Ashley Jain
2014CS50280

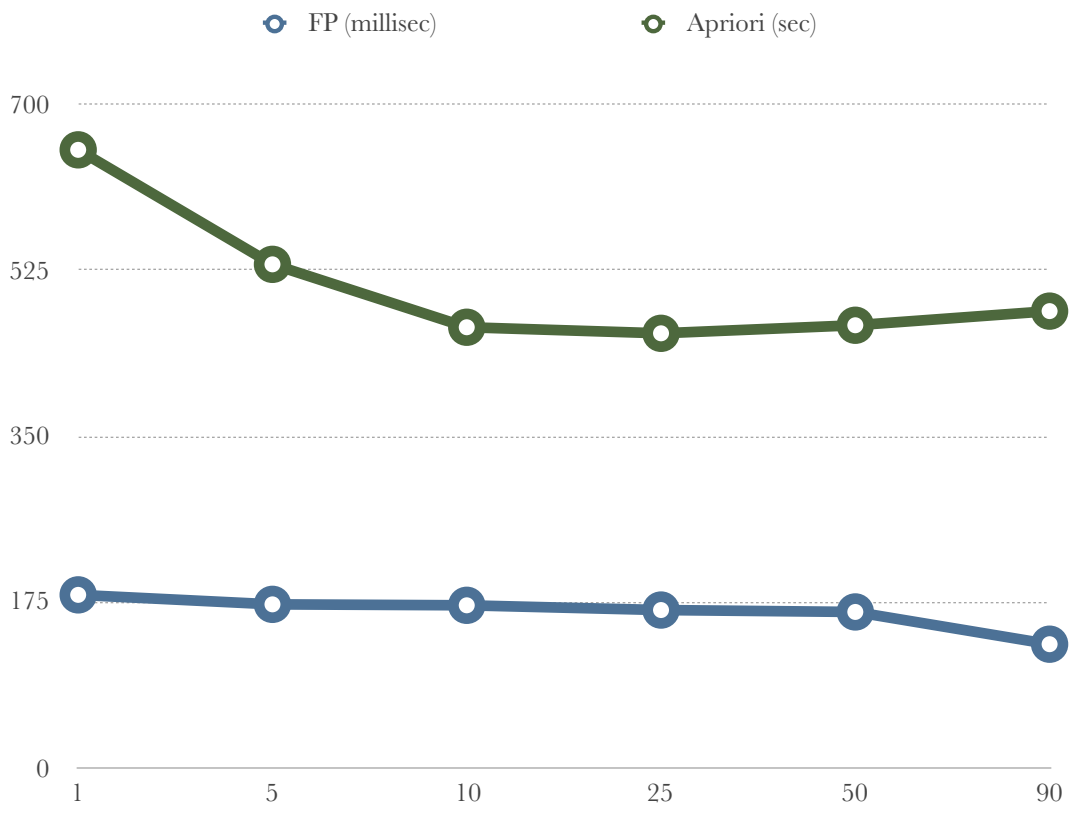
Homework 2

1. b.

(Using FP tree Library:- http://www.borgelt.net/fpgrowth.html)

Graph I :-

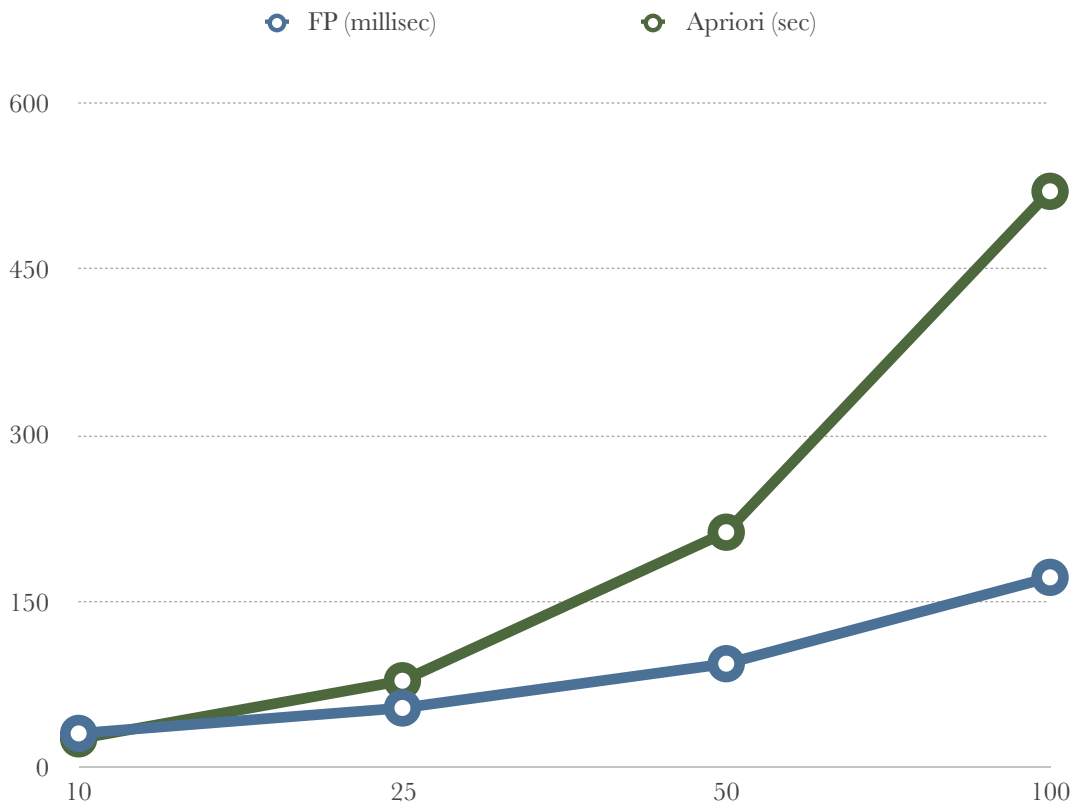
running time as y-axis with support threshold as x-axis



	FP growth (in millisec)	Apriori (in sec)
1	183	652.1
5	173	531.4
10	172	465.1
25	167	458.7
50	165	467.2
90	131	482.1

Graph II :-

running time as y-axis with dataset size as x-axis



Dataset size	FP growth (in millisec)	Apriori (in sec)
10	31	26.1
25	54	78.6
50	94	212.8
100	172	520.8

Explanation:-

Result from first graph :- The time execution of Apriori largely reduce in initial then goes down in small relative further while that of FP growth decrease almost linearly.

Second graph :- In apriori algorithm, the time increase exponentially with increase in dataset size and time increase linearly in FP growth algorithm.

This is due to O(n) complexity in FP growth as it removes the need to calculate pairs to be counted (heavy processing) while using FP tree. Execution time is lesser in FP growth than apriori algorithm(more time executed in generating candidate every time).

Moreover, apriori algo read the transaction file (Database) once for every iteration and FP-Growth algo needs to read transaction file twice only.