

ISA 444: Business Forecasting

02 - Introduction to Time Series Analysis

Fadel M. Megahed

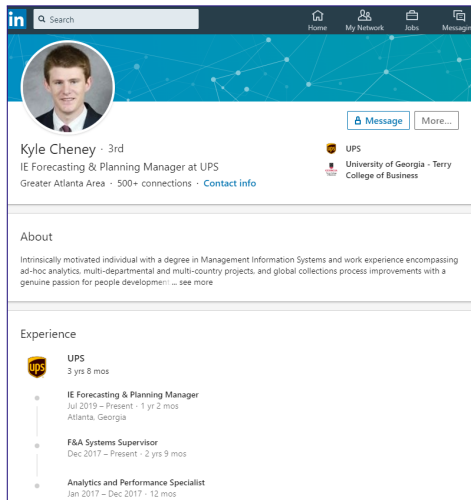
Associate Professor
Department of Information Systems and Analytics
Farmer School of Business
Miami University
Email: fmegahed@miamioh.edu
Office Hours: [Click here to schedule an appointment](#)

Fall 2020

Outline

- 1 **Preface**
- 2 Types of Data Over Time
- 3 Components of a Time Series
- 4 Recap

Announcement: Second Speaker Confirmed for Sept. 14, 2020



The image shows a LinkedIn profile for Kyle Cheney. The profile header includes the LinkedIn logo, a search bar, and navigation icons for Home, My Network, Jobs, and Messaging. The profile picture is a circular headshot of a young man with short brown hair. To the right of the picture are buttons for 'Message' and 'More...'. Below the picture, the name 'Kyle Cheney · 3rd' is displayed, followed by his current role 'IE Forecasting & Planning Manager at UPS' and location 'Greater Atlanta Area · 500+ connections · Contact info'. To the right of this information is a logo for UPS and text indicating his education: 'University of Georgia - Terry College of Business'. The 'About' section describes him as an 'Intrinsically motivated individual with a degree in Management Information Systems and work experience encompassing ad-hoc analytics, multi-departmental and multi-country projects, and global collections process improvements with a genuine passion for people development... see more'. The 'Experience' section lists three roles: 'IE Forecasting & Planning Manager' at UPS (Jul 2019 - Present, 1 yr 2 mos), 'F&A Systems Supervisor' (Dec 2017 - Present, 2 yrs 9 mos), and 'Analytics and Performance Specialist' (Jan 2017 - Dec 2017, 12 mos).

Kyle Cheney · 3rd
IE Forecasting & Planning Manager at UPS
Greater Atlanta Area · 500+ connections · [Contact info](#)

About
Intrinsically motivated individual with a degree in Management Information Systems and work experience encompassing ad-hoc analytics, multi-departmental and multi-country projects, and global collections process improvements with a genuine passion for people development... see more

Experience

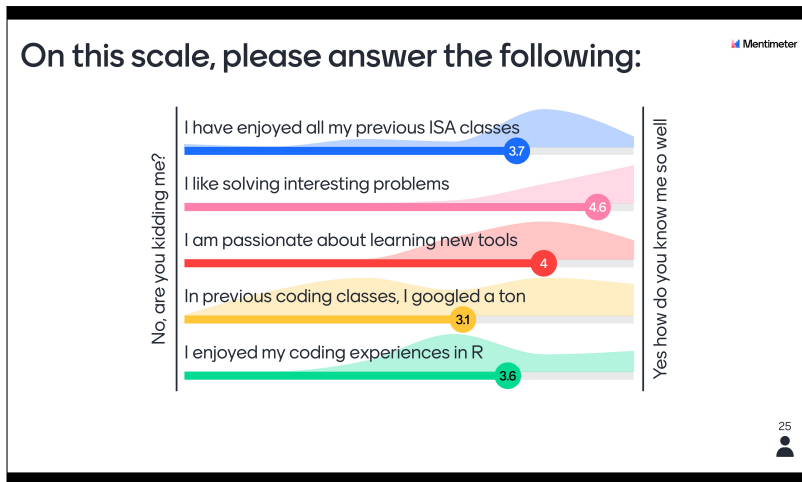
- UPS**
3 yrs 8 mos
 - IE Forecasting & Planning Manager**
Jul 2019 – Present · 1 yr 2 mos
Atlanta, Georgia
 - F&A Systems Supervisor**
Dec 2017 – Present · 2 yrs 9 mos
 - Analytics and Performance Specialist**
Jan 2017 – Dec 2017 · 12 mos

