## Assignment Project Exam Help

https://powcoder.com

Forecasting MeChat powcoder



# Agenda

Start	End	Item
		What is Forecasting?
		Naïve Forecasting Methods
		Time Series Decomposition
	As	signments Projectin Exam Help
		Appendix https://powcoder.com
		mtps.//poweoder.com
		Add WeChat powcoder



#### Main ideas

- Forecast future values of a time series
- Distinction between forecasting (main focus) and describing/explaining
  ASSIGNMent Project Exam Help

  • Before forecasting understand "periodicity"
- Four component to the following of the following the fol
  - Level
  - Add WeChat powcoder Trend
  - Seasonality
  - Noise
- Time series data is great for enrichment and engineering e.g. "event flags" but can be modeled as a standalone vector due to the meta data "inside" the vector.

#### Difference between ML Data Setup & Time Series Data

#### **ML Data Frame**

dataID	HHuniqueID	Communication	LastContactDay	LastContactMonth	NoOfContacts	DaysPassed	PrevAttempts	Outcome	CallStart	CallEnd	Y_AccetpedOffer
1	HHd4d0af8c72	telephone	28	jan	2	-1	0	NA	13:45:20	13:46:30	0
2	HH8d3e87c164	NA A	26	may 1 Dr	0100±5		a Llai		14:49:03	14:52:08	0
3	HHdd53ef1db6	cellular 🔼	921111	nent Pr		CXab		alure	16:30:24	16:36:04	1
4	HH6fa0de6516	cellular	11	may	2	-1	0	NA	12:06:43	12:20:22	1
5	HHeb436ca7cf	cellular	3	jun	1	-1	0	NA	14:35:44	14:38:56	0
6	HH5119beb3cd	cellular	h++23	may //	roodd	209	1	failure	14:58:08	15:11:24	1
7	HH3a1172a4ab	cellular	111119	mar // DOV	/ COUG	r.COj		NA	13:00:02	13:03:17	1

## Add We Chat powcoder

- Each record is a standalone observation of a phenomena you are trying to predict, classify or cluster.
- Records have defined attributes for each data row.
- Notice time is not interacting between rows (or it had not better be) but is present at the observational row.

HARVARD UNIVERSITY

### Difference between ML Data Setup & Time Series Data

#### **Time Series Data**

Assignments Project Tways in Lange vector with each value being in sequence to the next.

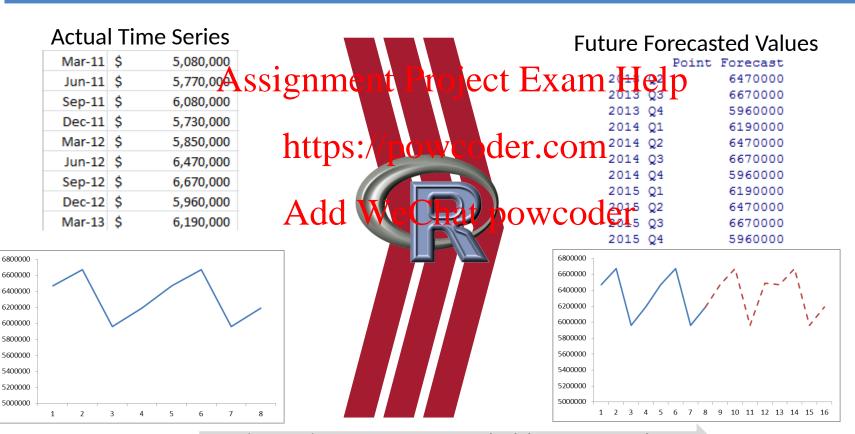
httptributes may not be present because temporal information is held "within" the vector due to the

Add WeChat powcoder



## What is forecasting?

Forecasting is the process of applying mathematical tools on time series data to create future time series values, doesn't have to explain the reason for observed changes.



Times Series Data > Forecast Methodology> Future Values

HARVARD UNIVERSITY

## Explain Vs. Predict

#### **Explanation** is the goal of "time series analysis"

Models are based on causal argument Models are not "black-box"

## Example Explansionment Project Exam Help

"The housing crisis reduced the expected bank revenue over a 2yr period" "Inclement weather netgative yperfected holiday chapping at Target by 5%."

## Add WeChat powcoder

Forecasting (our focus) seeks to predict future values

#### Example outcomes:

"Next quarter bank revenue is forecasted to rise to \$XYZ" "Wal-Mart's 3<sup>rd</sup> quarter revenue will be \$130B."

HARVARD UNIVERSITY

### The data inside the time series data

**Level** – an average of the observations "steady state"

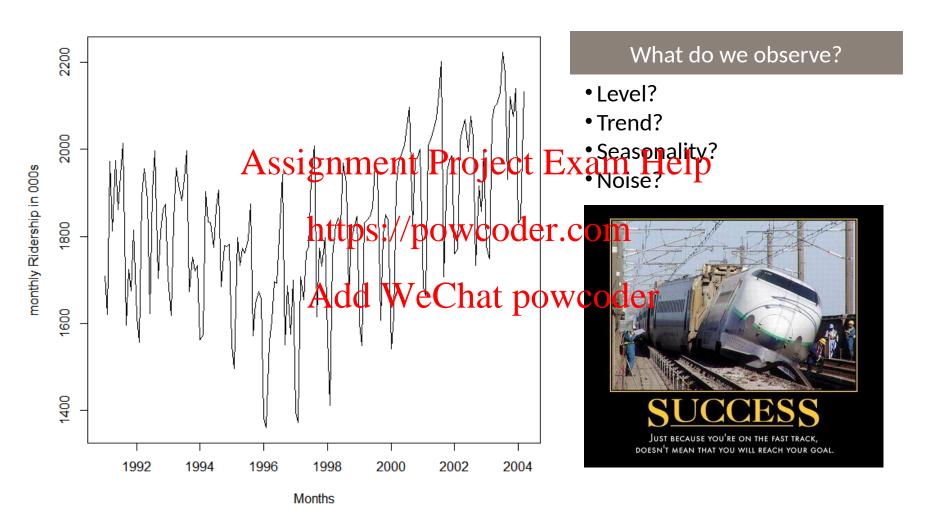
Trend – are values increasing, decreasing or stationary Assignment Project Exam Help

Seasonality - is ther het representation the periodicity

Noise - unexplained And Sere February Work adding "trend", "seasonality" and "level" together. Basically its what left, and unaccounted for.

