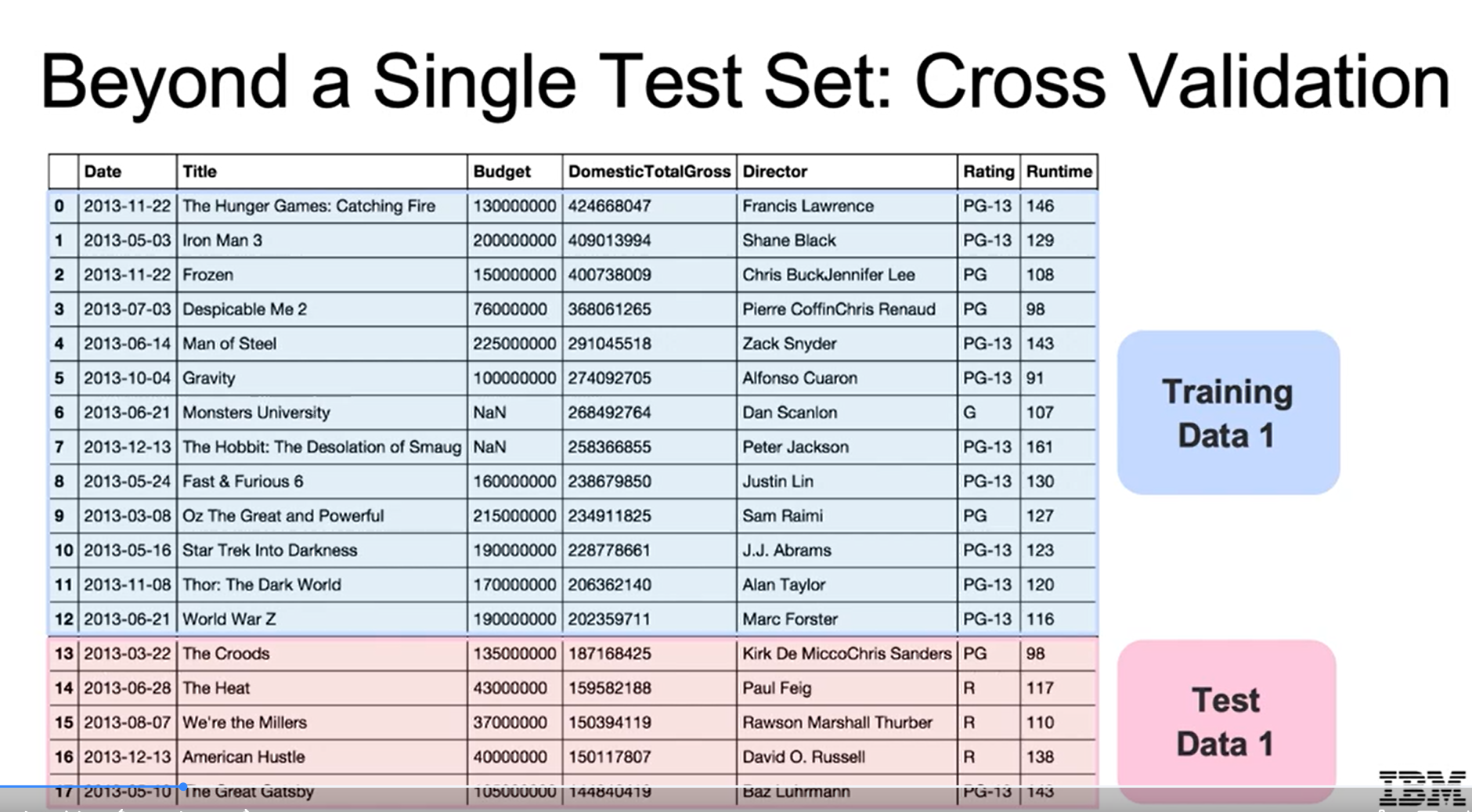
