In relation **A**, we are loading the dataset using CSVExcelStorage because of its effective technique to handle double quotes and header.

In relation **B**, we are generating the columns which are required for processing and explicitly typecasting each of them.

In relation**C1**, we are removing the null values fields present if any.

In relation **D1**, we are grouping the data based on column “origin.”

In relation **E1**, we are finding average delay from each unique origin.

Relations named **Result** and **Top\_ten** are ordering the results in descending order and printing the top ten values.

These steps are good enough to find the top ten origins with the highest average departure delay.

However, rather than generating just the code of origin, we will be following a few more steps to find some more details like country and city.

In the relation **Lookup**, we are loading another table to which we will look up and find the city as well as the country.

In the relation L**ookup1,** we are generating the destination, city, and country from the previous relation.

In the relation **Joined**, we are joining relation Top\_ten and Lookup1 based on common a column, i.e., “origin.”

In the relation**Final,**we are generating required columns from the Joined table.

Finally, we are ordering and printing the results.

**Below are the screenshots for each steps and output.**















