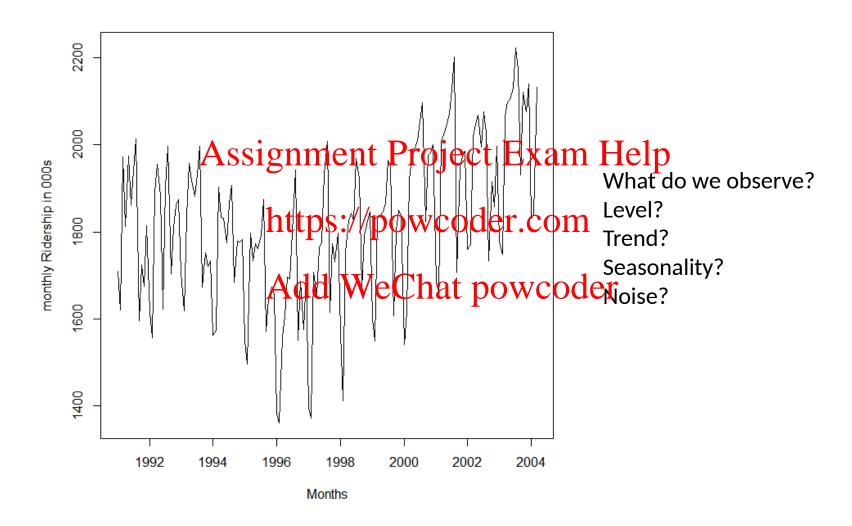


## **Amtrak Actual Riders**



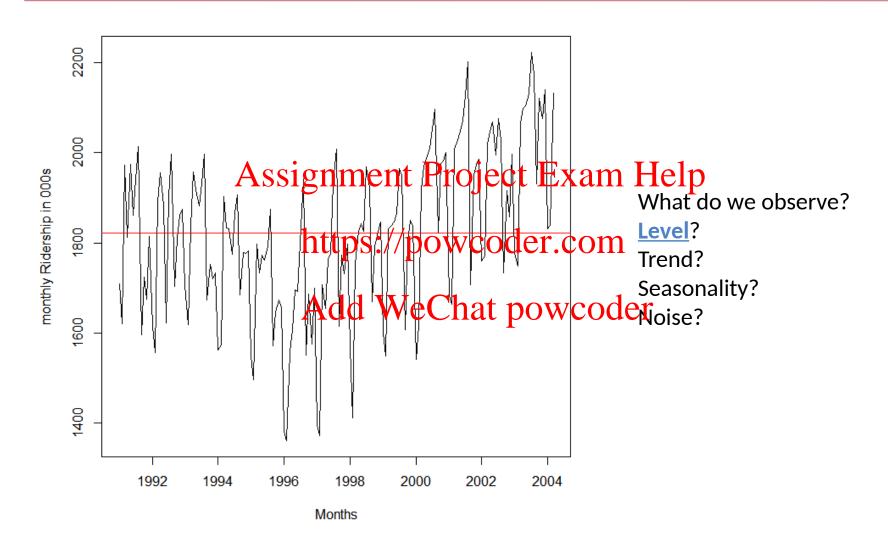


## **Amtrak Actuals**



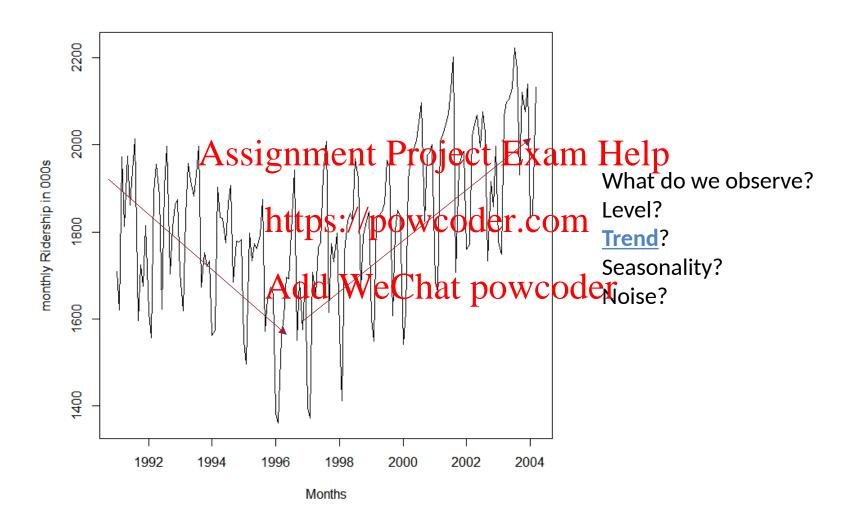


### **Amtrak Actuals**





### **Amtrak Actuals**





#### Seasonality\* appears:

Each year traffic peaks in summer Assignment Project Exam Help

#### Noise:

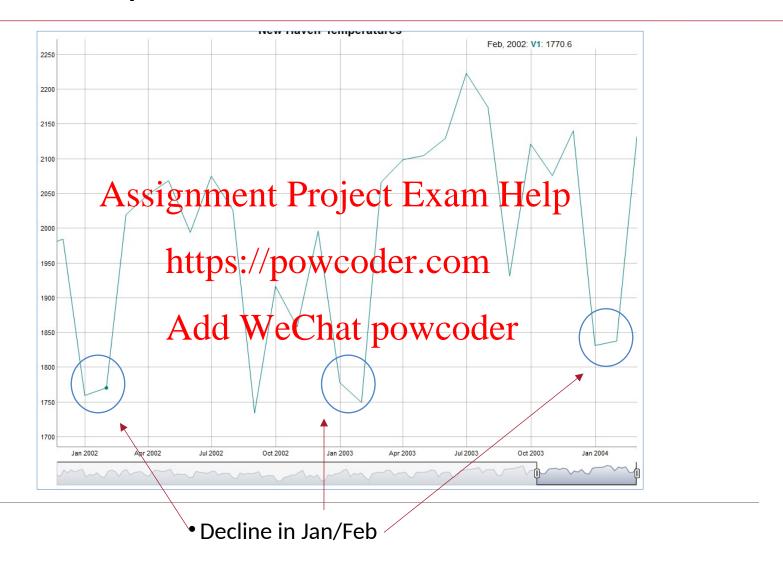
Departure from the general level that is neither trend nor seasonality

Put another way, if you add the seasonal, trend and level values, the

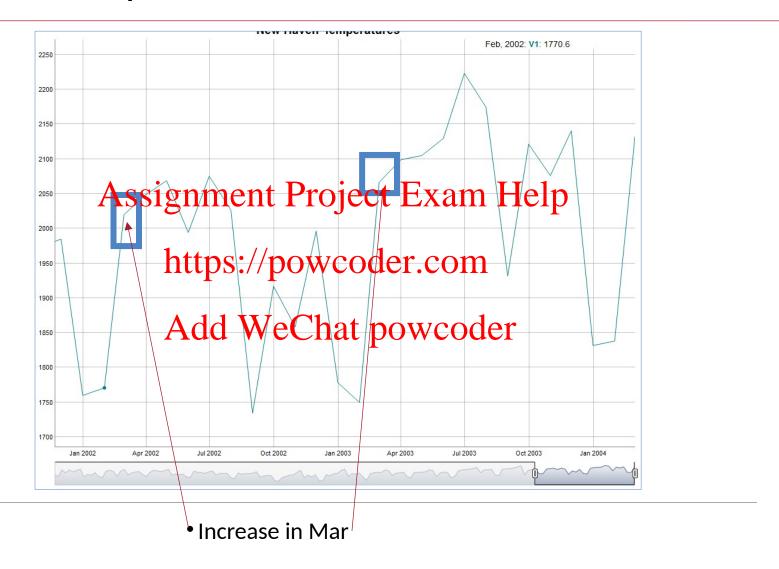
difference is the You'se We Chat powcoder

\* Don't confuse the time series term "season," which is the period over which a cyclical pattern repeats (e.g. a year), with the standard English seasons of the year (fall, winter, etc.)







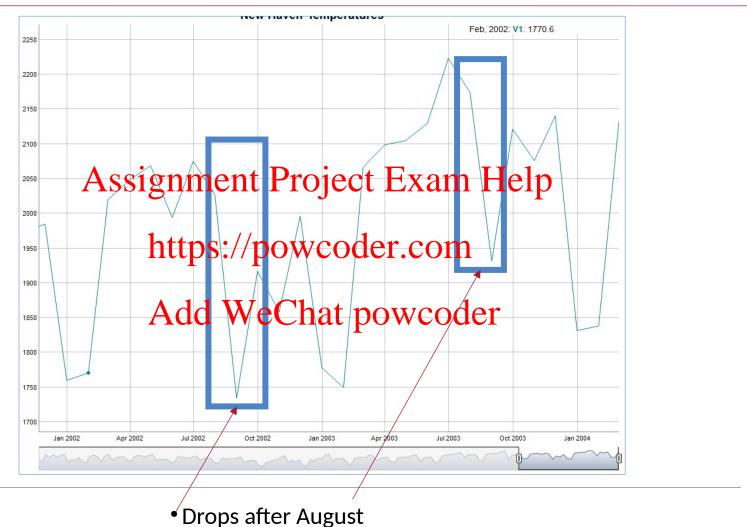






Sustained increase until Aug

HARVARD UNIVERSITY



