Machine Learning – Time Series Analysis

Week 12 – Part 1 – Introduction to Time Series Analysis

CS 457 - L1 Data Science

Zeehasham Rasheed

Lecture Objectives

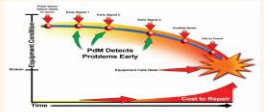
- Learn Basic Time Series Forecasting Process
- Methods
- Evaluation
- Real World Use Cases
- Discussion

Use Cases

retail sales forecasting

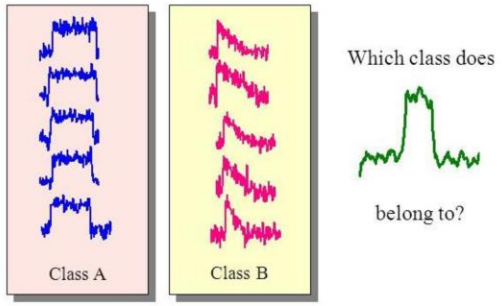
financial forecasting

web traffic forecasting

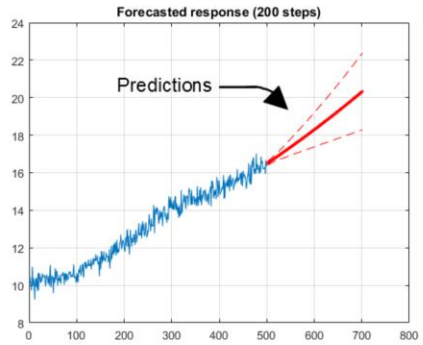
Type of Data	Use Case	
Financial Data	Analyze stock performance	
Census Data	Analyze population types and growth	
Sales Data	Analyze sales by region, product, etc.	
Industrial Data	Analyze machine performance	

Data Science / Machine Learning

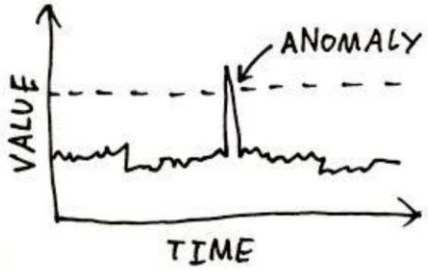
Classification



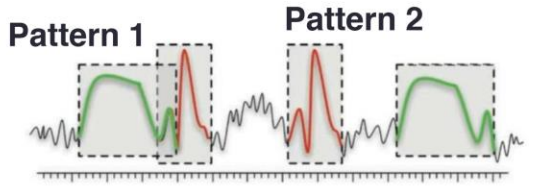
Prediction



Anomaly Detection



Pattern Discovery



Time Series Objectives

- 1. Compact description of data.
- 2. Interpretation.
- 3. Forecasting.

Forecasting Process

- Look at the data (Patterns in a graph)
- Forecast (choose one or more methods)
- Evaluate (examine errors)

Types of Forecasting

- Qualitative methods – judgmental methods
  - Forecasts generated subjectively by the forecaster
  - Educated guesses

- Quantitative methods:
  - Forecasts generated through mathematical modeling
  - More Data Driven (Data Science)

Qualitative Methods

Type	Characteristics	Strengths	Weaknesses
Executive opinion	A group of managers meet & come up with a forecast	Good for strategic or new-product forecasting	One person's opinion can dominate the forecast
Market research	Uses surveys & interviews to identify customer preferences	Good determinant of customer preferences	It can be difficult to develop a good questionnaire
Delphi method	Seeks to develop a consensus among a group of experts	Excellent for forecasting long-term product demand, technological	Time consuming to develop

Source: Wiley

Quantitative Methods

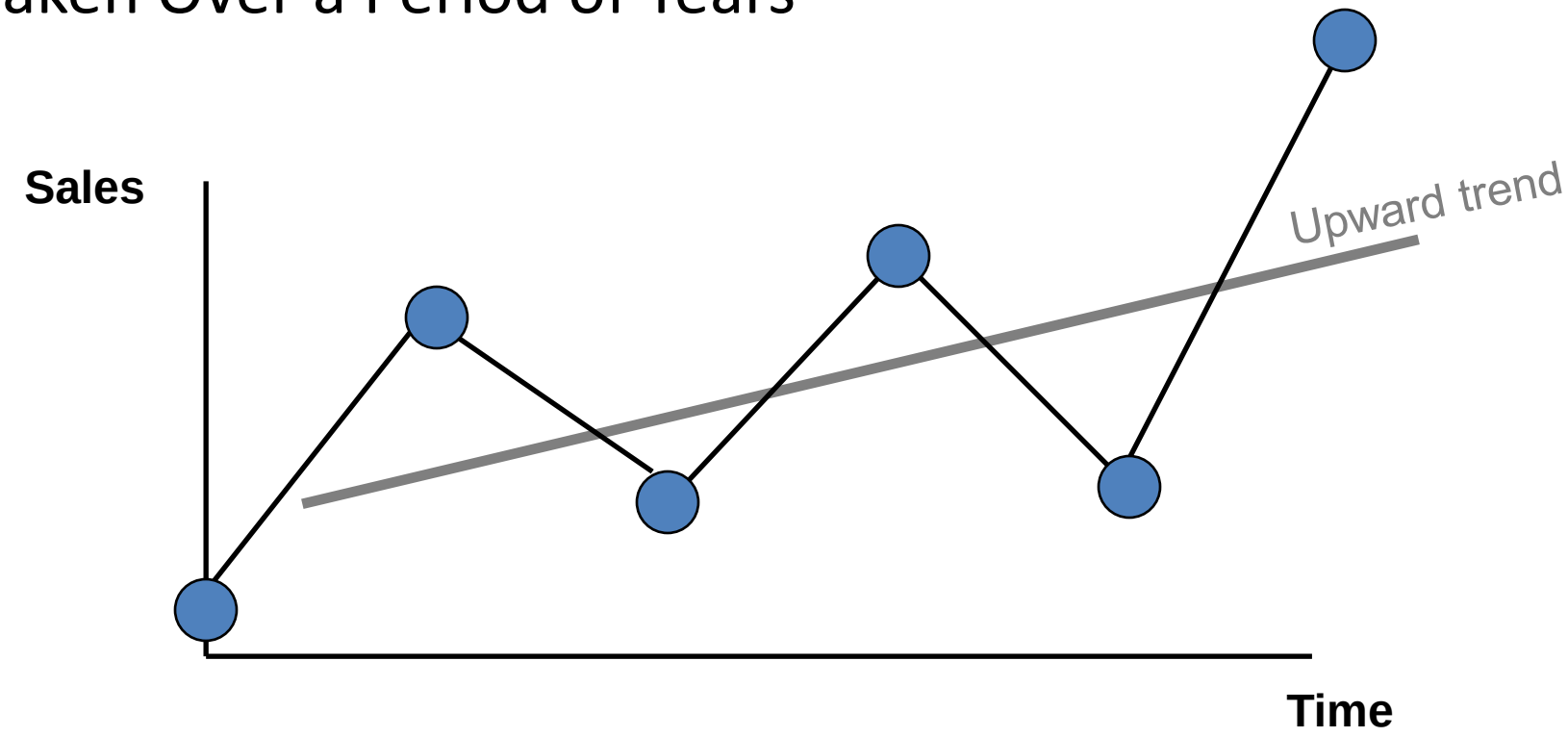
- Time Series Models:
  - Assumes information needed to generate a forecast is contained in a time series of data
  - Assumes the future will follow same patterns as the past

Time Series Models

- Forecaster looks for data patterns as
  - Data = historic pattern + random variation
- Historic pattern to be forecasted:
  - Trend – data exhibits an increasing or decreasing pattern
  - Seasonality – any pattern that regularly repeats itself and is of a constant length
  - Cycle – patterns created by economic fluctuations
- Random Variation cannot be predicted

