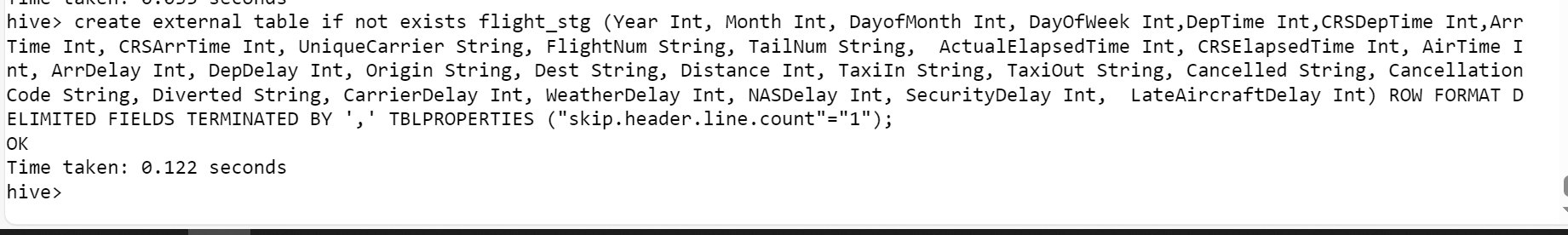
Airports, Carrier, Plane Info,Flights

A close up of a text

Description automatically generated with medium confidence



1. Data Setup in HDFS

Step 1: Copy the provided csv files in a directory named AirlinesOntimeAnalysis

A screenshot of a computer program

Description automatically generated

Step 2: Unzip the contents for the bz2 files. Note => compressed to decompressed

A number symbols with numbers

Description automatically generated with medium confidence

Step 3: Create the directories in hdfs where your data will reside

Step 4: copy the respective files to the above created hdfs structure

A screenshot of a computer

Description automatically generated

2. Create tables using Hive

A close-up of a document

Description automatically generated

1. Improving the performance of Hive tables - Parquet format

A close up of a text

Description automatically generated

Loading the data =>

A close up of a document

Description automatically generated

A white screen with black text

Description automatically generated

What is the average arrival delay and average departure delay in each month of the year 2004?

A white screen with black text

Description automatically generated

A close-up of a computer screen

Description automatically generated