_Sortation_Area
CANE CORSO	SEVERE	2017-03-18	M4N
GOLDEN RETR	SEVERE	2017-03-23	M9W
SHEPHERD	NON SEVERE	2017-03-29	M1E

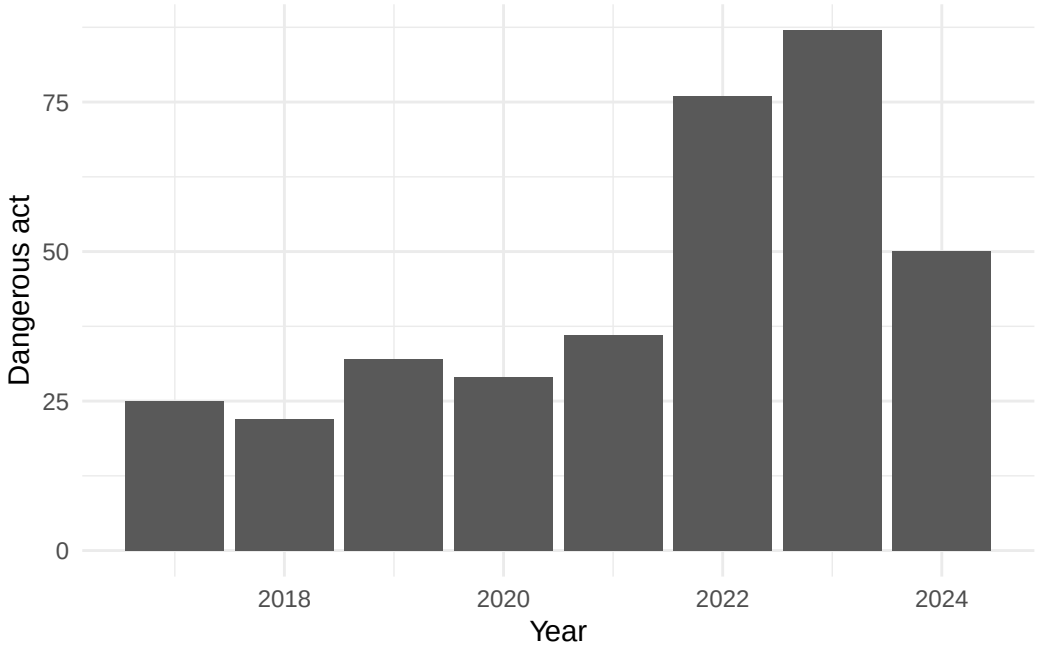

Figure 1: Number of dogs that have been issued as a dangerous dog in Toronto

How many dogs are issued as dangerous dog each year are the first thing we are interested in. From the graph in Figure 1, we see that more dogs are being issued as dangerous dog after 2022 which is the year of pandemic. People choose to own a dog as pet to bring happiness and reduce the feeling of lonely and anxiety about the pandemic period. The frequency of dog attack increased since then. There could be a lot of different reasons, the number increase of dogs, lack of training for dogs, more pressure for dogs than before. Since the data contains a huge amount of different observations, we do not need to worry about the issue caused by lack of observations.

