



Time Series Regression

FinTech
Lesson 10.3



Class Objectives

By the end of today's class you will understand:



Linear Regression



Time Series Linear Regression



Regression Metrics



Train Test Split



Rolling Out-of-Sample

Linear Regression

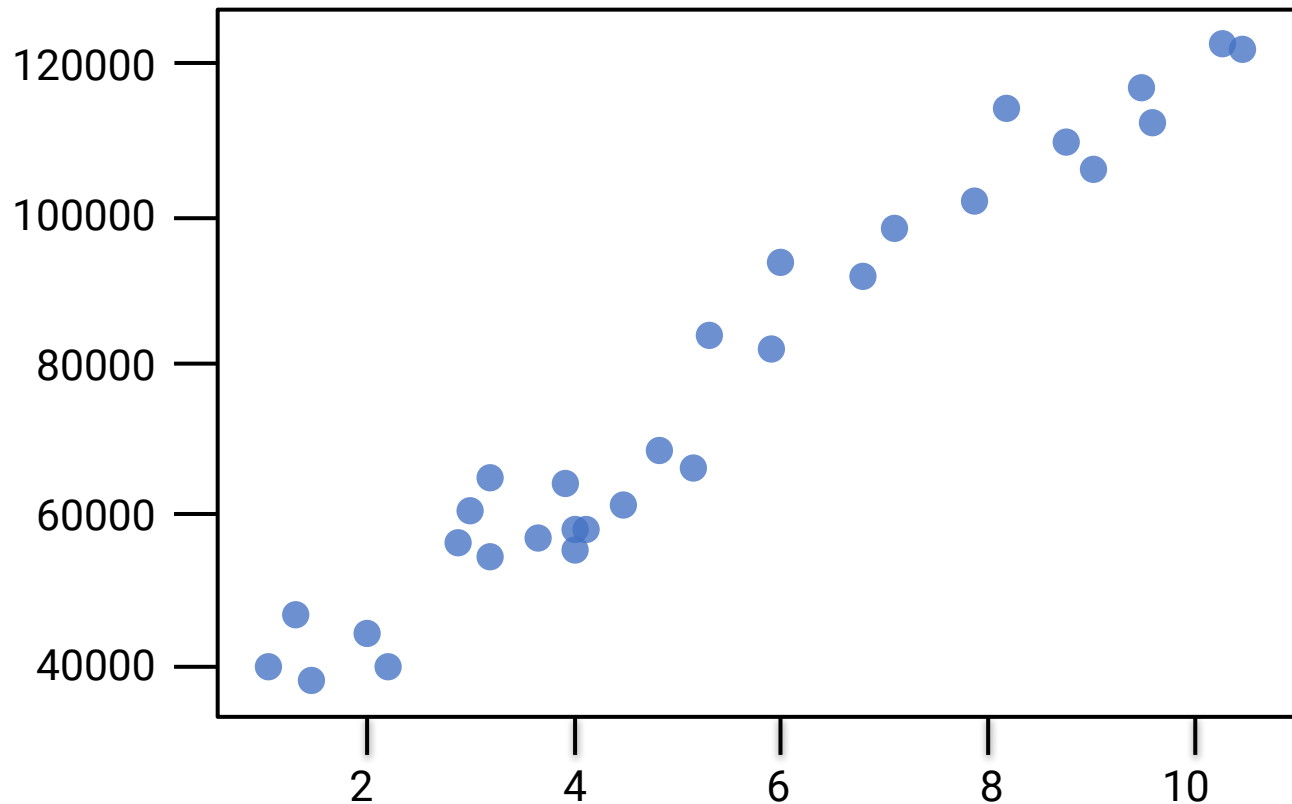
Line Equation

$$y = mx + b$$

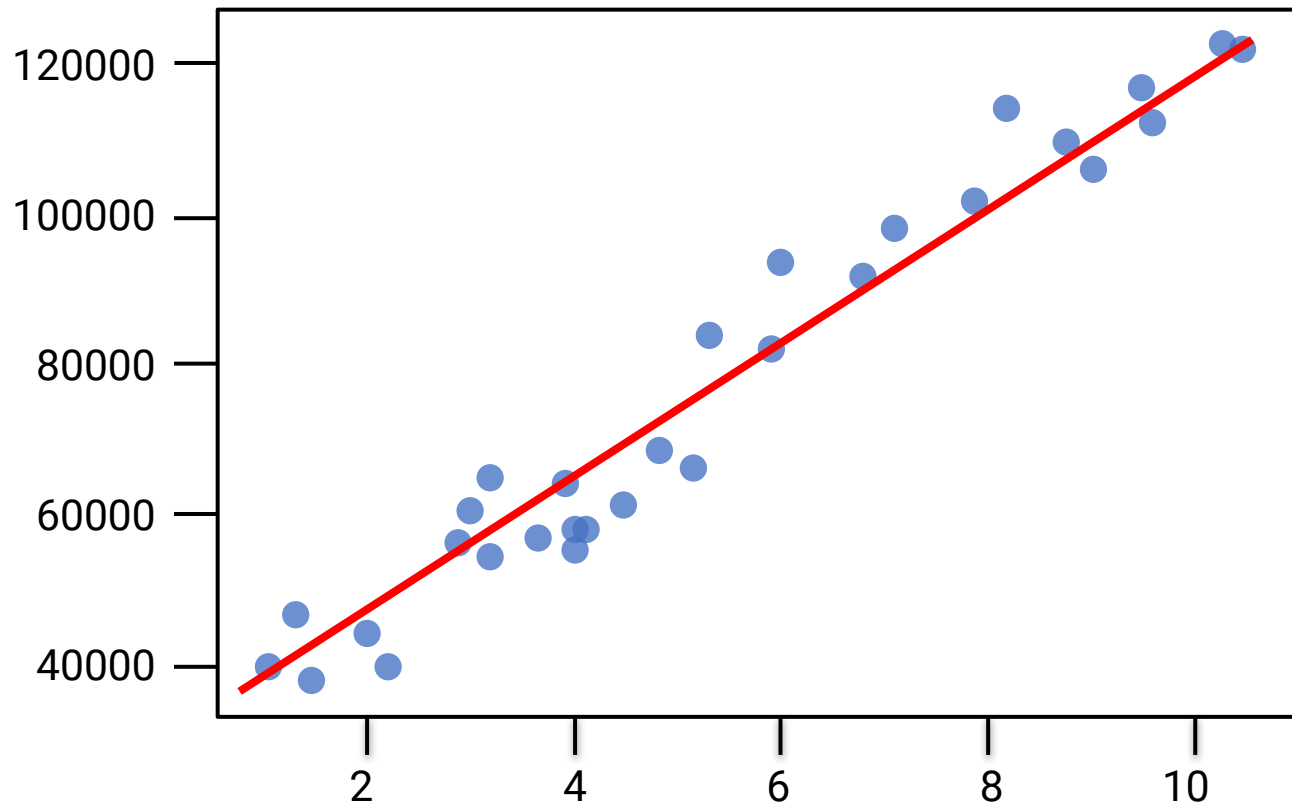
$$m = \text{slope}$$

$$b = \text{y-intercept (the value of } y \text{ when } x = 0)$$