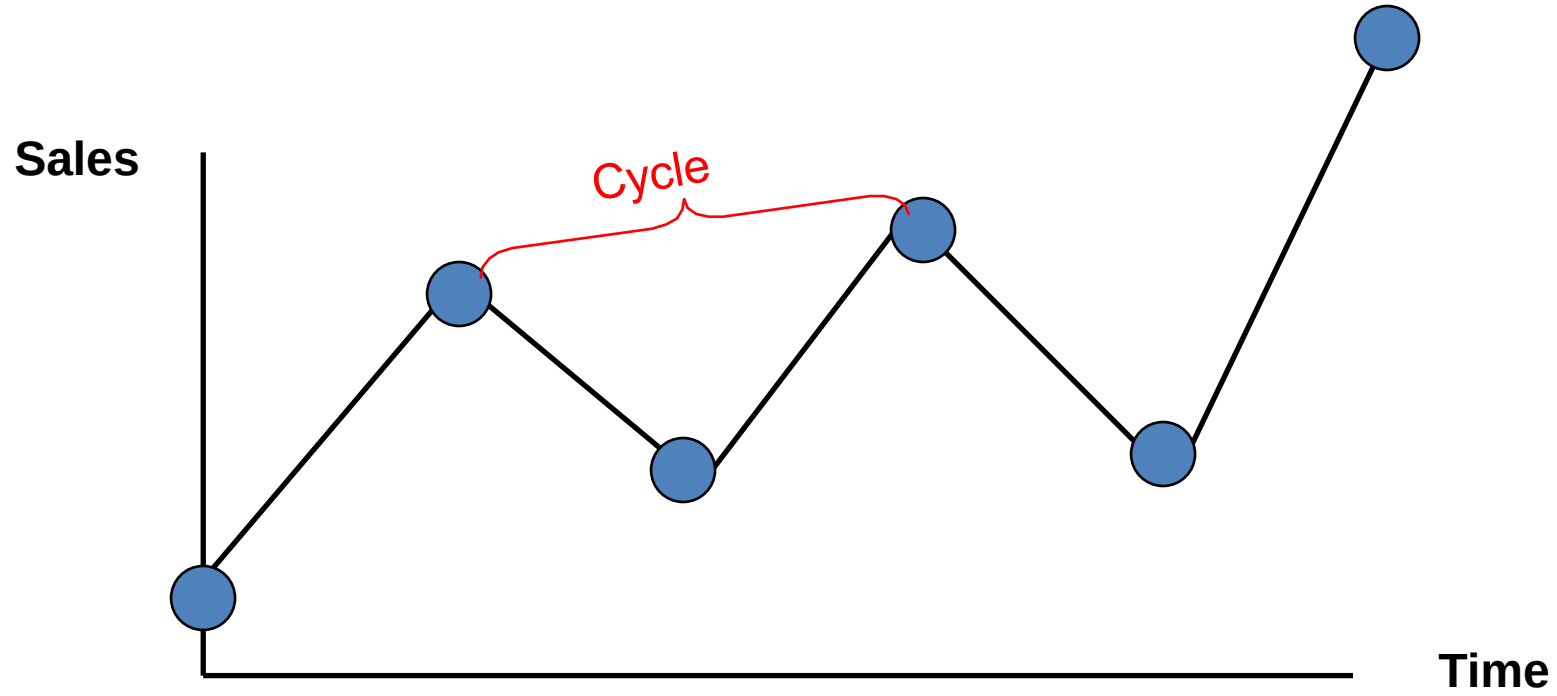
Patterns in Data (1)

- Trend Component
  - Overall Upward or Downward Movement
  - Data Taken Over a Period of Years



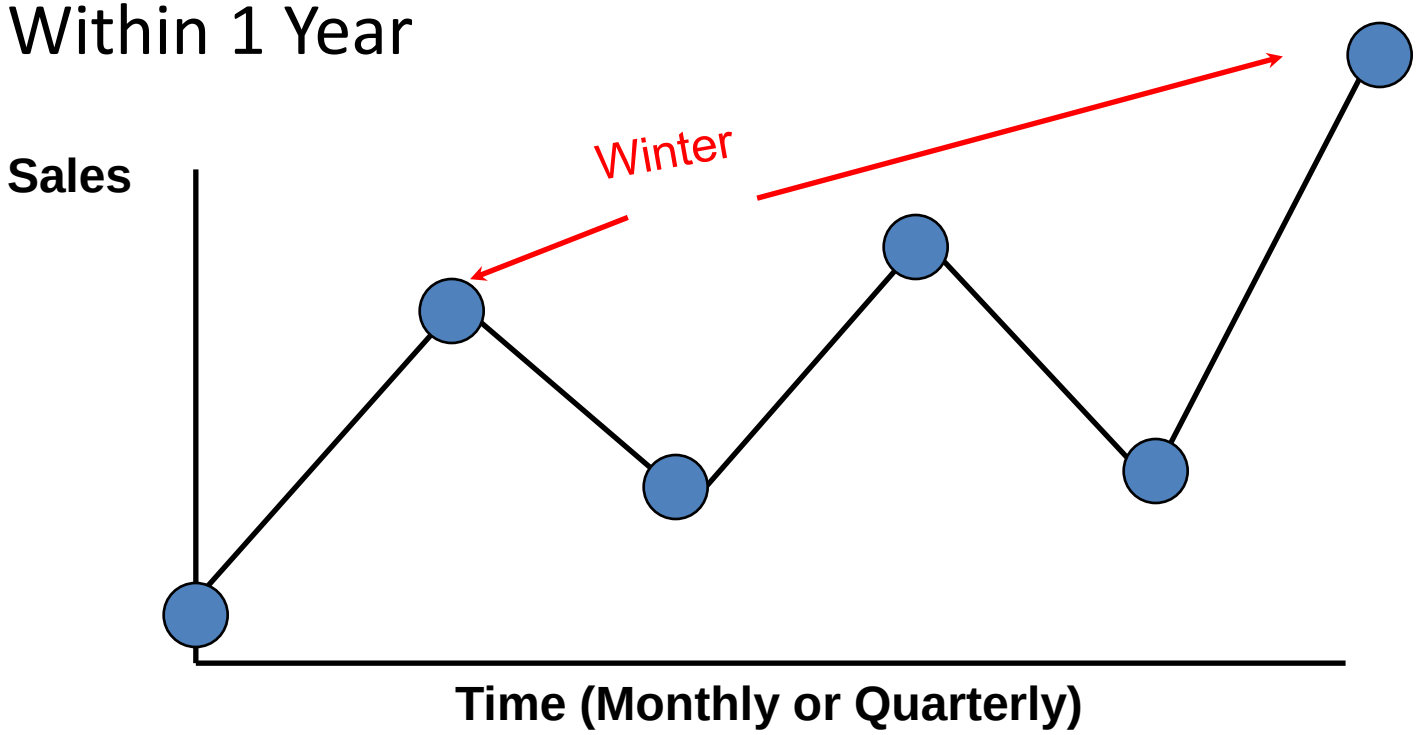
Patterns in Data (2)

- Cyclic Component
  - Upward or Downward Swings
  - May Vary in Length



Patterns in Data (3)

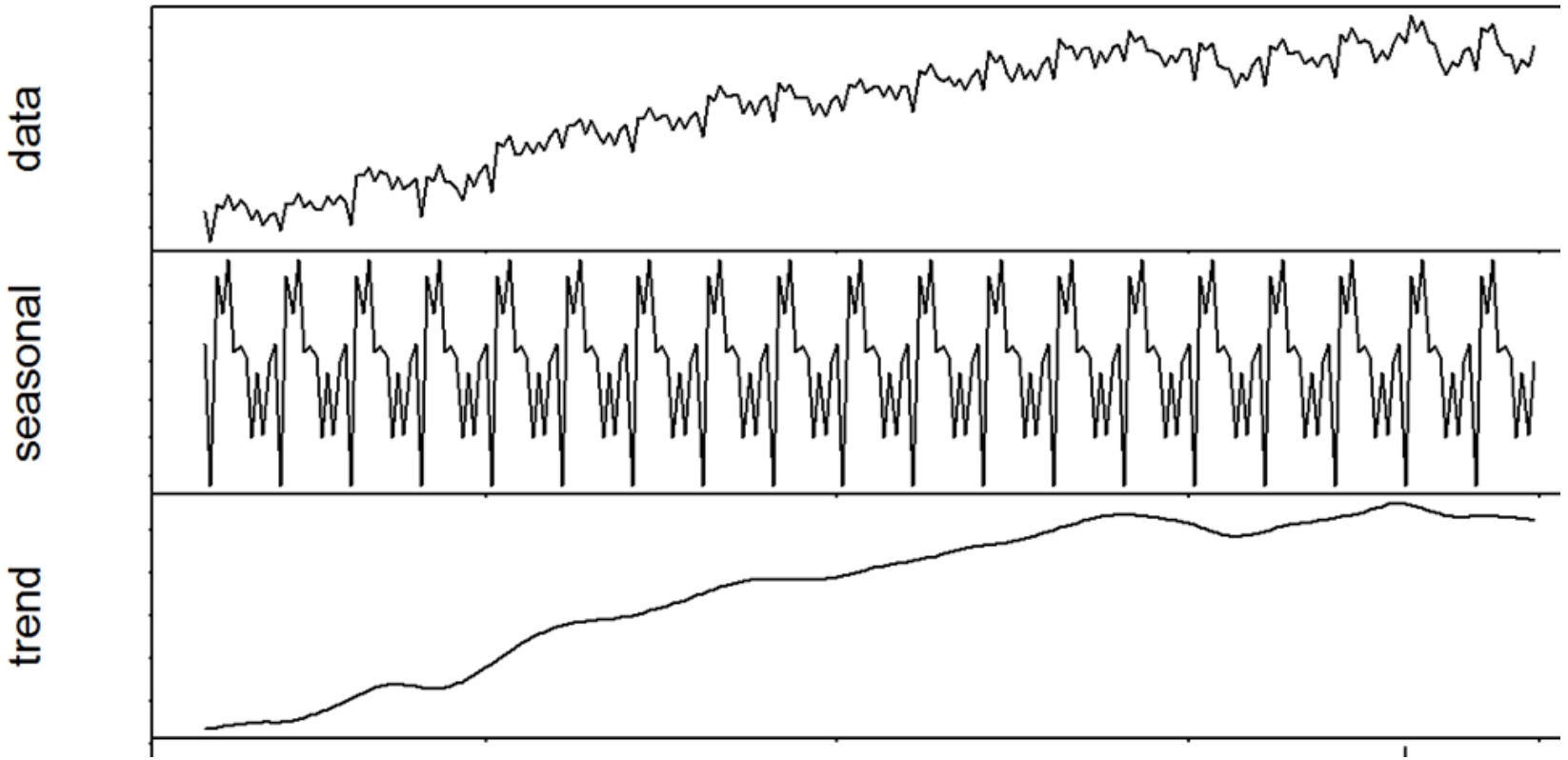
- Seasonal Component
  - Upward or Downward Swings
  - Regular Patterns
  - Observed Within 1 Year



