What is Econometrics?

EC 320: Introduction to Econometrics

Philip Economides
Winter 2022

Prologue

Who am I?

Philip Economides

- Doctoral student in economics
- TA in 'metrics EC320 (UG), EC423/EC523 (UG/MSc) and EC607 (PhD)
- Former RA at ESRI/Trinity College Dublin, Ireland
- Focus: International Economics, Applied Econometrics

Where can you find me?

- Office: 520 PLC
- Office hours: Tuesday 15:00-16:00, Thursday 10:00-11:00, or by appointment
- Email: peconomi@uoregon.edu **(use EC 320 in the subject line)**

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.

Read through the prescribed readings listed in the Tentative Schedule of the syllabus **before each lecture**.

What will we do?

Coursework

- **W1 & W2:** Statistics Review + Fundamental Problem
- **W3:** Logic of Regressions
- **W4:** SLR - Estimation
- **W5:** SLR - Classical Assumptions
- **W6:** **Midterm**, SLR - Inference
- **W7:** MLR - Estimation & Inference
- **W8:** Non-linear & Qualitative variables
- **W9:** Interaction terms & Model Specifications
- **W10:** Heteroskedasticity & **Finals** Review

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

What will we do?

Grade

- 25% - Problem Sets (5pp each)
- 10% - Quizzes (5pp each)
- 20% - Midterm Exam
- 10% - Data Project
- 35% - Final Exam

What will we do?

TA: **Micaela Wood**

Venue: **442 MCK**, Thurs 16:00-17:20

Click here to see location

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 and data project.

Good Classroom Citizenship

As you all know very well by now:

- Wear your mask and make sure it fits you well
- Stay home if you're sick
- Get to know your neighbors in class, let them know if positive
- Watch for signs and symptoms with the daily symptom self check
- Wash your hands frequently or use hand sanitizer
- Complete the UO COVID-19 case and contact reporting form if you test positive or are a close contact of someone who tests positive.

UO Corona Corps: 541-346-2292; Careteam@uoregon.edu

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?

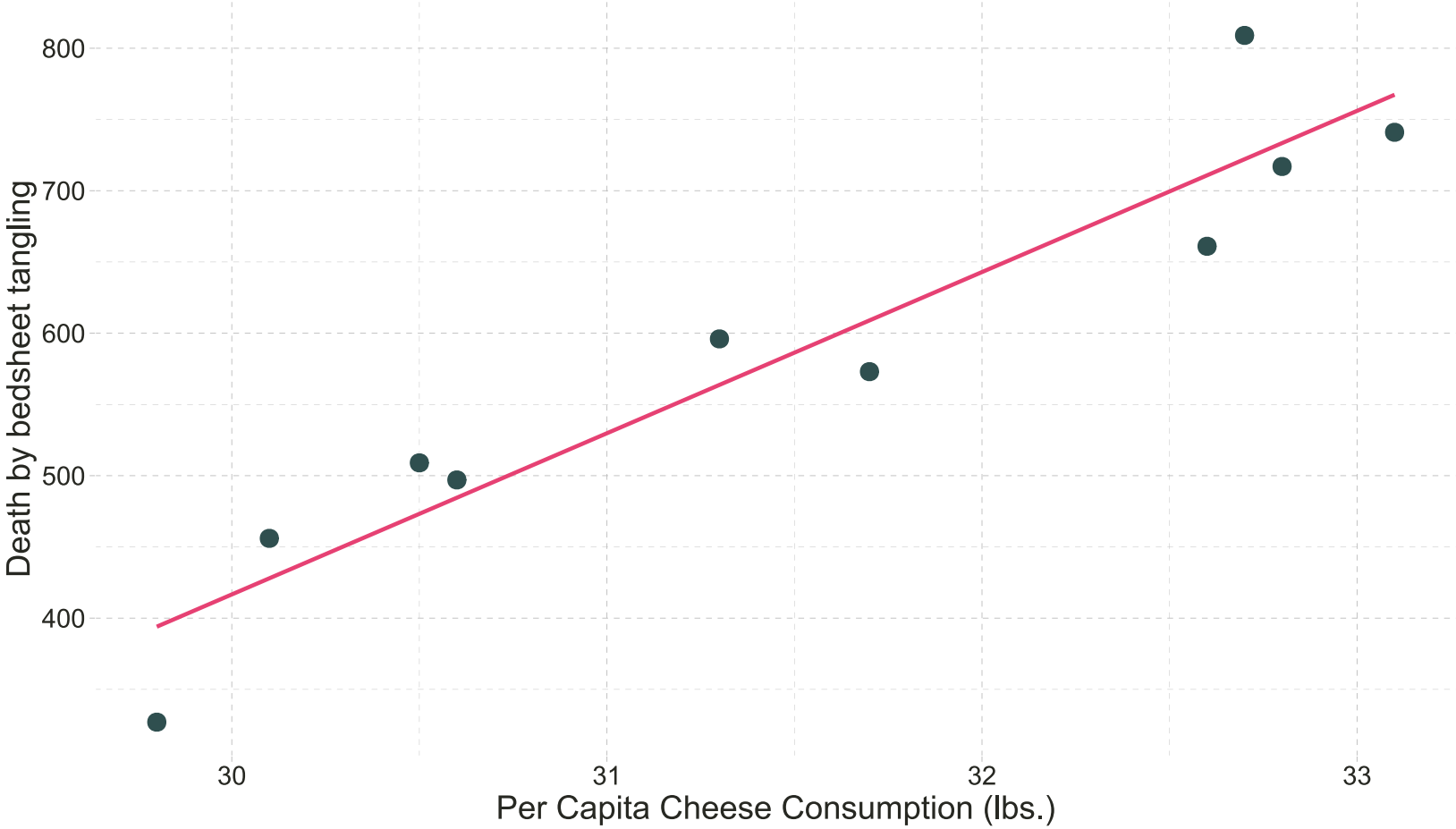
The main focus of EC 320 and EC 421 is causal identification.