Linear Regression: Find the Line That Best Describes the Data



Best Fit Line



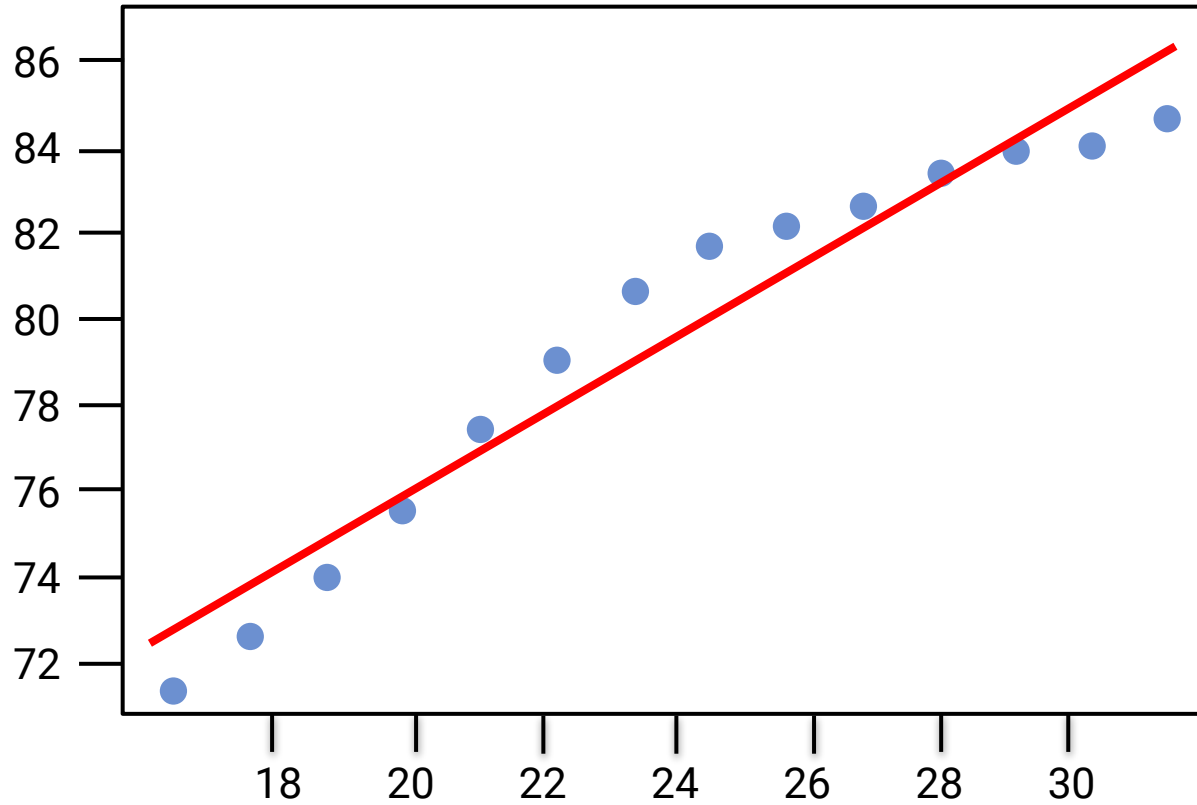
Multiple Regression

Each day (X) is assigned its weight, or coefficient.

