An Introduction to plotly

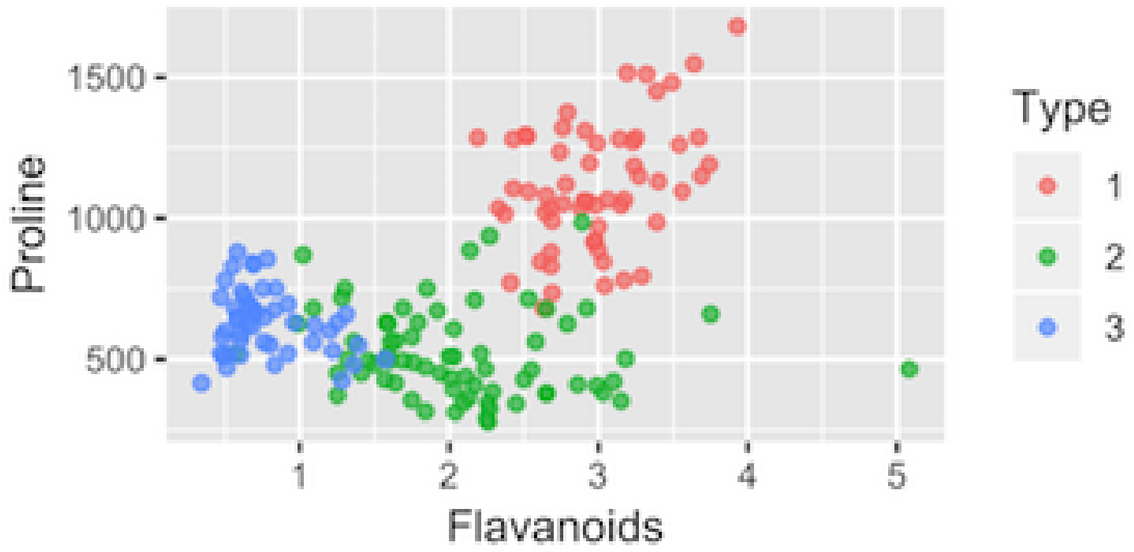
INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

