PRU LearnR - Pedestrian to Novice Series

Prospect RNinjas

Invalid Date

Table of contents

1 This is the index

just a test

```
# set a random seed and generate data
set.seed(123)
x <- rnorm(100)

# calculate mean
mean(x)
```

2 ggplot

> **⚠ Warning**
>
> Installing and loading ggplot2 on webR takes a little while. The install is happening in the background. Don't worry, once you've waited to load the package everything else will be quick.

2.0.1 ggplot

- Very popular plotting package
- Good plots quickly
- Declarative - describe what you want not how to build it
- Contrasts w/Imperative - how to build it step by step

2.0.2 Basics

- Load the package and some data

```
library(ggplot2)
```

- To build a plot using `ggplot` we start with the `ggplot()` function

```
ggplot()
```

- `ggplot()` creates a base ggplot object that we can then add things to - like a blank canvas
- We can also add optional arguments for information to be shared across different components of the plot
- The two main arguments we typically use here are
- `data` - which is the name of the data frame we are working with, so `acacia`
- `mapping` - which describes which columns of the data are used for different aspects of the plot

- We create a `mapping` by using the `aes` function, which stands for "aesthetic", and then linking columns to pieces of the plot
- We'll start with telling ggplot what value should be on the x and y axes
- Let's plot the relationship betwen the circumference of an acacia and its height

```
ggplot(data = mtcars, mapping = aes(x = mpg, y = hp))
```

- This still doesn't create a figure, it's just a blank canvas and some information on default values for data and mapping columns to pieces of the plot
- We can add data to the plot using layers
- We do this by adding a `+` after the the `ggplot` function and then adding something called a `geom`, which stands for `geometry`
- To make a scatter plot we use `geom_point`

```
ggplot(data = mtcars, mapping = aes(x = mpg, y = hp)) +
  geom_point()
```

- To change things about the layer we can pass additional arguments to the `geom`
- We can do things like change

  - the `size` of the points, we'll set it to `3`
  - the `color` of the points, we'll set it to `"blue"`
  - the transparency of the points, which is called `alpha`, we'll set it to 0.5

```
ggplot(data = mtcars, mapping = aes(x = mpg, y = hp)) +
  geom_point(size = 3, color = "blue", alpha = 0.5)
```

- Try changing these values to make the graph look like you want it to

- To add labels (like documentation for your graphs!) we use the `labs` function

```
ggplot(data = mtcars, mapping = aes(x = mpg, y = hp)) +
  geom_point(size = 3, color = "blue", alpha = 0.5) +
  labs(x = "Miles per Gallon", y = "Horse Power",
       title = "Car stuff")
```

2.0.3 Grouping

- Group on a single graph
- Look at influence of experimental treatment

```
ggplot(data = mtcars, mapping = aes(x = mpg, y = hp, color = cyl)) +
  geom_point(size = 3, alpha = 0.5)
```

- Try changing the above code to color based on the `gear`

- We can also split each group into different subplots (known as facets)

```
ggplot(data = mtcars, mapping = aes(x = mpg, y = hp)) +
  geom_point(size = 3, alpha = 0.5) +
  facet_wrap(~cyl)
```

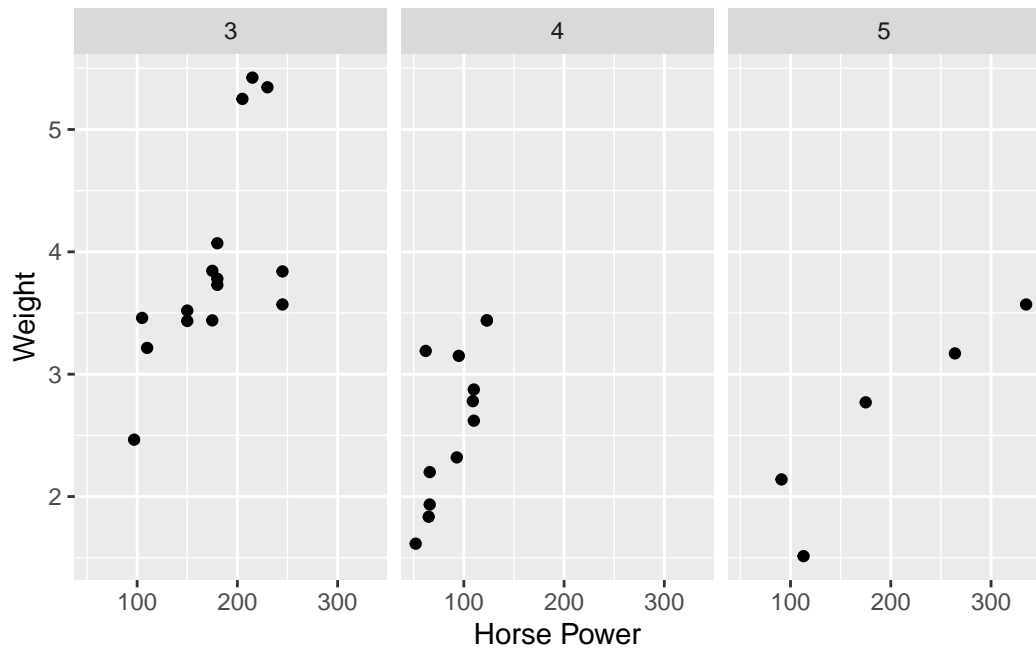- Try changing this code to create a subplot for each value in `vs` with points of size 4

> **i** Exercise
>
> Make a scatter plot with `hp` on the x axis and `wt` on the y axis. Label the x axis "Horse Power" and the y axis "Weight". Make one subplot for each value in `gear`.
>
> ```
> # Add you code here
> ```
>
> Your result should look like the plot below

Expected Result

Solution Code

```r
library(ggplot2)
ggplot(mtcars, aes(x = hp, y = wt)) +
  geom_point() +
  labs(x = "Horse Power", y = "Weight") +
  facet_wrap(~gear)
```