- But...both rely on a common set of statistical techniques.

- For those interested, Professor Tim Duy teaches forecasting (EC 422) this same term.

Econometrics

Not all relationships are causal



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.

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.

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.

Example: *Blue Paradox*

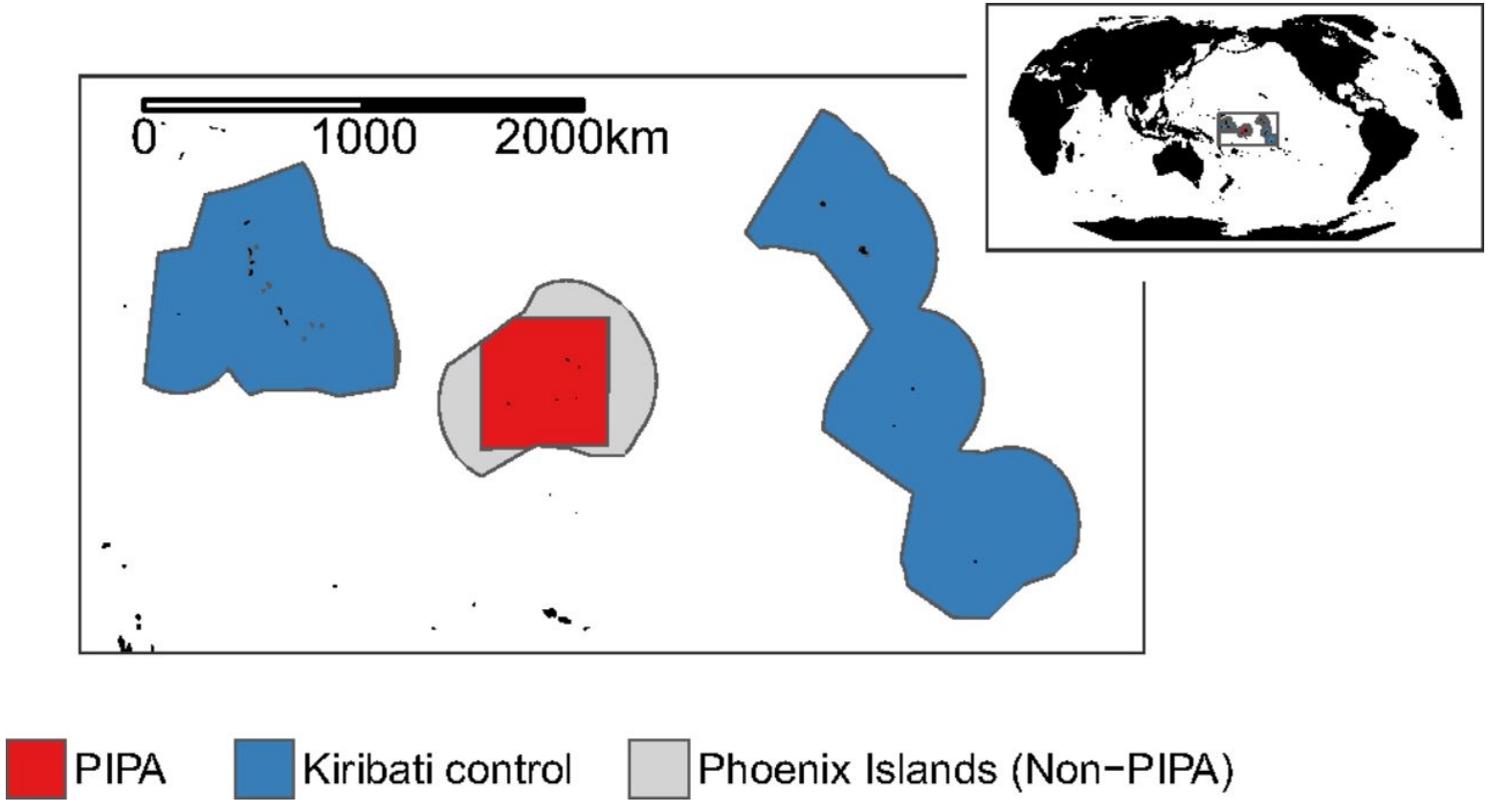
Natural Experiment

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.

Example: *Blue Paradox*

Results



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 $\neq$ 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?

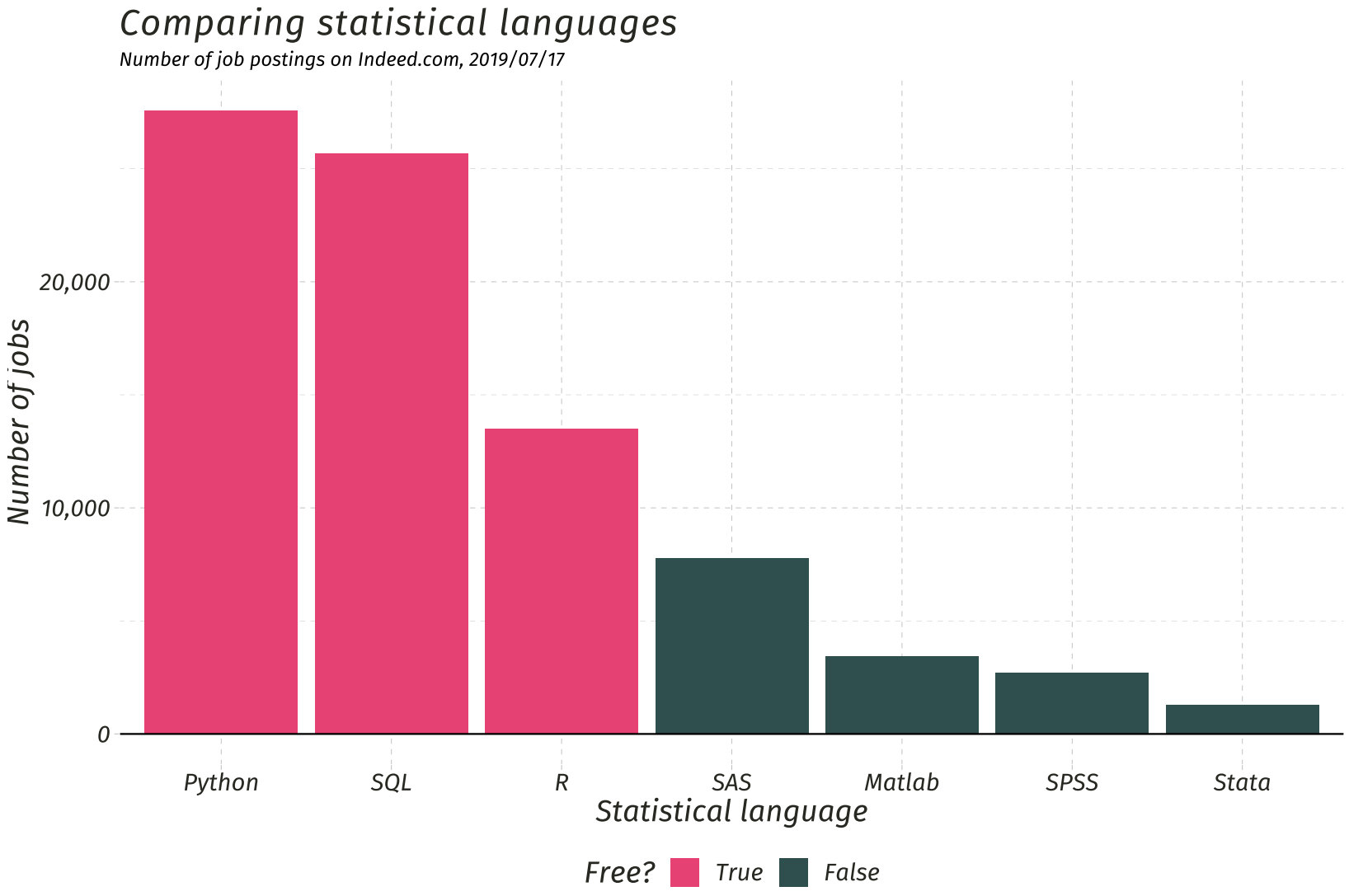
1. R is **free**.