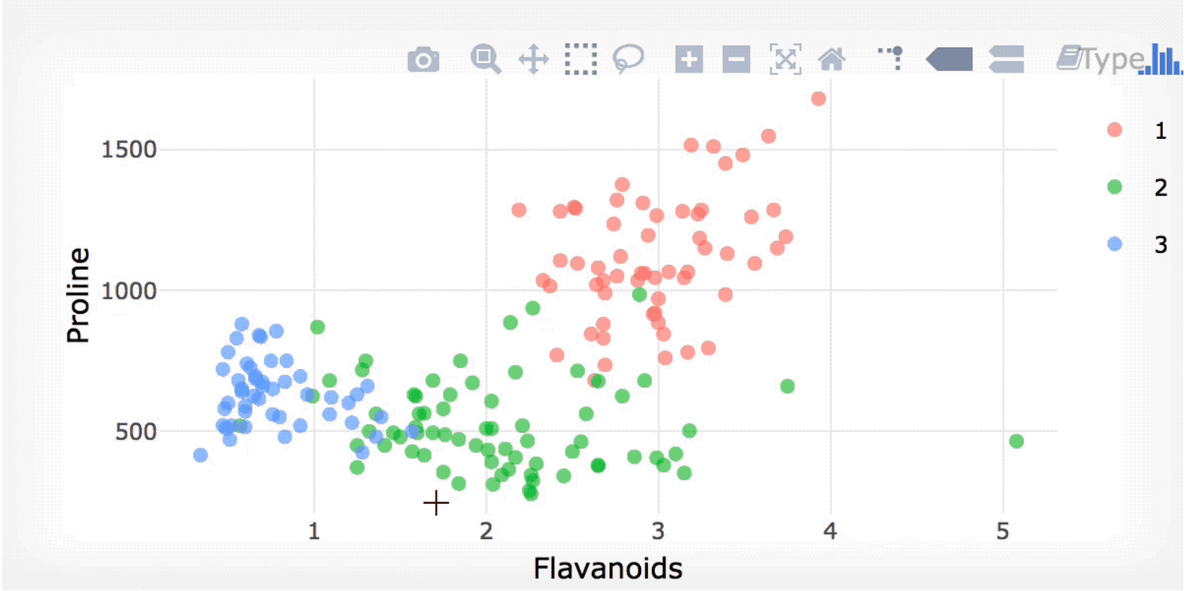
Adam Loy
Statistician, Carleton College

plotly

- Visualization library for interactive and dynamic web-based graphics
- Plots work in multiple formats
  - viewer windows
  - R Markdown documents
  - shiny apps
- Active development + supportive community

Static vs. Interactive graphics



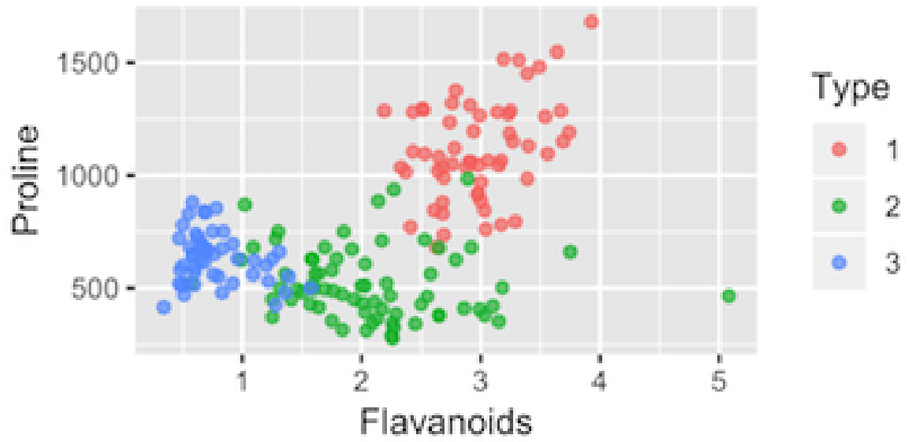


Wine data

```
library(dplyr)
glimpse(wine)
```

```
Observations: 178
Variables: 14
$ Type            <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1...
$ Alcohol         <dbl> 14.23, 13.20, 13.16, 14.37, 13.24, 14.20, 14.3...
$ Malic           <dbl> 1.71, 1.78, 2.36, 1.95, 2.59, 1.76, 1.87, 2.15...
...
$ Hue             <dbl> 1.04, 1.05, 1.03, 0.86, 1.04, 1.05, 1.02, 1.06...
$ Dilution        <dbl> 3.92, 3.40, 3.17, 3.45, 2.93, 2.85, 3.58, 3.58...
$ Proline         <int> 1065, 1050, 1185, 1480, 735, 1450, 1290, 1295,...
```