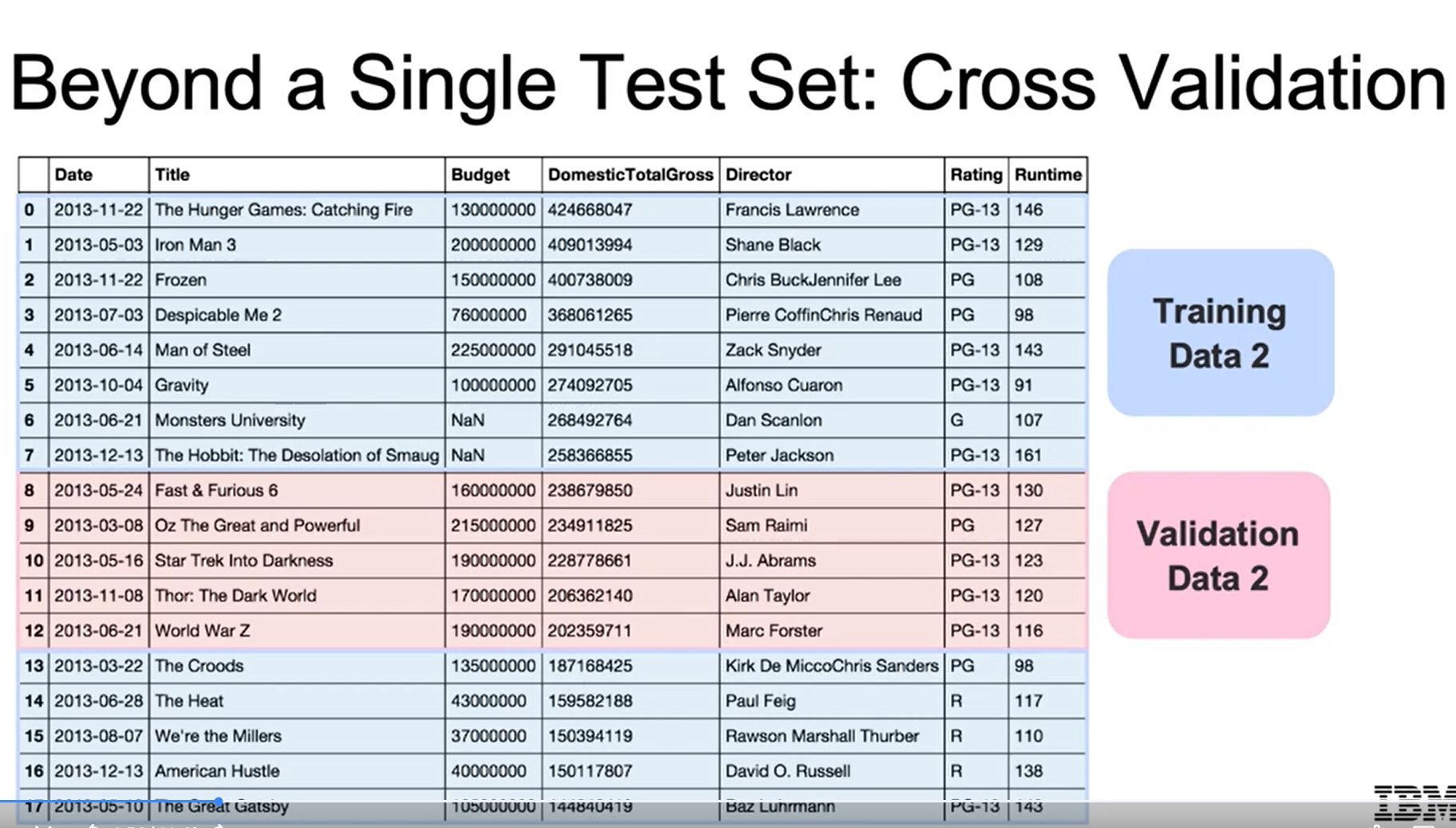
