# ISA 401: Business Intelligence & Data Visualization

## 18: Charts Used Time-Series, Spatial and Spatiotemporal Data

Fadel M. Megahed, PhD

Associate Professor
Department of Information Systems and Analytics
Farmer School of Business
Miami University

Twitter: FadelMegahed
GitHub: fmegahed
Email: fmegahed@miamioh.edu
Office Hours: Automated Scheduler for Virtual Office Hours

Spring 2022

# Learning Objectives for Today's Class

- Understand main goals behind visualizing time-series data
- Explain the different types of charts for univariate and multivariate time-series
- Explain the different types of spatial plots
- Select suitable spatial graphs for different scenarios
- Understand how spatiotemporal plots can help in storytelling (what makes BI special)

# Types of Data Over Time

# Cross Sectional Data

**Cross Sectional Data:** Measurements on multiple units, recorded in a single time period.

**Example 1:** H1B 2020-2022 Data for Senior Data Scientists at Netflix

| | START DATE | JOB TITLE | BASE SALARY | LOCATION |
|---|---|---|---|---|
| 1 | 2021-08-11 | SENIOR DATA SCIENTIST | 118,955 | LOS GATOS, CA |
| 2 | 2021-06-14 | SENIOR DATA SCIENTIST | 143,291 | LOS GATOS, CA |
| 3 | 2021-06-14 | SENIOR DATA SCIENTIST | 143,291 | LOS GATOS, CA |
| 4 | 2021-09-09 | SENIOR DATA SCIENTIST | 143,291 | LOS GATOS, CA |
| 5 | 2021-10-18 | SENIOR DATA SCIENTIST | 143,562 | LOS GATOS, CA |

Showing 1 to 5 of 18 entries Previous 1 2 3 4 Next

Data scraped from https://h1bdata.info/index.php?em=Netflix+Inc&job=Senior+Data+Scientist&city=LOS+GATOS&year=All+Years on March 30, 2022 using the rvest 🤗. The printing was limited to those individuals who started on/after January 01, 2020, with the filters specified in the URL.

# Cross Sectional Data

**Cross Sectional Data:** Measurements on multiple units, recorded in a single time period.

**Example 2:** NBA 2021-2022 Leaders - Top Players in PTS/Game

| | Player | Pos | Age | Tm | G | FG | FG% | eFG% | PTS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | LeBron James | SF | 37 | LAL | 55 | 11.4 | .523 | .589 | 30.1 |
| 2 | Joel Embiid | C | 27 | PHI | 61 | 9.5 | .489 | .523 | 29.9 |
| 3 | Giannis Antetokounmpo | PF | 27 | MIL | 61 | 10.2 | .550 | .579 | 29.7 |
| 4 | Kevin Durant | PF | 33 | BRK | 48 | 10.5 | .519 | .567 | 29.5 |
| 5 | Trae Young | PG | 23 | ATL | 69 | 9.3 | .458 | .532 | 28 |

Showing 1 to 5 of 791 entries

Previous 1 2 3 4 5 ... 159 Next

**Source:** Data scraped from https://www.basketball-reference.com/leagues/NBA_2022_per_game.html}{Basketball-Reference on March 30, 2022 using the rvest 🤗. The printing was limited to the selected variables.

# Time Series Data

**Time Series Data:** Comparable measurements recorded on a single (or a few) variables over time (usually a long period of time).

**Example 2:** Stock prices of U.S. Airlines

# Panel Data

**Panel Data:** Cross sectional measurements (usually many variables) repeated over time (usually over a few time periods).

**Example:** World Bank's Data

| | iso3c | date | NY.GDP.MKTP.KD.ZG | SH.DYN.NMRT | SH.HIV.INCD.ZS | SH.MED.BE |
|---|---|---|---|---|---|---|
| 1 | CHN | 2018 | 6.7 | 4.1 | | |
| 2 | CHN | 2019 | 5.9 | 3.7 | | |
| 3 | CHN | 2020 | 2.3 | 3.5 | | |
| 4 | EGY | 2018 | 5.3 | 11.1 | 0 | |
| 5 | EGY | 2019 | 5.6 | 10.7 | 0 | |

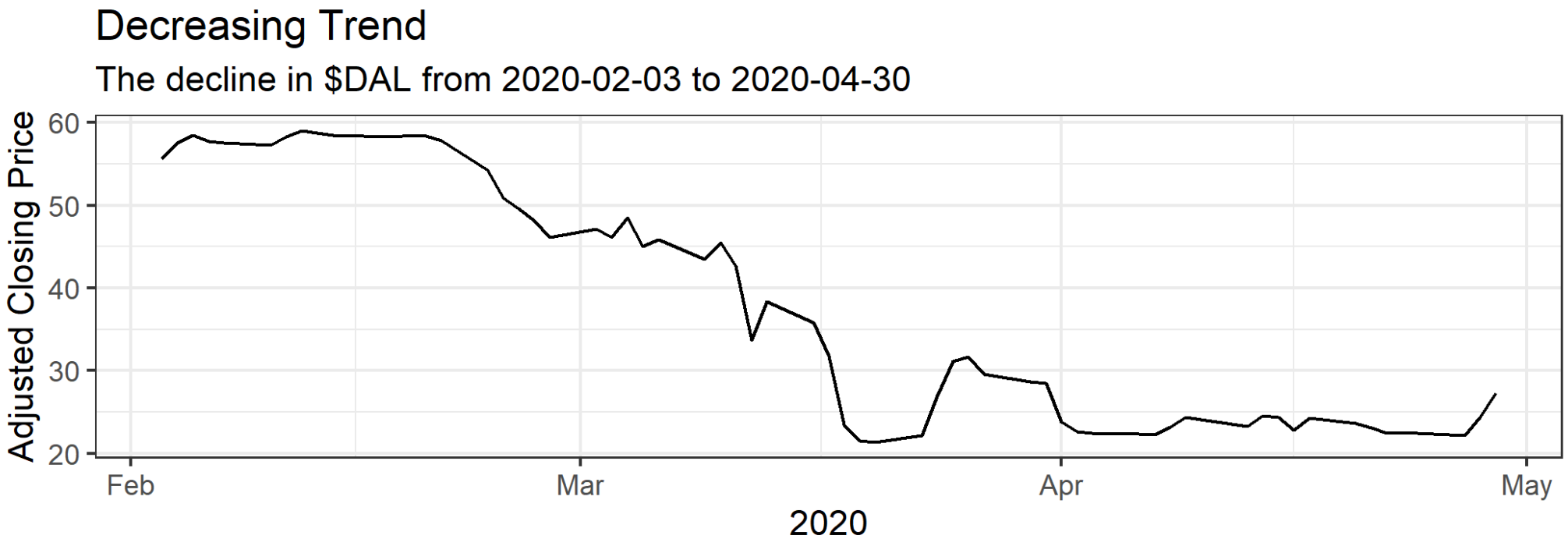
Showing 1 to 5 of 9 entries Previous 1 2 Next

