Precept 2

Jordan Klein (Demography PhD Student)
jdklein@Princeton.edu
POL 345

Agenda

- Check in
- Logistics
- R tips/best practices
- Difference-in-means
- Precept Questions

Check In

- This class is picking up its pace very fast
- It’s okay to be confused/overwhelmed
- Problems with turning in work/missing precept are okay
  - 1 requirement = **Please communicate with me!**
  - (*Thanks to those who have been doing so*)

Let’s Share

1. 1 thing that’s confusing you about the logistics or material of this class
2. Nerdy guilty internet pleasure

General Logistics

No longer need to use Apporto!

Problem Set 1 Logistics

- Due Wednesday 10/2 at 3 PM EDT
- Work in pairs
  - If looking for a partner post here
  - Your precept-mates make great partners!
- *May **only look at/share code with your partner***
- *May **only discuss** problem set **with others***
  - Cite them if you do!

Do not plagiarize!

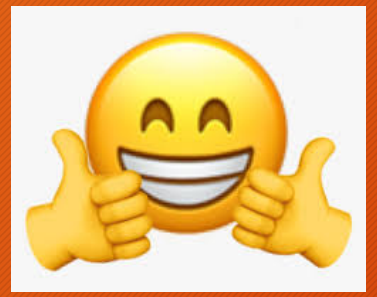
R Tips/Best Practices

Split code into different lines

- Comments & commands can be very long
- Writing them on multiple lines is more readable than using the same line
- **Ignore "+" signs that start new lines in the handouts**
  - This is how the R console prints new lines of the same command

Split code into different lines (comments)

```
2   # Welcome to precept number 2. Today we will be working or
```



```
4   # Welcome to precept number 2.
5   # Today we will be working on several practice problems.
```

Split code into different lines (commands)

Script

```
7  mean(c(5, 4, 6, 7, 8, 9, 23, 45, 34, 23, 56, 78, 45, 34,
8          2, 1, -10))
```

Console

```
> mean(c(5, 4, 6, 7, 8, 9, 23, 45, 34, 23, 56, 78, 45, 34,
+        2, 1, -10))
[1] 21.76471
```

Tips for error messages

- Google is your best friend!
  - Simply copy paste
  - *Not just for error messages*
- *Work backwards from the code that returns the error message to identify where it's coming from*

More R tips/best practices

- Run all previous lines of code before the current one
  Code is cumulative!
- Order of arguments in a function don't matter
  func(x = a, y = b) is the same as *func(y = b, x = a)*
- Make generous use of R Documentation to learn how function work
  ?func

Difference-in-means

- How to find the average total causal effect across all your data?
  -> **Average Treatment Effect** (ATE):

$$\frac{1}{N}\sum_{i=1}^{N}(Y_i(1) - Y_i(0))$$

  - This is the difference in mean outcomes between the 2 treatment groups!
    - In R:

mean(data$outcome[data$treatment == 1]) - mean(data$outcome[data$treatment == 0])

Precept Questions

- Oil, Islam, & female labor force participation across countries
  - Does oil or Islam cause low female participation?
- Paper
- Blog post
- Data: RossOilWomenIslamData.csv
- Variables we'll use (z-scaled):
  - *islam: Proportion of the country that is Muslim*
  - *femlabor: A measure of female participation in the workforce*
  - *oil: Oil rents per capita, logged*

Question 1

1. Calculate the difference-in-means for *femlabor b*etween countries with high and low levels of Islam (i.e. between countries with positive and negative or zero values of *islam*)
2. Calculate the difference-in-means for *femlabor b*etween countries with high and low levels of oil (i.e. between countries with positive and negative or zero values of *oil*)

- Is higher female labor force participation associated with higher or lower levels or Islam/oil?
  - Is this causal? Why or why not?

Question 2

- Calculate the difference-in-means for *femlabor*
  - Between countries with high and low levels of oil only for countries with a **high** level of Islam
  - Between countries with high and low levels of oil only for countries with a **low** level of Islam
- How does the relationship between oil and female labor participation vary between countries with a high and low level of Islam?

Question 3a

Create 2 figures- density plots of female labor participation in countries with high levels vs. low levels of Islam.

1. Countries with low oil rents
2. Countries with high oil rents

Graphs should:

i. Contain the density for female labor participation for countries with a high level of Islam (in red) and a low level of Islam (in blue).

ii. The figure should have a red or blue dashed vertical line at the mean level of female labor participation, corresponding with each density.

iii. Each plot should have informative axis labels and titles.

iv. Each plot should have a legend. If necessary, you should lengthen the Y-axis so that the legend does not overlap with the density plot.

Question 3b

- Is the effect of Islam on female labor force participation higher in countries with low or high oil rents?
- Why are we treating oil as a confounder? How do we control for it?