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.

R


Comparing statistical languages
Number of job postings on Indeed.com, 2019/07/17
Number of jobs
20,000
10,000
0
Python
SQL
R
SAS
Matlab
SPSS
Stata
Statistical language
Free?
True
False

R + [Examples]

R + Regression

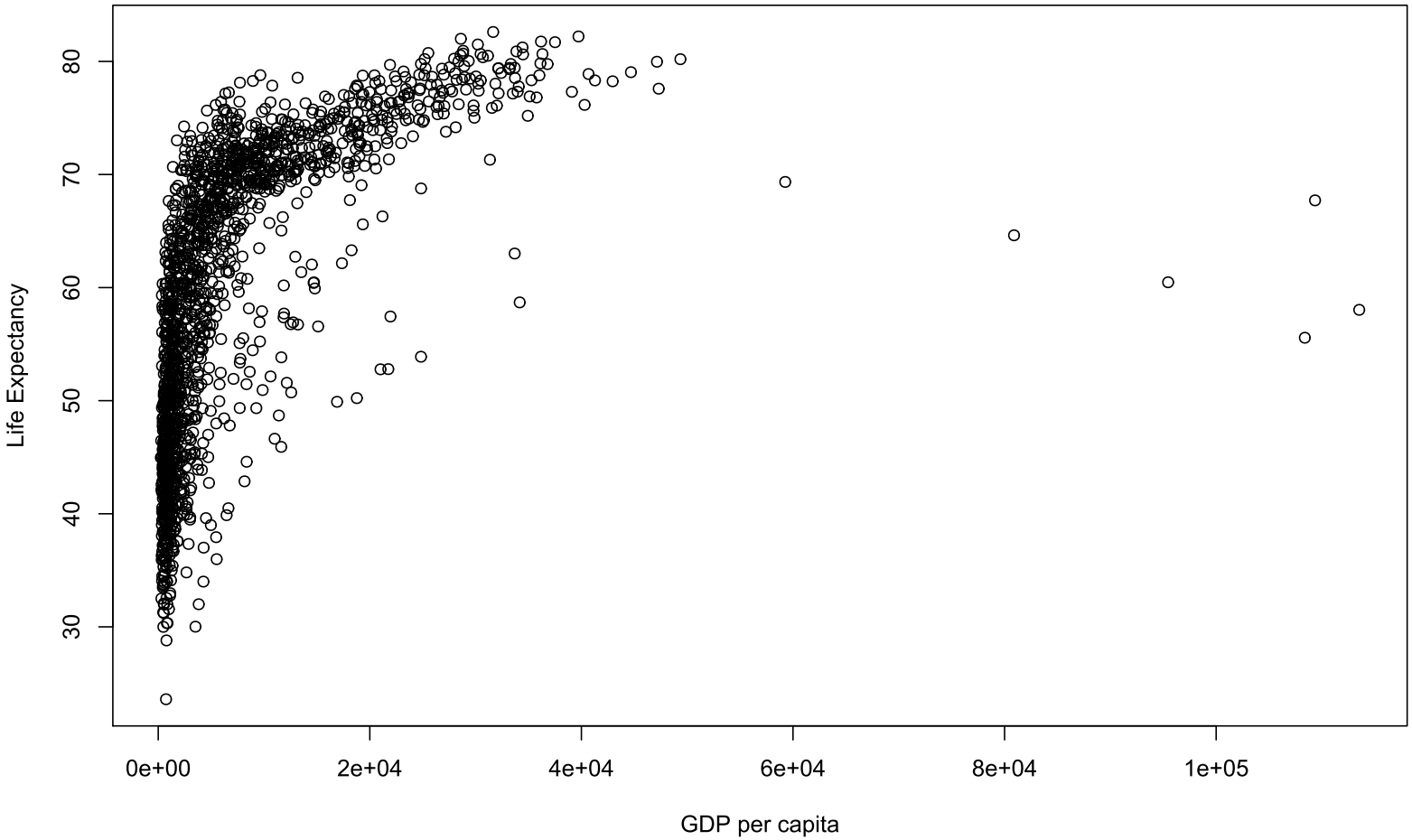
```r
# A simple regression
fit ← lm(mpg ~ 1 + wt, data = mtcars)
# Show the coefficients
coef(summary(fit))
```

```
#>              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
```

