Week 15 - Tidy Tuesday 6, an Introduction to Quarto Articles (not a true publication)

Keanu Rochette-Yu Tsuen

Generated by ChatGPT as placeholder: The United States boasts one of the largest and most extensive highway systems in the world, facilitating efficient travel across its vast landscapes. The Interstate Highway System, established in 1956, is the backbone of long-distance road travel. These highways are numbered using a systematic scheme: odd-numbered interstates run north-south (e.g., I-5 on the West Coast), while even-numbered interstates run east-west (e.g., I-10 in the southern U.S.). Three-digit interstates typically serve metropolitan areas or act as bypasses and spurs. Speed limits on highways vary by state but generally range between 55 mph (88 km/h) and 75 mph (120 km/h). In rural areas, some highways allow speeds up to 80 mph (130 km/h), particularly in states like Texas and Utah. Urban interstates often have lower limits to accommodate heavier traffic and ensure safety.Safety on highways is paramount. Drivers should maintain a safe following distance, as high speeds reduce reaction times. Always use seat belts and adhere to posted speed limits. Stay vigilant in construction zones where conditions can change rapidly. Avoid distractions such as texting, and never drive under the influence of alcohol or drugs. In case of breakdowns, pull over safely, turn on hazard lights, and stay inside your vehicle if traffic is heavy.Planning your route with updated GPS and weather forecasts can make your highway travel smoother and safer. With proper precautions, highways can be an efficient and enjoyable way to explore the U.S.

Introduction

This week, I decided to learn how to use RStudio and Quarto files to make publishable and reproducible journal articles. I chose the template from PLOS Global Public Health.
We'll be looking at the **National Highway Traffic Flow** data from *Tidy Tuesday*.

Materials and methods

Setup

First, we load the required libraries

```
library(tidyverse)
library(gt)
library(janitor)
library(ghibli)
```

Then, we download the data from the Tidy Tuesday Github.

```
tuesdata <- tidytuesdayR::tt_load('2024-12-03')
traffic <- tuesdata$A64_traffic
```

Data Clean-Up

First, we will clean up the data setup by doing the following:

- clean up the column names with *Janitor*

- select the columns of interest

- rename the size categories to be clearer

- pivot the size categories to long format for plotting later

- pivot the speed categories longer for plotting later

- rename the speed categories (I wanted to rename columns in different ways)

```
traffic_clean <- traffic %>% clean_names() %>%
  select(-site_id, -site_name, -report_date, -time_period_ending, -time_interval,
         -name, -longitude, -latitude) %>%
  rename(`0-520` = x0_520_cm,
         `521-660` = x521_660_cm,
         `661-1160` = x661_1160_cm,
         `Over 1160` = x1160_cm) %>%
  pivot_longer(cols = `0-520`:`Over 1160`,
               names_to = "size_class_cm",
               values_to = "num_by_size") %>%
  pivot_longer(cols = x0_10_mph:x80_mph,
```

```
              names_to = "speed_cat_mph",
              values_to = "num_by_speed") %>%
  mutate(speed_cat_mph = str_replace(speed_cat_mph, "x", ""),
         speed_cat_mph = str_replace(speed_cat_mph, "_mph", ""),
         speed_cat_mph = str_replace(speed_cat_mph, "_", "-"),
         speed_cat_mph = ifelse(speed_cat_mph == "80", "Over 80", speed_cat_mph),
         size_class_cm = as.factor(size_class_cm))
```

Graphing

For the size distribution, I am plotting a boxplot graph to show which size class of cars are
more represented on the highway.

```
size <- traffic_clean %>%
  ggplot(aes(x= size_class_cm, y= num_by_size, fill = size_class_cm))+
  geom_boxplot() +
  labs(title = "Distribution of vehicules by size class",
       x = "Size Classes (cm)",
       y = "Number of Vehicles") +
  theme_bw() +
  theme(plot.title = element_text(size=14, face = "bold"),
        plot.subtitle = element_text(size=12),
        axis.title = element_text(size = 12, face = "bold"),
        axis.text = element_text(size = 12),
        strip.text.x = element_text(size = 12, face = "bold"),
        legend.title=element_text(size=12, face = "bold"),
        legend.text=element_text(size=12),
        legend.position = "none",
        panel.background = element_rect(fill = "azure1"))+
  #custom the colors
  scale_fill_ghibli_d("MarnieMedium1", direction =-1)
```