.

HARVARD UNIVERSITY

## Open 0\_amtrak.R

#### Let's Practice!

Time Series Basics

Dynamic Plottinssignment Project Exam Help Lubridate Functions

https://powcoder.com

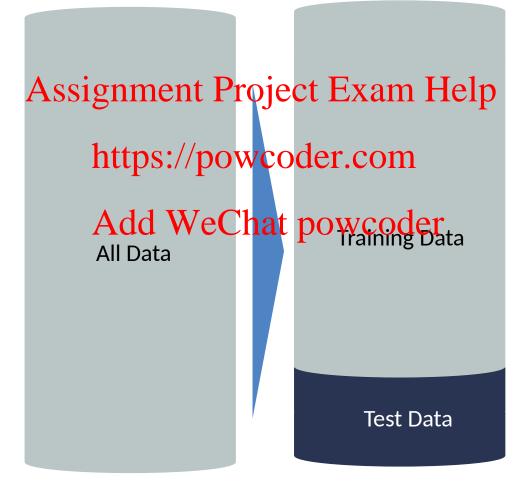
Add WeChat powcoder



18

## **Machine Learning Partitioning**

Divide data into training portion and validation portion Test model on the validation portion





## Time Series Partitioning is not random

Random partitioning would leave holes in the data, which causes problems

Forecasting methods assume regular sequential data

Instead of random selection, divide data into two parts

Train on early data
Validate on later data ignment Project Exam Help

• Performance can be assessed against the "naïve benchmark" & historical accuracy raiters & prosimply to the first of the profile of the prof



Sub-setting data with regard to time is called "out of time" sampling.

