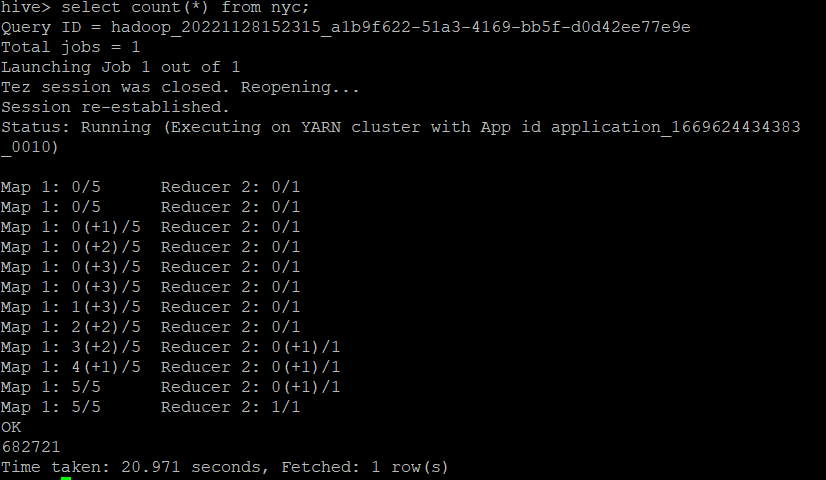
Name – Nishad Purohit

Email – [Nishad.purohit20@gmail.com](mailto:Nishad.purohit20@gmail.com)