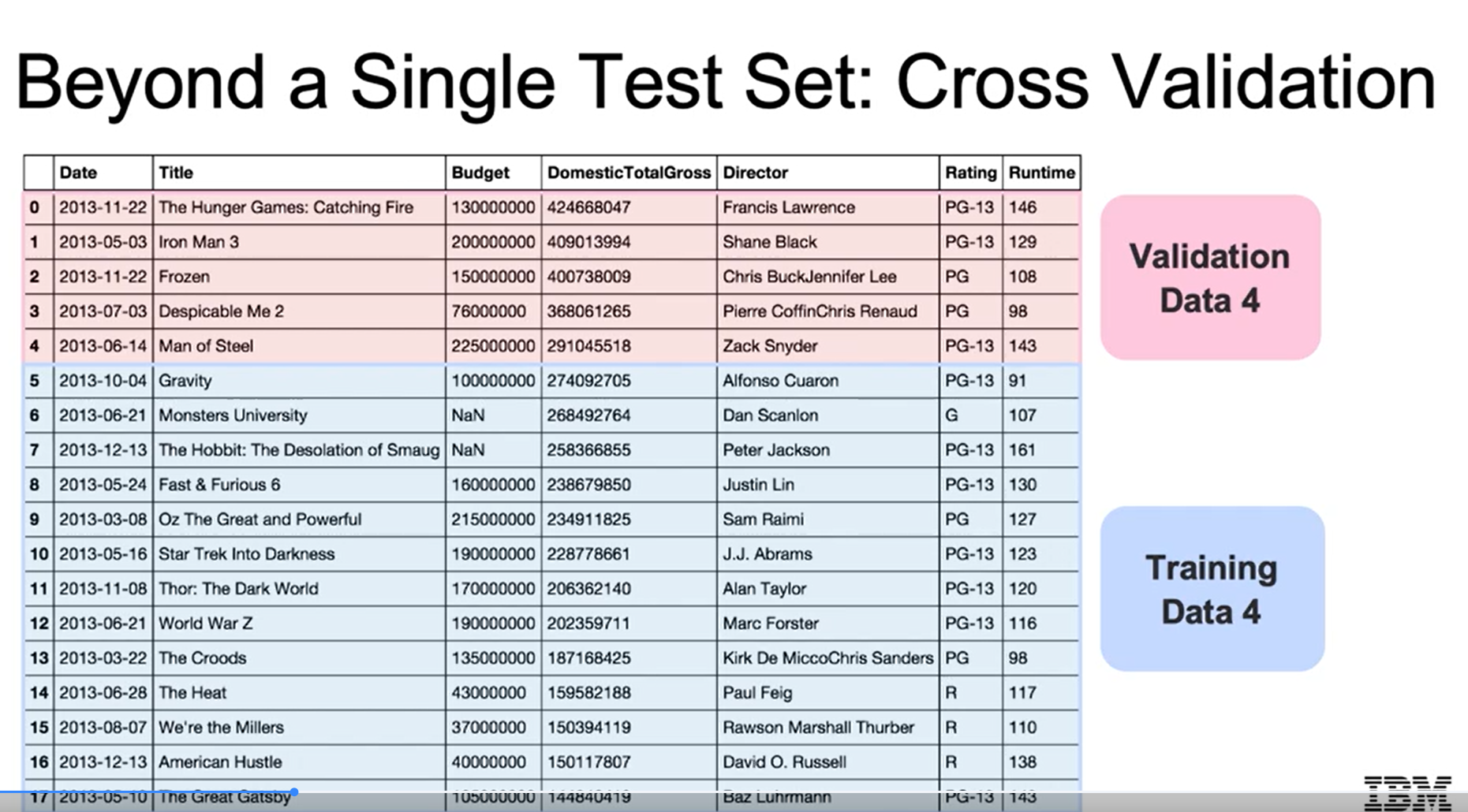
**MOOC 2-MODULE 3**

