

BI / read / 1

BI 1

BI 2

BI 3

BI 4

BI 5

BI 6

BI 7

BI 8

BI 9

BI 10

BI 11

BI 12

BI 13

BI 14

BI 15

BI 16

BI 17

BI 18

BI 19

BI 20

query	BI / read / 1				
title	Posting summary				
pattern	<div><div>message: Message</div><div>creationDate < \$datetime</div><div>length year(creationDate)</div></div>				
desc.	<p>Given a datetime, find all Messages created before that moment. Group them by a 3-level grouping:</p> <ol style="list-style-type: none">by year of creationfor each year, group into Message types: is Comment or notfor each year-type group, split into four groups based on length of their content<ul style="list-style-type: none">0: $0 \leq \text{length} < 40$ (short)1: $40 \leq \text{length} < 80$ (one liner)2: $80 \leq \text{length} < 160$ (tweet)3: $160 \leq \text{length}$ (long)				
params	<div>1</div>	datetime	DateTime	For later microbatches, later datetime parameters are selected keep the variance low (<0.5%)	
result	<div>1</div>	year	32-bit Integer	R	year(message.creationDate)
	<div>2</div>	isComment	Boolean	M	True for Comments, False for Posts
	<div>3</div>	lengthCategory	32-bit Integer	C	0 for short, 1 for one-liner, 2 for tweet, 3 for long
	<div>4</div>	messageCount	32-bit Integer	A	Total number of Messages in that group
	<div>5</div>	averageMessageLength	32-bit Float	A	Average length of the Message content in that group
	<div>6</div>	sumMessageLength	32-bit Integer	A	Sum of all Message content lengths
	<div>7</div>	percentageOfMessages	32-bit Float	A	Number of Messages in group as a percentage of all messages created before the given date
sort	<div>1</div>	year	↓		
	<div>2</div>	isComment	↑	False < True, i.e. Posts come first and Comments second	
	<div>3</div>	lengthCategory	↑	order based on the lengthCategory value	
limit	n/a				
CPs	1.2, 3.2, 4.1, 4.2, 8.5				