

Predicting Lyft's Revenues

A CASE STUDY USING LINEAR REGRESSION

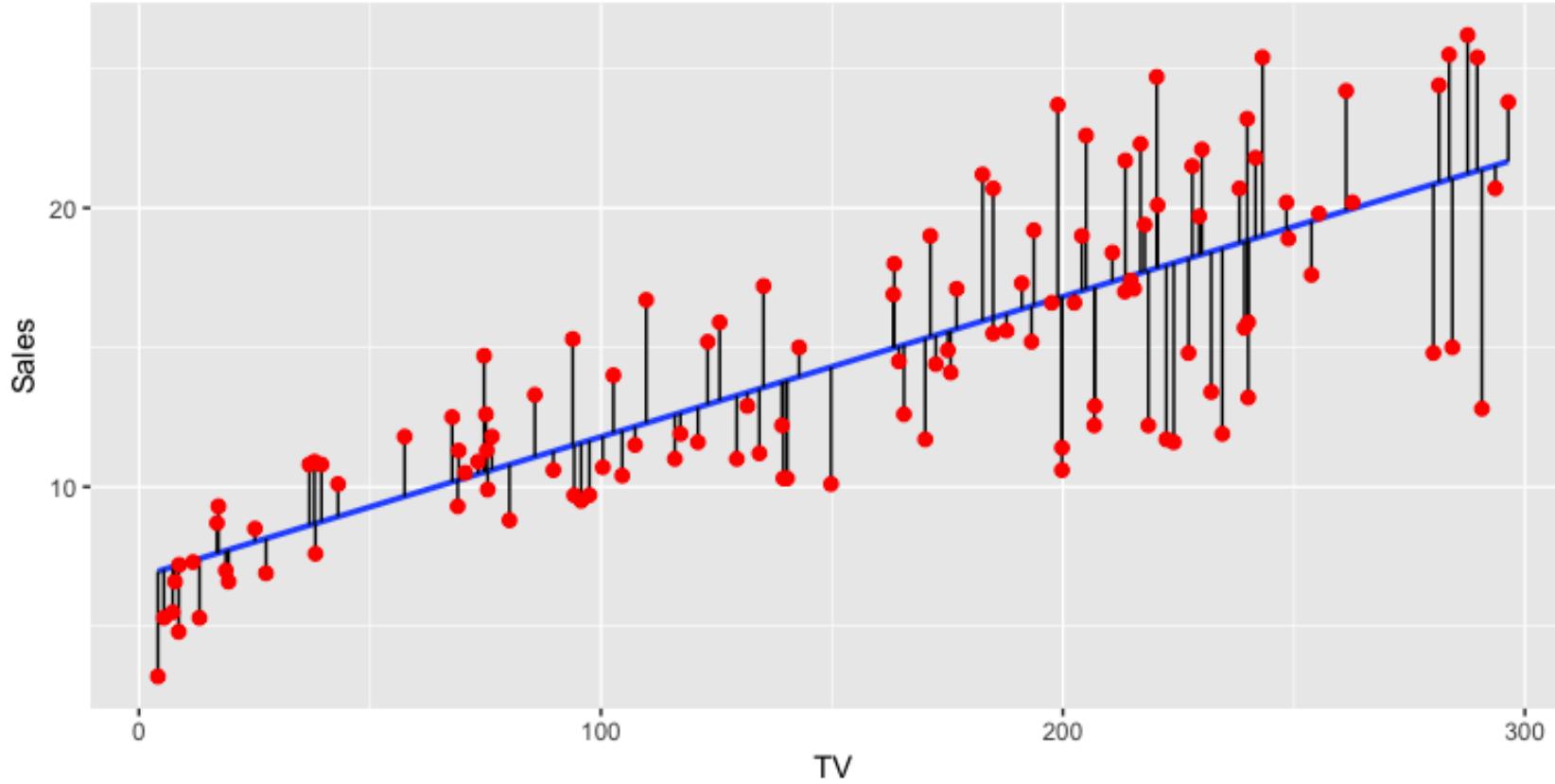
The Problem:

- ▶ Lyft debuted its IPO and the reception was lackluster
- ▶ Potential investors are going to be focused on a number of Key Performance Indicators (KPIs) such as sales and profits
- ▶ Since Lyft won't see profitability for quite some time, sales are very important and for various reasons, specifically, sales
- ▶ Our goal in this exercise is to predict Lyft's revenues based on a dataset that includes features & data related to Lyft's business.
- ▶ While historical data and past sales are not always an indicator of future sales, holding all external factors constant, we should be able to get in the right ballpark.

How We'll Do It:

- ▶ Set up our Jupyter notebook
- ▶ Import our data
 - ▶ Note: Our data is compiled from a Lyft simulation and not actual driver data
- ▶ Do some exploratory data analysis (EDA)
- ▶ Randomly split our dataset into two parts: one for training and one for testing
- ▶ Build our models
 - ▶ Regression models are usually the “go-to” for revenue forecasting and there are many different types of models that are increasingly sophisticated.
 - ▶ Again, for our purposes of learning about regression models, we are going to work with a multivariate regression model.
- ▶ We're going to build two models using the scikit-learn and statsmodels OLS libraries
- ▶ Evaluate its performance

Quick Linear Regression Review



Source: [UC Business Analytics R Programming Guide](#)



So let's head to our Jupyter notebook!.....