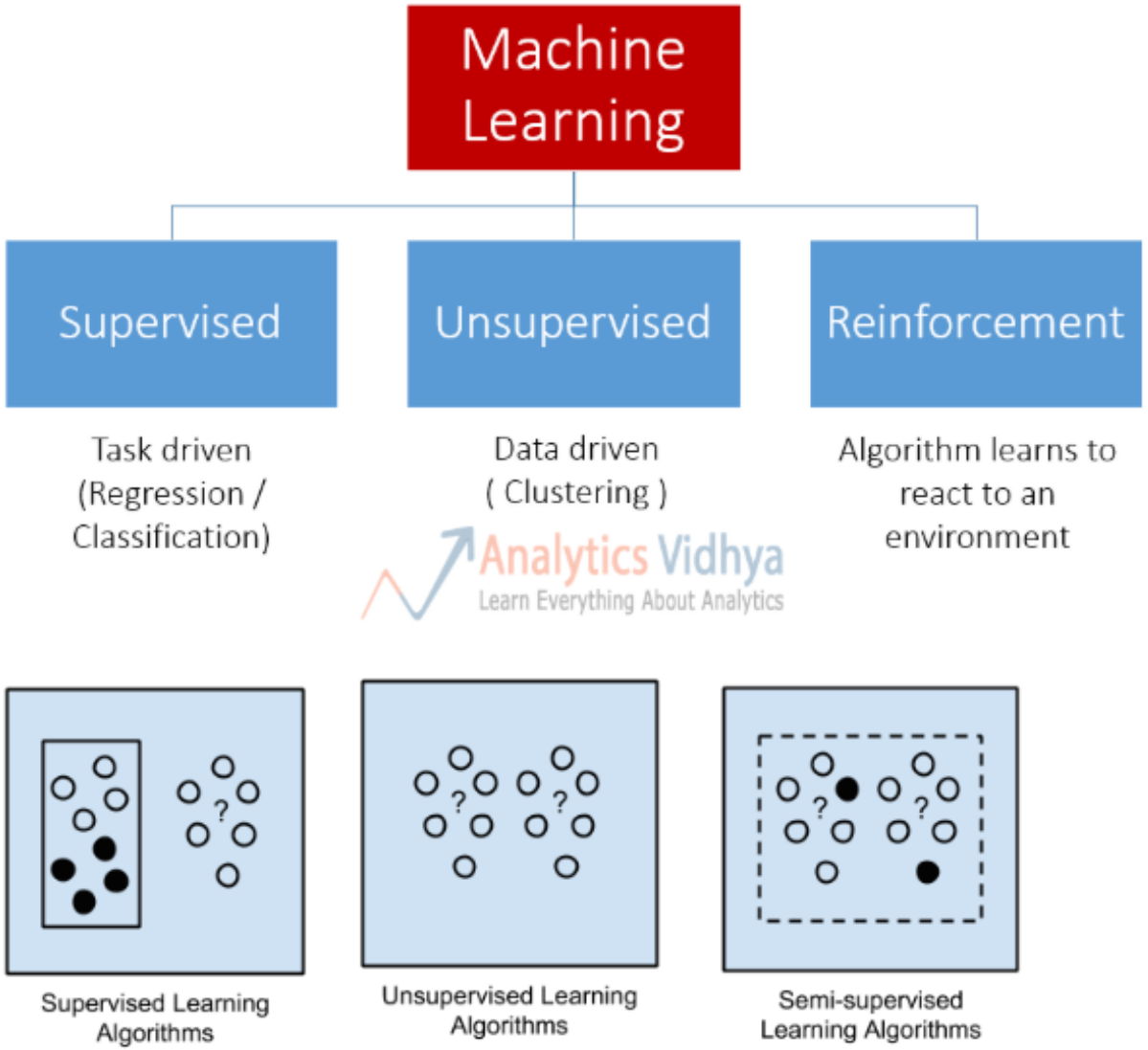
