

Does Technology Advancement Decrease Distracted Driving?*

Harrison Huang

January 24, 2024

1) what was done, 2) what was found, and 3) why this matters (all at a high level). Likely four sentences. Abstract must make clear what we learn about the world because of this paper.

Table of contents

1	Introduction	2
2	Data	2
2.1	Below is a table of the cleaned and filtered table of all distracted driving data .	2
2.2	Violations by age groups	2
3	Results	3
3.1	Violations by youth group per year	3
3.2	Violations by adult group per year	4
3.3	Overall violation by both groups per year	5
4	Discussion	6
5	Conclusion	6
	References	7

*Code and data are available at: <https://github.com/lemonface88/Distracted-Driving-Analysis>

Table 1: Sample of Cleaned ticket Statistics Data

Year	Offence Type	Age Group	Ticket Count
2021	Distracted Driving	Adult	3
2021	Distracted Driving	Adult	5
2021	Distracted Driving	Adult	4
2021	Distracted Driving	Adult	1
2015	Distracted Driving	Adult	158

1 Introduction

2 Data

Data used in this paper are retrieved from Open Data Toronto Portal through the library `opendatatoronto` (Gelfand 2022). Data was cleaned and analyzed using the open source statistically programming language R (R Core Team 2023), using functionalities from `tidyverse` (Wickham 2023), `ggplot2` (Wickham, Chang, et al. 2023), `dplyr` (Wickham, François, et al. 2023), `readr` (Wickham, Hester, and Bryan 2024), `tibble` (Müller and Wickham 2023), `janitor` (Firke 2023), and `knitr` (Xie 2023). Details of the data extraction and cleaning processes are discussed in the subsections below.

2.1 Below is a table of the cleaned and filtered table of all distracted driving data

The table is a representation of the selected columns that remain in the data set after cleaning. This table gives insight to the first 5 rows that are left in the distracted driving tickets handed out.

2.2 Violations by age groups

So what about distracted driving violations? To dive deeper into to the data analysis, the table is then separated into a table that only includes either all youth violations or adult violations. Table 2 is the version that only includes youth distracted driving data, and Table 3 is the other version that only includes distracted driving.

