**SUMMARY**

select (std(d.dist)/avg(d.dist))\*100 from(

select Extract(hour from activity\_time) as hours, count(\*) as dist

from UserActivity inner join User on User.user\_id = UserActivity.user\_id and User.email = 'dr\_abdul.musa01@msn.com' group by hours) as d;

This is the code I am using to identify the cheaters.

The above code will give me the Percentage of standard deviation from the average of the hourly activity of a particular user.

Here I am using the email of the user to identify this value.

We can also use the user\_id as the identifier.

**Code with user\_id as identifier.**

select (std(d.dist)/avg(d.dist))\*100 from(

select Extract(hour from activity\_time) as hours, count(\*) as dist

from UserActivity inner join User on User.user\_id = UserActivity.user\_id and User.user\_id = 100005

group by hours) as d;

Here, I am calculating the standard deviation of the hourly distribution of a particular user and also the average of the hourly activity distribution of them. Then I am trying to find out the percentage of standard devation from the average value. If this value is small like in the range of (0-10), then the user can be identified as a cheater or a fraud user as this means that the hourly distribution is uniform throughout the hours. On the other hand, if the value is large like in the range of (35+) then that means the hourly distribution of the user is not uniform, hence it can be considered as a non-cheating user.