Weather Analysis and Prediction

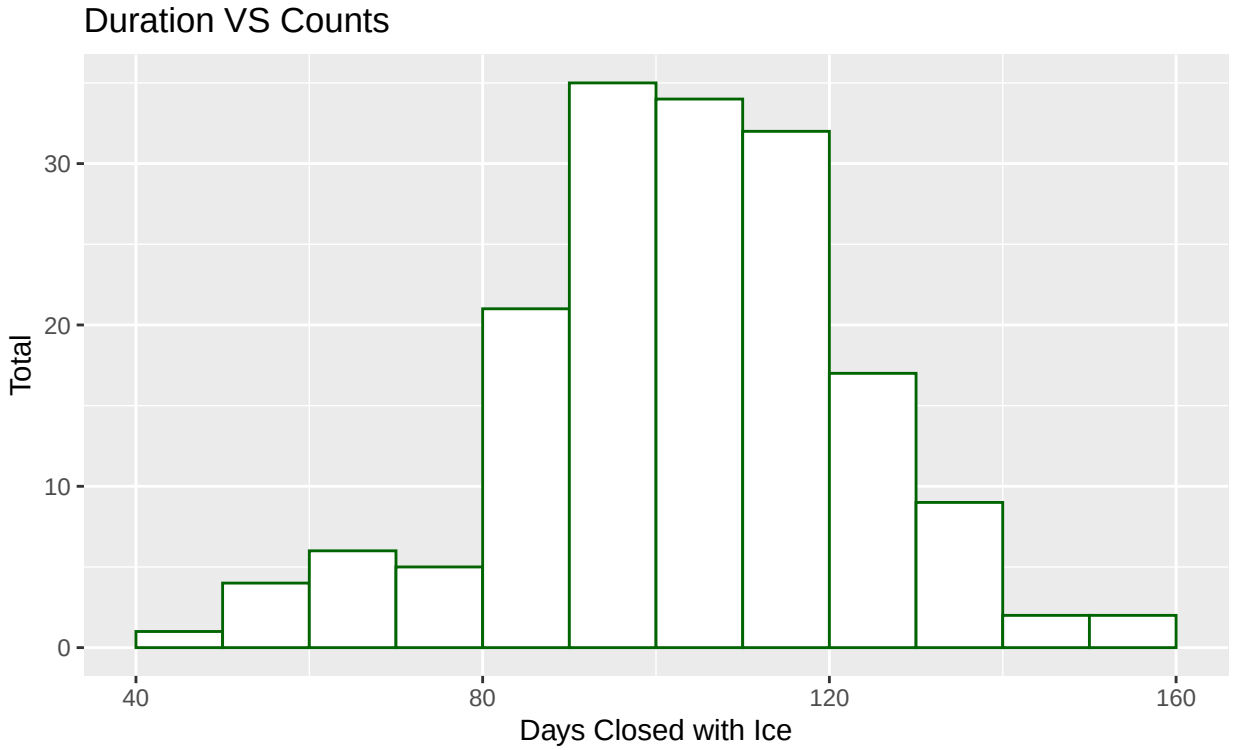
Muhao Li(Jimmy)

```
#Reading the data
monona = read_csv("../../data/lake-monona-winters-2023.csv") %>%
  mutate(ff_cat = reorder(ff_cat, ff_x))
```

In this project, I use a dataset containing the data for a local lake at Madison, Wisconsin on the observation of the lake being frozen to perform time-series analysis of climate patterns.