Quick Refresher based on Last Class

Main Learning Outcomes

- ✓ Describe course objectives & structure.
- ✓ Describe what do we mean by **forecasting** and explain the **PIVASE** framework.
- ✗ Explain the differences between cross sectional, time series, and panel datasets.
- ✗ Identify and describe the basic components of a time series including trends, seasonal components, and cycles.

Self-Assessed Skills from Last Class's Mentimeter Poll




Results from Last Class's Mentimeter Poll for Question A

Self-Described Goals for Taking the Class

What are you hoping to get out of ISA 444? 

required of my major	I would like to expand on the topics that were covered in 225. This class is also required for my major	I love any form of predictive modeling - looking for more methods/techniques to experiment with/expand on my own
A requirement for my minor but also a stronger understanding of forecasting	Required for my minor	Apply new skills to marketing and consumer research
Another professor recommended it based on my interests	Required for minor. Learn basics of forecasting for potential job use	Curious about the best methods for time series analysis

What are you hoping to get out of ISA 444? 

Forecasting is something extremely useful. Seemed more interesting than other ISA classes that were options for my minor	I am trying to develop the ability to make predictions and forecast the future as it will help with career opportunities	I chose this course as my ISA elective I enjoy forecasting more than database management.
Forecasting something with tech is cool!	I hope to increase my skills within R, but more importantly, I want to gain knowledge that will allow me to succeed in my career post-grad.	I'm someone who can't help but always look ahead and try to predict what will happen in my life. I have an innate interest in forecasting and this class was one that I was really looking forward to as I minor in analytics.

Learning Objectives for Today's Class

Main Learning Outcomes

- Explain the differences between cross sectional, time series, and panel datasets.
- Identify and describe the basic components of a time series including trends, seasonal components, and cycles.

Outline

- 1 Preface
- 2 Types of Data Over Time**
- 3 Components of a Time Series
- 4 Recap

Cross Sectional Data [1]

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

Example 1: H1B 2020 Data for Senior Data Scientists at Netflix¹

EMPLOYER	JOB TITLE	BASE SALARY	LOCATION
NETFLIX INC	SENIOR DATA SCIENTIST	375,000	LOS GATOS, CALIFORNIA
NETFLIX INC	SENIOR DATA SCIENTIST	400,000	LOS GATOS, CALIFORNIA
NETFLIX INC	SENIOR DATA SCIENTIST	420,000	LOS GATOS, CALIFORNIA
NETFLIX INC	SENIOR DATA SCIENTIST	420,000	LOS GATOS, CALIFORNIA
NETFLIX INC	SENIOR DATA SCIENTIST	450,000	LOS GATOS, CALIFORNIA
NETFLIX INC	SENIOR DATA SCIENTIST	600,000	LOS GATOS, CALIFORNIA

¹Data scraped from <https://h1bdata.info/index.php?em=NETFLIX+INC&job=SENIOR+DATA+SCIENTIST&city=LOS+GATOS&year=2020> on August 18, 2020 using the `rvest` package in R. The printing was limited to those individuals who started on/after January 01, 2020, with the filters specified in the URL.

