Real-Time Telecommunication Data Processing

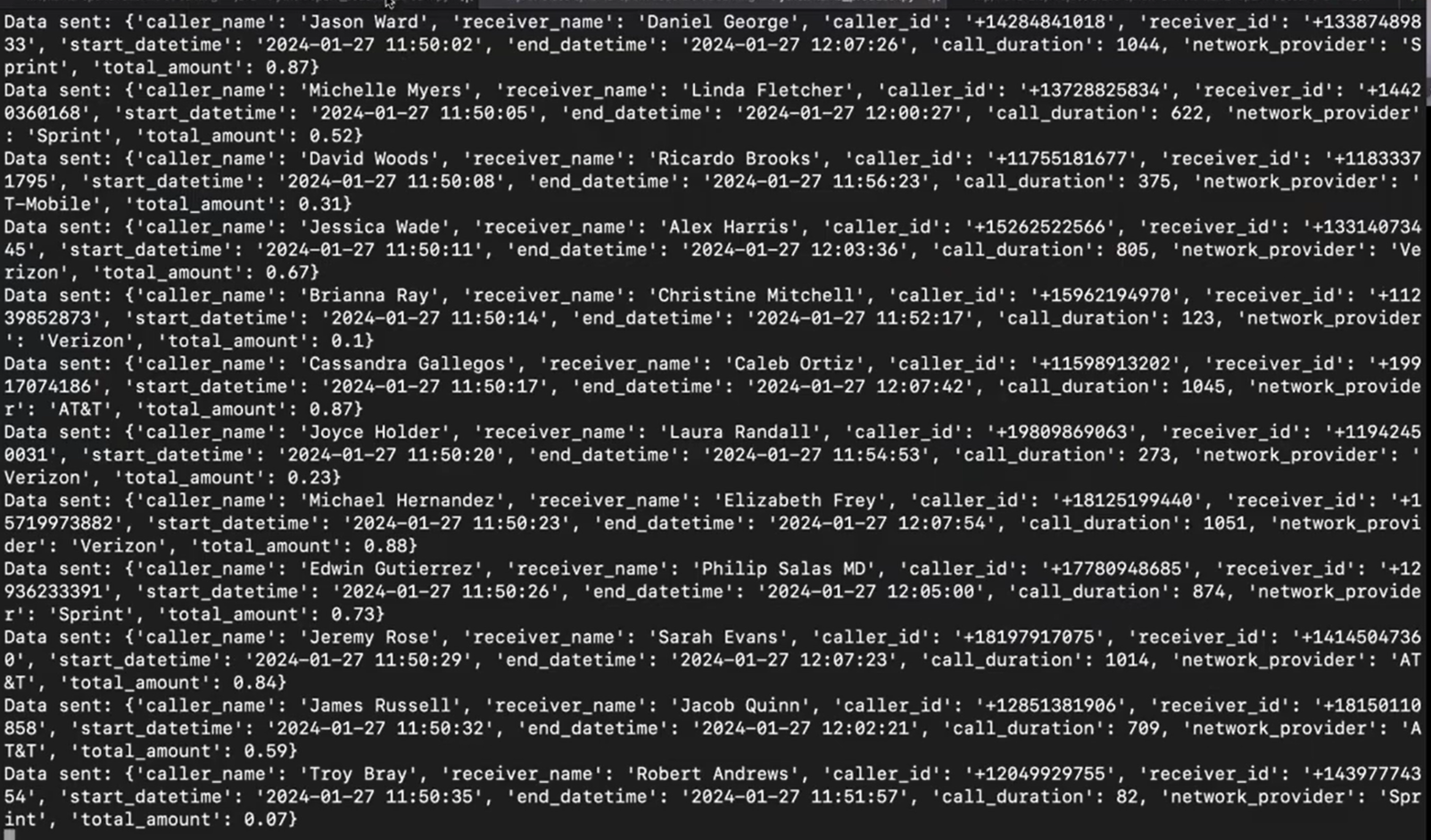
(Kafka-Redshift)

1. We had a producer script which was dumping telecommunication data on kafka topic. Then from there we had a streaming application which doing some filtrations and ingesting data to redshift warehouse in real time. We dumped data into warehouse & we had a quick sight dashboard on the top of it.