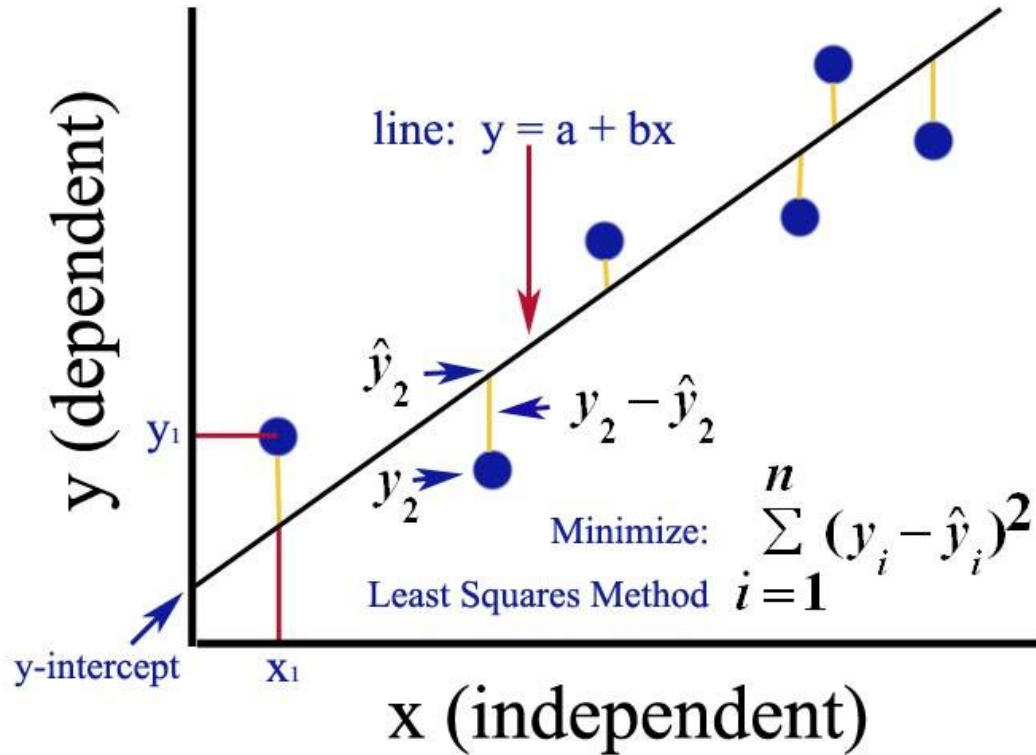
$$y = b_0 + b_1X_1 + b_2X_2...$$

Regression Metrics

Best Fit Line



Regression Metrics

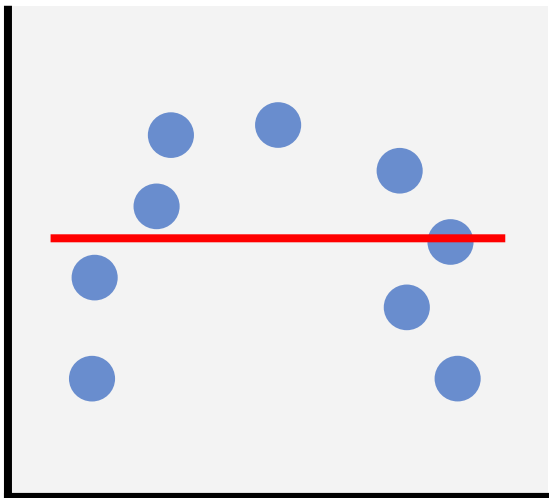




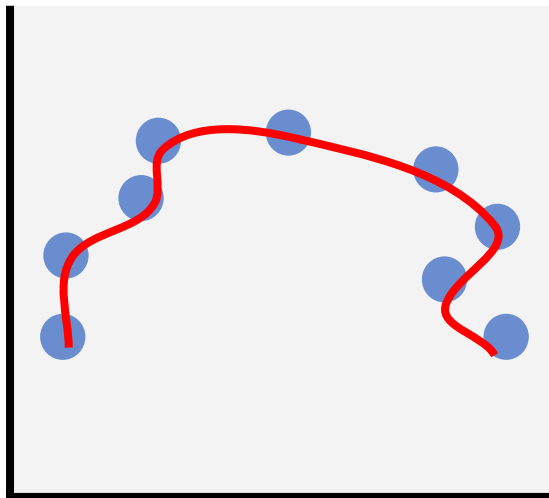
Overfitting

Overfitting

