

W5

Irlanda Ayon-Moreno

4/10/2018

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.4.2
```

```
library(ggplot2)
```

dplyr

filter: keep rows matching criteria

select: pick columns by name

arrange: reorder rows

mutate: add new variables

summarise: reduce variables to values

Structure

- 1st argument is a data frame
- next arguments say what to do with the data frame
- always returns a data frame

ggplot2

Automated things in ggplot2: * axis labels

* legends

* color for points

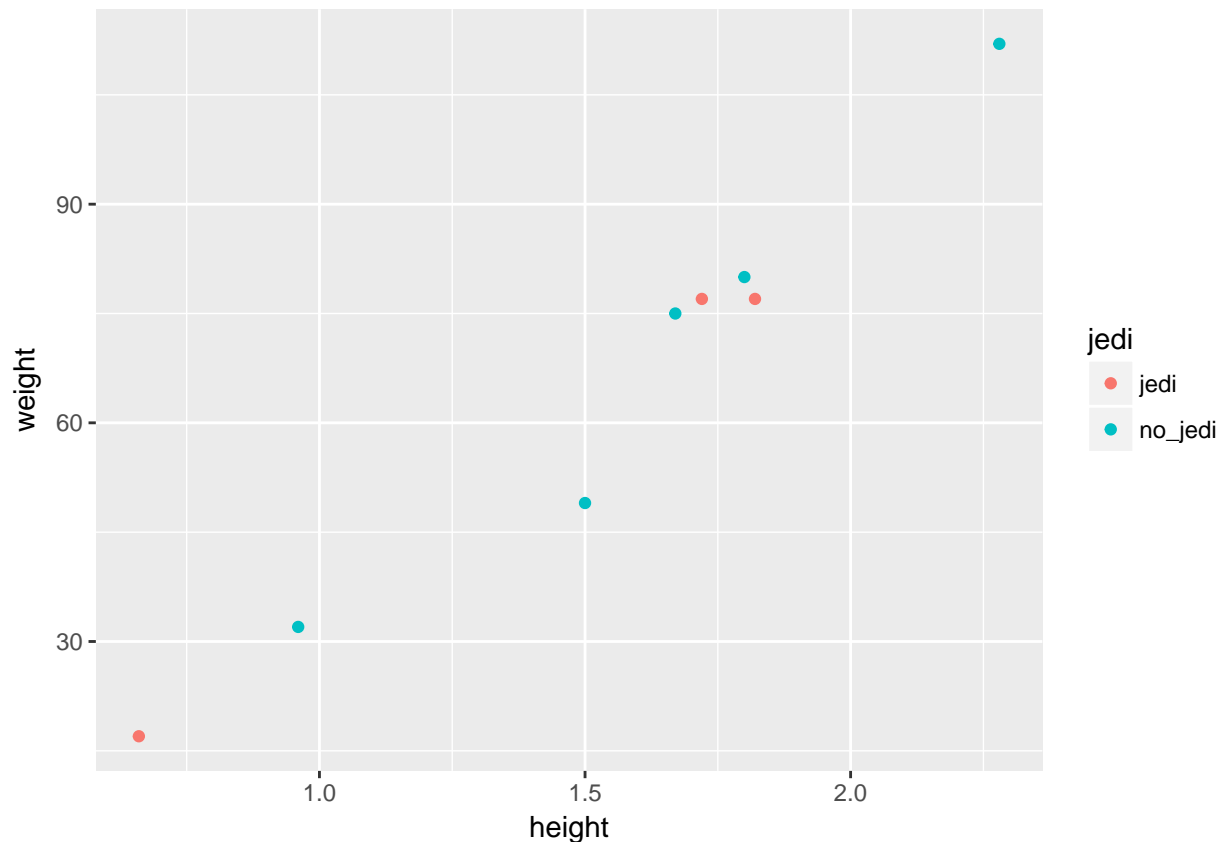
* grid lines

* axis tick marks

```
name <- c("Luke Skywalker", "Leia Skywalker", "Obi-Wan Kenobi", "Han Solo", "R2-D2", "C-3PO", "Yoda")
gender <- c("male", "female", rep("male", 6))
height <- c(1.72, 1.5, 1.82, 1.8, 0.96, 1.67, 0.66, 2.28)
weight <- c(77, 49, 77, 80, 32, 75, 17, 112)
jedi <- c("jedi", "no_jedi", "jedi", rep("no_jedi", 3), "jedi", "no_jedi")
species <- c(rep("human", 4), rep("droid", 2), "yoda", "wookie")
weapon <- c(rep(c("lightsaber", "blaster"), 2), rep("unarmed", 2), "lightsaber", "bowcaster")
starwars_df <- data.frame(name, gender, height, weight, jedi, species, weapon)
starwars_df
```

##		name	gender	height	weight	jedi	species	weapon
## 1	Luke Skywalker	male	1.72	77	jedi	human	lightsaber	
## 2	Leia Skywalker	female	1.50	49	no_jedi	human	blaster	
## 3	Obi-Wan Kenobi	male	1.82	77	jedi	human	lightsaber	
## 4	Han Solo	male	1.80	80	no_jedi	human	blaster	
## 5	R2-D2	male	0.96	32	no_jedi	droid	unarmed	
## 6	C-3P0	male	1.67	75	no_jedi	droid	unarmed	
## 7	Yoda	male	0.66	17	jedi	yoda	lightsaber	
## 8	Chewbacca	male	2.28	112	no_jedi	wookiee	bowcaster	

```
ggplot(data = starwars_df) +
  geom_point(mapping = aes(x = height, y = weight, color = jedi))
```



A graphic is the mapping from data to aesthetic attributes (color, shape, size) of geometric objects (points, lines, bars, etc)