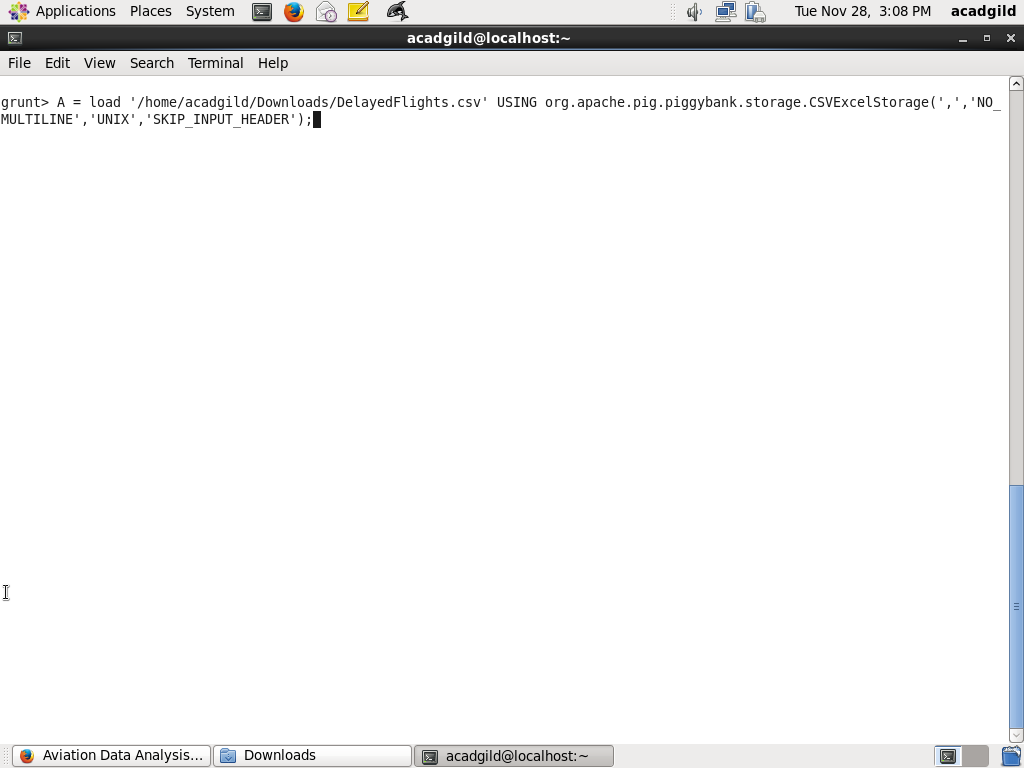
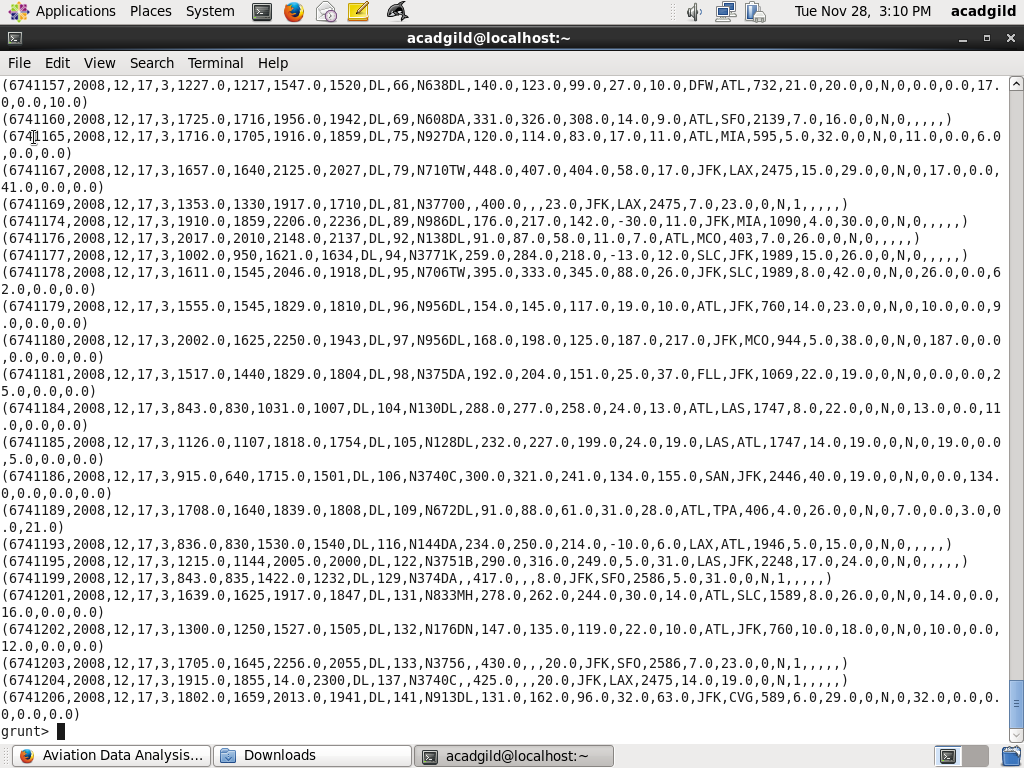
In relation **A**, we are loading the dataset using CSVExcelStorage because of its effective technique to handle double quotes and header.