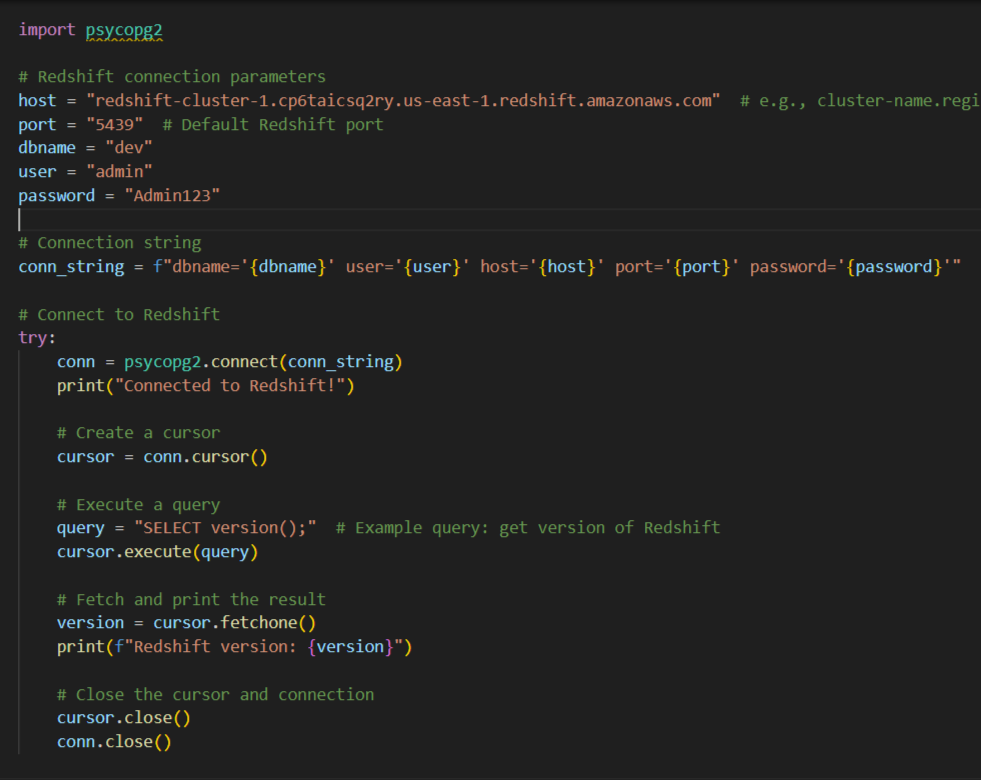
producer script:



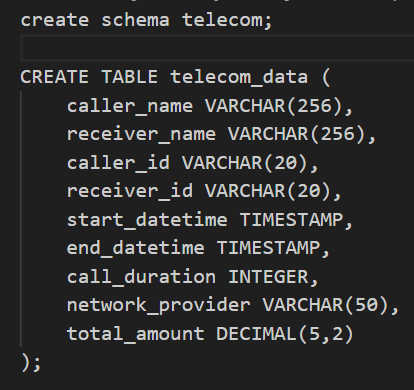
data getting produced to kafka topic:



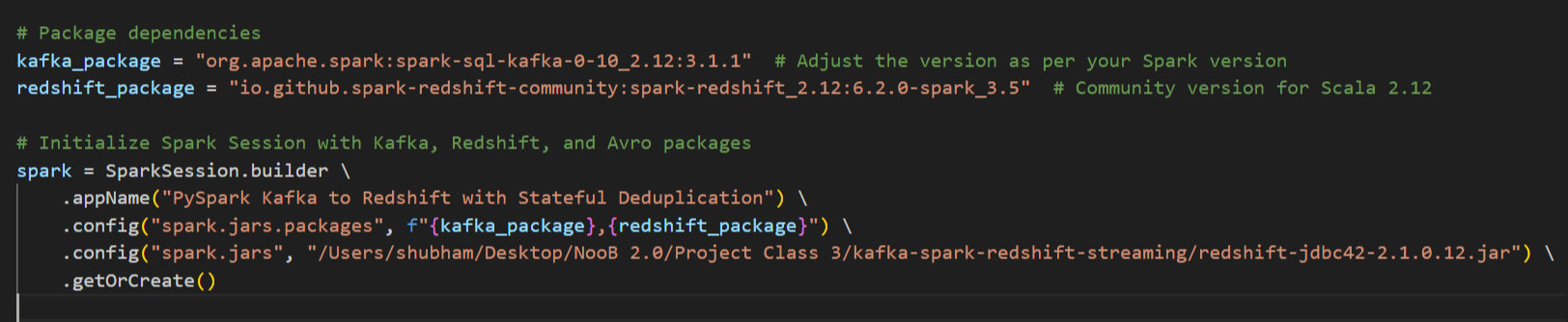
1. we created inbound rule like from where the inbound traffic can come to the port. So actually we connected our local script with the aws redshift.