2022/11/23 Kwartler CSCI 96 19 HARVAR UNIVERSITY

## Summary

Focus is to predict (not describe/explain)

Four component symmetry Project Exam Help

Level

Trend https://powcoder.com

Seasonality

Noise Add WeChat powcoder

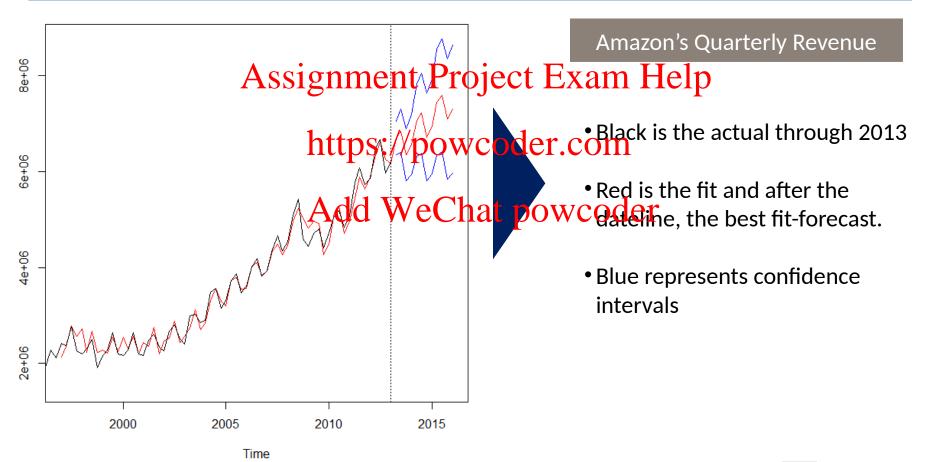
Partition data by dividing into early/late



20

## What types of business problems can be forecasted?

We will cover a set of common forecasting tools to make predictions. The goal is to create accurate future values and provide ranges of accuracy in real contexts.



HARVARD UNIVERSITY

## Open 1\_getRevenueData.R

• Let's grab a time series & plot it

Assignment Project Exam Help

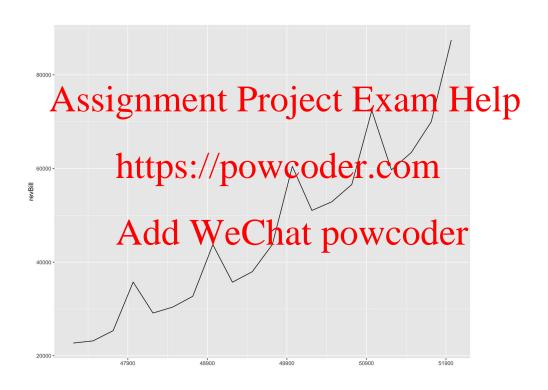
https://powcoder.com

Add WeChat powcoder



## Inspecting meta data.

#### What is the meta data for AMZN?



Does the meta data exhibit a discernable pattern?

Do you think historical values are a basis for future values?

HARVARD UNIVERSITY

# Agenda

Start	End	Item
		What is Forecasting?
		Naïve Forecasting Methods
	As	Time Series Decomposition Signments Projecto Exam Help
		Appendix https://powcoder.com
		Add WeChat powcoder



#### **5 Common Methods**

#### We will cover 3 common methods to forecasting.

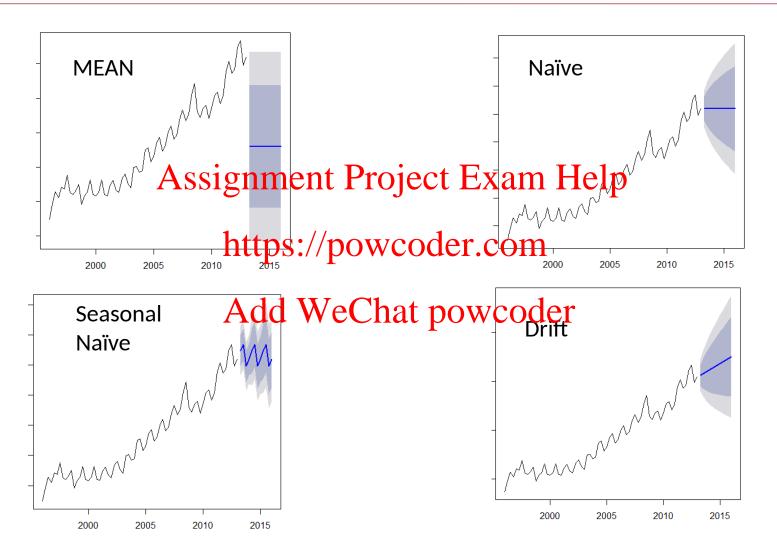
Linear Naive **Holt Winters TSD ARIMA** Modeling -Assignment Project Exam Help Essentially Most Smoothing into meta- Good with linear common • Good https://powcoder.comtrends/trend Easy/Fast regression components Usually not w/Seasonality reversals Requires Add We@phfprplowcodegod with additional very accurate seasoning & noisy series data investigating Good manipulation Not really w/Seasonality & effort • "black box" forecasting future

ARIMA & Linear Model Forecasting is in the Appendix presentation.