**Source:** Data queried from the World Bank Data using the wbstats in R. The printed results show a snapshot of 7 variables (out of a much larger panel dataset). You can think of panel data as a cross-sectional dataset with a longitudinal/time component.

# Components of a Time Series

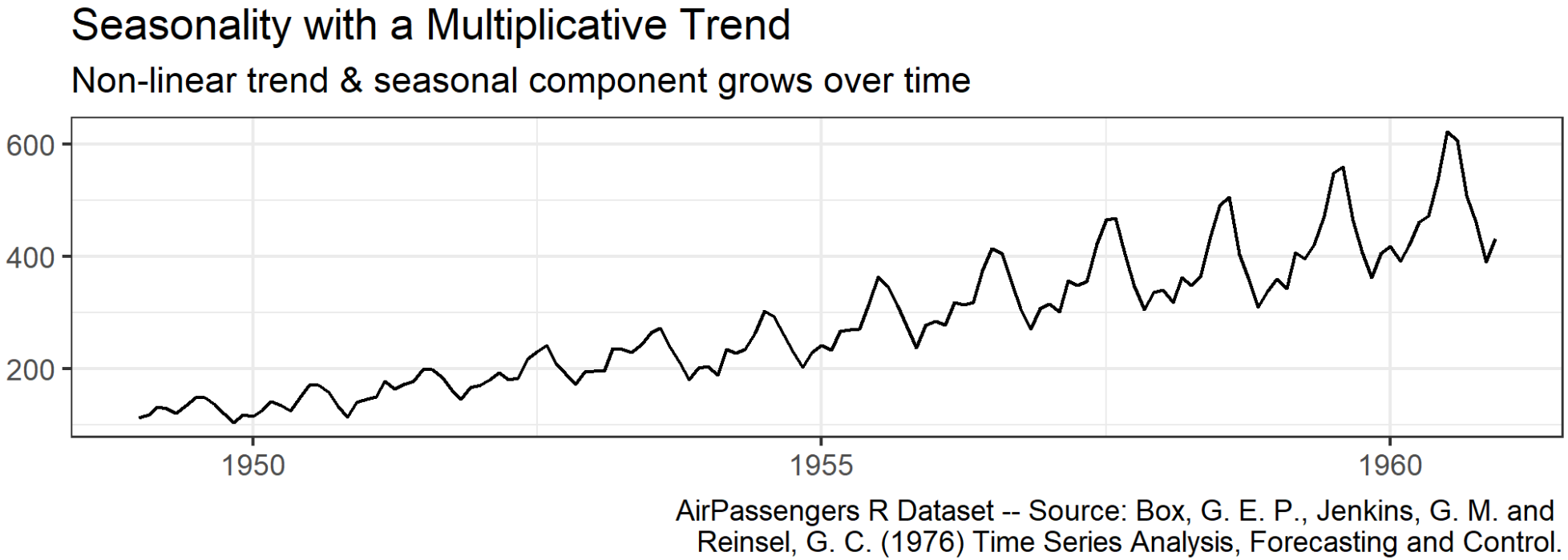
# Trend

A **trend** is an increasing or decreasing pattern over time.