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

```
#> # A tibble: 2 x 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
```

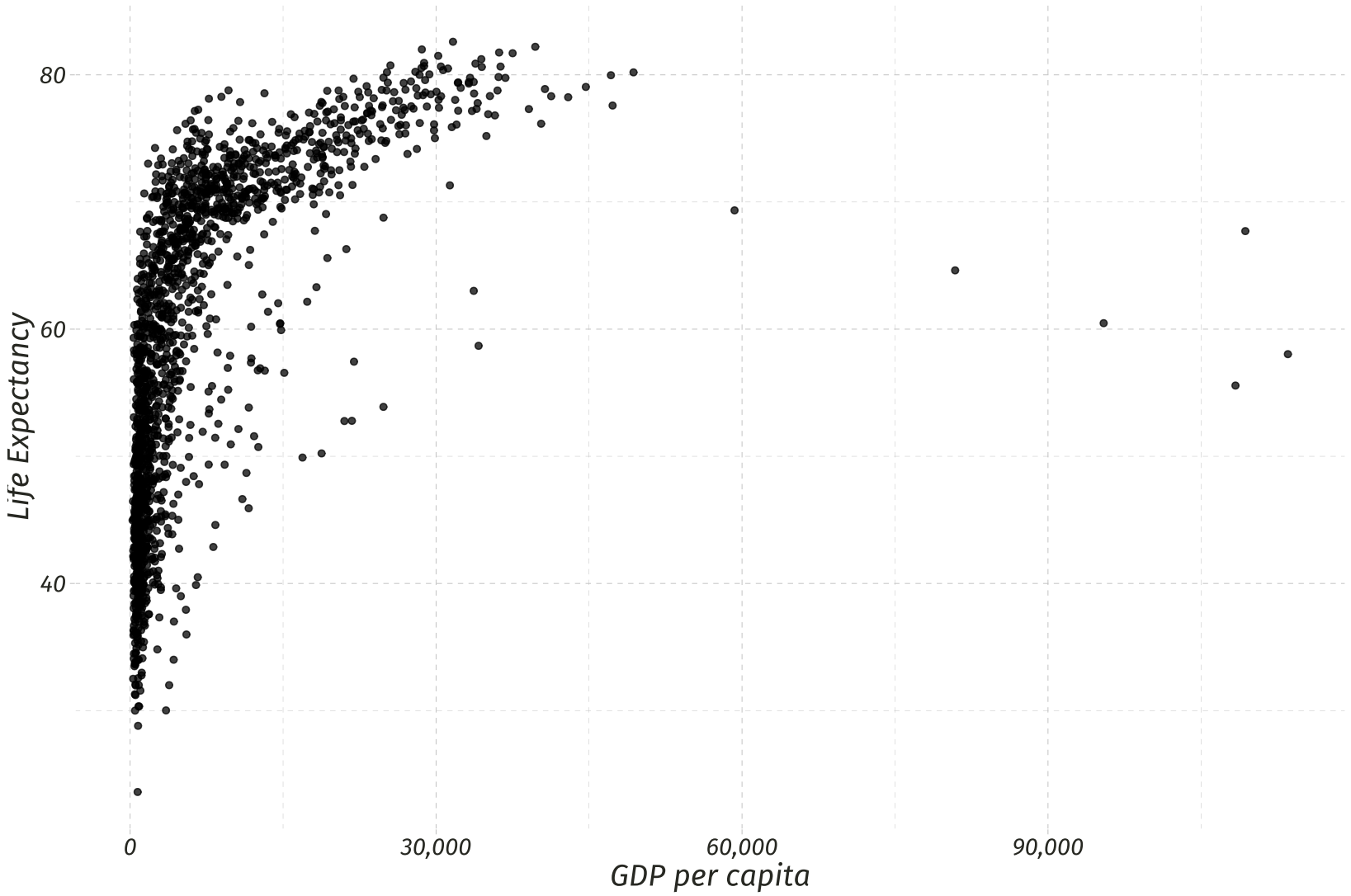
R + Plotting (w/ `plot`)