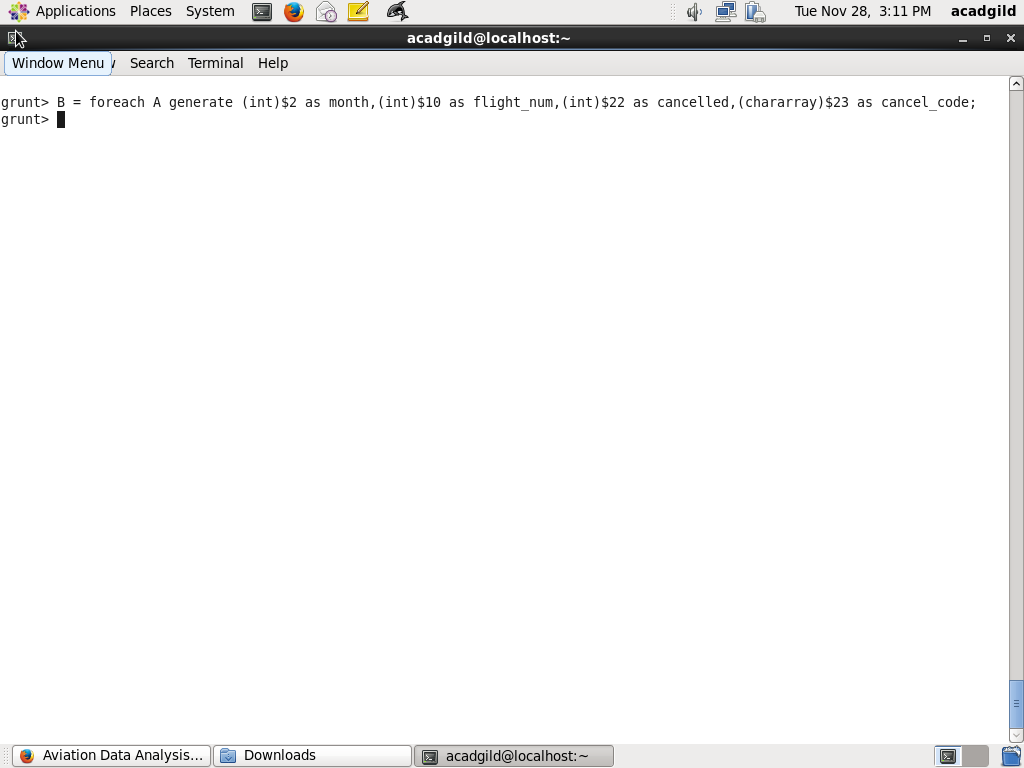
Below it screenshot of step and derived output

