STA160-allcom_plots

Shih-Chi Chen

6/1/2022

```
#import weather data combined fire data
joined_weather <- read.csv("~/Downloads/ClimateAnalysis-main/joined_weather.csv")
```

```
#change type to date
joined_weather$Date <- as.Date(joined_weather$Date)
```

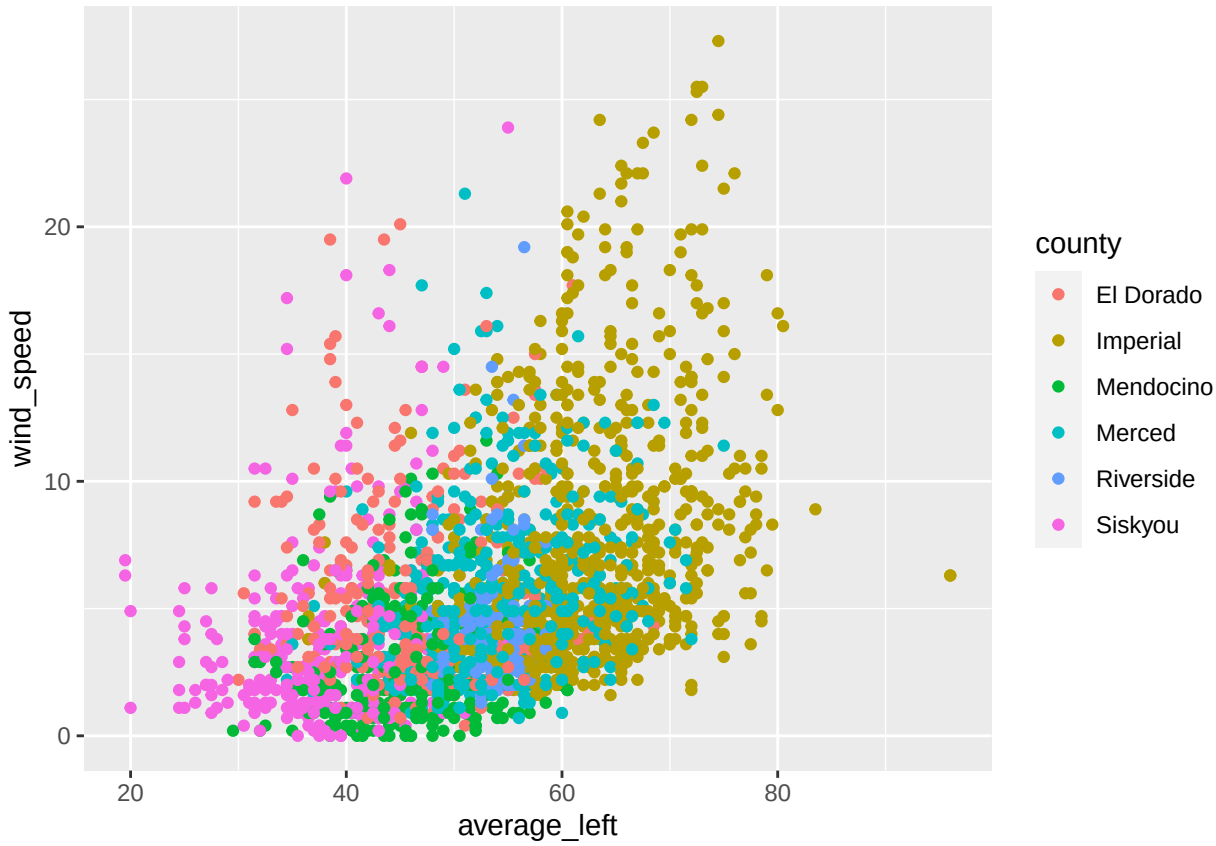
EDA for weather data

```
# load the library
library(ggplot2)
library(ggridges)
```