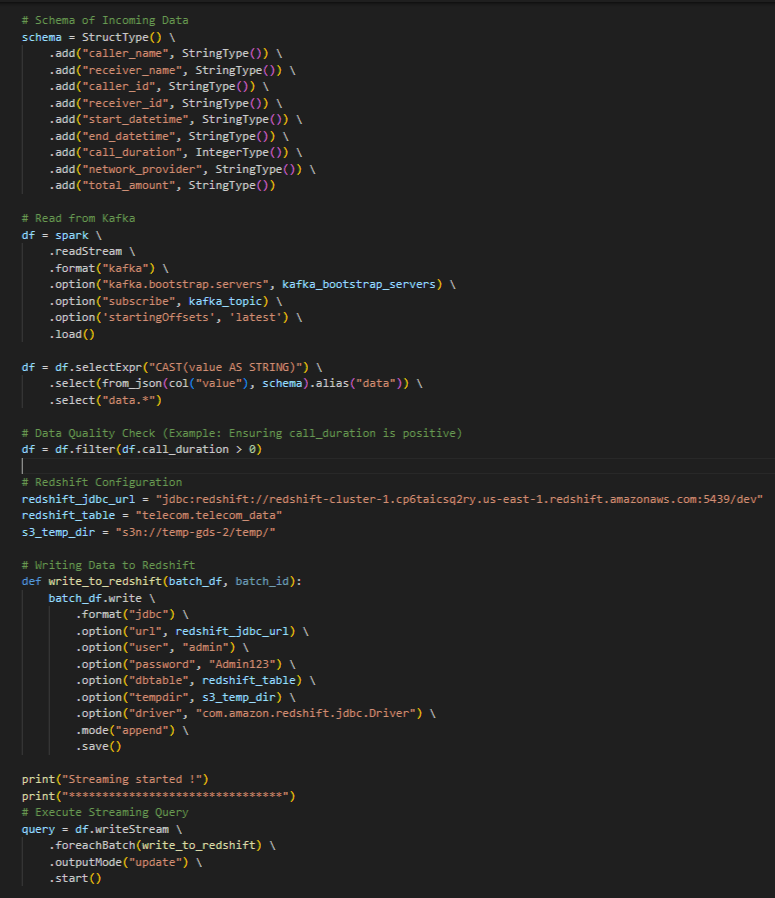
1. On the other hand, we created then redshift schema and table where we will be ingesting the data.



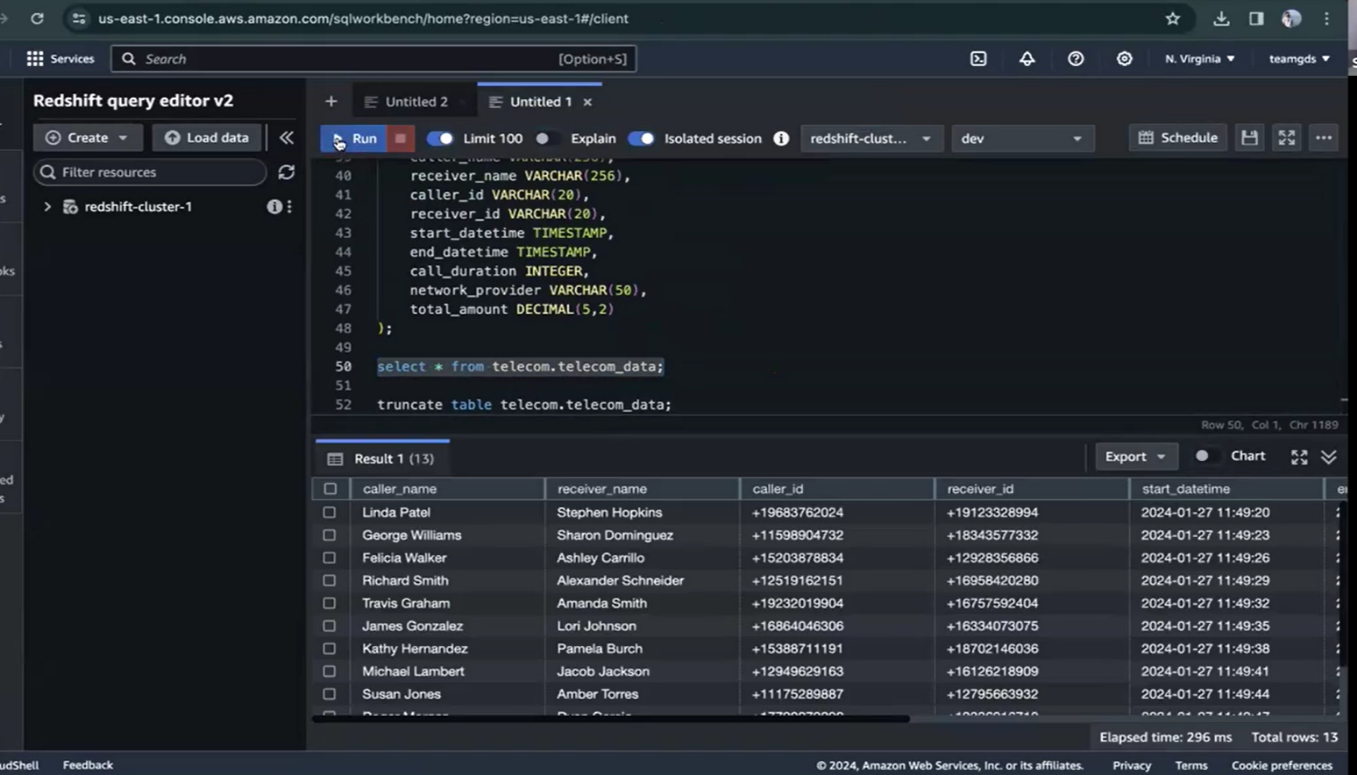
1. Now we had streaming app which will be connecting to kafka and redshift so we included kafka-spark & redshift-spark connector jar file respectively. We also needed redshift jdbs connector because we were writing data in jdbc format as per our spark version



