

What is Econometrics?

EC 320: Introduction to Econometrics

Emmett Saulnier
Spring 2022

Prologue

Who am I?

Emmett Saulnier: emmetts@uoregon.edu (use EC 320 in the subject line)

Office hours: Wed 11a-12p, Fri 10a-11a, or by appointment in **PLC 520**

- 3rd year Doctoral student in the economics department
- First time teaching as a sole instructor! But I taught the lab section for EC 421 (Upper level 'metrics class) last term
- My research falls at the intersection of the environment and public policy

Before grad school...

- I grew up in Cincinnati, OH
- Studied economics and business at the University of Virginia
- Worked as a data analyst for Carmax (Largest used car retailer in the US), mostly as a part of their marketing team

Who am I?



Who am I?

Before I started grad school, I hiked 2,650 miles from Mexico to Canada on the Pacific Crest Trail



Who am I?

In January I found a sick 5 week old kitten all alone in Hendricks Park. My partner and I decided to keep her, say hello to Rainbow!



What will we do?

Coursework

- **W1:** Introduction and Review
- **W2:** The fundamental Econometric Problem
- **W3:** SLR - Estimation
- **W4:** SLR Assumptions
- **W5: Midterm**
- **W6:** SLR - Inference
- **W7:** MLR - Estimation & Inference
- **W8:** Nonlinearity & transformations
- **W9:** Quantitative Variables
- **W10:** Exogeneity & **Finals** Review

Content will be available on Canvas. If you wish to download them faster, see the course [Github](#) page.

What will we do?

GE: **Giorgi Nikolaishvili**

Office hours are **Mon 2-3pm** in PLC
417 or **Wed 3-4pm** on Zoom

Labs - Learning R

Develop practical skills based on the material we cover.

Lab sessions will teach you how to conduct data cleaning, visualize data and produce regression analysis based on simulations and empirical questions.

Crucial for computational portions of assignments.

What will we do?

We will use two textbooks:

- Introduction to Econometrics, 5th Ed. by C. Dougherty (**ItE**)
- Mastering 'Metrics: the Path from Cause to Effect, by Angrist & Pischke (**MM**)

Lectures are not a substitute for these resources, but rather the finishing touch on solidifying your knowledge.

I will post readings for each week on Canvas, do them **before each lecture**. In the next two lectures, we will be covering the **Review** chapter from **ItE**.

What will we do?

Grade

- 20% - Analytical Problem Sets
- 20% - Computational Problem Sets
- 30% - Midterm Exam
- 30% - Final Exam

Good Classroom Citizenship

- Don't distract others around you by goofing off on your laptop or phone
- Don't disrupt the class by showing up 5 minutes late
- The slides will be posted after class -- it is more important to write things down that we talk about which are **not** on the slides
- Stay home if you are feeling *symptomatic*
- Wear a mask if it makes you more comfortable

Motivation

Why study econometrics?

1. Develop **skills that employers value.**
2. Cultivate **healthy skepticism.**
3. Learn about the world using **data.**

Motivation

Why study econometrics?

Provide answers to important questions

- Do minimum wage policies **reduce poverty?**
- Does the death penalty **deter violent crime?**
- Does recreational marijuana **cross state lines?**
- Are recessions **good for your health?**
- How will global warming **affect the economy?**
- What **explains the gender pay gap?**
- Does corruption **impact exporting prospects in developing countries?**

Econometrics

Most econometric inquiry concerns one of two distinct goals:

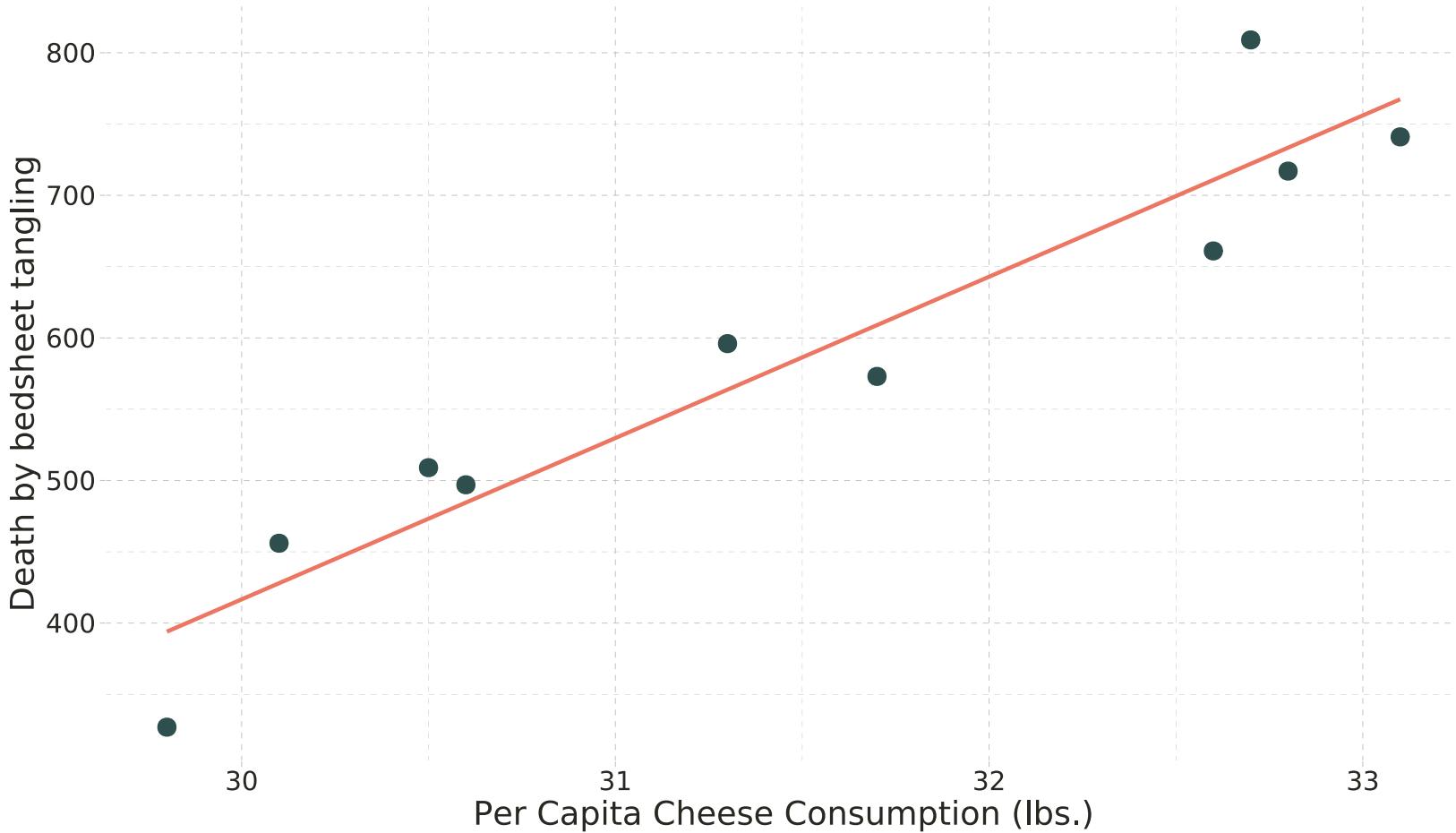
1. **Prediction:** Accurately *predict* or *forecast* an outcome given a set of predictors. *Given what we know about x , what values do we expect y to take?*
2. **Causal identification:** Estimate the effect of an intervention on an outcome. *How does y change when we change x ?*

The main focus of EC 320 and EC 421 is causal identification.

- But...both rely on a common set of statistical techniques.
- For those interested in prediction, Professor Tim Duy teaches forecasting (EC 422)

Econometrics

Not all relationships are causal



Econometrics

Correlation vs. Causation

Common refrain: "*Correlation doesn't necessarily imply causation!*"

- **Q:** Why might correlation fail to describe a causal relationship?
- **A:** Omitted-variables bias, selection bias, simultaneity, reverse causality.

Correlation can imply causation.

- Requires strong assumptions.
- **Real life often violates these assumptions!**
- **Solutions:** Conduct an experiment or find a natural experiment.

Econometrics

Example: *Blue Paradox*

Recent study by UO economist [Grant McDermott](#) and coauthors.

Question: Do commercial fishers preempt fishing bans by increasing their fishing effort before the bans go into effect?

Motivation

- Recent conservation efforts seek to preserve aquatic habitat and increase fish stocks.
- Policy lever: Restrict fishing activity in marine protected areas.
- Concern: Preemptive behavior could *decrease* fish stocks.

Data: Vessel-level data on fishing effort/intensity.

Econometrics

Example: *Blue Paradox*

Natural Experiment

Phoenix Islands Protected Area (PIPA)

- First mentioned on 1 September 2014; implemented 1 January 2015.
- *Treatment group*: PIPA.
- *Control group*: Outlying Kiribati islands.