Scatter plots

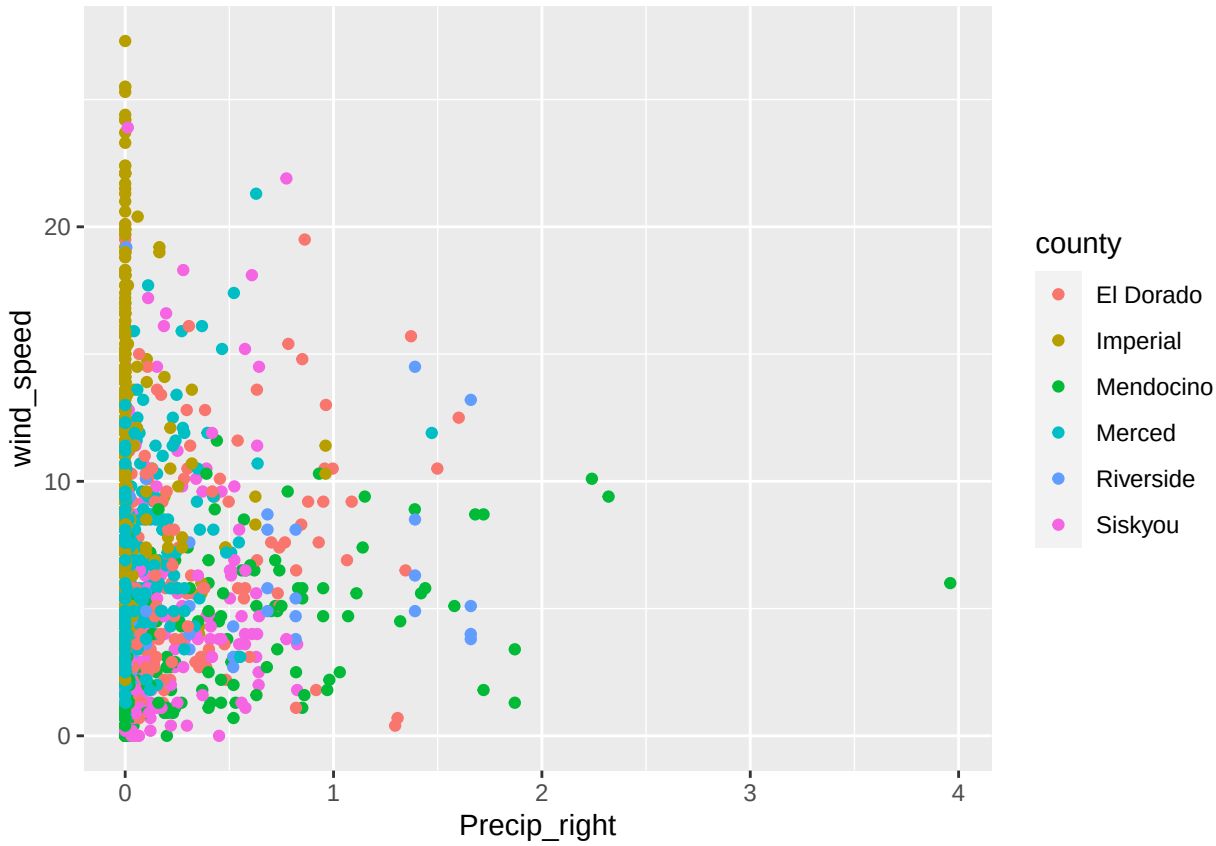
```
#wind vs temp
ggplot(data = joined_weather) +
  geom_point(mapping = aes(x = average_left, y = wind_speed, colour = county))
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```

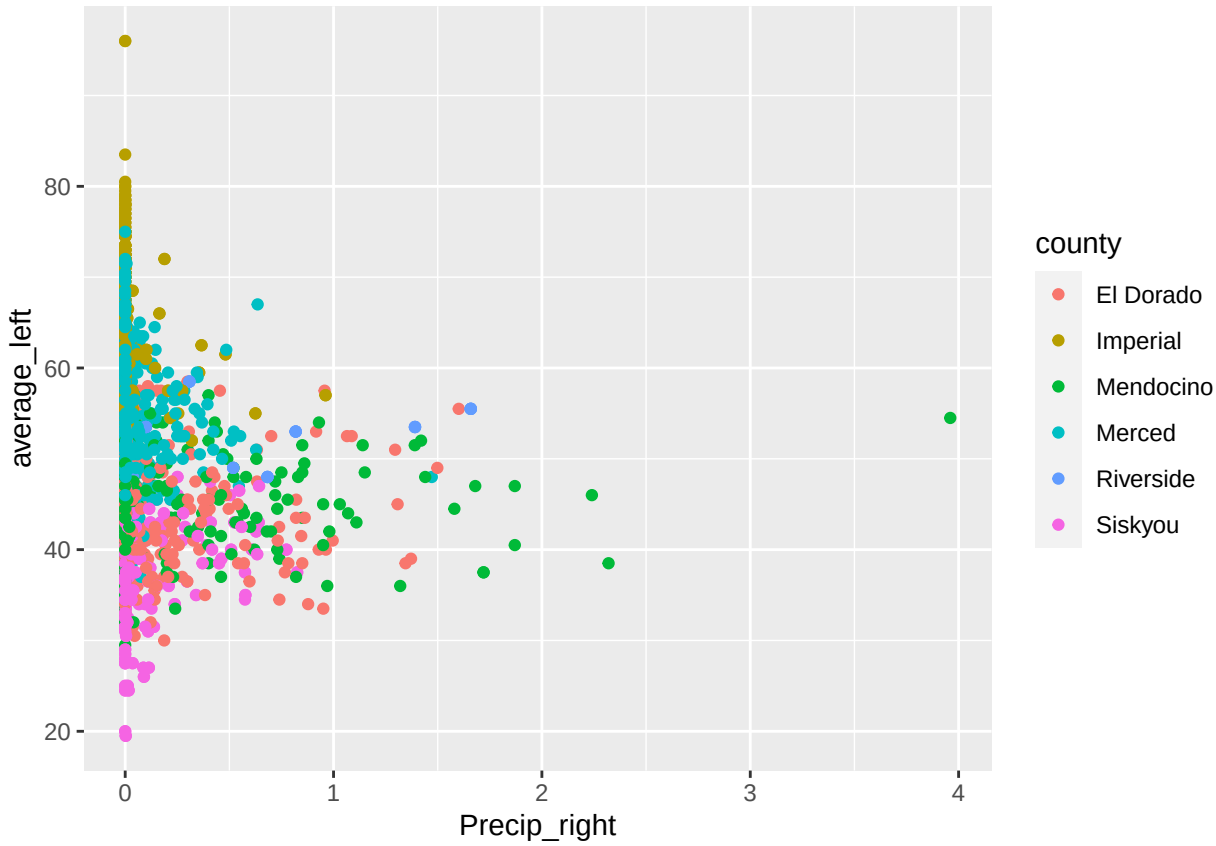