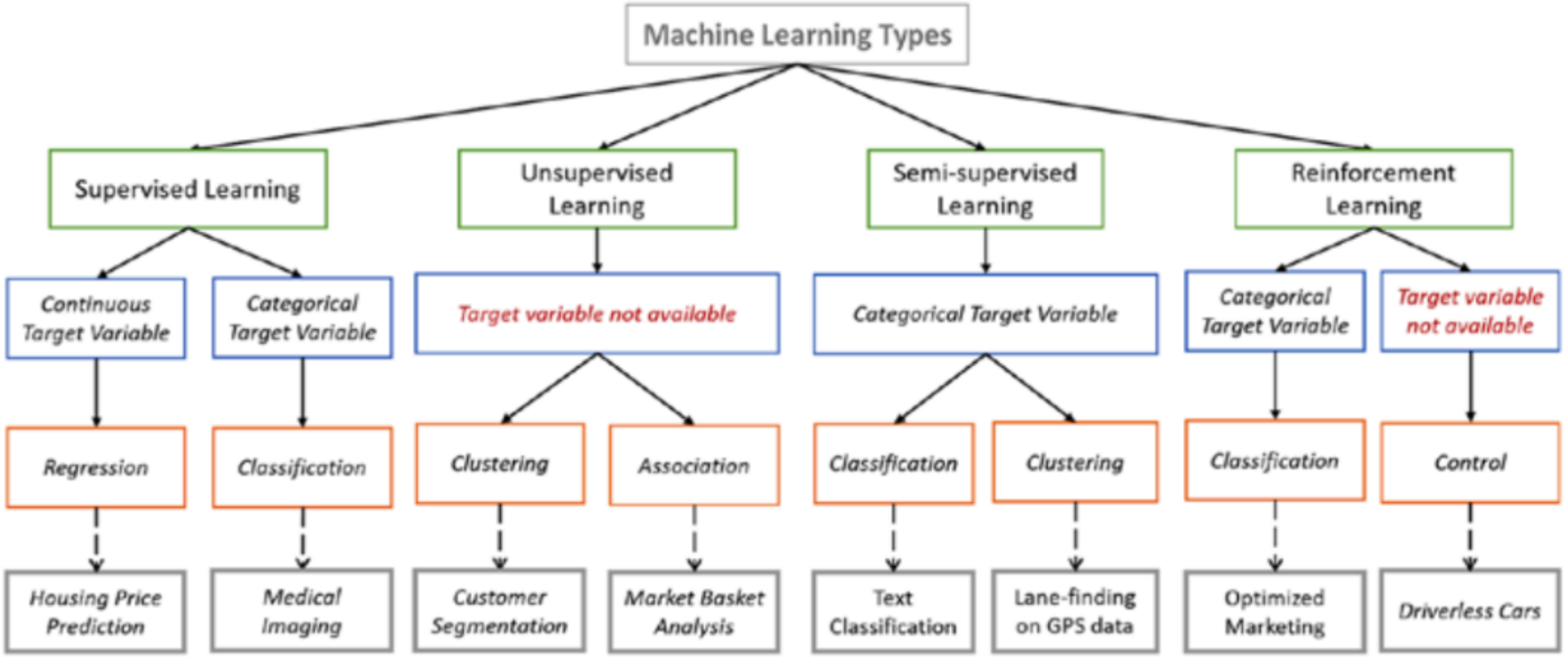
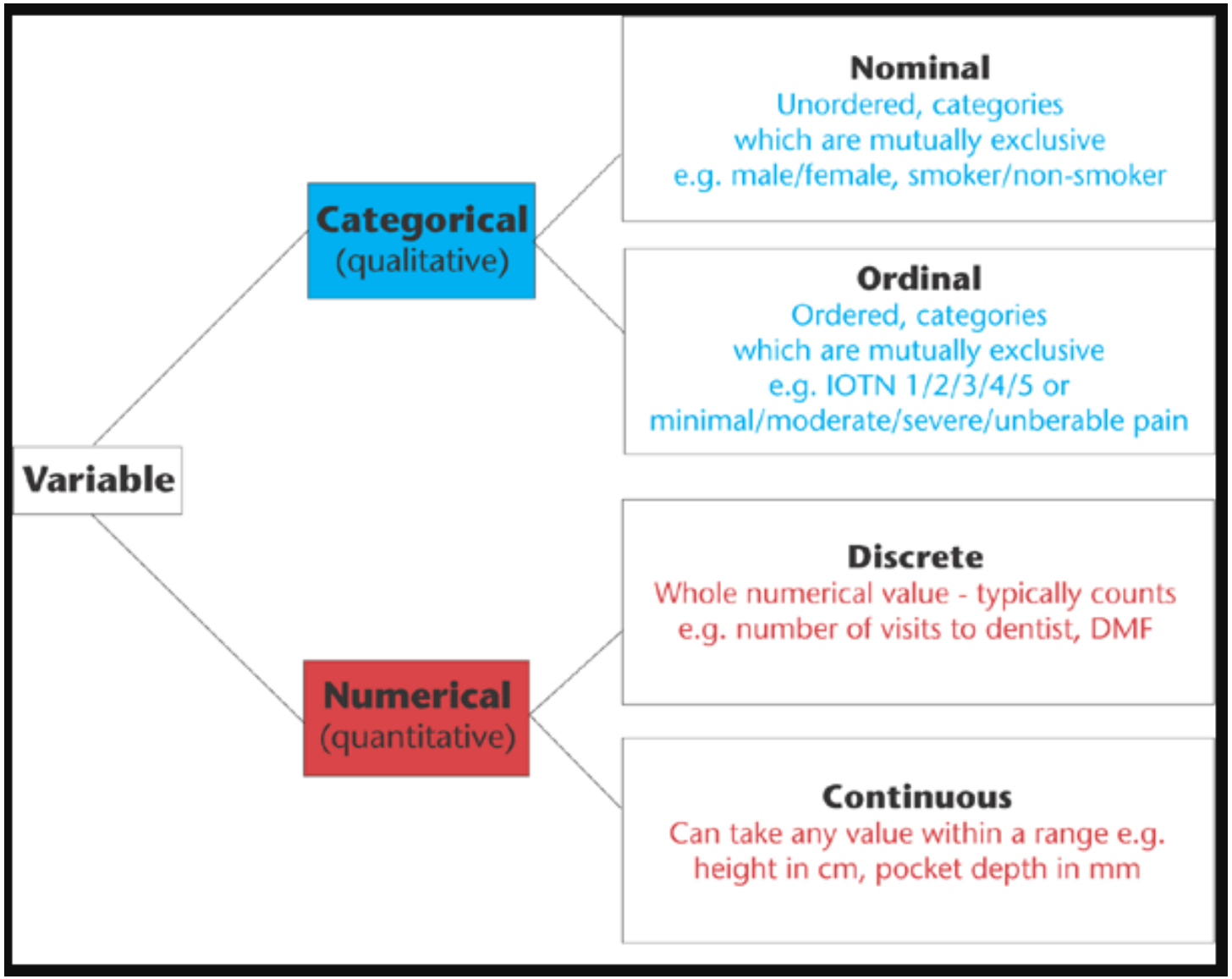
Types of Machine Learning Algos