Econometrics

Example: *Blue Paradox*

Natural Experiment

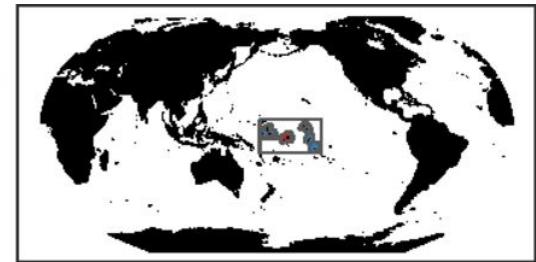
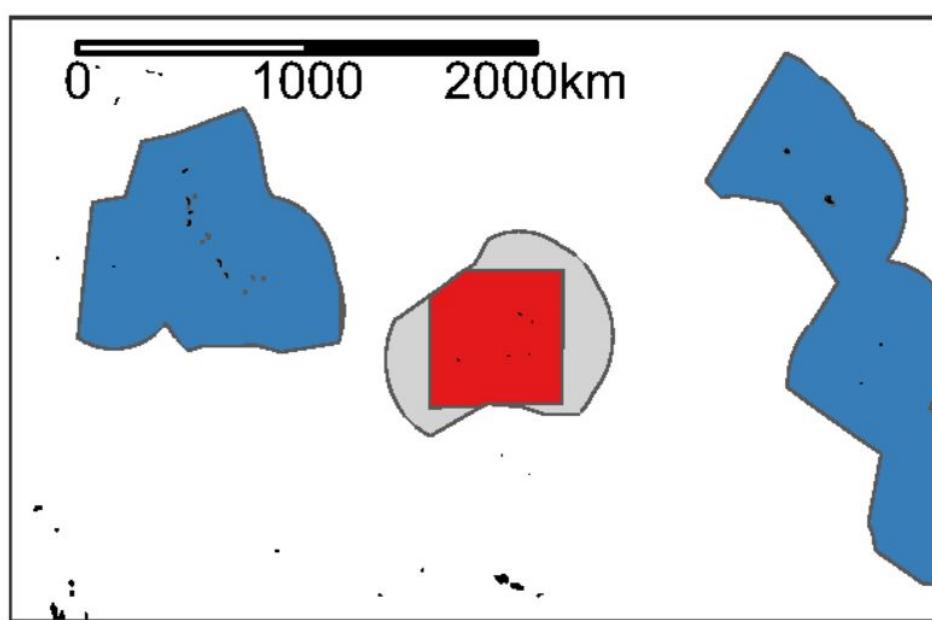
Measure the causal effect of the fishing ban by comparing fishing effort in treatment and control regions, before-and-after PIPA.

- A *difference-in-differences* comparison.
- **Assumption: Parallel trends.** If we believe this assumption, then the observed change supports a causal interpretation. If not, then the change could reflect other factors and thus fail to isolate the causal effect of the ban.

Econometrics

Example: *Blue Paradox*

Results



PIPAs

Kiribati control

Phoenix Islands (Non-PIPA)

Phoenix Islands (Non-PIPA)

Econometrics

Example: *Blue Paradox*

Discussion

Results provide causal evidence that commercial fishers engage in preemptive behavior in response to conservation policy changes.

Results are *consistent* with economic theory, but *cannot prove* that the theory is correct.

- **Science cannot prove anything.**
- Science can **falsify or reject** existing hypotheses or **corroborate** existing evidence.

Econometrics

Also...the causal statement rests on a critical assumption.

- Cannot prove that the assumption is true, but can falsify it.
- Failure to falsify ≠ assumption is true.

What is R?

According to the [R project website](#),

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS.

What does that mean?

