Analyzing Credit Scores with tidymodels in R
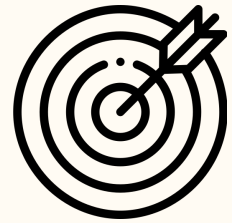
2023-05-09

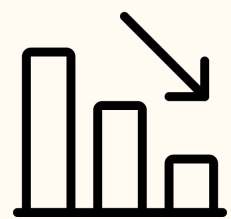datacamp

About this case study

Data: [Credit score classification](#) from Kaggle

Primary Goal: Explain why credit scores differ.

Secondary Goal: Learn about dimensionality reduction using UMAP and `tidymodels` workflows.

Why? We live in a world of complex data, simplifying it helps us understand it.

Dimensionality Reduction

Purpose: simplify complex datasets (i.e., reduces the # of columns)
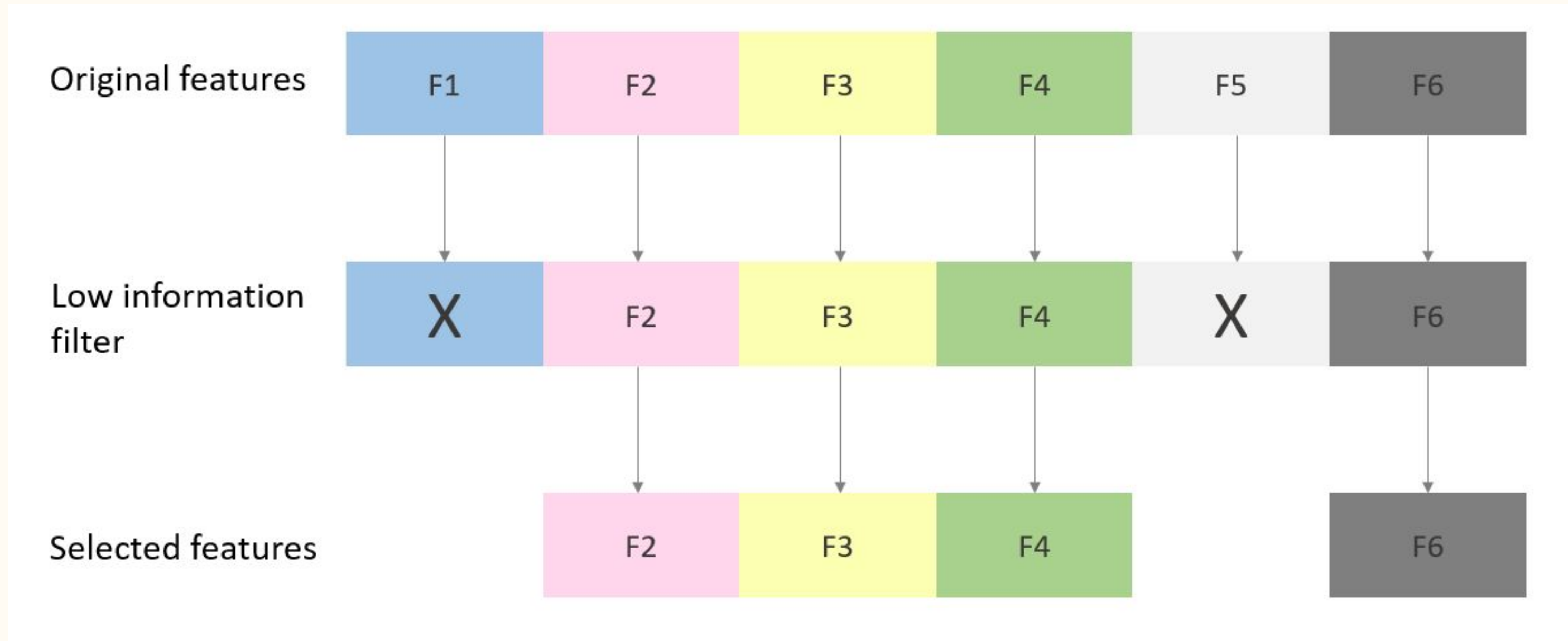
Two Types:
- **feature selection** – removes low-information features.
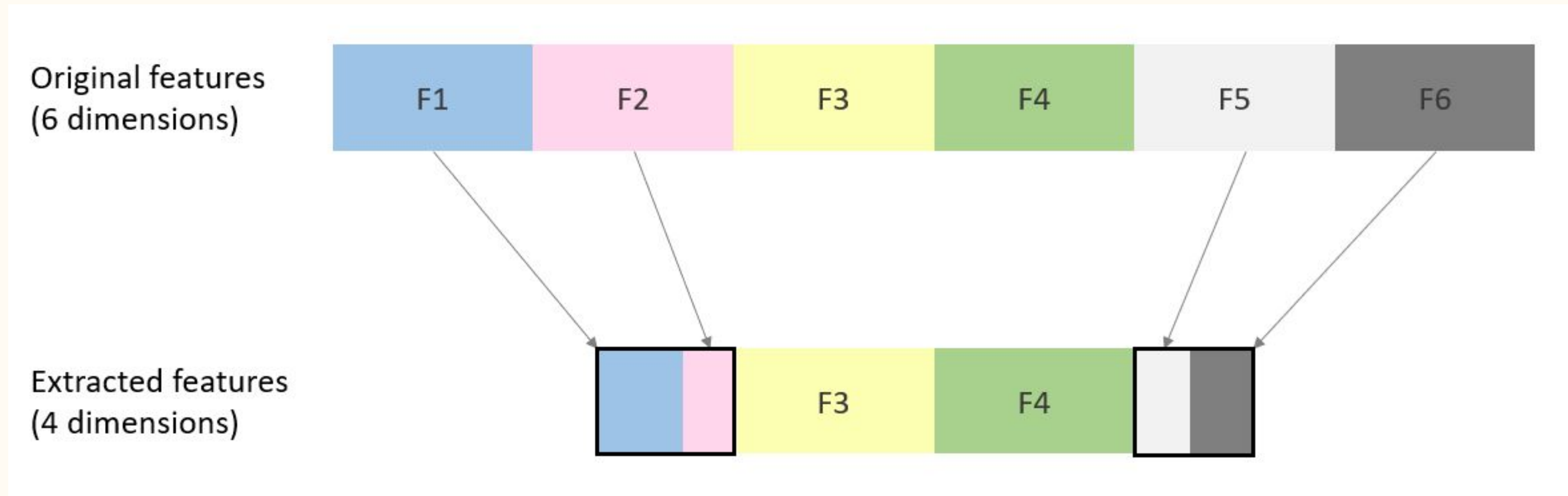- **feature extraction** – combines information from several features.