Cross Sectional Data [2]

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

Example 2: Heart Disease Dataset²

age	sex	restingBP	maxHR	label
60.00	1.00	130.00	132.00	2
63.00	0.00	108.00	169.00	2
59.00	1.00	178.00	145.00	1
57.00	1.00	152.00	88.00	2
60.00	0.00	158.00	161.00	2
52.00	1.00	125.00	168.00	2
45.00	1.00	128.00	170.00	1
51.00	1.00	140.00	122.00	2
58.00	1.00	140.00	165.00	1
51.00	1.00	100.00	143.00	1
65.00	0.00	160.00	151.00	1
57.00	0.00	120.00	163.00	1
66.00	0.00	178.00	165.00	2

²Data sampled from [this UCI Machine Learning Repository](#).

Cross Sectional Data [3]

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

Example 3: NBA 2019-2020 Leaders - Top 12 in PTS/Game³

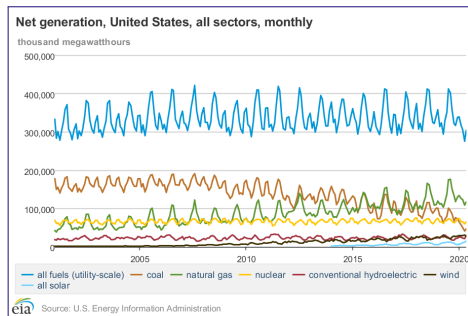
Player	Pos	Age	Tm	G	FG	FG%	eFG%	PTS
James Harden	SG	30	HOU	68	9.9	.444	.543	34.30
Bradley Beal	SG	26	WAS	57	10.4	.455	.520	30.50
Damian Lillard	PG	29	POR	66	9.5	.463	.563	30.00
Trae Young	PG	21	ATL	60	9.1	.437	.519	29.60
Giannis Antetokounmpo	PF	25	MIL	63	10.9	.553	.589	29.50
Luka Doncic	PG	20	DAL	61	9.5	.463	.531	28.80
Kyrie Irving	PG	27	BRK	20	10.0	.478	.546	27.40
Russell Westbrook	PG	31	HOU	57	10.6	.472	.493	27.20
Kawhi Leonard	SF	28	LAC	57	9.3	.470	.524	27.10
Devin Booker	SG	23	PHO	70	9.0	.489	.544	26.60
Karl-Anthony Towns	C	24	MIN	35	9.0	.508	.600	26.50
Anthony Davis	PF	26	LAL	62	8.9	.503	.536	26.10

³Data scraped from [Basketball-Reference](#) on August 18, 2020 using the [rvest](#) package in R. The printing with limited to the top 12 players and the selected variables.

Time Series Data [1]

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

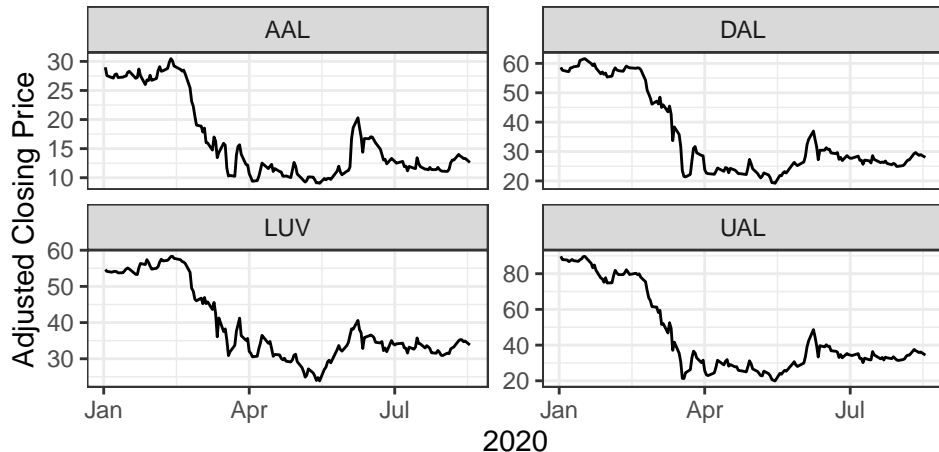
Example 1: Net Power Generation in the U.S.



Net monthly electricity generation in the U.S. from the U.S. Energy Information Administration.