Table 2: Sample of Cleaned youth distracted driving

Year	Ticket Count
2014	2
2014	1
2014	1
2014	2
2014	2

Table 3: Sample of Cleaned adult distracted driving

Year	Ticket Count
2014	41
2014	1
2014	10
2014	4
2014	156

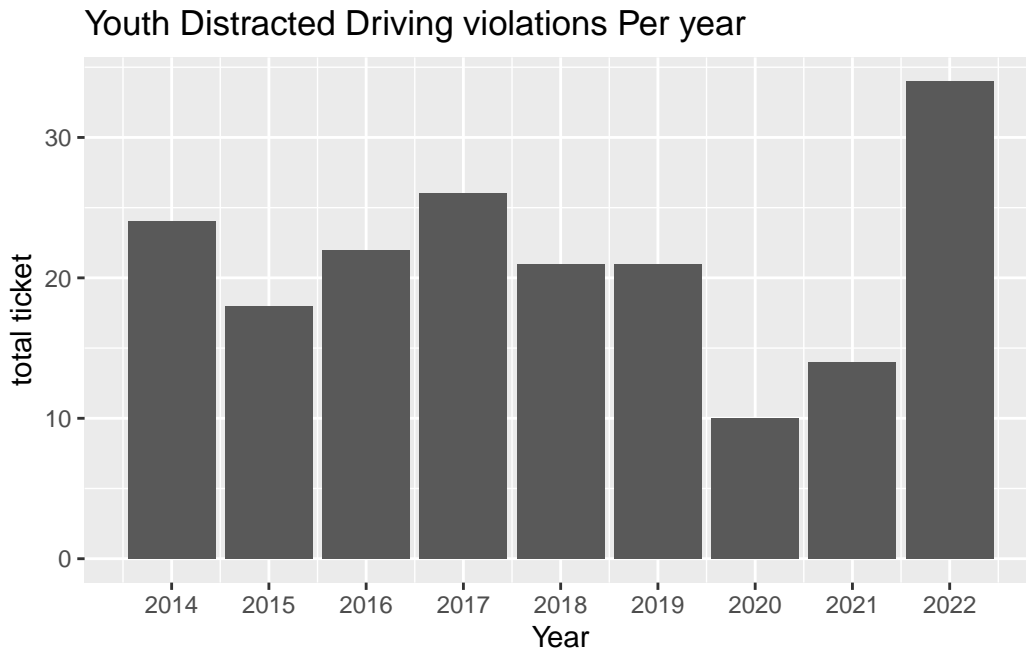
3 Results

3.1 Violations by youth group per year

The graph below shows the total distracted driving violations cases in each year, which is inferred from the total number of tickets given for distracted driving. Overall, the trend sees a big dip in the year 2020 and 2021, this could be due to the effect of covid. As we all know, the lock down that happened during the pandemic prevented people from leaving the house, thus might have caused the overall decrease in the number of violations in those years.

```
only_youth_data |>
  summarize(sum_per_year = sum(ticket_count, na.rm = TRUE), .by = offence_year) |>

ggplot(mapping = aes(x = offence_year, y = sum_per_year))+
  geom_bar(stat= 'identity') +
  scale_x_continuous(breaks =c(2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022))+
  labs(title = "Youth Distracted Driving violations Per year", x = "Year", y = "total tick
```