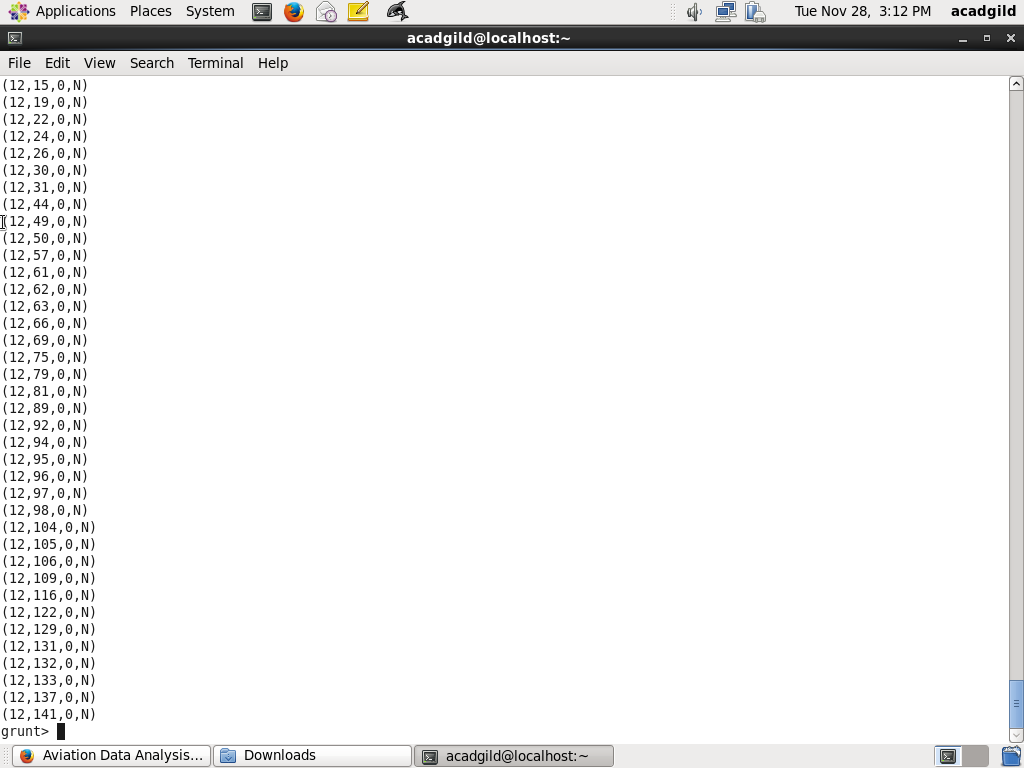




In relation **B**, we are generating the columns which are required for processing and explicitly typecasting each of them.

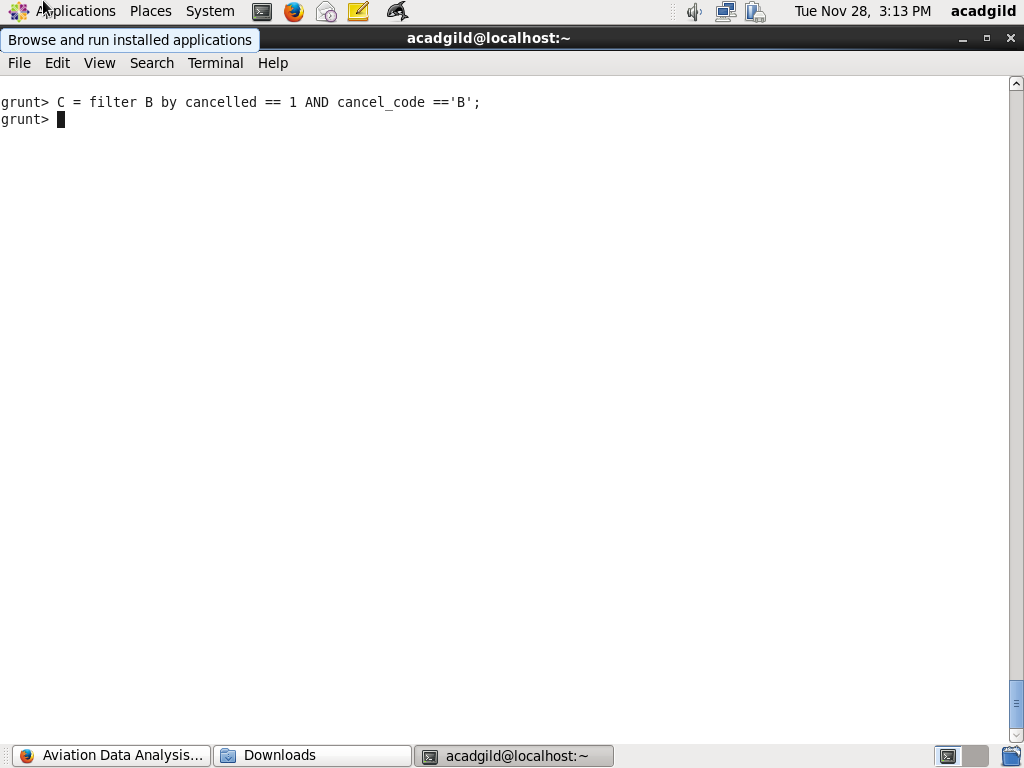
Below it screenshot of step and derived output