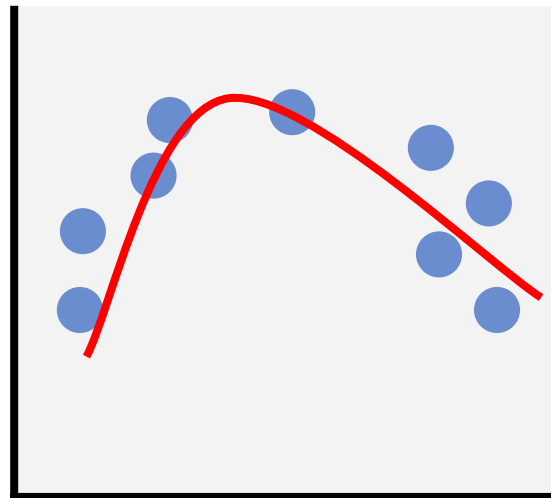
Underfit



Overfit

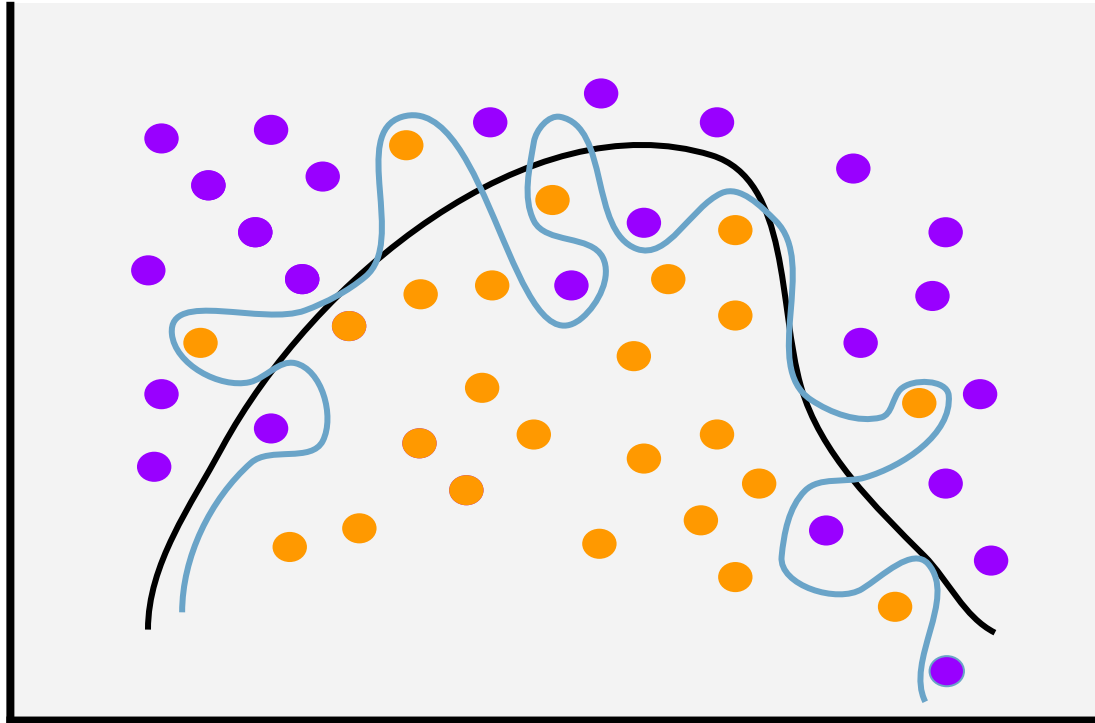


Ideal

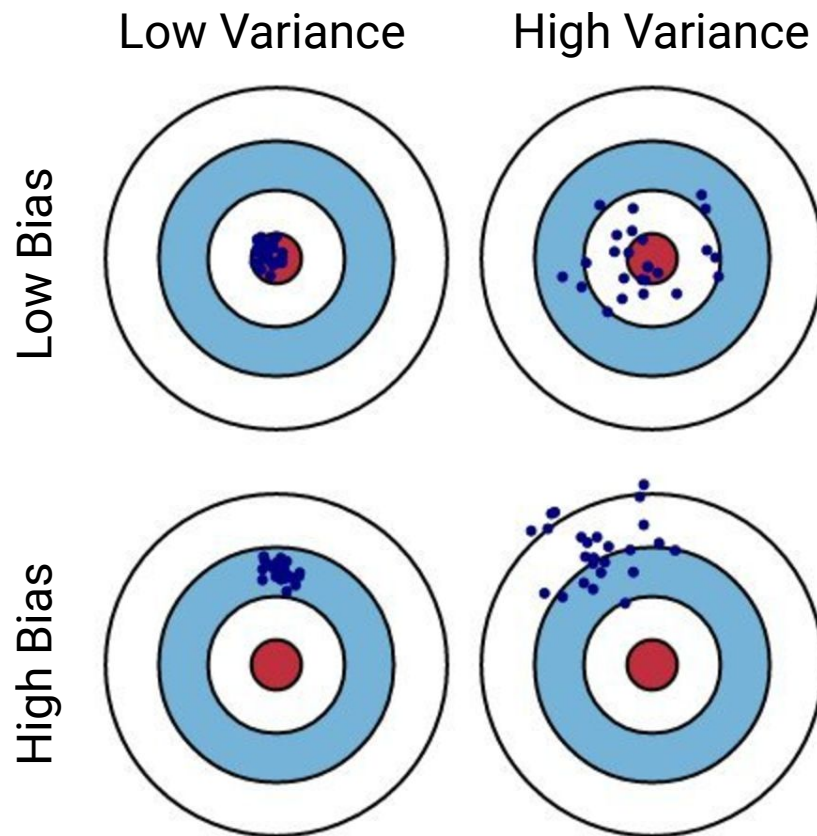


Overfitting

Overfit models learn the 'noise' found in the training data, rather than just the 'signal'