Types of Machine Learning Algos



Types of Variables



What is sci-kit learn?

- Simple and efficient tools for data mining and data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

https://scikit-learn.org/stable/

ML Pipeline

Train & Test

Collect > Clean > Featurize & transformed > Split into Train, Test > Model w Training > Test with Test Data > Put into Production (Inference)

Inference

Actual data > Clean > Featurize & transformed > Inference

Install sci-kit learn

- !pip install -U scikit-learn – from jupyter notebook cell <shift+enter>

or

- conda install scikit-learn

- https://scikit-learn.org/stable/install.html

Regression using scikit-learn

- https://scikit-learn.org/stable/supervised_learning.html#supervised-learning

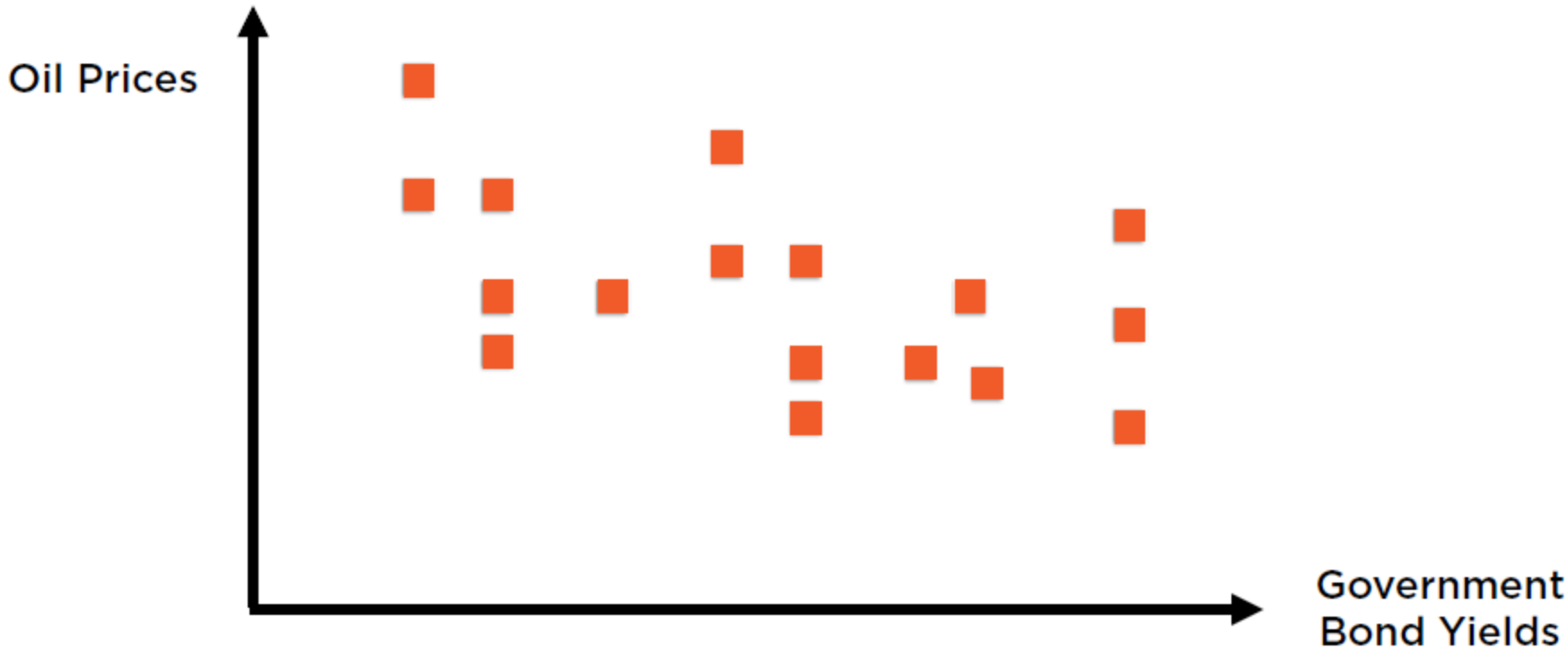
Data in One Dimension

Unidimensional data points can be represented using a line, such as a number line

