Lecture 1: Introduction to Labor Economics and R

Economía Laboral

Junghanss, Juan Cruz

Universidad del CEMA

2nd Semester 2023

Introduction

Labor Economics Exercise Sessions

- Martes 13-15h

Prof. Junghanss Juan Cruz

- Email: jcjunghans22@ucema.edu.ar , junghanss.jc@gmail.com.ar
- Telegram @jcjunghanss or WhatsApp
- LinkedIn

Feel free to contact me if you have questions regarding the course or even for anything related to the Uni student's life.

Introduction

Slides and Material available at Webcampus and own website junghanss.github.io

Bibliography of the course:

- Borjas, George J.. (2016). *Labor Economics* (Ed. 7th). New York: McGraw Hill.
- Grolemund, G., & Wickham, H. (2017). *R for Data Science*. O'Reilly Media.

Additional Bibliography (recommended):

- James, G., Witten, D., Hastie, T., Tibshirani, R., (2013). *An introduction to statistical learning : with applications in R*. New York. Springer.

Introduction

Calification:

Let $E_{TPi} \in [0, 10]$ be one of our TP's grade. The final grade E of our sessions is:

$$E = \frac{1}{4} \sum_{i=1}^{4} E_{TPi}$$

We'll have a total of 4 TPs. The exercise sessions' grade accounts for the 0.2 of the whole course.

Introduction

Lecture Plan 2S 2023:[1]

Day	Nr.	Topic
01-08	1	Introducción a Econ Laboral & R.
08-08	2	EPH & R para TP Nº1.
15-08	3	Data Wrangling en R.
22-08	4	Entrega TP Nº1. Intro a GitHub.
29-09	5	Data Wrangling: Dates and Time.
05-09	6	Matematica para TP Nº2.
12-09	7	Entrega TP Nº2. Labor Supply Review.
19-09	8	Labor Supply & Demand Review.

[1] It is subject to change without prior notice.

Introduction

Lecture Plan 2S 2023:[2]

Day	Nr.	Topic
26-09	9	**Primer Parcial.**
03-10	10	Matemática TP Nº3.
10-10	11	R para TP Nº3. Vectores y Factors en R.
17-10	12	Strings & Regex en R.
24-10	13	Entrega TP Nº3. Python & Stata.
31-10	14	Diferencias Igualadoras y Distribución del Ingreso.
07-11	15	Applied topic Nº1: Sindicatos
14-11	16	Entrega TP Nº4. Applied topic Nº2
21-11	17	Repaso para Segundo Parcial

[2] It is subject to change without prior notice.

Introduction

Objetivos del curso:

- Refinar la visión sobre el mercado laboral (INDEC, indicadores, etc.)
- Ganar una visión más práctica de la programación, data analysis y R.
- Aprender sobre el uso de herramientas útiles como GitHub.

Appendix: How to study for the course

For this course (and the rest of your subjects) I recommend you the following study techniques:

- **Exercises**: best way to gain understanding on what you're doing
- **Active Recall**: work effectively on your memory
- **Study groups**: discuss the topics with others

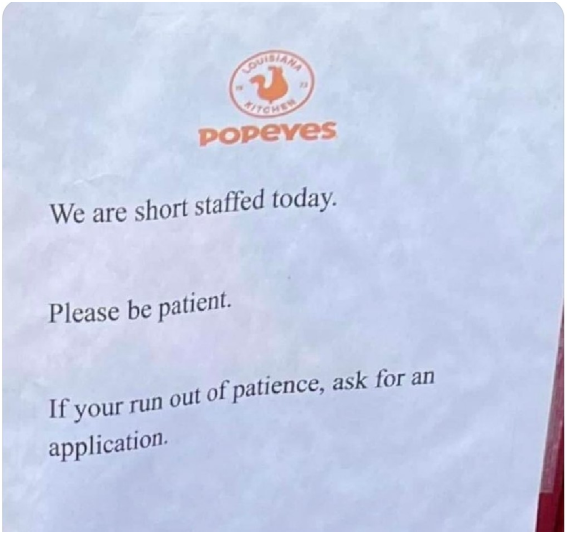
YouTube video on Evidence-based study methods

What is Labor Economics

How do labor markets work? Main questions:

- How is income distribution determined?
- Economic impact of unions
- Allocation of a worker's time to the market
- Hiring and firing decisions of firms
- Labor market discrimination
- Determinants of unemployment
- Worker's decision to invest in education (human capital)

