# Lecture 15.5 Final Report Tips

UNIVERSITY OF
SAN FRANCISCO

Abbie M Popa

BSDS 100 - Intro to Data Science with R

# Outline

- Updates
- "New" Final Project
- Review of Control Flow Lab
- Functions

# Updates

- Final Project is now a report (we will discuss this at length) **due December 7 at midnight**
- Nov 27 (today): Function Lecture
- Nov 29 (Thursday): Function Lab
- Dec 4 (Tuesday): Extra review, wrap-up, teaching evals

# Update - No Oral Presentations

- Because project is now a report, no oral presentation
- If you would still like to see presentation tips, they are on github with final materials
- Can schedule an ungraded presentation if you are too sad

# Update - Extra Review

- Optional Assignment on Canvas
- "Due" Sunday December 2 at midnight

# Update - Final Project

- Now due December 7 at midnight
- Now report
  - Introduction
  - Body
  - Conclusion
- Example report in canvas announcement

# Final Project - Introduction

- What is the big overarching question or problem?
- Is there any background information we need to understand your data?
- What dataset are you using to answer the question?
- What analyses will you be completing to answer the question?

# Final Project - Body

- For EACH analysis
  - What did you do?
  - Why did you do it?
  - Some code and output (graph, table, number, etc.)
  - What were your findings?

# Final Project - Conclusion

- Remind me what your big question or problem was
- Summarize the analyses you completed and what you found
- Relate these findings to your big question
- If applicable, what actions might someone take based on these conclusions?

# Final Project - Formatting

- Formatting is 15/75 points, or 20% of the grade
- File must be a knit pdf or you will lose substantial points
- Please make use of headers, plain text, and code chunks (see example)
- Code should be neat and use the multi-line format where possible

# Review of Control Flow Lab

Data located at:

https://raw.githubusercontent.com/abbiepopa/BSDS100/master/Data/titanic.csv

## titanic.csv

1. Using a `for()` loop and an `if()` conditional, recode the entries in the `Survived` variable with `"Survived"` and `"Perished"` into a new column `survived_text`
2. Using the `if()` command and loop, create a new variable of type `ordered factor` in the data frame called `ageClass`, and map `Age` to: `"Minor"` if less than 18 yrs; 18 yrs $\leq$ `"Adult"` $\leq$ 65 yrs; and `"Senior"` if older than 65 yrs
3. Using a `switch()` statement, identify each passenger class, `Pclass`, as either `"First Class"`, `"Business Class"` or `"Economy"`, and print the results to the console