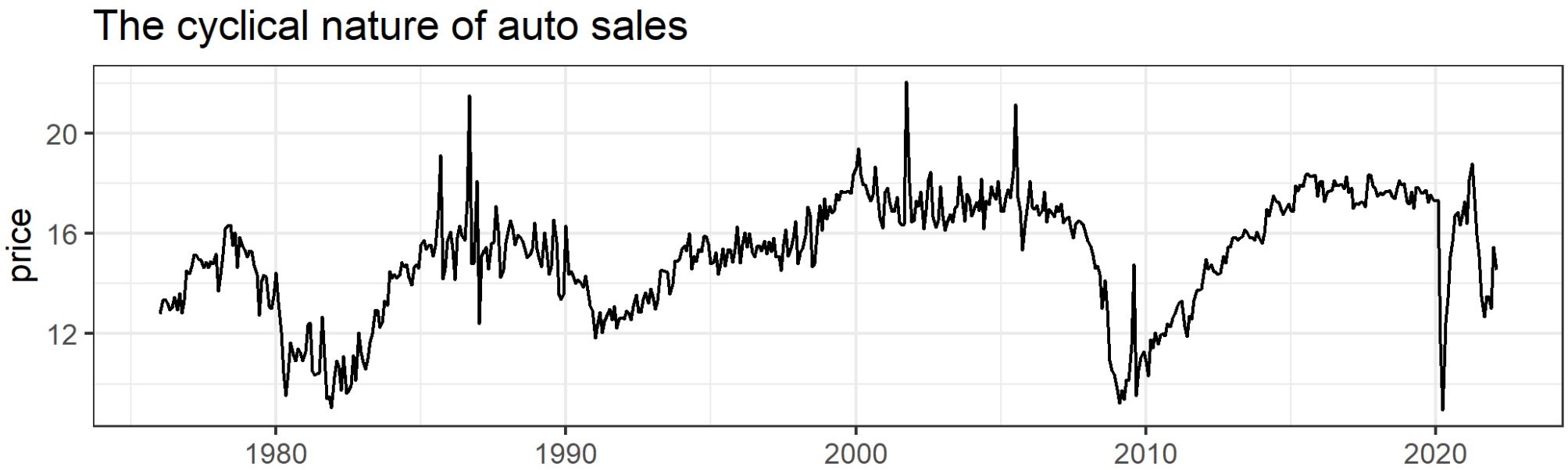
# Seasonality

**Seasonality** refers to the property of a time series that displays REGULAR patterns that repeat at a constant frequency (*m*).



# Cycle

**Cyclical fluctuations** are somewhat irregular (unknown duration).



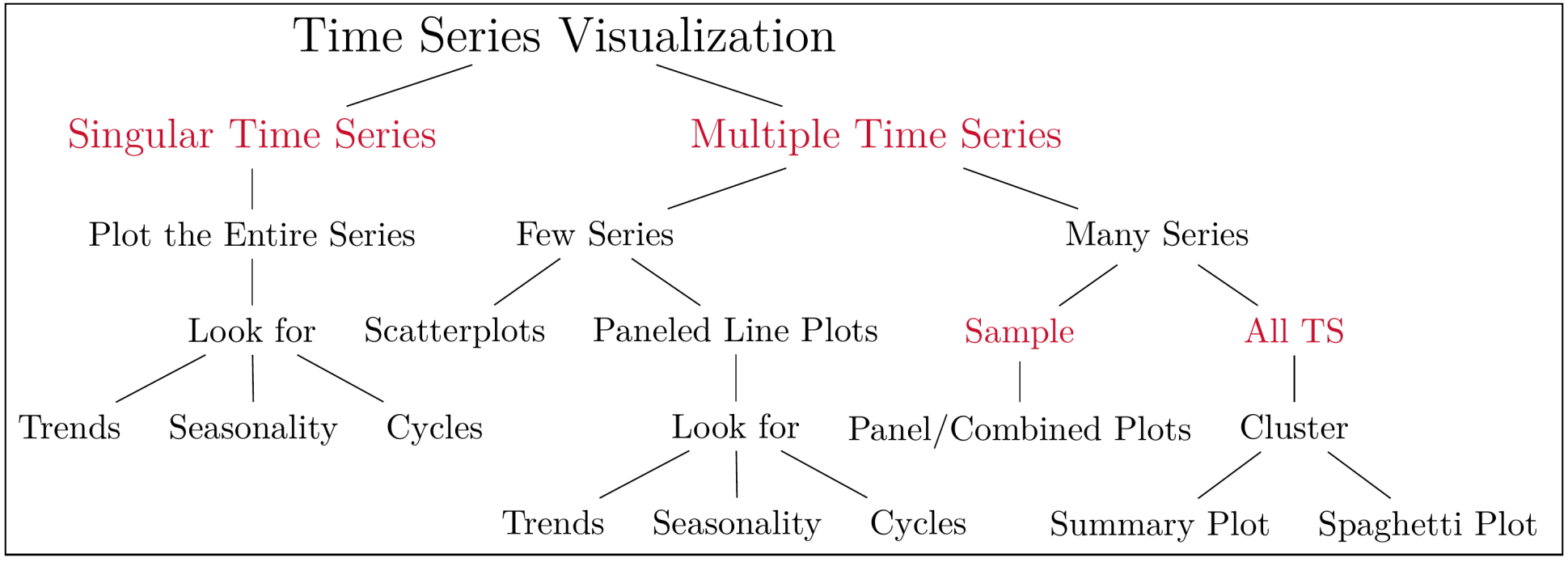

U.S. Bureau of Economic Analysis, Total Vehicle Sales [TOTALSA], retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/TOTALSA, on March 29, 2022

# The Goals Behind Visualizing (Time Series) Data

# Exploratory vs Confirmatory Viz Goals

> *Visualizations can be used to explore data, to confirm a hypothesis, or to manipulate a viewer. . . In exploratory visualization the user does not necessarily know what he is looking for. This creates a dynamic scenario in which interaction is critical. . . In a confirmatory visualization, the user has a hypothesis that needs to be tested. This scenario is more stable and predictable. System parameters are often predetermined.*
> -- (Grinstein and Ward 2001, 22)