ggplot2 scatterplot



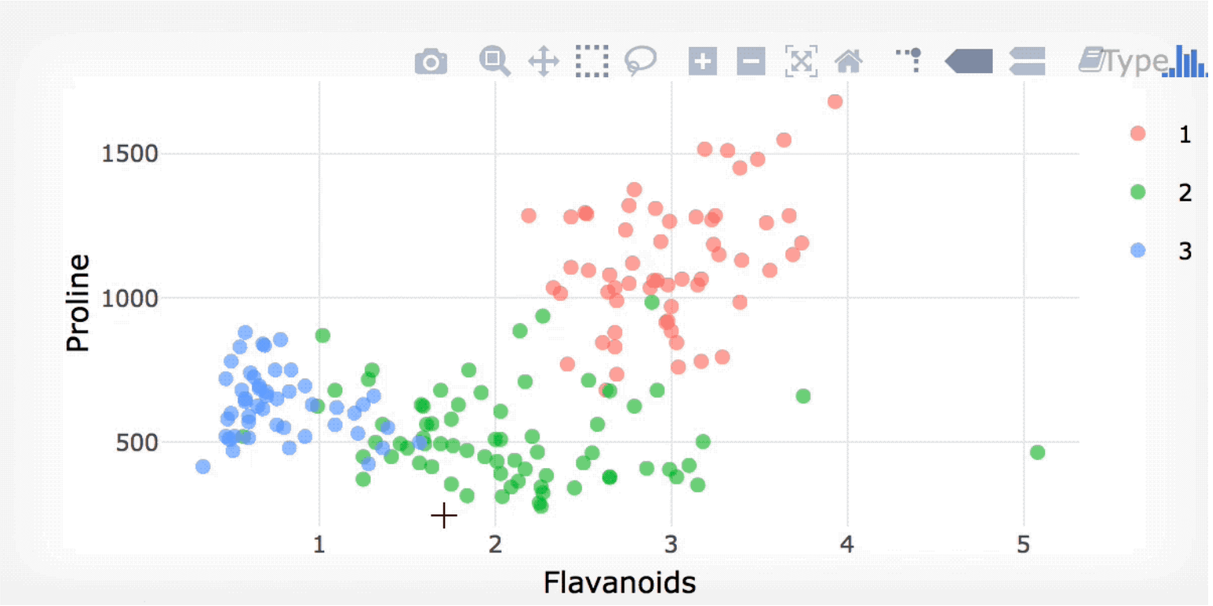

```
library(ggplot2)
static <- wine %>%
  ggplot(aes(x = Flavanoids, y = Proline, color = Type)) +
  geom_point()
```

- Dataset, `wine`
- Aesthetics, `aes()`
- Add a layer, `geom_point()`

ggplotly()

```
library(plotly)
```

```
ggplotly(static)
```



Remarks

- Interactive ? Good
  - Bad design = bad interactive graphic
  - Follow data-viz best practices
- ggplotly() is only the beginning

Let's practice!

INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

Plotting a single variable

INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

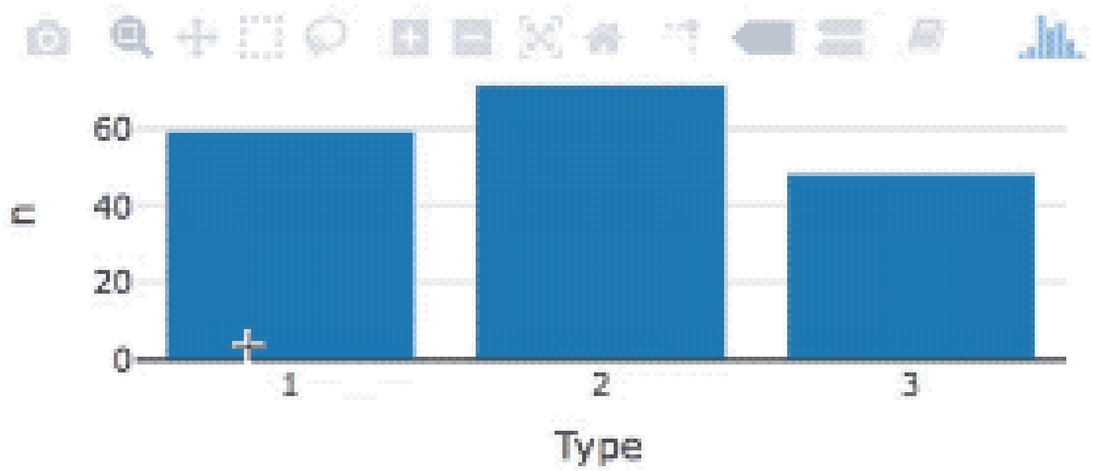
Adam Loy
Statistician, Carleton College

Exploring the wine data

```
library(dplyr)
glimpse(wine)
```

```
Observations: 178
Variables: 14
$ Type            <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1...
$ Alcohol         <dbl> 14.23, 13.20, 13.16, 14.37, 13.24, 14.20, 14.3...
$ Malic           <dbl> 1.71, 1.78, 2.36, 1.95, 2.59, 1.76, 1.87, 2.15...
...
$ Hue             <dbl> 1.04, 1.05, 1.03, 0.86, 1.04, 1.05, 1.02, 1.06...
$ Dilution        <dbl> 3.92, 3.40, 3.17, 3.45, 2.93, 2.85, 3.58, 3.58...
$ Proline         <int> 1065, 1050, 1185, 1480, 735, 1450, 1290, 1295,...
```

Bar charts with plotly

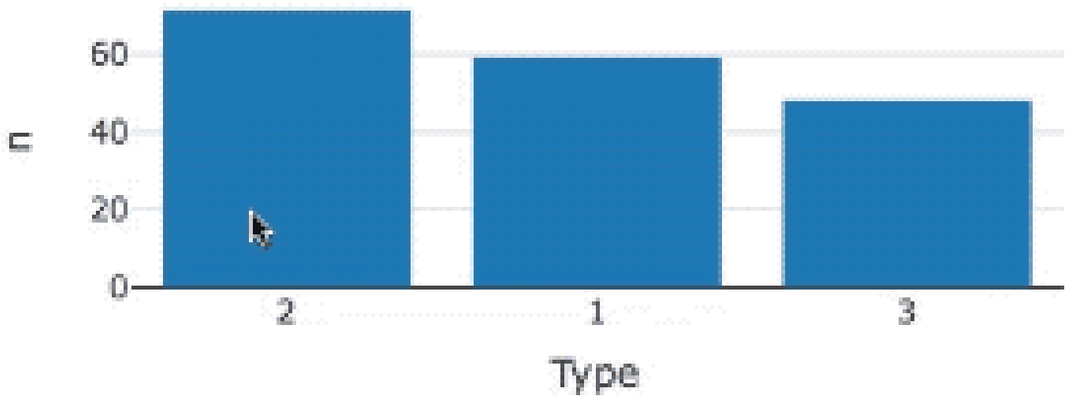



```
library(plotly)

wine %>%
   count(Type) %>%
   plot_ly(x = ~Type, y = ~n) %>%
   add_bars()
```

- Create a frequency table with `count()`
- Specify aesthetics using `~`
- Add the bars trace with `add_bars()`

Reordering the bars

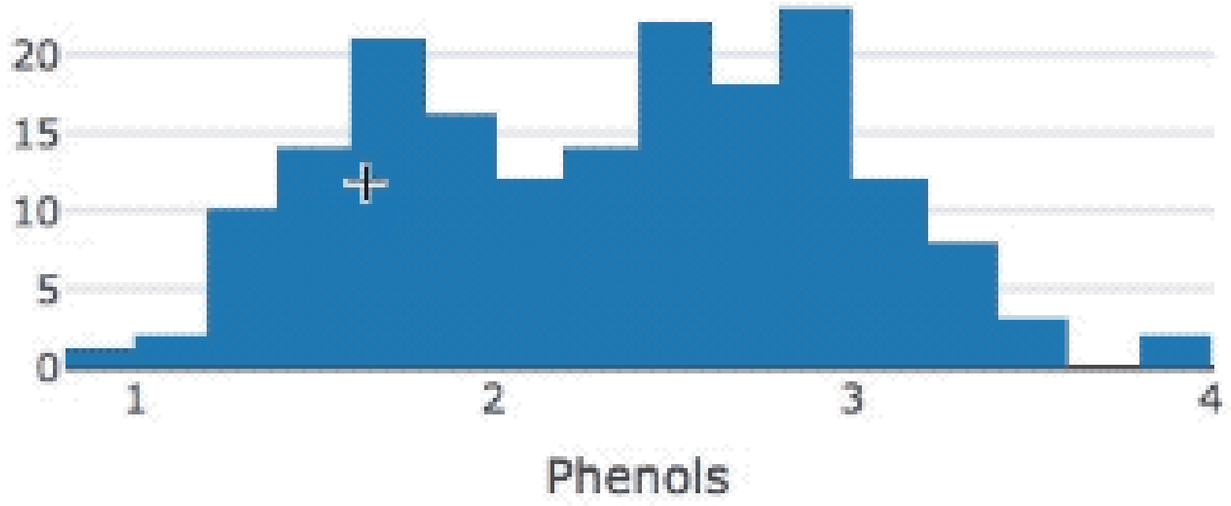



```
library(forcats)
```

```
wine %>%
  count(Type) %>%
  mutate(Type = fct_reorder(Type, n, .desc = TRUE)) %>%
  plot_ly(x = ~Type, y = ~n) %>%
  add_bars()
```