Data in One Dimension

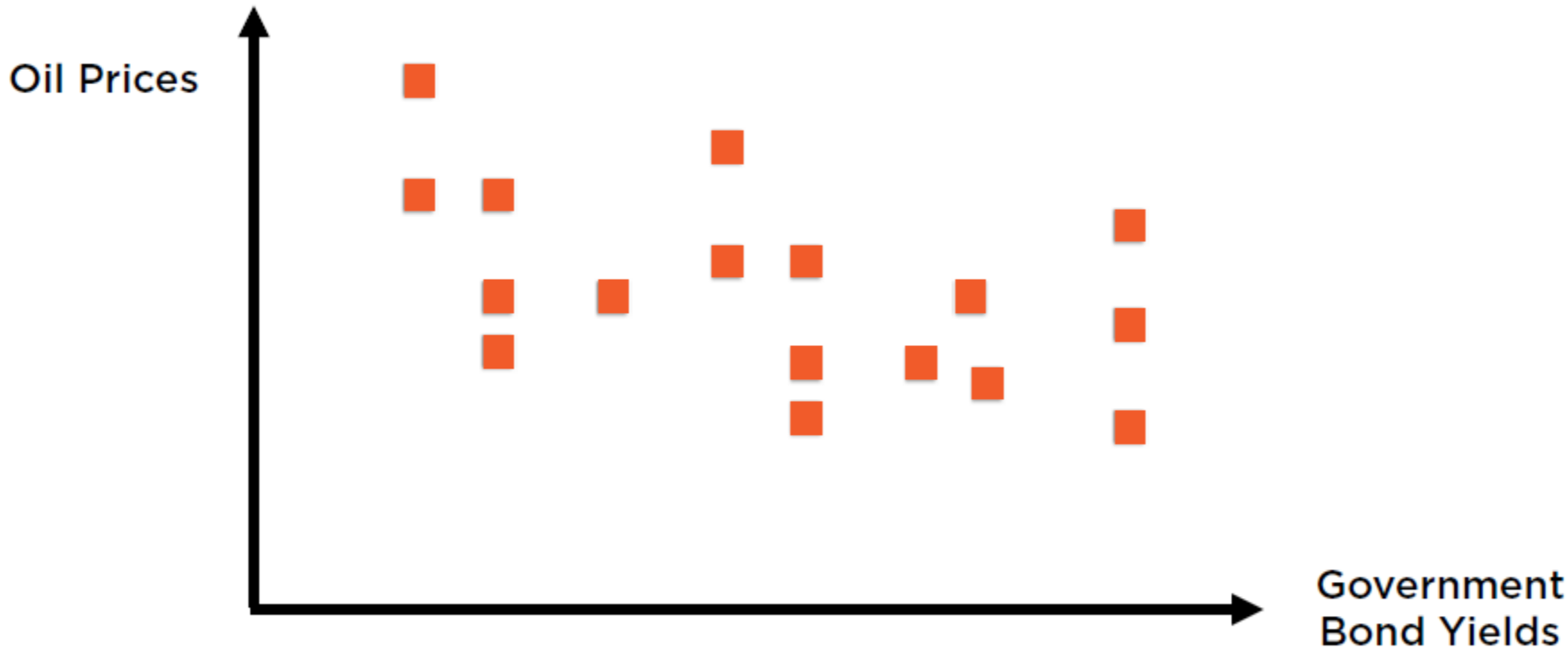
Unidimensional data is analysed using statistics such as mean, median, standard deviation

Data in Two Dimensions



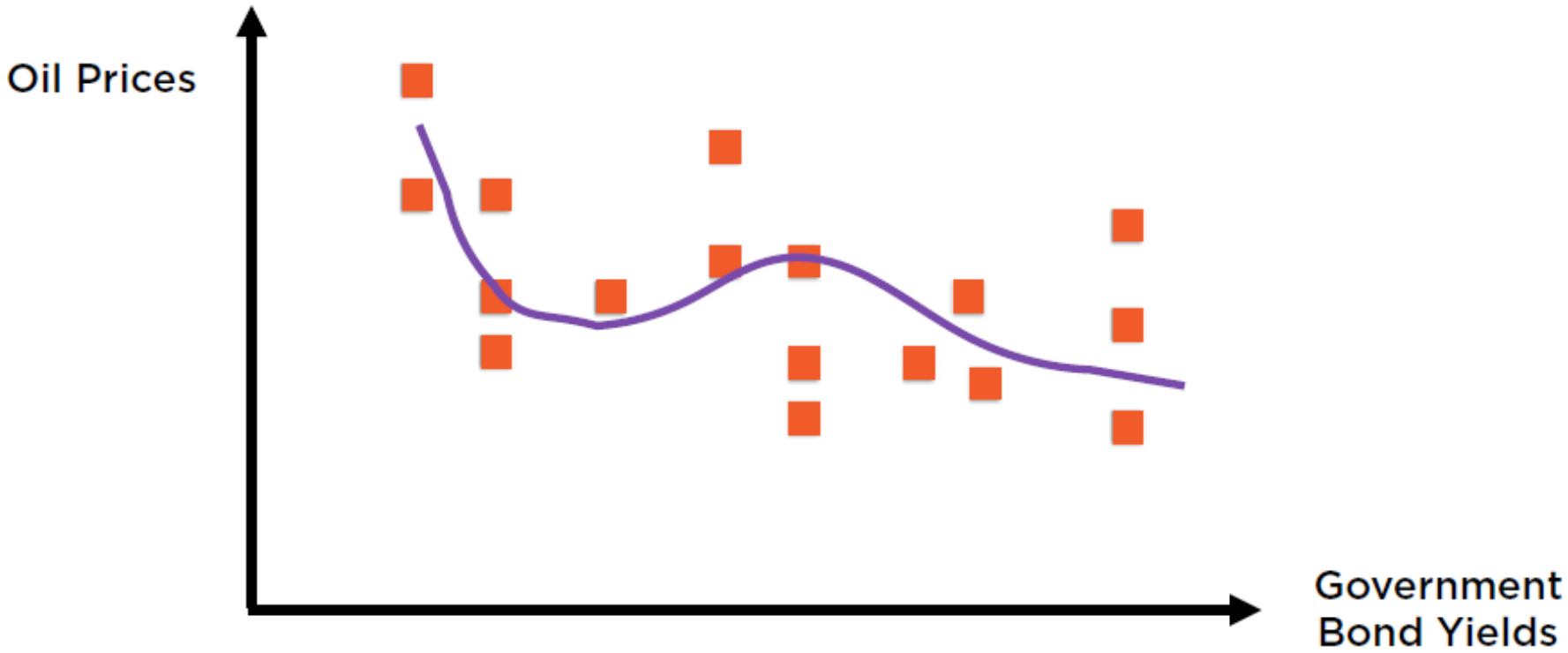
Its often more insightful to view data in relation to some other, related data

