Airflow Project :

Task 1: read csv file from s3 bucket

Task 2: run spark job with that csv file using TimeSensor

Task 3: send o/p to s3 bucket.

Dependency:

Task 3 is dependent on task 2 and task 2 is dependent on task 1.

Dag Name: Timberland stock analysis

Timberland\_stock.csv:

**What day had the Peak High in Price?**

#### What is the mean of the Close column?

#### What is the max and min of the Volume column?

#### How many days was the Close lower than 60 dollars?

#### What percentage of the time was the High greater than 80 dollars ?

#### What is the Pearson correlation between High and Volume?

#### What is the max High per year?

#### What is the average Close for each Calendar Month?