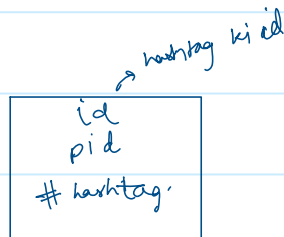
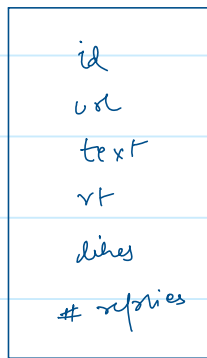


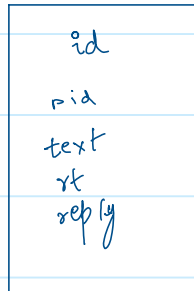
# Project

Saturday, January 26, 2019 2:19 PM

## twitter schema



tid → tweet id  
 hid → hash id  
 cid → comment id  
 pid → parent id  
 rt → retweet



hashtag select

↓

pid ai

↓

text

Hash tag

f: string → n

#divi	1
#go	2

Tweet id

↓ text

1	
2	
3	
4	
5	

hashtag id

tweet id

1	
1	
2	
2	

n is to be memory efficient by being small.