Data Visualization with `ggplot2`

Allison Theobold

This reference card provides you with the recipes for exploratory graphics using **`ggplot2`**. For each recipe you will need to consider what ingredients need to be modified (added or subtracted), such as the data set (`data_set`), the variable name(s) (`variable`, `group_variable`, `x_variable`, `y_variable`), and any axis labels you wish to change.

Univariate Graphics

Bar Chart

```
ggplot(data = <data_set>, mapping = aes(x = <variable>)) +
  geom_bar() +
  labs(x = "Variable Name")
```

Histogram

```
ggplot(data = <data_set>, mapping = aes(x = <variable>)) +
  geom_histogram(binwidth = <NUMBER>, color = "COLOR") +
  labs(x = "Variable Name")
```

Density Plot

```
ggplot(data = <data_set>, mapping = aes(x = <variable>)) +
  geom_density() +
  labs(x = "Variable Name")
```

Histogram with Density Plot

```
ggplot(data = <data_set>) +
  geom_density(mapping = aes(x = <variable>)) +
  geom_histogram(mapping = aes(x = <variable>, y = ..density..),
                 binwidth = <NUMBER>, color = "COLOR")
  labs(x = "Variable Name")
```

Multivariate Graphics

Scatterplot

```
ggplot(data = <data_set>, mapping = aes(x = <x_variable>, y = <y_variable>)) +
  geom_point() +
  labs(x = "Name of x Variable (with units)",
       y = "Name of y Variable (with units)")
```

Scatterplot with (Linear) Smoother

```
ggplot(data = <data_set>, mapping = aes(x = <x_variable>, y = <y_variable>)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(x = "Name of x Variable (with units)",
       y = "Name of y Variable (with units)")
```

Boxplot

```
ggplot(data = <data_set>, mapping = aes(x = <group_variable>, y = <variable>)) +
  geom_boxplot() +
  coord_flip() +
  labs(x = "Group Name",
       y = "Variable Name (with units)")
```

Violin Plot

```
ggplot(data = <data_set>, mapping = aes(x = <group_variable>, y = <variable>)) +
  geom_violin() +
  labs(x = "Group Name",
       y = "Variable Name (with units)")
```

Facetted Histograms

```
ggplot(data = <data_set>, mapping = aes(x = <variable>)) +
  geom_histogram() +
  labs(x = "Variable Name") +
  facet_wrap(~ <group_variable>)
```

Side-by-Side Boxplots

```
ggplot(data = <data_set>, mapping = aes(x = <group_variable>, y = <variable>)) +
  geom_boxplot() +
  labs(x = "Group Name",
       y = "Variable Name (with units)")
```