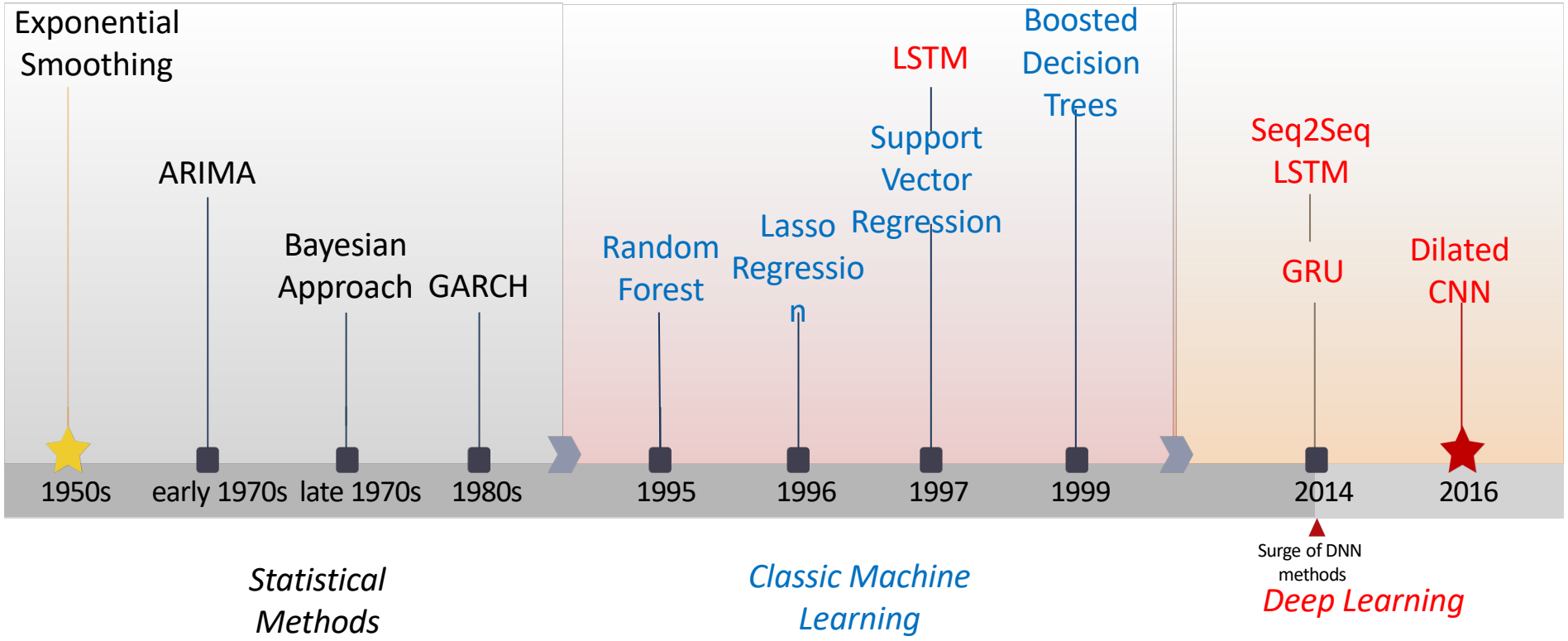
Patterns in Data (4)

- Random Irregular Component
  - Erratic, Nonsystematic, Random, ‘Residual’ Fluctuations
  - Due to Random Variations of
    - Nature
    - Accidents
  - Short Duration and Non-repeating

Pattern Graph



Methods/Techniques/Algorithms



J. Gooijer and R. Hyndman. 25 Years of Time Series Forecasting.

End of Part 1

Machine Learning – Time Series Analysis

Week 12 – Part 2 – Basic Time Series Methods

CS 457 - L1 Data Science

Zeehasham Rasheed

Basic Methods

- Statistical Methods
  - Moving Averages
  - Exponential Smoothing
- Classic Machine Learning
  - Least Square Trend Fitting (Regression)

Talk About Evaluation First

- Mean Absolute Deviation (MAD)
  - measures the total error in a forecast without regard to sign
- Cumulative Forecast Error (CFE)
  - Measures any bias in the forecast
- Mean Square Error (MSE)
  - Penalizes larger errors

$$\mathbf{MAD} = \frac{\sum |\mathbf{actual} - \mathbf{forecast}|}{\mathbf{n}}$$

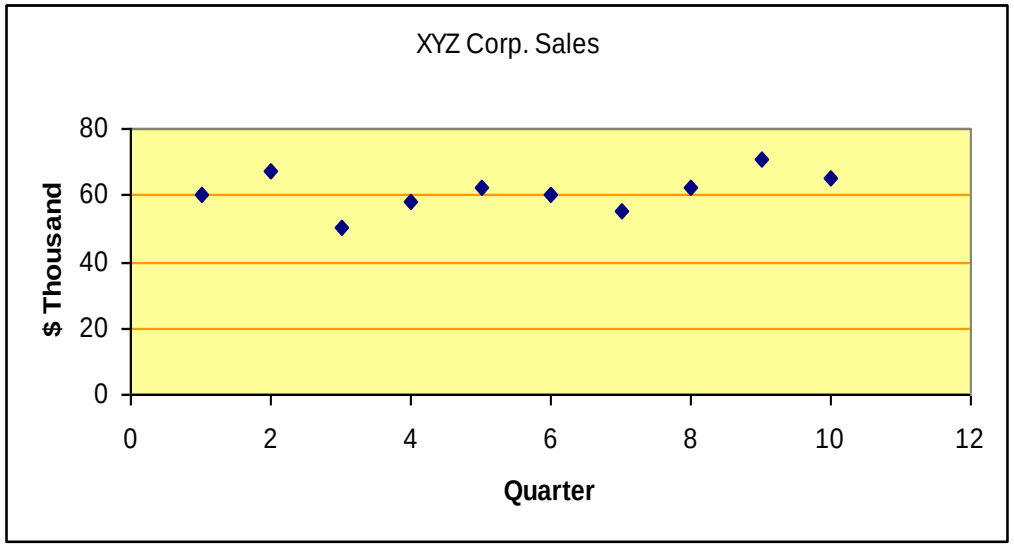
$$\mathbf{CFE} = \sum (\mathbf{actual} - \mathbf{forecast})$$

$$\mathbf{MSE} = \frac{\sum (\mathbf{actual} - \mathbf{forecast})^2}{\mathbf{n}}$$

Problem Statement

- Consider the following sales data for 10 time periods (quarters)

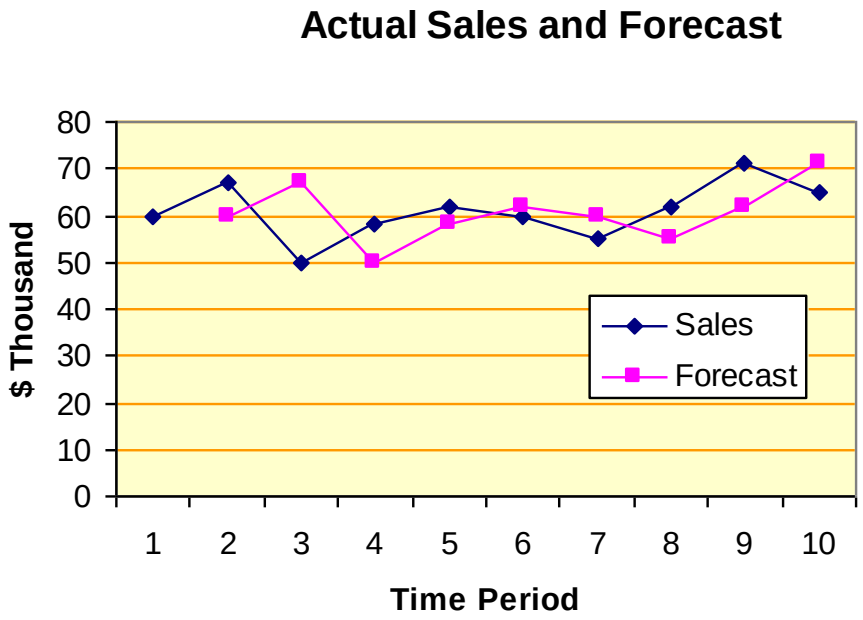
Period	Sales
1	60
2	67
3	50
4	58
5	62
6	60
7	55
8	62
9	71
10	65



What is a good forecast for Sales for the next period?

Statistical Methods: Naive Forecast

Period	Sales	Naive Forecast
1	60	N/A
2	67	60
3	50	67
4	58	50
5	62	58
6	60	62
7	55	60
8	62	55
9	71	62
10	65	71
11		**65**



$$\hat{y}_t = y_{t-1}$$

How good is this forecast?