[**Supervised Machine Learning: Regression**](https://www.coursera.org/learn/supervised-machine-learning-regression/home/welcome)

**Cross Validation**

****



****

### Cross-Validation Overview

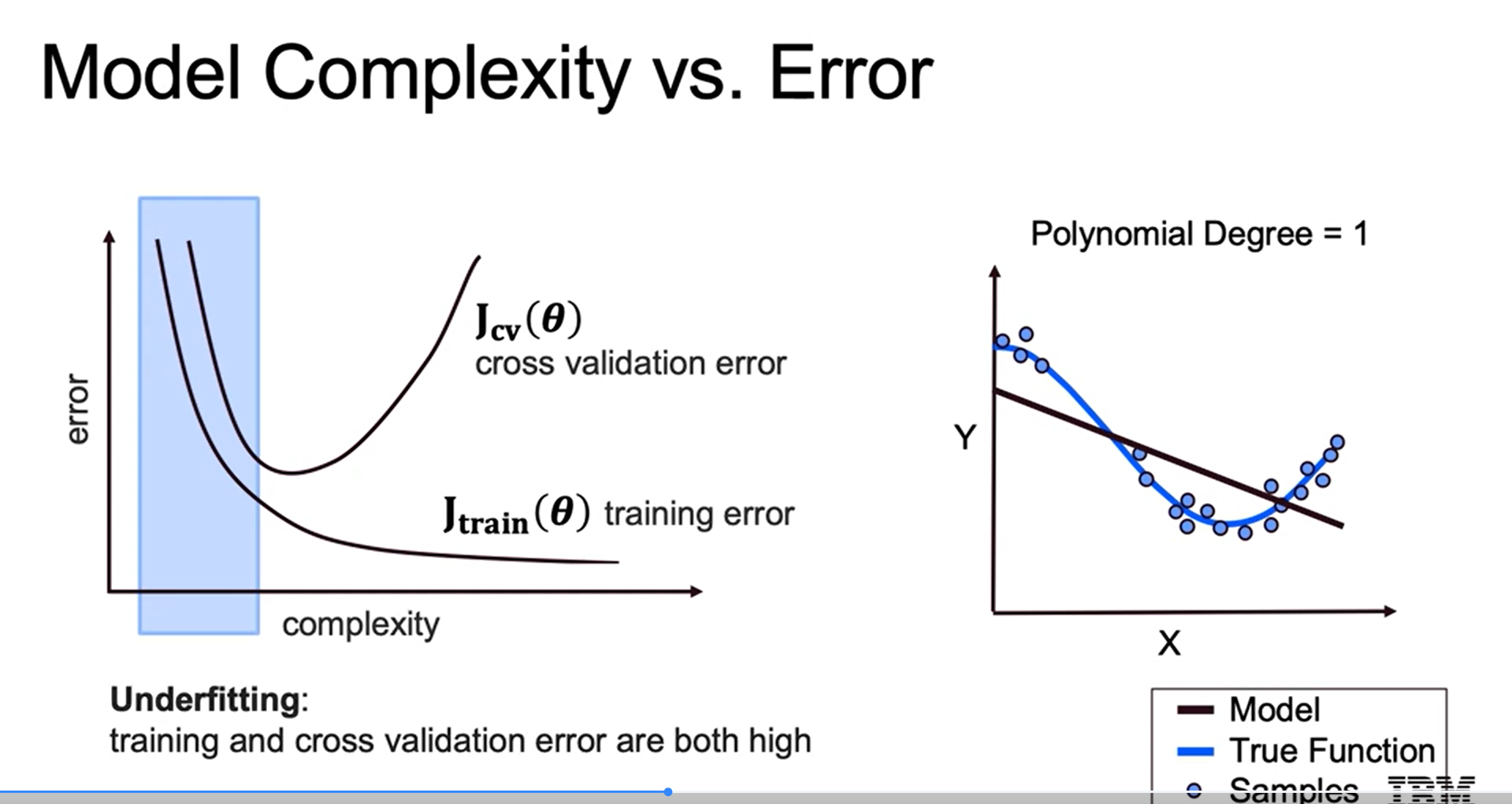
### Cross-validation involves splitting data into multiple train-test pairs to assess model performance more reliably.