Data in Two Dimensions



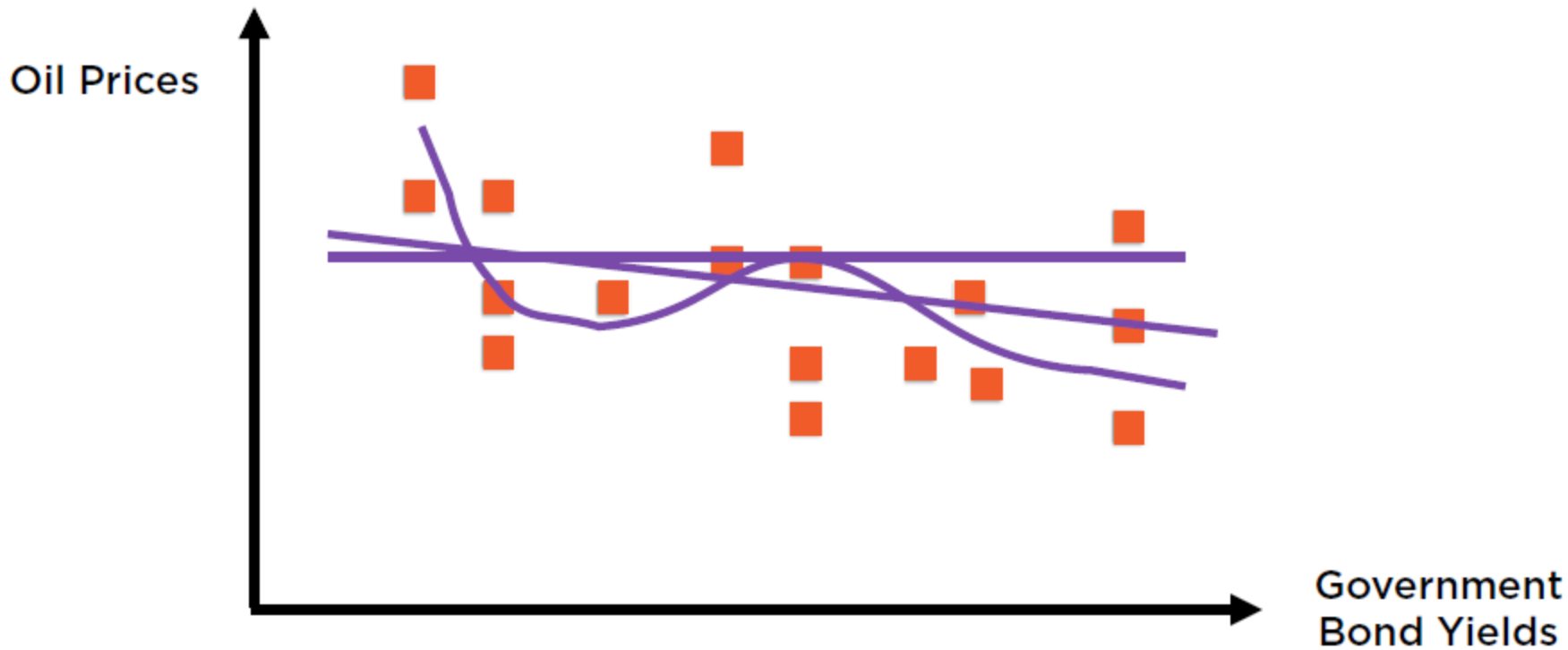
Bidimensional data can be represented in a plane

