

Week 4

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

Dr. Ozlem Tuncel

Georgia State University

otuncelgurlek1@gsu.edu



Today's agenda:

- 1. Finish week 2 R code**
- 2. Short intro to tidyverse**
- 3. Finish week 4 R code**

What is the Tidyverse?

An Ecosystem of Packages

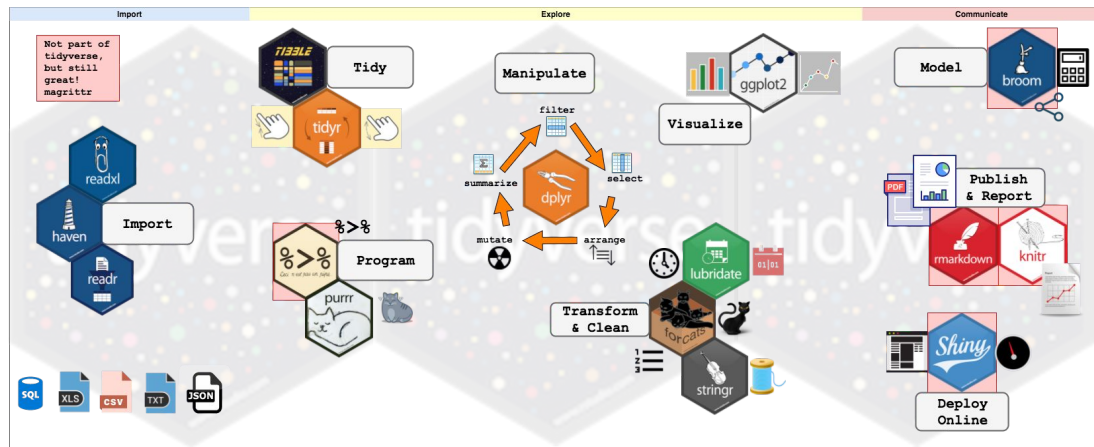
Instead of a single tool, the **tidyverse** is a collection of interconnected R packages designed specifically for data science.

Shared Philosophy

Every package shares a common grammar and data structure.

Core Packages We Will Use:

- **readr**: For importing data into R (e.g., CSV files).
- **dplyr**: For manipulating, filtering, and sorting data frames.
- **ggplot2**: For building beautiful data visualizations.



Anatomy of a Spreadsheet in R

Working with Tidy Data

In data science, we work with tidy, rectangular data frames where structure defines meaning.

Rows are Observations

A row is usually a specific unit at a specific time (e.g., *Turkey in 2025* or *Hungary in 2010*).

Columns are Variables

Characteristics or measurements for those observations (e.g., **country**, **year**, **libdem_score**).

VISUAL REPRESENTATION (DATA FRAME)

The diagram illustrates a data frame structure. A blue label 'Columns' with arrows points to the column headers: Name, Team, Number, Position, and Age. An orange label 'Rows' with arrows points to the row indices: 0, 1, 2, 3, 4, 5, and 6. A pink label 'Data' with a bracket points to the data cells, specifically highlighting the 'Number' column for rows 2 through 5.

	Name	Team	Number	Position	Age
0	Avery Bradley	Boston Celtics	0.0	PG	25.0
1	John Holland	Boston Celtics	30.0	SG	27.0
2	Jonas Jerebko	Boston Celtics	8.0	PF	29.0
3	Jordan Mickey	Boston Celtics	NaN	PF	21.0
4	Terry Rozier	Boston Celtics	12.0	PG	22.0
5	Jared Sullinger	Boston Celtics	7.0	C	NaN
6	Evan Turner	Boston Celtics	11.0	SG	27.0

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

If we have time: Tracking the Third Wave of Autocratization

01. The Third Wave

According to political scientists Lührmann and Lindberg (2019), we are living through a "Third Wave of Autocratization".

02. Empirical Study

To study this empirically, we cannot just look at individual numbers in isolation; we need a historical panel dataset tracking countries over time.

03. Practical Application

Today, we will learn how to read our curated V-Dem file directly into R so we can explore these trends across multiple columns and rows simultaneously.

Visualizing the Three Waves