3 Results

3.1 Finding the most dangerous breed by comparison

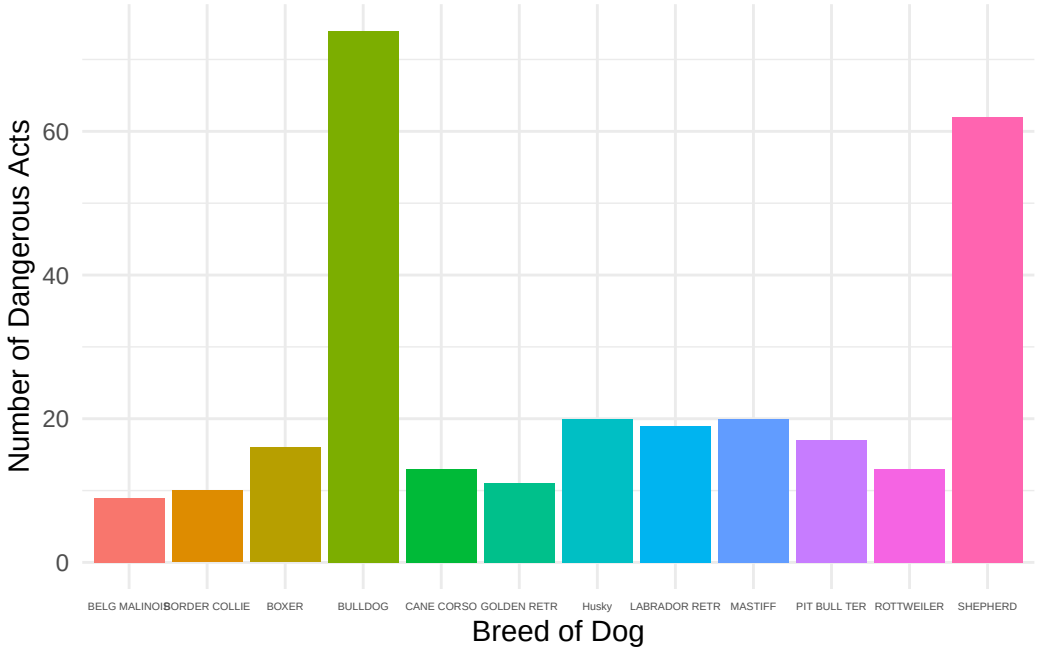

Figure 2: Breed of dogs that has issued as dangerous dog above the mean number

By integrate all the mix breed into the same breed, we have reduced the number of breeds in the data set that we could find out the what kind of breed are showing higher aggressive desire overall. From Figure 2, we see dogs with bulldog or shepherd gene are being assigned as dangerous dogs more than other dogs with other genes. There are over 70 dogs with bulldog gene have been issued as dangerous dog and over 60 dogs with shepherd gene are issued as dangerous dog.

3.2 Comparison on bite circumstances

Through Figure 3 we see that over 200 bite circumstances are severe and over 25 of them are very severe. Bite circumstance is the point how we are going to compare each dogs and getting the result of the most dangerous dog breed.

By doing the further comparison in bite circumstances, we focused on severe and very severe bite circumstances to see what breed are causing most of these situation. The comparison only consider the breed that has been issued as dangerous dogs for severe and very severe bite