Evaluating the Forecast

X	Y	Naive		Abs	Percent	Squared
Period	Sales	Forecast	Error	Error	Error	Error
1	60					
2	67	60	7	7	10.45%	49.0
3	50	67	-17	17	34.00%	289.0
4	58	50	8	8	13.79%	64.0
5	62	58	4	4	6.45%	16.0
6	60	62	-2	2	3.33%	4.0
7	55	60	-5	5	9.09%	25.0
8	62	55	7	7	11.29%	49.0
9	71	62	9	9	12.68%	81.0
10	65	71	-6	6	9.23%	36.0
11		**65**	**0.56**	**7.22**	**12.26%**	**68.1**
			BIAS	**MAD**	**MAPE**	**MSE**
		Standard Error (Square Root of MSE) =				**8.3**

Error = $y - \hat{y}$

Bias = Avg (Errors)
MAD = Avg (Abs Errors)
MAPE = Avg (Percent Errors)
MSE = Avg (Squared Errors)

Moving Averages

Period	Sales	Moving Avg. Forecast	Error	Abs. Error	Percent Error	Squared Error
1	60	N/A				
2	67	N/A				
3	50	N/A				
4	58	59.0	-1.0	1.0	1.72%	1.0
5	62	58.3	3.7	3.7	5.91%	13.4
6	60	56.7	3.3	3.3	5.56%	11.1
7	55	60.0	-5.0	5.0	9.09%	25.0
8	62	59.0	3.0	3.0	4.84%	9.0
9	71	59.0	12.0	12.0	16.90%	144.0
10	65	62.7	2.3	2.3	3.59%	5.4
11		66.0	**2.62**	**4.33**	**6.80%**	**29.86**
			BIAS	**MAD**	**MAPE**	**MSE**
		Standard Error (Square Root of MSE) =				**5.5**

How does this *3-period moving average* forecast compare to the Naive forecast?

Moving Average can be
2-period
3-period
4-period and so on

Exponential Smoothing

- Most frequently used time series method because of ease of use and minimal amount of data needed
- Need just three pieces of data to start:
  - Last period's forecast (F_t) (or average)
  - Last periods actual value (A_t)
  - Select value of smoothing coefficient
    - between 0 and 1.0
    - Higher values (e.g. 0.7 or .8) may place too much weight on last period's random variation

$$F_{t+1} = \alpha A_t + (1 - \alpha) F_t$$

Simple Exponential Smoothing

alpha= 0.3

Period	Sales	Exponential Smoothing	Error	Abs. Error	Percent Error	Squared Error
1	60	N/A				
2	67	60.0	7.0	7.0	10.45%	49.0
3	50	62.1	-12.1	12.1	24.20%	146.4
4	58	58.5	-0.5	0.5	0.81%	0.2
5	62	58.3	3.7	3.7	5.92%	13.5
6	60	59.4	0.6	0.6	0.95%	0.3
7	55	59.6	-4.6	4.6	8.37%	21.2
8	62	58.2	3.8	3.8	6.10%	14.3
9	71	59.4	11.6	11.6	16.40%	135.6
10	65	62.8	2.2	2.2	3.31%	4.6
		63.5	**1.29**	**5.11**	**8.50%**	**42.79**
			BIAS	**MAD**	**MAPE**	**MSE**

Standard Error (Square Root of MSE) = 6.5

Brain Hack 1

- **Forecasting trend problem:** a company uses exponential smoothing with trend to forecast usage of its lawn care products. At the end of July the company wishes to forecast sales for August. July demand was 62. The trend through June has been 15 additional gallons of product sold per month. Average sales have been 57 gallons per month. The company uses alpha 0.2
- Forecast the sales for August.

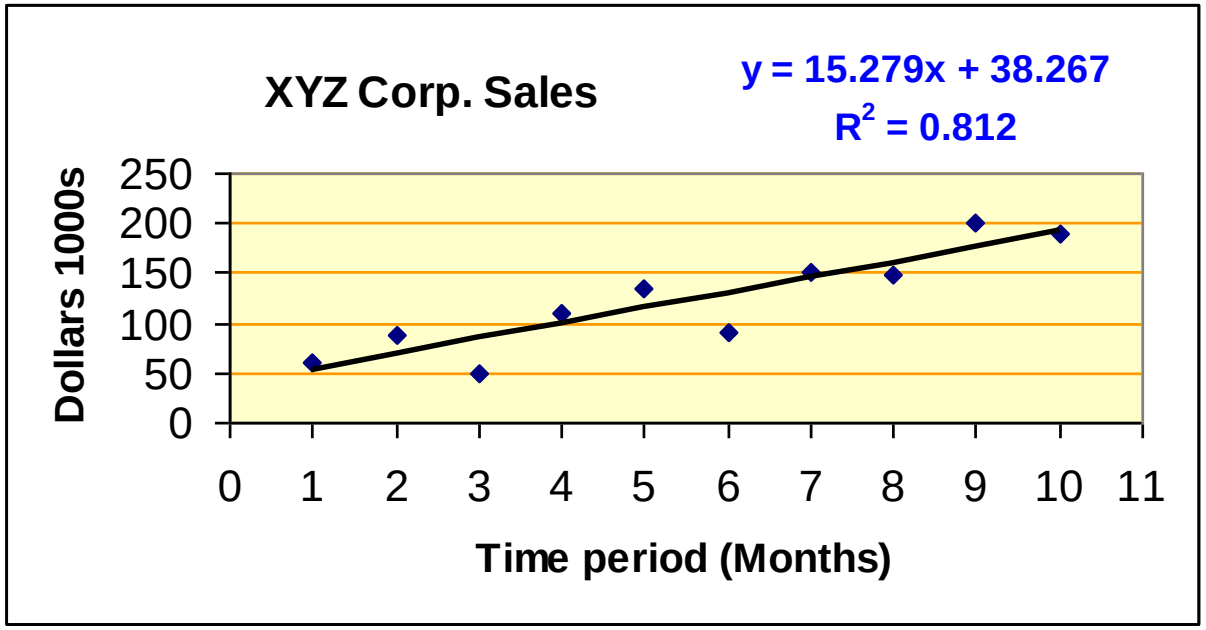
Brain Hack 1 (Answer)

- **Forecasting trend problem:** a company uses exponential smoothing with trend to forecast usage of its lawn care products. At the end of July the company wishes to forecast sales for August. July demand was 62. The trend through June has been 15 additional gallons of product sold per month. Average sales have been 57 gallons per month. The company uses alpha 0.2
- Forecast the sales for August.

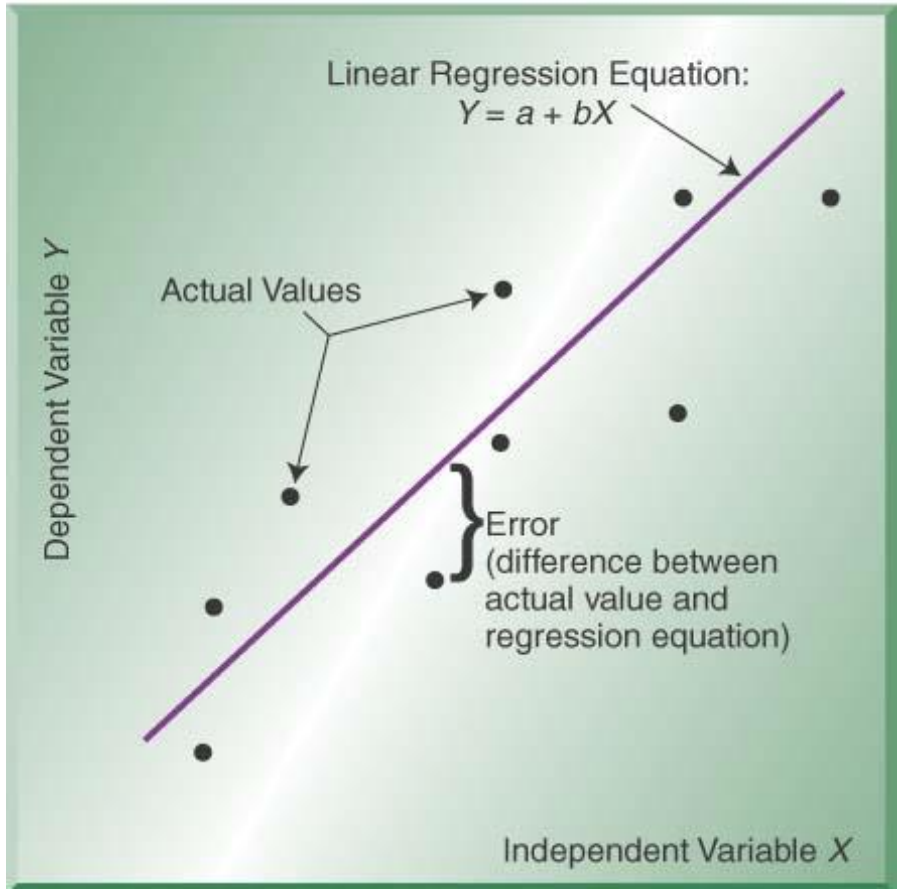
$$S_{August} = \alpha A_t + (1 - \alpha)(S_{t-1} + T_{t-1}) = (0.2)(62) + (0.8)(57 + 15) = 70$$

Classic ML Methods: Regression

- Finding an equation of a trend (line in case of Linear Regression) that fits the data
  - y is the dollar value and x is the month
    - Using this equation, we can predict dollar value at any given time



Regression Steps



- Identify dependent (y) and independent (x) variables
- Solve for the slope of the line

$$\mathbf{b} = \frac{\sum \mathbf{XY} - \mathbf{n}\overline{\mathbf{X}}\,\overline{\mathbf{Y}}}{\sum \mathbf{X}^2 - \mathbf{n}\overline{\mathbf{X}}^2}$$

- Solve for the y intercept

$$\mathbf{a} = \overline{\mathbf{Y}} - \mathbf{b}\overline{\mathbf{X}}$$

- Develop your equation for the trend line
- **Y = a + bX**

Brain Hack 2

- **Linear Regression Problem:** A maker of golf shirts has been tracking the relationship between sales and advertising dollars. Use linear regression to find out what sales might be if the company invested $53,000 in advertising next year.
- Find the expected sale for the upcoming year