R + Plotting (w/ `plot`)

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

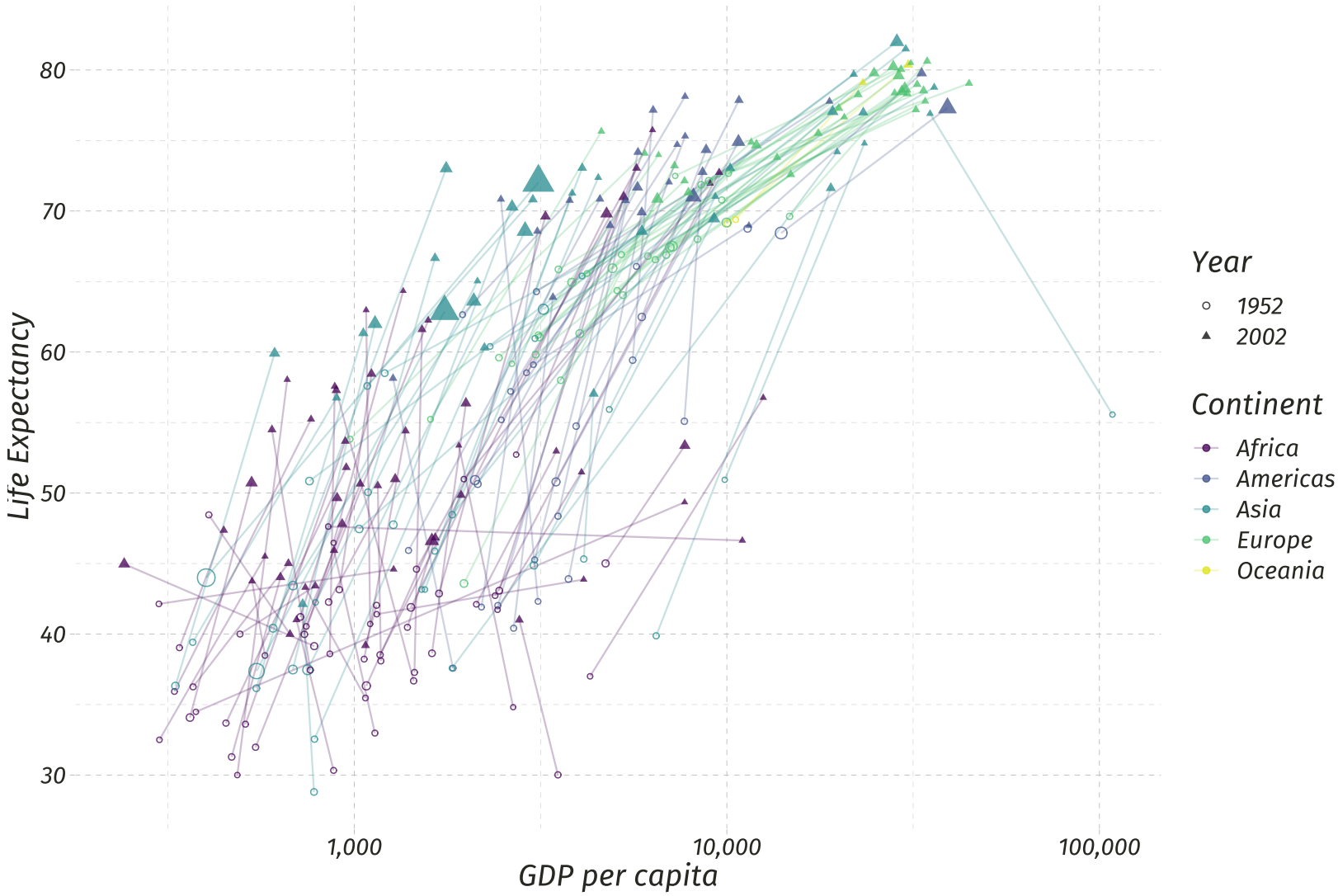
R + Plotting (w/ ggplot2)