Hive assignment

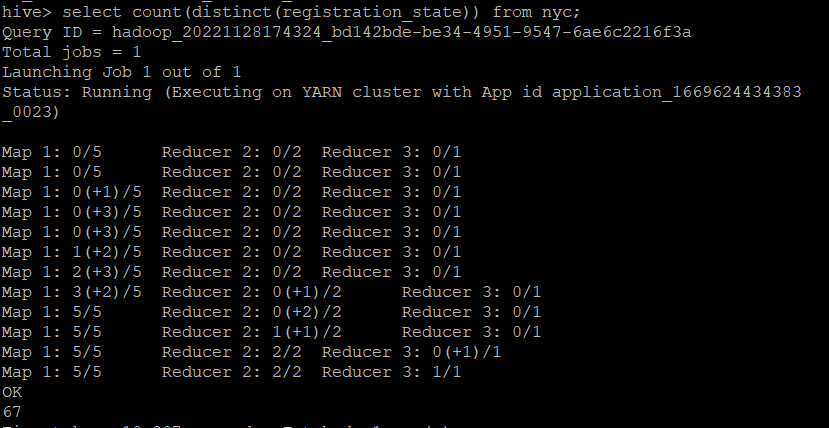
* Part-I: Examine the data

1. Find the total number of tickets for the year.



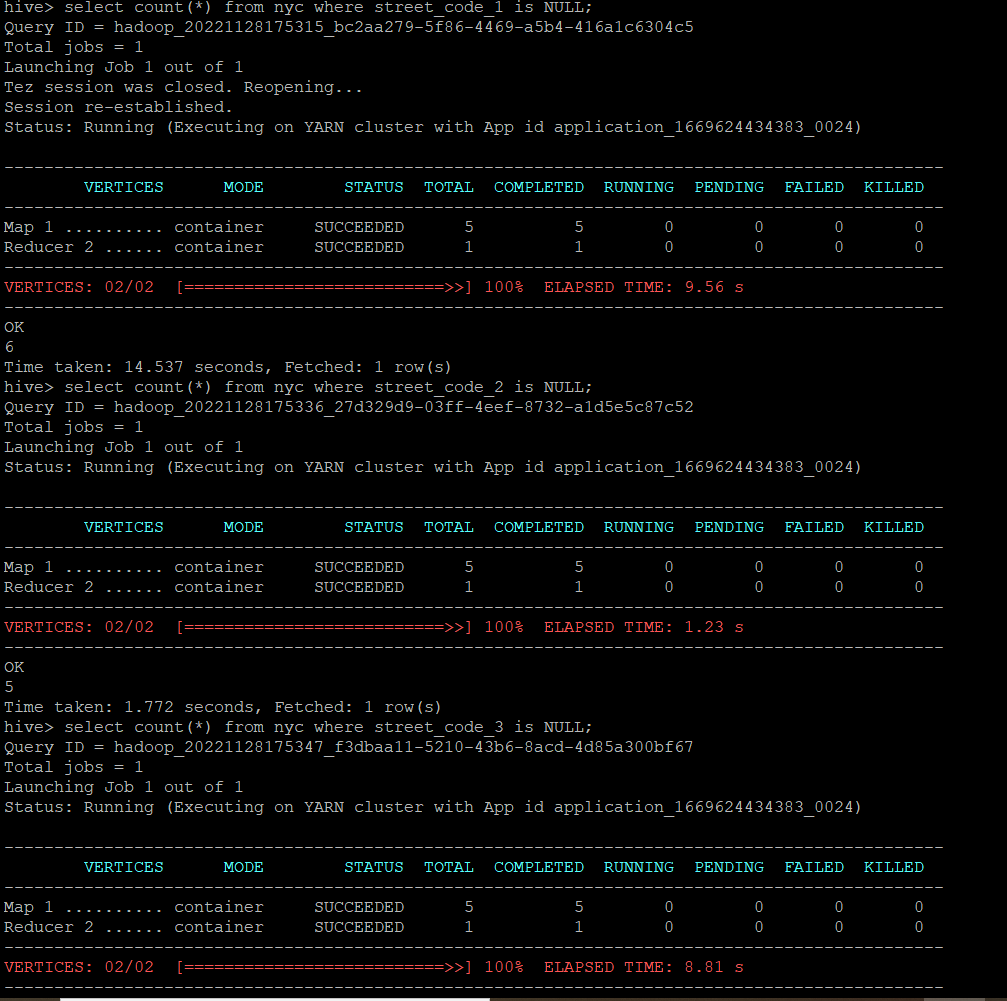
The total number of tickets is 682721.

1. Find out the total number of states to which the cars with tickets belong.



There are total 67 states with tickets.

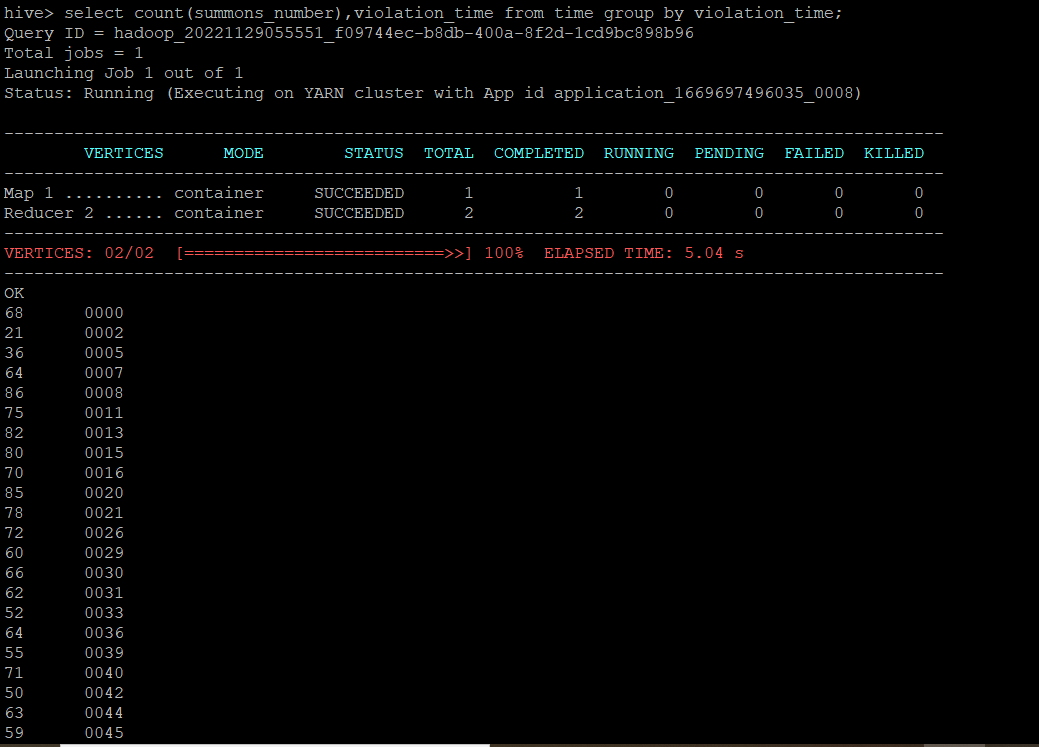
1. Some parking tickets don’t have addresses on them, which is a cause for concern. Find out the number of such tickets which have no addresses.