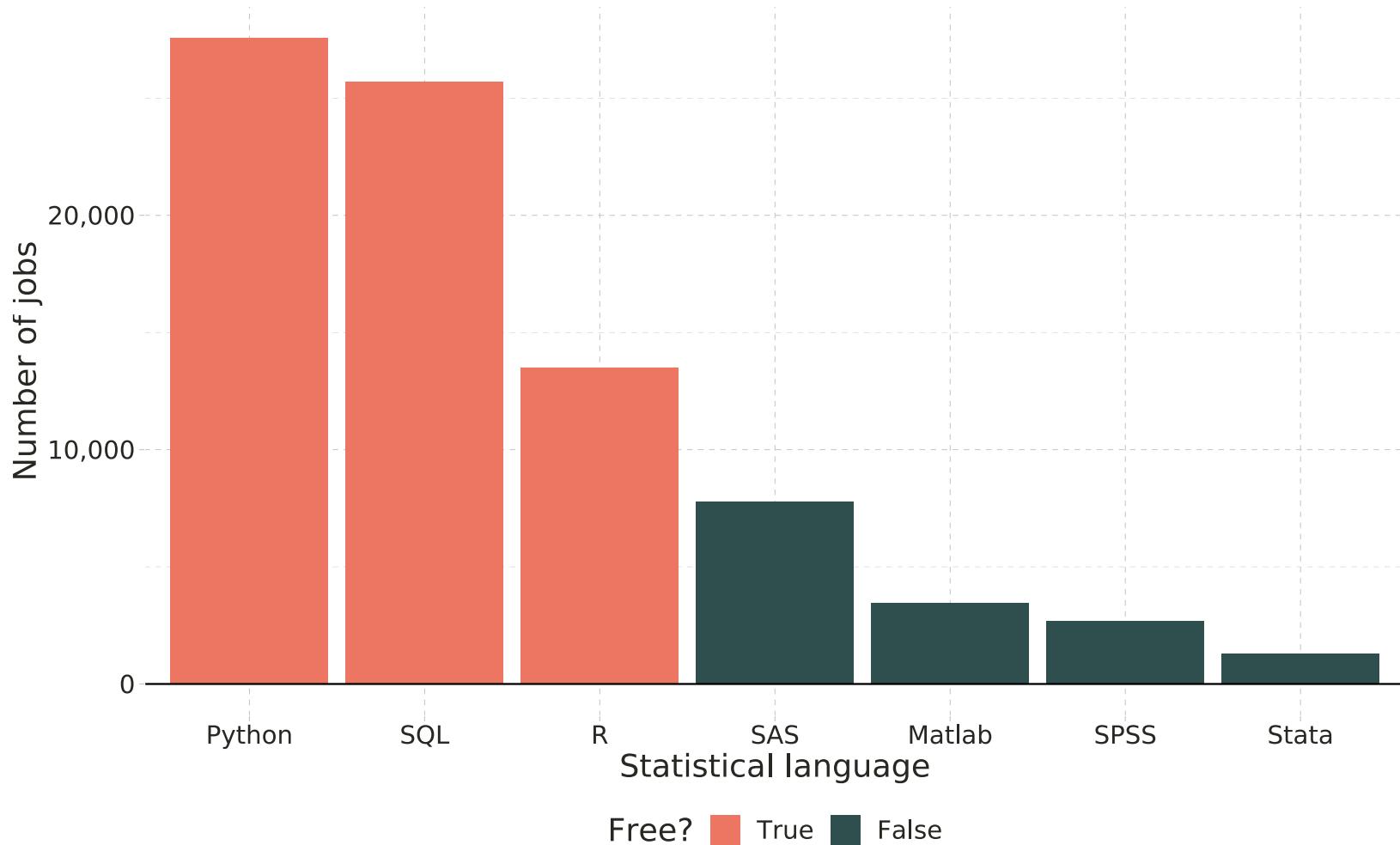
- R is **free** and **open source**.
- R executes a variety of statistical techniques and produces beautiful graphs.
- R has a vibrant, thriving online community (see [stack overflow](#)).

Why are we using R?

1. R is **free**.
2. **R is popular** among economists, political scientists, psychologists, sociologists, geographers, anthropologists, biologists, data scientists, and statisticians.
3. **Employers prefer R** over most competing software environments.
4. R can **adapt to nearly any task**: 'metrics, spatial data analysis, machine learning, web scraping, data cleaning, website building, teaching.

Comparing statistical languages

Number of job postings on Indeed.com, 2019/07/17



R + [Examples]