What is Labor Economics

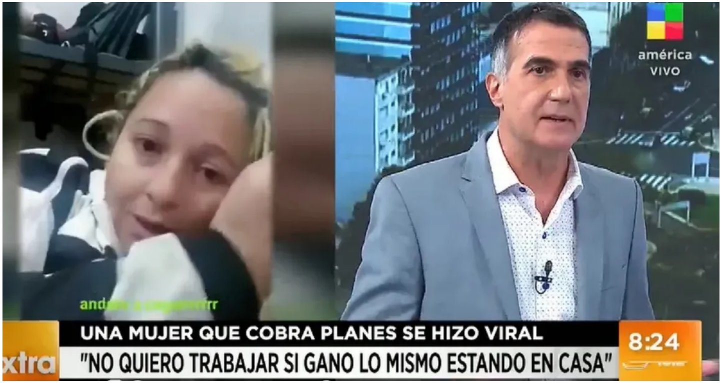


We are short staffed today.

Please be patient.

If your run out of patience, ask for an application.

What is Labor Economics


américa
vivo
UNA MUJER QUE COBRA PLANES SE HIZO VIRAL
"NO QUIERO TRABAJAR SI GANO LO MISMO ESTANDO EN CASA"
8:24
xtra

Basic structure of the Labor Market

Actors on the labor market:

- Labor supply: determined by individuals, i.e. workers
- Labor demand: determined by firms
- Government: responsible of markets regulations

How to derive the labor supply?

You can derive the labor supply by solving the individual's optimization problem. From now on you'll always have **leisure** ("ocio") in your objective function of your micro problems.

Main concepts derived from the problem's solution:

- **Working hours**: how many hours can I offer in the market
- **Salario de reserva**: for which minimum wage do I want to work?
- **Elasticity of labor supply**: percent change in amount of labor due to a percent change in wages

Labor Market Classifications

There are for sure concepts that you must know before working with labor market data. For example, how would you classify yourself in the labor market?

- Clasificaciones (ver archivo)
- Mercado de Trabajo - 2022 - INDEC
- Encuesta de Uso de Tiempo - 2021 - INDEC

Introduction to R - Installation

R can be downloaded from: http://cran.r-project.org/

It's recommended that you run R within an integrated development environment (IDE) such as "RStudio", which can be freely downloaded from: http://rstudio.com

Getting help and learning

1. Search for **help** related to the dataset, package, etc. directly in R with the command “?”. For example:

   ```
   ?mpg
   ```

2. Read the **documentation** of the package on the official website.
3. If you get stuck, start with **Google**. Typically, adding “R” to a query is enough to restrict it to relevant results: if the search isn’t useful, it often means that there aren’t any R-specific results available. Google is particularly useful for error messages. If you get an error message and you have no idea what it means, try googling it!
4. If Google doesn’t help, try **stackoverflow**. Start by spending a little time searching for an existing answer; including [R] restricts your search to questions and answers that use R.

Best practices

How to write good and clean code? It's important for you to learn it from the very beginning:

1. Documentate everything
2. Be explicitely precise and clear with your objects and functions names.
3. If possible, write always functions instead of repeting code more than twice.

Object names must start with a letter, and can only contain letters, numbers, _, and .. You want your object names to be descriptive, so you'll need a convention for multiple words.

Best practices - Fundamentals

(1) Documentate everything:

It's extremely important that you always comment (#) why are you doing something (and maybe also how). Doesn't matter if sometimes you write simple sentences or larger texts, the documentation helps you and others to review the code after several time and understand it quickly.

Professional code is always well documentated. An important use of comments is to break up your file into easy readable chunks. Use long lines of "-", for example:

Example

```
# Load data ----------------------
```

Best practices - Fundamentals

(2) Precision by naming objects, functions, etc:

The rules(conventions) are simple:

- Objects begin with lowercase. For example: “unemployment”
- Functions names should be verbs and evoke what they do. Better clear than short.
- You never ever use symbols like “-” to separate words, but underscore “_” or Uppercases. For example: “calculate_working_hours” or “calculateWorkingHours”

Best practices - Fundamentals

(3) Writing functions instead of repeating:

If you copy-paste code more than twice, then it's worth to write a function. Why?

- Functions make your code easier to understand.
- When requirements change, you only need to update the function code once instead of making several updates.
- You eliminate the potential chance of making mistakes by copying and pasting.

Exploratory Data Analysis

Exploratory Data Analysis (EDA) is an iterative cycle:

1. Generate questions about your data.
2. Search for answers by visualizing, transforming, and modeling your data.
3. Use what you learn to refine your questions and/or generate new questions.

“Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.”
-John Tukey

Exploratory Data Analysis

Example of questions to be made:

- What type of variation and covariation occurs within my variables?
- Which values are the most common? Why?
- Which values are rare? Why? Does that match my expectations?
- Can I see any unusual patterns? What might explain them?

Normative and Positive Economics

We'll focus on the positive side of economics since we want to describe how markets work. The main question is “What is it?”

Forget about normative economics in a strict way for this course, we don't care if the government intervention is good or bad for the market (although it's always good to think about it).

Data

Argentina:

- INDEC
- Buenos Aires Data

US Data:

- Bureau of Labor Statistics (BLS)
- Bureau of Census
- Statistical Abstract of the United States

International:

- OECD
- National Bureau of Economic Research (NBER)
- IZA
- World Bank

Data Science related sources

Extremely recommended for the rest of your lives:

- Kaggle: private datasets for data science projects

Introduction to R - Packages

You'll need to install some R packages. An R package is a collection of functions, data, and documentation that extends the capabilities of base R. The most important for us will be **Tidyverse**.

You can install each complete package with a single line of code:

Example

```
install.packages("tidyverse")
```

Introduction to R - Packages

You will not be able to use the functions, objects, and help files in a package until you load it with library(). Once you have installed a package, you can load it with the library() function:

Example

```
library(tidyverse)
```

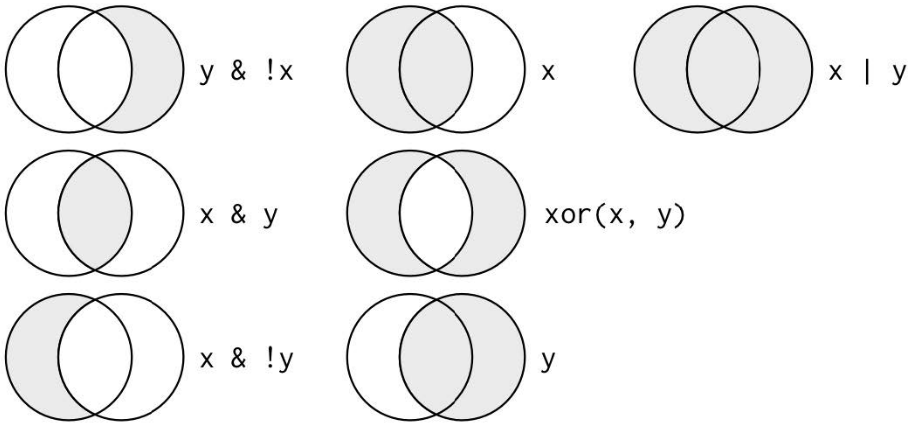
On this example, tidyverse will load the **ggplot2**, **tibble**, **tidyr**, **readr**, **purrr**, and **dplyr** packages. These are considered to be the core of the tidyverse because you'll use them in almost every analysis.

Basics - Variables

How variables are described

- **int** stands for integers.
- **dbl** stands for doubles, or real numbers.
- **chr** stands for character vectors, or strings.
- **dttm** stands for date-times (a date + a time).
- **lgl** stands for logical, vectors that contain only TRUE or FALSE.
- **fctr** stands for factors, which R uses to represent categorical variables with fixed possible values.
- **date** stands for dates.

Basics - Logical Operators



Basics - The pipe

The pipe operator, written as $\% > \%$, takes the output of one function and passes it into another function as an argument. Hence, this allows us to link a sequence of analysis steps.

You can use the pipe to rewrite multiple operations in a way that you can read left-to-right, top-to-bottom. The pipe is very useful, however, you shouldn't use it in some cases. We'll explore when.

The pipe comes from the package **magrittr**, but the tidyverse environment load it for you automatically.

Basics - The pipe

When not to use the pipe?

- You have multiple inputs or outputs, i.e. two or more objects are being transformed.
- Your pipes are longer than approx. 10 steps. It's better to create some significant intermediate objects because it will make debugging easier.
- You want to express complex relationships, i.e. directed graphs with dependency structures.

Review Questions

1. What is labor economics? Which types of questions do labor economists analyze?
2. Who are the key actors in the labor market?
3. Why do we need a theory to understand real-world labor market problems?
4. What is the difference between positive and normative economics? Why are positive questions easier to answer than normative questions?
5. What is exploratory data analysis?
6. What are the most important packages of the Tidyverse?