

# ISA 444: Business Forecasting

## 01: Introduction to Time Series Analysis and Forecasting

Fadel M. Megahed, PhD

Endres Associate Professor  
Farmer School of Business  
Miami University

 @FadelMegahed

 fmegahed

 fmegahed@miamioh.edu

 Automated Scheduler for Office Hours

Spring 2023

# Learning Objectives for Today's Class

- Describe **course motivation** and **structure**.
- Explain the differences between **cross sectional**, **time series** and **panel** datasets.
- Describe the **components of time series** datasets.
- Explain the **forecasting steps**.

# Course Motivation and Structure

# The Analytics Journey: Pre-analytics

**Pre-Analytics/Data Management:** where one attempts to **extract** the needed *data* for analysis.

# The Analytics Journey: Pre-analytics

**Pre-Analytics/Data Management:** where one attempts to **extract** the needed *data* for analysis.

Show 

6

 entries

Search:

	DATE	RSCCASN
1	Jan 1992	6,938
2	Feb 1992	7,524
3	Mar 1992	8,475
4	Apr 1992	9,401
5	May 1992	9,558
6	Jun 1992	9,182

Showing 1 to 6 of 372 entries

# The Analytics Journey: Descriptive

**Descriptive Analytics:** where one attempts to **understand** the data through **visualizations**, **descriptive statistics**, and **exploratory statistical models**.

Activity	Viz 1	Viz 2	Viz 3	Viz 4	Key Points
----------	-------	-------	-------	-------	------------

Over the next 3 minutes, please identify what you have learned from the charts in each tab.

- Write down your answers in the last tab (it is editable).
- Discuss your answers with your neighboring classmates.
- Be prepared to share these answers with class.

# The Analytics Journey: Descriptive

**Descriptive Analytics:** where one attempts to **understand** the data through **visualizations**, **descriptive statistics**, and **exploratory statistical models**.

Activity

Viz 1

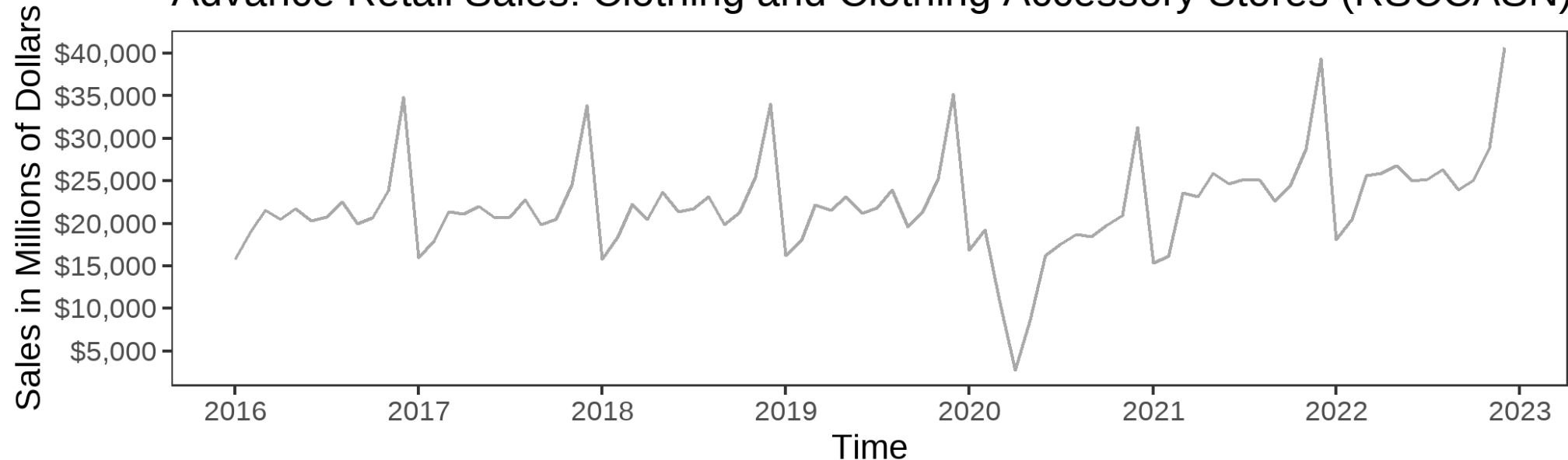
Viz 2

Viz 3

Viz 4

Key Points

Advance Retail Sales: Clothing and Clothing Accessory Stores (RSCCASN)



# The Analytics Journey: Descriptive

**Descriptive Analytics:** where one attempts to **understand** the data through **visualizations**, **descriptive statistics**, and **exploratory statistical models**.

Activity    Viz 1    **Viz 2**    Viz 3    Viz 4    Key Points





# The Analytics Journey: Descriptive

**Descriptive Analytics:** where one attempts to **understand** the data through **visualizations**, **descriptive statistics**, and **exploratory statistical models**.