R + Regression

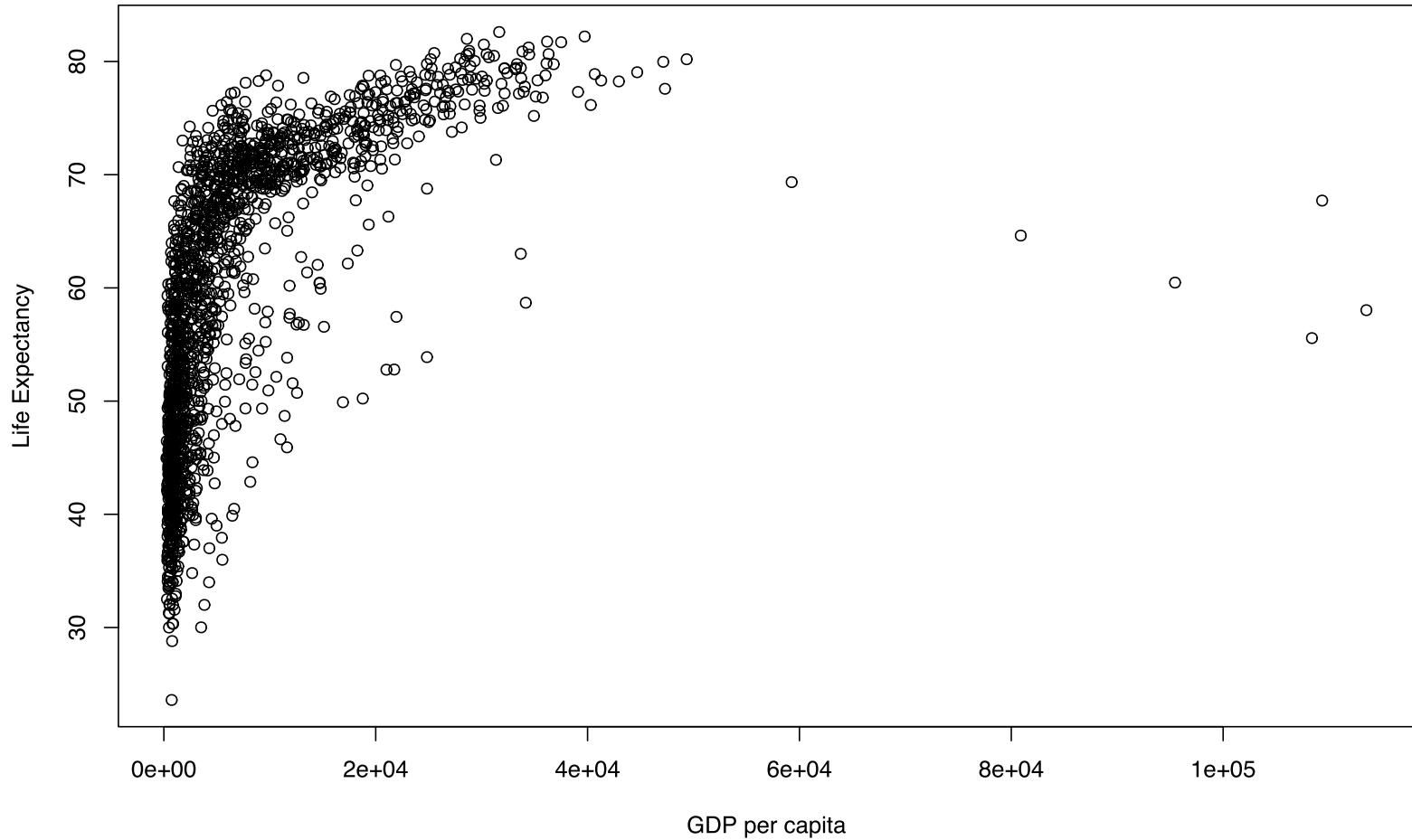
```
# A simple regression
fit <- lm(mpg ~ 1 + wt, data = mtcars)
# Show the coefficients
coef(summary(fit))
```

```
#>              Estimate Std. Error    t value    Pr(>|t| )
#> (Intercept) 37.285126   1.877627 19.857575 8.241799e-19
#> wt           -5.344472   0.559101 -9.559044 1.293959e-10
```

```
# A nice, clear table
library(broom)
tidy(fit)
```

```
#> # A tibble: 2 × 5
#>   term      estimate std.error statistic p.value
#>   <chr>      <dbl>     <dbl>      <dbl>     <dbl>
#> 1 (Intercept) 37.3       1.88      19.9  8.24e-19
#> 2 wt          -5.34      0.559     -9.56  1.29e-10
```