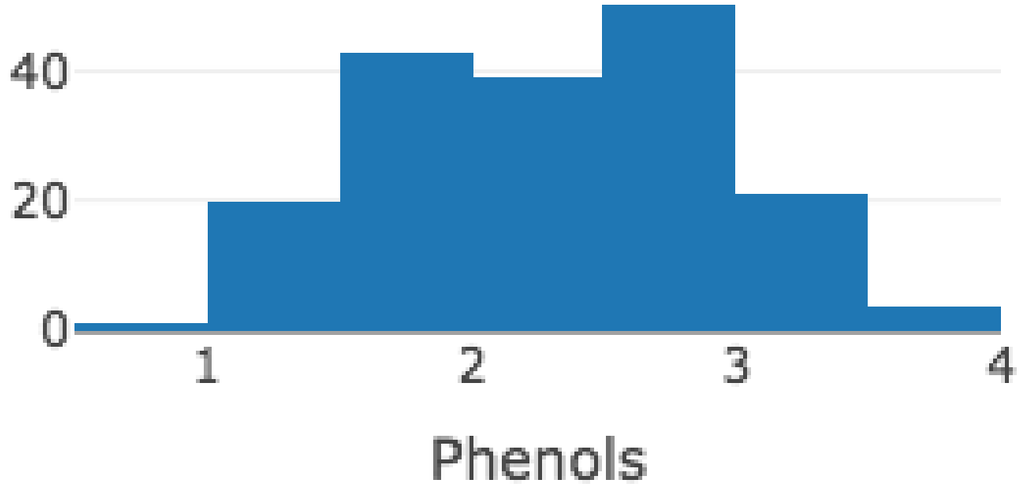
- `fct_reorder()` to rearrange the bars
- set `.desc` argument to `TRUE`

Histograms with plotly



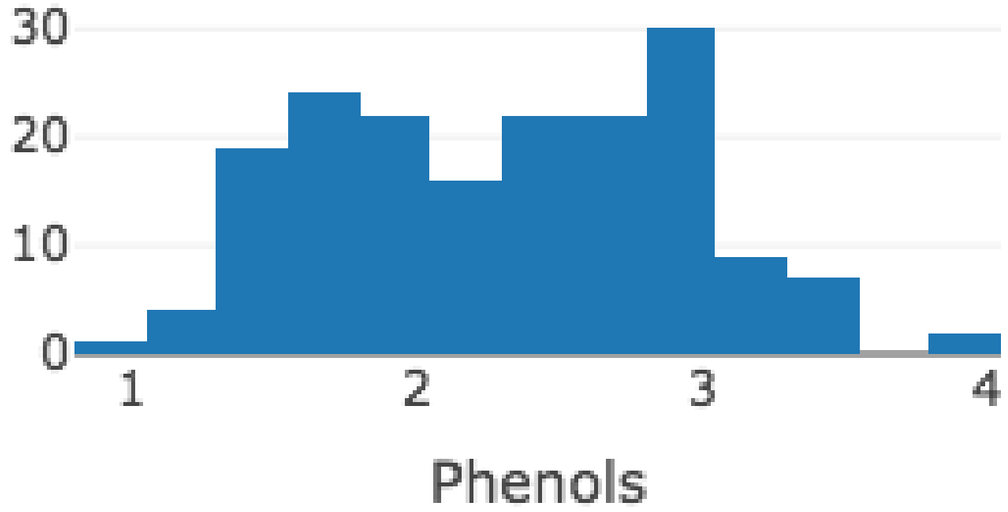

```
wine %>%
   plot_ly(x = ~Phenols) %>%  # specify aesthetics
   add_histogram()            # add the histogram trace
```