The count of values with null as address for street 1 is 6, street 2 is 5 and street 3 is 5 so the total of these values is 16.

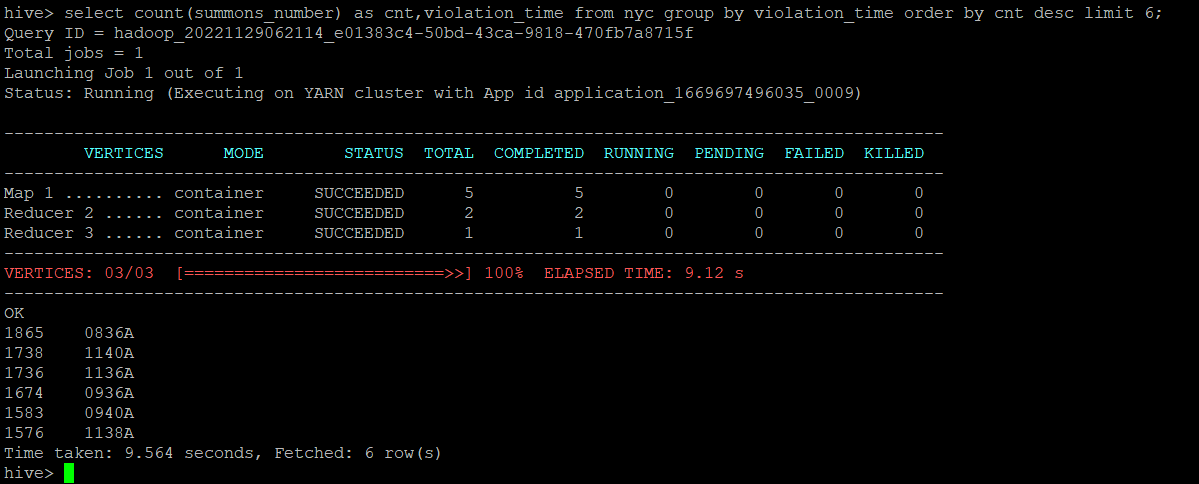
* Part-II: Aggregation tasks

1. Find out the frequency of parking violations across different times of the day:

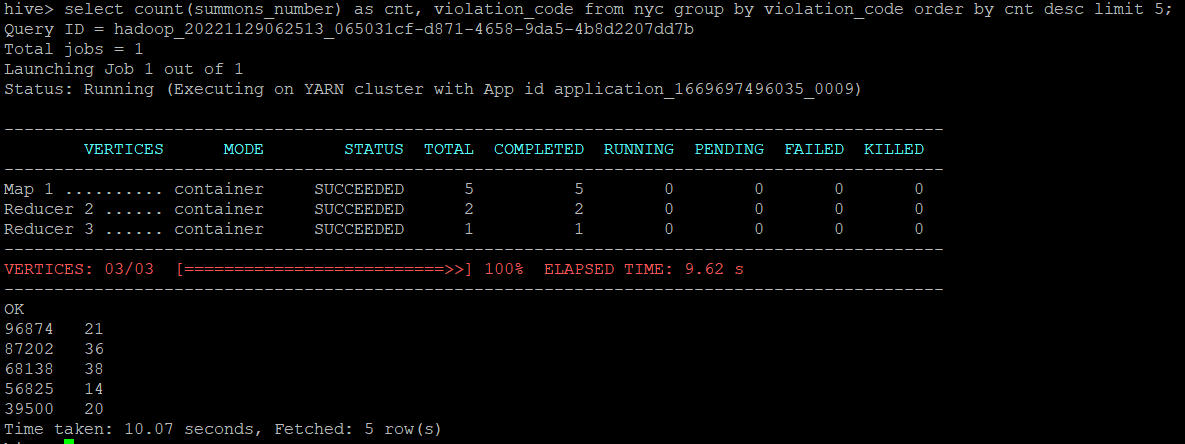


This the count of tickets issued on different times of the day.