For the speed data, I am plotting a bar graph to show distribution of the vehicle speed on the
highway. It made a nice bell curve.

```
speed <- traffic_clean %>% group_by(speed_cat_mph) %>%
  summarize(avg_by_speed = mean(num_by_speed, na.rm = T)) %>%
  ggplot(aes(x= speed_cat_mph, y= avg_by_speed, fill = speed_cat_mph))+
  geom_bar(stat= "identity")+
  labs(title = "Distribution of vehicles by speeding category",
       x = "Speed Categories",
```

```
      y = "Average Number of Vehicles") +
theme_bw() +
theme(plot.title = element_text(size=14, face = "bold"),
      plot.subtitle = element_text(size=12),
      axis.title = element_text(size = 12, face = "bold"),
      axis.text = element_text(size = 12),
      axis.text.x = element_text(angle = 30, vjust = 0.8, hjust=0.5),
      legend.title=element_text(size=12, face = "bold"),
      legend.text=element_text(size=12),
      legend.position = "none",
      panel.background = element_rect(fill = "azure1"))
```

Results

We can show the plots now. We can see that they are nicely labeled.

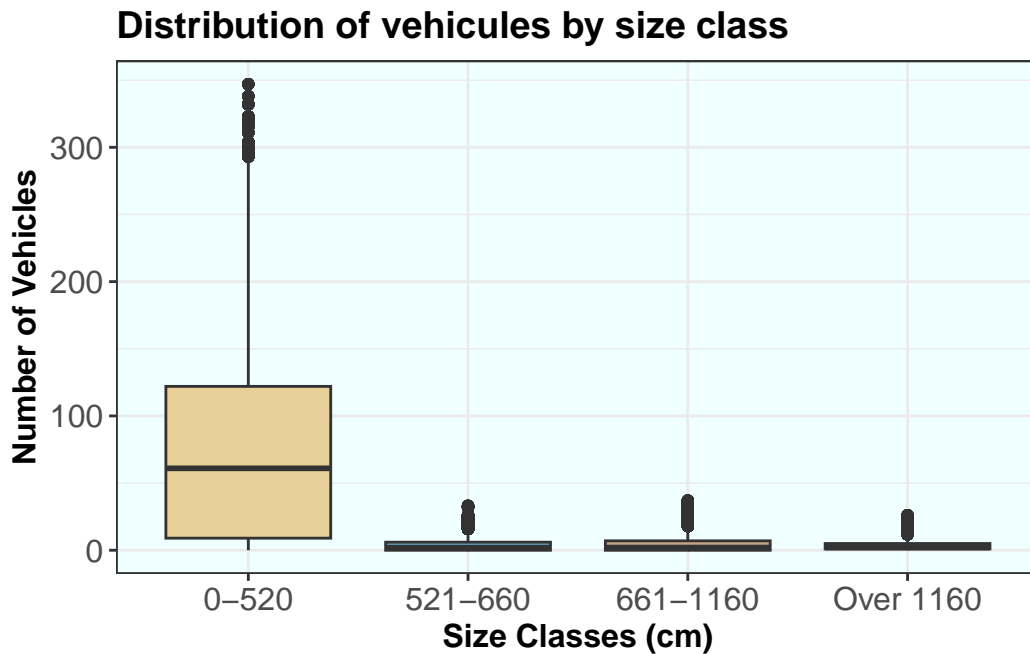In the size plot, cars from 0-520cm are most represented on the road.

```
size
```