Adjusting the number of bins




```
wine %>%
    plot_ly(x = ~Phenols) %>%
    add_histogram(nbinsx = 10)
```

Adjusting the bin width




```
wine %>%
    plot_ly(x = ~Phenols) %>%
    add_histogram(xbins = list(start = 0.8, end = 4, size = 0.25))
```

Let's practice!

INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

Bivariate graphics

INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

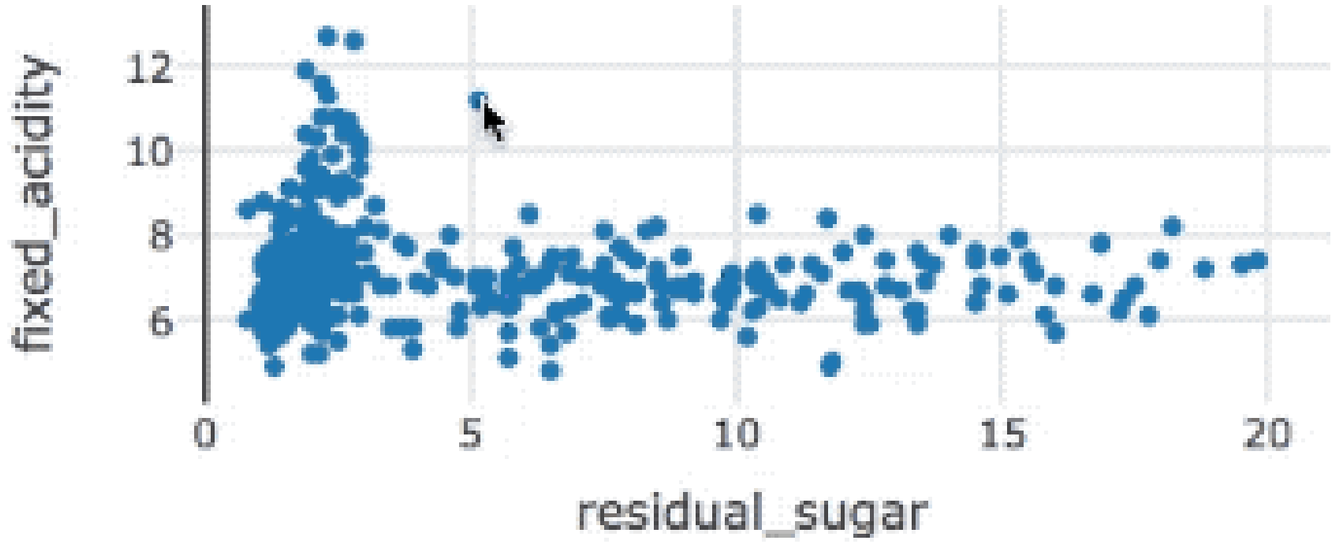
Adam Loy
Statistician, Carleton College

Wine quality data

```
glimpse(winequality)
```

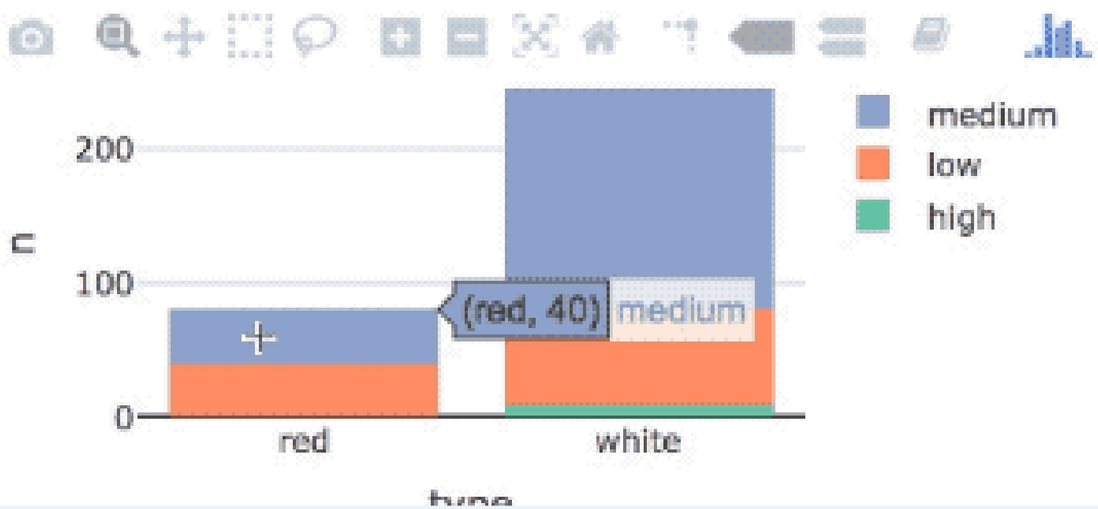
```
Observations: 325
Variables: 14
$ type             <chr> "red", "red", "red", "red", "red", "red", ...
$ fixed_acidity    <dbl> 8.2, 8.2, 8.0, 10.2, 8.6, 6.1, 10.7, 9.1, 7.2...
