

Assignment Overview

Welcome again to the Data Engineering Capstone Project.

This video will give you an overview of the sixth module – Big Data Analytics with Spark.

In this assignment, you will perform a number of tasks to analyze search terms on the e-commerce web server. You will work in Watson Studio within a Jupyter notebook to run your analysis against a CSV file containing the webserver data. You will load this file into a Spark data frame and print the results of your queries against this data set. You will then load a pre-trained sales forecasting model and use this to predict the sales for next year. After performing each task, take a screenshot of the code you used and its output and name the screenshot.

Good luck and let's get started!