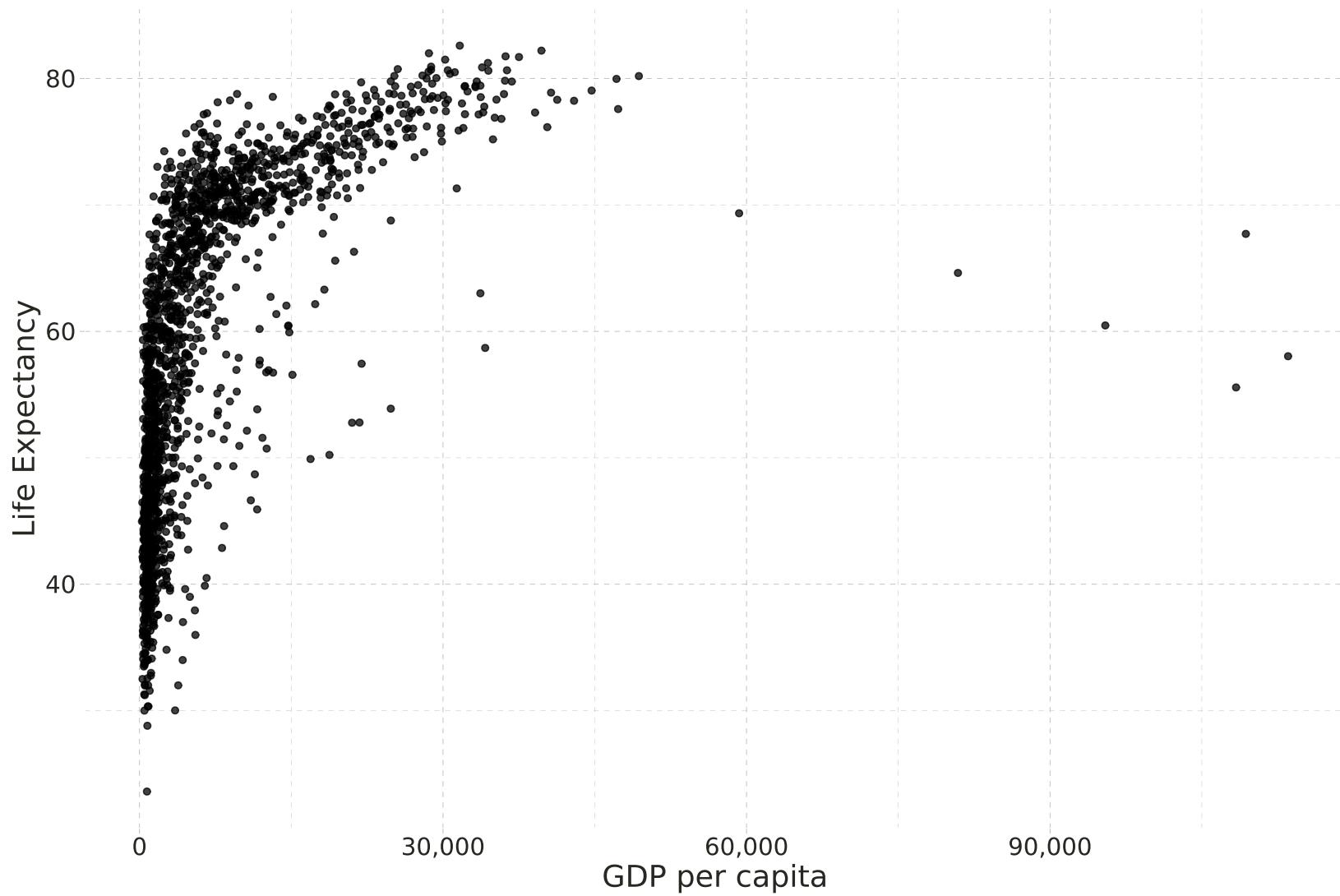
R + Plotting (w/ plot)



R + Plotting (w/ plot)

```
# Load packages with dataset
p_load(gapminder)
# Create dataset
plot(
  x = gapminder$gdpPercap, y = gapminder$lifeExp,
  xlab = "GDP per capita", ylab = "Life Expectancy"
)
```

