

# Class 14: Midterm project conferences and R questions

---

June 8, 2018



# General

# Announcements

- Reading 11: **Introductory Statistics with Randomization and Simulation**
  - From chapter 2: from the beginning through to the end of section 2.3
- Homework 3 due by 11:59pm on Tuesday, June 12th
- Rough draft of your question responses due on Monday

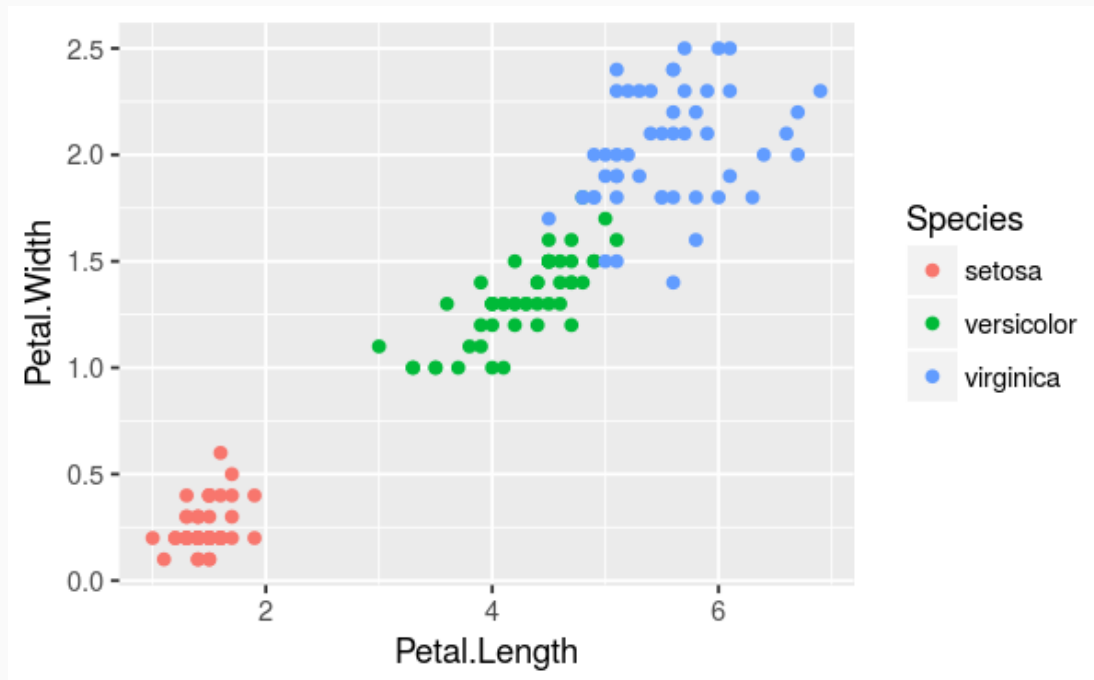
# Midterm project: group conferences

- I will meet with Team 1 first, then Team 2.
- While your team is not conferencing with me, use the time to work on your project.
- Your team should have a plan on how you will work on the project over the weekend, so that everyone has a rough draft for their part of the **Exploratory Data Analysis** section of the report