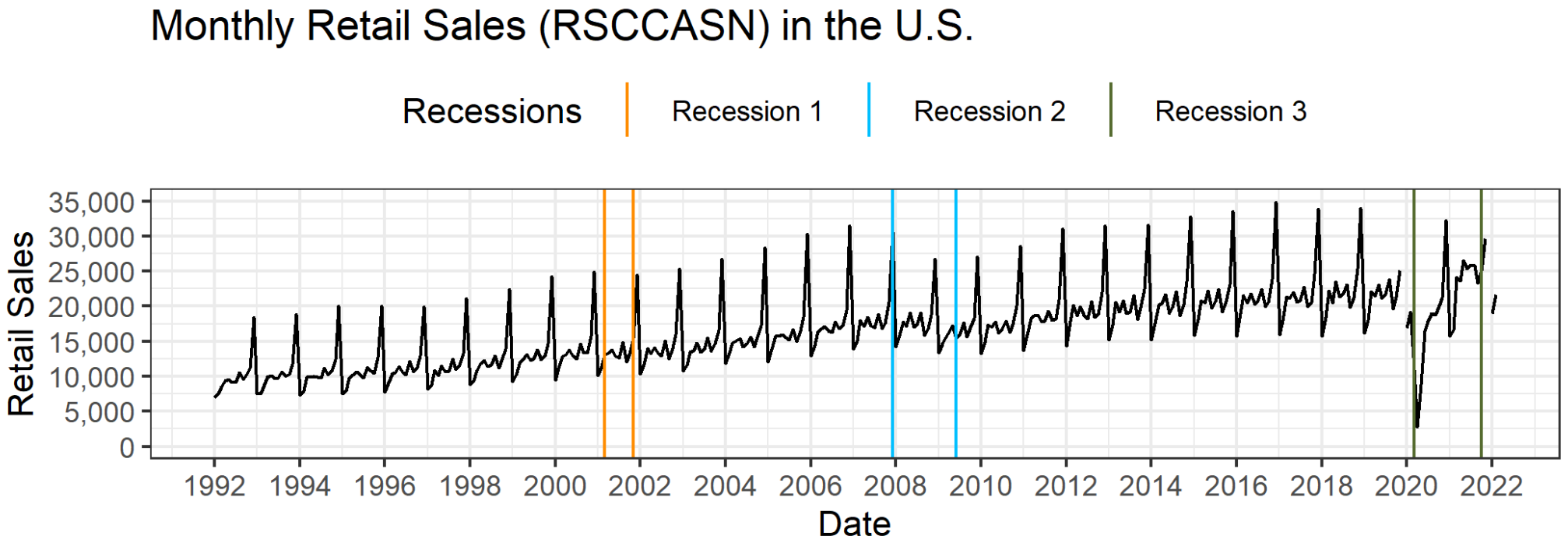
# A Structured Approach for Time Series Viz



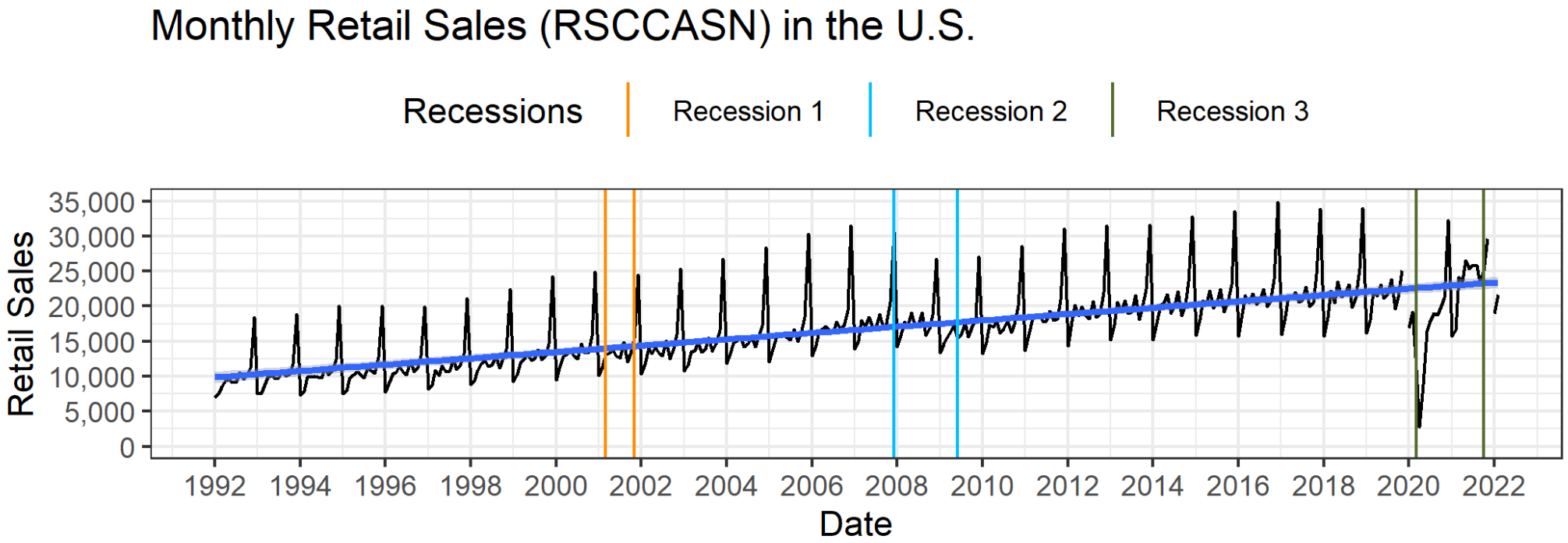

A Potential Framework for Time Series Visualization

This is my best attempt to improve on the general advice provided in the previous slide. Many of the suggestions, presented in this flow chart, stem from my past and current research/consulting collaborations. They are by no means a comprehensive list of everything that you can do.

