

Unit-II

Association

FP-growth

krraocse@gmail.com

Improving the efficiency of Apriori



Association Rule Mining

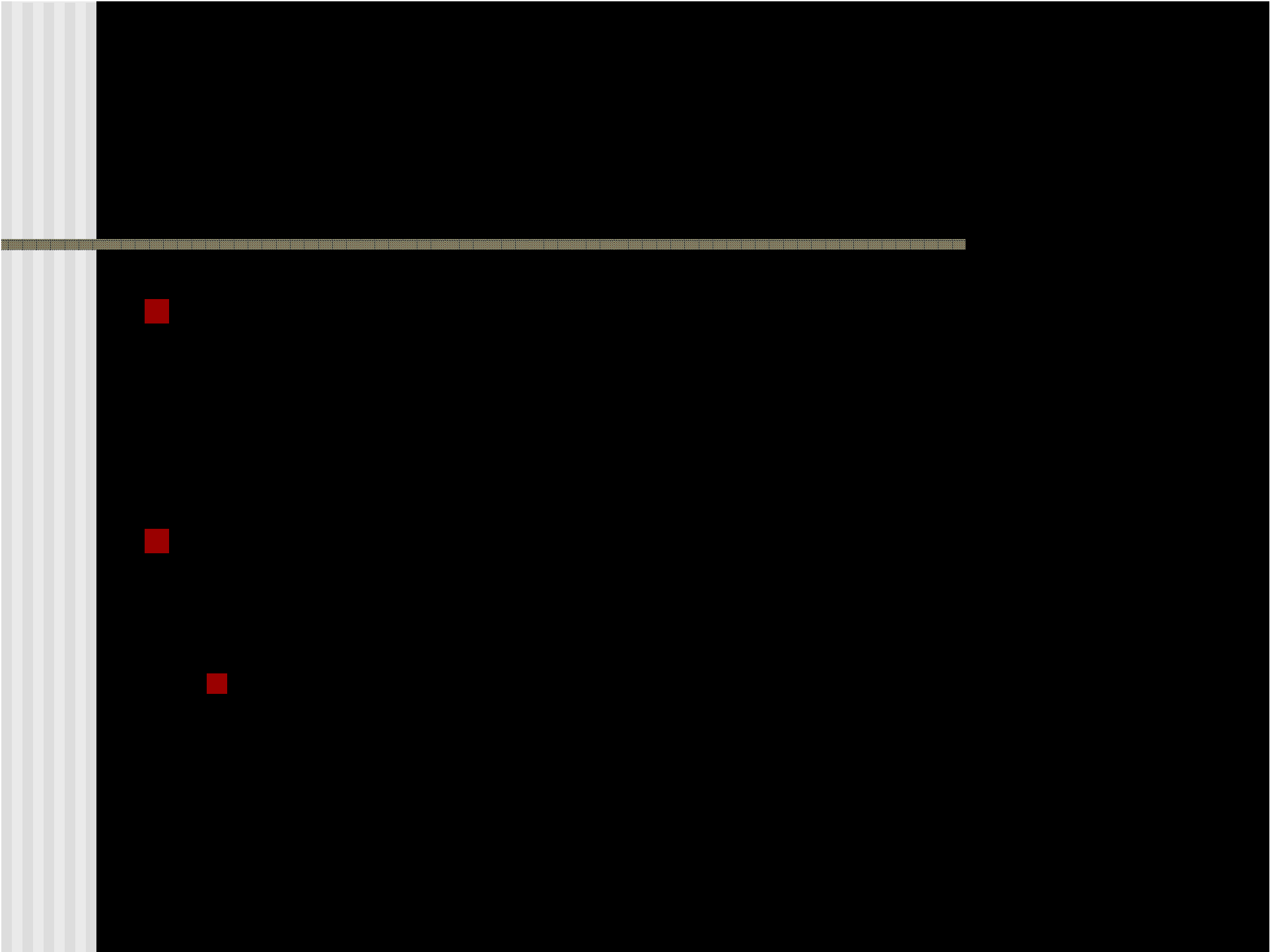


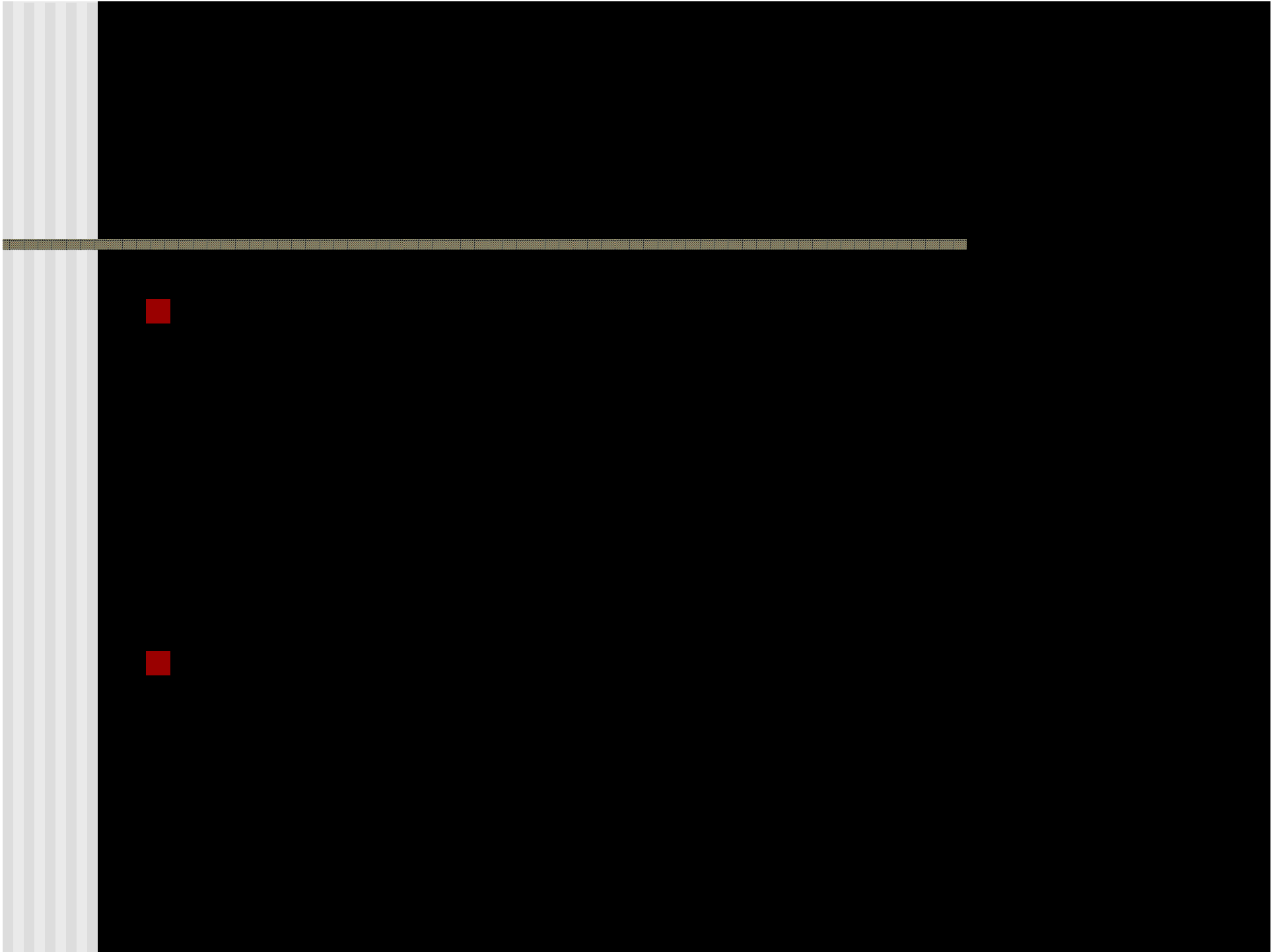
Improving Apriori



Hash-based technique







- A
su
fro

H_2

Create hash table H_2
using hash function
 $h(x, y) = ((order\ of\ x) \times 10$
 $+ (order\ of\ y))\ mod\ 7$



bucket address	0	1	2	3	4	5	6
bucket count	2	2	4	2	2	4	4
bucket contents	{I1, I4} {I3, I5}	{I1, I5}	{I2, I3} {I2, I3} {I2, I3}	{I2, I4}	{I2, I5}	{I1, I2} {I1, I2} {I1, I2}	{I1, I3} {I1, I3} {I1, I3}