```
#precip vs wind
ggplot(data = joined_weather) +
  geom_point(mapping = aes(x = Precip_right, y = wind_speed, colour = county))
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```



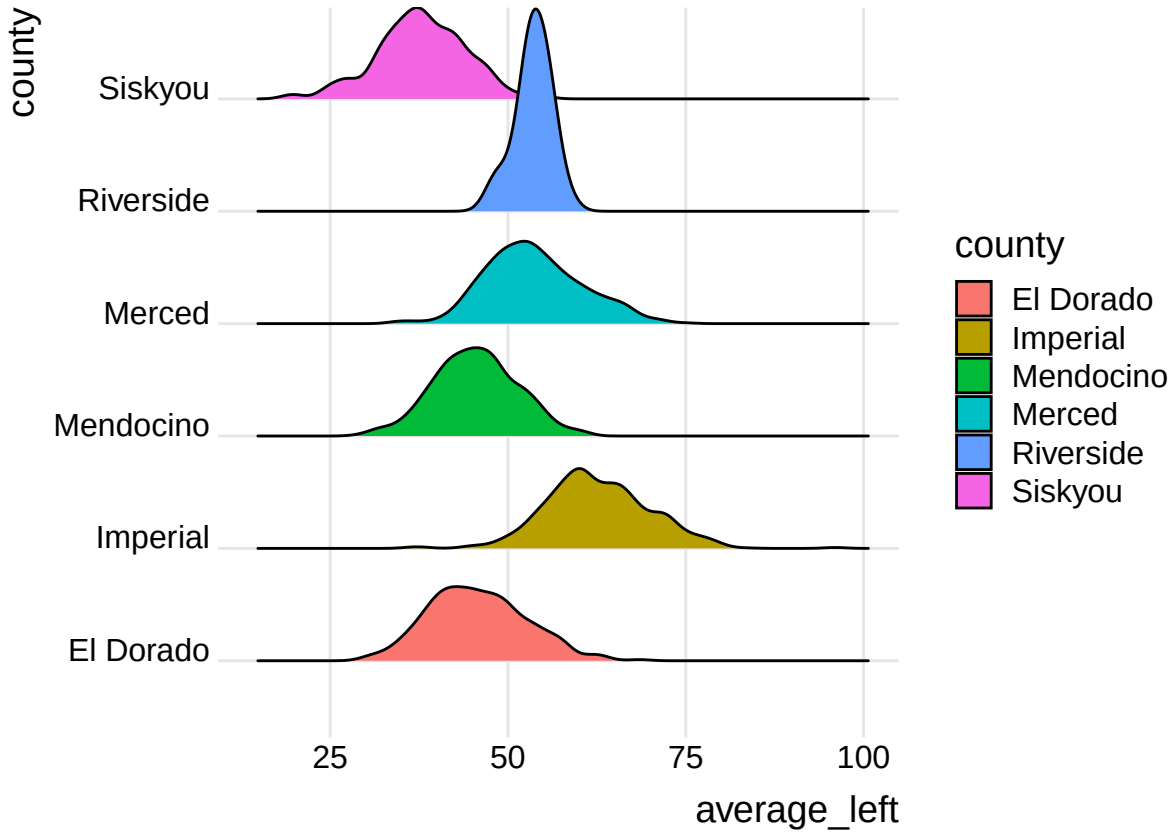

```
#precip vs avg temp
ggplot(data = joined_weather) +
  geom_point(mapping = aes(x = Precip_right, y = average_left, colour = county))
```



Distribution plots

```
#county and average temp
theme_set(theme_ridges())
ggplot(joined_weather, aes(x = average_left, y =county )) +
  geom_density_ridges(aes(fill = county))
```