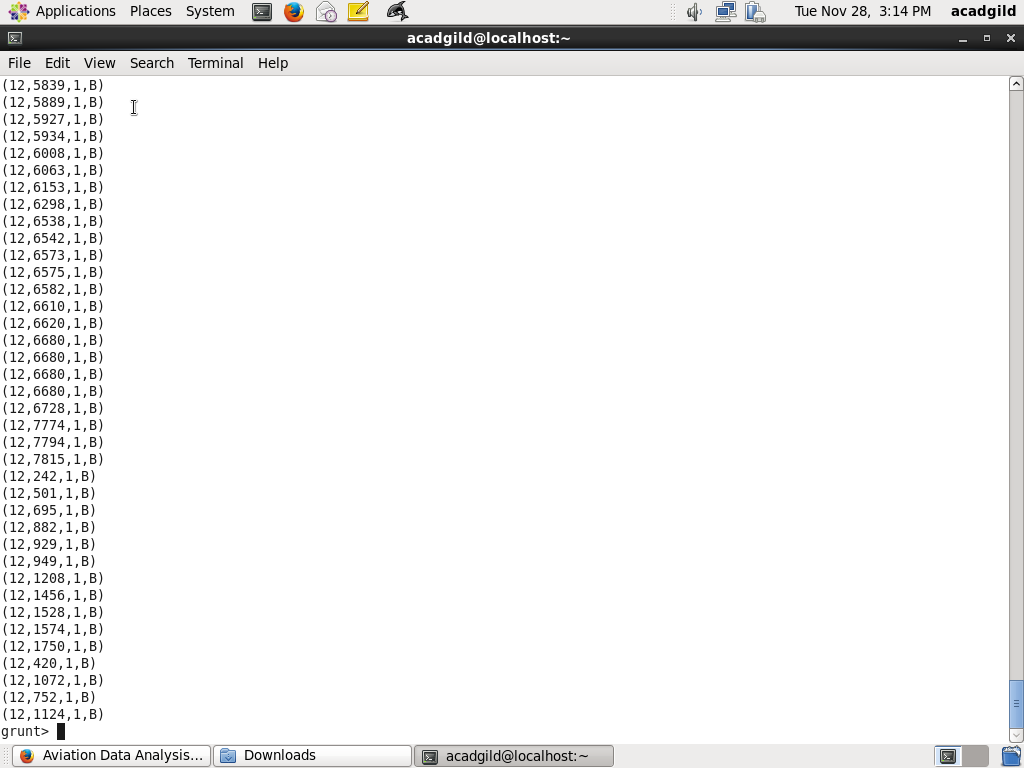




In relation **C**, we are filtering the data based on cancellation and cancellation code, i.e.,  canceled = 1 means flight have been canceled and cancel\_code = ‘B’ means the reason for cancellation is “weather.” So relation C will point to the data which consists of canceled flights due to bad weather

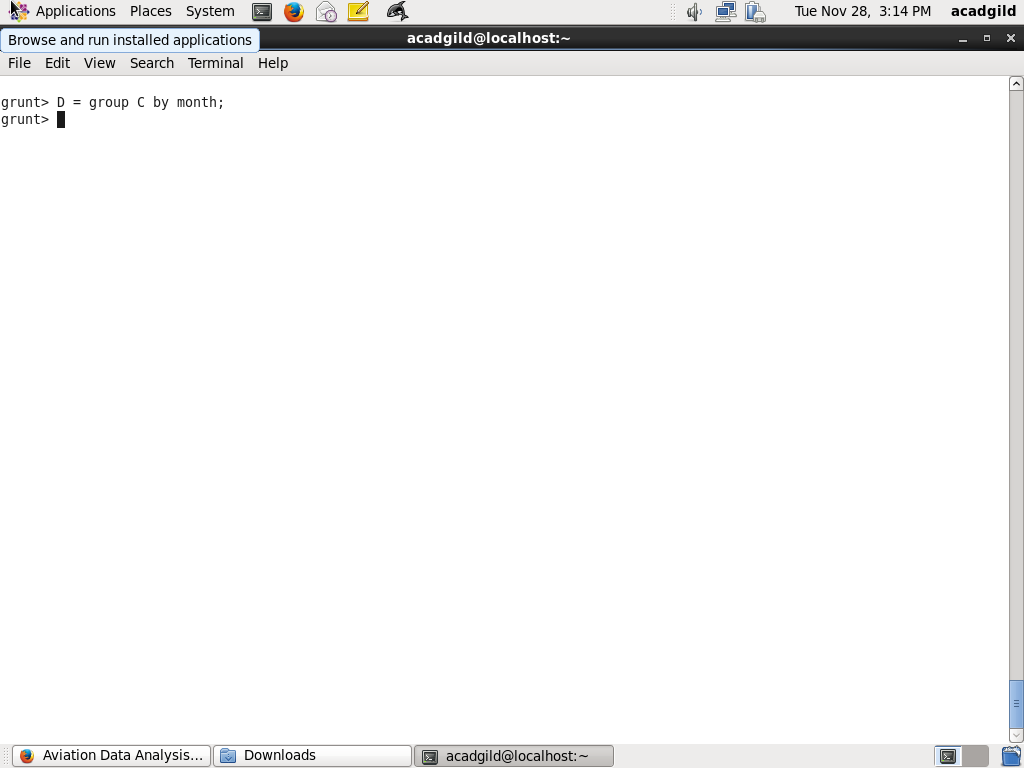
Below it screenshot of step and derived output