Feature Selection

Summary: Removes low-information features.

Feature Extraction

Summary: Combines information in features.

Comparison of Feature Extraction Techniques

Summary: PCA is easier to interpret than t-SNE and UMAP

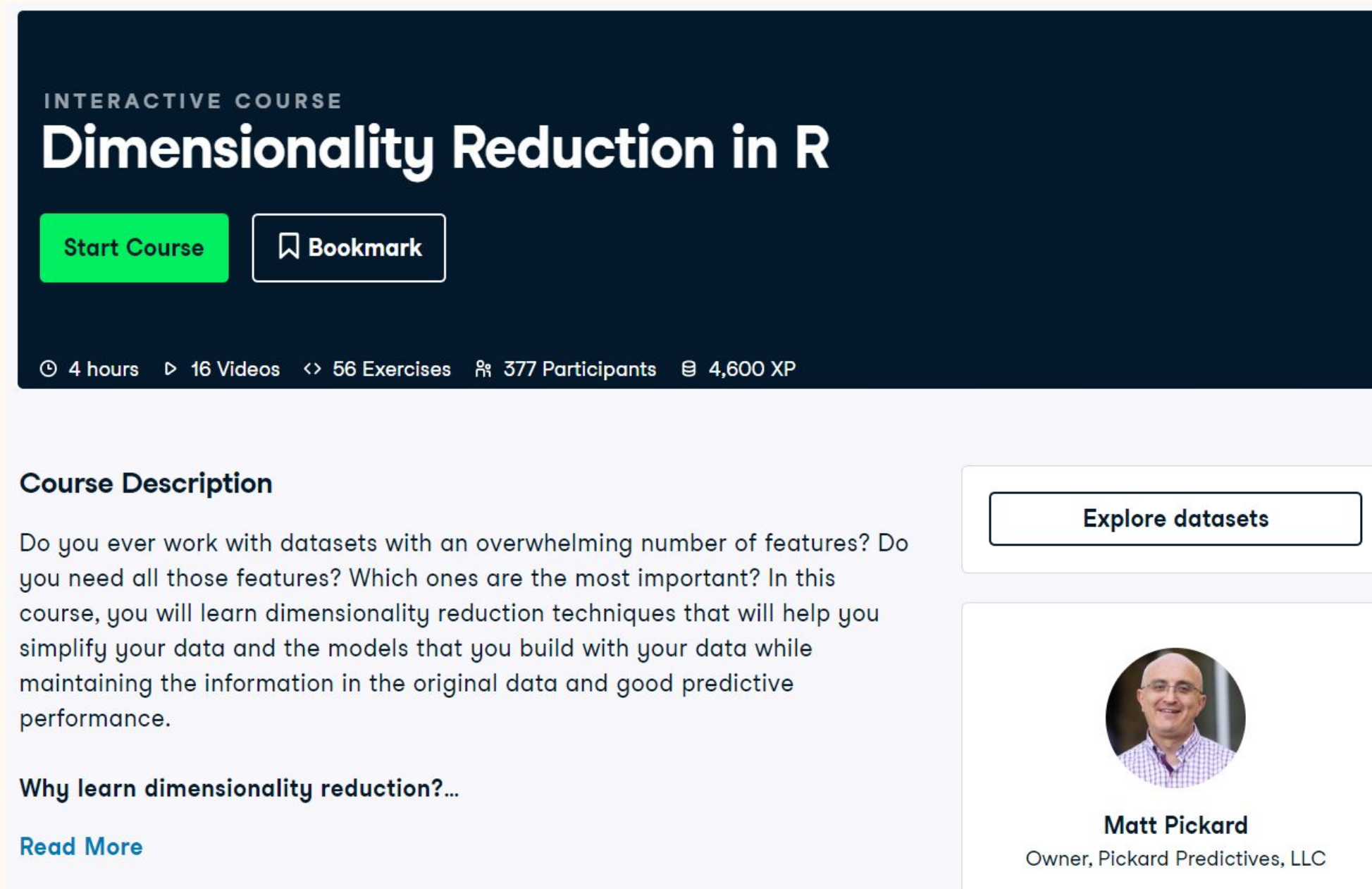| PCA	t-SNE	UMAP
Linear	Non-linear	Non-linear
Deterministic	Non-deterministic	Non-deterministic
Computationally cheap	Computationally expensive	Computationally efficient
Preserves global structure	Preserves local structure	Preserves local and global structure
No hyperparameters	Hyperparameters	Hyperparameters