Their impact on the final will be minimized but some/few multiple choice questions from the book may appear.

25

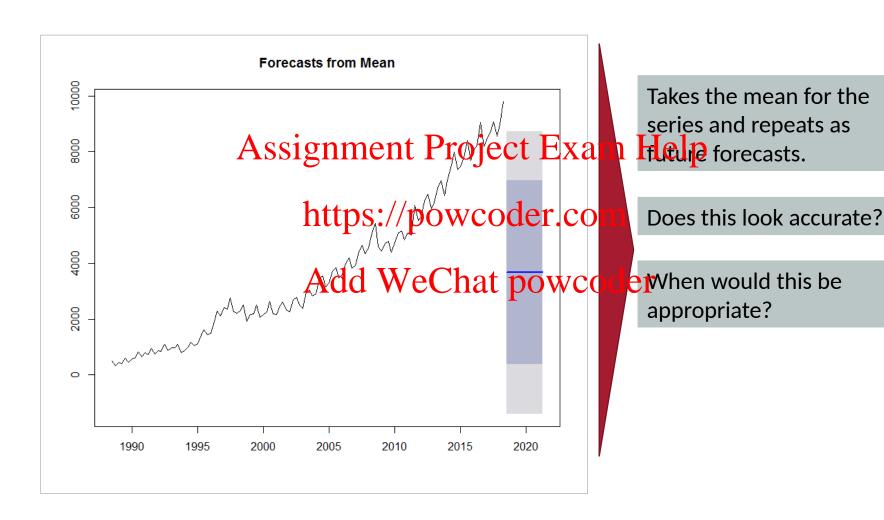
## 4 Methods of Naïve Forecasting





26

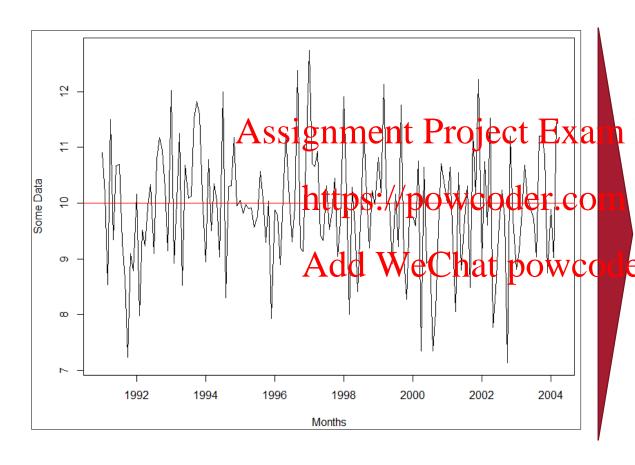
#### Naïve Forecast - Mean



HARVARD UNIVERSITY

28

### Naïve Forecast - Mean



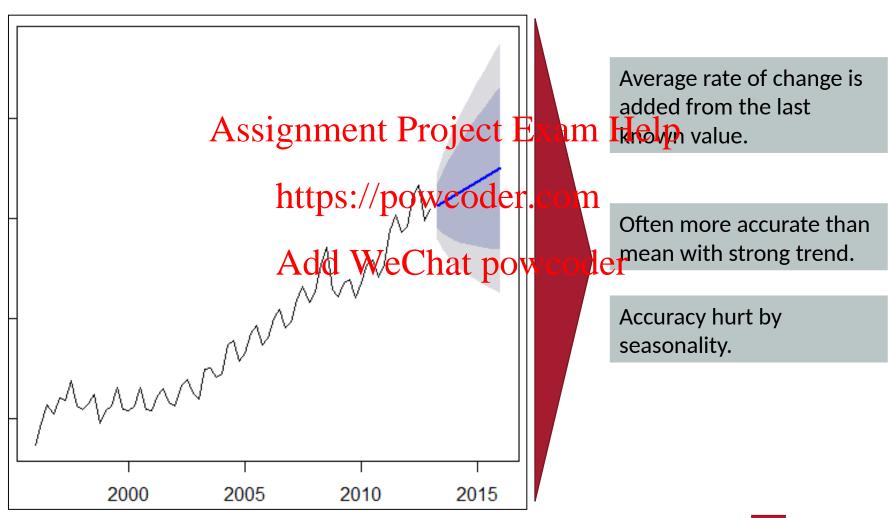
Takes the mean for the series and repeats as forecasts.

Does this look accurate?

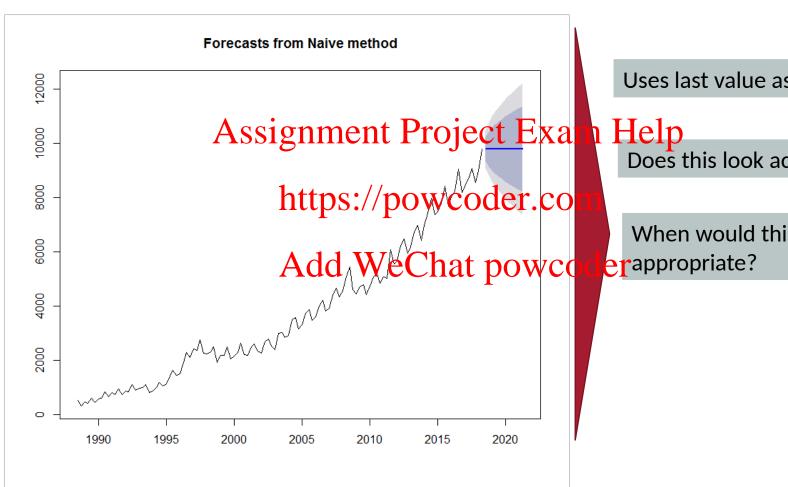
et When would this be appropriate?

