



**Computer Science And Engineering  
University of Dhaka**

**Title:** Comparison of data mining on different data sets

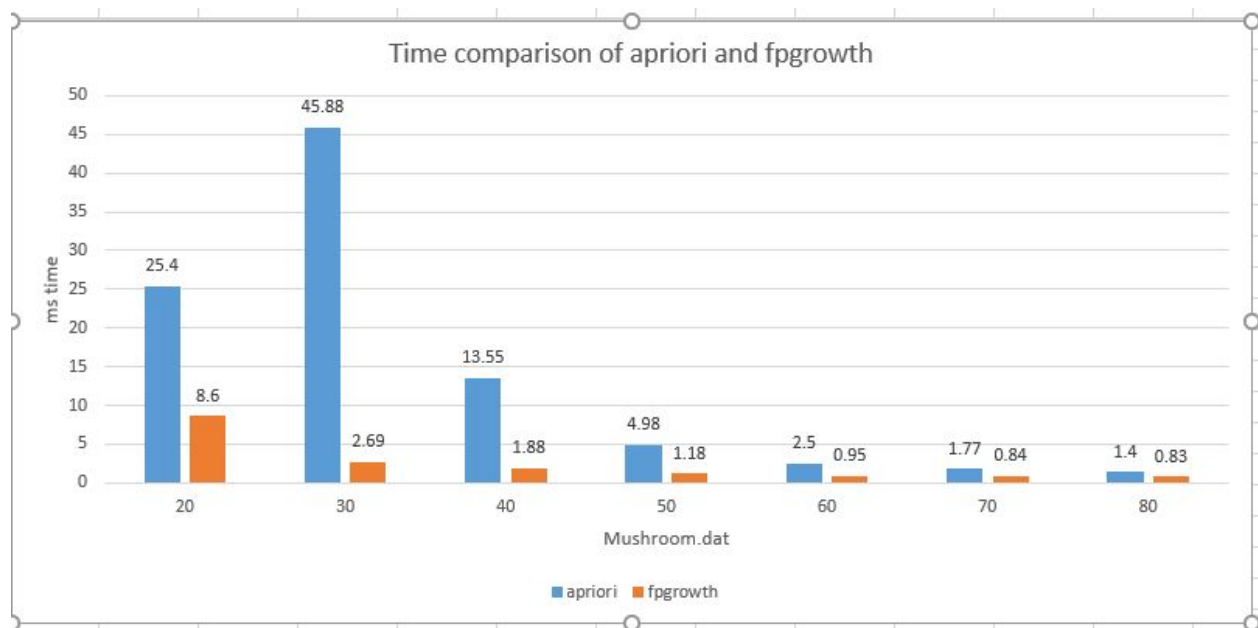
**Submitted by:**

Md. Mizanur Rahman  
Roll: 60

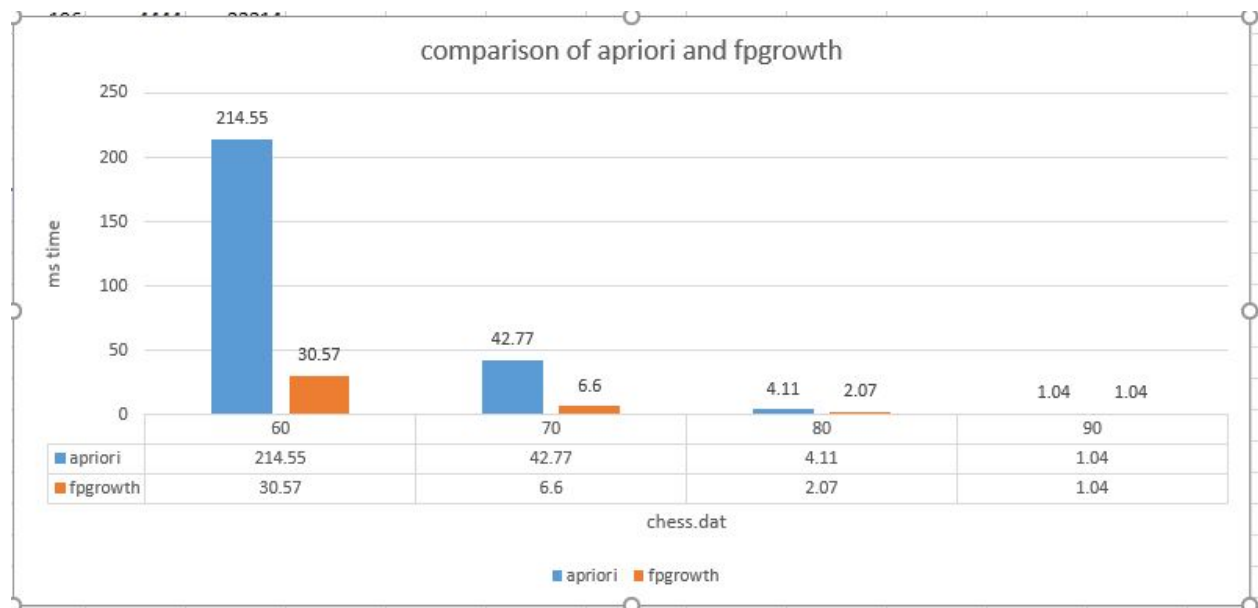
**Submitted to: Chowdhury Farhan Ahmed**

**Professor, University of Dhaka**

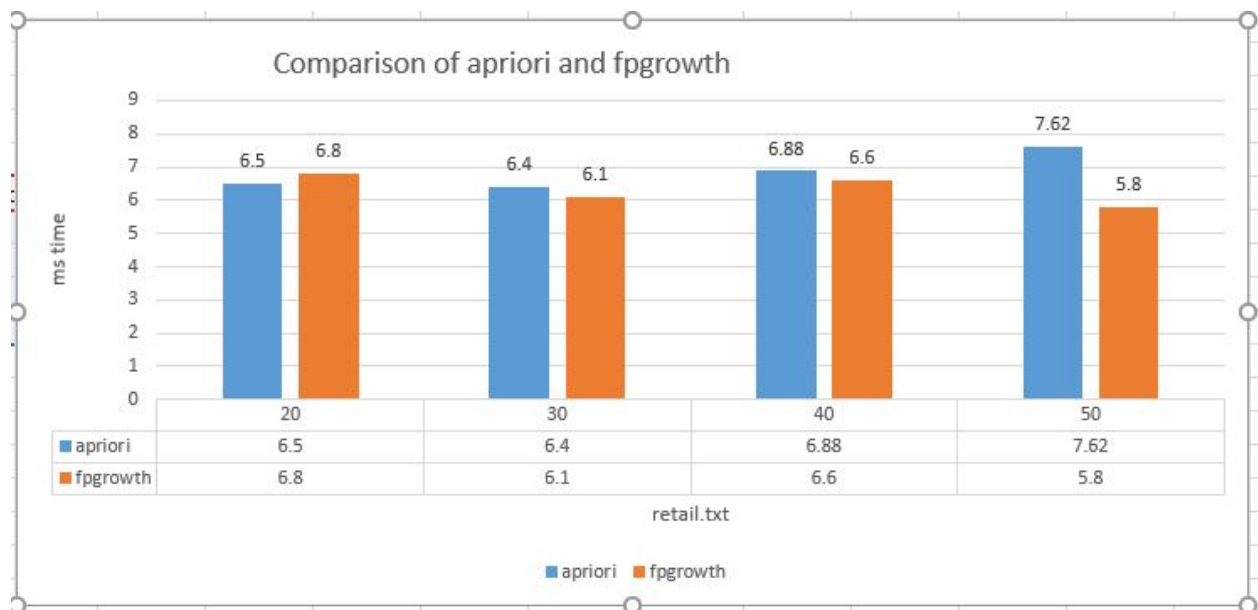
**Department of computer science and engineering**



Here, this statistical graph denotes the comparative analysis of apriori and fpgrowth with respect to execution time efficiency. From the value labeled it is clear that when the minimum support is 60-80 then there is little different in the execution time. This is because the min support is pretty big and the scanning time is less here. On the other hand when the min support is 40-80 then there is significant difference in the execution time. The reason is the algorithm. Fp growth scans the dataset twice on the contrary apriori scan the data n times. That is why for moderately dense mushroom data set the fpriori algorithm takes much less time than the apriori algorithm.



Here, this statistical graph denotes the comparative analysis of apriori and fpgrowth with respect to execution time efficiency for a dense data set chess.dat. From the value labeled it is clear that when the minimum support is 80-90 then there is little different in the execution time. This is because the min support is pretty big and the scanning time is less here. On the other hand when the min support is 60-70 then there is significant difference in the execution time. The reason is the algorithm. Fp growth scans the dataset twice on the contrary apriori scan the data n times. That is why for moderately dense mushroom data set the fpriori algorithm takes much less time than the apriori algorithm. On addition with that for 20-50 percent of the min support the execution time is too high which was not included in the graph.



Here, this statistical graph denotes the comparative analysis of apriori and fpgrowth with respect to execution time efficiency for a sparse data set retail.txt. From the value labeled it is clear that when the minimum support is 20-50 the execution time is somewhat similar. Because the data set min support is big which makes the output like this.

## Conclusion:

Fpgrowth algorithm is much faster than the apriori algorithm. This is because of the algorithm. Fpgrowth scans the data set twice while apriori scans the data n times. But depending on the min support the output may vary.