# Questions on course material to date

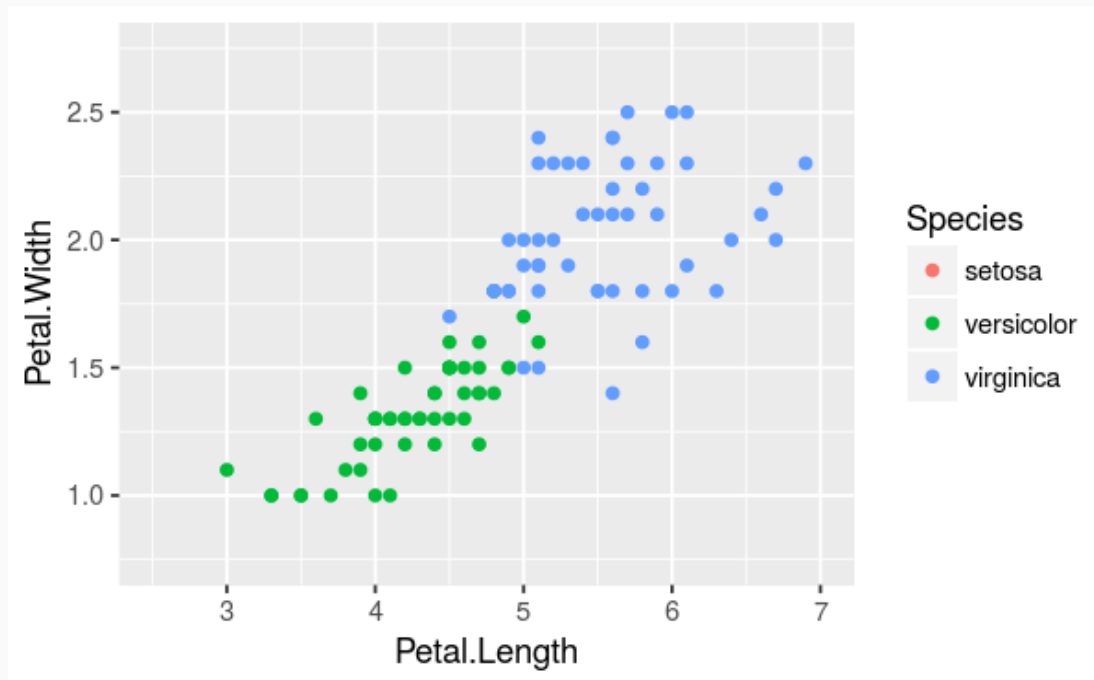
# ggplot2: Changing the viewing window

```
as_data_frame(iris) %>%  
  ggplot() +  
  geom_point(  
    mapping = aes(x = Petal.Length, y = Petal.Width, color = Species)  
  )
```



# ggplot2: Changing the viewing window

```
as_data_frame(iris) %>%  
  ggplot() +  
  geom_point(  
    mapping = aes(x = Petal.Length, y = Petal.Width, color = Species)  
  ) +  
  coord_cartesian(xlim = combine(2.5, 7), ylim = combine(0.75, 2.75))
```



# ggplot2: Save the image to disk

Immediately after you create a plot:

```
as_data_frame(iris) %>%  
  ggplot() +  
  geom_point(  
    mapping = aes(x = Petal.Length, y = Petal.Width, color = Species)  
  ) +  
  coord_cartesian(xlim = combine(2.5, 7), ylim = combine(0.75, 2.75))
```

Run `ggsave()` as follows to save the image to disk:

```
ggsave("iris_petal_width_vs_length.png", dpi = 120)
```

Increase the `dpi` input to increase the figure's resolution.



# dpLyr: Rename columns

```
iris_df <- as_data_frame(iris) %>%  
  rename(  
    petal_length = Petal.Length,  
    petal_width = Petal.Width,  
    sepal_length = Sepal.Length,  
    sepal_width = Sepal.Width,  
    species = Species  
  )
```

sepal_length	sepal_width	petal_length	petal_width	species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa

# dp<sub>lyr</sub>: Recode categorical levels

```
college
```

UNITID	OPEID	OPEID6	INSTNM	CITY	STABBR	ZIP
100654	00100200	001002	Alabama A & M University	Normal	AL	35762
100663	00105200	001052	University of Alabama at Birmingham	Birmingham	AL	35294-0110
100690	02503400	025034	Amridge University	Montgomery	AL	36117-3553
100706	00105500	001055	University of Alabama in Huntsville	Huntsville	AL	35899
100724	00100500	001005	Alabama State University	Montgomery	AL	36104-0271

# dpLyr: Recode categorical levels

```
college %>%  
  mutate(  
    STABBR = recode(STABBR, AL = "Alabama")  
  )
```

UNITID	OPEID	OPEID6	INSTNM	CITY	STABBR	ZIP
100654	00100200	001002	Alabama A & M University	Normal	Alabama	35762
100663	00105200	001052	University of Alabama at Birmingham	Birmingham	Alabama	35294-0110
100690	02503400	025034	Amridge University	Montgomery	Alabama	36117-3553
100706	00105500	001055	University of Alabama in Huntsville	Huntsville	Alabama	35899
100724	00100500	001005	Alabama State University	Montgomery	Alabama	36104-0271