Author: Samuel Rettig

A)

Graphical user interface, Word

Description automatically generated

Query 1:

Num of bytes per record: X

X = 1(bytes in char) \* 80(max char amount for screen\_name) +

1(bytes in char) \* 80(max char amount for user\_name) +

1(bytes in char) \* 80(max char amount for category) +

1(bytes in char) \* 80(max char amount for sub\_category) +

1(bytes in char) \* 80(max char amount for state) +

4 (bytes in int) + 4 (bytes per int) = X

X = 408 bytes/record

Number of records per page = Y

Y = floor(4000/X) = floor (4000 / 408) = 9 (9.8… if not floored)

Number of pages = Z

Z = ceil (5000 / Y) = (5000 / 9) = 556 pages

Since we are not using the header page, the final result is 556

Disk I/O = 556

A picture containing Word

Description automatically generated

X is any index on User that contains screen\_name as the only/first element

Y is any index on Tweets that contains posting\_user as the only/first element

Query 2:

Disk IO Cost = Disk I/O cost of full scan + Disk I/O cost of rows

Full table:

Number of bytes per record is X

X = 8 (bytes (bigint) for tid) + 4 (bytes for post\_day) +

4(bytes for post\_month) + (4 bytes for post\_year) + 4 (bytes for rwtweetCT) +

1 \* 80 ( char posting user)

X = 104 byes/record

Number of records per page = Y

Y = floor(4000 / 104) = 38 records/page

Total number of pages = Z

Z = ceil(10000 / 38) = 264 pages

Disk I/O cost for full scan is 264

Row cost =

Number of rows to write is = X

X = 10000 (amount of tuples) \* 0.1 (selectivity factor) = 1000

Number of bytes per row = Y

Y = 8 bytes/row

Number of rows per page = Z

Z = floor(4000 / 8) = 500

Z = 500 rows/ page

Number of pages to write to: (following x,y,x, ) W

W = ceil(1000 / Z) = (1000 / 500) = 2 pages

W = 2 pages

Disk I/O cost is 2 pages

Final overall cost = 264 + 2 = 266