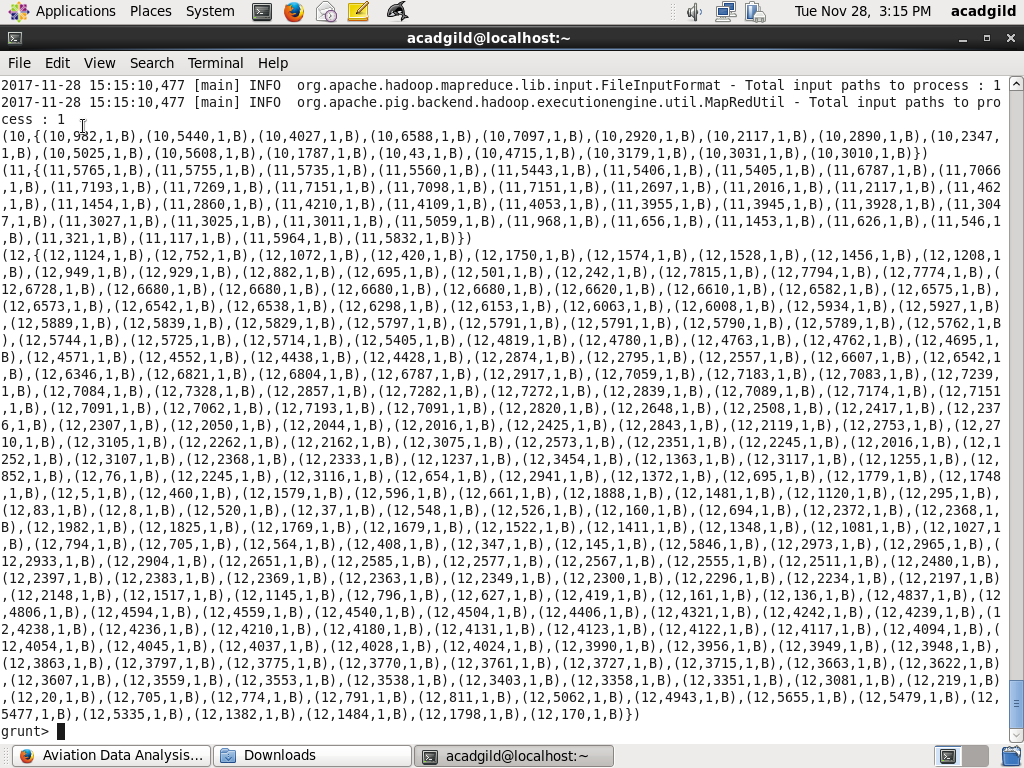




In relation**D**, we are grouping the relation C based on every month.