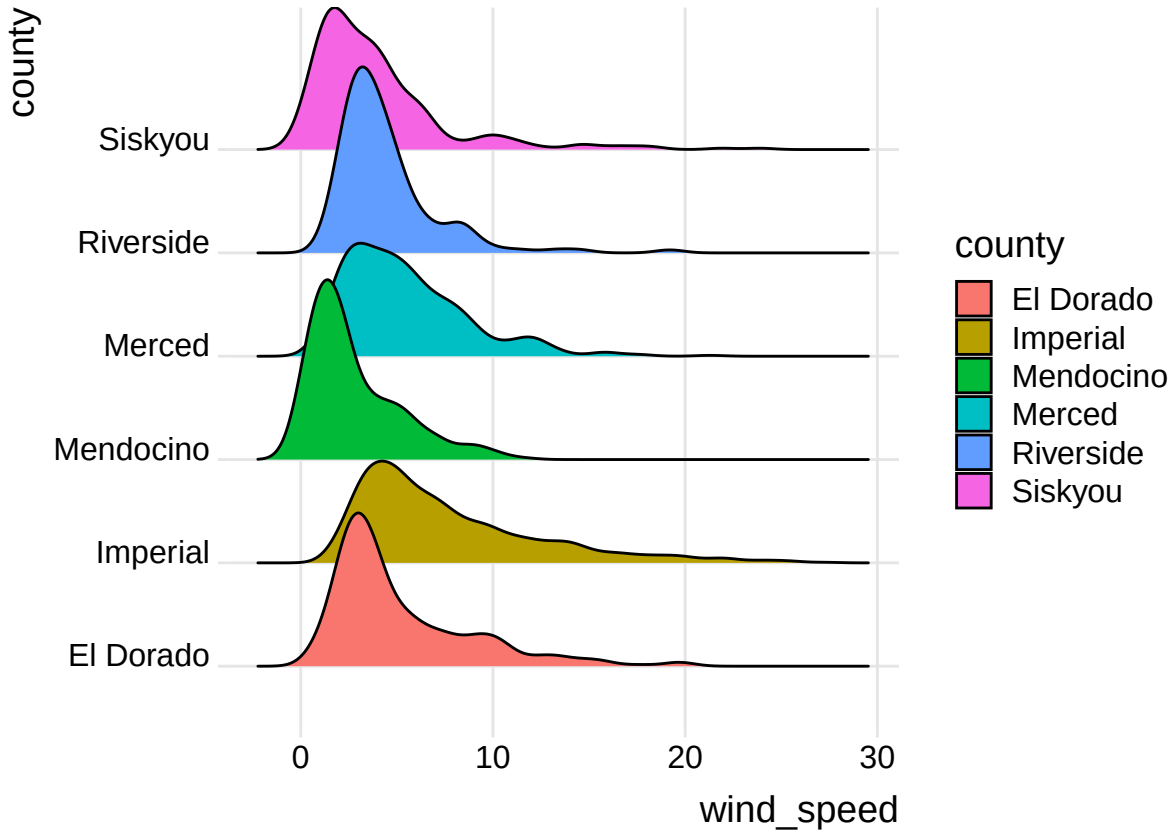
```
## Picking joint bandwidth of 1.56
```




```
#county and wind speed
theme_set(theme_ridges())
ggplot(joined_weather, aes(x = wind_speed, y =county )) +
  geom_density_ridges(aes(fill = county))
```

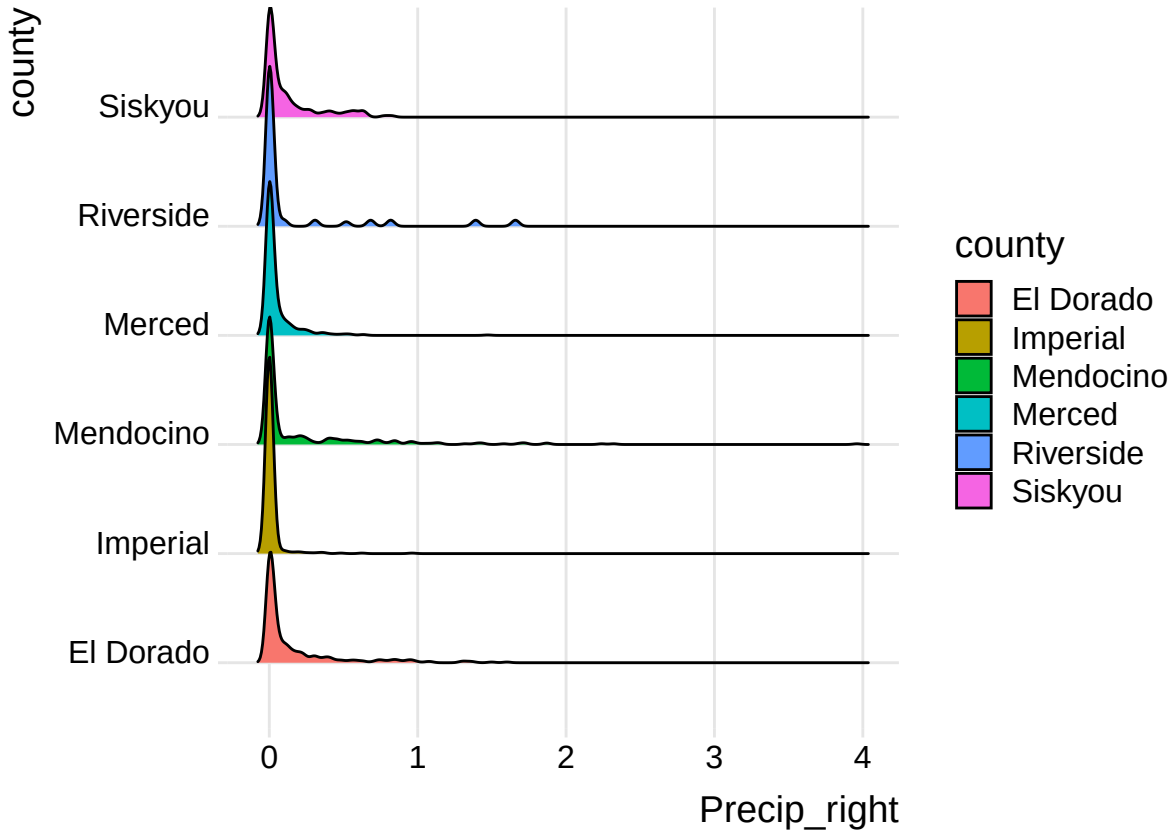
```
## Picking joint bandwidth of 0.744
```

```
## Warning: Removed 6 rows containing non-finite values (stat_density_ridges).
```




```
#county and Precip
theme_set(theme_ridges())
ggplot(joined_weather, aes(x = Precip_right, y =county )) +
  geom_density_ridges(aes(fill = county))
```

```
## Picking joint bandwidth of 0.0258
```