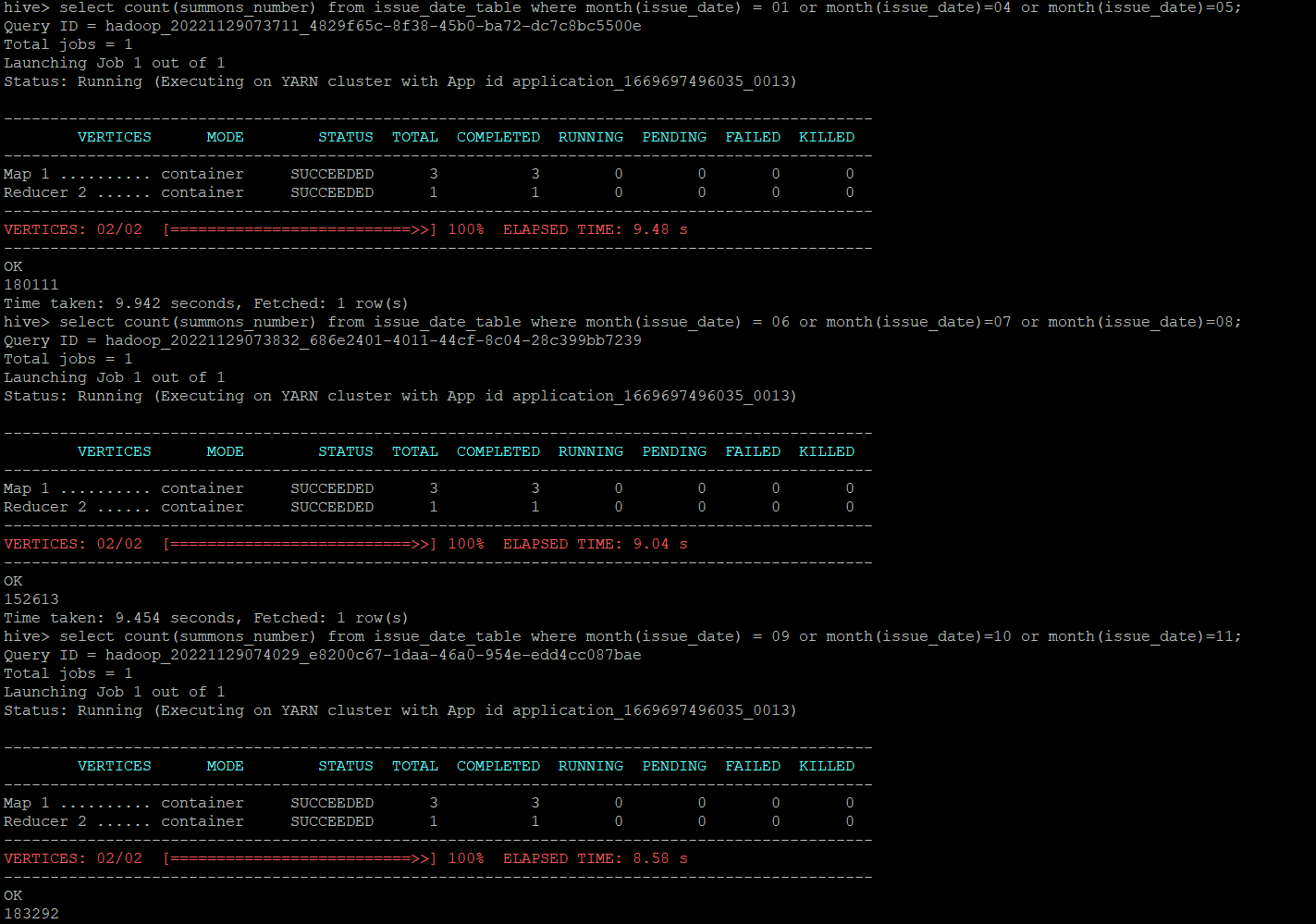
1. Divide 24 hours into six equal discrete bins of time. The intervals you choose are at your discretion. For each of these groups, find the 3 most commonly occurring violations.