Activity

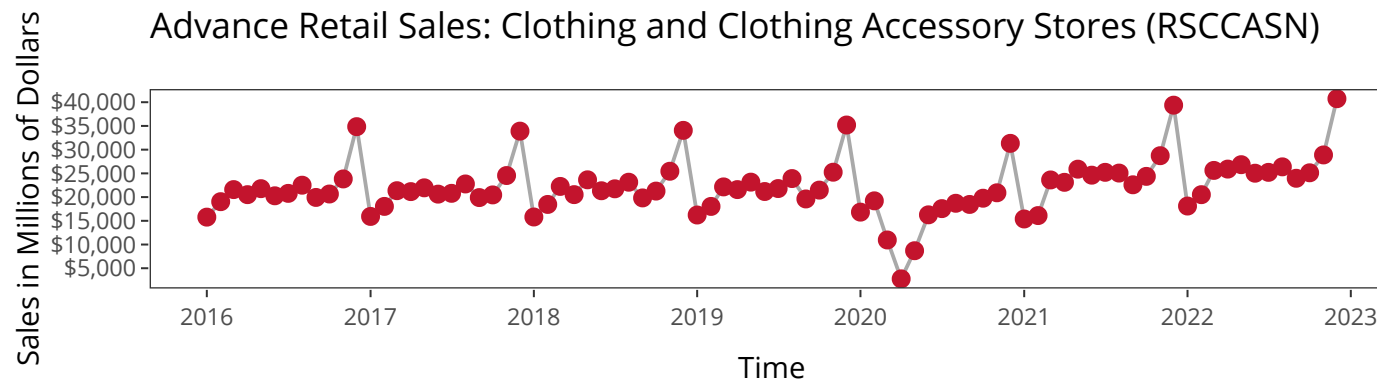
Viz 1

Viz 2

Viz 3

Viz 4

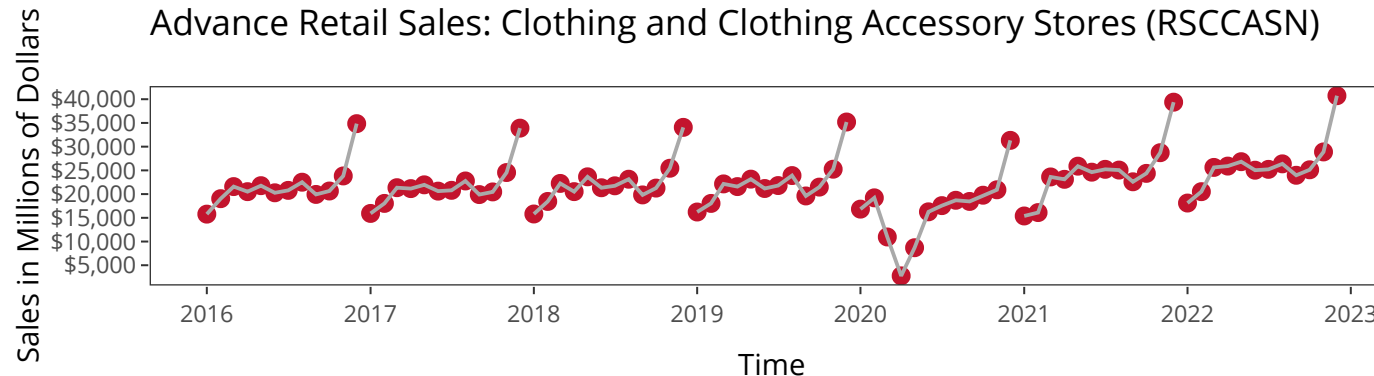
Key Points



# The Analytics Journey: Descriptive

**Descriptive Analytics:** where one attempts to **understand** the data through **visualizations**, **descriptive statistics**, and **exploratory statistical models**.

Activity    Viz 1    Viz 2    Viz 3    **Viz 4**    Key Points



# The Analytics Journey: Descriptive

**Descriptive Analytics:** where one attempts to **understand** the data through **visualizations**, **descriptive statistics**, and **exploratory statistical models**.

Activity	Viz 1	Viz 2	Viz 3	Viz 4	Key Points
----------	-------	-------	-------	-------	------------

**Main Insight(s):** (Insert below)

- Edit me
- ...
- ...

# The Analytics Journey: Descriptive

**Descriptive Analytics:** where one attempts to **understand** the data through **visualizations**, **descriptive statistics**, and **exploratory statistical models**.

Show 

6

 entries

Search:

Month	Seasonal Factor
Jan	-4,423.84
Feb	-2,845.91
Mar	-761.87
Apr	-1,231.37
May	-91.67
Jun	-1,103.19

Showing 1 to 6 of 12 entries

Previous

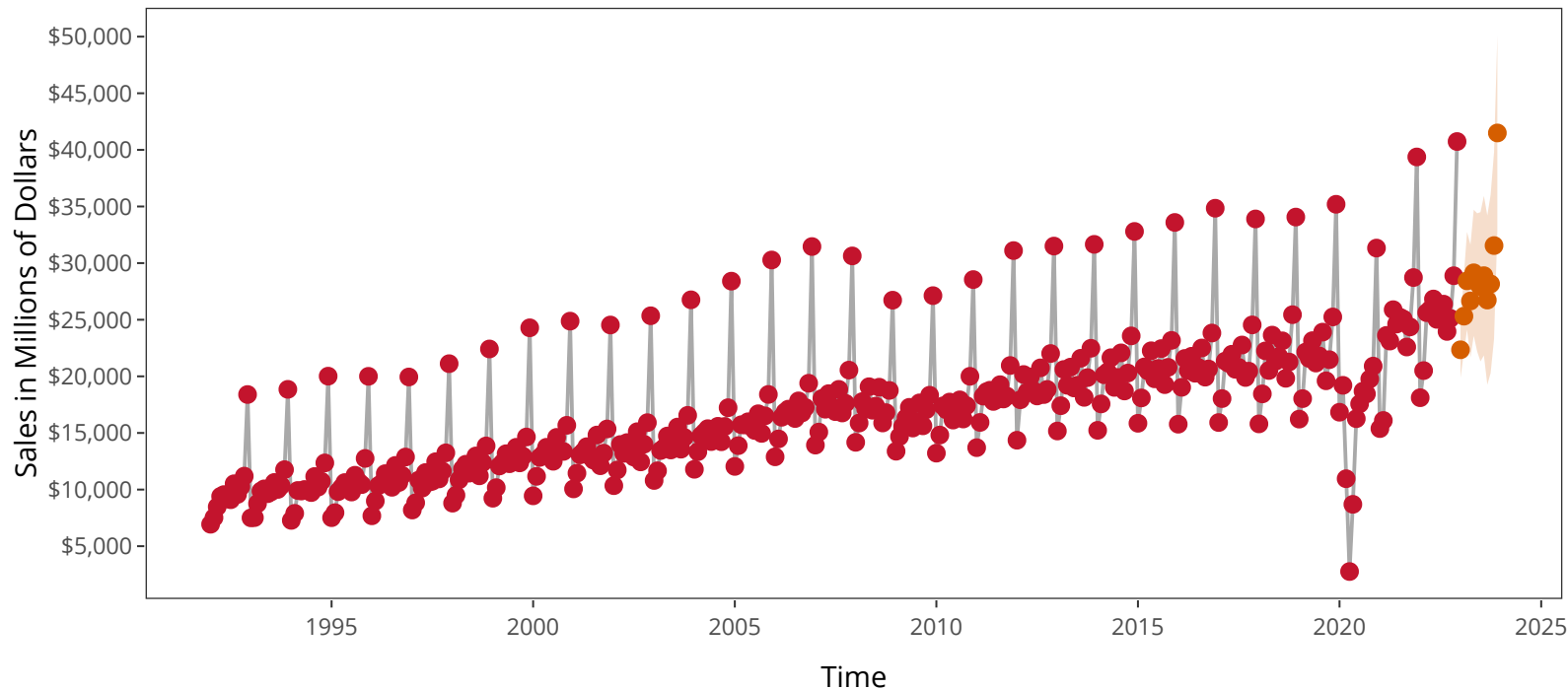
1

2Next

# The Analytics Journey: Predictive

**Predictive Analytics:** where **statistical** and **machine learning** models are used to help us utilize independent variable[s] to predict an outcome variable of choice.