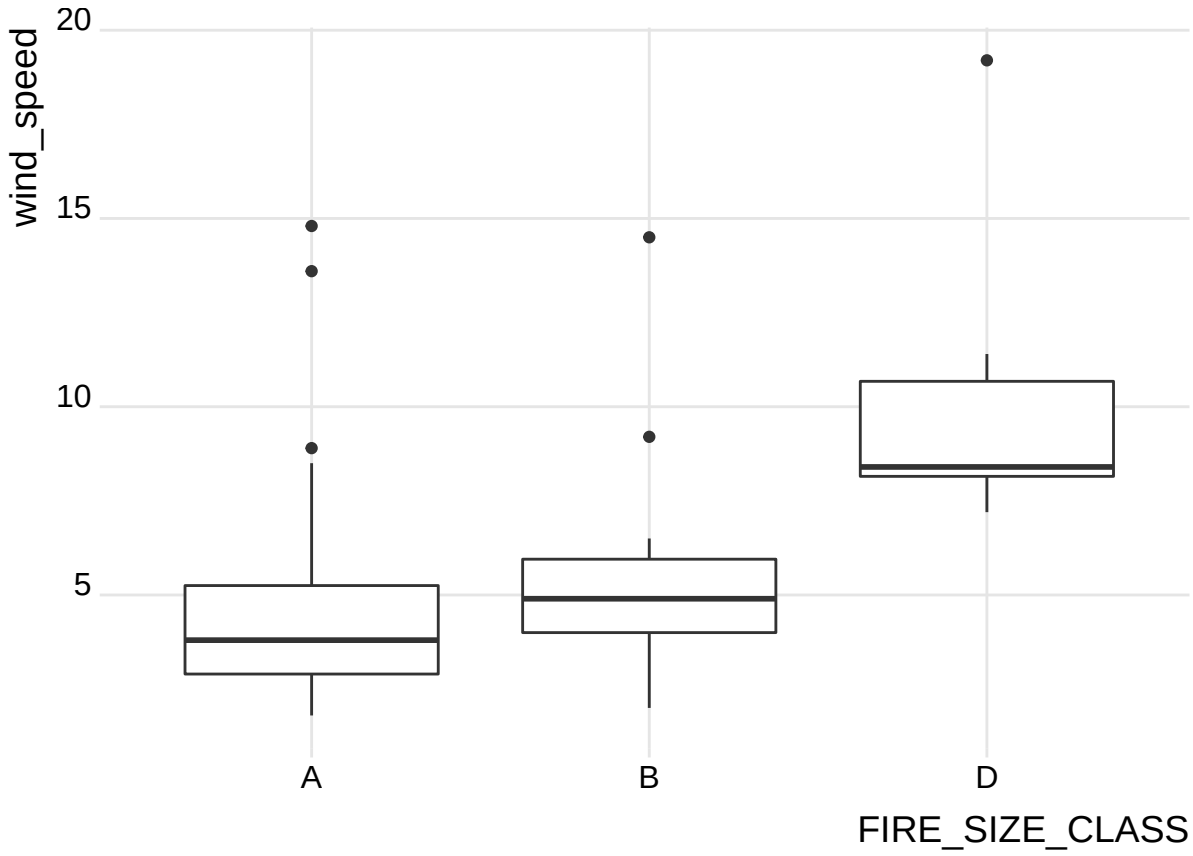


EDA for fire and weather data

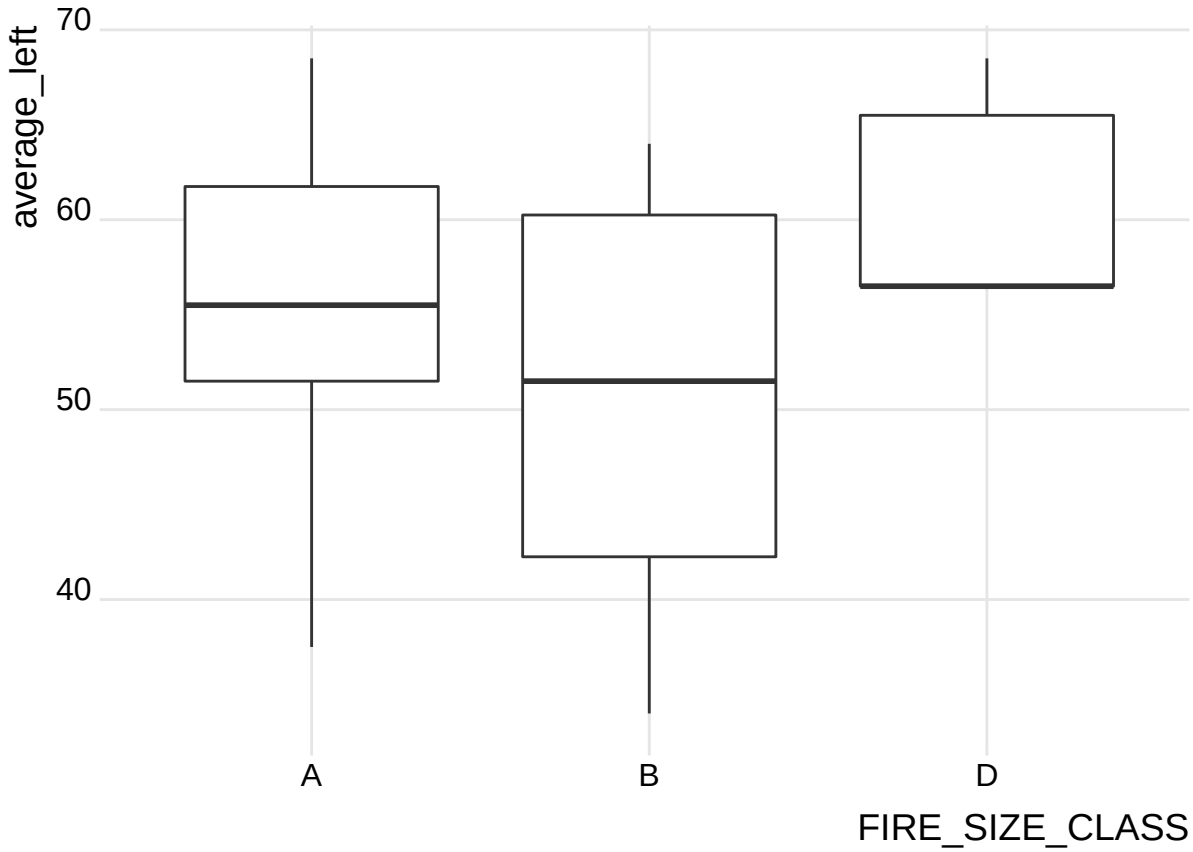
```
#import weather data combined fire data
weather_fire <- read.csv("~/Downloads/ClimateAnalysis-main/weather_fire.csv")
#change type to date
weather_fire$Date <- as.Date(weather_fire$Date)
weather_fire$date.CONT_DATE.<- as.Date(weather_fire$date.CONT_DATE.)

#fire days
days=weather_fire$date.CONT_DATE.-weather_fire$Date
#add to df
weather_fire$cts_days=days
```

#Boxplots From the boxplot, it shows that the large fire tend to have high wind speed.

```
#wind speed vs fire size class
ggplot(data = weather_fire, mapping = aes(x = FIRE_SIZE_CLASS, y = wind_speed)) +
  geom_boxplot()
```