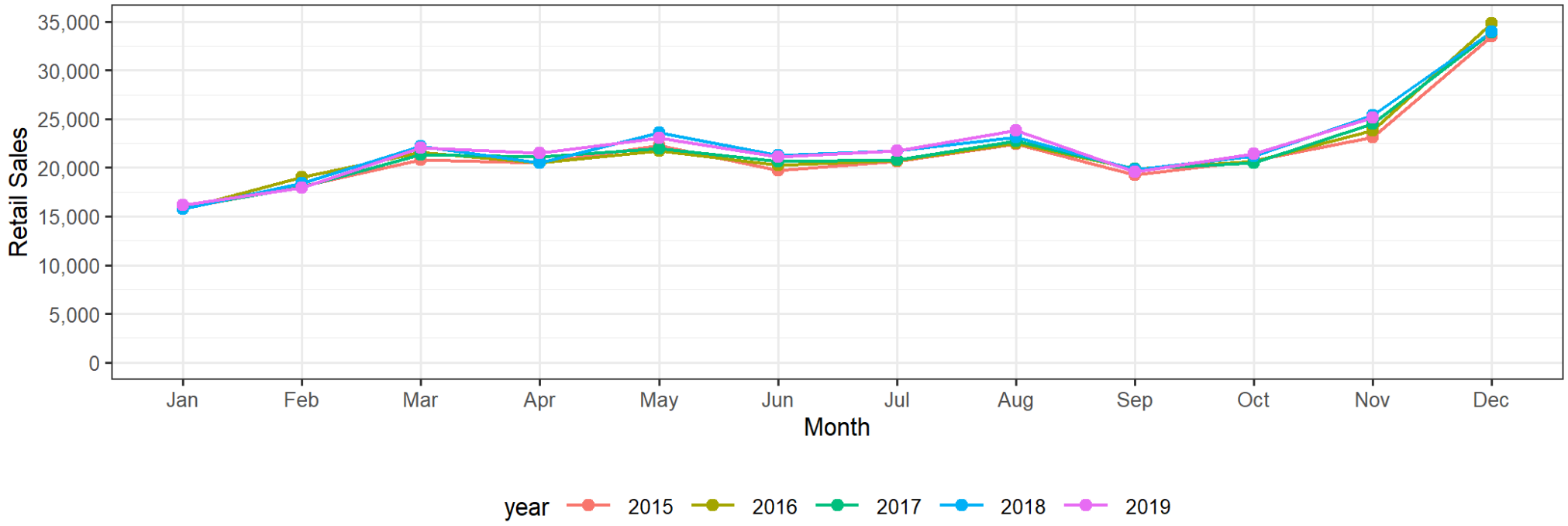
# Singular TS


Monthly Retail Sales (RSCCASN) in the U.S.
Recessions
Recession 1
Recession 2
Recession 3
Retail Sales
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
Date