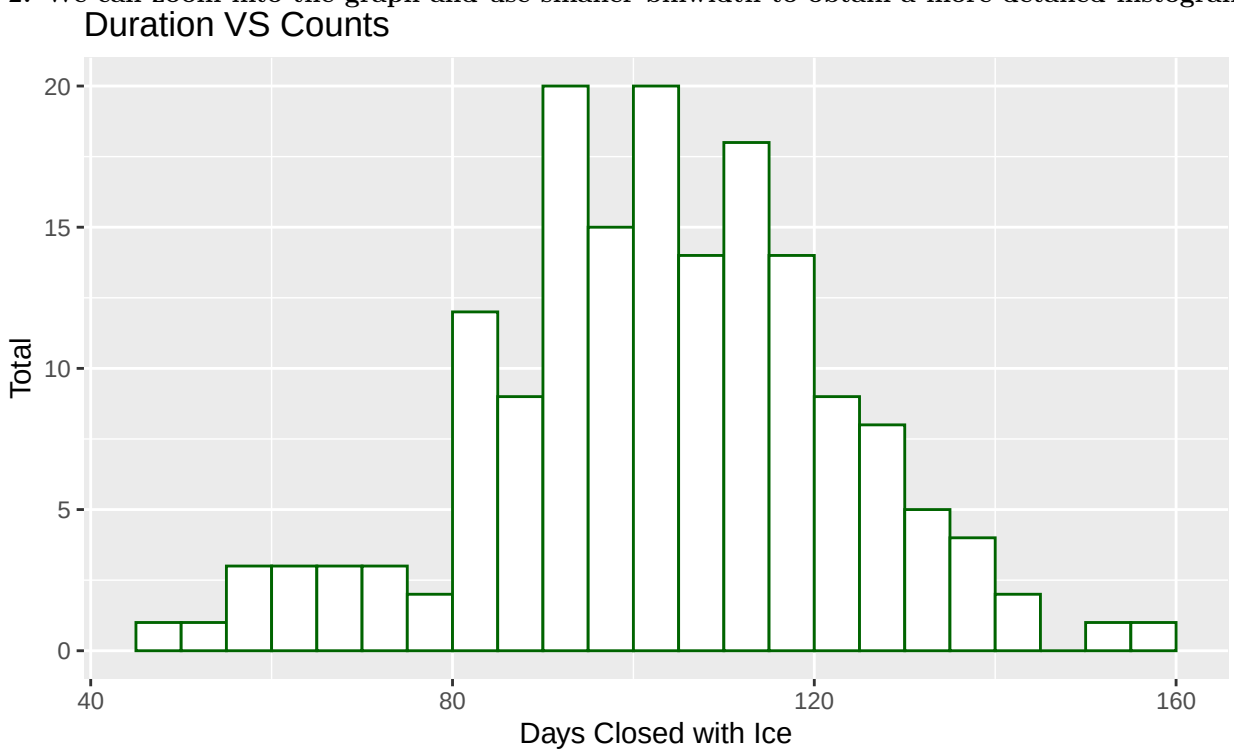
1. Starting off with a simple histogram of Duration of winter for Lake Monona, to visualize the overall distribution to obtain general understanding of the data.

```
ggplot(monona, aes(x = duration)) +
  geom_histogram(boundary=0, binwidth=10, color='darkgreen',fill='white') + labs(x='Days Closed with Ic
```



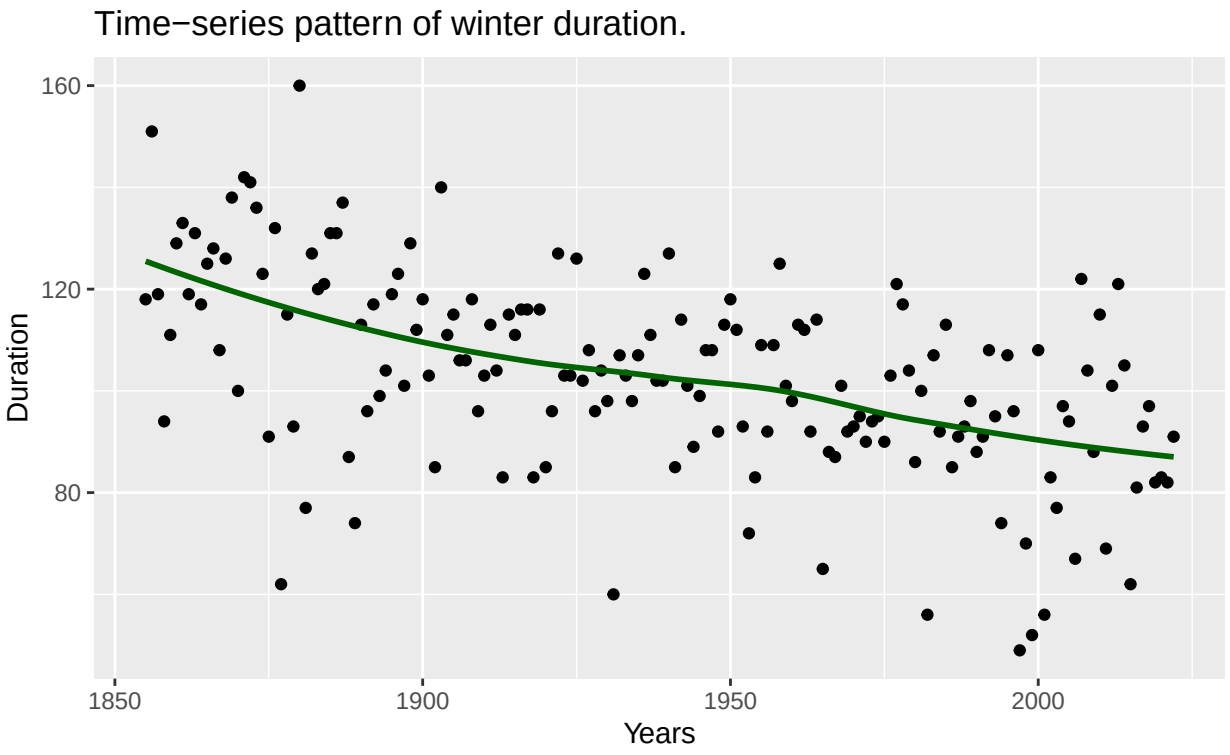

```
ggplot(monona, aes(x = duration)) +
  geom_histogram(boundary=70, binwidth=5, color='darkgreen',fill='white') + labs(x='Days Closed with Ic
```