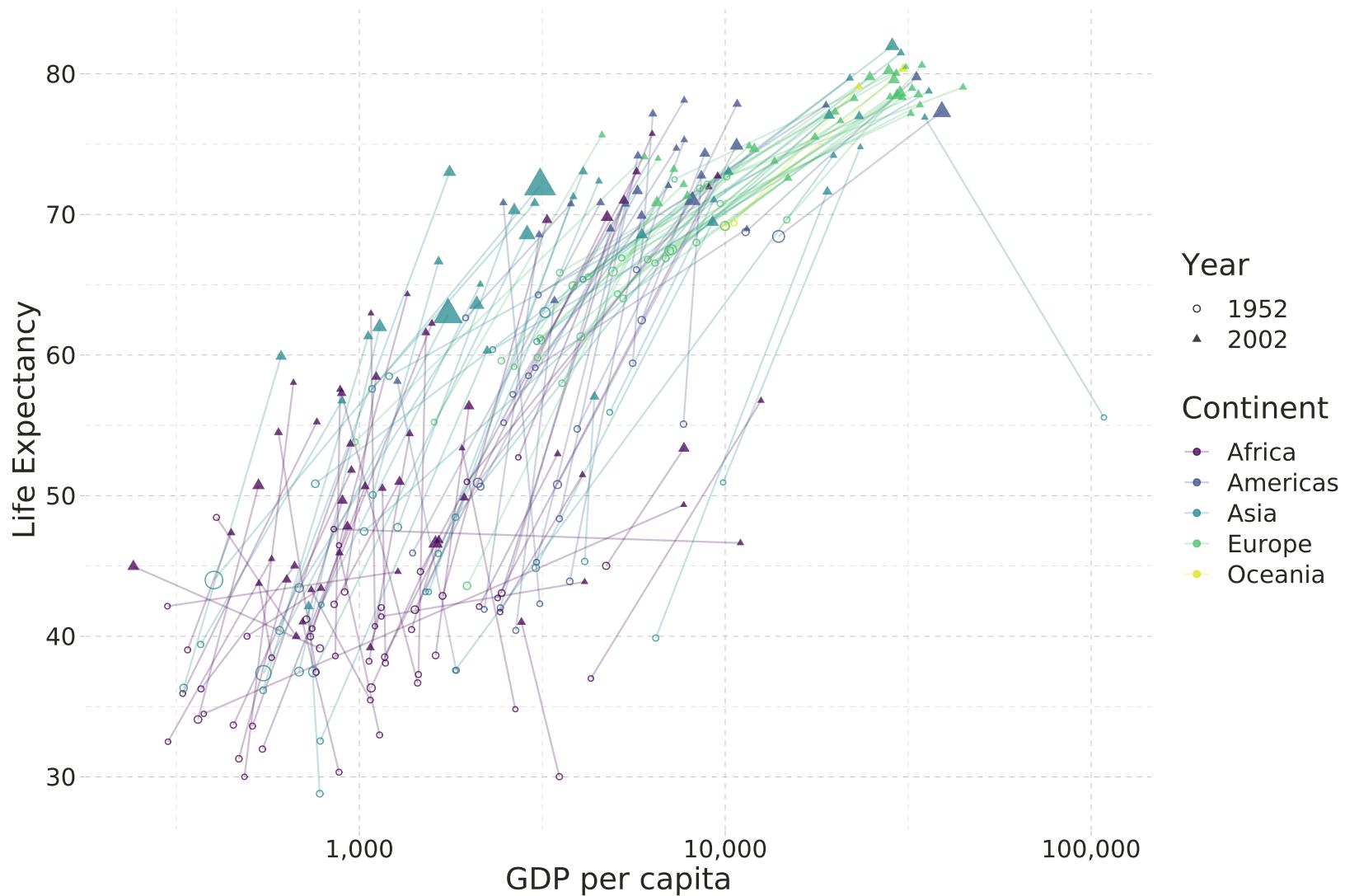
R + Plotting (w/ ggplot2)



R + Plotting (w/ ggplot2)

```
# Load packages
p_load(gapminder, dplyr)
# Create dataset
ggplot(data = gapminder, aes(x = gdpPercap, y = lifeExp)) +
  geom_point(alpha = 0.75) +
  scale_x_continuous("GDP per capita", label = scales::comma) +
  ylab("Life Expectancy") +
  theme_pander(base_size = 17, base_family = "Fira Sans", fc = met_slate)
```

R + More plotting (w/ ggplot2)