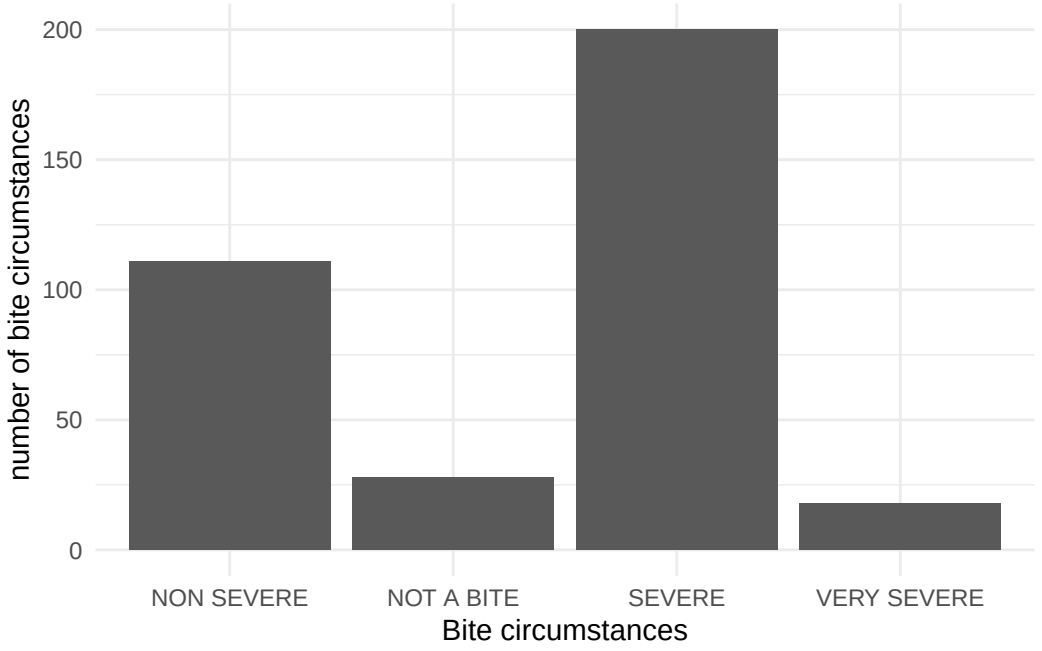

Figure 3: Number of different attack circumstance

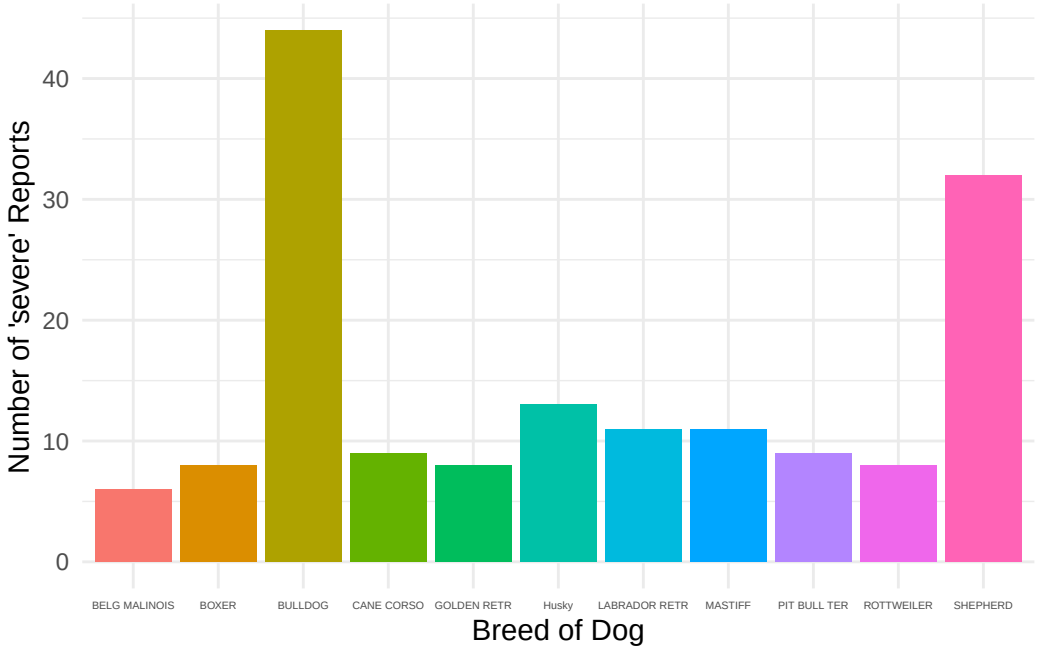

Figure 4: Number of severe each breed of dog caused above mean

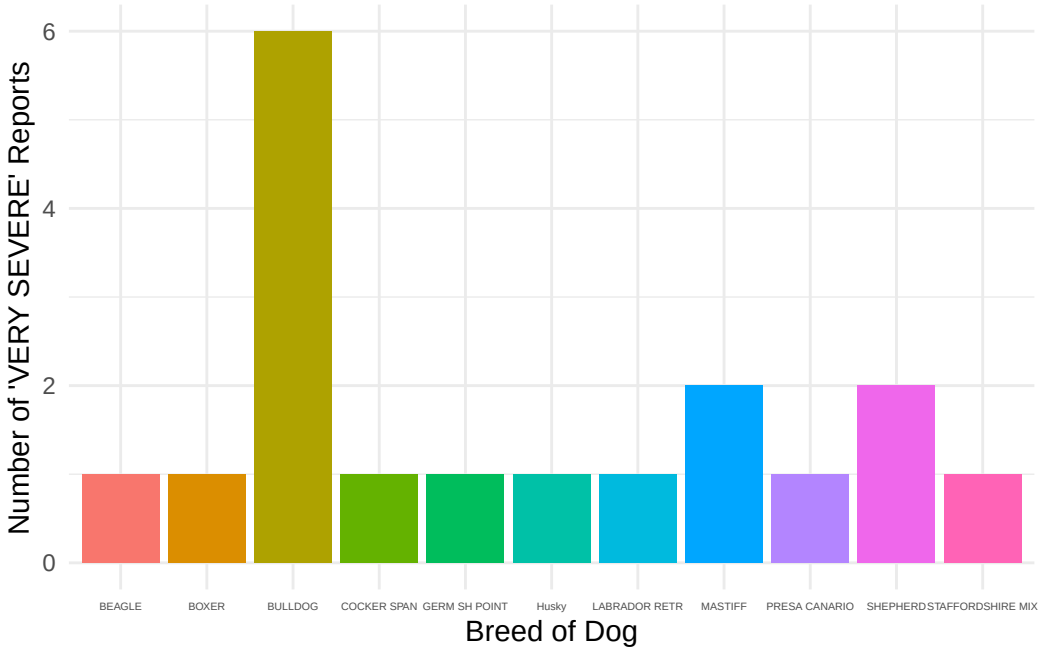

Figure 5: Number of very severe each breed of dog caused above mean

circumstances above the mean number of dogs that have been issued. From Figure 4 we see that bulldog have been report for severe bite circumstances over 40 times and shepherd over 30 times. Bulldog is also having the most number of very severe reports as very severe in Figure 5 for 6 times. Mastiff and shepherd are the second highest as 2 times.

3.3 Finding the area with highest number dangerous dog issued

Table 2: Areas of dog attack above the mean

Forward_Sortation_Area	dangerous_dogs
M1E	12
M1W	8
M2J	12
M4J	12
M4L	14
M6M	8

Finally, we are doing comparison between areas with first 3 digits of postal code, Figure 6 shows the first 3 digits postal code has more than 7 dangerous dogs and “M4L” is having the