Advance Retail Sales (RSCCASN) with **forecast for 2023 in orange**

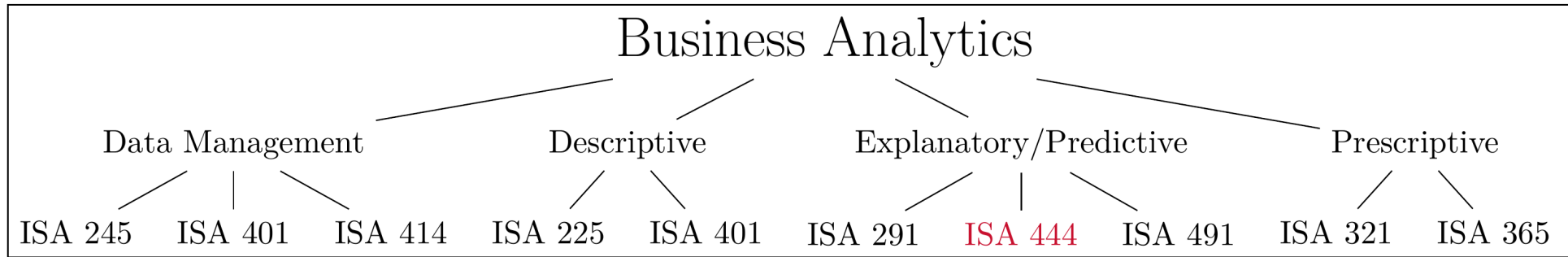


# The Analytics Journey: Prescriptive

**Prescriptive Analytics:** where **mathematical models** are used to make recommendations for business actions.

- Our **overarching goal** behind data/business analytics, is to **make informed decisions based on what we have learned from the data**. Hence, this stage is where we build on what we learned during the *descriptive* and *predictive* stages to make more informed decisions.
- Imagine that you are L Brands or Express Inc.; how would you use the information pertaining to the **U.S. trends in clothing sales for staffing and operational decisions**?

# How does our Curriculum at Miami University Prepare you for this Journey?

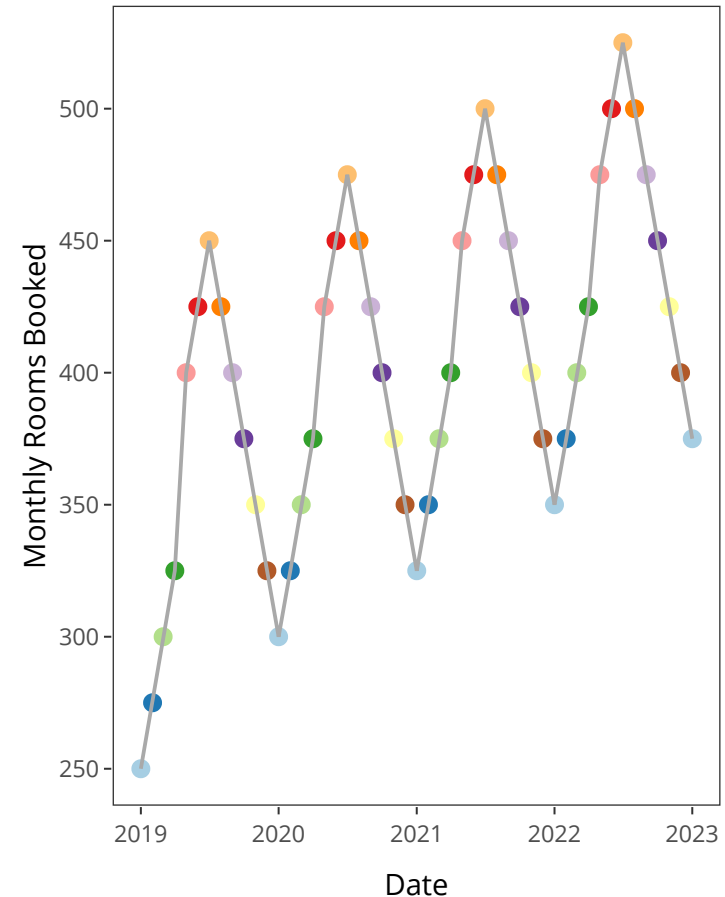


My take on the courses within the business analytics major/minor at Miami University

# Why are time series data different?

- Time series (TS) data **cannot be** analyzed using standard regression techniques.
- In regression, you assume that your data are **randomly sampled, resulting in independent observations**.
- In TS, we want to capitalize on the **structure of the data**, i.e., observations are typically **correlated over time (autocorrelated)** to build our models.

ChatGPT-based data of resort hotel bc





# Course Objectives


By the **end of this course**, you should be able to:

- Explain the purpose of forecasting in a business setting.
- Use the basic tools of forecasting including plots, summary measures, transformations, measures of forecast accuracy, and prediction intervals.
- Use an appropriate smoothing-based method for time series analysis and forecasting.
- Use an appropriate ARIMA model to forecast a time series.
- Use simple and multiple linear regression models to forecast a time series.

# Why Should You Care?

- **Sales forecasting:** Businesses use time series analysis to forecast future sales based on historical sales data, in order to make informed decisions about inventory management, staffing, and marketing efforts.
- **Financial forecasting:** Companies use time series analysis to forecast future financial performance, such as revenue, expenses, and profits, in order to make strategic business decisions and plan for future growth.
- **Demand forecasting:** Manufacturers and retailers use time series analysis to forecast demand for their products, in order to optimize production and inventory levels.
- **Resource planning:** Businesses use time series analysis to forecast resource usage, such as energy consumption, in order to plan for future infrastructure needs and reduce costs.

# Why Should You Care?

Find jobs Company reviews Find salariesUpload your resume Sign in Employers / Post Job

What forecasting r \$85,000

Where United States

Search Jobs

Date Posted

Remote

\$85,000+

Job Type

Encouraged to Apply

Location

Company

Entry Level

Education

Upload your resume - Let employers find you

forecasting r \$85,000 jobs in United States

Sort by: relevance - date 118 jobs

Data Scientist - Analytics

Garner Health

Hybrid remote in New York, NY

\$90,000 - \$135,000 a year

- Develop forecasting models for critical business KPIs that drive strategy.
- Partner with Engineering, Product Management, & Customer Success to ensure Garner has...

Posted 9 days ago · More...

Advanced Analytics Entry-Level Opportunities

Verizon 3.9 ★

Boston, MA 02114 (West End area)

Estimated \$85.2K - \$108K a year Full-time

- Curating data products to accelerate value and speed to decision.
- Conducting automated post-mortems and provide diagnostic/descriptive analytics to explain...

Posted 30+ days ago · More...