HARVARD UNIVERSITY

### Naïve Forecast - Drift



## Naïve Forecast - Naïve (true)



Uses last value as future.

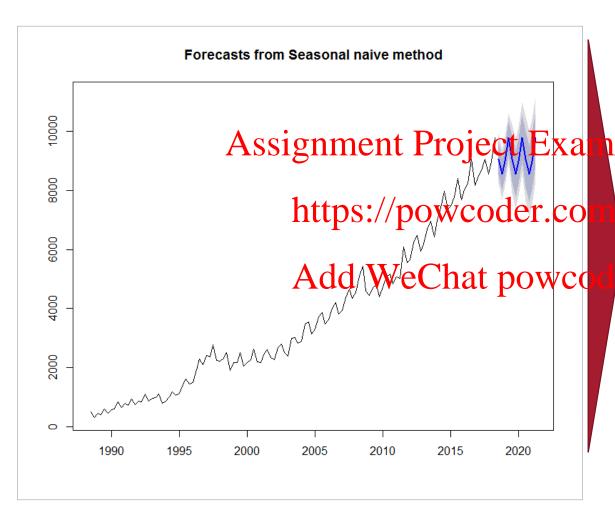
Does this look accurate?

When would this be

31



## Naïve Forecast - Naïve Seasonal



Uses last corresponding seasonal values in a repeting pattern. Good if no trend.

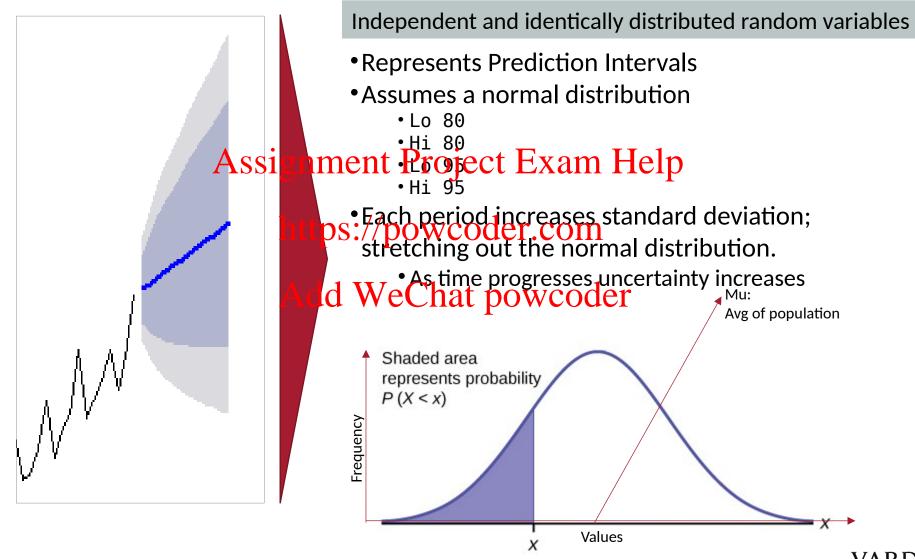
Does this look accurate?

er

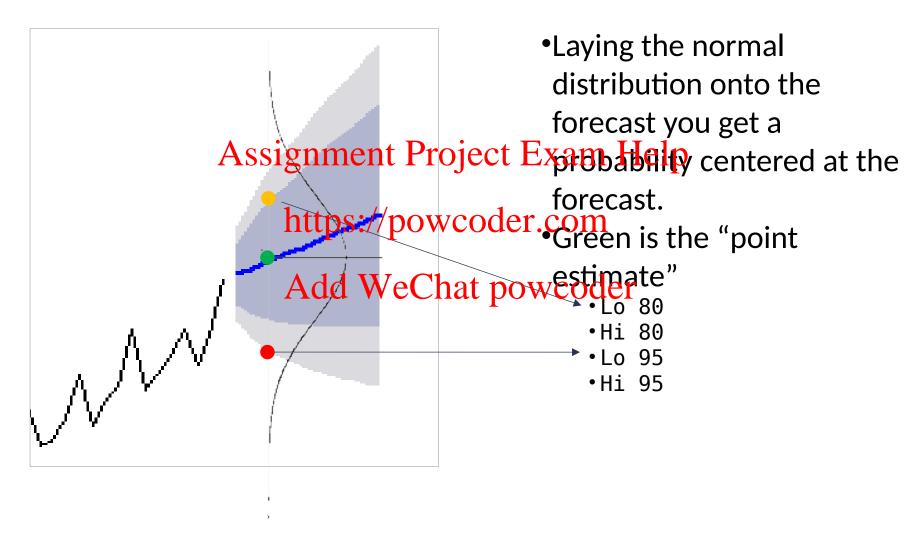
When would this be appropriate?



## **Shaded Forecast Area?**



### **Shaded Forecast Area?**



## Open 2\_NaiveNike.R

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

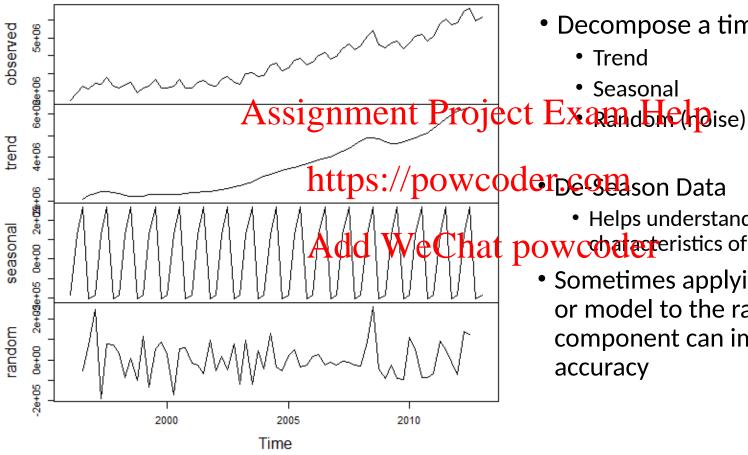


# Agenda

Start	End	Item
		What is Forecasting?
		Naïve Forecasting Methods
	As	Time Series Decomposition Signments Projecto Exam Help
		Appendix https://powcoder.com
		Add WeChat powcoder



