Lab 1

CS 327E
September 18, 2017

1) Amazon Athena uses a query engine called Presto under the hood.

A)True
B)False

2) Presto uses ANSI SQL as its query language.

A)True
B)False

3) Athena "tables" are metadata definitions of delimited files (e.g. csv, tsv, etc.) or other structured files. These metadata definitions are called Hive DDL statements.

A)True
B)False

4) Results of queries in Athena are stored in S3 buckets on AWS.

A)True
B)False

5) Why did Netflix choose Presto for data exploration?

A)To solve their problems at scale
B)Well integrated with Hive metastore
C)Easy to integrate with S3
D)All of the above

Demo Time
(LucidChart, Git, Stache)

Lab 1

Assignment Sheet: http://www.cs.utexas.edu/~scohen/projects/lab1.pdf
Rubric: http://www.cs.utexas.edu/~scohen/projects/lab1-rubric.pdf