	Sales $ (Y)	Adv.$ (X)	XY	X^2	Y^2
1	130	32	4160	2304	16,900
2	151	52	7852	2704	22,801
3	150	50	7500	2500	22,500
4	158	55	8690	3025	24964
5	?	53			
Tot					
Avg					

Brain Hack 2 (Answer)

- **Linear Regression Problem:** A maker of golf shirts has been tracking the relationship between sales and advertising dollars. Use linear regression to find out what sales might be if the company invested $53,000 in advertising next year.
- Find the expected sale for the upcoming year

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$b = \frac{28202 - 4(47.25)(147.25)}{9253 - 4(47.25)^2} = 1.15$$

$$a = \bar{Y} - b\bar{X} = 147.25 - 1.15(47.25)$$

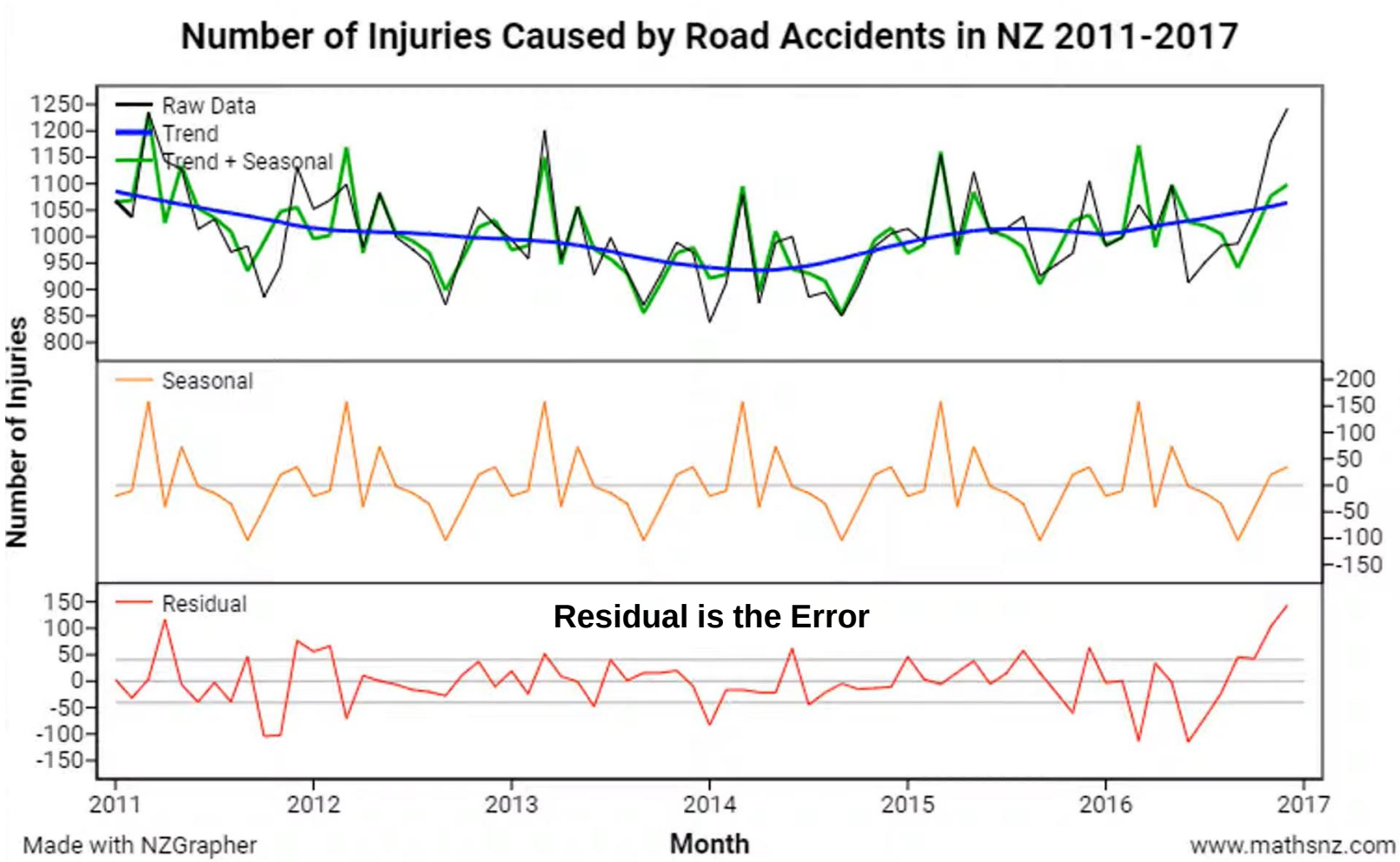
$$a = 92.9$$

$$Y = a + bX = 92.9 + 1.15X$$

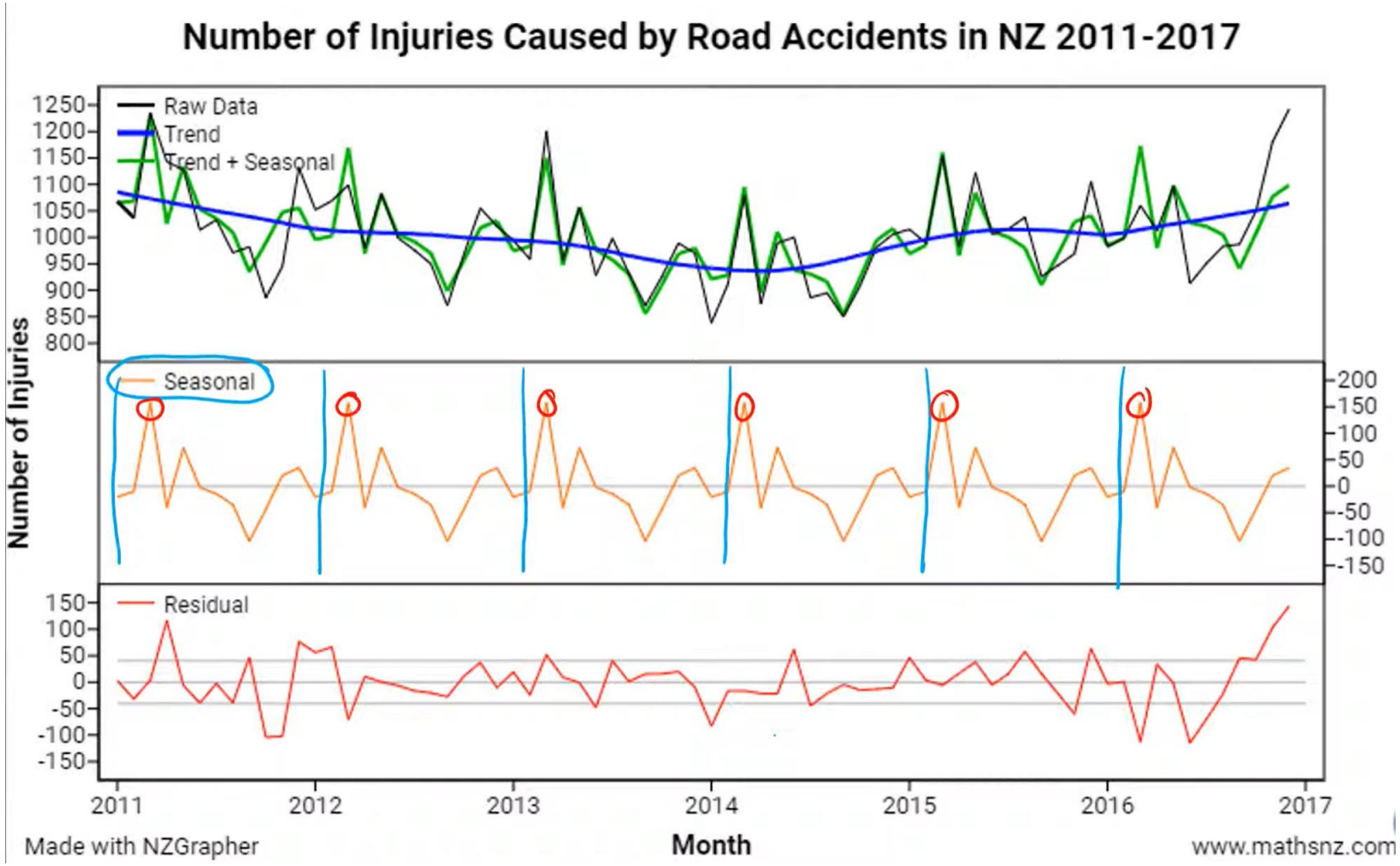
$$Y = 92.9 + 1.15(53) = 153.85$$

	Sales $ (Y)	Adv.$ (X)	XY	X^2	Y^2
1	130	32	4160	2304	16,900
2	151	52	7852	2704	22,801
3	150	50	7500	2500	22,500
4	158	55	8690	3025	24964
5	**153.85**	53			
Tot	589	189	28202	9253	87165
Avg	**147.25**	**47.25**			

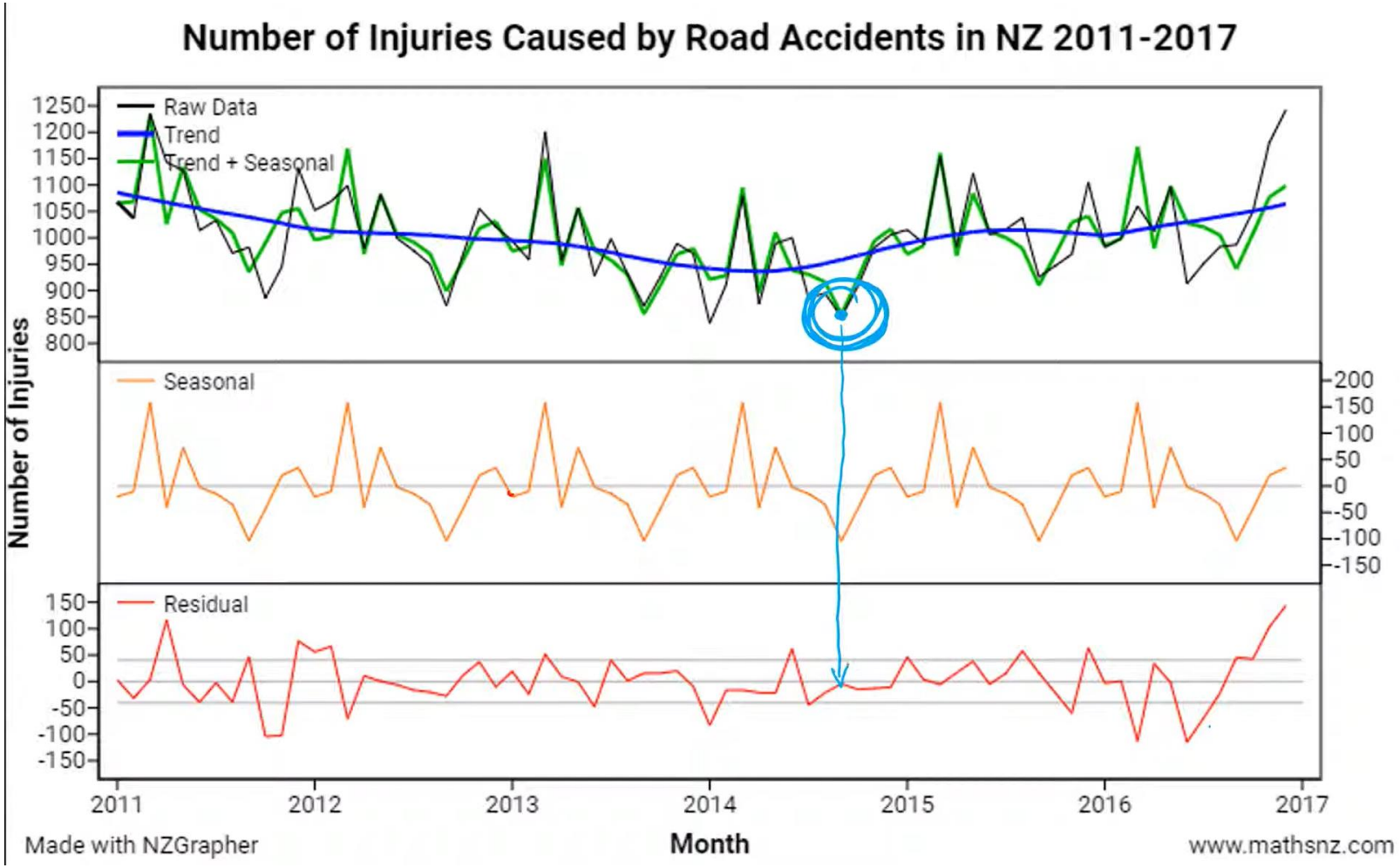
Interpretation Process - Trend



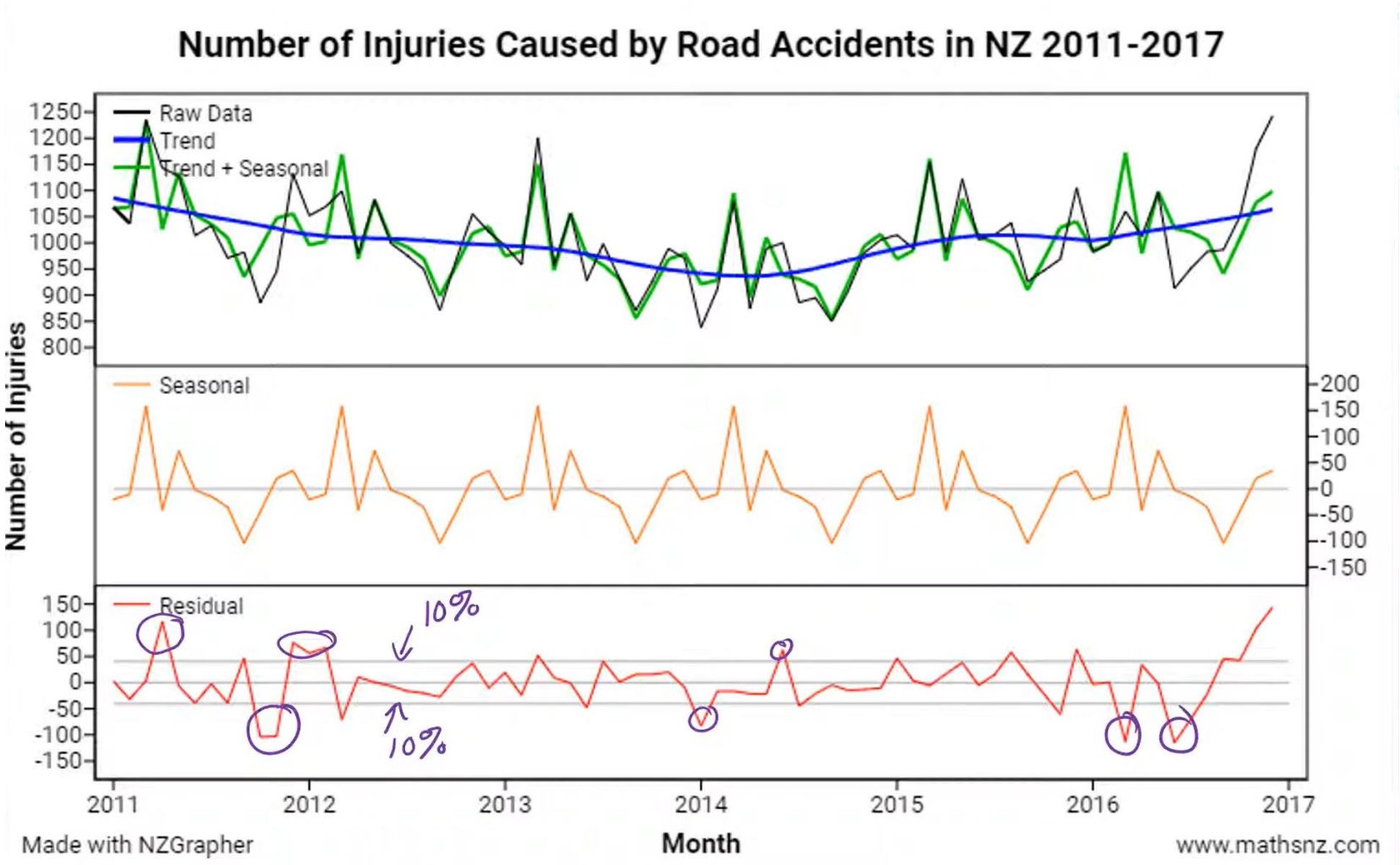
Interpretation Process - Seasonality



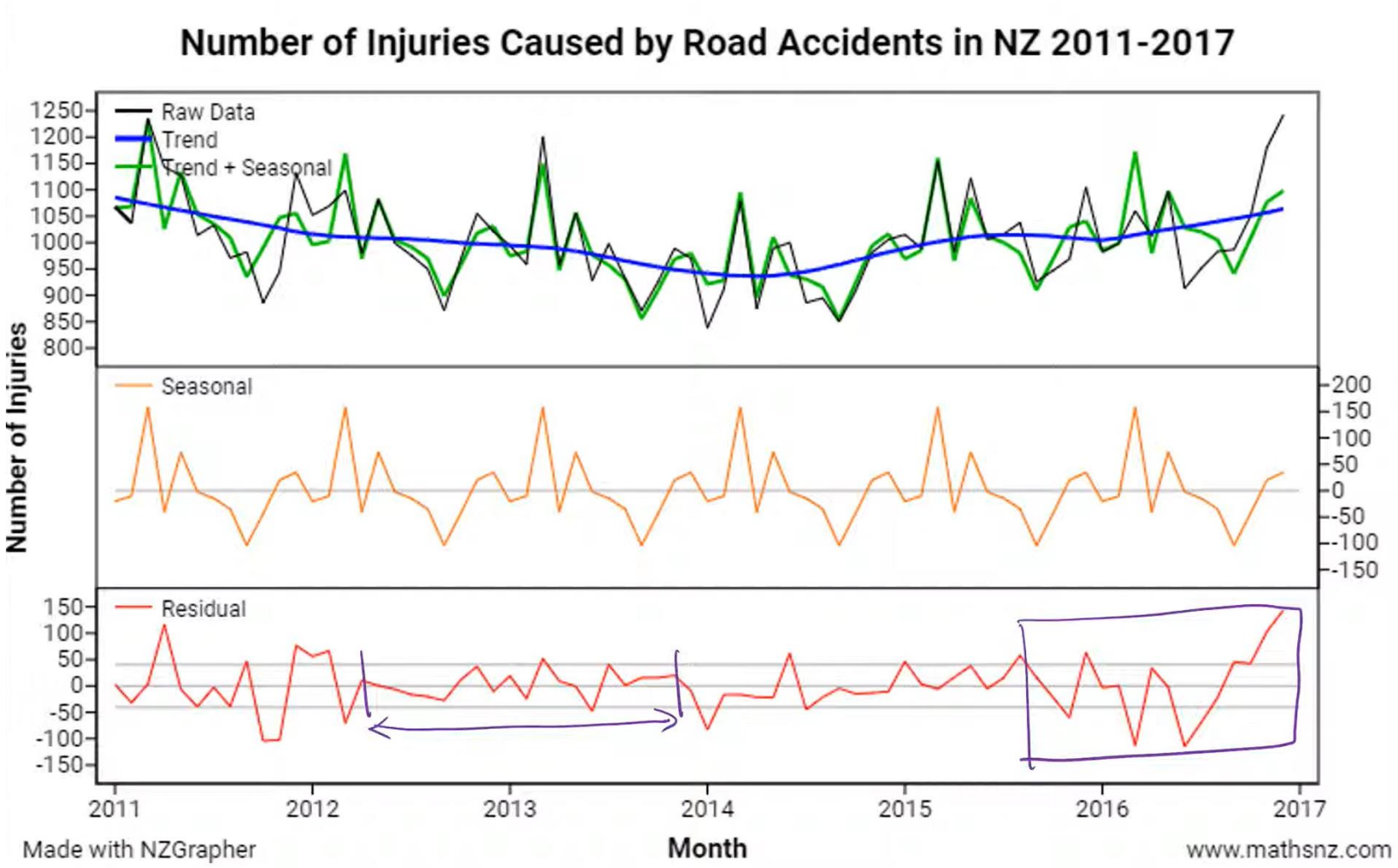
Interpretation Process - Accuracy



Interpretation Process – Unusual Values



Interpretation Process – Model Stability



End of Part 2

Machine Learning – Time Series Analysis

Week 12 – Part 3 – Time Series Analysis
Use Cases

CS 457 - L1 Data Science

Zeehasham Rasheed

Real World Use Cases

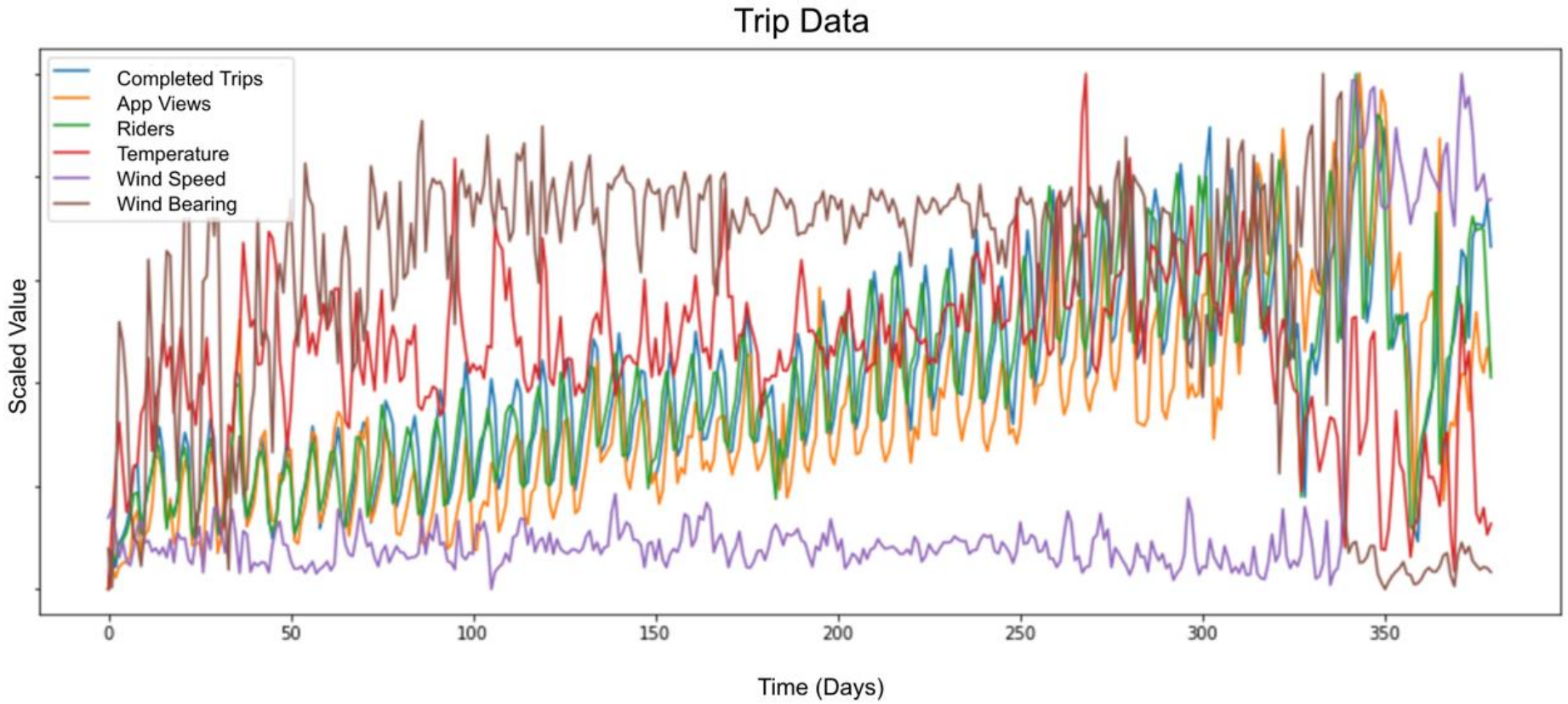
- Uber / Bykea
  - Predict Number of Trips per Day
  - Predict the number of riders at given time and location
- Verizon
  - Telecommunication company