Time Series Data [2]

Example 2: Stock prices of U.S. Airlines



As of August 18, 2020: COVID-19 had a substantial impact on these airline stocks.

Panel Data

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

Example: World Bank's Data⁴

iso3c	date	NY.GDP.MKTP.KD.ZG	SH.DYN.NMRT	SH.HIV.INCD.ZS	SH.MED.BEDS.ZS	SH.MED.PHYS.ZS	SH
CHN	2017.00	6.95	4.60			1.98	
CHN	2018.00	6.75	4.30			1.98	
CHN	2019.00	6.11	4.30			1.98	
EGY	2017.00	4.18	11.60	0.06	1.60	0.80	
EGY	2018.00	5.31	11.20	0.06		0.45	
EGY	2019.00	5.56	11.20	0.06		0.45	
USA	2017.00	2.22	3.60			2.61	
USA	2018.00	2.93	3.50			2.61	
USA	2019.00	2.33	3.50			2.61	

⁴Data queried from the [World Bank Data Catalog](#) using the [wbstats](#) package 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.

Outline

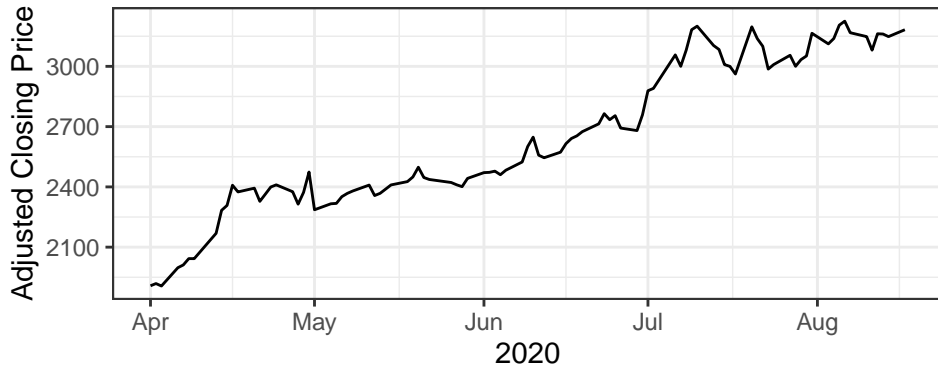
- 1 Preface
- 2 Types of Data Over Time
- 3 Components of a Time Series**
- 4 Recap

Trend [1]

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

Increasing Trend

The meteoric rise of \$AMZN from 2020-04-01 to 2020-08-18

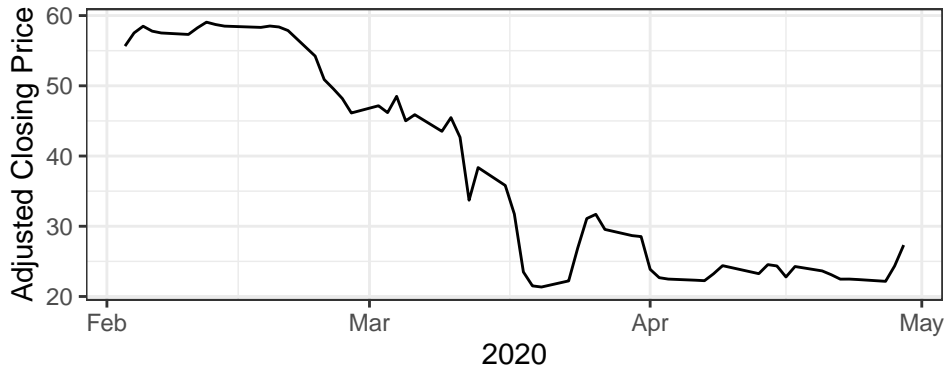


Trend [2]