1. We created next schema of incoming data from kafka topic and used readStream to read data from kafka topic. Hence did some transformations and writing data to redshift using writeStream.

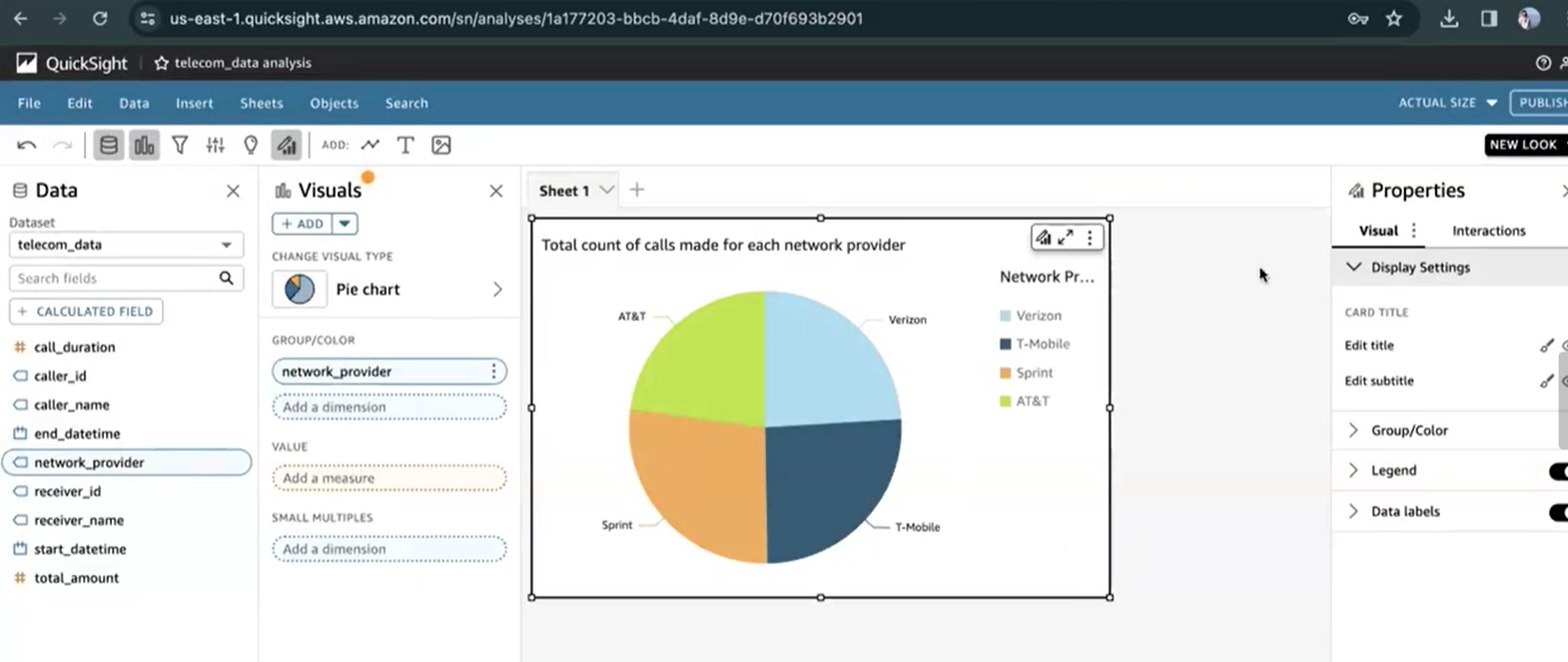


Redshift table results:



1. next we worked on aws quicksight. Created dataset and made dashboard on the top of it like network provider wise calls, network provider wise total call duration

network provider wise calls:



network provider wise total call duration:

