

# Road map!

- Module 1- Demystifying Timeseries Data and Modeling
- Module 2- Setting up Deep Forecasting Environment
- Module 3- Exponential Smoothing
- Module 4- ARIMA models
- Module 5- Machine Learning for Time series Forecasting
- Module 6- Deep Neural Networks
- Module 7- Deep Sequence Modeling (RNN, LSTM)
- Module 8- Prophet (Forecasting at Scale)



Pedram.Jahangiry

# Module 1

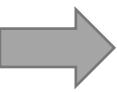
## Demystifying Timeseries Data and Modeling



TS Basics  
Modeling  
Forecasting



Pedram Jahangiry



# Key topics

## 1. Timeseries Basics

- What is sequence data?
- Time series tasks
- Trend, Seasonality, Residual
- ACF and PACF
- Stationarity

## 2. Forecasting strategies

- One-step, Multi-step, Multi-output
- Univariate vs Multivariate
- Benchmarks

## 3. Modeling

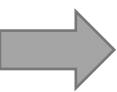
- White Noise
- Econometrics
- Machine Learning
- Deep Learning

## 4. Can we beat wall street?

- Is stock price Random Walk? (EMH)
- Does Uni-variate Model work?
- Can Multi-variate Model help?
- So What?



Pedram.Jahangiry



# Module 1- Part 1

## 1. Timeseries Basics

- What is sequence data?
- Time series tasks
- Trend, Seasonality, Residual
- ACF and PACF
- Stationarity

## 3. Modeling

- White Noise
- Econometrics
- Machine Learning
- Deep Learning

## 2. Forecasting strategies

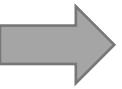
- One-step, Multi-step, Multi-output
- Univariate vs Multivariate
- Benchmarks

## 4. Can we beat wall street?

- Is stock price Random Walk? (EMH)
- Does Uni-variate Model work?
- Can Multi-variate Models help?
- So What?

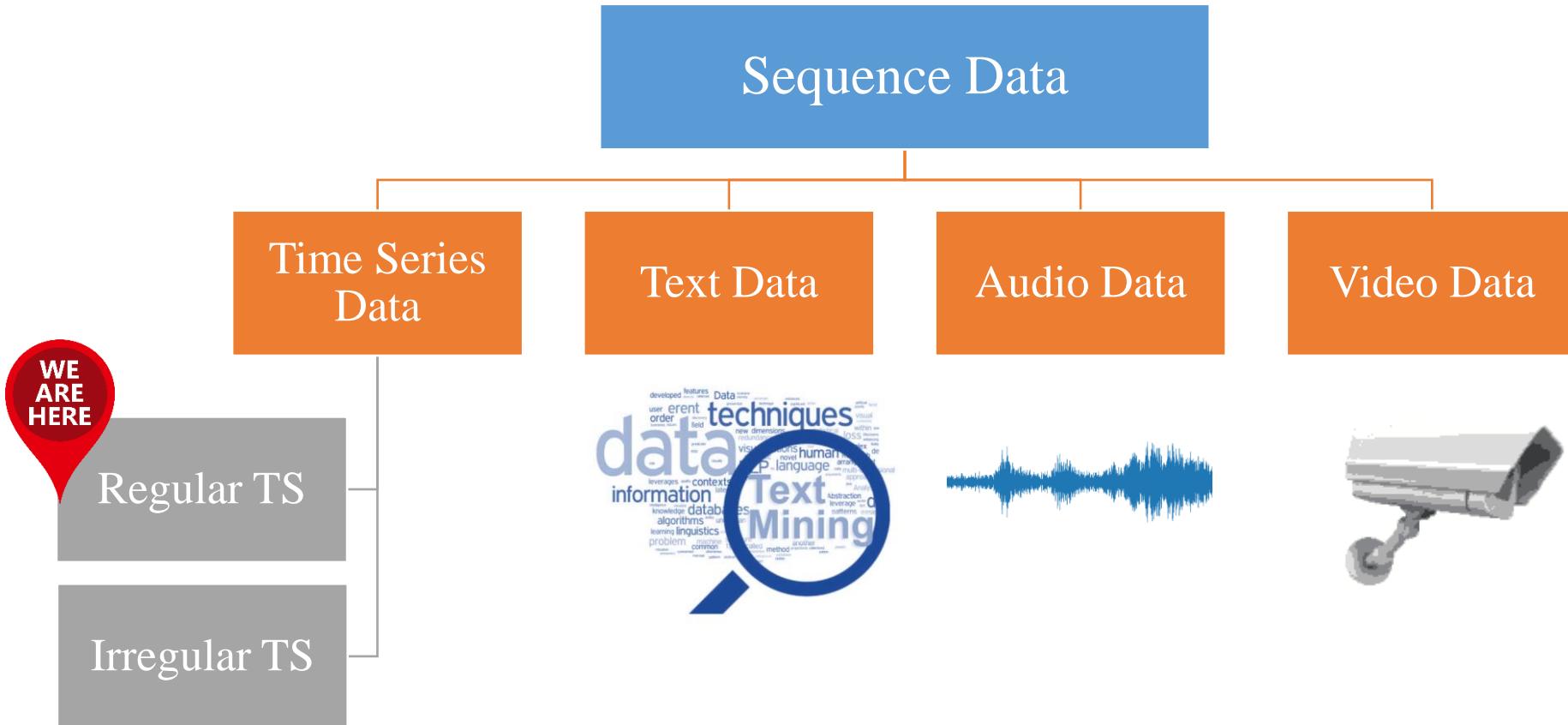


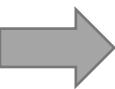
Pedram.Jahangiry



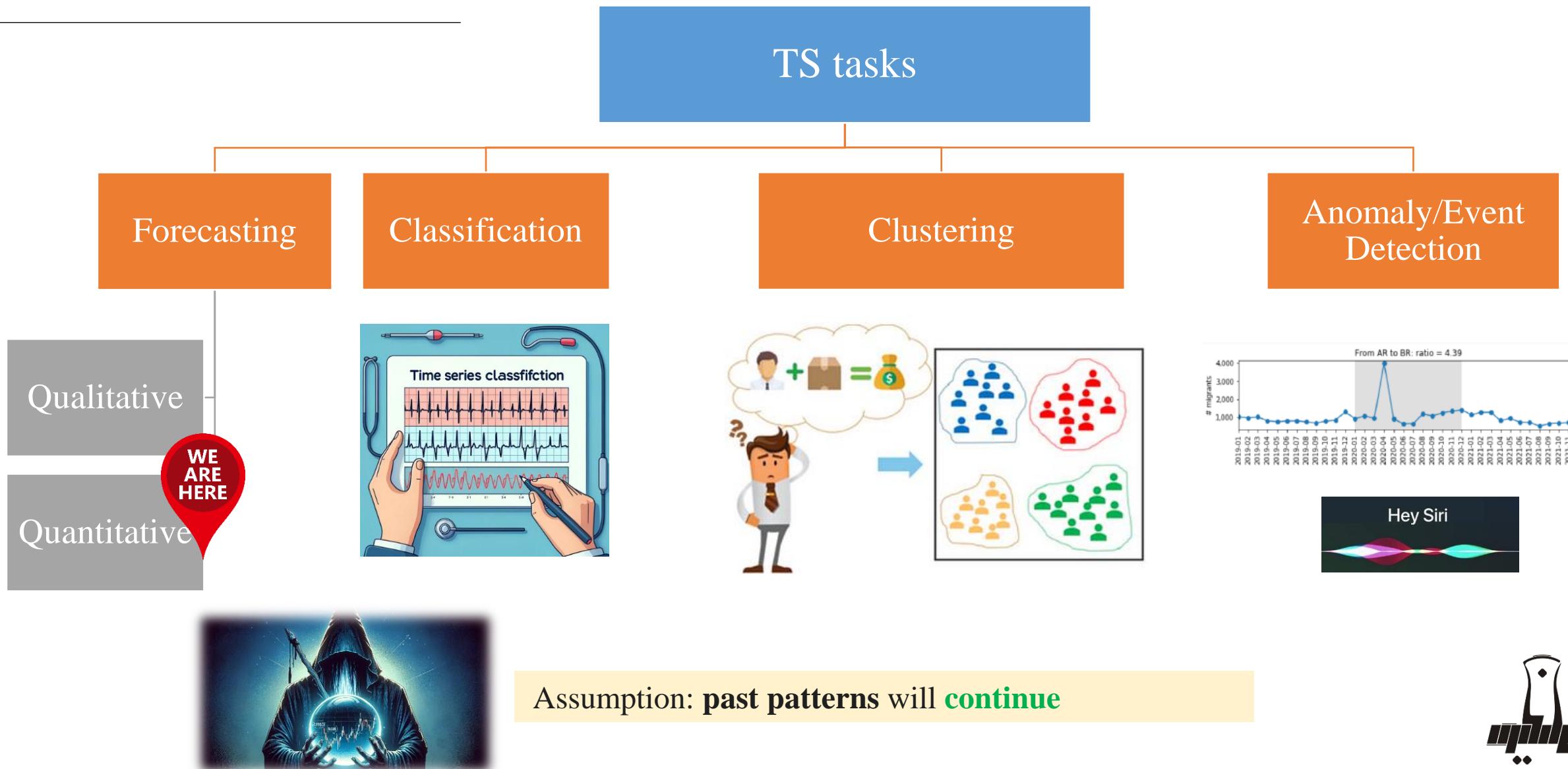
# What is Sequence Data?

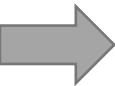
Sequence data refers to any data that has a specific **order** or sequence to it!





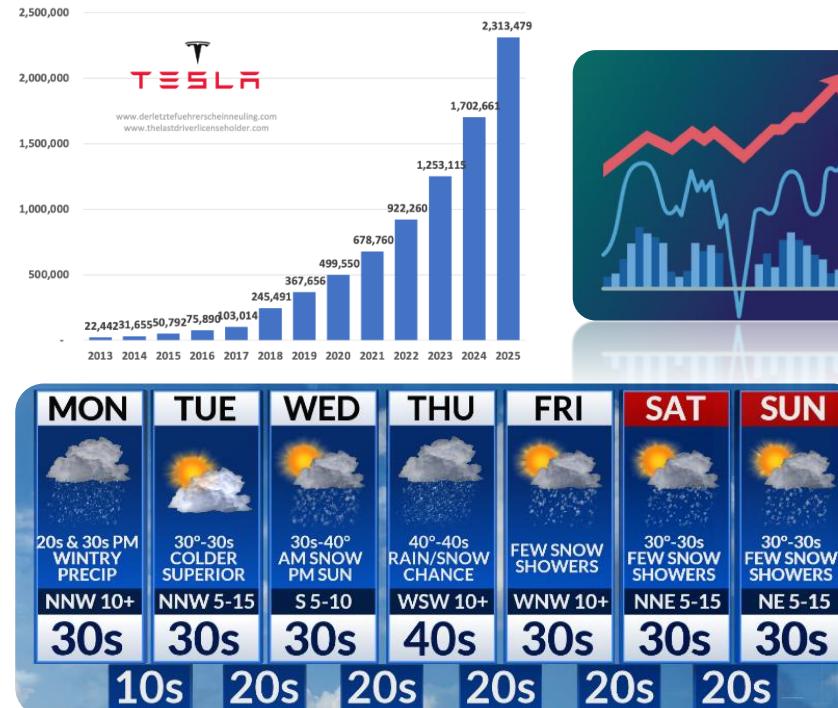
# Time series Tasks

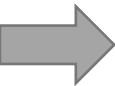




# More Forecasting examples

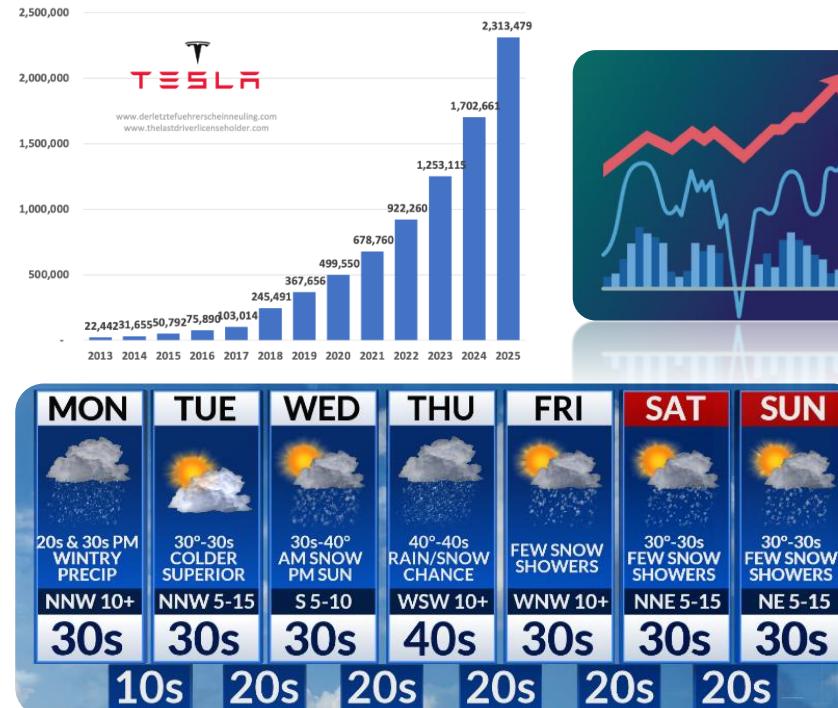
- Which of the following examples are easier to forecast?
  1. Time of sunset this day next month
  2. Apple stock price tomorrow
  3. Airline ticket demand/price next year
  4. New car model sales in the first quarter
  5. Monthly rainfall in Utah next winter
- What impacts Forecastability?
  - Data Availability
  - How similar the future is to the past!
  - Good understanding of the underlying factors



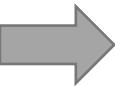


# More Forecasting examples

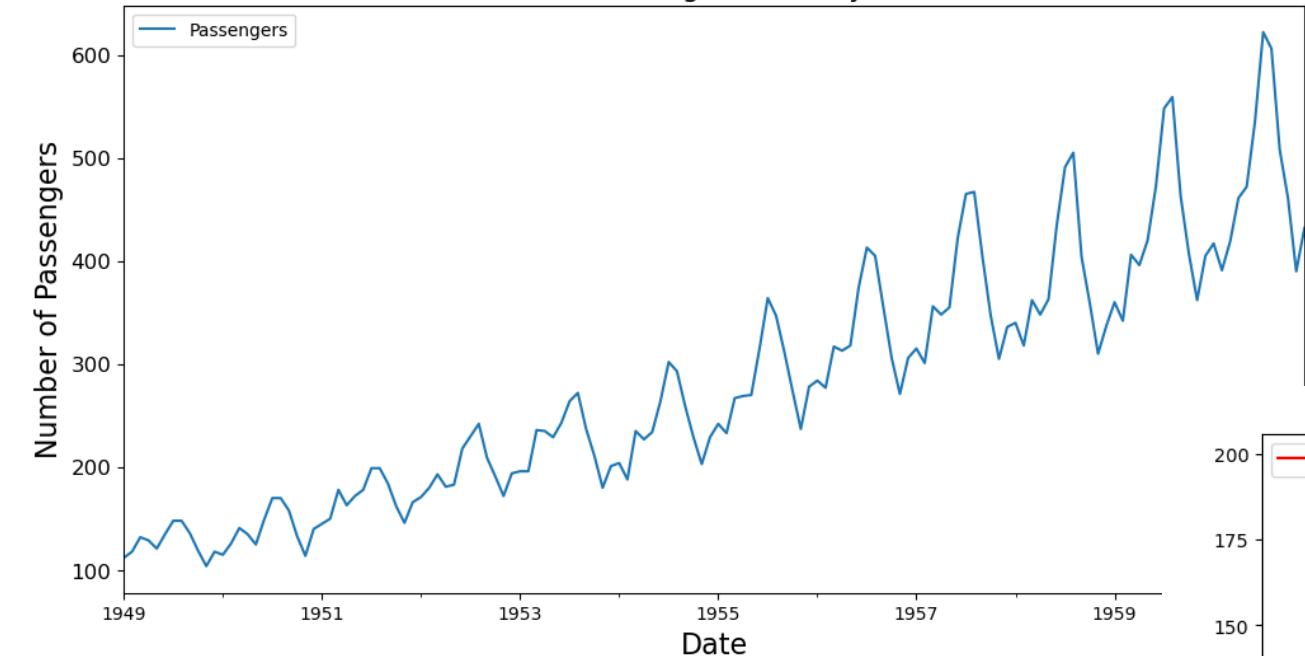
- Which of the following examples are easier to forecast?
  1. Time of sunset this day next month
  2. Apple stock price tomorrow
  3. Airline ticket demand/price next year
  4. New car model sales in the first quarter
  5. Monthly rainfall in Utah next winter
- What impacts Forecastability?
  - Data Availability
  - How similar the future is to the past!
  - Good understanding of the underlying factors



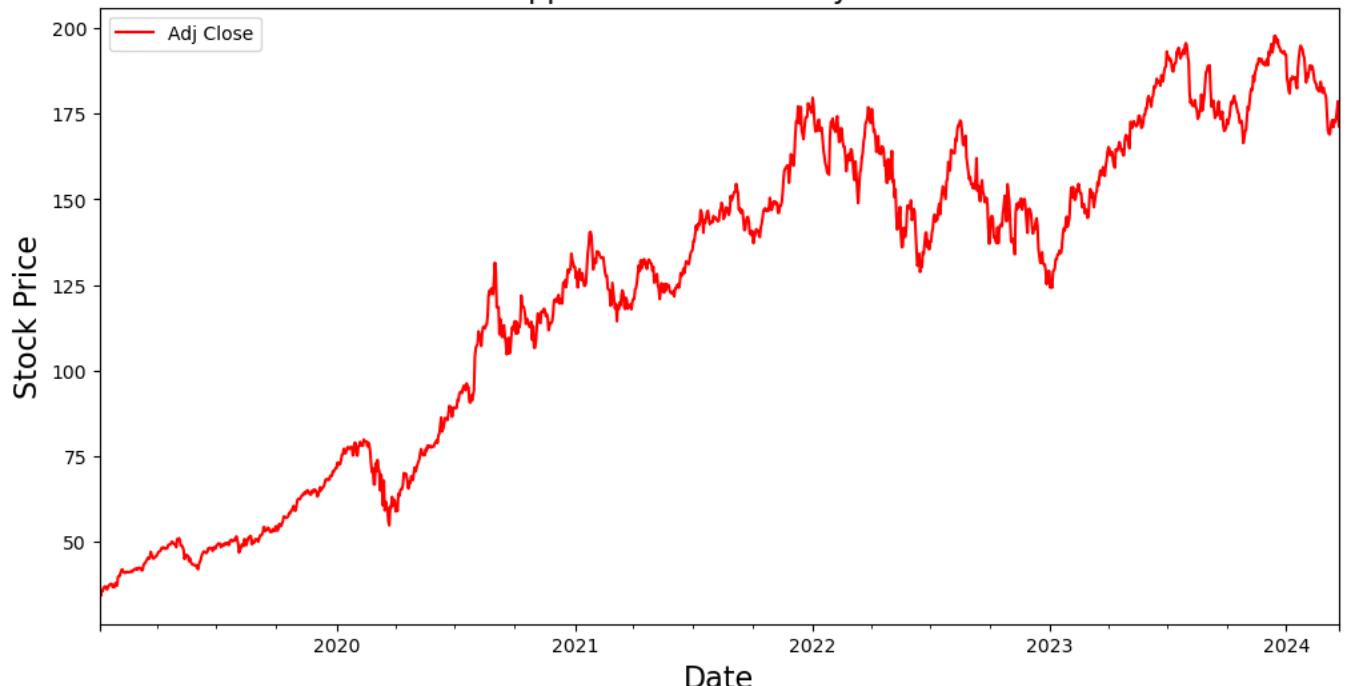
# Quantitative Forecasting



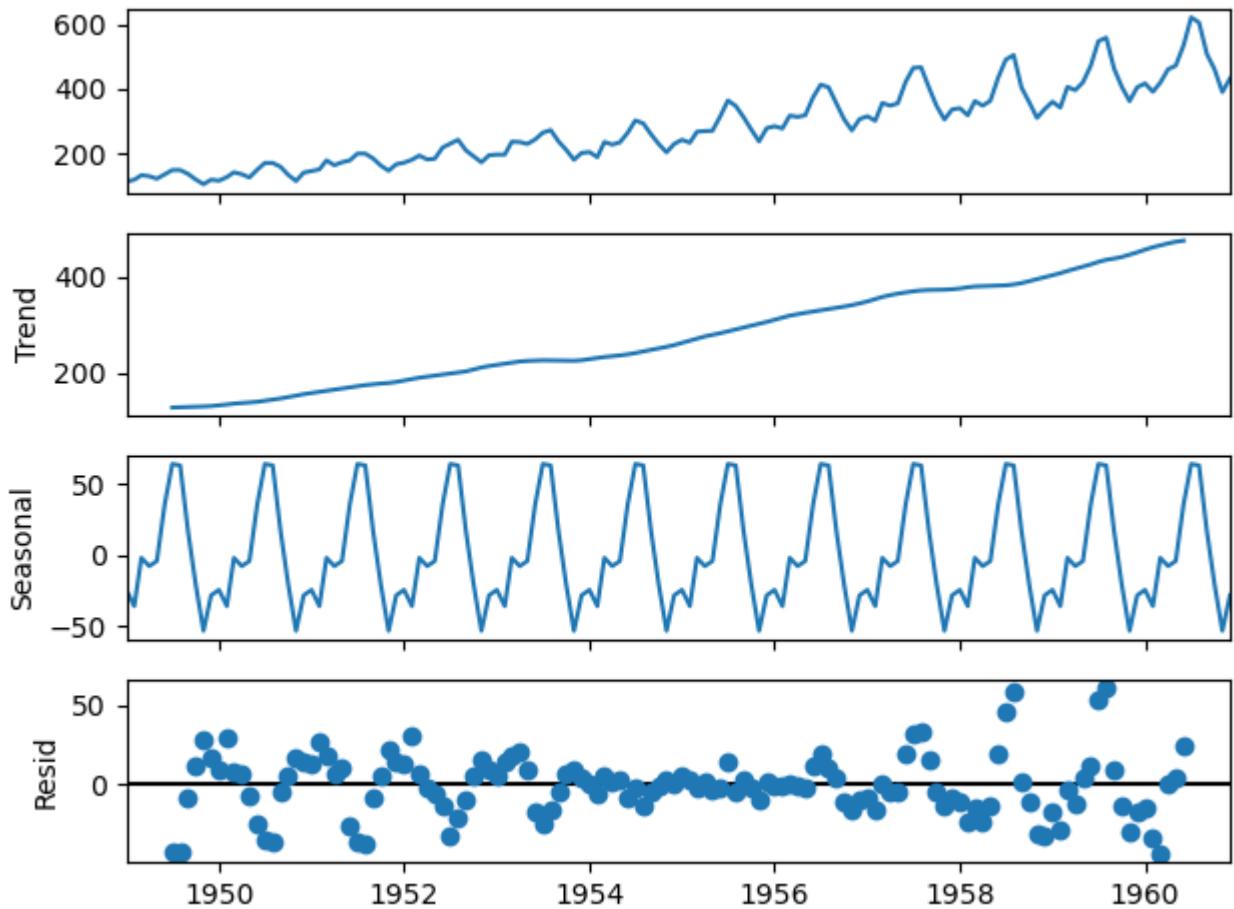
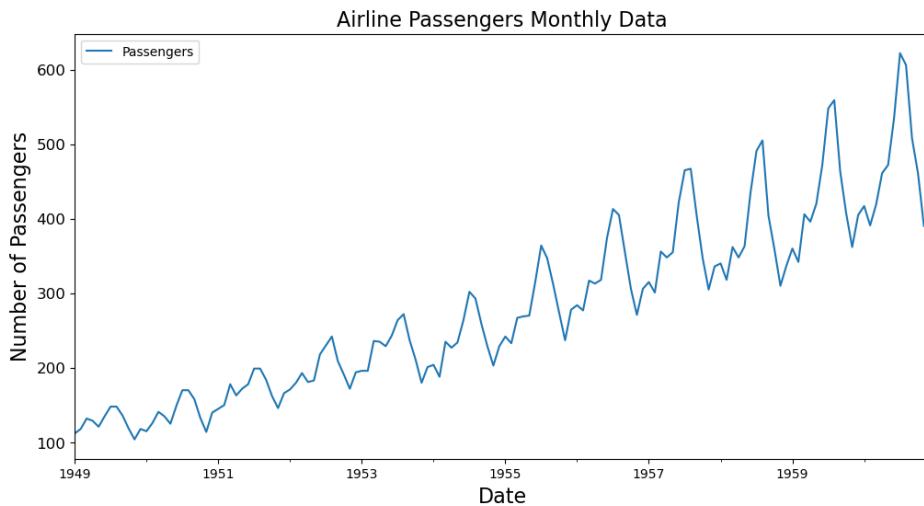
Airline Passengers Monthly Data

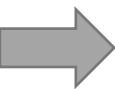


Apple Stock Price Daily Data

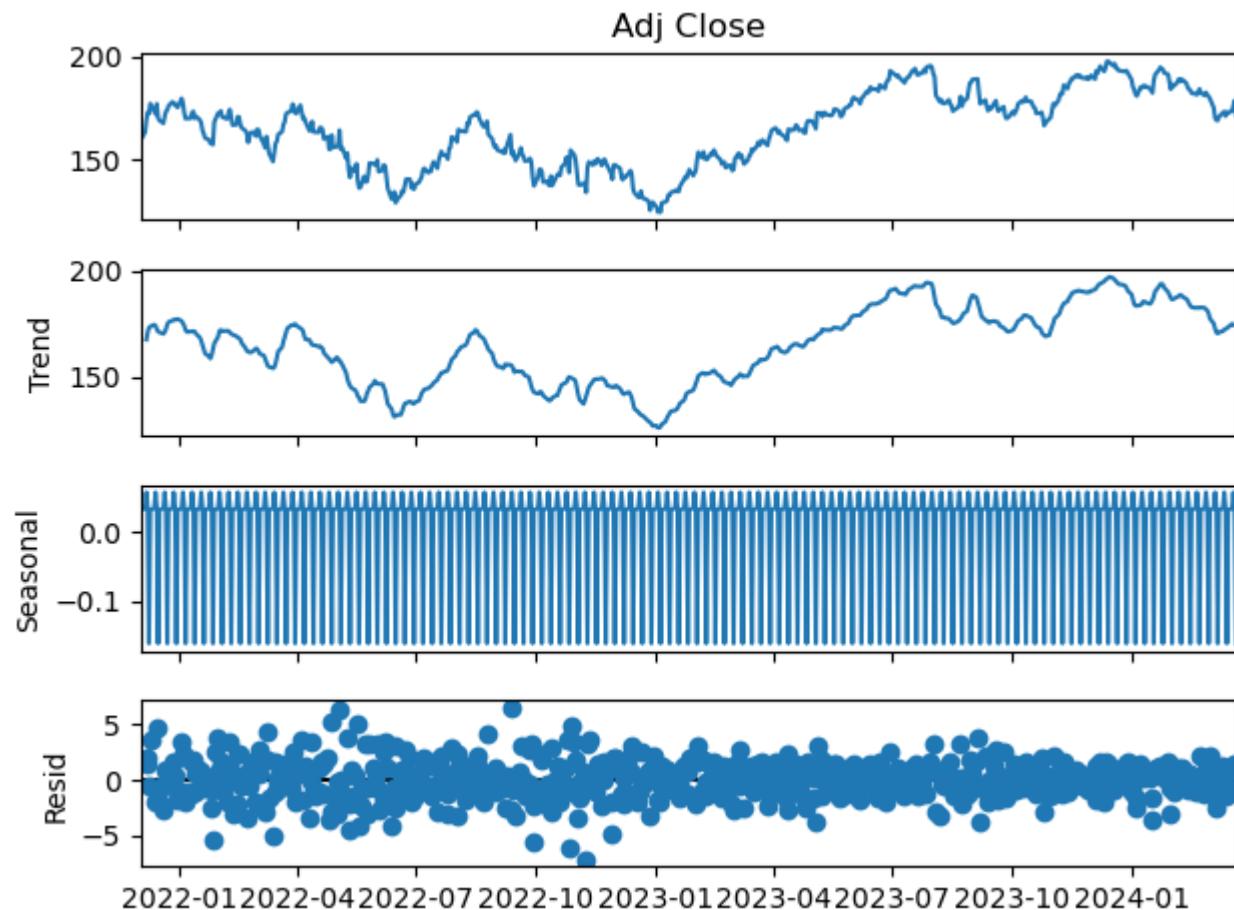


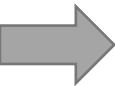
# TS Decomposition: Trend, Seasonality, Residual





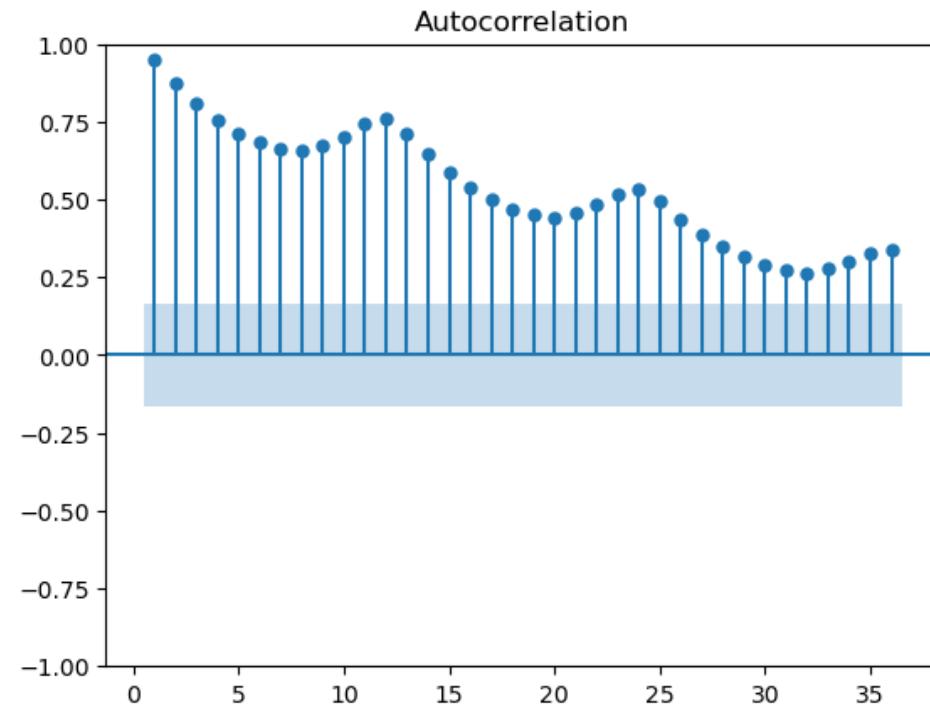
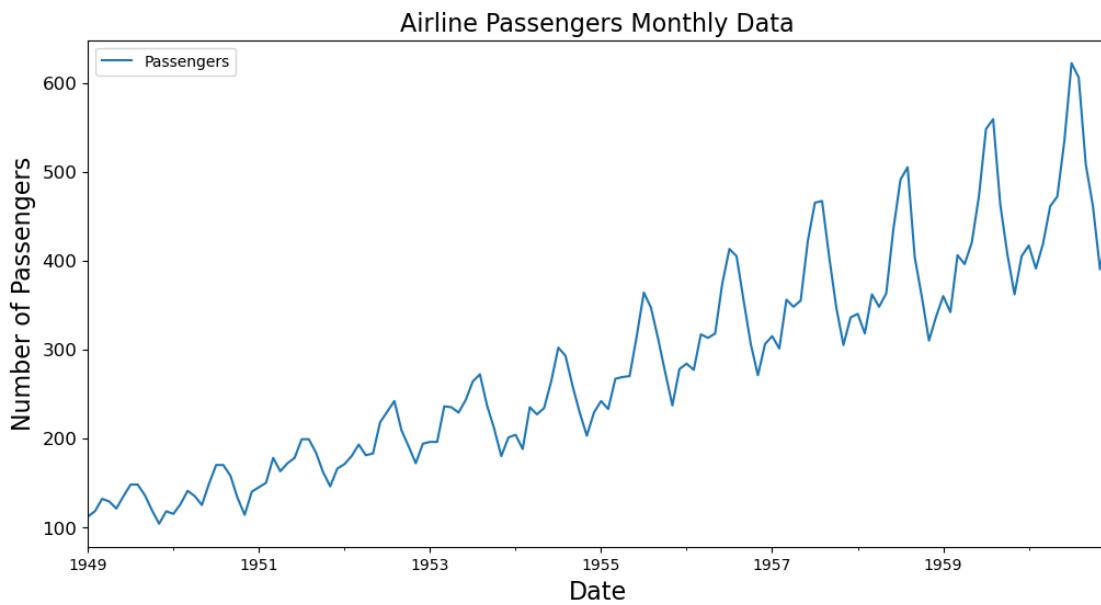
# TS Decomposition: Trend, Seasonality, Residual

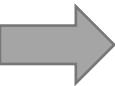




# ACF: Autocorrelation Function

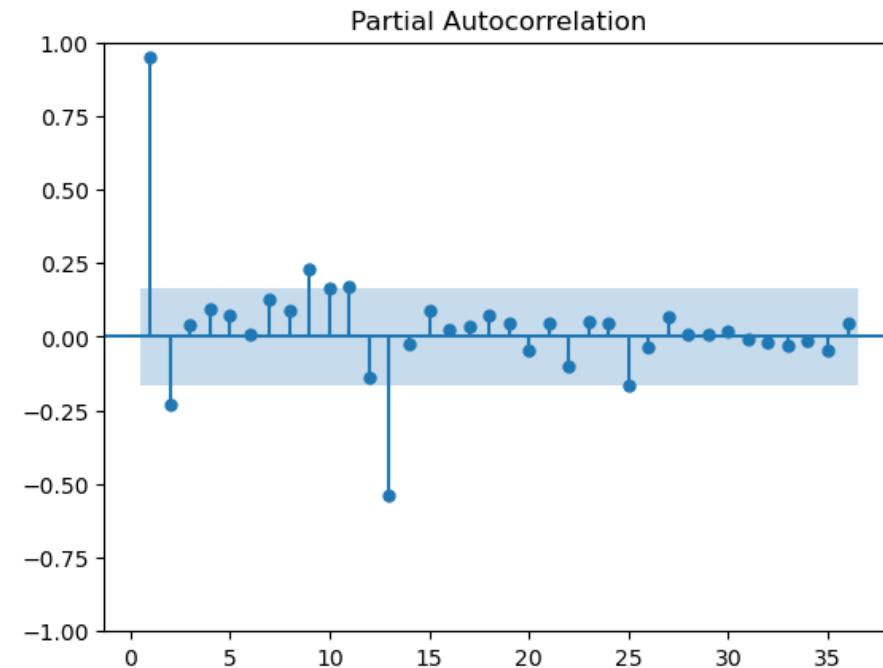
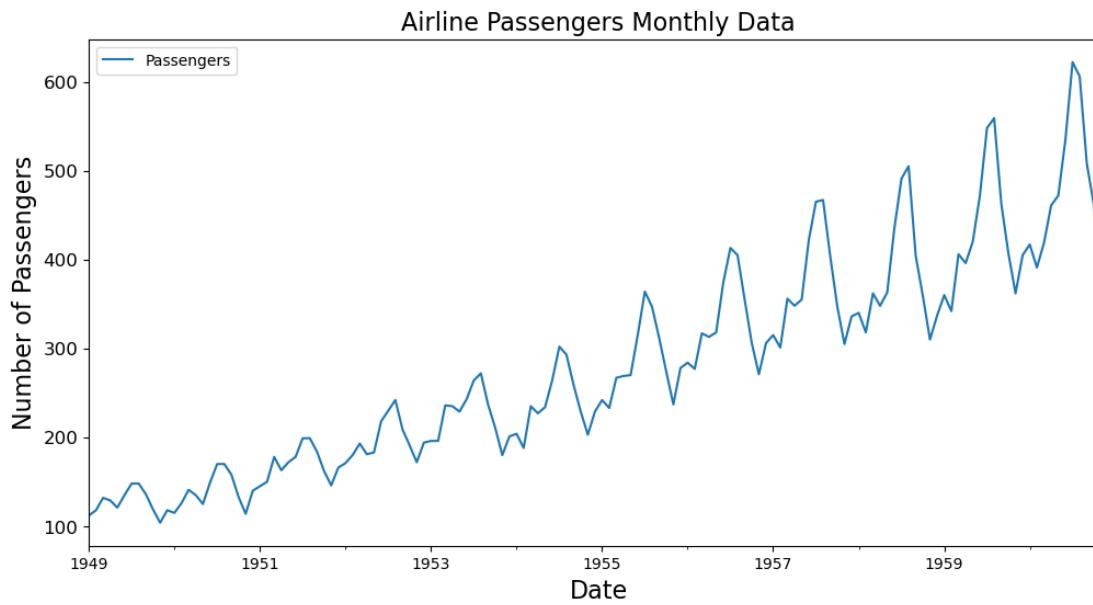
- ACF is a statistical tool that can be used to measure the autocorrelation of a time series.
- Autocorrelation (Serial correlation) is a measure of the correlation between a time series and a lagged version of itself.
- ACF shows the **degree to which the past values are predictive** of future values.



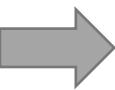


# PACF: Partial Autocorrelation Function

- PACF measures the direct correlation between time series and its lag, **controlling for intermediate lags**.
- $y_t$  and  $y_{t-2}$  might be correlated, simply because they are both connected to  $y_{t-1}$

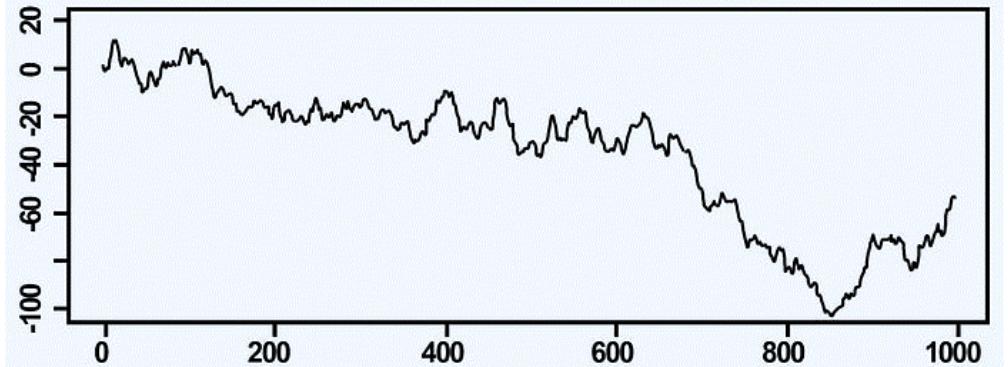
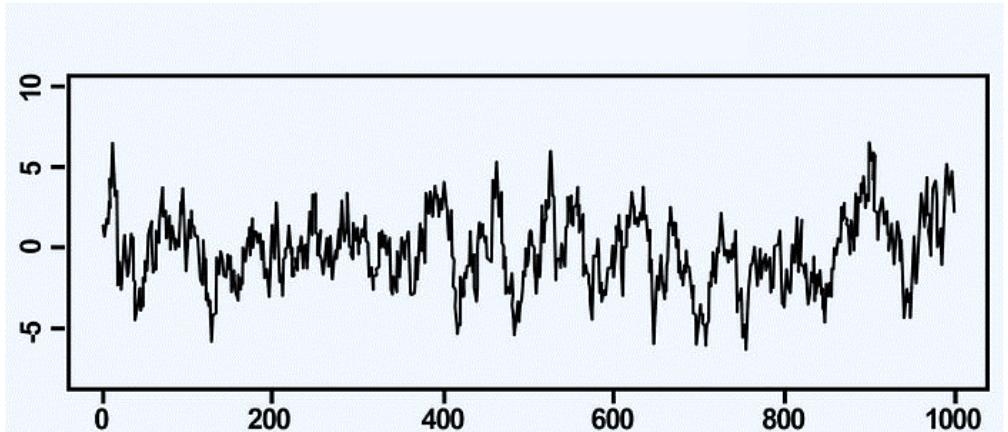


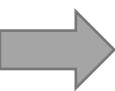
Pedram.Jahangiry



# Stationarity

- Stationary vs Non-Stationary Data. What makes a data set **Stationary**?
- In a stationary timeseries, the statistical properties **do not depend on the time**
- **Predictability (long-horizon)**: Stationary time series are easier to forecast because you can assume that future statistical properties will not change.
- This doesn't mean we cannot forecast non-stationary data!



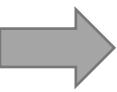


# Differencing

- Differencing: Computing the difference between consecutive observations.
- Differencing helps to stabilize the mean of a time series by removing changes in the level

Original data		1 <sup>st</sup> differenced		2 <sup>nd</sup> differenced	
Time1	10	Time1	NA	Time1	NA
Time2	12	Time2	2	Time2	NA
Time3	8	Time3	-4	Time3	-6
Time4	14	Time4	6	Time4	10
Time5	7	Time5	-7	Time5	-13





# Module 1- Part 1

## 1. Timeseries Basics

- What is sequence data?
- Time series tasks
- Trend, Seasonality, Residual
- ACF and PACF
- Stationarity

## 3. Modeling

- White Noise
- Econometrics
- Machine Learning
- Deep Learning

## 2. Forecasting strategies

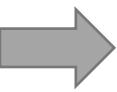
- One-step, Multi-step, Multi-output
- Univariate vs Multivariate
- Benchmarks

## 4. Can we beat wall street?

- Is stock price Random Walk? (EMH)
- Does Uni-variate Model work?
- Can Multi-variate Models help?
- So What?



Pedram.Jahangiry



# Module 1- Part 2

## 1. Timeseries Basics

- What is sequence data?
- Time series tasks
- Trend, Seasonality, Residual
- ACF and PACF
- Stationarity

## 3. Modeling

- White Noise
- Econometrics
- Machine Learning
- Deep Learning

## 2. Forecasting strategies

- One-step, Multi-step, Multi-output
- Univariate vs Multivariate
- Benchmarks

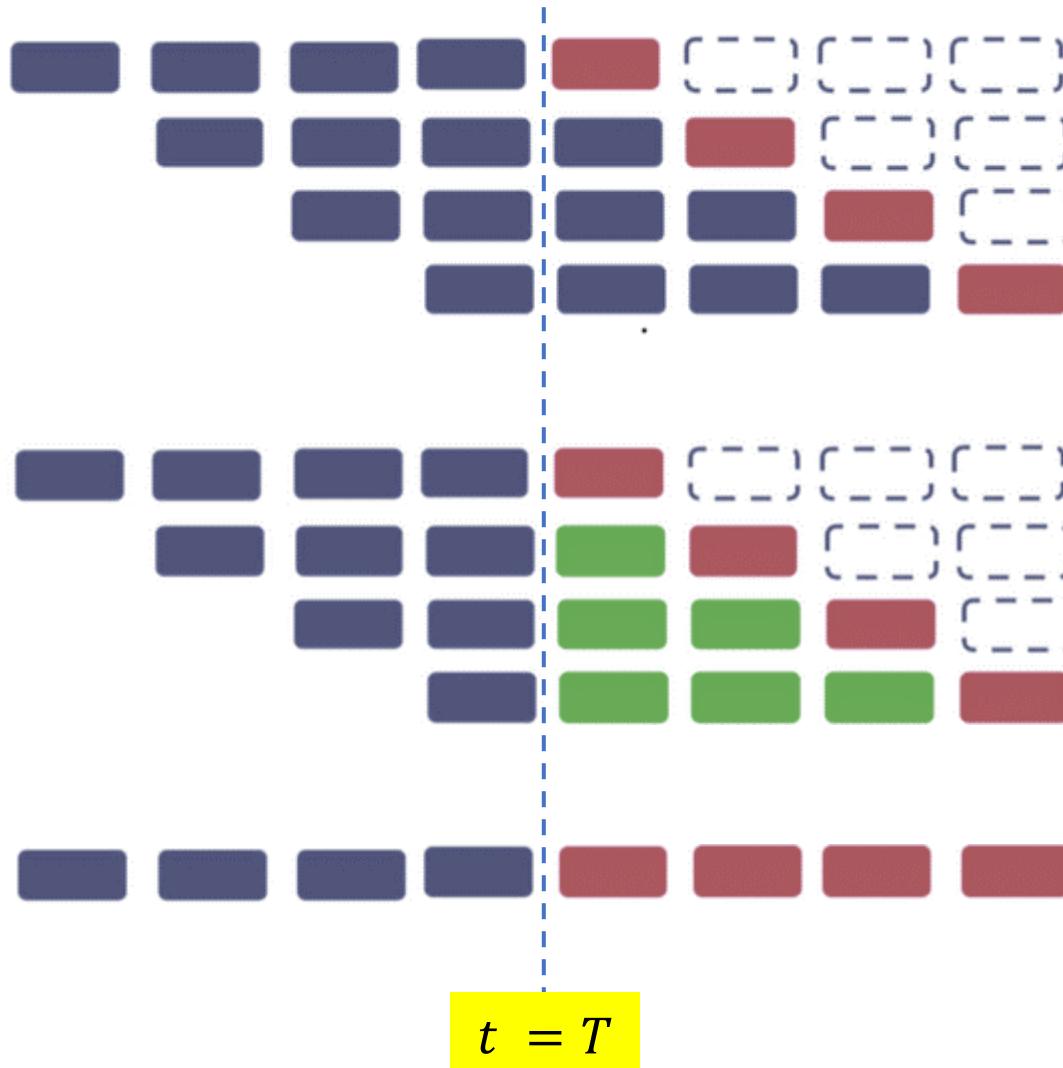
## 4. Can we beat wall street?

- Is stock price Random Walk? (EMH)
- Does Uni-variate Model work?
- Can Multi-variate Models help?
- So What?



Pedram.Jahangiry

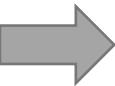
# One-step, Multi-step, Multi-output Forecasts



**Model comparison is  
highly sensitive to  
the chosen  
forecasting horizon.**

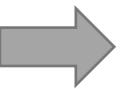


# Univariate vs Multivariate



- Using a single variable's historical data to predict future values for that same variable
- Incorporates multiple variables to forecast future values, considering the relationship between them.
- **Cannot** be used for **multi-step** forecasting
- Less stable in LR due to increased forecast error.





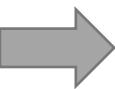
# Forecasting Benchmarks

- Some forecasting methods are **extremely simple** and **surprisingly effective**.

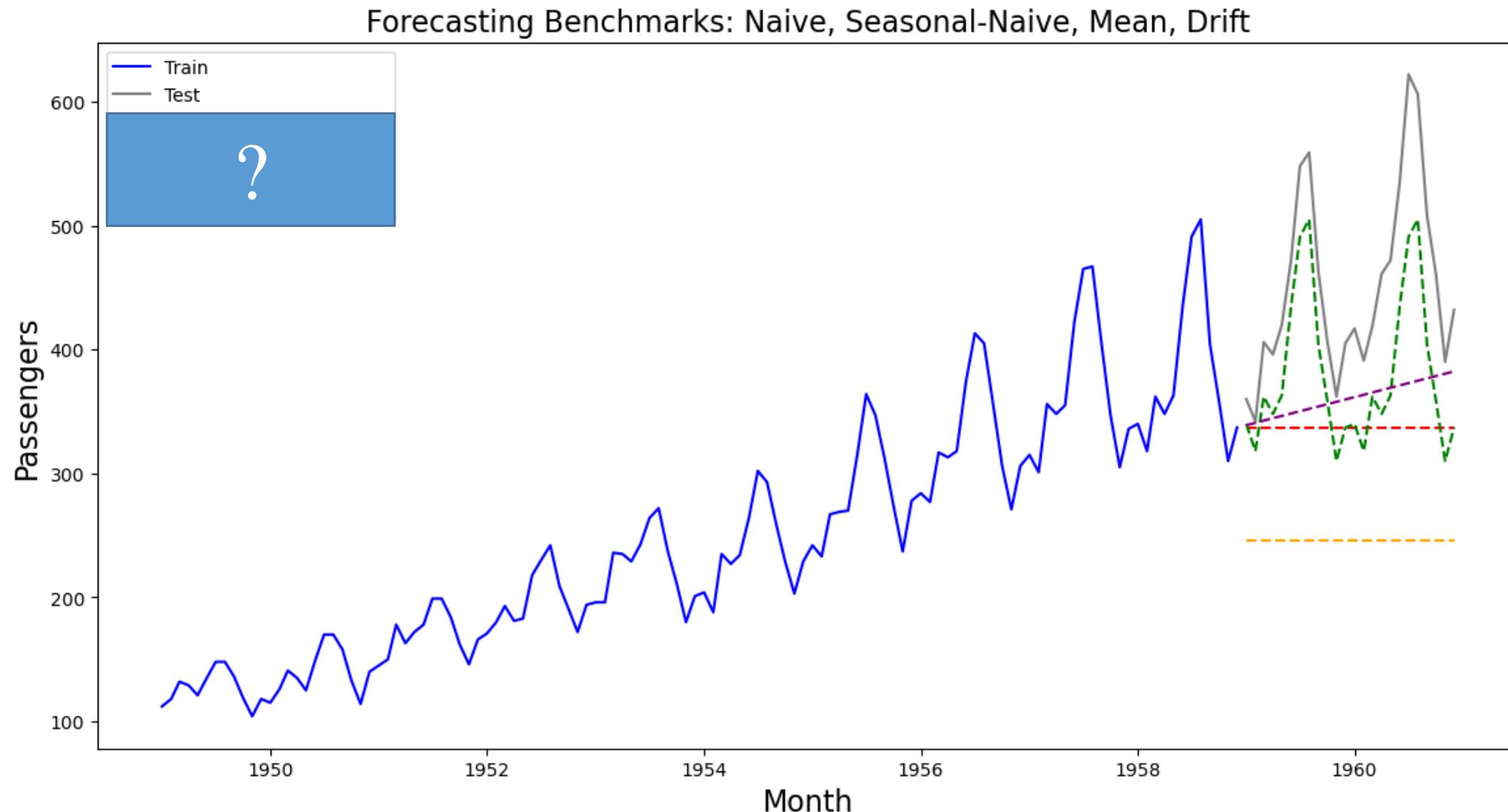
1. Mean method: Predicts all future values to be the average of the historical data.
2. Naïve method: Predicts the next value to be the current value
3. **Seasonal Naïve**: Predicts the next value by using the same value from the previous season
4. **Drift** method: Predicts future values by extending a linear trend fitted to the historical data.

- These methods will serve as benchmarks rather than the method of choice.
- If our model cannot beat the benchmark, **it is not** worth considering!!!

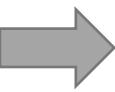




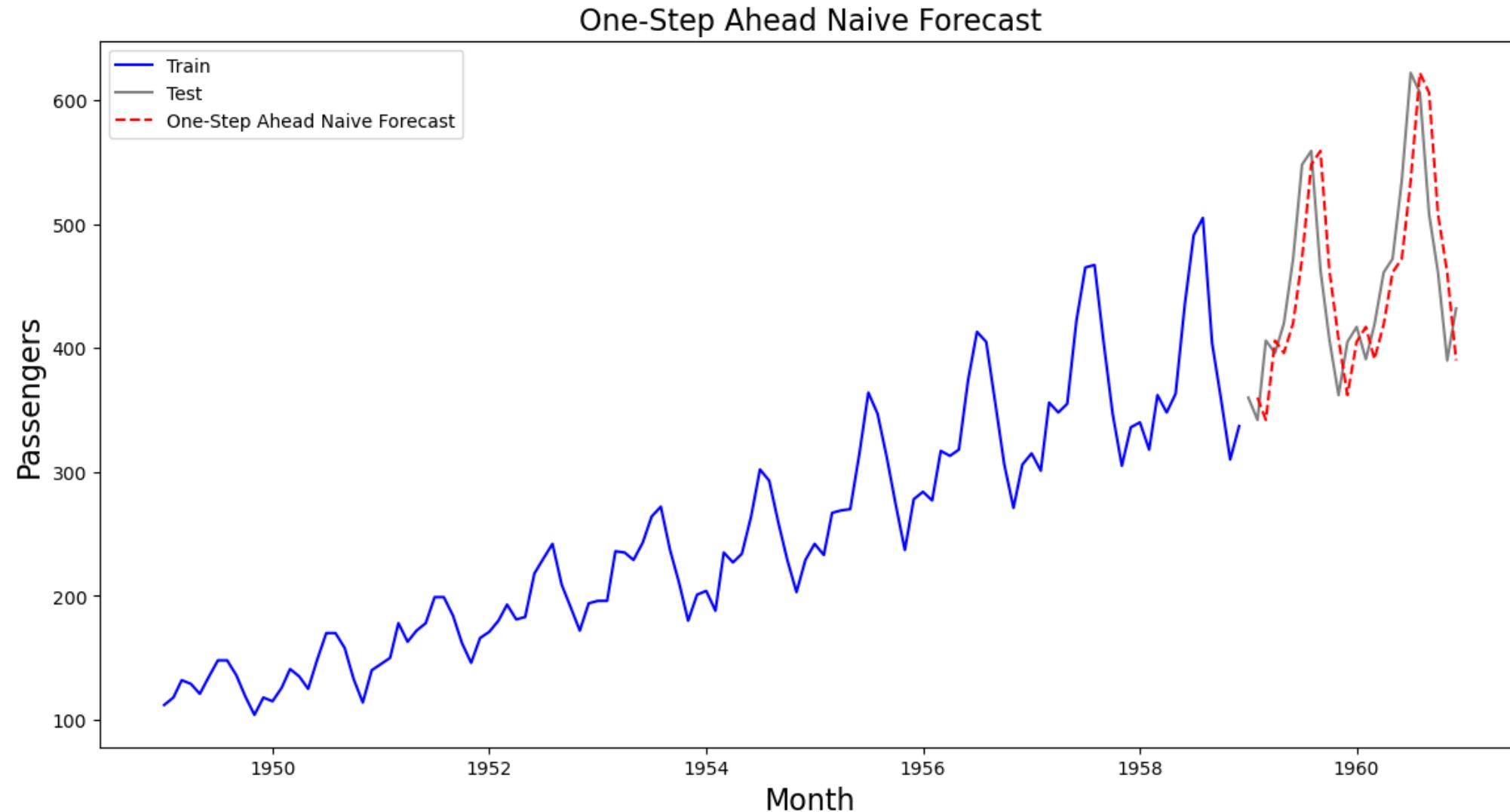
# Forecasting Benchmarks (Multi-Horizon)