Data Scientist - Analytics

Garner Health

New York, NY · Hybrid remote

\$90,000 - \$135,000 a year

Apply now

Benefits

Pulled from the full job description

Health insurance

Garner's mission is to transform the healthcare economy, delivering high quality and affordable care for all. By helping employers restructure their healthcare benefit to provide clear incentives and data-driven insights, we direct employees to higher quality and lower cost healthcare providers. The result is that patients get better health outcomes while doctors are rewarded for practicing well, not performing more procedures. We are backed by top-tier venture capital firms, are growing rapidly and looking to expand our team.

Garner is headquartered in NYC and this person would be expected to work from the office 3 days a week.

As a Data Scientist - Analytics, you will...

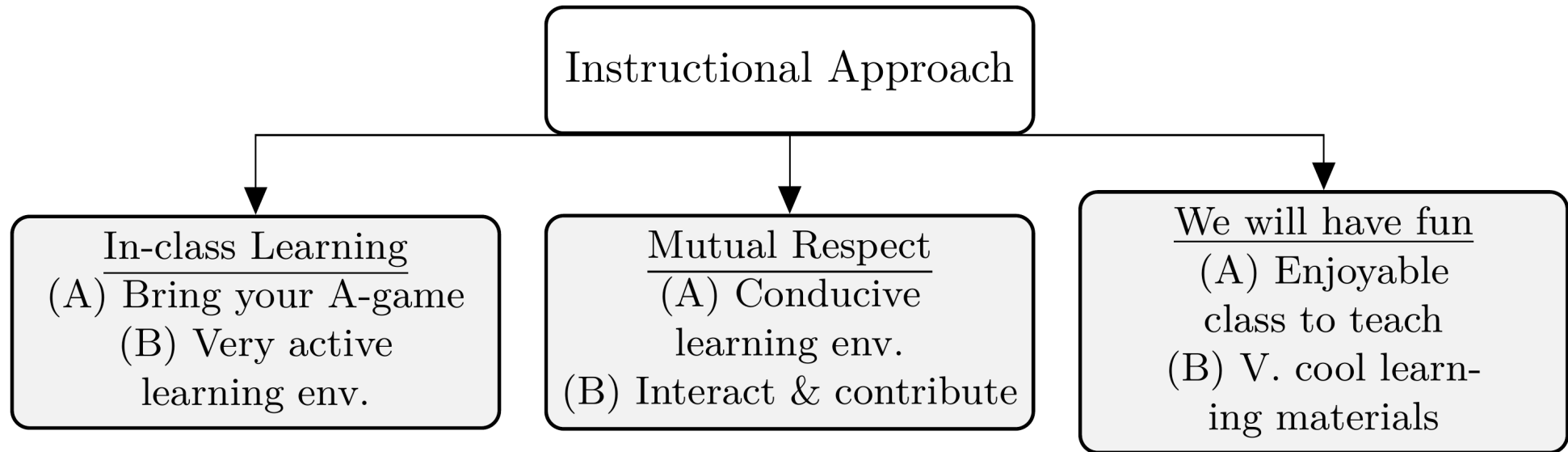
- Partner with Engineering, Product Management, & Customer Success to ensure Garner has the right infrastructure and tools in place to make data-driven decisions across all departments
- Design and build business intelligence dashboards that produce client facing reports at scale
- Uncover answers to open-ended questions to inform decisions across departments such as Product, Data, Marketing, and Account Management
- Develop forecasting models for critical business KPIs that drive strategy

The ideal candidate has:

- 1+ years of experience in a fast-paced environment

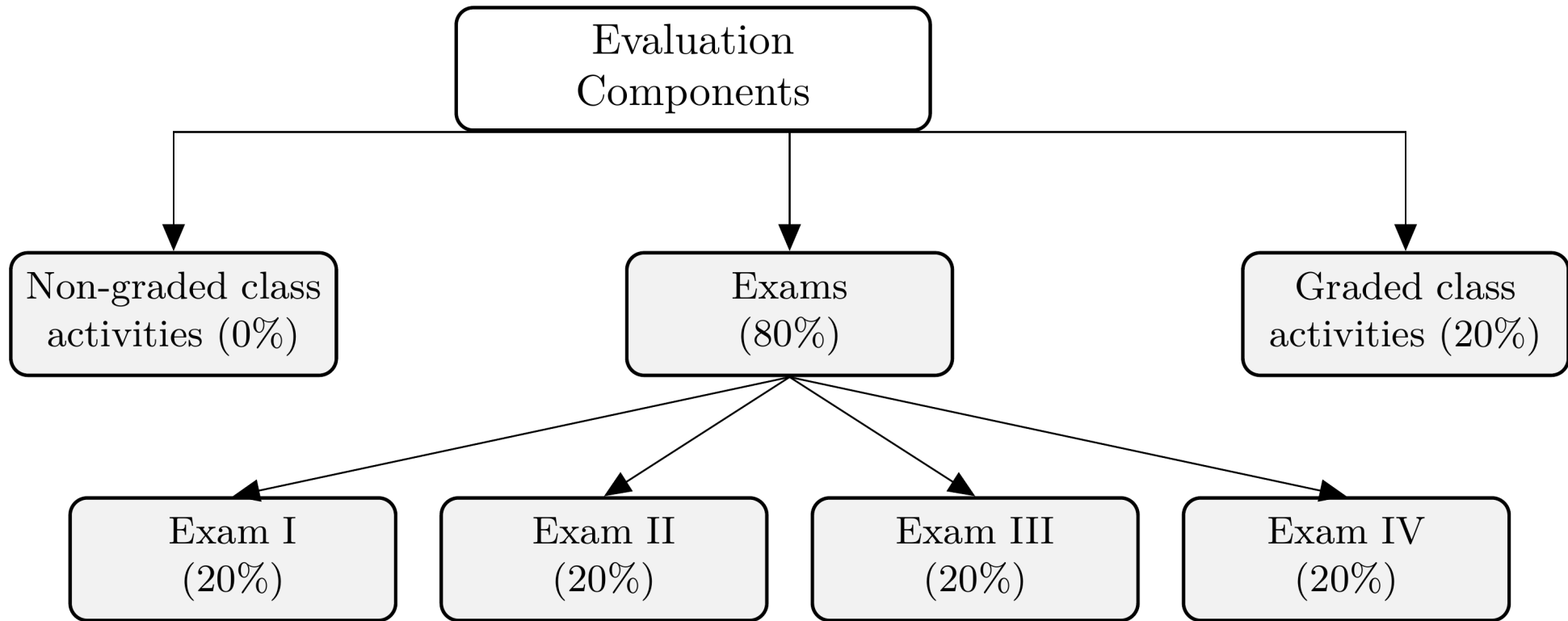
11 / 32

# Instructional Approach



An overview of the instructional approach for ISA 444.