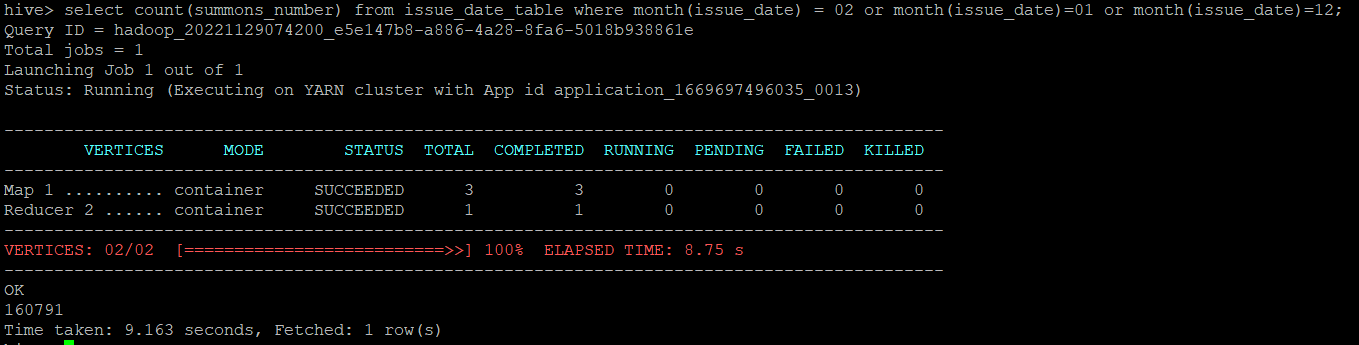


1. Now, try another direction. For the 3 most commonly occurring violation codes, find the most common times of day



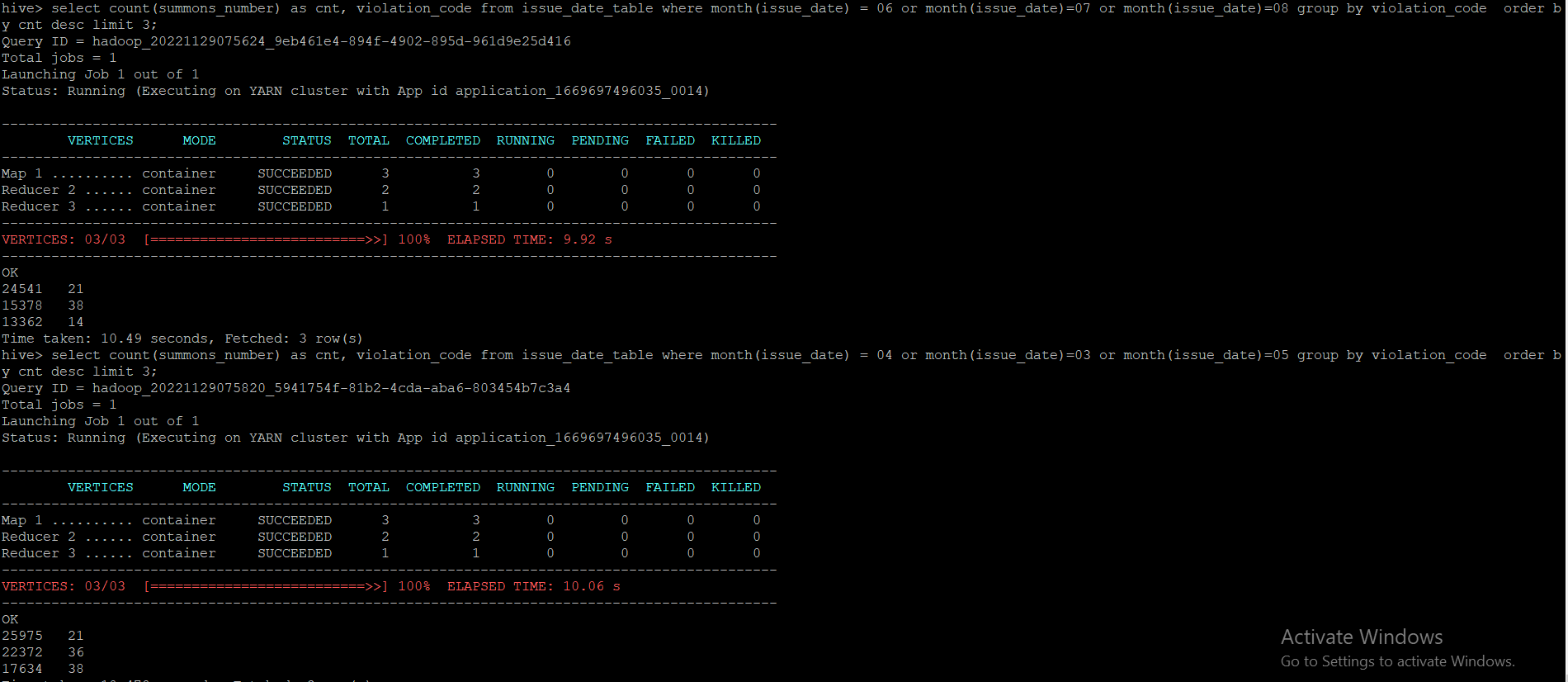
1. Let’s try and find some seasonality in this data:
2. First, divide the year into seasons, and find the frequencies of tickets for each season.

