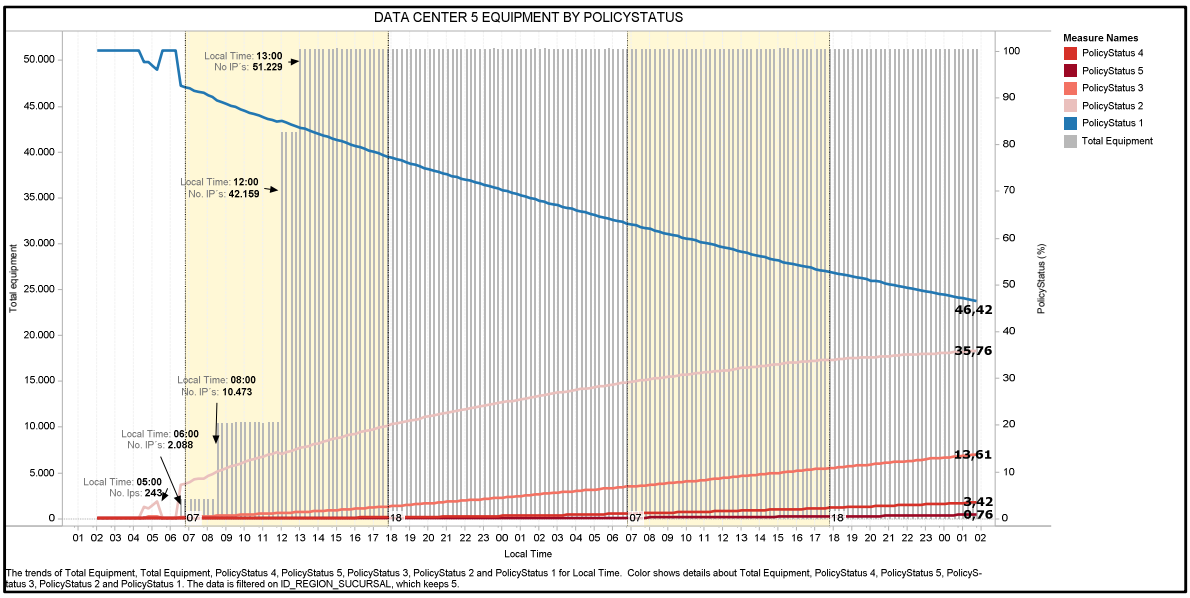
QUESTION 2

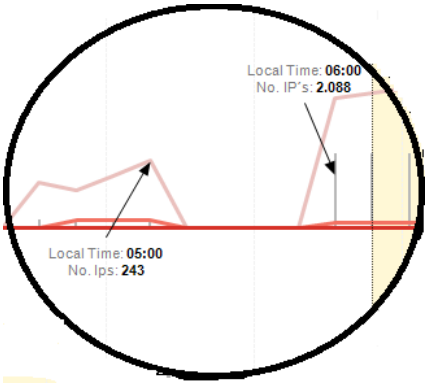
For answering question 2, first we generated a unique code for each branch using business unit and facility. Then we summarized data in terms of policy status, activity flag and number of connections. For policy status and activity flag, we calculated the percent of machines that were active by category for each branch. For number of connections we calculated the maximum, minimum, average and standard deviation by branch. Then we explore data using tables and graphs. After, we selected graphs that clearly identify the anomalies found in the Bank of Money.

Anomaly 1

For detecting this anomaly, first we obtained the hours worked by machine; then we summarize by branch. After we calculated maximum, minimum, mean and deviation by each branch. Data center 5 register the less number of hours worked.