# How will I Evaluate your Learning?



An overview of the evaluation components for ISA 444.

# 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 23 entries

Previous

1

2

3

4

5

Next

# Cross Sectional Data

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

**Example 2:** NBA 2022-2023 Leaders - Top Players in PTS/Game

	Player	Pos	Age	Tm	G	FG	FG%	eFG%	PTS
1	Luka Dončić	PG	23	DAL	41	11.3	0.497	0.559	33.7
2	Joel Embiid	C	28	PHI	34	11.2	0.537	0.563	33.6
3	Jayson Tatum	PF	24	BOS	44	10.2	0.468	0.544	31.2
4	Giannis Antetokounmpo	PF	28	MIL	35	10.9	0.524	0.541	31
5	Shai Gilgeous-Alexander	PG	24	OKC	41	10	0.501	0.526	30.5

Showing 1 to 5 of 503 entries

**Source:** Data scraped from [https://www.basketball-reference.com/leagues/NBA\\_2023\\_per\\_game.html](https://www.basketball-reference.com/leagues/NBA_2023_per_game.html) on January 23, 2023 using the [rvest](#) package. The printing was limited to the selected variables.

Previous

1

2

3

4

5

...

101

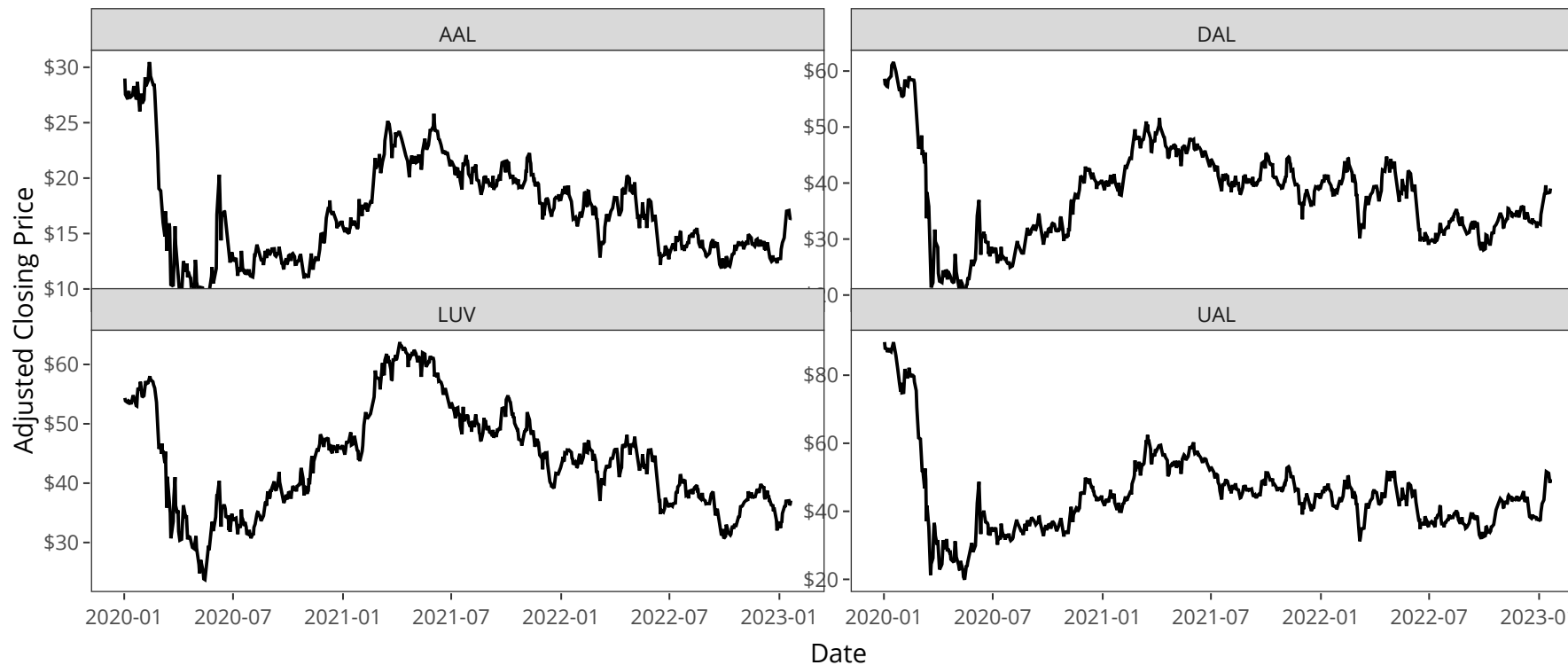
Next



# 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.BED
1	CHN	2019	6	3.7		
2	CHN	2020	2.2	3.5		
3	CHN	2021	8.1	3.5		
4	EGY	2019	5.6	10.7		
5	EGY	2020	3.6	10.3		