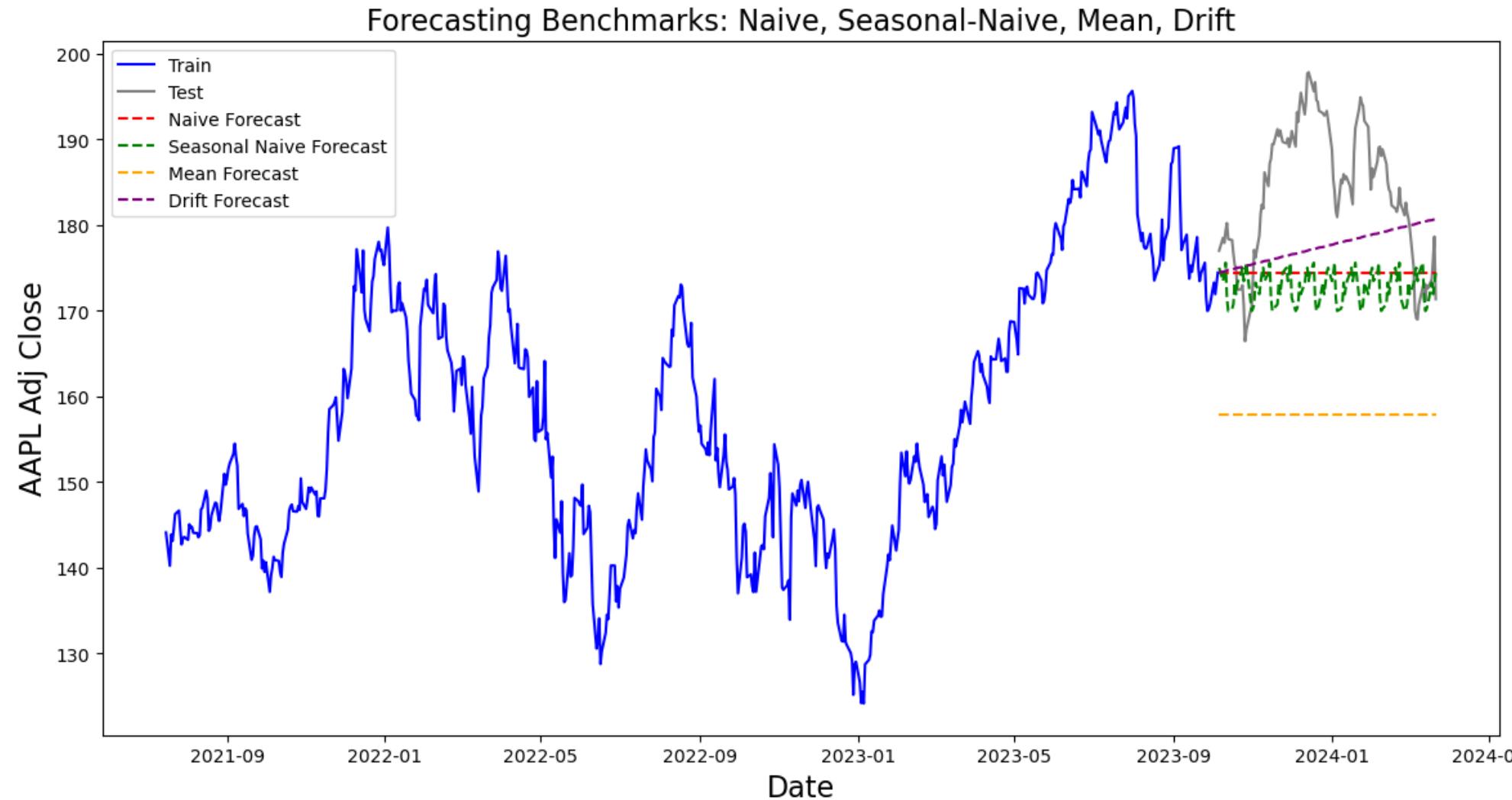
Pedram Jahangiry



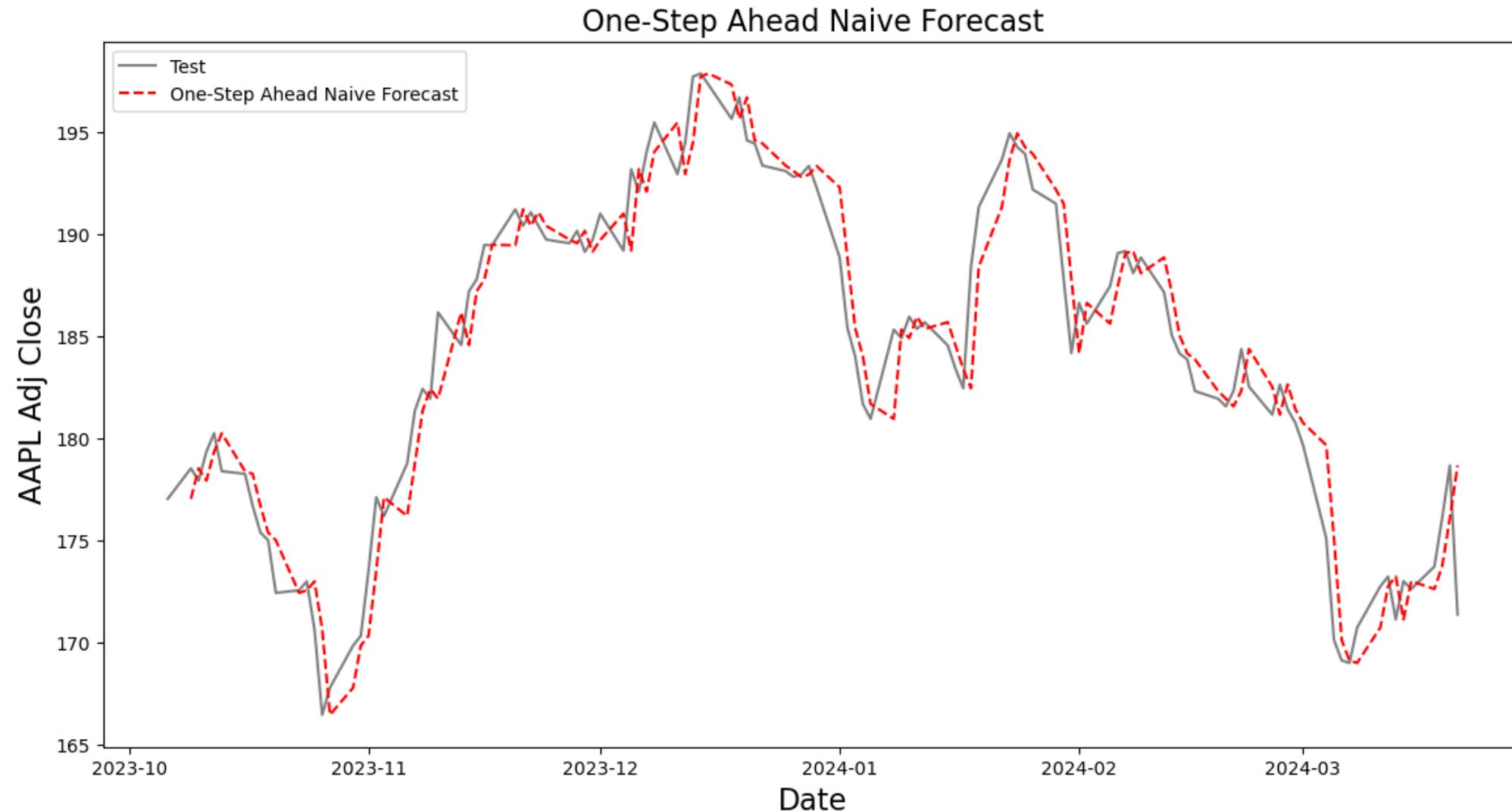
# Forecasting Benchmarks (One-step)

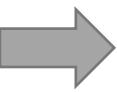


# Forecasting Benchmarks



# Forecasting Benchmarks





# Module 1- Part 3

## 1. Timeseries Basics

- What is sequence data?
- Time series tasks
- Trend, Seasonality, Residual
- ACF and PACF
- Stationarity

## 3. Modeling

- White Noise
- Econometrics
- Machine Learning
- Deep Learning

## 2. Forecasting strategies

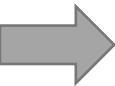
- One-step, Multi-step, Multi-output
- Univariate vs Multivariate
- Benchmarks

## 4. Can we beat wall street?

- Is stock price Random Walk? (EMH)
- Does Uni-variate Model work?
- Can Multi-variate Models help?
- So What?

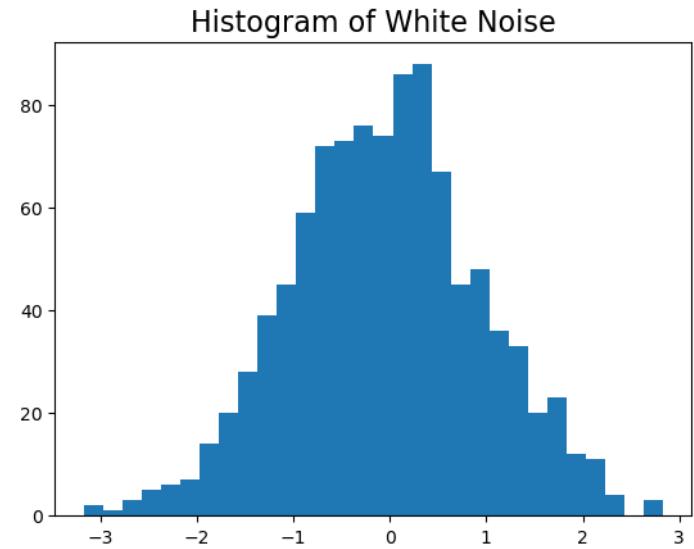
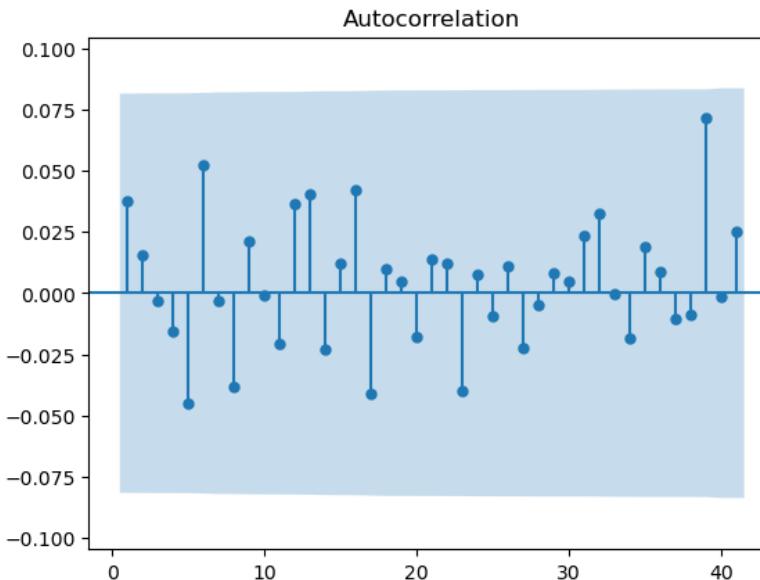
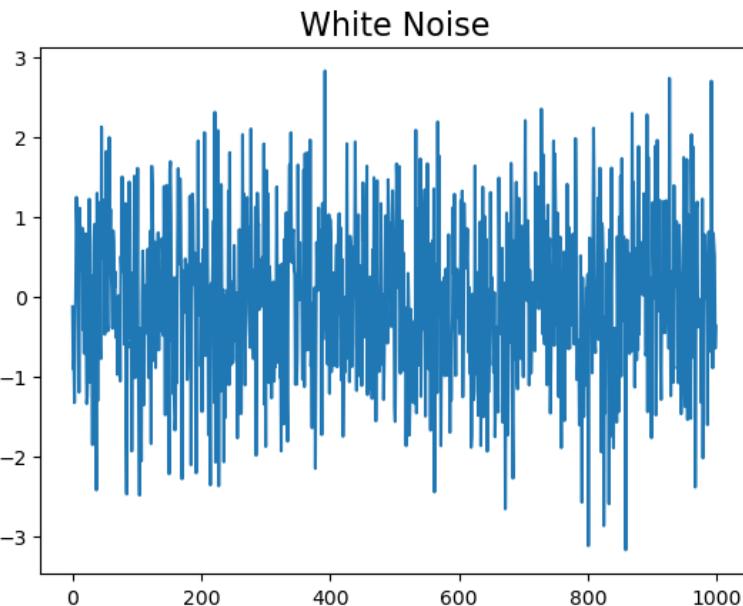


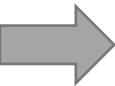
Pedram.Jahangiry



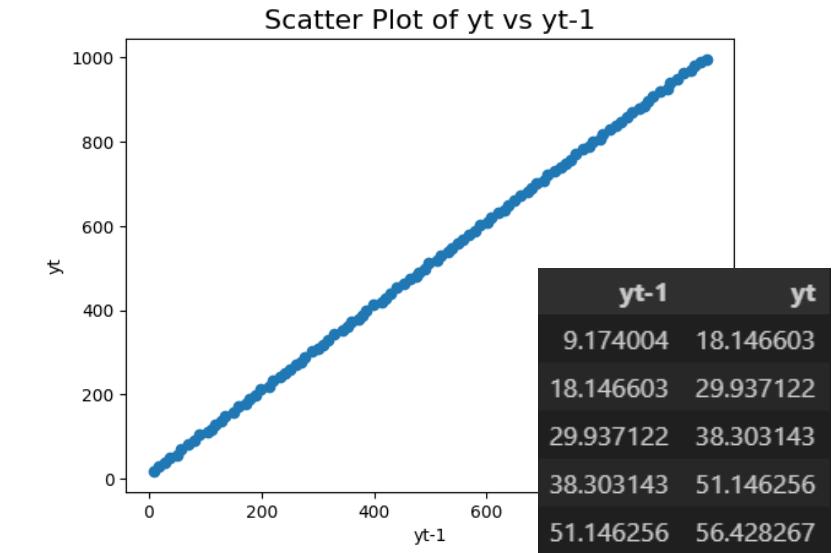
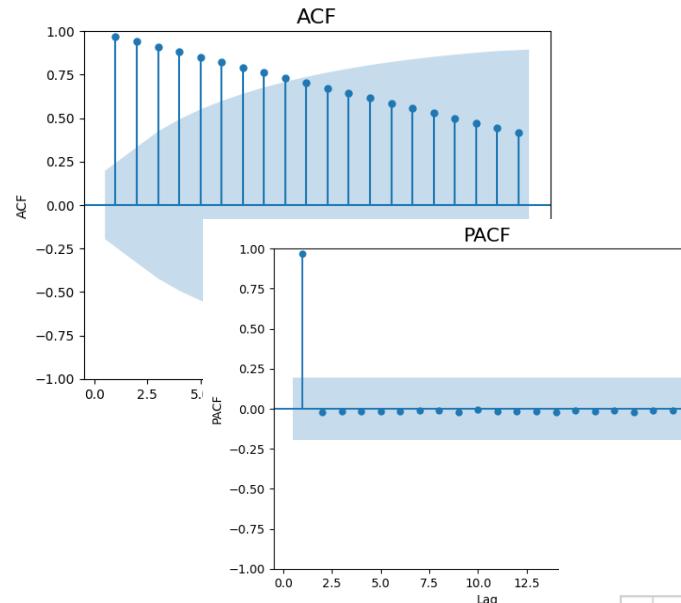
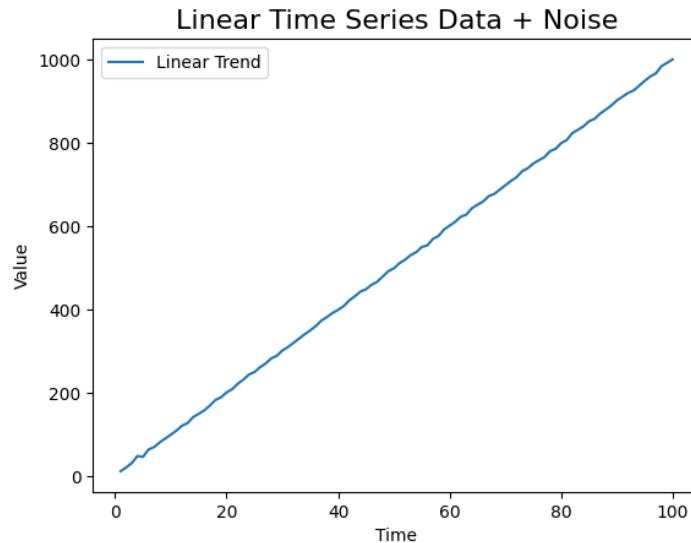
# The Importance of White Noise in Modeling

- White noise refers to a random process with:
- Zero mean, Constant variance and No autocorrelation.
- Role in timeseries modeling: **Diagnostic tool!** Indicator of a good fit!

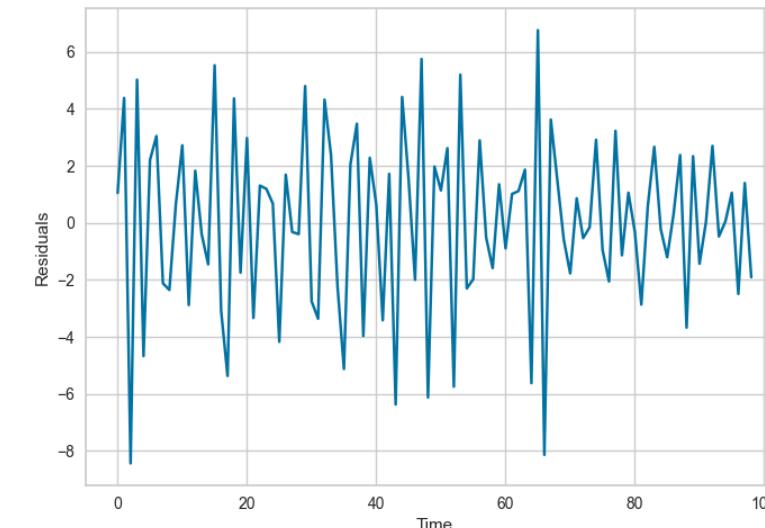




# Linear vs Non-Linear (data or model?)



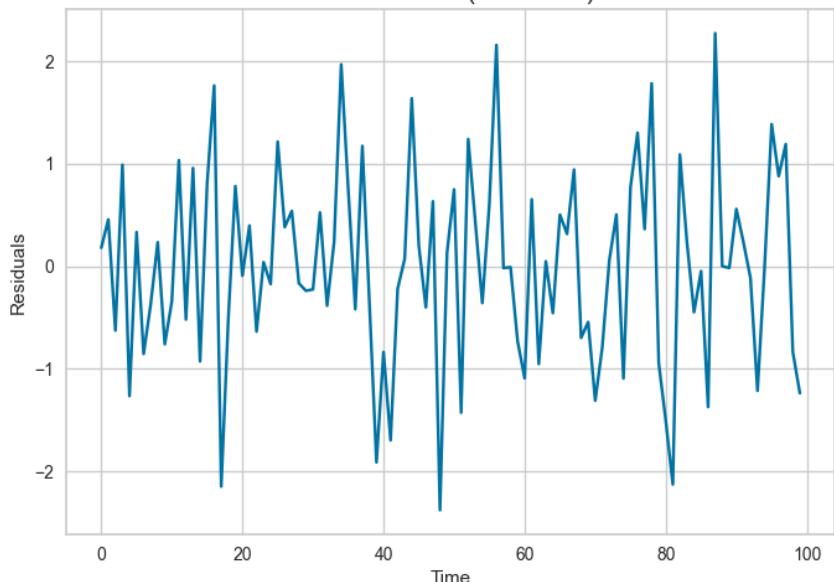
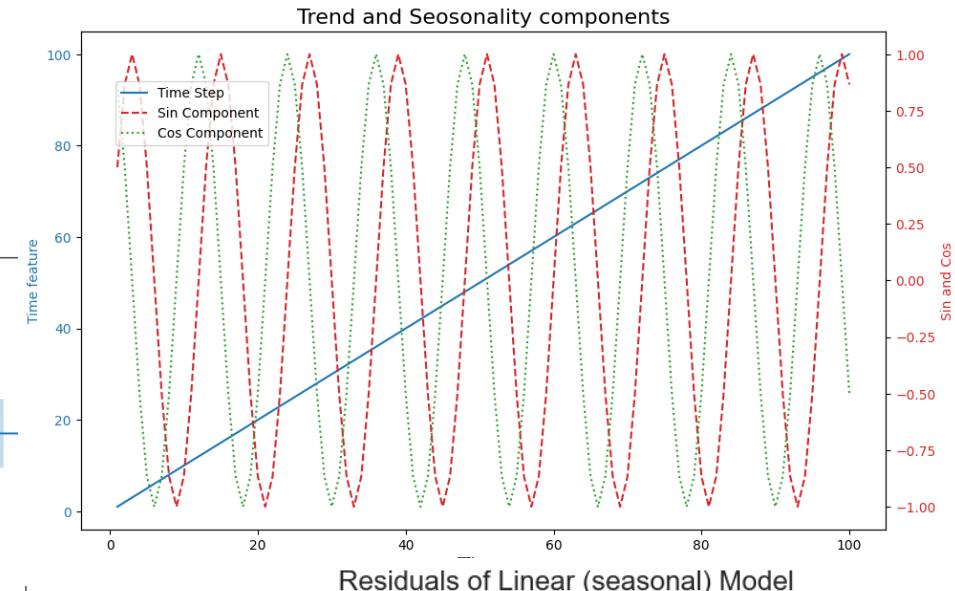
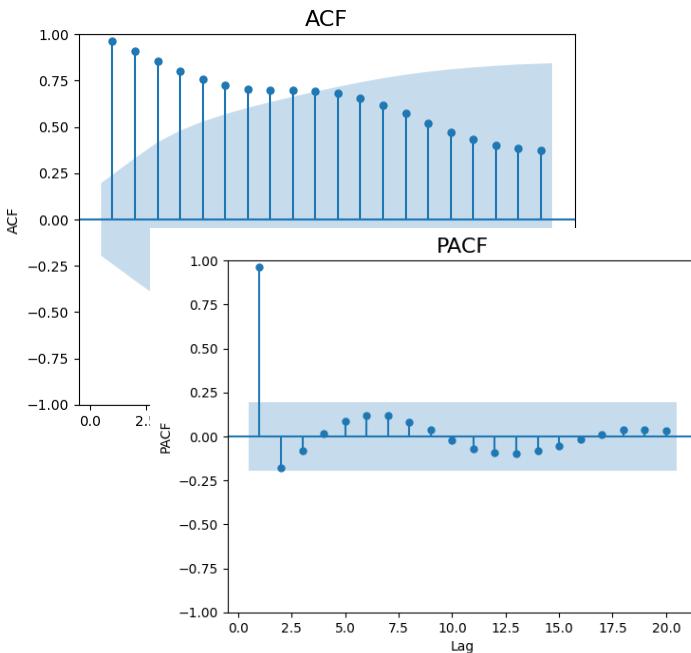
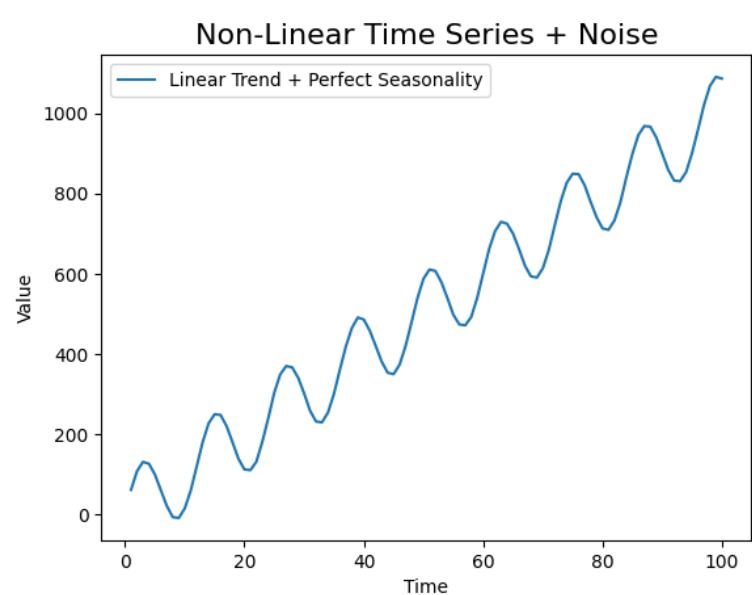
Residuals of Linear Model



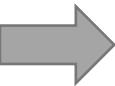
- Can you think of a simple linear model?
- True model:  $y_t = 10 + y_{t-1} + \epsilon_t$



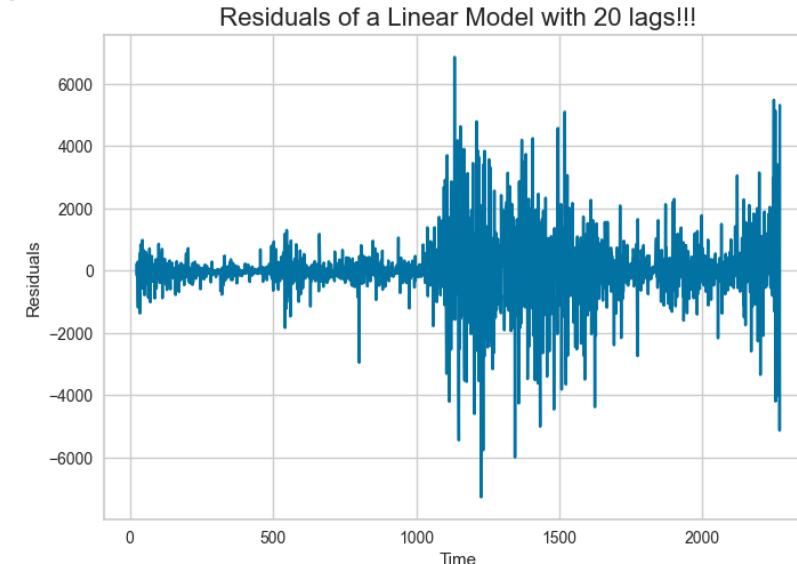
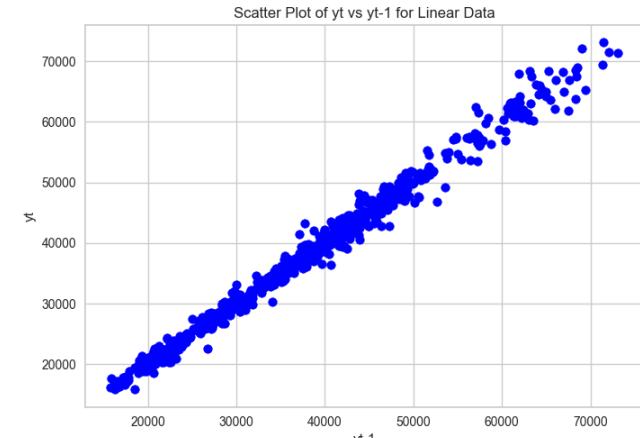
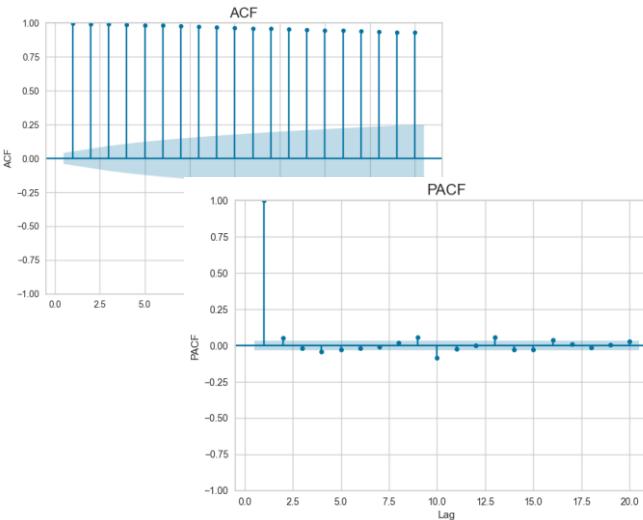
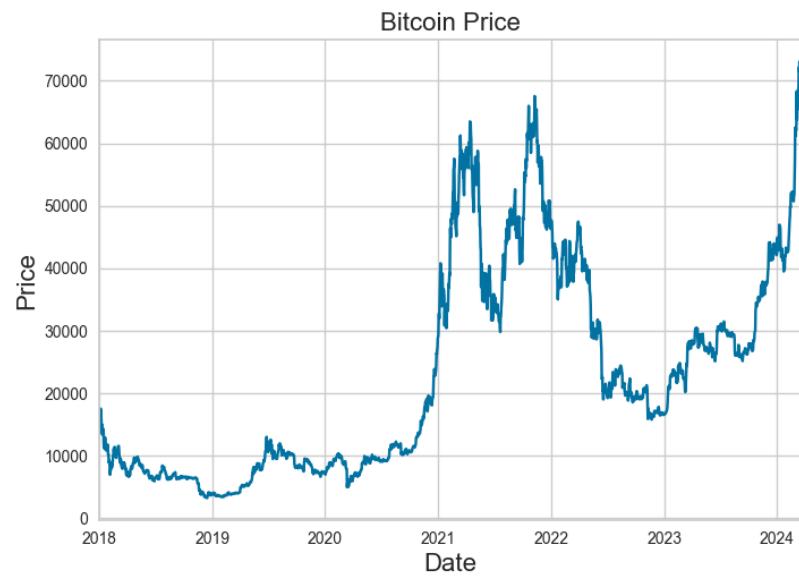
# Linear vs Non-Linear (data or model?)



- Can you think of a simple linear model?
- **True model:**  $y_t = 10t + 100\sin(..) + \epsilon_t$

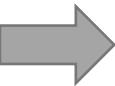


# Linear vs Non-Linear (data or model?)



- Can you think of a simple linear model?
- True model:  $y_t = I \text{ wish :)$





# Basic steps in a forecasting task

## Step 1: Problem definition

- Forecasting **type** and **horizon** (one-step, multi-step, multi-output forecasts), ...

## Step 2: Data Collection

- Time horizon, structural changes, **data type**, ...

## Step 3: Exploratory Analysis

- Trend, seasonality, outliers, ...

## Step 4: Model Selection and Training

- Traditional **vs** machine learning **vs** deep learning

## Step 5: Model Evaluation and Comparison

- MSE, RMSE, MAE, MAPE, ...



# → Modeling, a simple example

- Quantifying airline passengers!
- Let's build a model:
  - Cross sectional vs timeseries?
  - Univariate vs Multivariate?

$$\text{air passengers} = \beta_0 + \beta_1 gdp + \beta_2 unemp + \beta_3 saving + \dots + u$$

- ✓ Can you **interpret** this model?
- ✓ Can you make **predictions** using your model?
- ✓ Can you make **forecasts** into the future? What are the challenges?



# → Modeling, a simple example (cont'd)

- Focusing on Forecasting for one country:

$$\text{air passengers}_t = \beta_0 + \beta_1 \text{gdp}_t + \beta_2 \text{unemp}_t + \beta_3 \text{saving}_t + \dots + u_t$$

- This **explanatory** model is contemporaneous (static). How can we make it a **dynamic forecasting model**?

$$\text{air passengers}_{t+1} = \beta_0 + \beta_1 \text{gdp}_t + \beta_2 \text{unemp}_t + \beta_3 \text{saving}_t + \dots + u_t$$

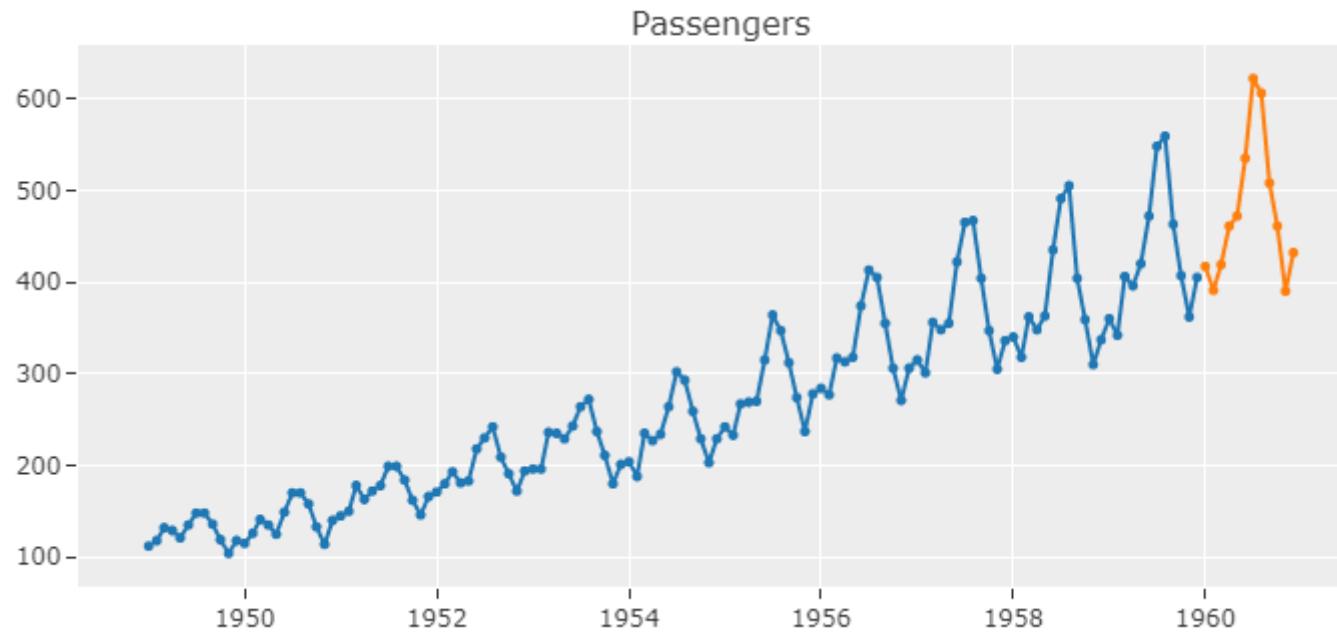
- Assumption: everything is reflected in passengers' number already! So, why not →

$$\text{passengers}_{t+1} = f(\text{passengers}_t, \text{passengers}_{t-1}, \dots)$$



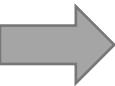
# → Modeling, a simple example (cont'd)

$$\text{passengers}_{t+1} = f(\text{passengers}_t, \text{passengers}_{t-1}, \dots)$$



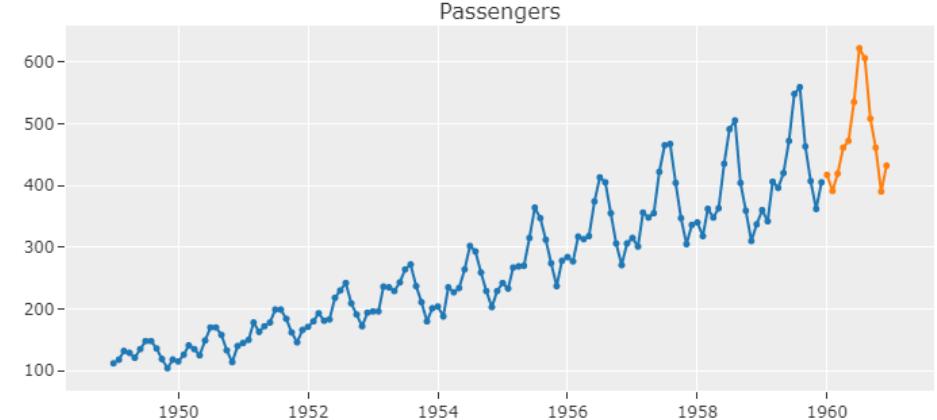
- Goal: Identify **existing patterns** in data to project **future patterns!**





# Why do we need ML/DL?

$$\text{passengers}_{t+1} = f(\text{passengers}_t, \text{passengers}_{t-1}, \dots)$$



- How do we identify the functional form of  $f( )$ ?
- How about this:

$$\text{passengers}_{t+1} = c + \phi_1 \text{passengers}_t + \phi_2 \text{passengers}_{t-1} + \dots + \epsilon_t$$

- What if we are looking for **more complex** relationships?
- Or a **non-parametric** approach?



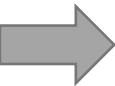
# → Predictive modeling vs forecasting?

$$\text{passengers} = \beta_0 + \beta_1 gdp + \beta_2 unemp + \beta_3 saving + \dots + u$$

	Predictive Modeling	Forecasting
Definition	Creating models to analyze data and predict an outcome.	Making informed predictions about future events based on past data.
Focus	Any point in time (past, present, future)	Primarily focused on future outcomes.
Variables	Multivariate	Typically, univariate

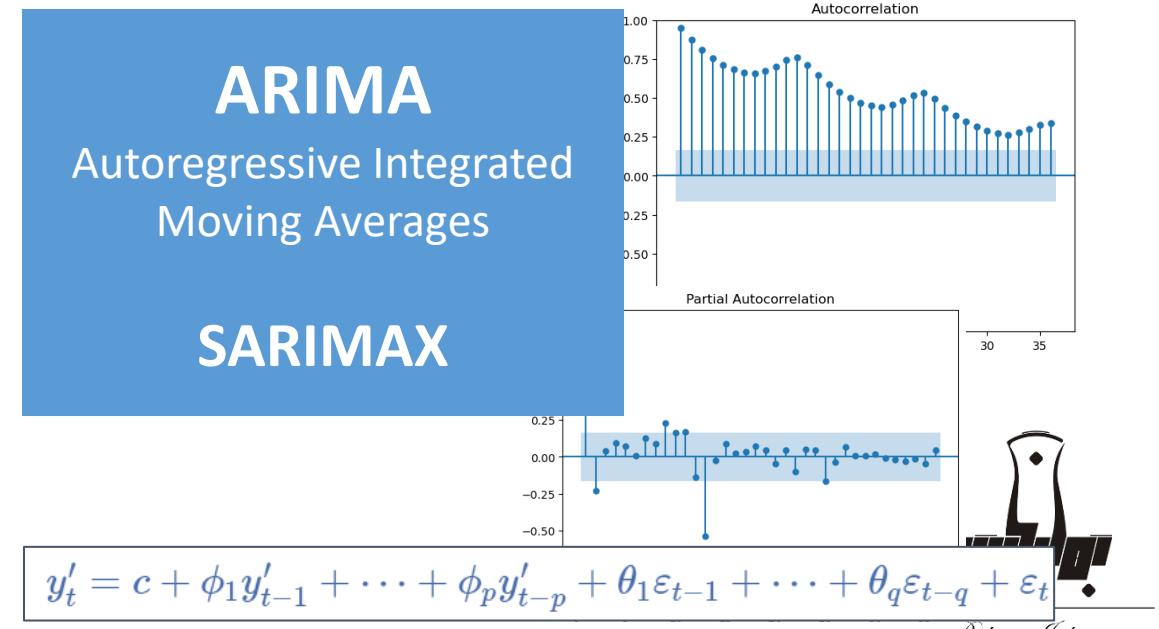
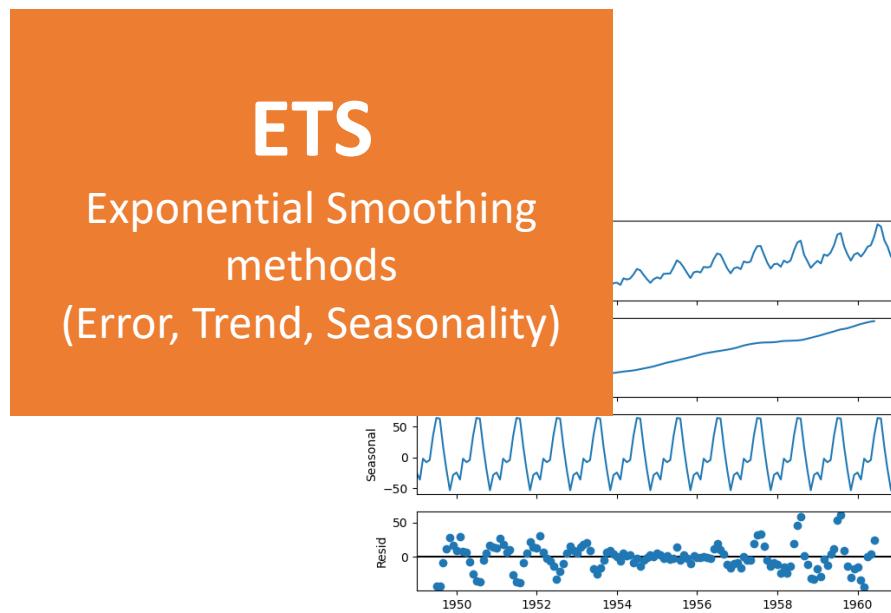
- Forecasting is a subset of predictive modeling.
- We can use econometrics, ML or DL for predictive modeling.

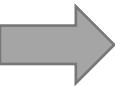




# Econometrics

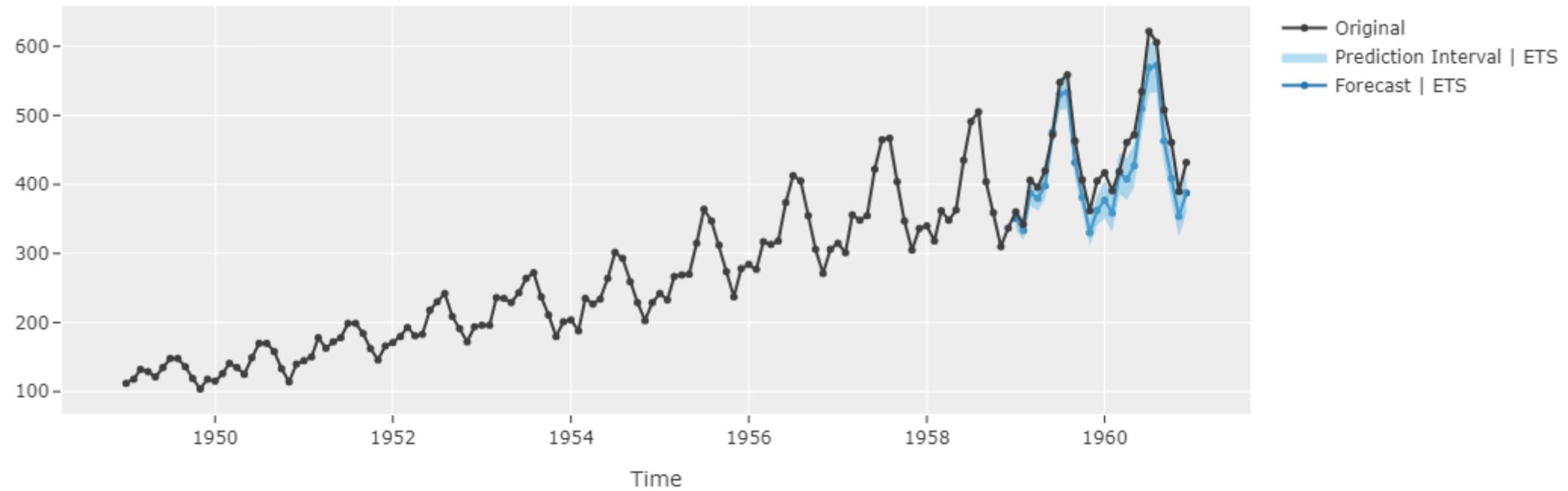
- **Decomposition-Based:** Focuses on modeling through **decomposition** and **autocorrelations**.
- **Parametric Structure:** Defined mathematical form; **interpretable** components.
- **Statistical Tools:** Confidence intervals, hypothesis testing for insights.



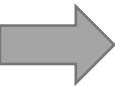


# Econometrics Example: ETS

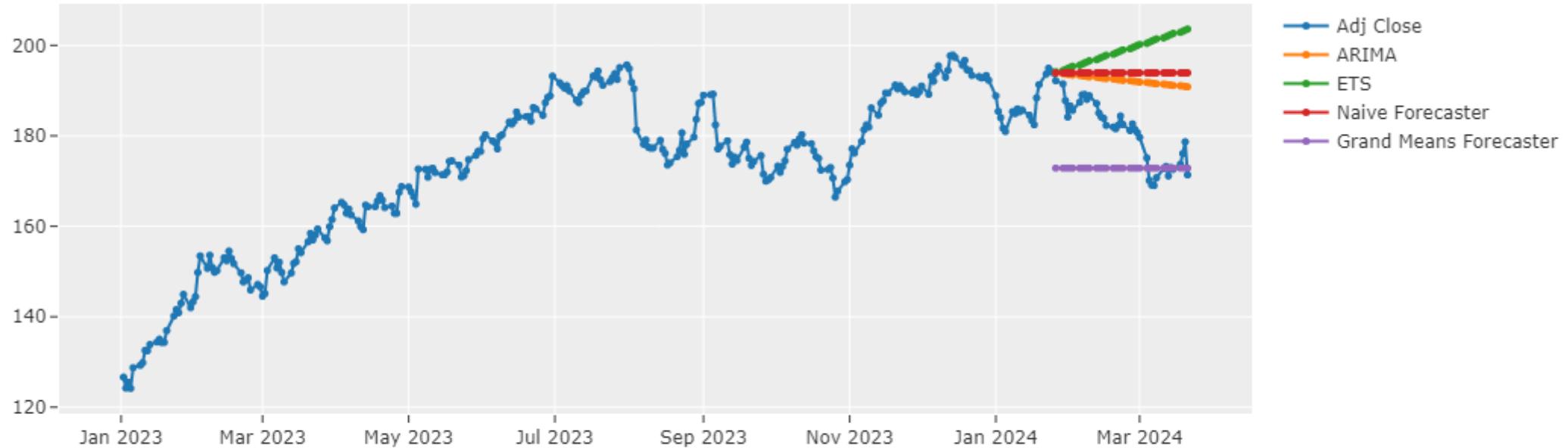
- ETS decomposes time series into Error, Trend, and Seasonal components.



Pedram.Jahangiry

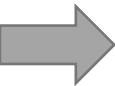


# Econometrics Limitations



- Cannot handle strong, complex non-linear relationships
- **Transition to Machine Learning**: ML/DL models are inherently designed to capture **complex non-linear patterns** in time series





# Machine Learning

- **Data Transformation:** preprocessing of time series data into a **supervised learning format**, creating features and a target variable.
- **Flexibility:** Can capture **complex nonlinear relationships** without explicit mathematical models for the underlying processes.
- **Feature Importance:** Offers insights into which features (lags, external variables) are most predictive.
- **Challenges:** Overfitting, plus Requires special cross validation and bootstrapping



## Supervised

- Regression
- Classification



## Unsupervised

- Clustering
- Anomaly detection

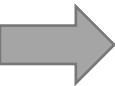


## Semi-supervised

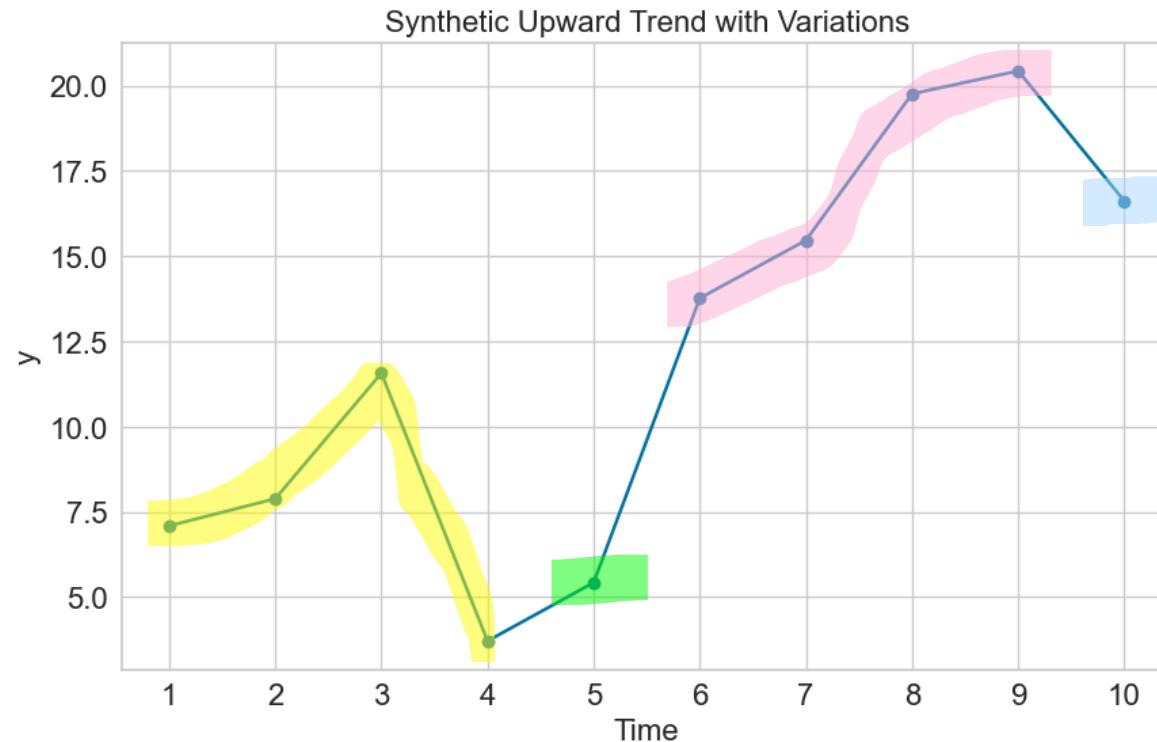


## Self-supervised





# ML Data Transformation (Single-output)



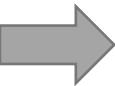
Features (X)				Target (y)
$y_{t-4}$	$y_{t-3}$	$y_{t-2}$	$y_{t-1}$	$y_t$
$y_1$	$y_2$	$y_3$	$y_4$	$y_5$
$y_2$	$y_3$	$y_4$	$y_5$	$y_6$
$y_3$	$y_4$	$y_5$	$y_6$	$y_7$
$y_4$	$y_5$	$y_6$	$y_7$	$y_8$
$y_5$	$y_6$	$y_7$	$y_8$	$y_9$
$y_6$	$y_7$	$y_8$	$y_9$	$y_{10}$

TS raw data

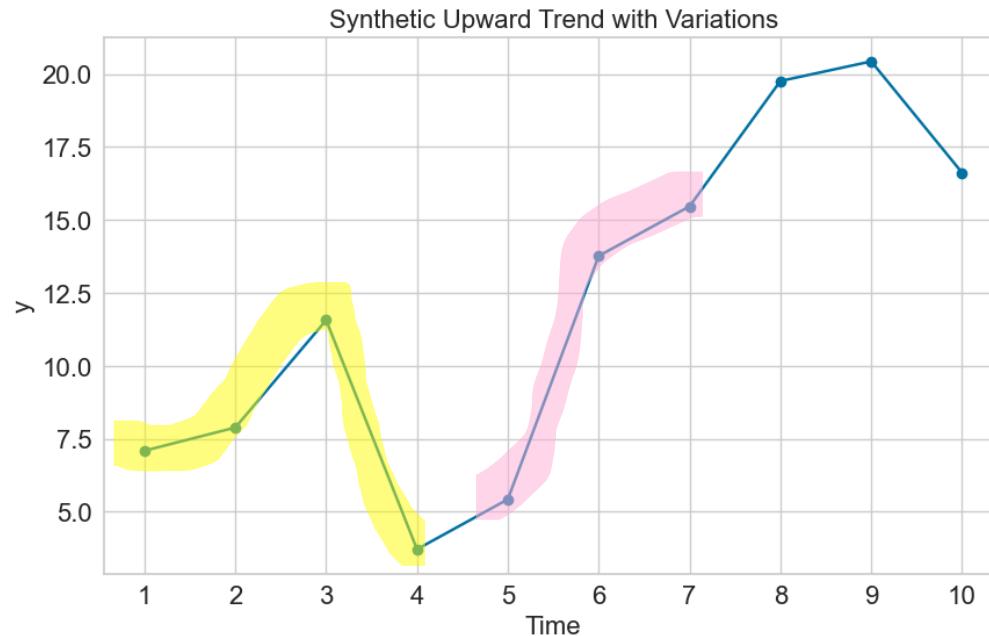
$y_1$	$y_2$	$y_3$	$y_4$	$y_5$	$y_6$	$y_7$	$y_8$	$y_9$	$y_{10}$
-------	-------	-------	-------	-------	-------	-------	-------	-------	----------

→ TS Supervised data

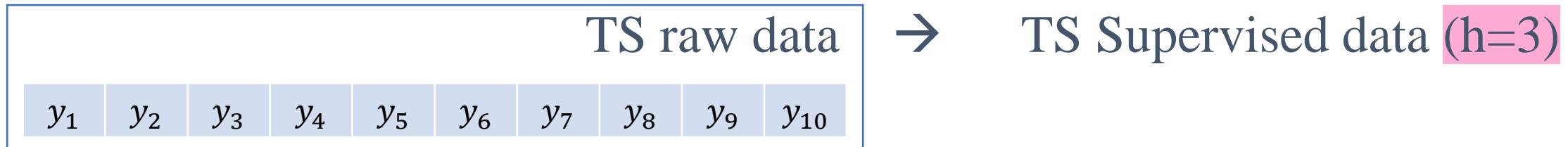


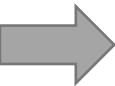


# ML Data Transformation (Multi-output)



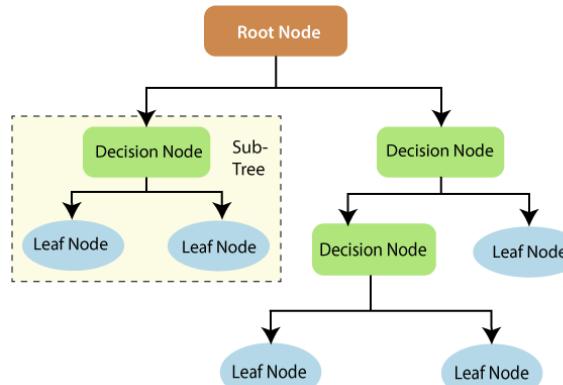
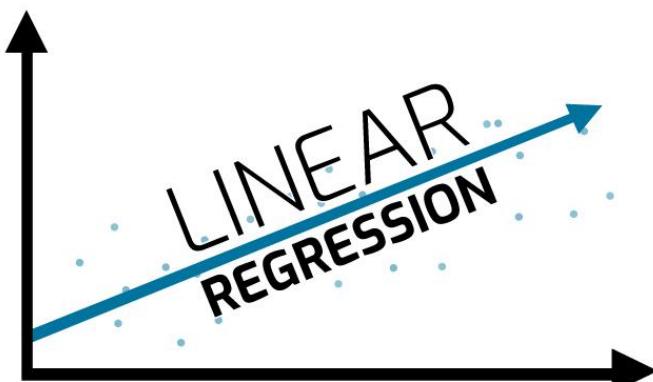
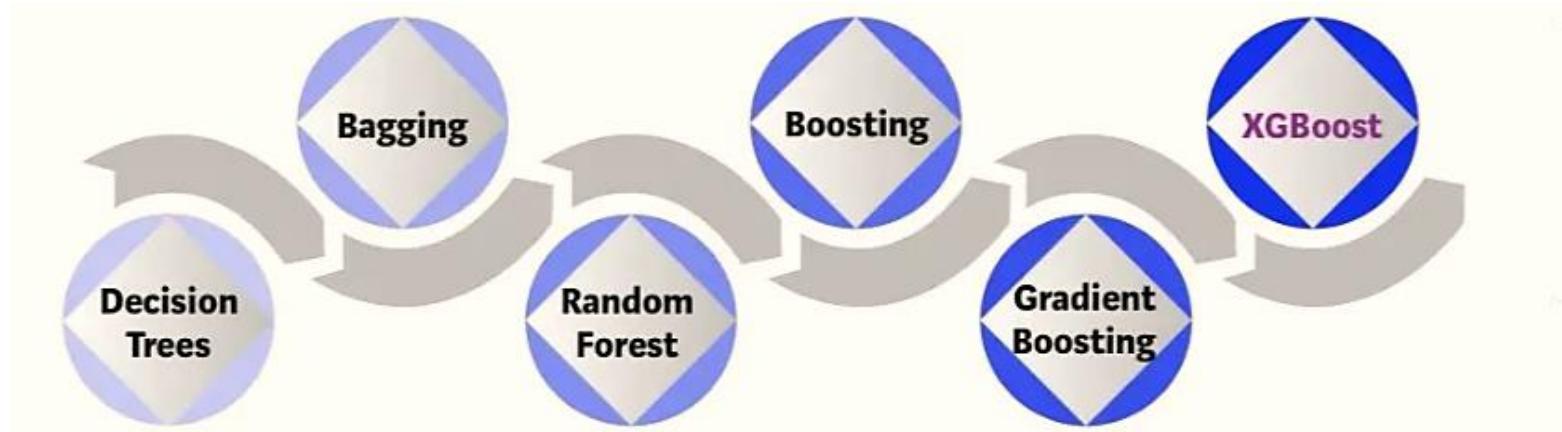
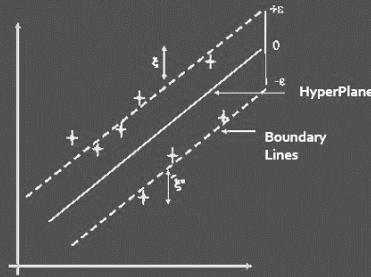
Features (X)				Target (y)		
$y_{t-4}$	$y_{t-3}$	$y_{t-2}$	$y_{t-1}$	$y_t$	$y_{t+1}$	$y_{t+2}$
$y_1$	$y_2$	$y_3$	$y_4$	$y_5$	$y_6$	$y_7$
$y_2$	$y_3$	$y_4$	$y_5$	$y_6$	$y_7$	$y_8$
$y_3$	$y_4$	$y_5$	$y_6$	$y_7$	$y_8$	$y_9$
$y_4$	$y_5$	$y_6$	$y_7$	$y_8$	$y_9$	$y_{10}$





