This is a broad outline for the Sentiment Analysis Project for a smart phone and tablet app developer. This developer is working with a government health agency to create a suite of smart phone medical apps for use by aid workers in developing countries. The government agency will be providing workers with technical support services, but they need to limit the support to a single model of smart phone and operating system. This will also help to limit purchase costs and ensure uniformity when training aid workers to use the device. After completing an initial investigation, the developer has created a short list of devices that are all capable of executing the app suite's functions.  
  
To narrow this list down to one device, I was tasked with conducting a broad-based web sentiment analysis to gain insight into the attitudes toward the devices.  
  
First, I set up and become familiar with the Amazon Web Services (AWS) computing environment. Next, I used the AWS Elastic Map Reduce (EMR) platform to run a series of Hadoop Streaming jobs that collected large amounts of smart phone-related web pages from a massive repository of web data called Common Crawl. By "massive" I mean mining approximately 150,000 web pages. Once gathered, I compiled the data into a Small Data Matrix where I used the machine learning package called WEKA to develop a predictive model that labeled the data with the web sites' sentiment toward the devices. Finally, I applied the predictive model to the data using Java.