Figure 1: Size Distribution of Vehicles

As of speed, cars mostly reach speed around 41-45 mph.

```
speed
```

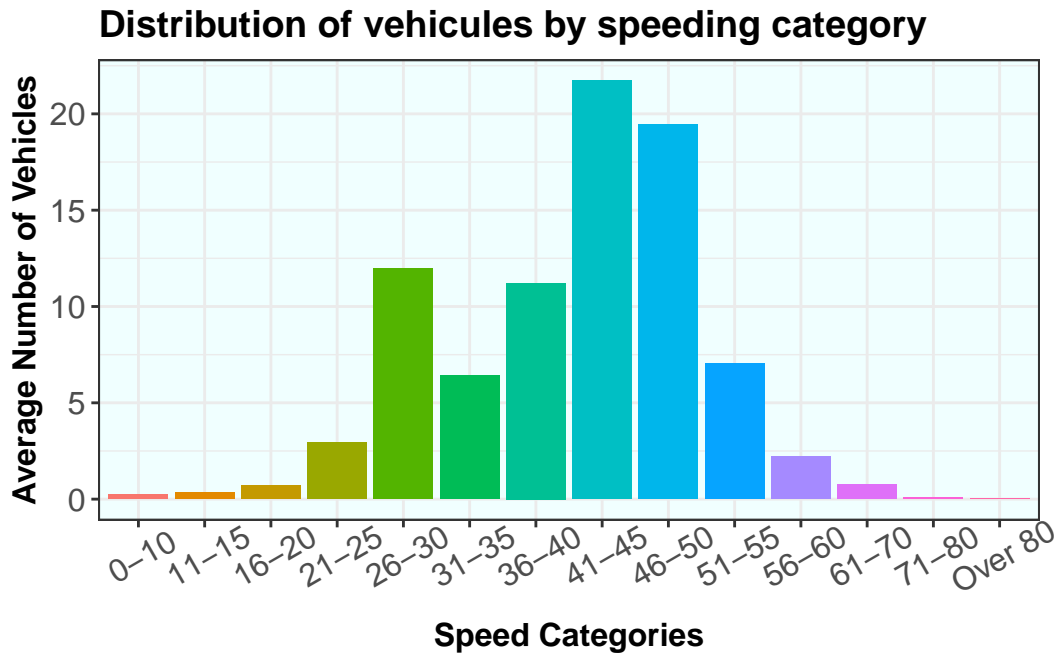**Distribution of vehicules by speeding category**

Figure 2: Speed Distribution of Vehicles

Discussion

The cool thing about quarto document is that I can reference figures like that, Figure 1.
Then I can talk some more to discuss Figure 2.

I can also add citations from my Zotero library like that:

- Leptospirosis in Hawaii : (Anderson & Minette, 1986)

- Agniel et al. (Agniel, 2018) said something about environmental reservoirs of pathogens.

All the citations will be added to the bibliography and the reference.

Unfortunately, I was not able to change the citation style of the document. I was told to download the CSL file for APA 6th Edition and set it in the YAML header as {csl: apa-6th-edition.csl}. But it didn't change the citation format. I believe that it is because I used a template from a specific journal which might use their own citation format and it overrides my file.

Conclusion

Overall, Quarto Articles are very awesome and it allows us to have our scientific writing, citations, and plots with features to both label and reference the plots already embeded in the file. Unfortutely, I was not able to change the citation format. I will have to practice some more with Quarto in order to better understand those mechanisms.

Acknowledgments

Thank you to *me* for taking the time to learn Quarto Articles and Nyssa for teaching us how to code and create those awesome outputs !

References

Agniel, D. (2018). *Les Marquises, archipel le plus isolé du monde, sont menacées par la pêche industrielle.* Retrieved from https://reporterre.net/Les-Marquises-archipel-le-plus-isole-du-monde-sont-menacees-par-la-peche

Anderson, B. S., & Minette, H. P. (1986). Leptospirosis in Hawaii: shifting trends in exposure, 1907-1984. *International journal of zoonoses, 13*(2), 76–88.