  - Predict number of circuits/routers/switches/devices needed for each customer
  - Predict the maintenance cost for those devices
- Predict = Forecast

Uber and Bykea

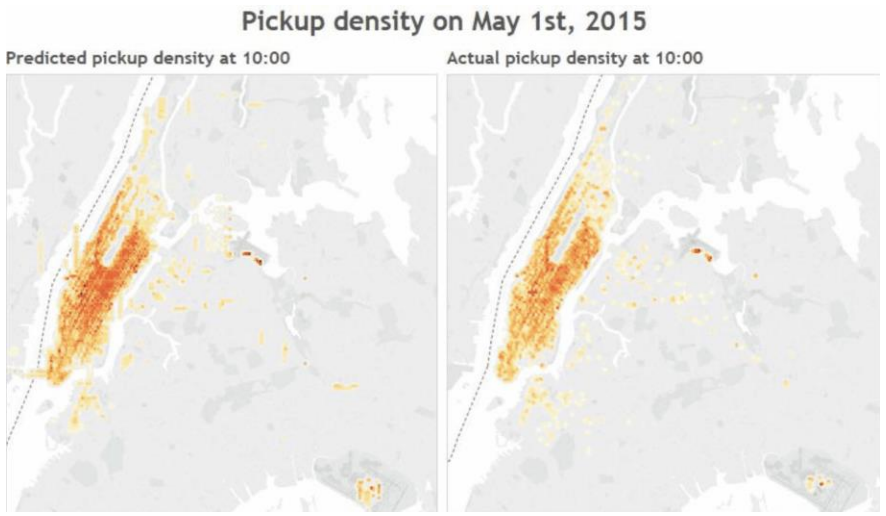
- Forecast number of drivers needed in advance so that they can schedule riders on specific time and location
  - Long wait time, customer quits the ride
- What makes forecasting (at Uber) challenging?
  - The Uber platform operates in the real, physical world
    - physical constraints, and unpredictability
    - Physical constraints like geographic distance and road throughput
    - Unpredictable weather conditions
    - Using spatio-temporal features instead of just temporal