2. We can zoom into the graph and use smaller binwidth to obtain a more detailed histogram.



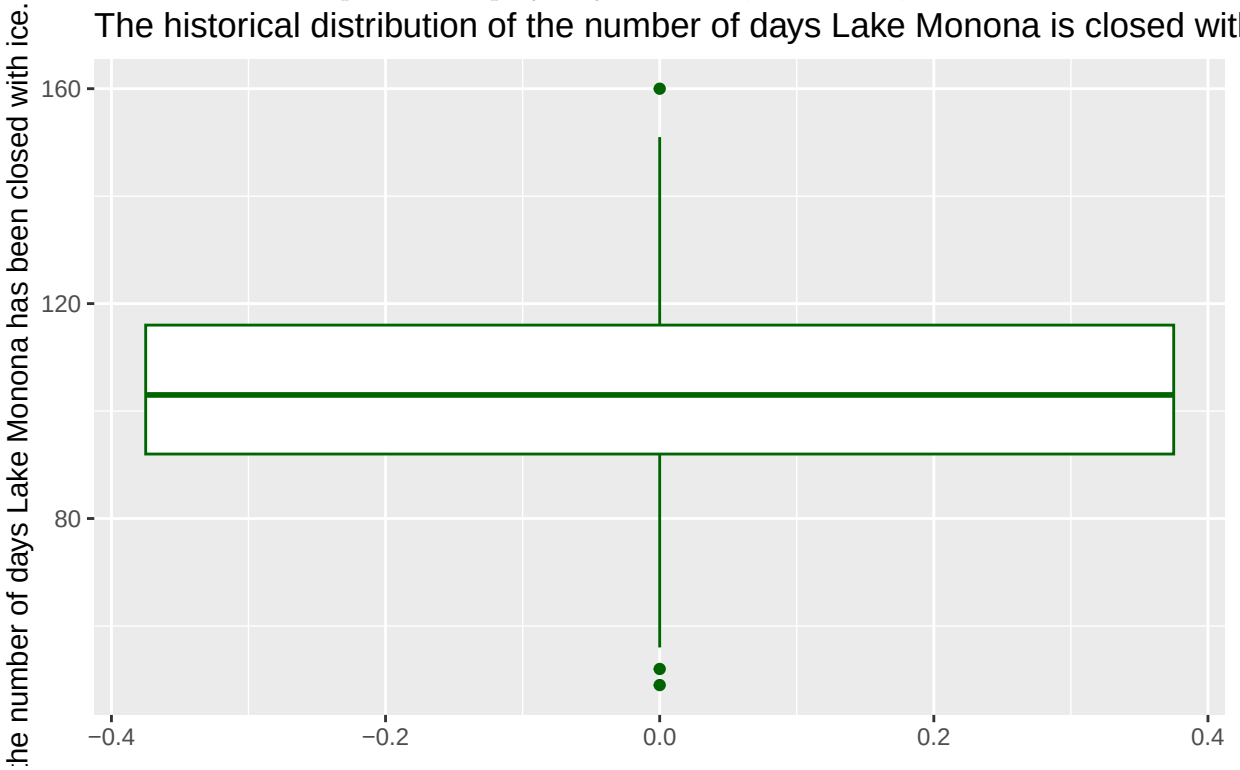

```
ggplot(monona, aes(x = year1, y = duration)) +
  geom_point(color='black') +
  geom_smooth(color='darkgreen',se=FALSE) +labs(x='Years', y="Duration", title="Time-series pattern of
```

3. To further visualize and simpify to improve understanding, I add a line of best fit to illustate the overall time-series pattern of the duration of winter across years.