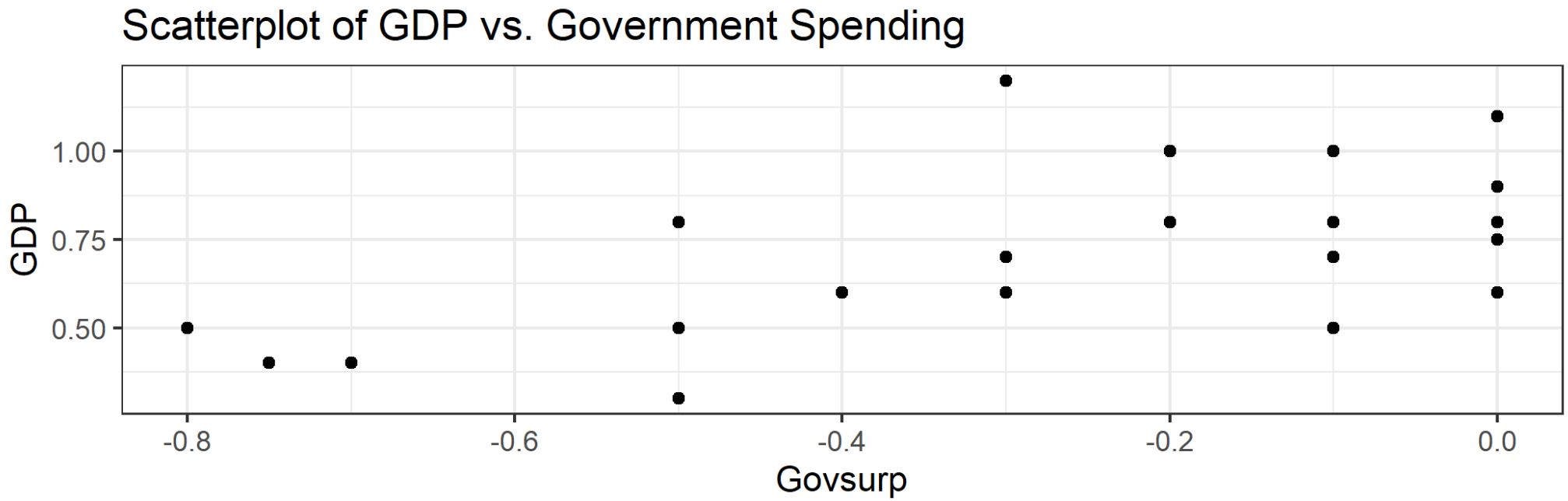
# Looking for Trends



# Looking for Seasonality

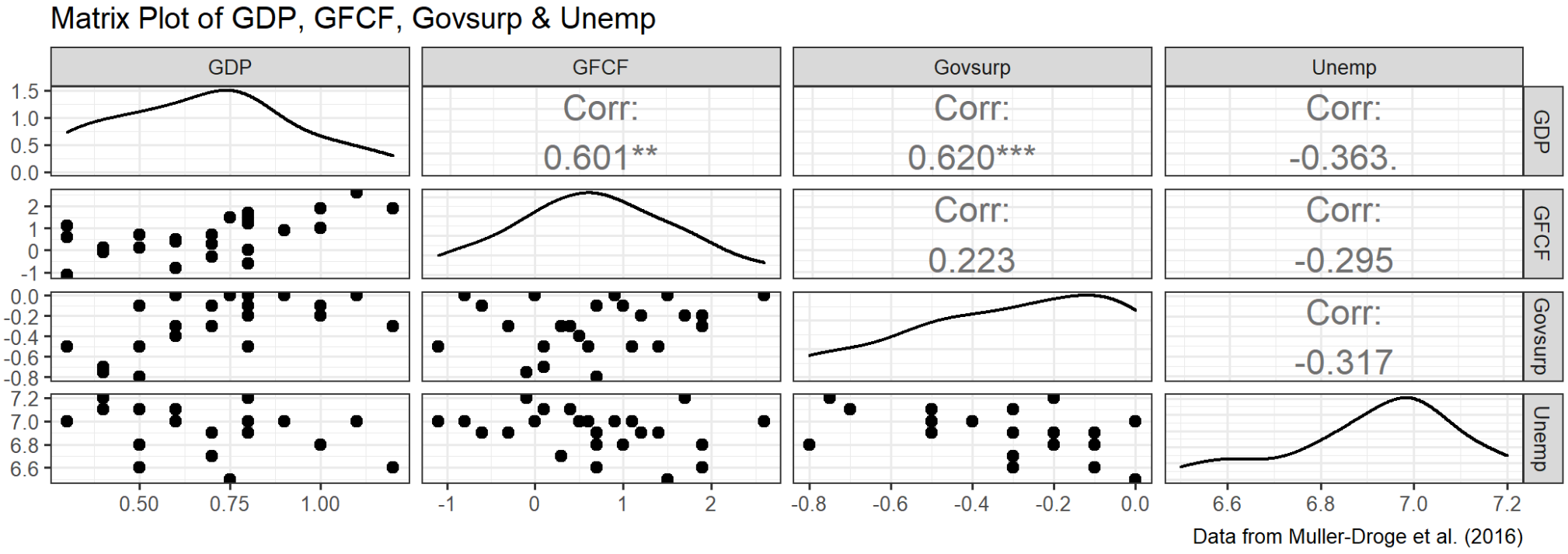


# Multiple TS: Scatterplots




Data from Muller-Droge et al. (2016)

# Multiple TS: Scatterplot Matrix



# Multiple TS: Panel Plots

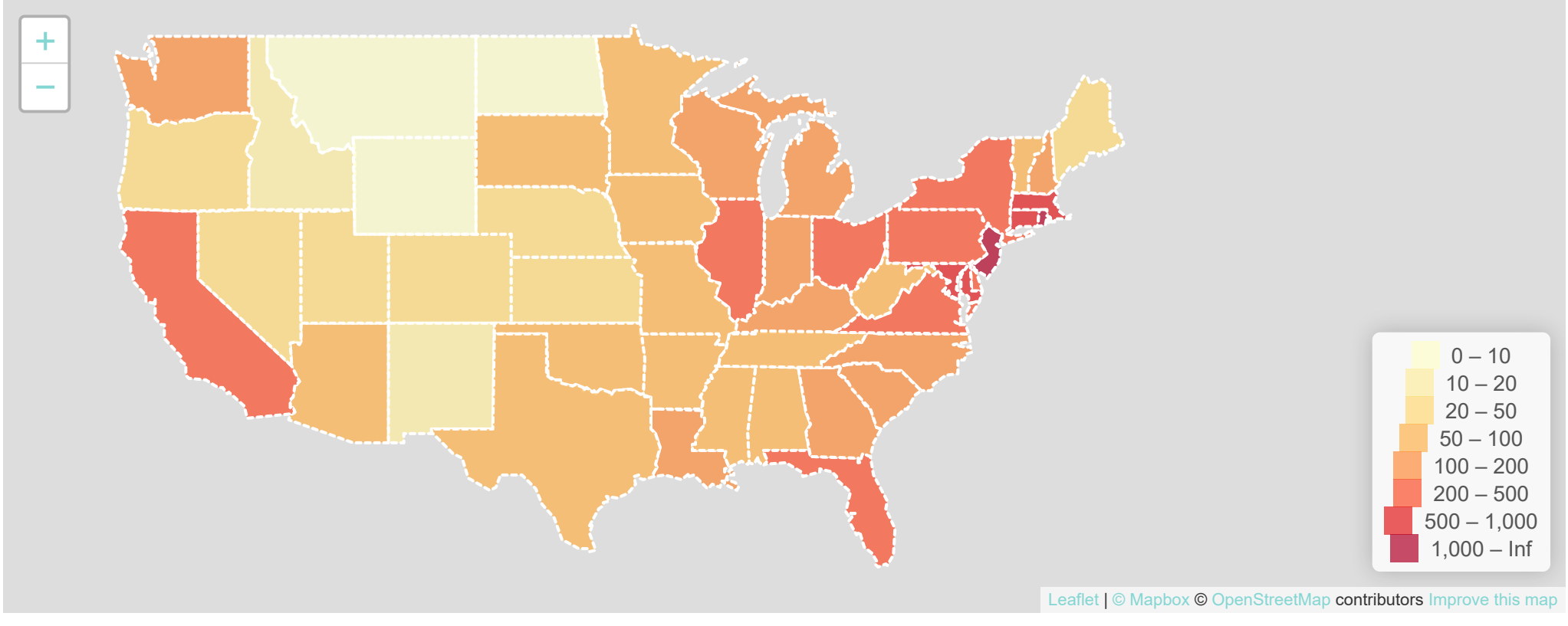
# Spaghetti Plots (Often w/ Clustering)

