# A Two Page Guide to ggplot2

Andy Grogan-Kaylor (http://www.umich.edu/~agrogan)

September 30, 2018

[Share on Twitter]

## 1   Why?

A great deal of data analysis and visualization involves the same core set of steps: get some data, clean it up a little, run some descriptive statistics, run some bivariate statistics, create a graph or a visualization. **ggplot2** can be an important part of a replicable, automated, documented workflow for complex projects.

    Given the fact that we often want to apply the same core set of tasks to new questions and new data, there are ways to overcome the steep learning curve and learn a replicable set of commands that can be applied to problem after problem.

    The same 5 to 10 lines of ggplot2 code can often be tweaked over and over again for multiple projects.

## 2   The Essential Idea Of ggplot2 Is Simple

There are 3 essential elements to any ggplot call:

1. An *aesthetic* that tells ggplot which variables are being mapped to the *x axis*, *y axis*, (and often other attributes of the graph, such as the *color fill*). Intuitively, the aesthetic can be thought of as **what you are graphing**.
2. A *geom* or *geometry* that tells ggplot about the basic structure of the graph. Intuitively, the geom can be thought of as **how you are graphing it**.
3. Other options, such as a *graph title*, *axis labels* and *overall theme* for the graph.

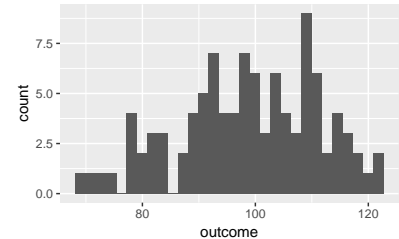## 3   Get Started

```
library(ggplot2)  # beautiful graphs

library(ggthemes)  # nice themes for ggplot2
```

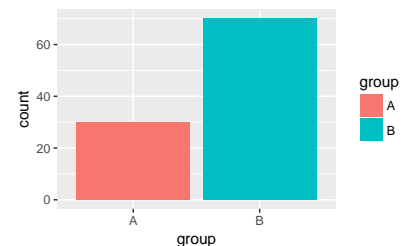## 4    Some Examples

### 4.1    One Continuous Variable

```
# anything that starts with a '#' is a comment

ggplot(mydata, # the data I am using
       aes(x = outcome)) + # the variable I am using
  geom_histogram() # how I am graphing it
```
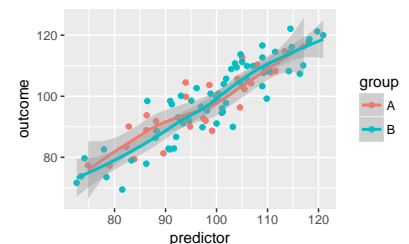


### 4.2    One Categorical Variable

```
ggplot(mydata, # the data I am using
       aes(x = group,
           fill = group)) + # the variable I am using
  geom_bar() # how I am graphing it
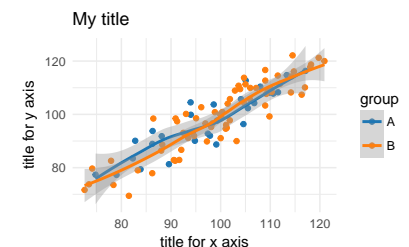```



### 4.3    Continuous by Continuous

```
ggplot(mydata, # the data I am using
       aes(x = predictor,
           y = outcome,
           color = group)) + # the variables I am using
  geom_point() + # how I am graphing it
  geom_smooth()
```



### 4.4    Add Some Options

```
ggplot(mydata, # the data I am using
       aes(x = predictor, y = outcome,
           color = group)) + # the variables I am using
  geom_point() + # how I am graphing it
  geom_smooth() +
  labs(title = "My title",
       x = "title for x axis", y = "title for y axis") +
  scale_color_tableau() + theme_minimal()
```



## 5    More Information

More information can be found here: http://www-personal.umich.edu/~agrogan/ggplot2