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

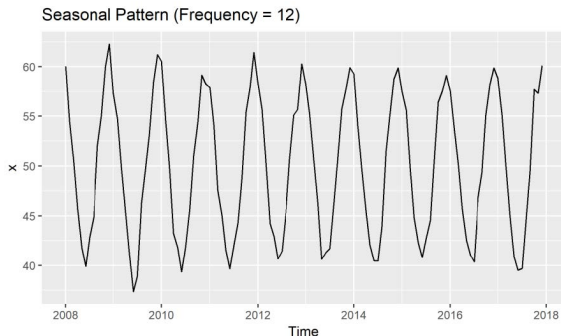
Decreasing Trend

The decline in \$DAL from 2020-02-03 to 2020-04-30



Seasonality [1]

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



A time series with a monthly seasonal pattern.⁵

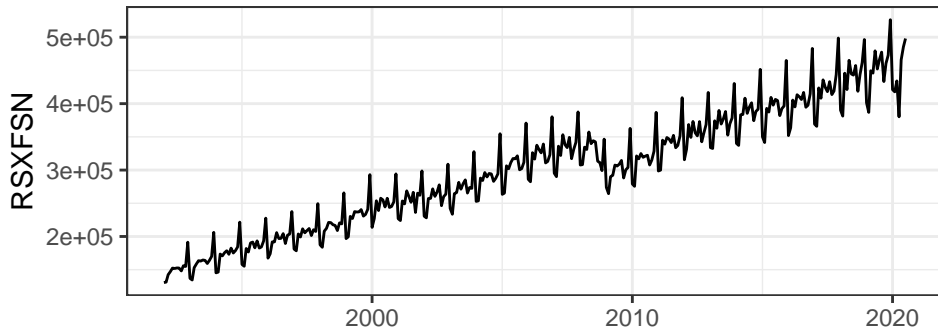
⁵Figure is from Dr. Allison Jones-Farmer's lecture notes, Miami University, Spring 2020.

Seasonality [2]

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

Seasonality with an Additive Trend

Retail (– Food Services) from 2010–01–01 to 2020–02–01

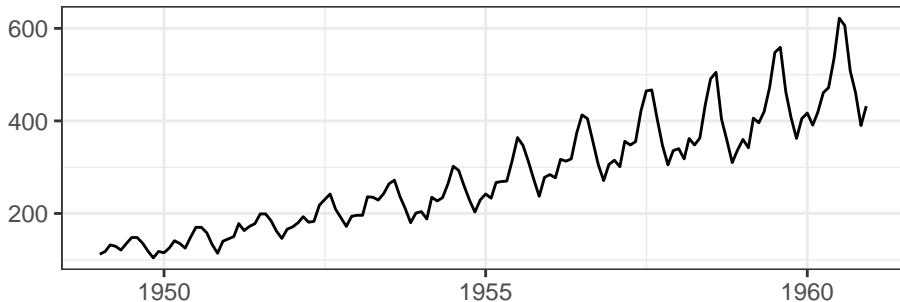


Data from FRED

Seasonality [3]

Seasonality with a Multiplicative Trend

Non-linear trend & seasonal component grows over time

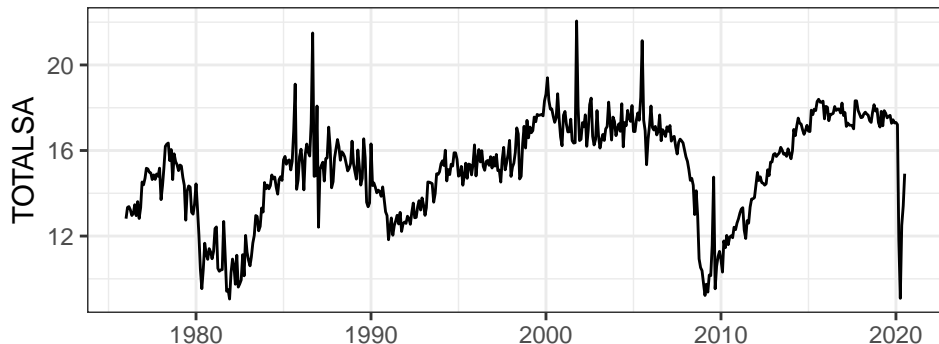


AirPassengers R Dataset -- Source: Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1976) Time Series Analysis, Forecasting and Control.