Data in Two Dimensions



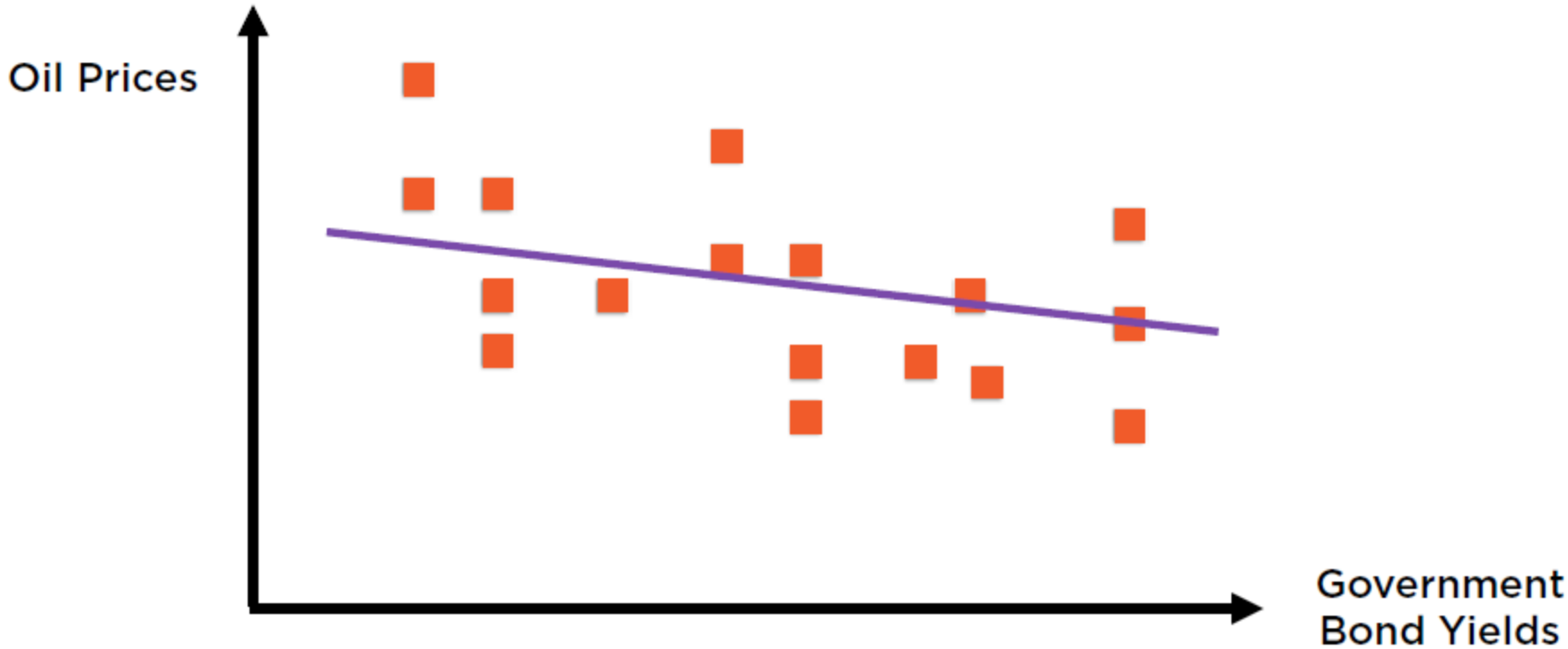
We can draw any number of curves to fit such data

Data in Two Dimensions



We can draw any number of curves to fit such data

Data in Two Dimensions



A straight line represents a linear relationship