#### Decomposition of additive time series



- Decompose a time series into
  - Trend
  - Seasonal

- Helps understand the underlying at powcodereristics of a time series
  - Sometimes applying a forecast or model to the random component can improve accuracy

37

$$Y_t = (S_t, T_t, E_t)$$

Data at a specific time period (t) is equal to a mix\* of seasonal values, trend values and what prefix the same period.

https://powcoder.com

Where (Math-speak version): WeChat powcoder

 $Y_t$  = data at period t

S<sub>t</sub> = seasonal component at period t

T<sub>+</sub> = trend component at period t

E<sub>+</sub> = remainder or residual component at period t



$$Y_t = (S_t, T_t, E_t)$$

Data at a specific time period (t) is equal to a mix\* of seasonal values, trend values and what prejisted that the same period.

https://powcoder.com

Additive - Y<sub>t</sub>= Seasonal effect + Trend + Residual

An additive model assumes that the difference between each time period is approximately the same

For example, Jan trend is +100, so next Jan trend would add another +100.

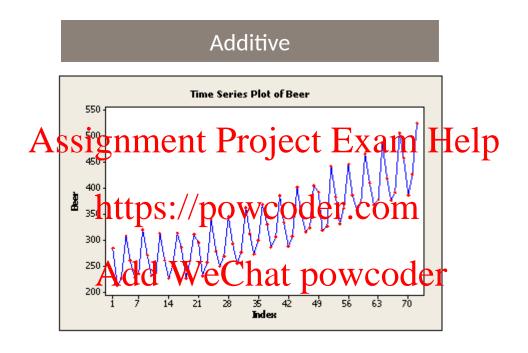
**Multiplicative** - Y<sub>t</sub>= Seasonal effect X Trend X Residual

A multiplicative model assumes changes are proportional and not constant.

For example Jan season is +100 as part of a 1,000 total (10%). The next Jan the total is 1500, and the seasonal adjustment would be 150 (10%).

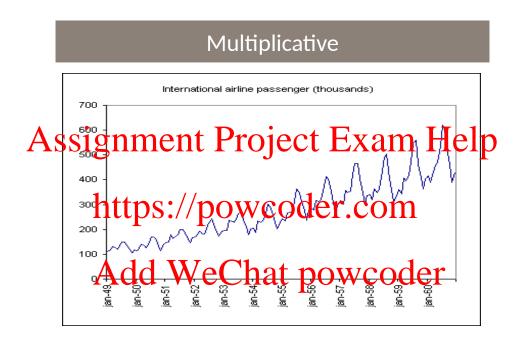
HARVARD UNIVERSITY

<sup>\*</sup>Mix can either be



- Use additive models if the seasonality (repeating pattern) is similar.
- Peak to trough looks the same, so additive is ok.

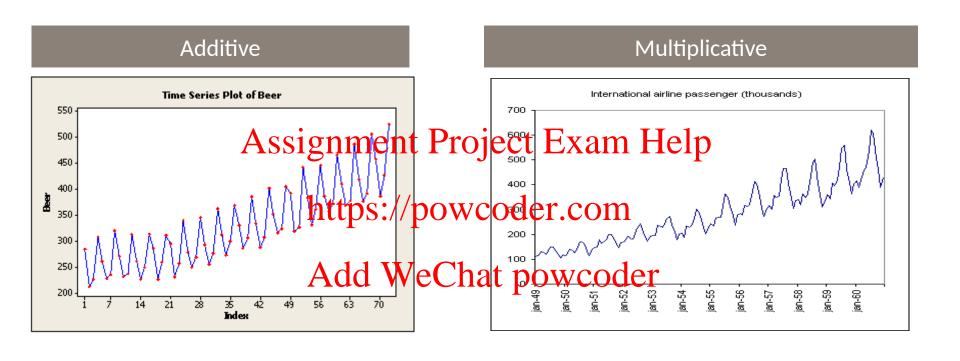
HARVARD UNIVERSITY



- Use multiplicative if the seasonality grows larger over time but is still the same proportion of the total
- The peak to trough looks exaggerated over time

HARVARD UNIVERSITY

### Time Series Decomposition - Side by Side



- Use additive models if the seasonality (repeating pattern) is similar.
- Use multiplicative if the seasonality grows larger over time but is still the same proportion of the total

HARVARD UNIVERSITY

### **Summary - Time Series Decomposition**

- Separates Trend, Seasonal and Random components of a time series
- Components are combined by
  - Additive Adding compend of isoperoprate if the less apal pattern is consistent
  - Multiplicative multiplying components is appropriate if the seasonal pattern changes over time but it proportional to the time series.
- TSD can help you understand the data and pan be contents part of EDA
- Each component can be forecasted separately to (sometimes) improve accuracy then each forecast can be combined to arrive at a final forecast
- De-seasoning data is possible by subtracting (additive TSD) or dividing (multiplicative TSD) it out of the time series – see the impact of an event from expected seasonal or trend changes; quantifies impact