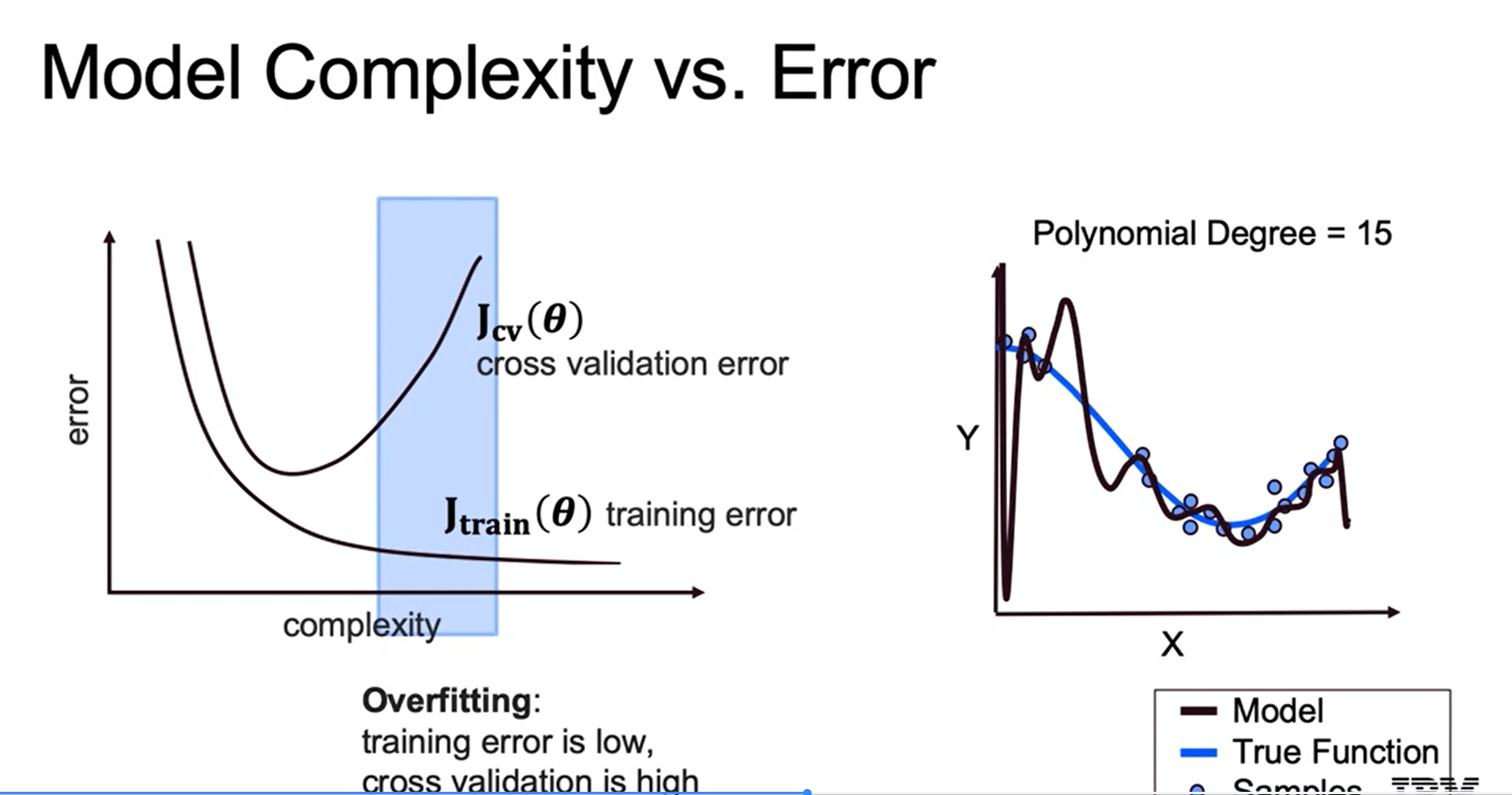
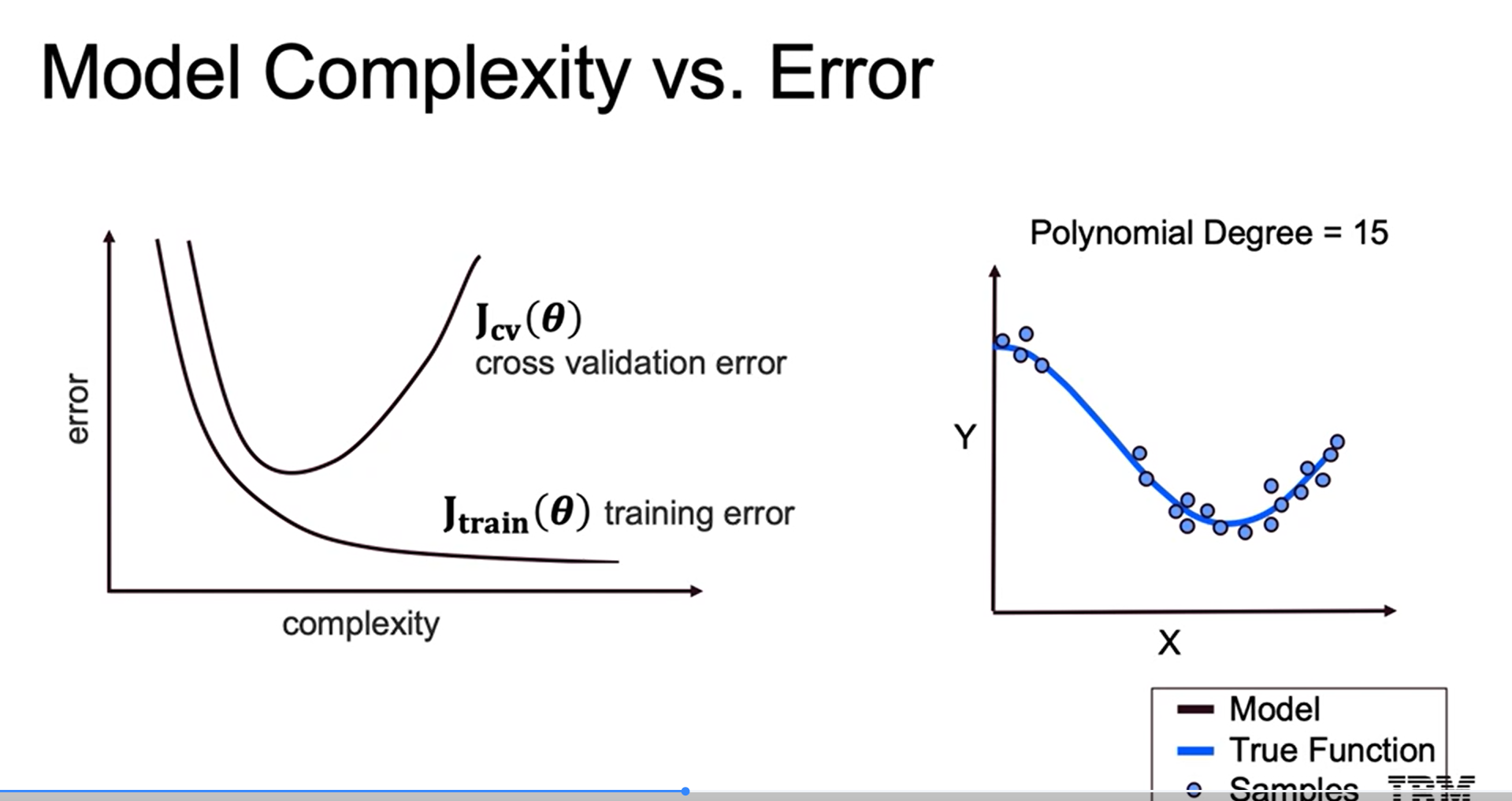
### It averages the error across different test sets, providing a statistically significant performance measure.

### Error Metrics and Model Complexity

### As model complexity increases, training error decreases, but overfitting can occur, leading to poor generalization on new data.

### The goal is to find a balance where both training and cross-validation errors are minimized.





### K-Fold Cross-Validation Techniques

### K-fold cross-validation splits data into K subsets, allowing for multiple evaluations of the model.

### Leave-one-out cross-validation uses a single row as the test set, providing many evaluations but requiring more training time.

### Stratified K-fold ensures that the distribution of categorical outcomes is maintained across splits, enhancing model reliability.

### 