Showing 1 to 5 of 9 entries

Previous

1

2

Next

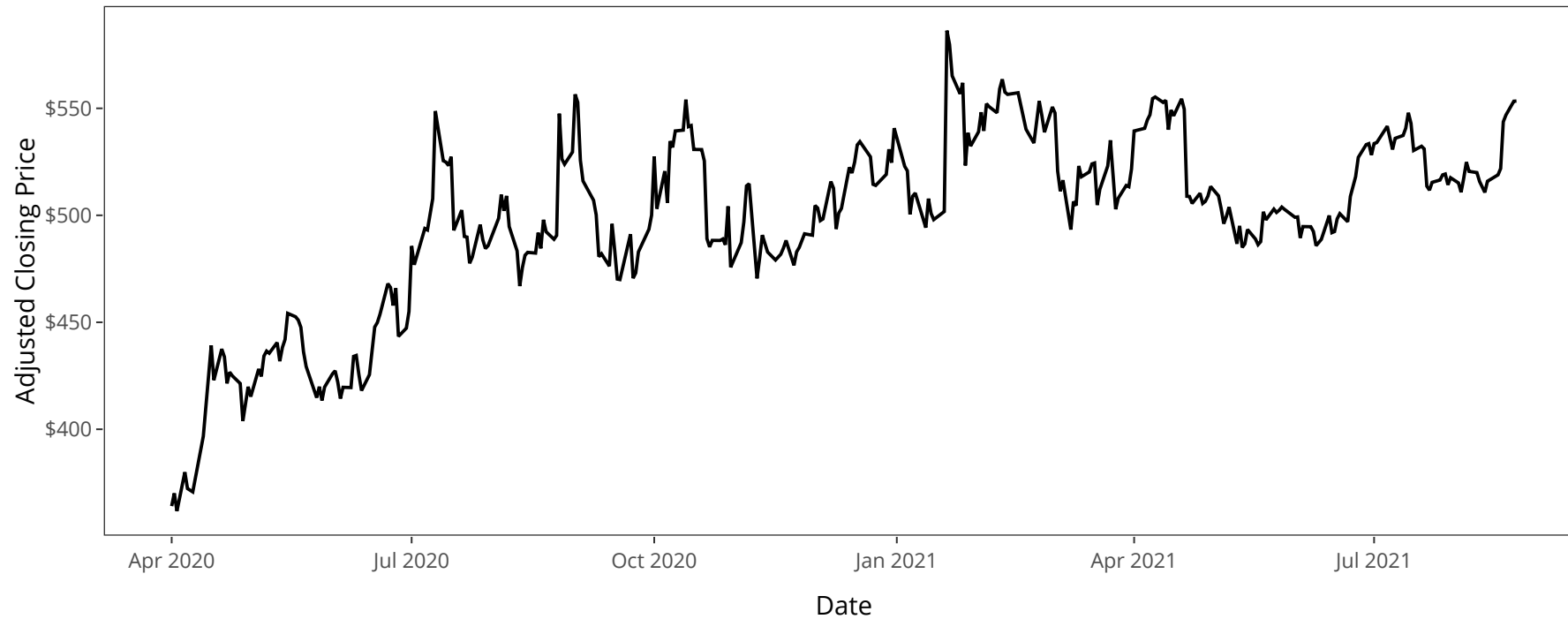
**Source:** Data queried from the [World Bank Data](#) using the [wbstats](#) in R. The printed results show a snapshot of 7 variables (out of a much larger panel dataset). You can think of panel data as a cross-sectional dataset with a longitudinal/time component.

# Components of a Time Series

# Trend

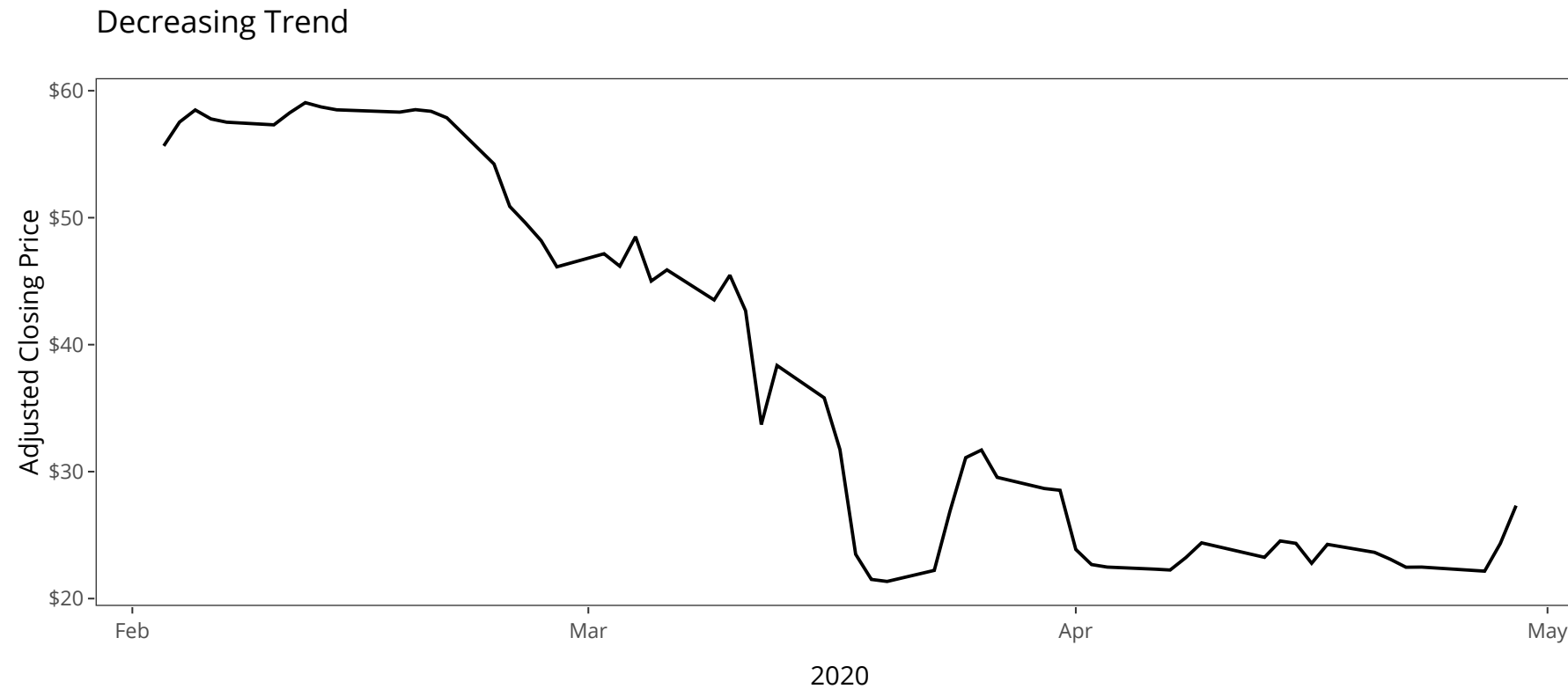
A **trend** is an increasing or decreasing pattern over time.