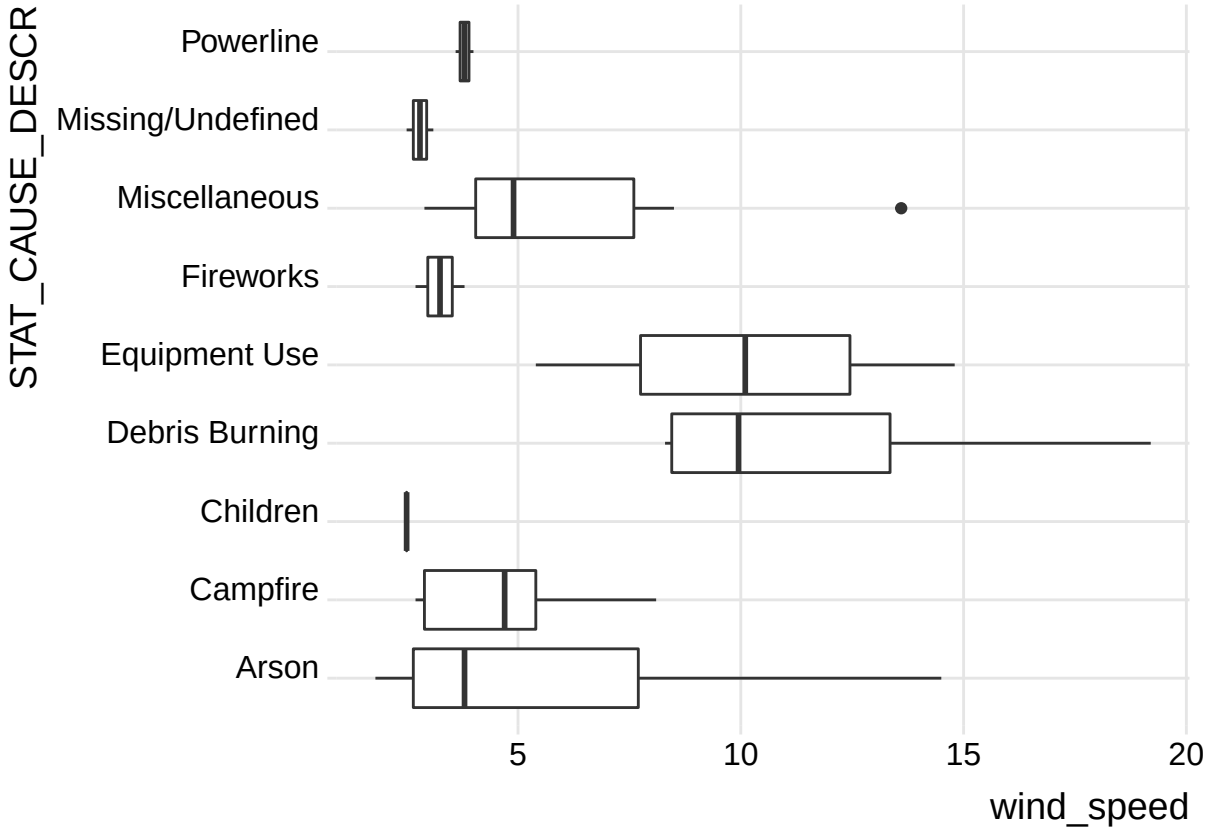



```
#average temp vs fire size class
ggplot(data = weather_fire, mapping = aes(x = FIRE_SIZE_CLASS, y =average_left )) +
  geom_boxplot()
```



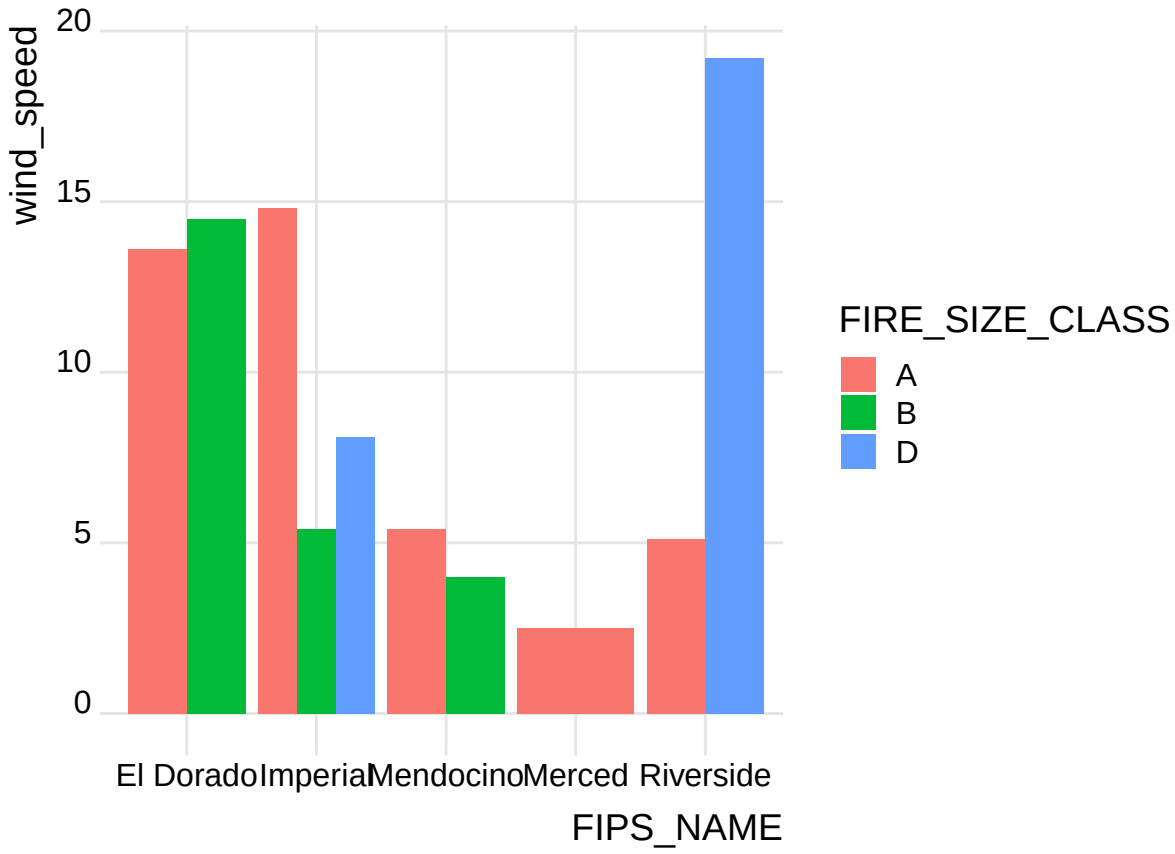
From the plot, it shows that most of the reason for fires occur in the low wind speed(<10). However, the Equipment Use and Debris Burning tend to causes fires when wind speed is large than 10. There also can be seen that the range of wind speed is large when the fire was caused by Arson and Debris Burning.

```
#wind speed vs fire cause
ggplot(data = weather_fire, mapping = aes(x = STAT_CAUSE_DESCR, y = wind_speed)) +
   theme(legend.position="none") +
  geom_boxplot()+
  coord_flip()
```