# Machine Learning Models

## Support Vector Regression



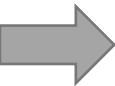
*dmlc*  
**XGBoost**

 **LightGBM**

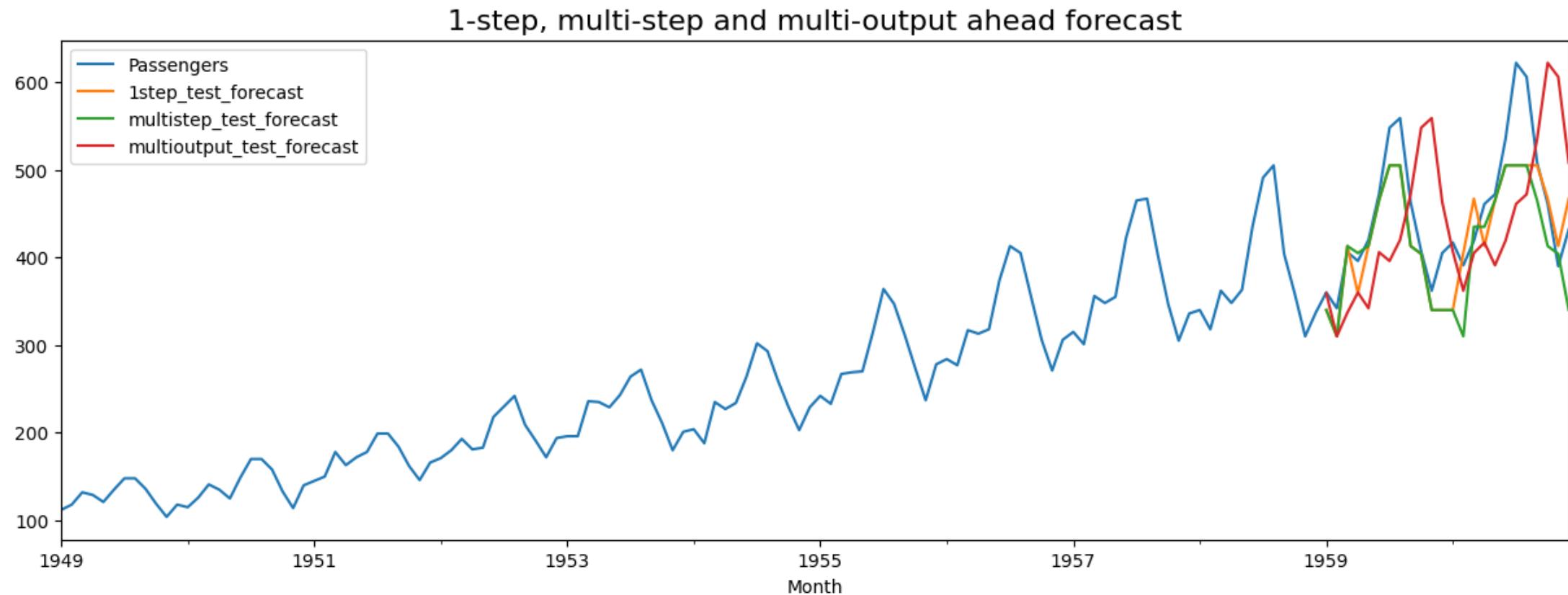
 **CatBoost**



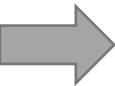
Pedram.Jahangiry



# ML Example: Random Forest with 12 lags

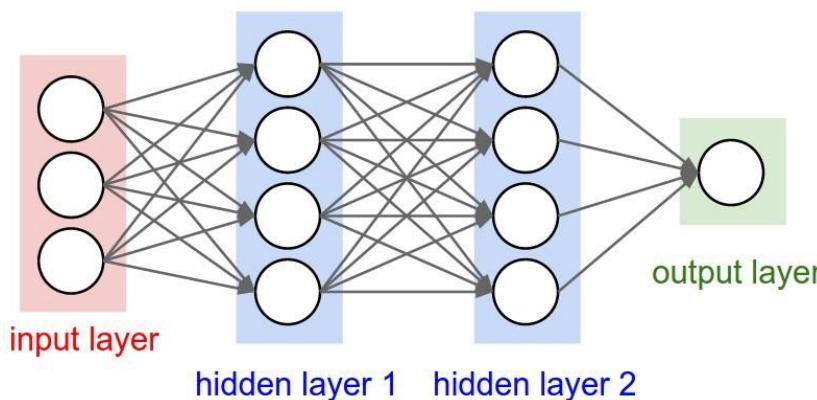


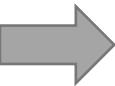
Pedram Jahangiry



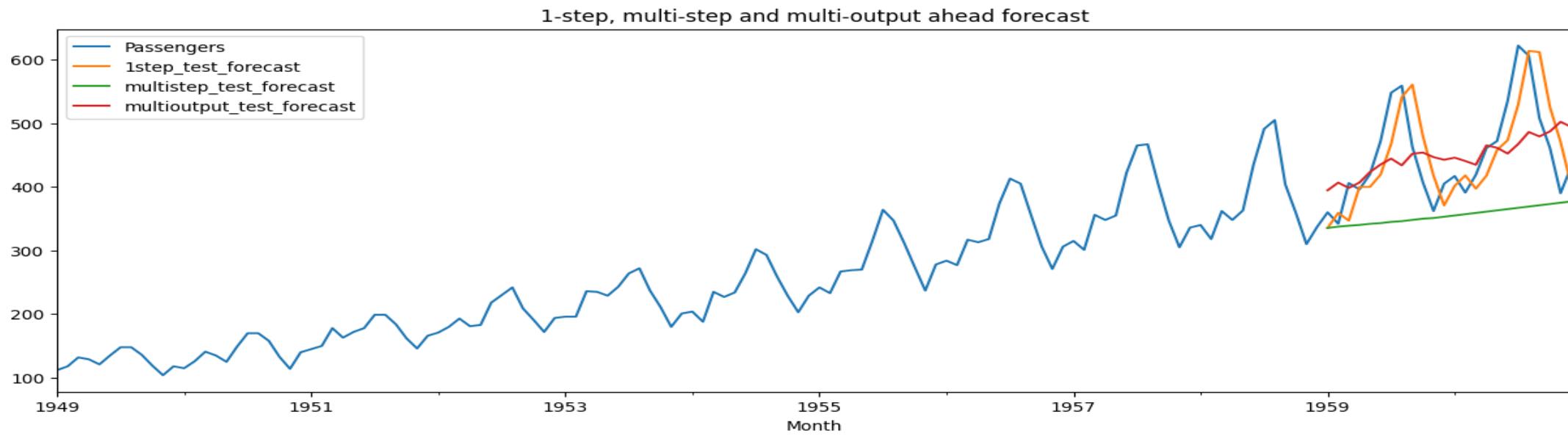
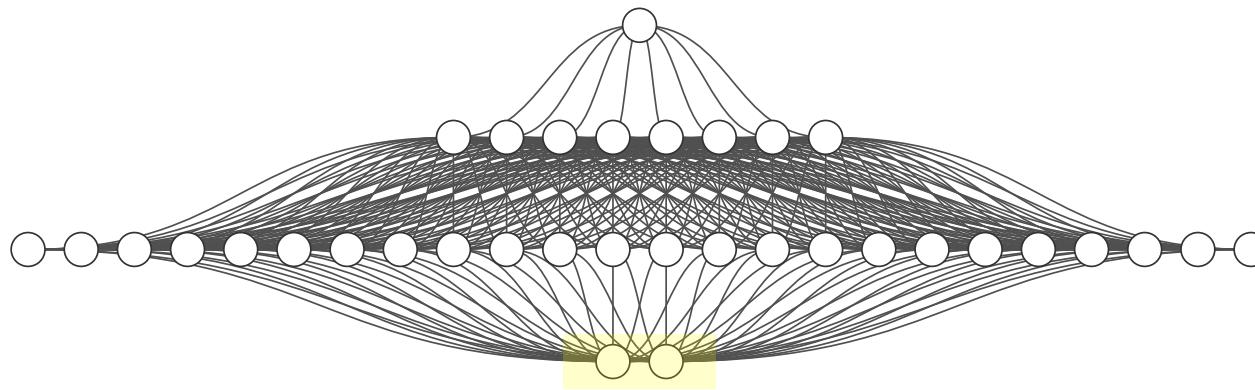
# Deep Learning

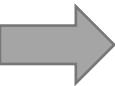
- Just like ML, DL models require Data Transformation.
- **Non-Parametric Nature:** DL models do not assume a predefined form for the time series model, allowing them to learn complex and highly nonlinear patterns directly from data.
- **High Capacity and Flexibility:** Can model very complex relationships and interactions, including multivariate and high-dimensional time series. Well-suited for large datasets.



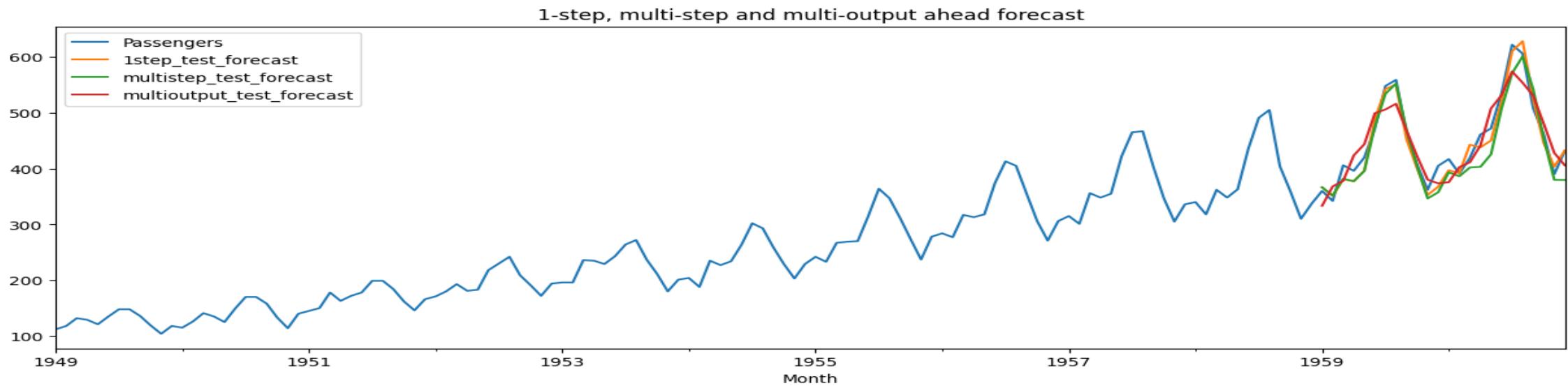
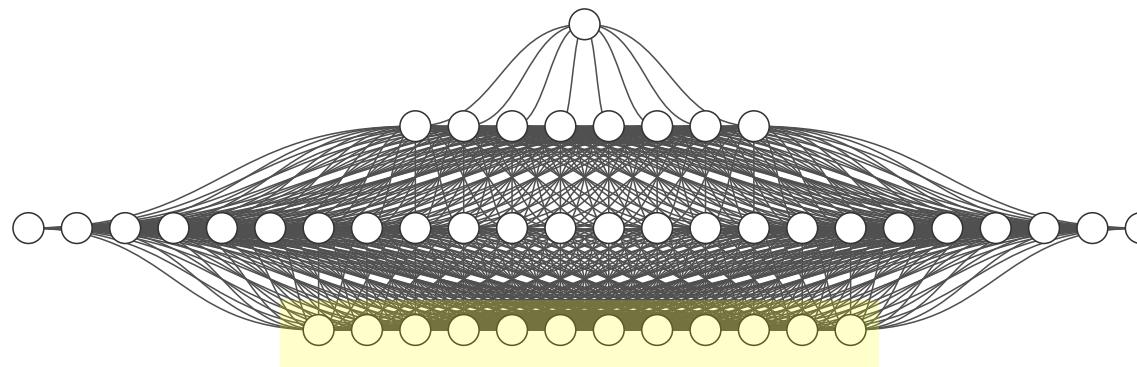


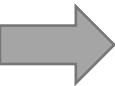
# DL Example: DNN with 2 lags





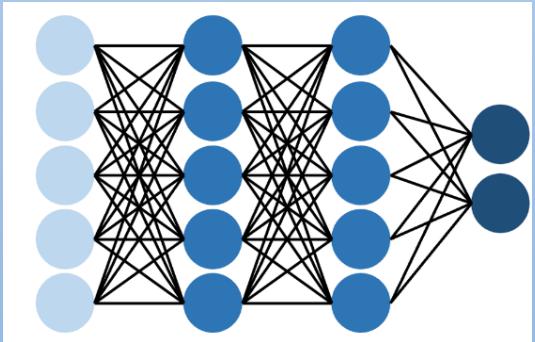
# DL Example: DNN with 12 lags



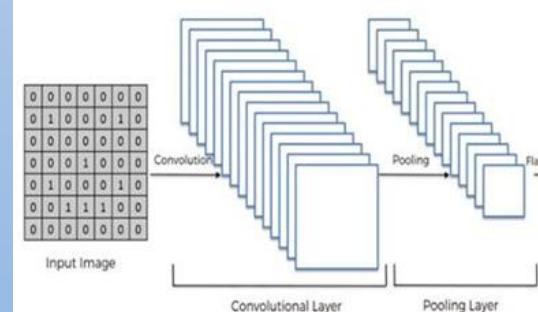


# Deep Learning Architectures

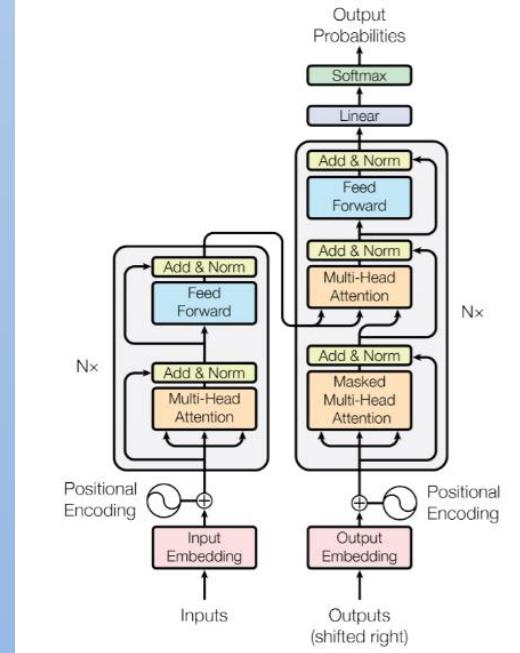
Standard NN



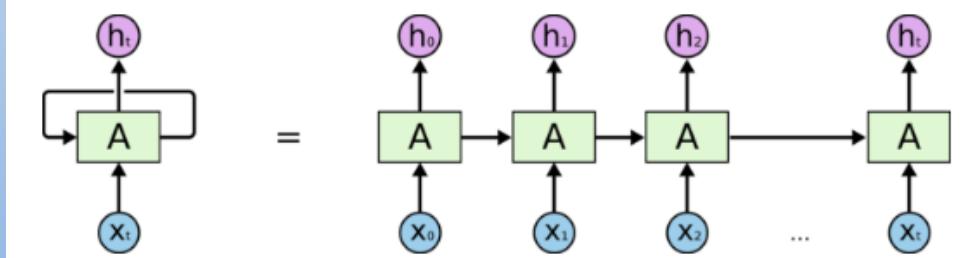
Convolutional NN



Transformers



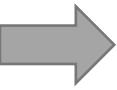
Recurrent  
NN



# Modeling Comparison

Aspect	Econometric Models	Machine Learning (ML)	Deep Learning (DL)
Feature Engineering	Requires <b>explicit modeling</b> of seasonality and trend	Captures complex patterns with <b>less need for manual engineering</b>	Often <b>automates feature engineering</b>
Model Complexity	<b>Lower</b> ; focuses on data generation process	<b>Moderate</b> ; can handle non-linear interactions	<b>High</b> ; suited for complex and high-dimensional data
Interpretability	<b>High</b> ; interpretable parameters, statistical tests	<b>Moderate</b> ; provides feature importance	<b>Low</b> ; considered a 'black box' approach
Computational Intensity	Generally <b>lower</b>	<b>Varies</b> ; dependent on model complexity and data size	<b>High</b> ; requires significant computational resources
Data Suitability	Works well when the underlying process is <b>well understood</b>	Effective for structured datasets with <b>complex relationships</b>	Ideal for large datasets, including <b>unstructured</b> data
Core Models	ARIMA, Exponential Smoothing (ETS)	Random Forest, Gradient Boosting	RNNs, LSTMs, CNNs



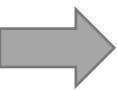


# Python packages

Package	Metrics	ML	DL	Focus
 statsmodels				Statistics, Econometrics
 scikit-learn				General Machine Learning
 Keras				General Deep Learning
 PyCaret				Auto ML, Rapid prototyping, Comparison
 Darts				Advanced timeseries and forecasting
 GluonTS				Large Scale probabilistic Models
 NIXTLA				Workflows, SOTA methods



Pedram.Jahangiry



# Module 1- Part 4

## 1. Timeseries Basics

- What is sequence data?
- Time series tasks
- Trend, Seasonality, Residual
- ACF and PACF
- Stationarity

## 3. Modeling

- White Noise
- Econometrics
- Machine Learning
- Deep Learning

## 2. Forecasting strategies

- One-step, Multi-step, Multi-output
- Univariate vs Multivariate
- Benchmarks

## 4. Can we beat wall street?

- Is stock price Random Walk? (EMH)
- Does Uni-variate Model work?
- Can Multi-variate Models help?
- So What?