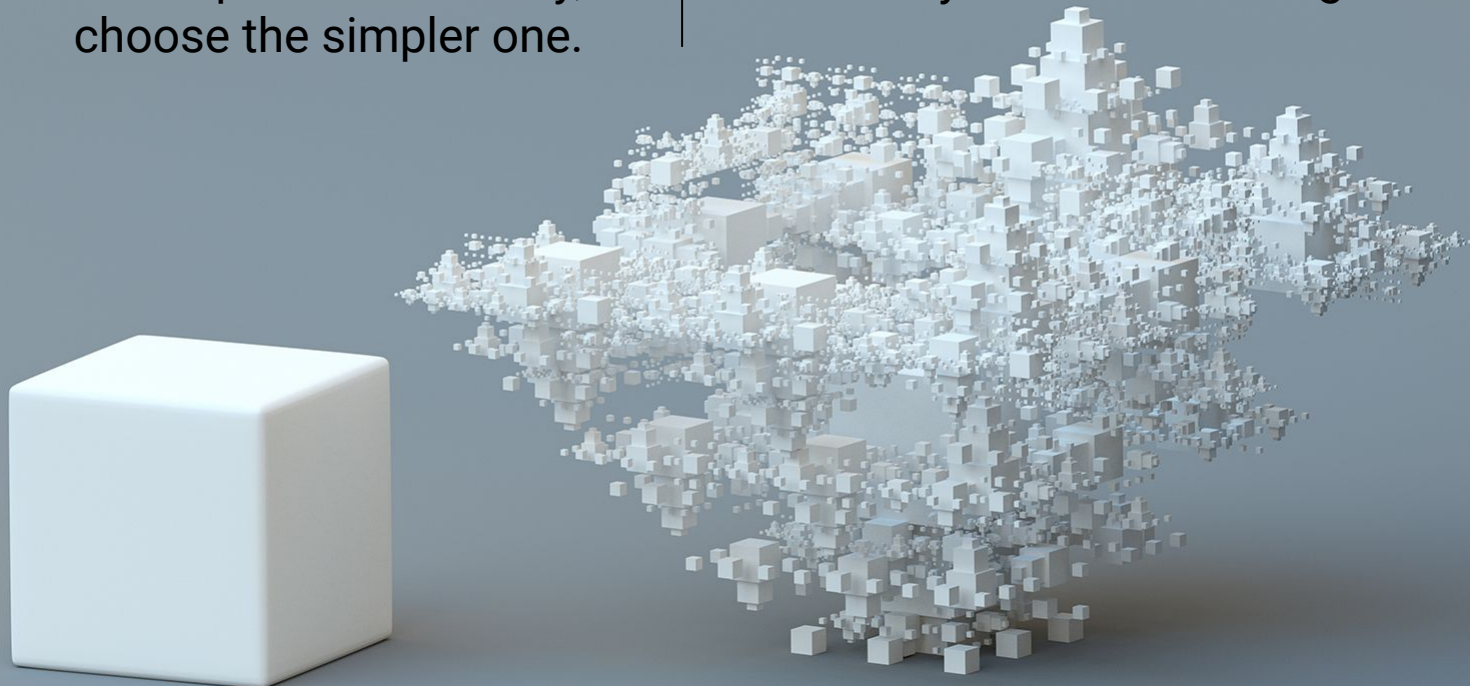
Variance vs Bias



Parsimony

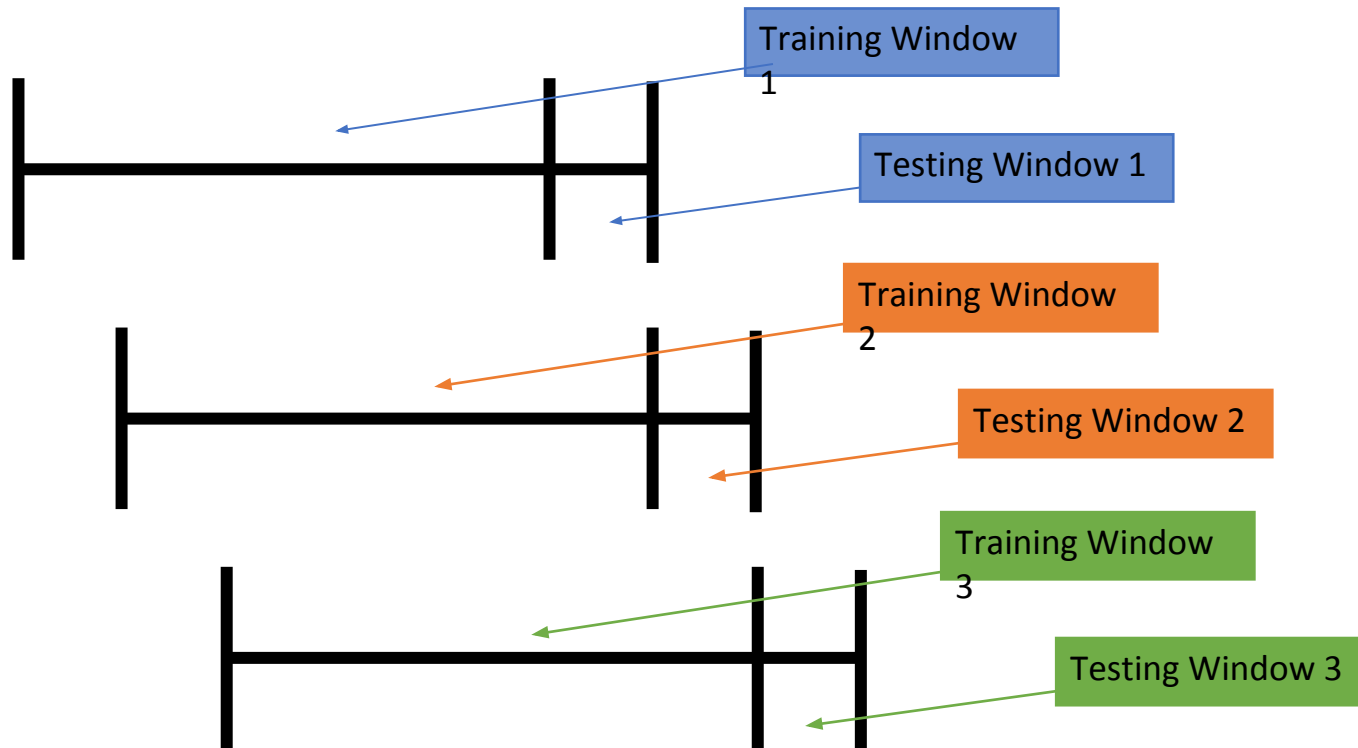
Statistical application of Occam's razor: when two models perform similarly, choose the simpler one.

Why? Needlessly complex models are harder to compute and may lead to overfitting.



Rolling Out-of-Sample

A Rolling Out-of-Sample Approach





Questions?