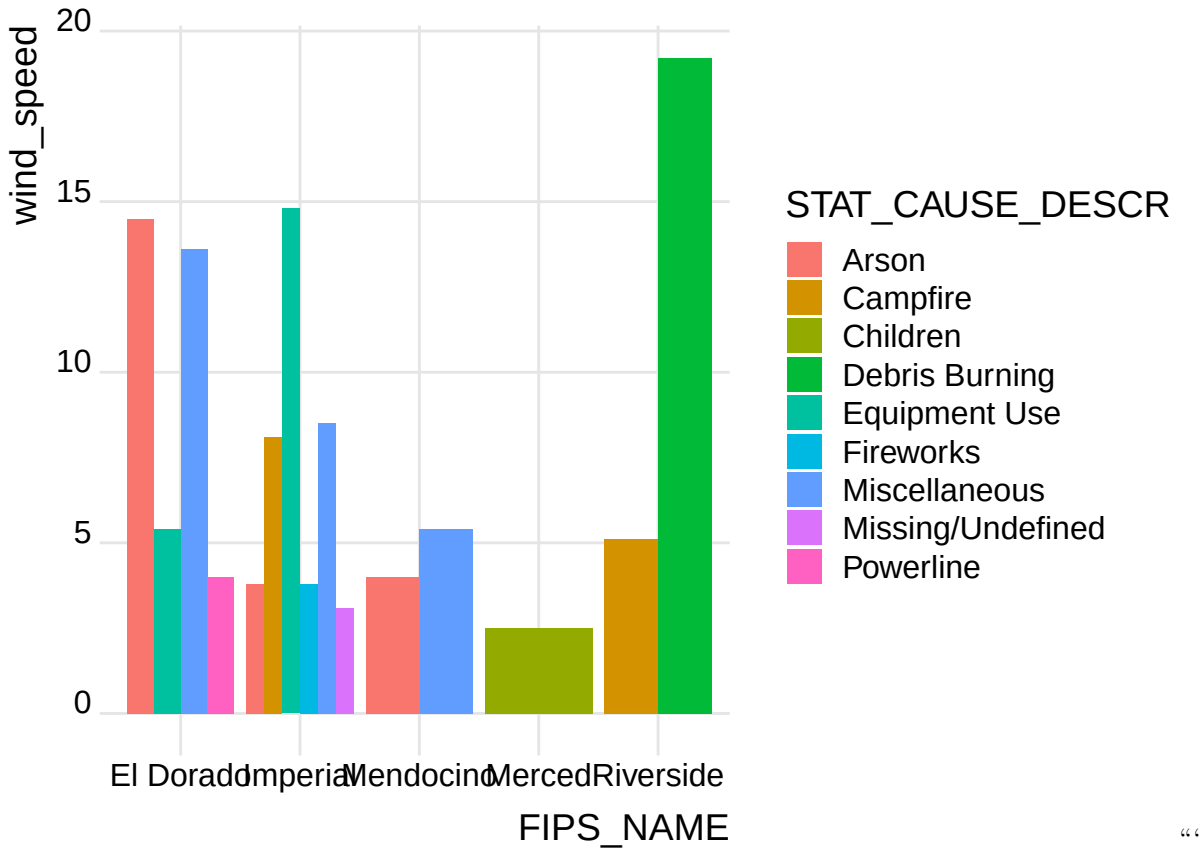


bar chart

```
#wind speed, fips and fire size class bar
ggplot(data=weather_fire, aes(x=FIPS_NAME, y=wind_speed, fill=FIRE_SIZE_CLASS)) +
  geom_bar(stat="identity", position=position_dodge())
```




```
# #wind speed, fips and cause
ggplot(data=weather_fire, aes(x=FIPS_NAME, y=wind_speed, fill=STAT_CAUSE_DESCR)) +
geom_bar(stat="identity", position=position_dodge())
```


wind_speed
20
15
10
5
0
El Dorado
Imperial
Mendocino
Merced
Riverside
FIPS_NAME
STAT_CAUSE_DESCR
Arson
Campfire
Children
Debris Burning
Equipment Use
Fireworks
Miscellaneous
Missing/Undefined
Powerline

“‘