

Week 2

Introduction to Data Science with R: Exploring Democracy and Autocracy Worldwide

Dr. Ozlem Tuncel

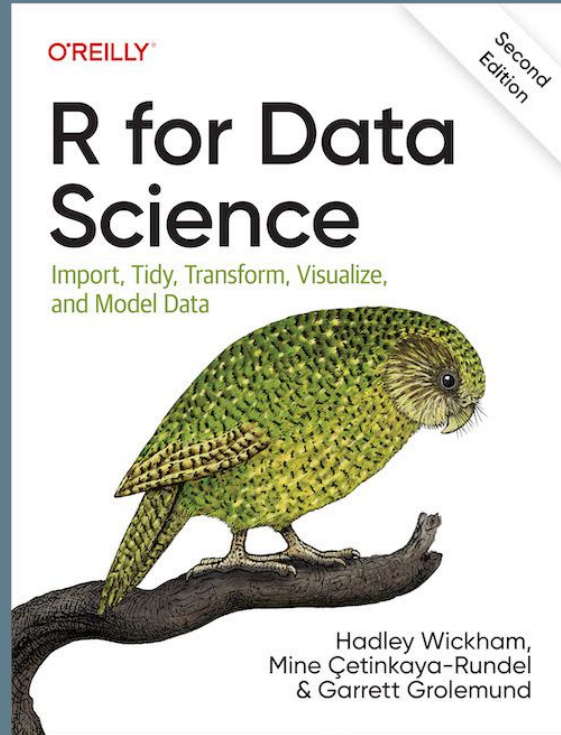
Georgia State University

otuncelgurlek1@gsu.edu



**3 small
announcements
before we start**

Readings?



Short polls?

- Remember slido exercise from last week
- *Would you be interested in answered democracy/autocracy related questions?*
- **Completely anonymous**
- Should not take more than 2-3 mins
- Not every week -- maybe 3-4 times a semester?
- Helps me gauge your knowledge/interest



Also, interested in doing research?

I am looking for Research Assistants (RA) to help me with my research projects on:

- Interested in democracy and autocracy research in the world
- *Current project:* Presidential impeachment in the world (specifically someone who can speak **Spanish**)
- It is voluntary, so not any form of payment
- But, I can write reference letters for jobs and other opportunities
- Contact Dr. Oz -- otuncelgurlek1@gsu.edu



**What is “data
analysis”?**

Let's talk about “Data Analysis”

DATA + ANALYSIS

Data analysis is the systematic process of inspecting, cleaning, transforming, and modeling raw data to discover useful information, draw conclusions, and support decision-making.

01

Data Collection

02

Cleaning &
Organization
OUR FOCUS

03

Data Analysis
OUR FOCUS

04

Interpretation

Step 1:
Data collection:
aka. How to find
“data”?

What is a data?

Data is a collection of raw, unprocessed facts, figures, numbers, symbols, or observations that describe the world.

Qualitative

Descriptive, non-numerical observations focused on characteristics.

"Football is my favourite game"

"My dog has white hair"

Quantitative

Numerical data that can be measured or counted.

Discrete

Countable, distinct values

5

Continuous

Infinite, measurable range

3.265...

But, where do we find data?

1. Collect It

Data does not exist — so you must gather it yourself.

- **Custom Creation:** Interviews, observation, archives, web scraping, etc.
- **Dr. Oz's Datasets:**
ozlemtuncel.github.io/datasets/
- **Autocratic Ruling Parties:**
[jstor.org/stable/48596916](https://www.jstor.org/stable/48596916)
- **Collective Efforts:**
 - <https://subjectguides.library.american.edu/c.php?q=1238130&p=9060337>
 -

2. Retrieve It

Data already exists — someone did the work for you!

- **Non-Governmental:** V-Dem, World Bank, etc.
 - *Organized Crime Index:*
<https://ocindex.net/2021/country/italy>
- **Governmental:** CDC ([example](#)), national statistics agencies (such as [NASA](#), [C-SPAN](#), [US HoR](#), [Elections](#)).
- Some authoritarian examples:
- **Arolsen Archives:**
arolsen-archives.org/en/
- **Stasi Records Agency:**
bundesarchiv.de/en/...