# Summary Plots (Often w/ Clustering)

# Spatial Plots

# Choropleth Maps

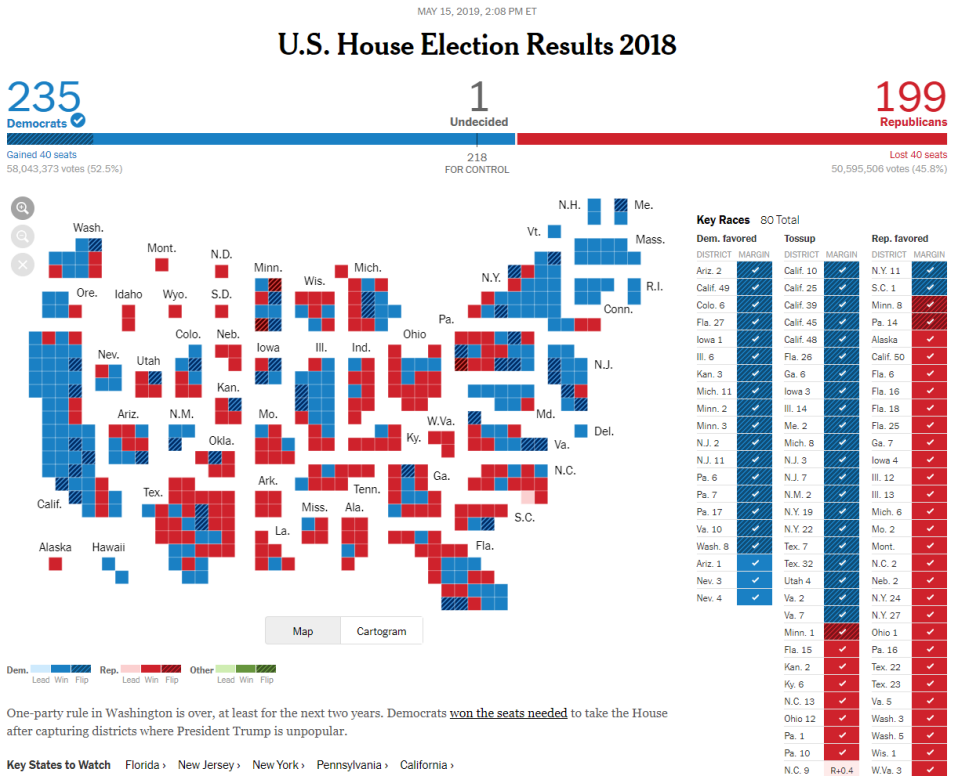
Maps where areas are **shaded, colored, or patterned** relative to a data attribute value.




Population Density in U.S.