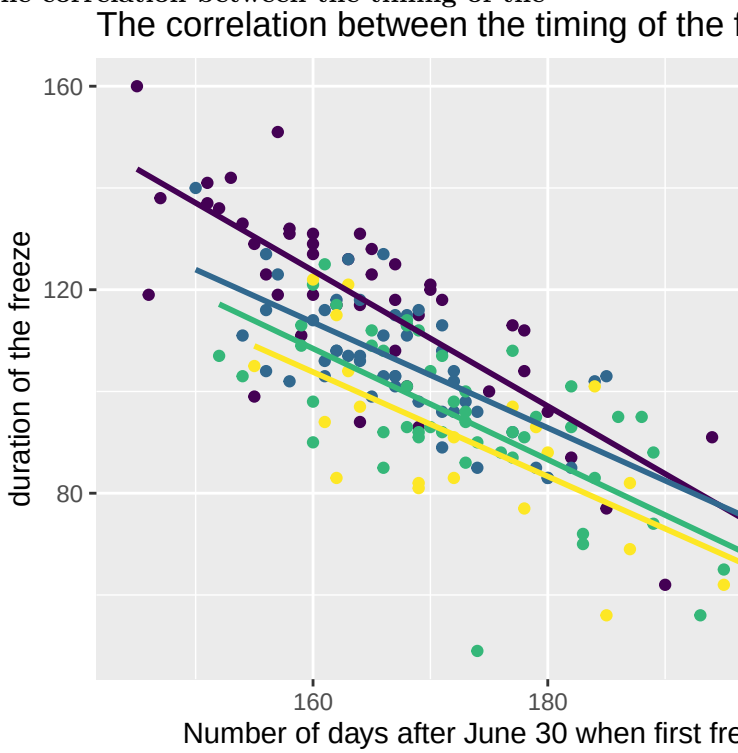

```
ggplot(monona, aes(y=duration)) + geom_boxplot(fill='white', color='darkgreen',)+labs( y="the number of
```

4. I also added a box plot to display any skewness, if existed, within the distribution.



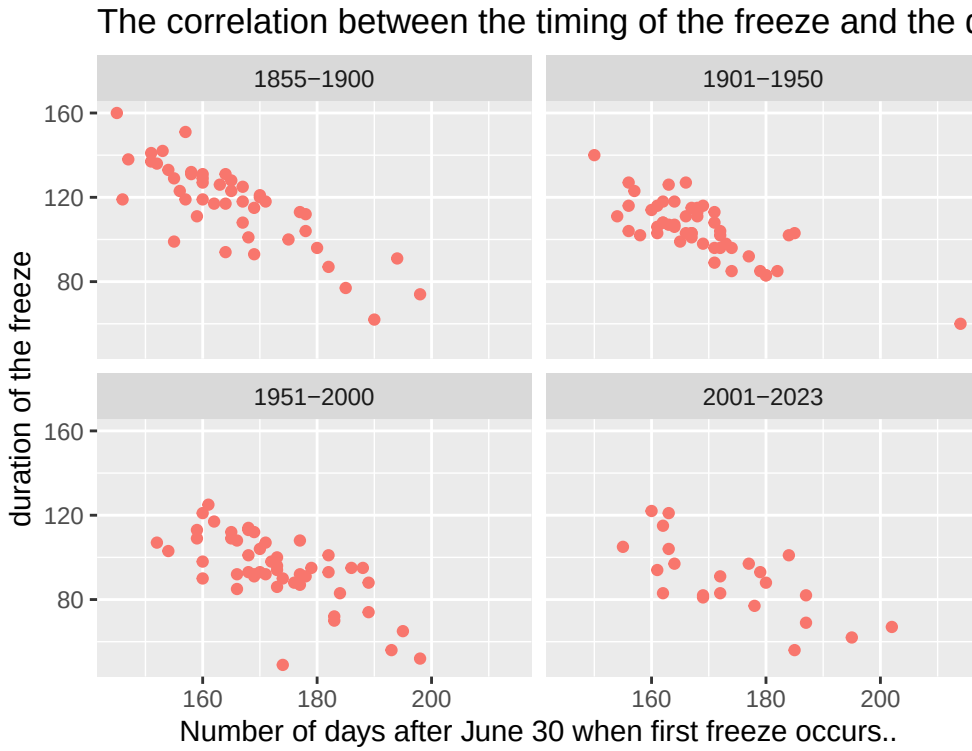
```
ggplot(monona, aes(x = ff_x, y = duration, color = period50)) +
  geom_point() +
  geom_smooth(se = FALSE, method = "lm") +
  scale_color_viridis_d() + labs(x='Number of days after June 30 when first freeze occurs..', y='duratio
```

5. To further study climate patterns, I looked into the correlation between the timing of the freeze and the duration of the freeze at Lake Monona.




```
ggplot(monona, aes(x = ff_x, y = duration, color = "Time Period")) +
  geom_point() +facet_wrap(facets = vars(period50))+labs(x='Number of days after June 30 when first free
```

8 And here is another visulization splitted into ranges of years. We can see the general pattern shfited downward from the 1850's to 2023. This indicate an overall shortening of the duration of



freeze at Lake Monona across time.