HARVARD UNIVERSITY

### Open 3\_TimeSeriesDecompositionAMZN.R

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder



## Agenda

Start	End	Item
		What is Forecasting?
		Naïve Forecasting Methods
		Time Series Decomposition
	As	signments Projecto Exam Help
		Appendix https://powcoder.com
		Add WeChat powcoder



#### Mean Average – good for population summary

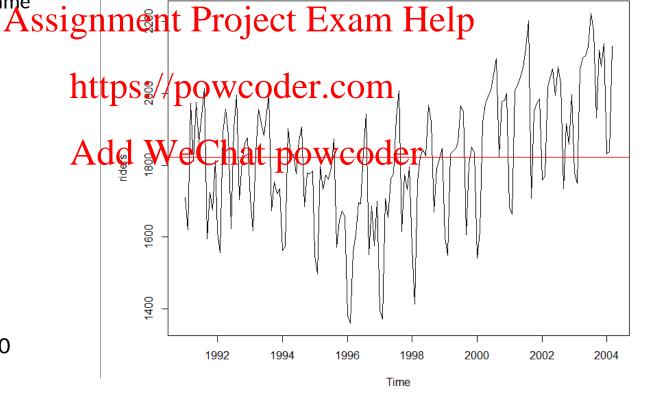
Add all values and divide by population

mean(riders)
[1] 1822.197

Each record has the same weight.

Number	Value
1	10
2	20
3	30
4	40
5	50

$$\frac{10+20+30+40+50}{5}$$
 =30



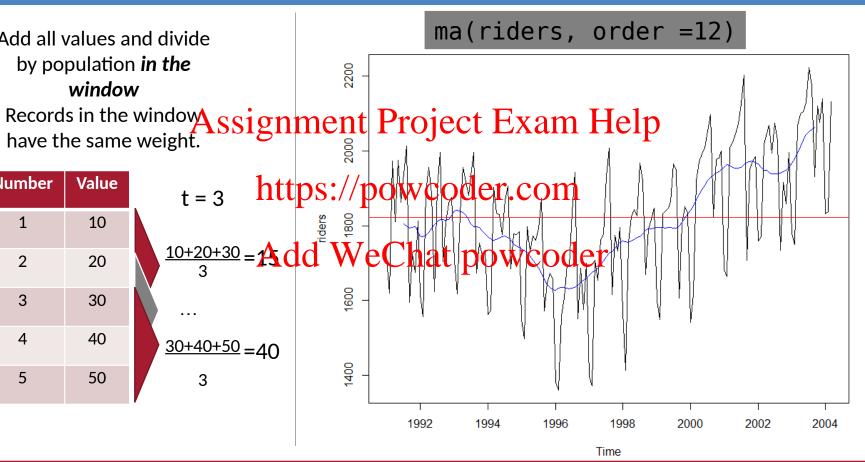
#### Centered Moving Average – smooths seasonality

Add all values and divide by population in the window

Number	Value	t=3 ht
1	10	
2	20	$\frac{10+20+30}{3} = 1$
3	30	
4	40	<del>30+40+50</del> =40

50

5

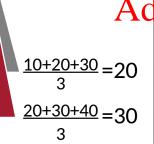


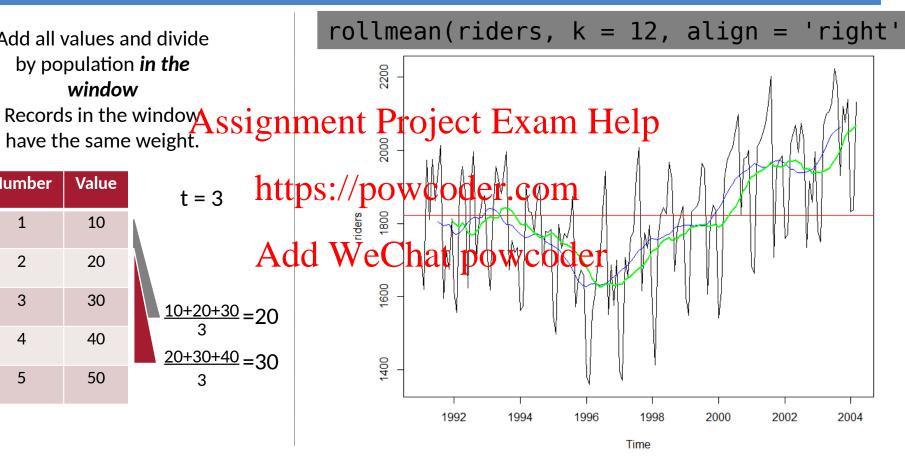
Descriptive because it uses values from the future so not good for forecasting.

#### Trailing Moving Average – smooths seasonality

Add all values and divide by population in the window

Number	Value	
1	10	١
2	20	١
3	30	
4	40	
5	50	





Uses preceding window values so ok for forecasts but lags for trend and seasonal.

#### **Exponential Smoothing**

#### Older records in the window have the diminishing weight

#### Assignment Project Exam Help

# α is a parameter between 0 and 1. 0 = more weight is given to observations from

- 0 = more weight is given to observations from the more distant past
- the more distant past
   Approaching 1 more Weight given to recent
- 1 = all weight given to the most recent (same as a true Naïve forecast)

Values are weighted so their impact diminishes in the average the farther back.

HARVARD UNIVERSITY

### Open 4\_HoltWintersWMT.R

HW applies exponential smoothing to **level**, **trend** and **seasonality** individually then combines them.

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder



#### Class announcement

- Stick to the existing assignments; reviewing workload options given the sudden change no promises at this time.
- Syllabus Changesignment Project Exam Help
  - No guest speakers later in the semester; having an optional and NOT on the examplesson on "growth models" used in new product forecasting & viral outbreaks

Add WeChat powcoder

Please stay mentally and physically healthy.