R + Plotting (w/ ggplot2)

```r
# 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)

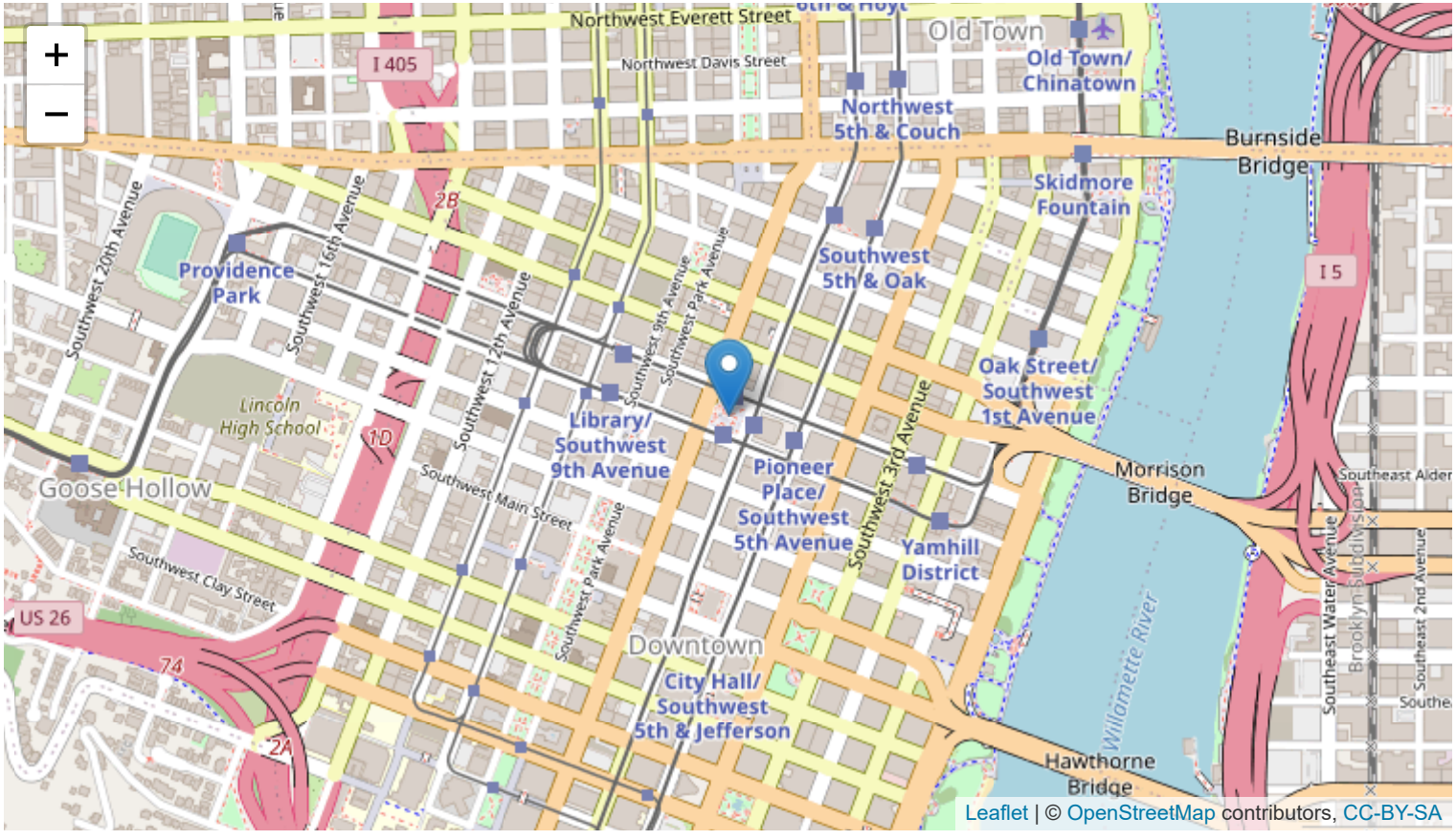


R + More plotting (w/ ggplot2)

```r
# 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/ gganimate)

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`