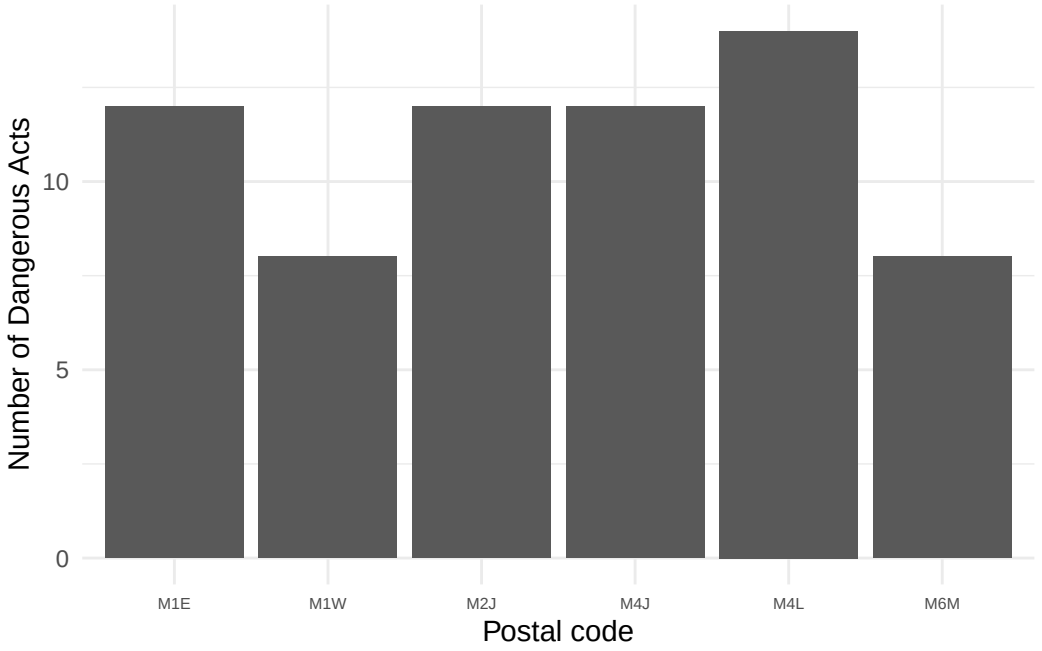

Figure 6: Areas of dog attack above the mean

most of these. Table 2 presents that “M4L“ has 14 dogs issued as dangerous dogs and”M1E”, “M2J”, “M4J” all have 12 dogs have been issued.

4 Discussion

Toronto has scaled level of bites from 0 to 5(City of Toronto 2024), where 0 is not a bite and 5 refers to victim is deceased as a result of bite or attack. level 4 stands for multiple-bite incident. Thus, we are connect level 1 as non severe, level 2 as severe and level 3 as very severe. level 2 stands for one to four punctures from a single bite with no deep punctures and level 3 is one to four punctures from a single bite with at least 1 deep puncture and dog clamps down without a quick release. Combining this evidence with Figure 4 and Figure 5, and the overall counts of dog attack above mean from Figure 2. We are having enough evidence to claim that dogs with bulldog gene are most dangerous and they are expected to have more aggressive desire among all of the breed of dogs. The shepherd breed is second highest which means the breed of shepherd is also having a higher chance to attack people or dogs. Interestingly, the breed of mastiff is not issued as dangerous dog in a huge number overall but it has been report 2 times of very severe bite circumstance. The owner of mastiff should also take more care about the dog. By comparison in different area in Figure 6 and Table 2, we find 4 areas that have been issued over 10 dangerous dog, people should be more careful when they are

around these areas. Government should take more actions on breed like bulldog and shepherd to prevent more dog attacks happen in Toronto and these areas with over 10 dangerous dog issued needs more attention when people are walking their dogs to avoid dogs get scared or nervous which ends as dog attack

5 Limitations and next step

When doing the data cleaning, since there are a lot of different breed mix in the data, these breeds are integrated into one like mentioned, this will cause inaccurate when compare the dangerous that is issued in Figure 2, Figure 4 and Figure 5. Which could do is compare these different mix of breed separate which makes the result more reliable.

In Figure 6, there could be a lot of reasons why that area is having more dangerous dog issued, higher population than area, more people are living there or there could be a dog park in the area that a lot of people bring their dog from other place. To make this analyze and result more accurate, more details about the area will be needed.

References

Ahmar Khan, Ryan Rocca &. 2023. "Woman Dragged to Ground, Suffers Life-Altering Injuries in Toronto Dog Attack." https://globalnews.ca/news/10284422/woman-dragged-to-ground-life-altering-injuries-toronto-dog-attack/.

Alexander, Rohan. 2023. *Telling Stories with Data*. Boca Raton: CRC Press. https://tellingstorieswithdata.com/.

City of Toronto. 2024. "Dog Bites or Attacks." https://www.toronto.ca/community-people/animals-pets/pets-in-the-city/dogs-in-the-city/dog-bites-or-attacks/.

Draaisma, Muriel. 2023. "Toronto's New Rules for Dangerous Dog Owners." https://www.cbc.ca/news/canada/toronto/dangerous-dog-owners-new-rules-city-toronto-1.7183803.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Global News. 2023. "Child Bitten by Dog at Midland and Lawrence in Scarborough." https://globalnews.ca/news/9798318/child-bit-dog-midland-lawrence-scarborough/.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Toronto, City of. 2023. "Dogs Issued Dangerous Dog Orders." https://open.toronto.ca/dataset/dogs-issued-dangerous-dog-orders/.

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.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.