$ volatile_acidity <dbl> 0.885, 0.640, 0.715, 0.360, 0.520, 0.590, 0.6...
$ citric_acid      <dbl> 0.20, 0.27, 0.22, 0.64, 0.38, 0.01, 0.22, 0.3...
$ residual_sugar   <dbl> 1.40, 2.00, 2.30, 2.90, 1.50, 2.10, 2.70, 2.1...
...
$ sulphates        <dbl> 0.46, 0.62, 0.54, 0.66, 0.52, 0.56, 0.98, 0.8...
$ alcohol          <dbl> 10.0, 9.1, 9.5, 12.5, 9.4, 11.4, 9.9, 11.2, 1...
$ quality          <int> 5, 6, 6, 6, 5, 5, 6, 6, 6, 7, 6, 5, 4, 6, 6, ...
$ quality_label    <chr> "low", "medium", "medium", "medium", "low", ...
```

Scatterplots with plotly



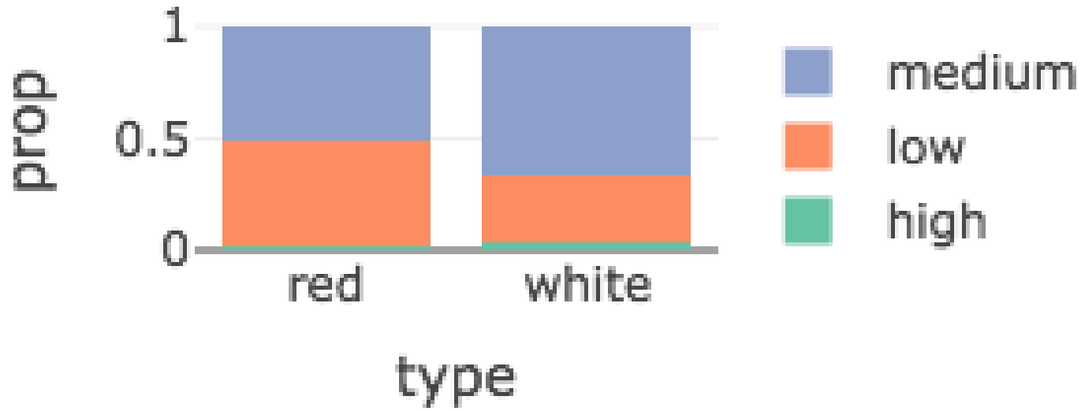

```
winequality %>%
  plot_ly(x = ~residual_sugar, y = ~fixed_acidity) %>%
  add_markers()
```