What is the V-Dem Dataset?



Varieties of Democracy (V-Dem)

One of the largest and most trusted social science data projects in the world.



The Strategy

Instead of just calling a country "free" or "not free," V-Dem uses expert surveys to score hundreds of specific indicators (e.g., freedom of the press, academic freedom, court independence).



The Tools

We will use R to map these scores and watch history unfold through data.



**For data analysis
course, we will
assume that you
already “collected”
and have your
data ready!**

Let's talk about “analysis” part of data analysis

OUR FOCUS

Descriptive

What happened?

Summarizes historical data to show current or past trends.

EXAMPLE

Tracking monthly sales spikes.

Causal

Why did it happen?

Finds the root causes, anomalies, or correlations behind the data.

EXAMPLE

Why sales spiked during specific holidays.

Predictive

What is likely?

Uses past data and trends to forecast future outcomes.

EXAMPLE

Estimating next year's holiday sales volume.

Prescriptive

What to do next?

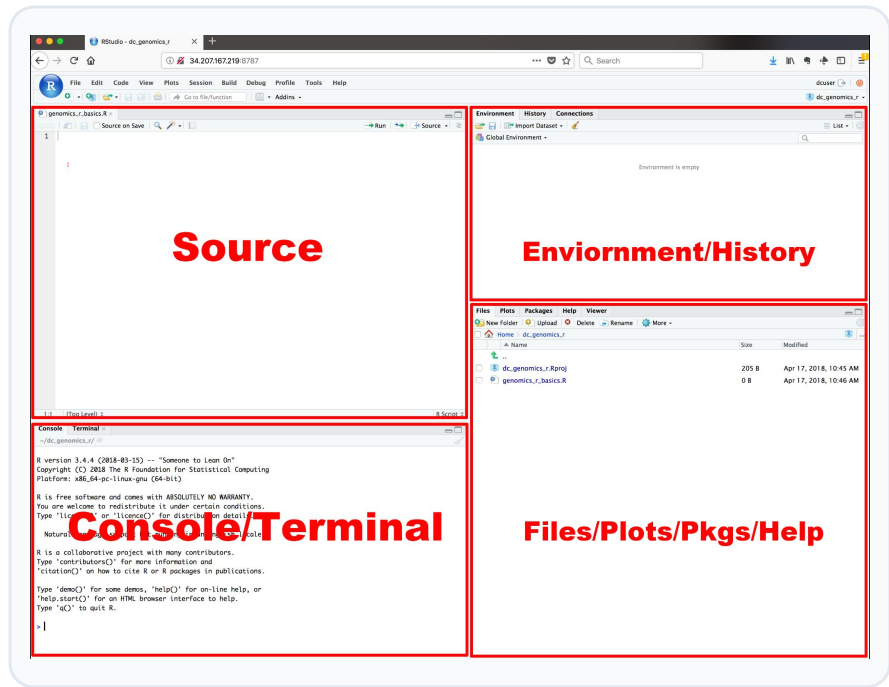
Recommends specific actions or decisions based on data models.

EXAMPLE

Scheduling targeted ads to maximize revenue.

Some R basics

Navigating the RStudio Interface



1. Source / Editor

Your script file (where you write recipes to save for later).

2. Console

The engine (where code executes immediately and talks back to you).

3. Environment

Your digital workbench (tracks variables and datasets currently loaded in memory).

4. Files / Plots / Help

Your multi-tool viewer (shows files, generated charts, and package help documentation).

Everything is an Object, Anything Happens via Functions

Objects

The Storage Box

Think of an object as a named storage box. We put data (numbers, names, entire data frames) inside it so we can reuse it later without typing it out.

Assignment

The Operator (<-)

The arrow puts things inside the box.

Example

```
gsu_color <- "blue"
```

Functions

The Machines ()

The machines that *do* things to your objects. They always end in parentheses.

Examples

```
round(3.1415)
```

```
install.packages()
```

**Let's switch to R
for practice!**

Next week to-do

Labor Day! -- no class :(

But, great time to finish R badges in one go!

**R Ready! Badge Due by Week 4 Monday
11:59 PM.**