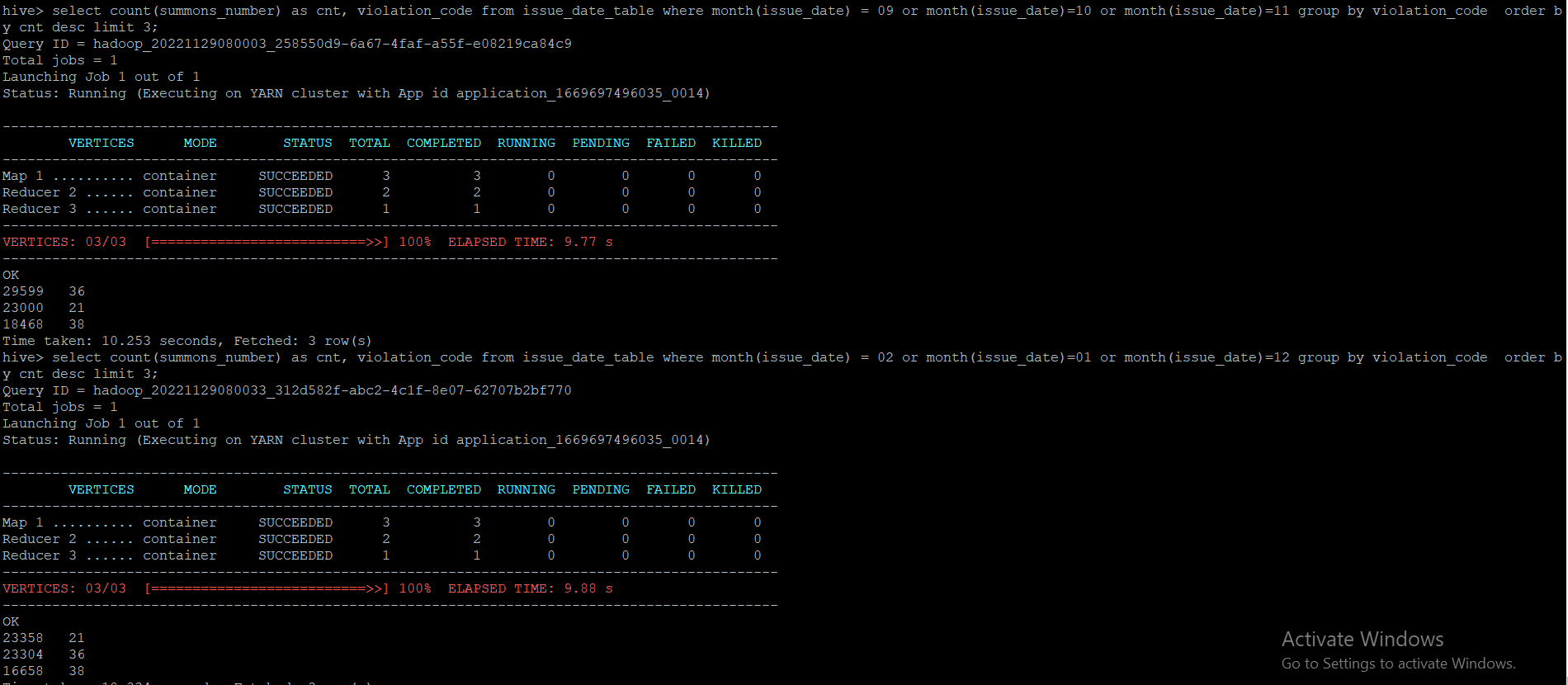




As we can see here is the count of tickets as per seasons.

1. Then, find the 3 most common violations for each of these seasons.





As we can see here are the top 3 violation codes as per seasons.