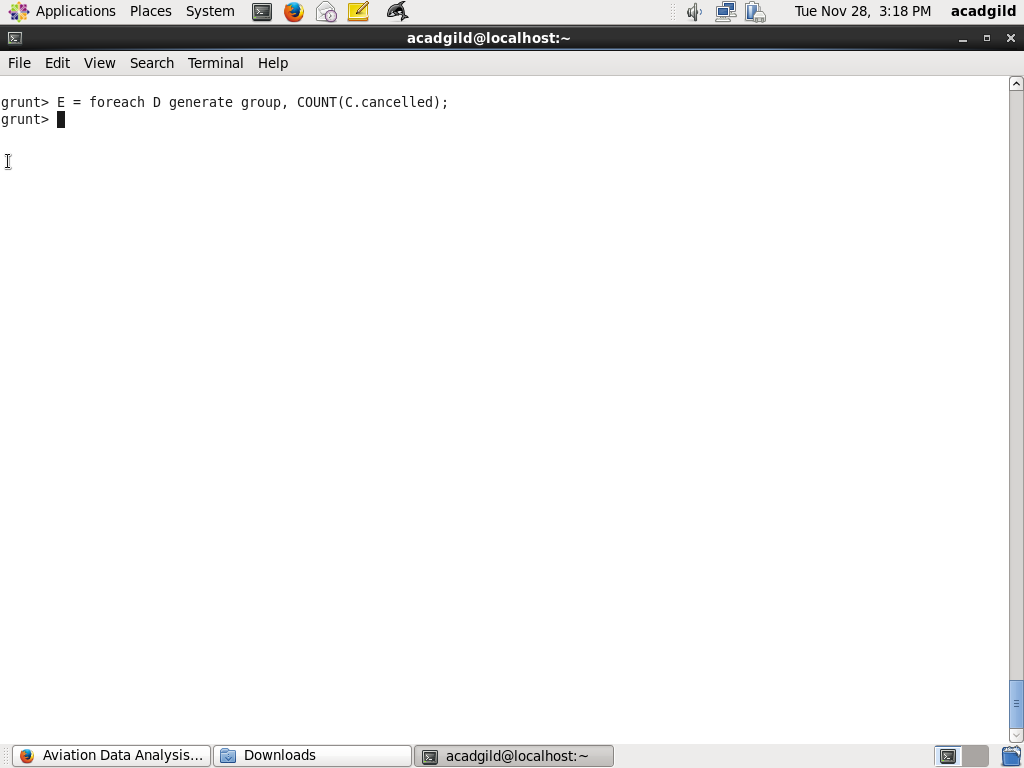
Below it screenshot of step and derived output

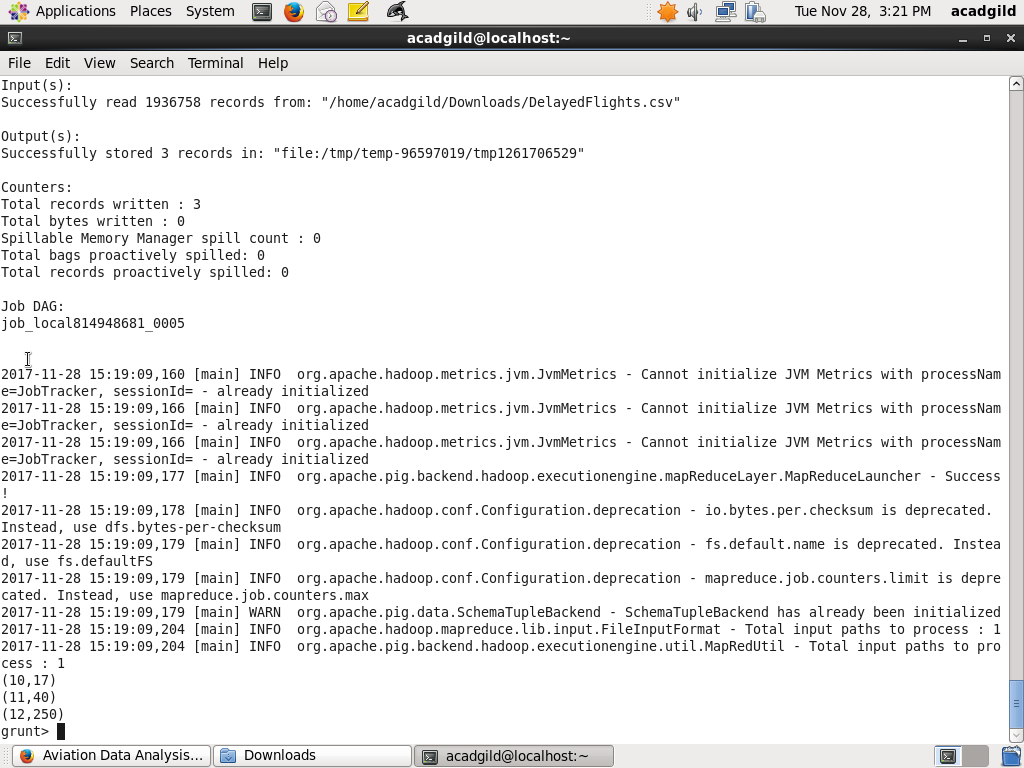




In relation **E**, we are finding the count of canceled flights every month.

Below it screenshot of step and derived output





Relation **F** and **Result** is for ordering and finding the top month based on cancellation.

Below it screenshot of step and derived output