Transaction reduction



Partitioning

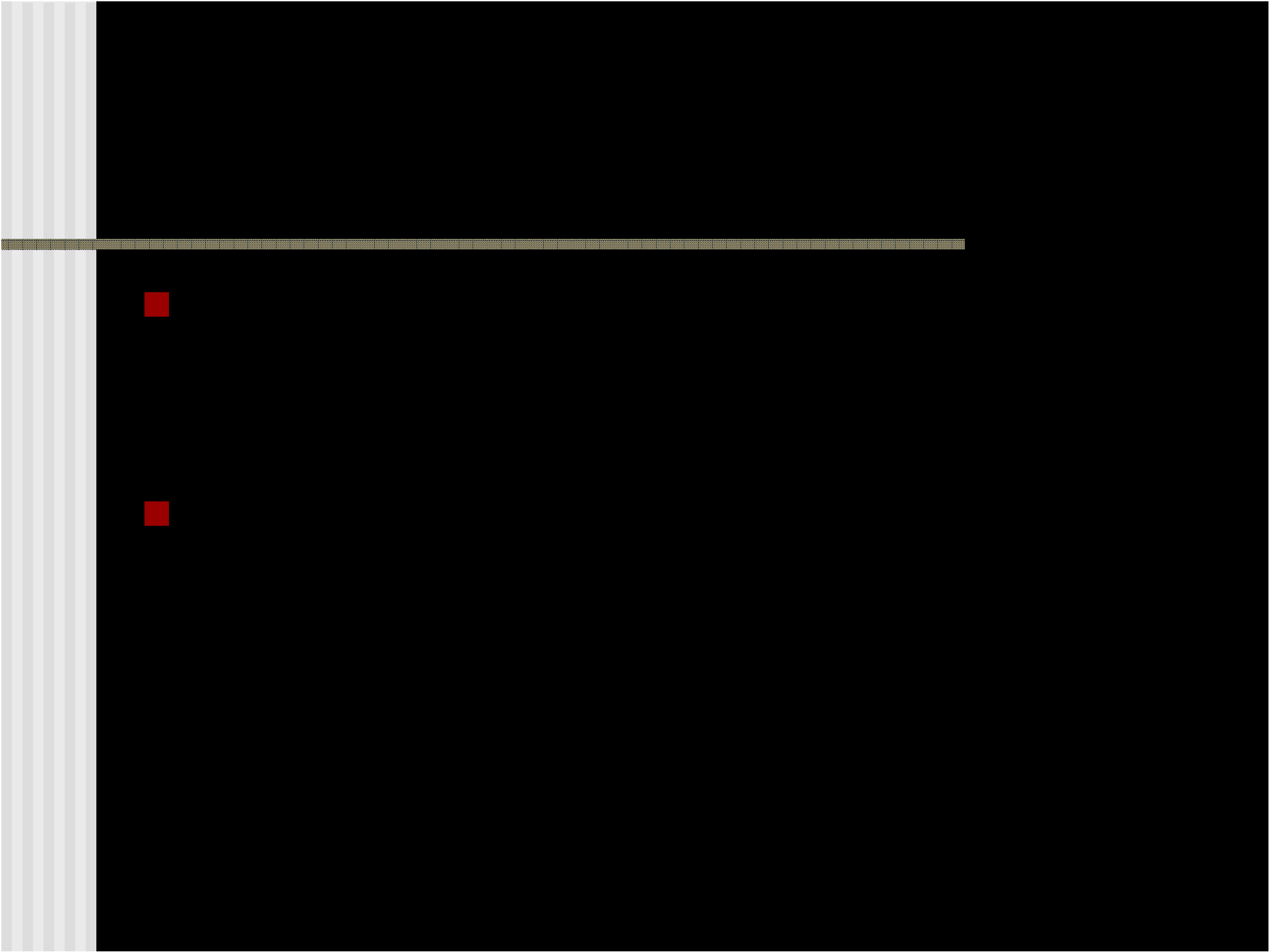