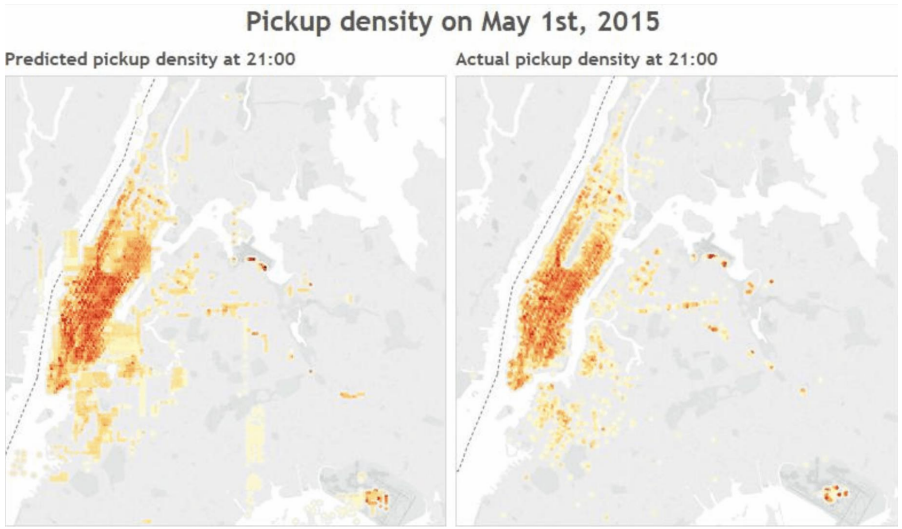
Uber Forecasting Model for Trips

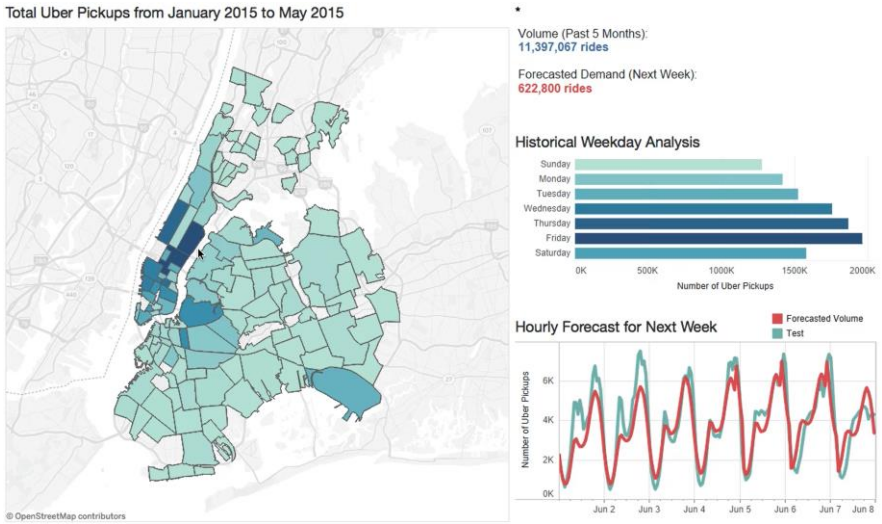
- Model was trained using a combination of exogenous variables, including weather (e.g., precipitation, wind speed, and temperature forecasts) and city-level information (e.g., trips in progress at any given time within a specific geographic area, registered Uber users, and local holidays or events



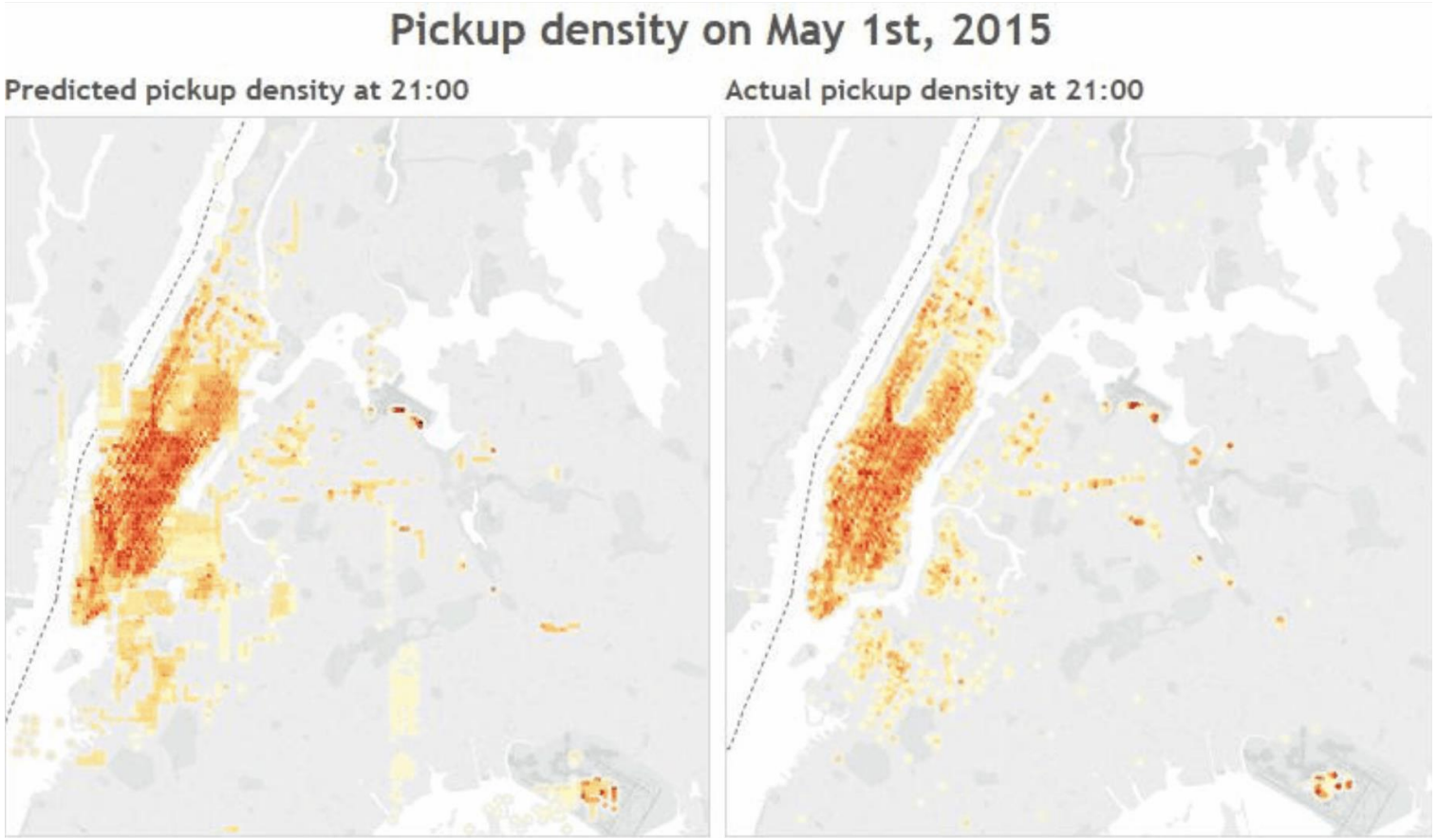
Uber NYC Pickup Prediction







http://sdaulton.github.io/TaxiPrediction/


Pickup density on May 1st, 2015
Predicted pickup density at 21:00
Actual pickup density at 21:00

Verizon

- Forecast circuit counts and circuit cost over a period of time
  - Company planning and budgeting for next years
  - Variables used (X)
    - Date (extra processing to get month, week and quarter of the year)
    - Product Type
    - Geo Location
    - Line of Business
    - Circuit Types
    - Foot Print (franchise/out-of-franchise)

Verizon (2)

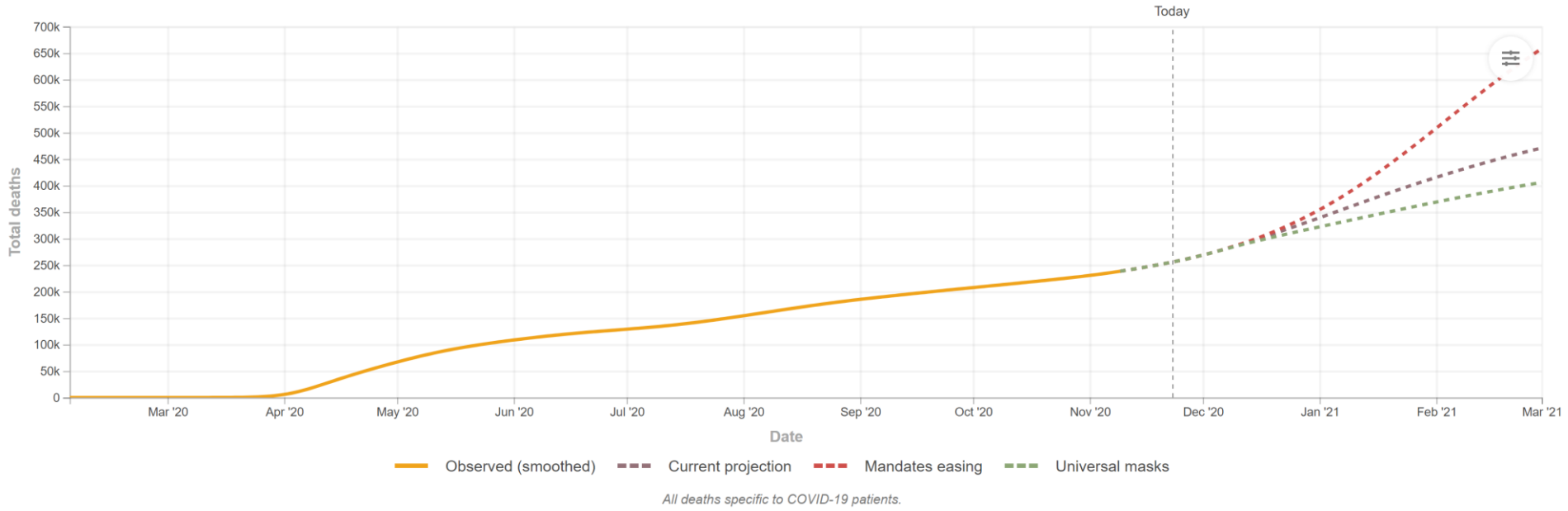
- Two models are created
  - Circuit Count (Y)
  - Circuit Cost (Y)
- Machine Learning Method
  - Random Forest Regressor
- Mean Square Error (converted to accuracy)
  - 87% accuracy for predicting circuit count
  - 83% accuracy for predicting circuit cost

Verizon (3)

- Variable Importance
  - Used other statistical measures (entropy, p-value etc.) to find most important variables contributing to these predictions
    - Foot-print, circuit type and geo location are top 3 input variables
- Benefit
  - Company can use these predictions for budgeting and improvements

Last but not Least: COVID 19

- United States COVID 19 Forecast
  - Current Projection
  - Projection with no lock-down
  - Projection with mask on



Review: TS Forecasting Steps

- Decide what needs to be forecast
  - Level of detail and time period
- Evaluate and analyze appropriate data
  - Identify needed data and whether it is available
- Select and test the forecasting model
  - Cost, ease of use and accuracy
- Generate the forecast
- Monitor forecast accuracy over time

End of Part 3