An increasing trend: rise of \$NFLX during the COVID-19 Pandemic



# Trend

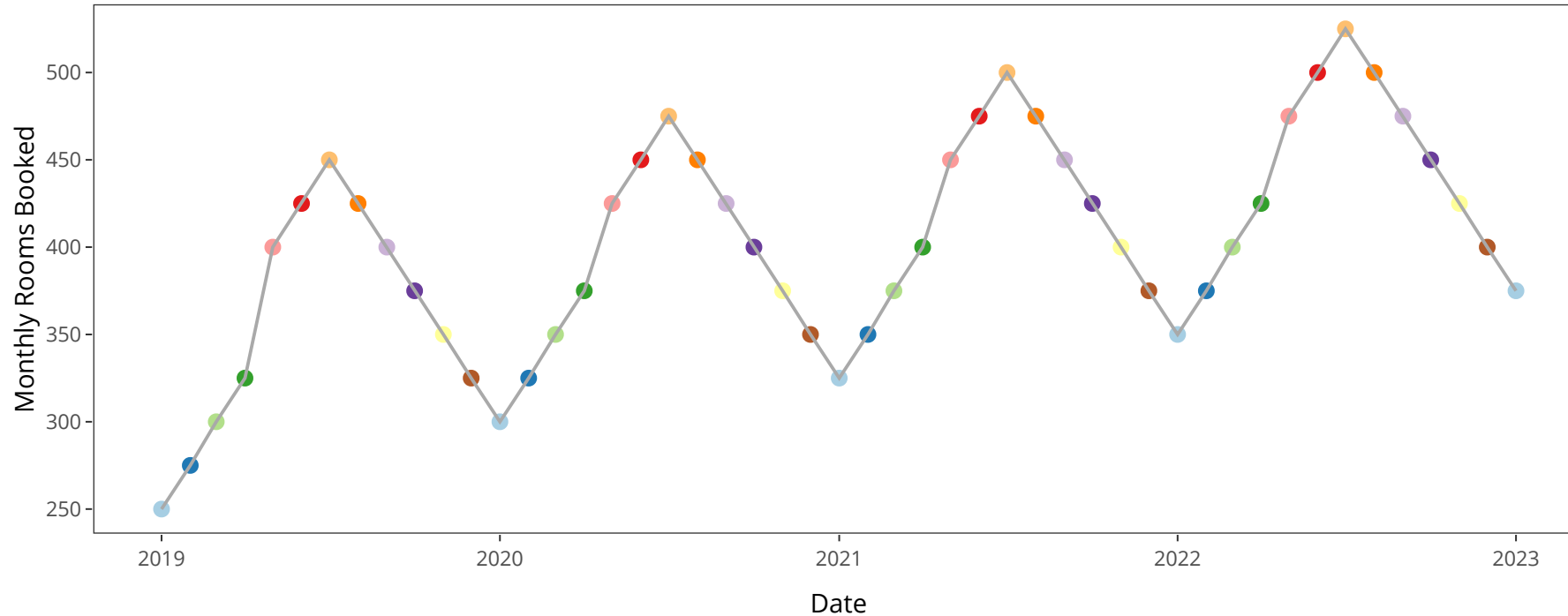
A **trend** is an increasing or decreasing pattern over time.



# Seasonality

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

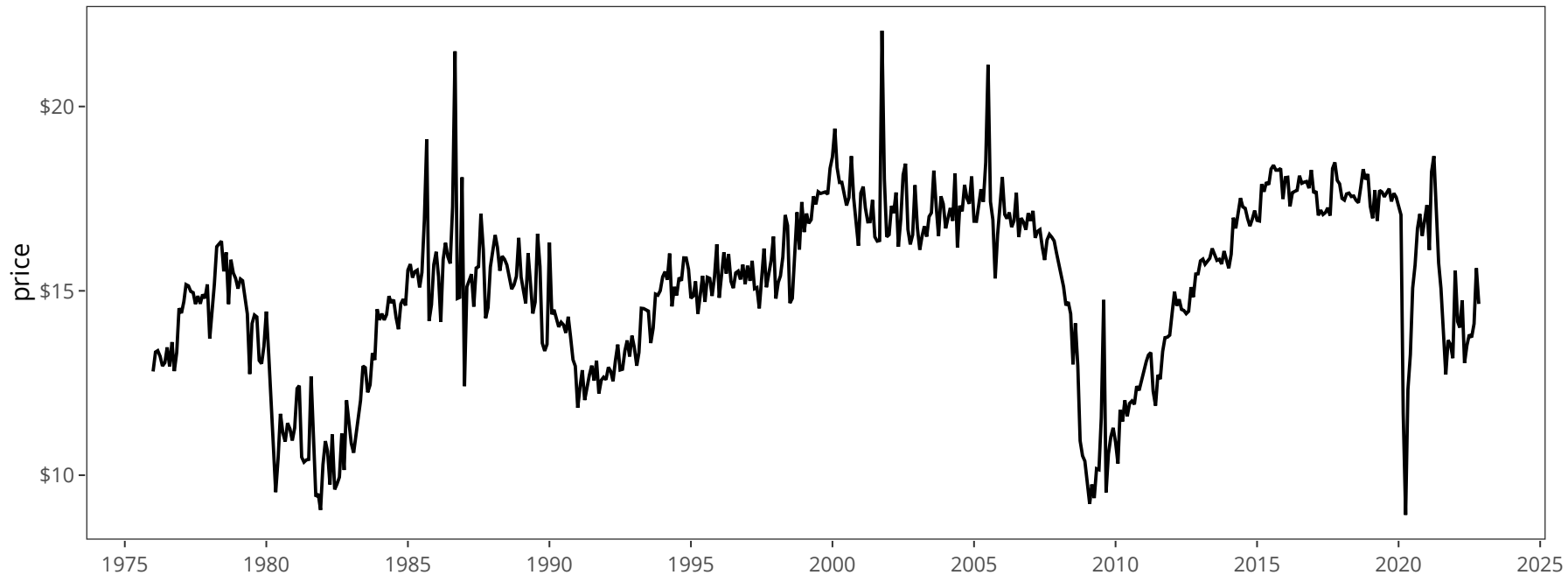
ChatGPT-based data of resort hotel bookings



# Cycle

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

The cyclical nature of auto sales



# Kahoot Competition

To assess your understanding and retention of the topics covered so far, you will **compete in a Kahoot competition (consisting of 5 questions)**:

- Go to <https://kahoot.it/>
- Enter the game pin, which will be shown during class
- Provide your first (preferred) and last name
- Answer each question within the allocated 20-second window (**fast and correct answers provide more points**)

**Winning the competition involves having as many correct answers as possible AND taking the shortest duration to answer these questions.** The winner 🏆 of the competition will receive a gift of their choosing (\$10 Starbucks gift card or a large chocolate bar). Good luck!!!



# Forecasting and Steps

# So What is Forecasting?

[ALL](#) [NEWS](#) [IMAGES](#) [VIDEOS](#) [MAPS](#) [SHOPPING](#) [MORE](#)


About 16,900,000 results [Any time](#) ▾

Dictionary

Data from [Oxford Languages](#)

Look it up

## fore·cast

[ˈfɔr,kɑst] 

**VERB**

1. predict or estimate (a future event or trend):  
*"rain is forecast for eastern Ohio" · "coal consumption is forecast to increase"*

SIMILAR: [predict](#) [prophecy](#) [prognosticate](#) [augur](#) [divine](#) [foretell](#) ▾

---

**NOUN**


1. a prediction or estimate of future events, especially coming weather or a financial trend.