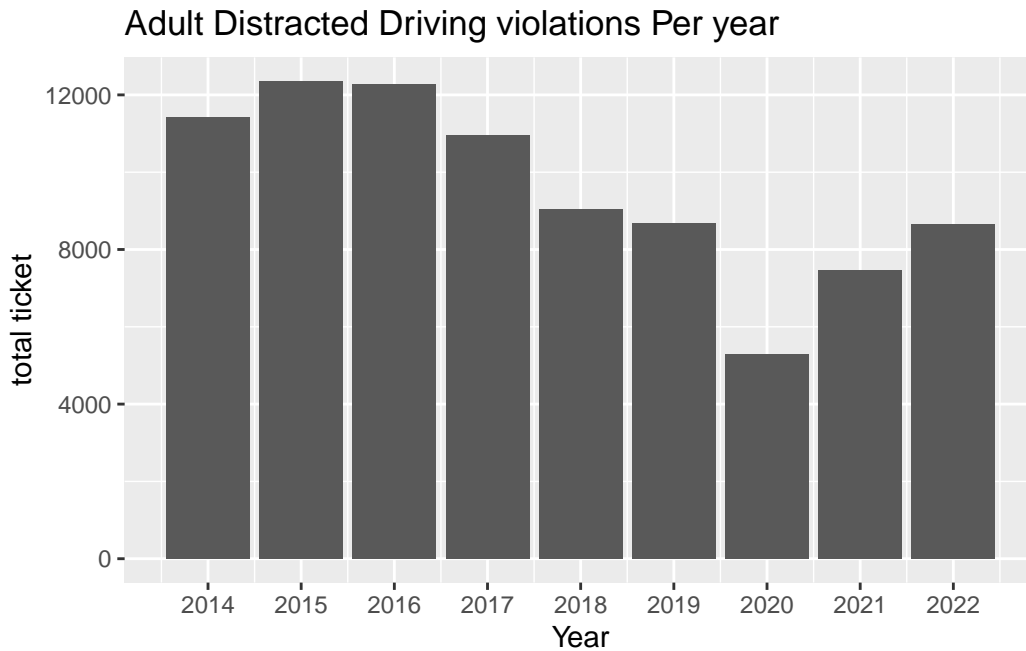


3.2 Violations by adult group per year

The graph below shows the total distracted driving violations cases committed by adults in each year, which is inferred from the total number of tickets given for distracted driving. Overall, the trend sees a big dip in the year 2020 and 2021, this displays a similar trend as the youth drivers. With the two groups of driving having overall lower violations in the year 2020 and 2021, it could suggest that those years might have some other contributing factors involved other than the advancement of technology.

```
only_adult_data |>
  summarize(sum_per_year = sum(ticket_count, na.rm = TRUE), .by = offence_year) |>

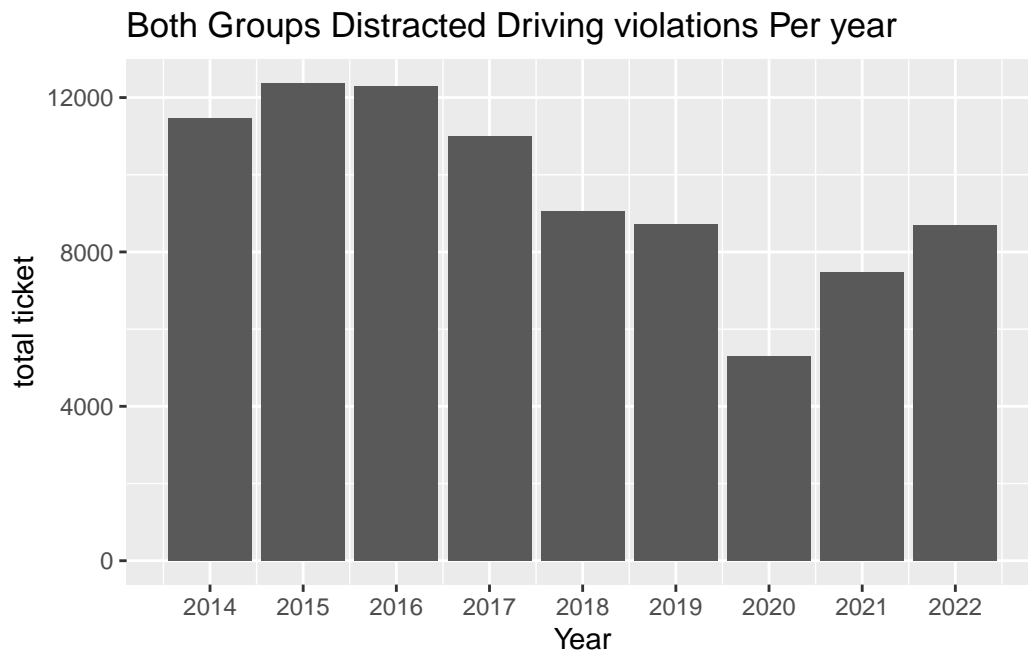
  ggplot(mapping = aes(x = offence_year, y = sum_per_year))+
  geom_bar(stat= 'identity') +
  scale_x_continuous(breaks =c(2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022))+
  labs(title = "Adult Distracted Driving violations Per year", x = "Year", y = "total tick
```



3.3 Overall violation by both groups per year

Upon looking at the summary, the graph below shows the total distracted driving violations cases committed by both adults and youth in each year, which is inferred from the total number of tickets given for distracted driving. As absorbed above, 2020 and 2021 both indicate a big decrease in violation numbers compared to all the other years. However, by combining both of the you and adult data, the overall trend observed is that distracted driving violations have steadily decreased during 2015 to 2019. Despite the numbers increasing again in 2022, the total violations compared to 2015 is still less by about 4000 cases.

```
sum_total_data |>
  ggplot(mapping = aes(x = offence_year, y = total_violations))+
  geom_bar(stat= 'identity') +
  scale_x_continuous(breaks =c(2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023))
  labs(title = "Both Groups Distracted Driving violations Per year", x = "Year", y = "total_violations")
```



4 Discussion

5 Conclusion

References

- Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://github.com/sfirke/janitor>.
- Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://sharlagelfand.github.io/opendatatoronto/>.
- Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. <https://tibble.tidyverse.org/>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley. 2023. *Tidyverse: Easily Install and Load the Tidyverse*. <https://tidyverse.tidyverse.org>.
- Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2023. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://ggplot2.tidyverse.org>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org>.
- Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2024. *Readr: Read Rectangular Text Data*. <https://readr.tidyverse.org>.
- Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.