Pedram.Jahangiry

# Module 1- Part 4

## Demystifying Timeseries Data and Modeling

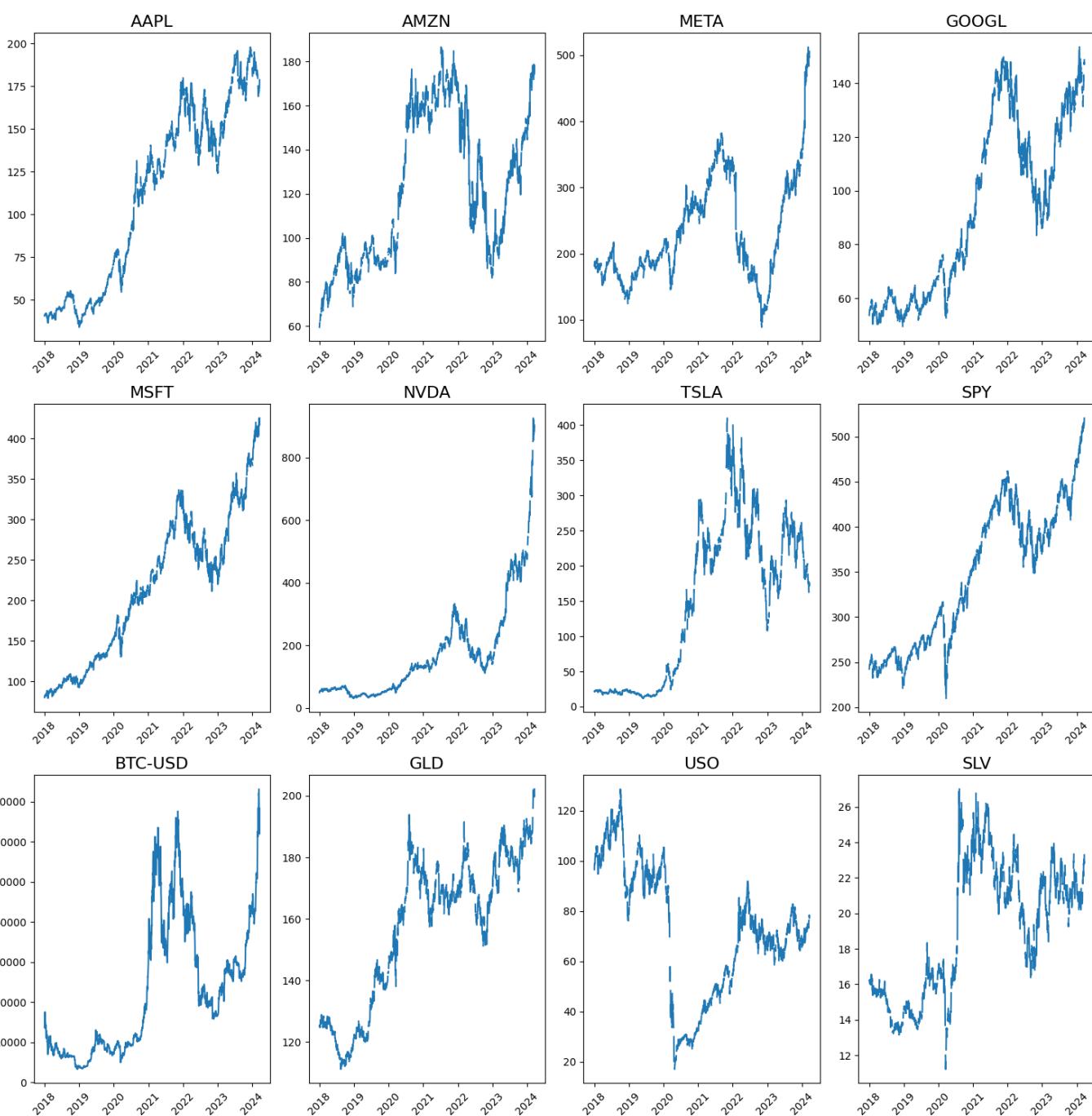


TS Basics  
Modeling  
Forecasting

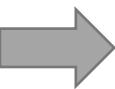


Pedram.Jahangiry

## 4. Beating Wall Street

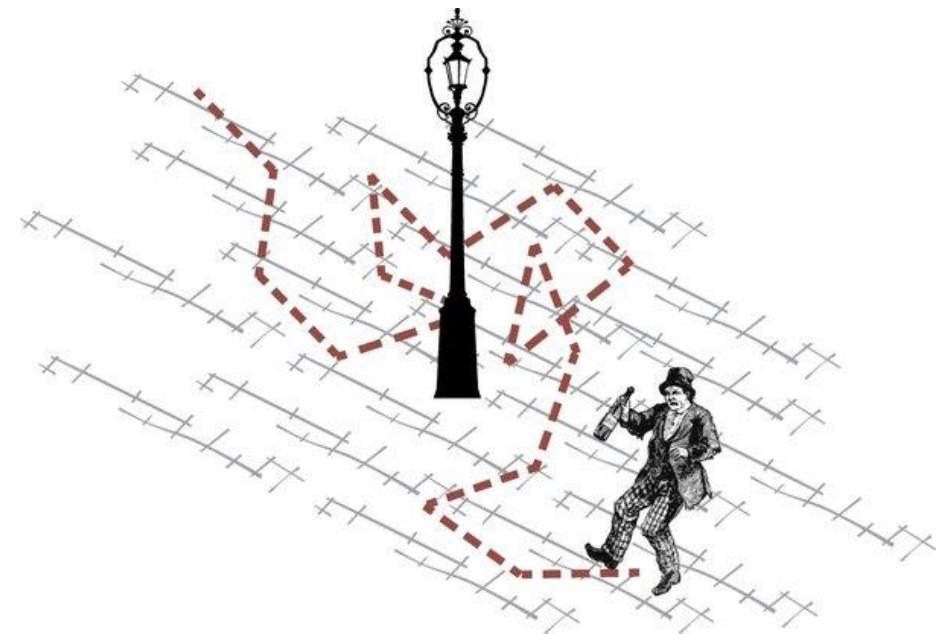


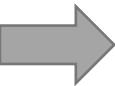
Pedram Jahangiry



# Is stock price Random Walk?

- Efficient Market Hypothesis (EMH): Stock prices reflect **all available information**.
- Empirical evidence: Statistical tests often fail to reject the random walk hypothesis. Many forecasting models struggle to **consistently** outperform **naive benchmarks**.
- Unpredictable Events: Markets are driven by news, economic shifts, and geopolitical events, many of which are inherently unpredictable and random.
- What are the arguments against random walk?





# Is stock price NOT Random Walk?

## Arguments Against Pure Random Walk

- **Predictability and Patterns:** Evidence of slight predictability and patterns in stock prices, with **technical analysis** exploiting these non-random elements.
- **Behavioral Finance:** Human psychology and **biases**, such as overreaction or herding, introduce **non-random** elements into market behavior.



Pedram Jahangiry

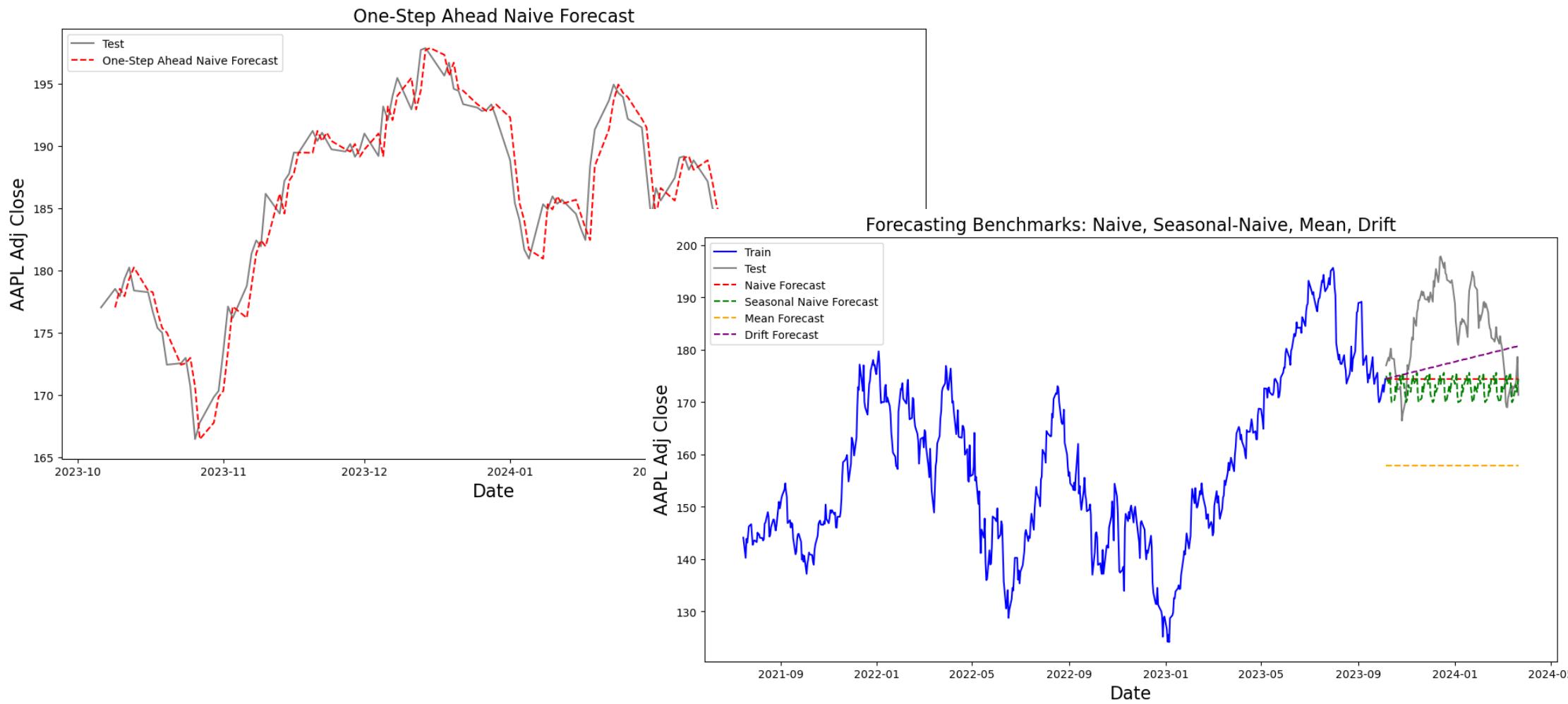
# The Middle Ground

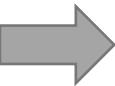
- **Not Perfectly Random:** Stock prices likely aren't purely random walks. There may be pockets of predictability or temporary trends.
- **Market efficiency** makes it very difficult to consistently exploit these deviations for significant gains.
- For long-term investors, a random walk model is a **reasonable approximation** →
- **Benchmark = Naïve forecaster**



Pedram Jahangiry

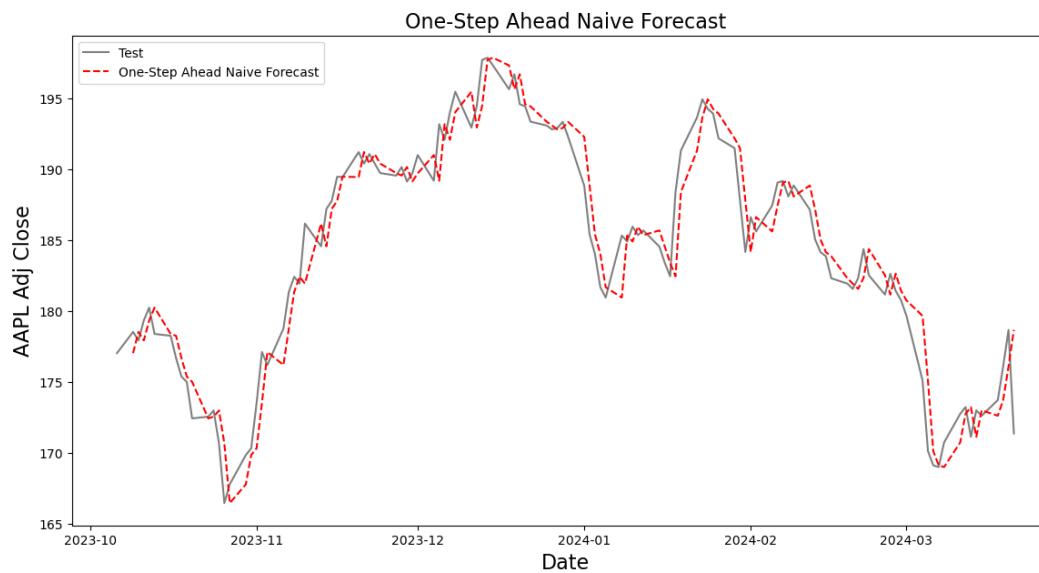
# Reminder: Naïve Forecaster

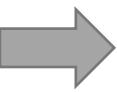




# Does Uni-variate Model Work?

- Hard to beat **naïve forecaster!**
- Complex models might offer marginal gains, but the effort/cost of **constant updating** is substantial.
- **Trading strategy** is a whole other challenge: bet sizing, transaction costs, and turning forecasts into profitable decisions.





# Can Multi-variate Models Help?

- Potential for Improvement by using **Alternative data**
- **Feature Engineering** is Key: The success of multi-variate models often hinges on identifying and extracting strong, predictive features from the diverse data.
- Increased Complexity and **accessibility challenge**: Building and maintaining multi-variate models is often computationally more demanding
- Trading strategies remains complex due to market efficiency and transaction costs.



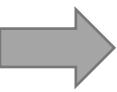
Pedram.Jahangiry



DALL-E

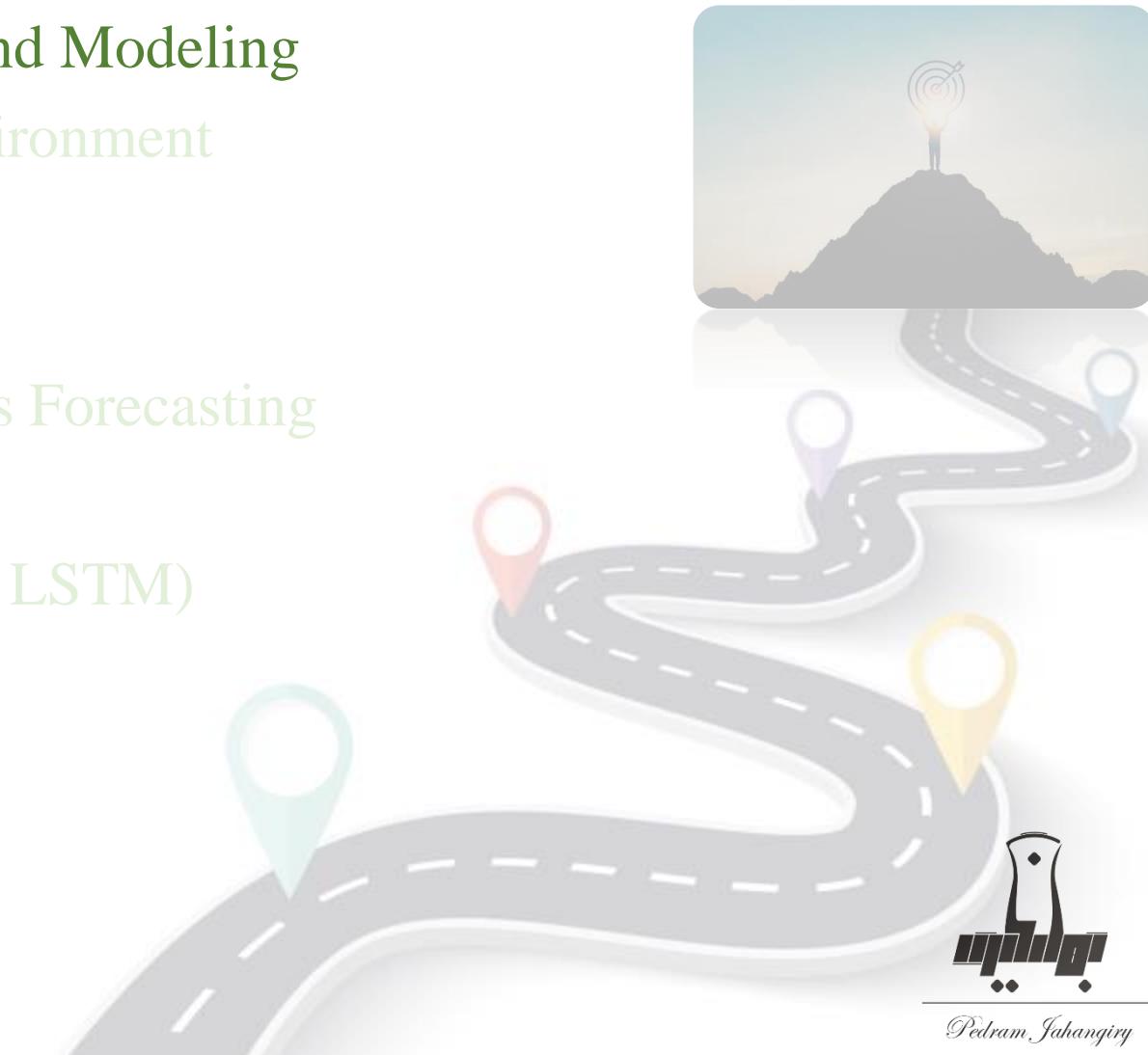


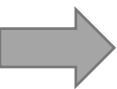
DALL-E



# Road map!

- ✓ Module 1- Demystifying Timeseries Data and Modeling
- Module 2- Setting up Deep Forecasting Environment
- Module 3- Exponential Smoothing
- Module 4- ARIMA models
- Module 5- Machine Learning for Time series Forecasting
- Module 6- Deep Neural Networks
- Module 7- Deep Sequence Modeling (RNN, LSTM)
- Module 8- Prophet (Forecasting at Scale)





# Where to find the materials?

Pinned

**Machine\_Learning-USU** Public

Github repository for machine learning course owned and maintained by prof. Jahangiry

• Jupyter Notebook ⭐ 29 📂 36

**Pedram Jahangiry**  
PJalgotrader

Professor at the Jon M. Huntsman School  
of Business, Utah State University.

Customize your pins

**Deep\_Learning-USU** Public

GitHub repository for deep learning courses owned and maintained by prof. Jahangiry

• Jupyter Notebook ⭐ 7 📂 6

**Deep\_forecasting-USU** Public

GitHub repository for deep forecasting courses owned and maintained by prof. Jahangiry

• Jupyter Notebook ⭐ 5 📂 6

**USU-Analytics-Solution-Center/Bruno.jl** Public

Repository for the Julia package with financial tools for data generation, asset pricing and strategy testing

• Julia ⭐ 35 📂 5

**ML-USU-SP21**

ECN 5090- Machine

• HTML ⭐ 68

**platforms-and-tools**

GitHub repository for platforms and tools

• Jupyter Notebook



Edit profile

**Pedram Jahangiry**

@PedramJahangiry

Professor of data analytics! Machine learning and AI enthusiast [youtube.com/@pedramjahangiry...](https://youtube.com/@pedramjahangiry)

Utah, USA Joined November 2022

143 Following 75 Followers



**Pedram Jahangiry**

@pedramjahangiry · 4.39K subscribers · 209 videos

Pedram Jahangiry, PhD, CFA >

[github.com/PJalgotrader](https://github.com/PJalgotrader) and 2 more links

Customize channel

Manage videos