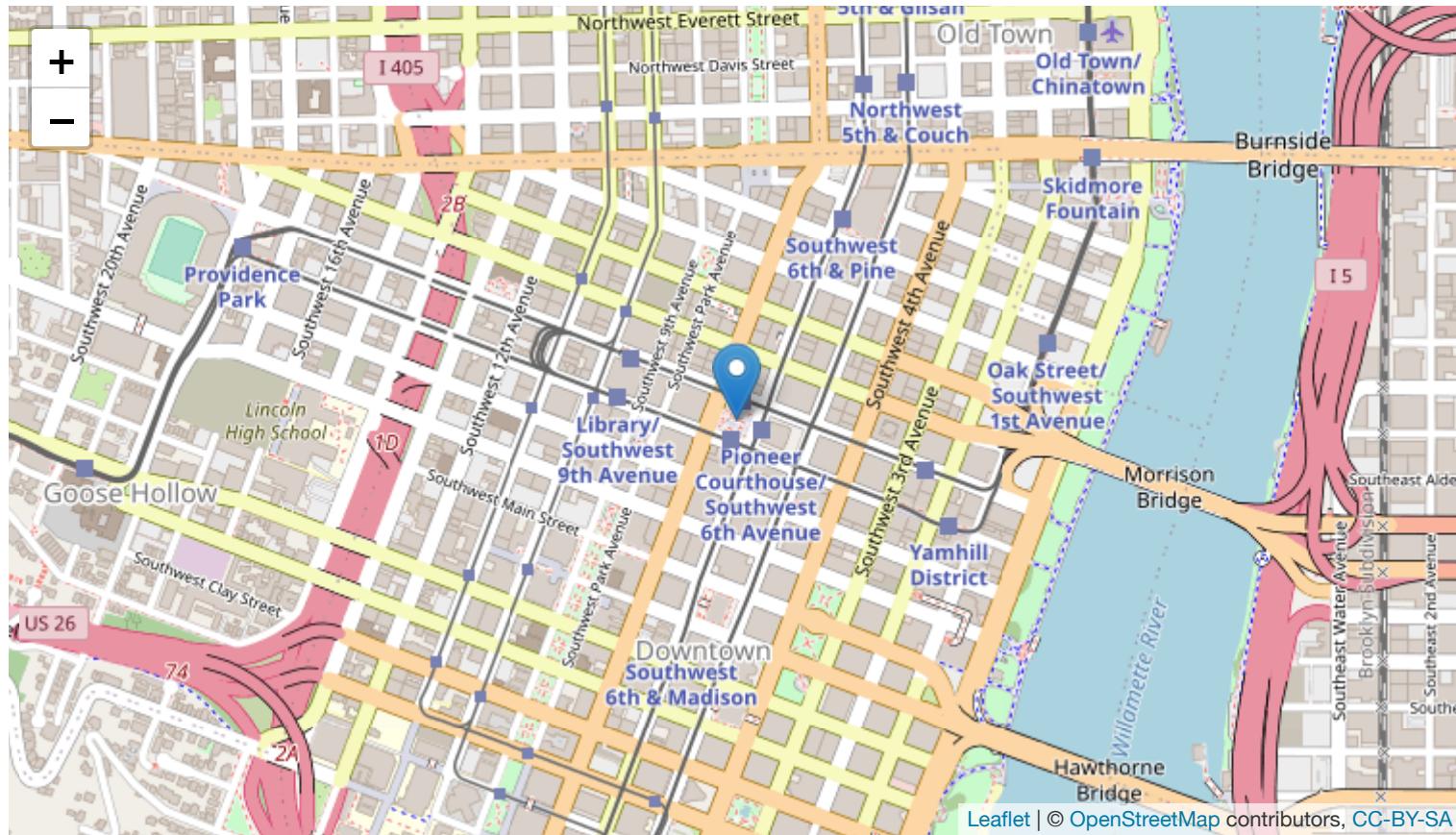


R + More plotting (w/ ggplot2)

```
# Load packages
p_load(gapminder, dplyr)
# Create dataset
ggplot(
  data = filter(gapminder, year %in% c(1952, 2002)),
  aes(x = gdpPercap, y = lifeExp, color = continent, group = country)
) +
  geom_path(alpha = 0.25) +
  geom_point(aes(shape = as.character(year), size = pop), alpha = 0.75) +
  scale_x_log10("GDP per capita", label = scales::comma) +
  ylab("Life Expectancy") +
  scale_shape_manual("Year", values = c(1, 17)) +
  scale_color_viridis("Continent", discrete = T, end = 0.95) +
  guides(size = F) +
  theme_pander(base_size = 17, base_family = "Fira Sans", fc = met_slate)
```

R + Maps

```
p_load(leaflet)  
leaflet() %> addTiles() %> addMarkers(  
  lng = -122.6793, lat = 45.51887, popup = "Pioneer Courthouse Square")
```



R + Animated maps (w/ ganimate)

Getting Started with R

Starting R

Installation

- Install R.
- Install RStudio.
- **Note:** All academic workstations at the UO have R, but having a copy of R on your computer will prove useful for the econometrics sequence and 400-level elective courses.

Resources

- Google and StackOverflow
- Time
- Your classmates
- Your GE
- Me

Starting R

R basics

1. Everything is an **object**.

```
foo
```

2. Every object has a **name** and **value**.

```
foo <- 2
```

3. You use **functions** on these objects.

```
mean(foo)
```

4. Functions come in **libraries (packages)**.

```
library(dplyr)
```

5. R will try to **help** you.

```
?dplyr
```

6. R has its **quirks**.

```
NA; error; warning
```