Stacked bar charts with plotly




```
winequality %>%
  count(type, quality_label) %>%
  plot_ly(x = ~type, y = ~n, color = ~quality_label) %>%
  add_bars() %>%
  layout(barmode = "stack")
```

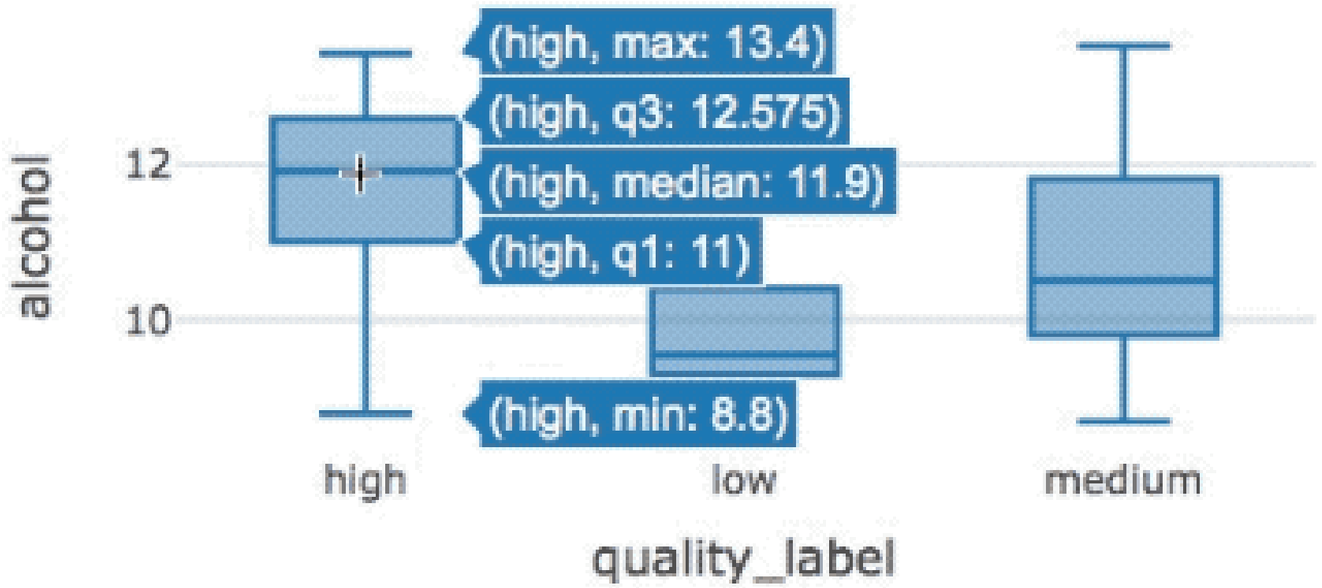
From counts to proportions




```
winequality %>%
  count(type, quality_label) %>%
  group_by(type) %>%
  mutate(prop = n / sum(n)) %>%
  plot_ly(x = ~type, y = ~prop, color = ~quality_label) %>%
  add_bars() %>%
  layout(barmode = "stack")
```

- Group the table with `group_by()`
- Calculate the proportions

Boxplots with plotly




```
winequality %>%
  plot_ly(x = ~quality_label, y = ~alcohol) %>%
  add_boxplot()
```

Let's practice!

INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R