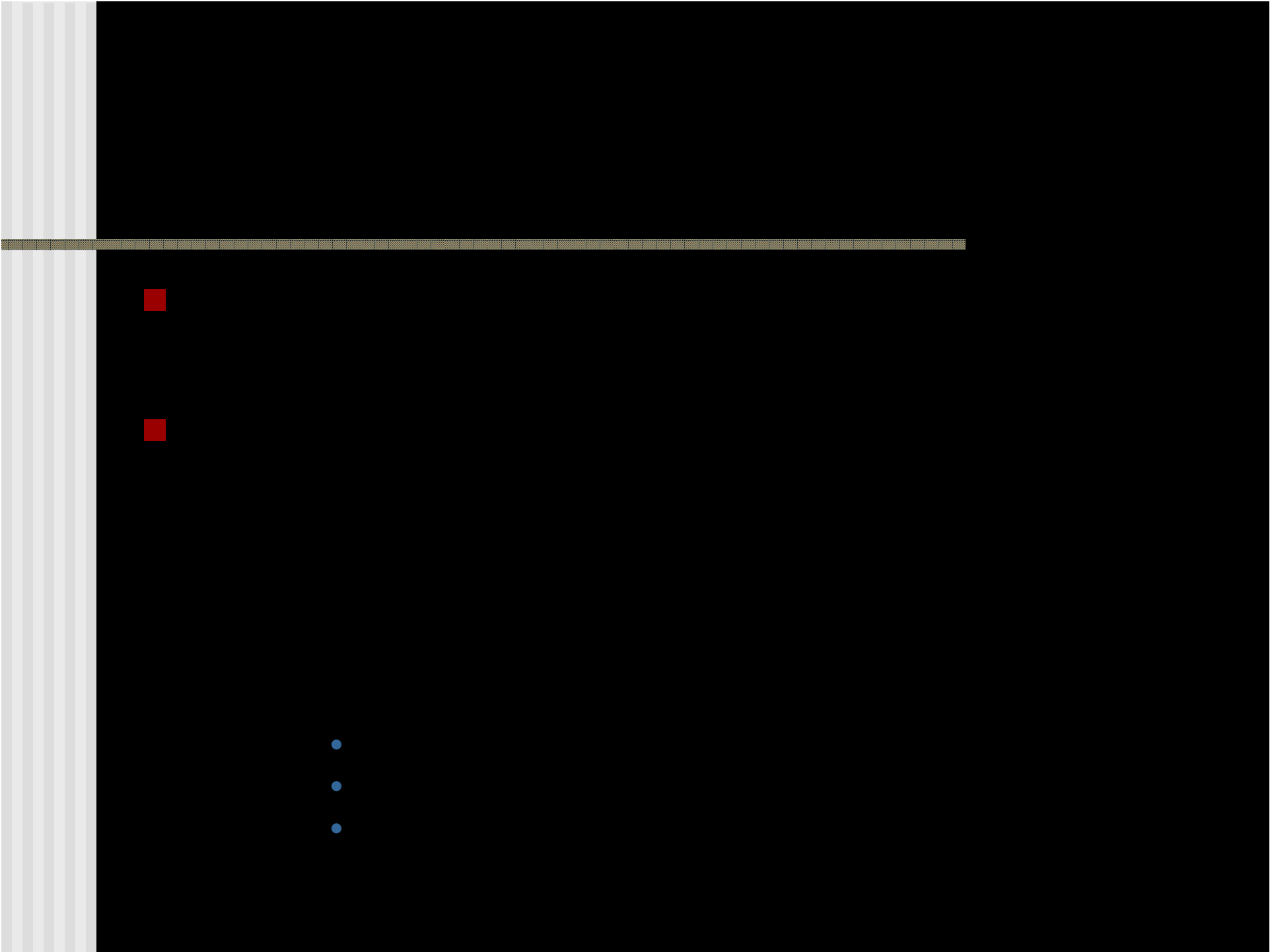
Sampling



Is Apriori fast enough?







Bottleneck of Apriori