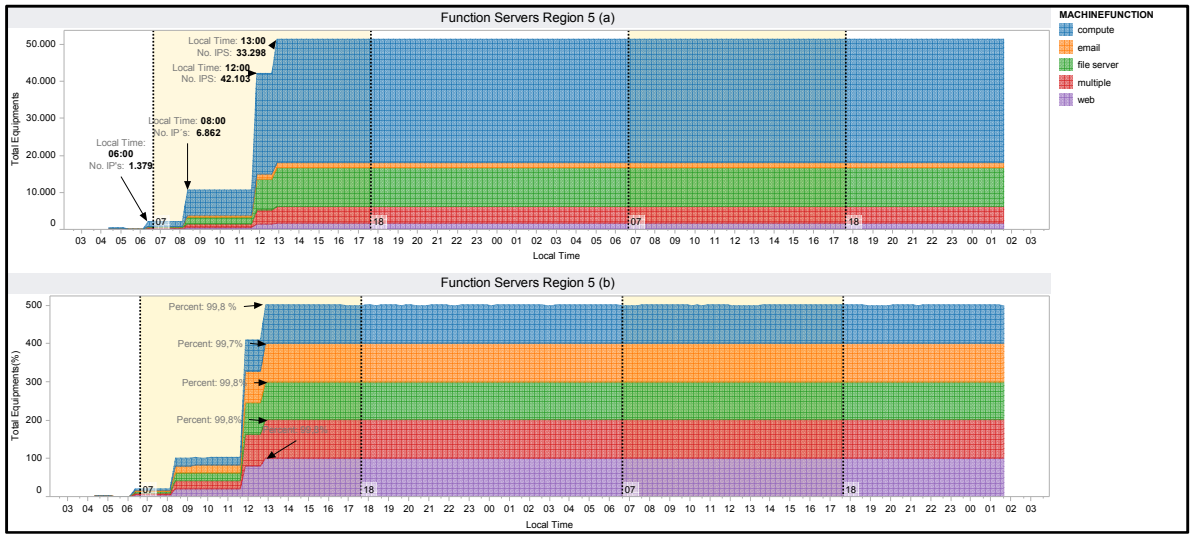
This graph is a representation of Datacenter 5, for which there are records since 2:00 am on February 2. Through the gray lines, we see how the number of active machines increases as the hours pass, so that just at 01:00 pm most machines are active; this behavior differs of the other data centers, as we can see in the other plots. In addition, there was an increase in the number of machines in use, between 4:30 and 5:30 am.





Anomaly 2

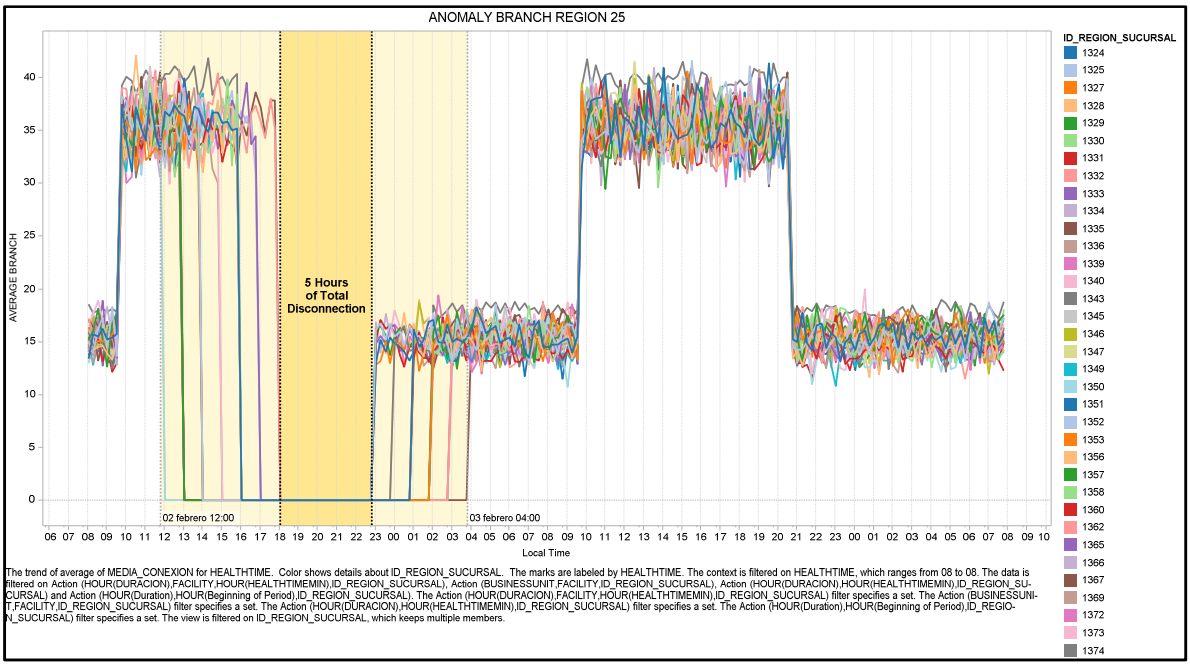
In data center 5, except 5 machines, the rest are servers. Thus, in the upper graph we see the number of servers according to their function, whereas in the lower part we see the percent of machines that are active by function, which shows that the machines are activated in the same proportion in each observed change. Therefore, it is assumed that there might be some planning to activate servers, but we does not distinguish what might be the reason why there are few servers running in the early business hours.

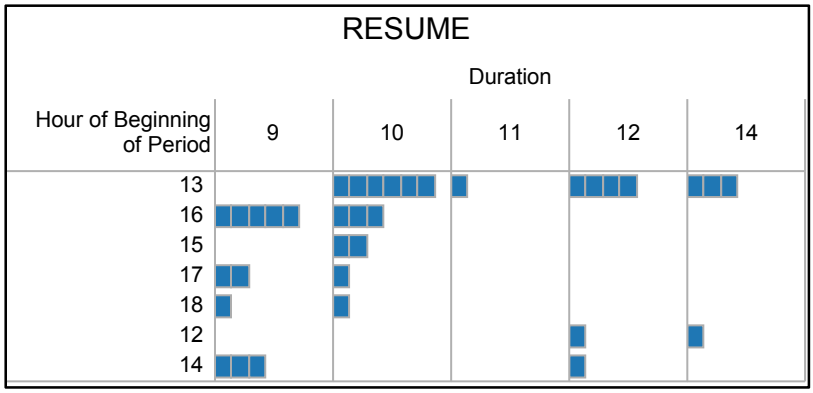


Anomaly 3

For each branch we missing hours and create intervals for each computer. With 35 branches intervals matching detected in 5 hours. This region is the 25.

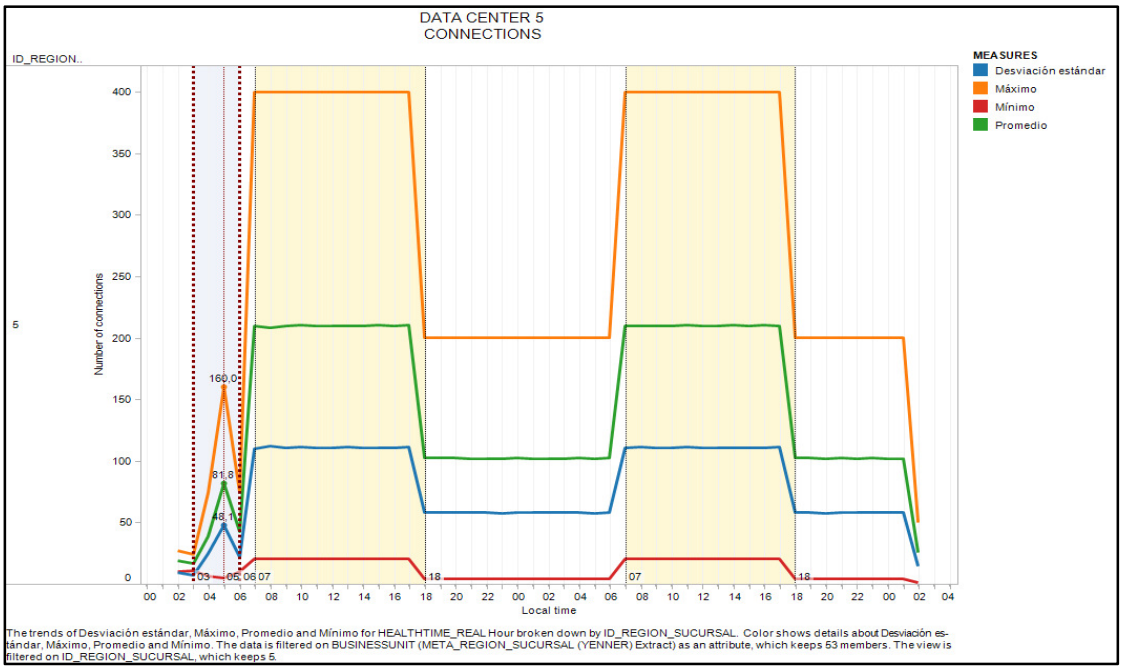
We observe as another anomaly that machines of 35 branches of region 25 are switched off for several hours. As you can see in the table, both start and duration of the break happened at different times for each branch. In addition, there is no previous evidence about machines that going down for maintenance. Apparently, it could be a sudden break.



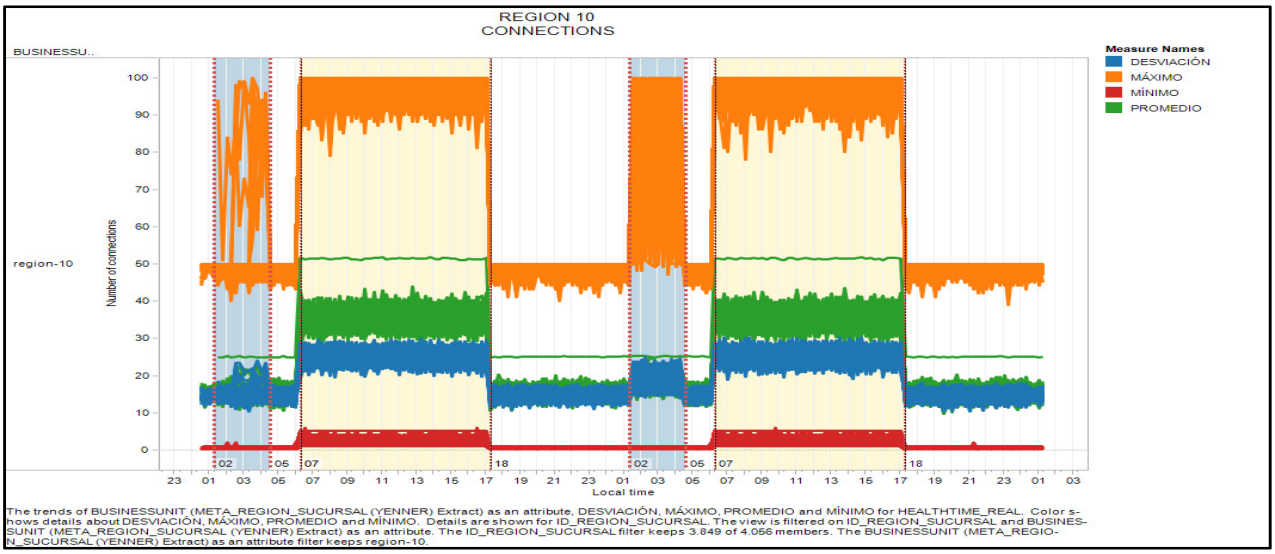


Anomaly 4

Considering the number of connections, we can see that in most branches of the Bank of Money, the number of these is constant outside of business hours. But in data center 5, on February 2, between 03:00 and 06:00 am, there is a variation in the number of connections. This change could be due to an automatic update of a program for example.



Anomaly 5

In region 10, on February 02, from 2:00 to 5:15 a.m., there was an increase in the maximum number of connections of various branches, which is determined as an anomaly. We see for all branches the same behavior, on 03 February, from 2:00 to 5:15.

This anomaly could be explained by a process that is scheduled to run in this time.