SIMILAR: [prediction](#) [prophecy](#) [forewarning](#) [prognostication](#) [augury](#) ▾

Translate forecast to 

Choose language ▾

More definitions and word origin



# Working Definitions

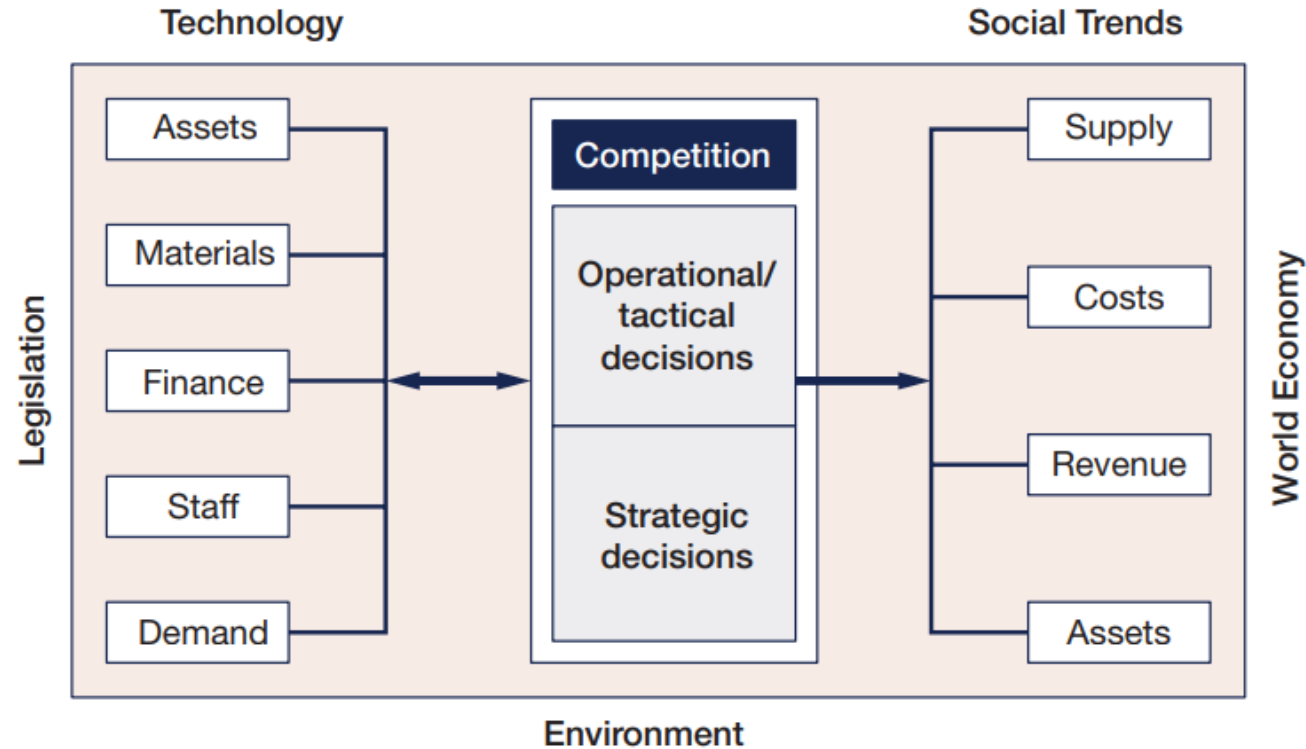
**Forecasting:** The process of predicting a future event. The objective of most time series analyses is to provide forecasts of future values of the time series.

**Time Series:** A time series is a sequence of observations on a variable measured at successive points in time or over successive periods of time. In the simplest cases, observations are evenly spaced at regular intervals such as hourly, daily, weekly, monthly, or yearly, or at any other regular interval

# Why do we Forecast? (PIVASE)

- **Purpose:** What do we hope to achieve by generating the forecast? That is, what plans are dependent upon the results of the forecasting exercise? How far ahead do we wish to forecast? We refer to this period as the forecasting horizon.
- **Information:** What do we know that may help us in forecasting. And when will we know it? Detailed data is only useful if it is available in timely fashion.
- **Value:** How valuable is the forecast? What would you pay for perfect knowledge?
- **Analysis:** From analyzing the data can we develop a model that captures its characteristics? And how does it perform on new (hold-out sample) data?
- **System:** What models and software are needed to meet the needs of the organization?
- **Evaluation:** How do we know whether a particular forecasting exercise was effective and what the potential is for improvement?

# Why do Businesses Forecast?



An Organization's Forecasting Needs

# Forecasting Step by Step

On the basis of our preliminary discussion of **PIVASE**, we can identify **seven major steps in the forecasting process**:

- Define the forecasting and planning problem, the forecast horizon and decide the value of better forecasts.
- Determine the resources to be devoted to providing the forecasts.
- Collect relevant information, whether from a survey, from company records, or from information generated by other agencies (e.g., government figures).
- Conduct an initial analysis of the data.
- Select an appropriate forecasting method.
- Generate forecasts.
- Evaluate the forecasting exercise by checking forecasts against actual outcomes.

# Recap

# Summary of Main Points

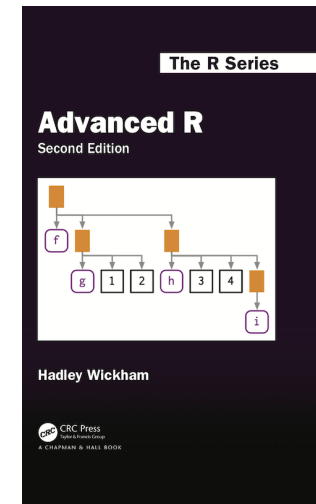
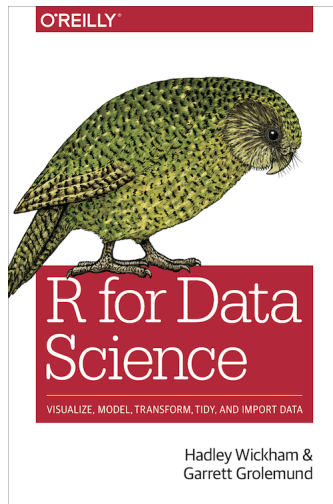
By now, you should be able to do the following:

- Describe **course motivation** and **structure**.
- Explain the differences between **cross sectional**, **time series** and **panel** datasets.
- Describe the **components of time series** datasets.
- Explain the **forecasting steps**.



# Things to Do to Prepare for Our Next Class

- Go over your notes and complete [Assignment 01](#) on Canvas.
- Read through [Chapter 01](#) of our reference book and skim the following:



- [Workflow: basics](#)
- [Workflow: scripts](#)
- [Workflow: project](#)
- [Names and values](#)
- [Vectors](#)
- [Subsetting](#)