tidymodels

What is it? A framework (like Python's scikit-learn) incorporates tidy principles into machine learning and modeling.
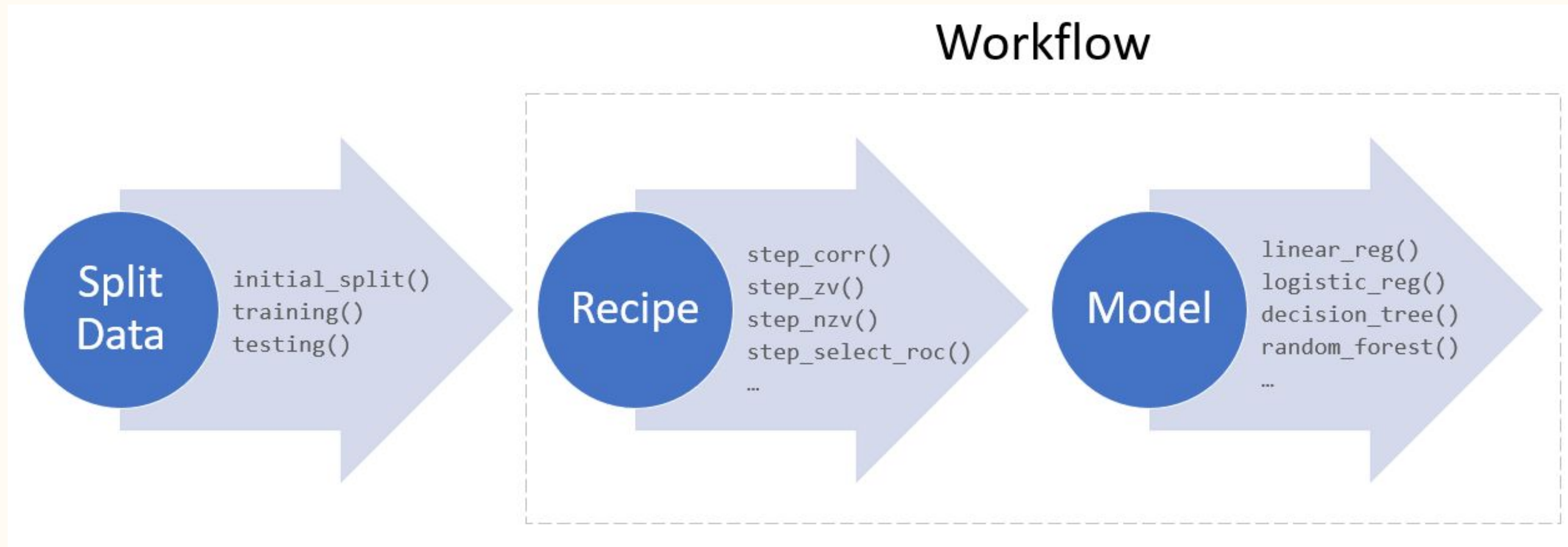
Learning more

We explore dimensionality reduction and `tidymodels` further in Datacamp's [Dimensionality Reduction in R](#) course.

INTERACTIVE COURSE

Dimensionality Reduction in R

[Start Course] [🔖 Bookmark]

⏱ 4 hours   ▷ 16 Videos   <> 56 Exercises   👥 377 Participants   🗄 4,600 XP

Course Description

Do you ever work with datasets with an overwhelming number of features? Do you need all those features? Which ones are the most important? In this course, you will learn dimensionality reduction techniques that will help you simplify your data and the models that you build with your data while maintaining the information in the original data and good predictive performance.

Why learn dimensionality reduction?...

Read More

[Explore datasets]

Matt Pickard
Owner, Pickard Predictives, LLC