Cycle [1]

Cyclical fluctuations are somewhat irregular (unknown duration).

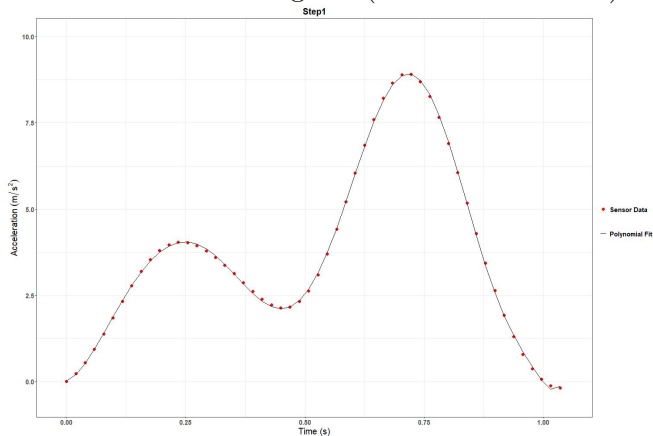
The cyclical nature of auto sales



Total Vehicle Sales [TOTALSA], retrieved from FRED, Federal Reserve Bank of St. Louis
<https://fred.stlouisfed.org/series/TOTALSA>, August 18, 2020.

Cycle [2]

Cyclical fluctuations are somewhat irregular (unknown duration).



A gait "cycle" from an IMU attached to the ankle. Joint work with the University at Buffalo.

Outline

- 1 Preface
- 2 Types of Data Over Time
- 3 Components of a Time Series
- 4 Recap**

Summary of Main Points

Main Learning Outcomes

- Explain the differences between cross sectional, time series, and panel datasets.
- Identify and describe the basic components of a time series including trends, seasonal components, and cycles.

Things to Do

- Thoroughly read Chapter 1 of our book, which can be downloaded from the [Publisher](#) (if you have not gotten your book yet).
- Go through the slides, examples and make sure you have a good understanding of what we have covered.
- Complete the two graded assignment (refer to next slides or Canvas) for more details.

Graded Assignment 01: Evaluating your Retention/Focus

Please go to [Canvas \(click here\)](#) and answer the four questions. **Due August 24, 2020 [2:51 PM, Ohio local time]**.

What/Why/Prep? The purpose of this assignment is to evaluate your understanding and retention of the material covered in Class 01. In order to prepare for this, you should have either actively attended class and/or watched the recording from WebEx.

General Guidelines:

- Individual assignment.
- This is a timed assignment (i.e. once you start the assignment you will have 25 minutes to complete 4 questions). The purpose of the time limit is to help you evaluate how much you have retained/understood from class. If the concepts we covered today are well-understood, this should take 10-15 minutes.
- Proctorio is NOT required for this assignment.
- You will need to have R installed (or accessible through the [Remote Desktop](#))

Graded Assignment 02: Evaluating your Understanding

Please go to [Canvas \(click here\)](#) and answer the four questions. **Due August 24, 2020 [2:51 PM, Ohio local time]**.

What/Why/Prep? The purpose of this assignment is to evaluate your understanding and retention of the material covered up to the End of Class 02. In order to prepare for this, you should have either actively attended class and/or watched the recording from WebEx. Furthermore, you should have thoroughly read [Chapter 01 of the book](#).

General Guidelines:

- Individual assignment.
- This is **NOT** a timed assignment. If the concepts we covered are well-understood, this should take 20-35 minutes.
- Proctorio is NOT required for this assignment.
- You will need to have R installed (or accessible through the [Remote Desktop](#))

ISA 444: Business Forecasting

02 - Introduction to Time Series Analysis

Fadel M. Megahed

Associate Professor
Department of Information Systems and Analytics
Farmer School of Business
Miami University
Email: fmegahed@miamioh.edu
Office Hours: [Click here to schedule an appointment](#)

Fall 2020