

# Introduction to Linear Regression - Recap

## Introduction

This short lesson summarizes the topics we covered in this section and why they'll be important to you as a data scientist.

## Key Takeaways

In this section, the nominal focus was on how to perform a linear regression, but the real value was learning how to think about the application of machine learning models to data sets. Key takeaways include:

- Statistical learning theory deals with the problem of finding a predictive function based on data
- A loss function calculates how well a given model represents the relationship between data values
- A linear regression is simply a (straight) line of best fit for predicting a continuous value ( $y = mx + c$ )
- The Coefficient of Determination (R Squared) can be used to determine how well a given line fits a given data set
- Certain assumptions must hold true for a least squares linear regression to be useful - linearity, normality and heteroscedasticity
- Q-Q plots can check for normality in residual errors
- The Jarque-Bera test can be used to test for normality - especially when the number of data points is large
- The Goldfeld-Quant test can be used to check for homoscedasticity