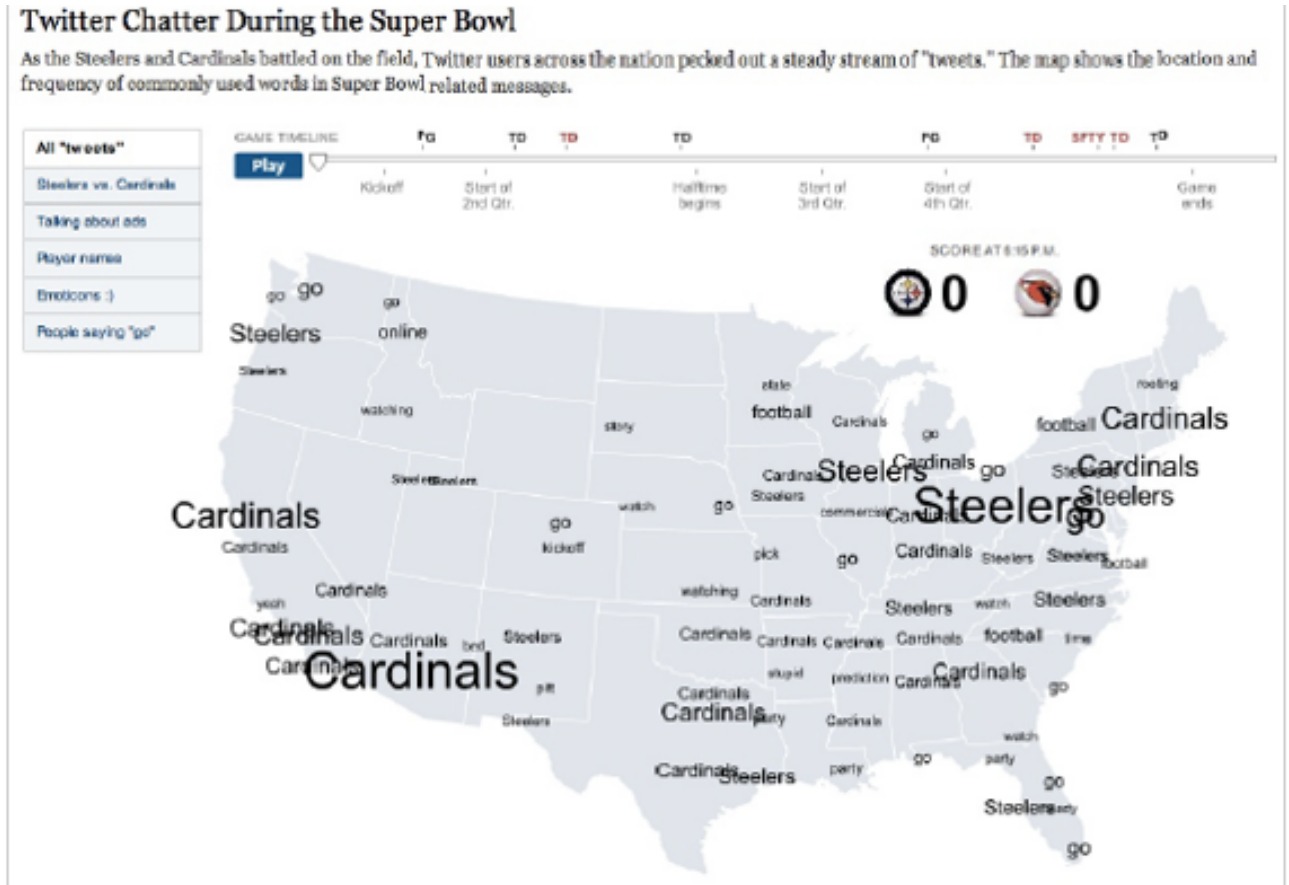
# Cartograms

A cartogram is a map in which **areas are scaled and distorted relative to a data attribute value**




The NYT's U.S. House Election Results 2018

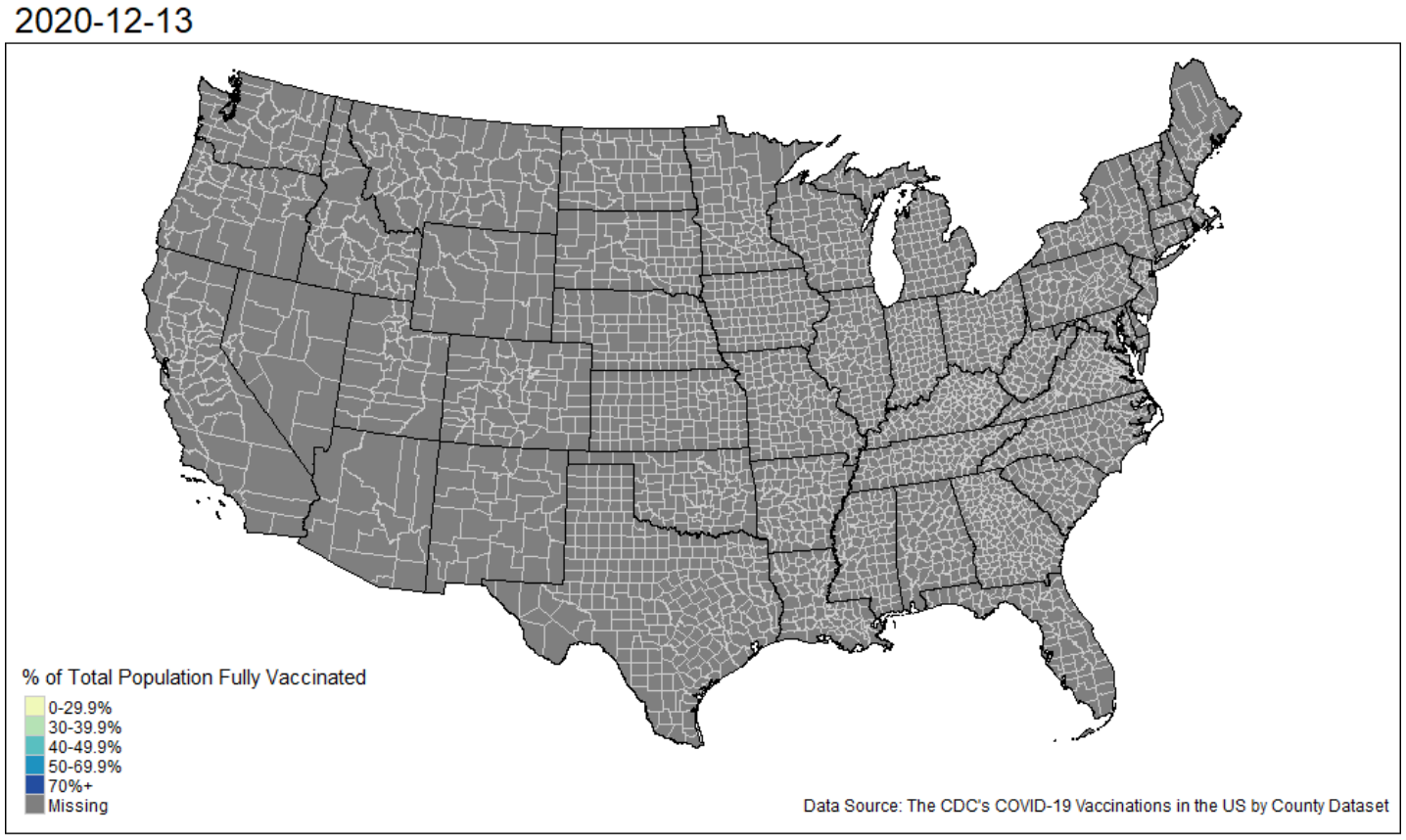
# Proportional Symbols Map




The NYT's U.S. Coverage of 2009 Super Bowl

# Spatiotemporal Maps

# Spatiotemporal Maps



# Recap

# Summary of Main Points

- Understand main goals behind visualizing time-series data
- Explain the different types of charts for univariate and multivariate time-series
- Explain the different types of spatial plots
- Select suitable spatial graphs for different scenarios
- Understand how spatiotemporal plots can help in storytelling (what makes BI special)