

## BI / read / 1

|                              |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
|------------------------------|-------------------------|--|----------|---|--|------------------|---------------------------|---|------------------------------------|----------------------------|---|-----------|---------|---|------------------------------------|---|----------------|----------------|---|---|---|--------------|----------------|---|--|---|----------------------|--------------|---|---|---|------------------|----------------|---|------------------------------------|---|----------------------|--------------|---|---|
| BI 1                         | query                   | BI / read / 1  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 2                         | title                   | Posting summary  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 3                         | pattern                 | <table><tr><td>message: Message</td></tr><tr><td>creationDate &lt; \$datetime</td></tr><tr><td>length<br/>year(creationDate)</td></tr></table>   |          |   |  | message: Message | creationDate < \$datetime | length<br>year(creationDate)                          |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| message: Message             |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| creationDate < \$datetime    |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| length<br>year(creationDate) |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 4                         |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 5                         |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 6                         | desc.                   | <p>Given a datetime, find all Messages created before that moment. Group them by a 3-level grouping:</p> <p>1. by year of creation</p> <p>2. for each year, group into Message types: is Comment or not</p> <p>3. for each year-type group, split into four groups based on length of their content</p> <ul style="list-style-type: none"><li>• 0: <math>0 \leq \text{length} &lt; 40</math> (short)</li><li>• 1: <math>40 \leq \text{length} &lt; 80</math> (one liner)</li><li>• 2: <math>80 \leq \text{length} &lt; 160</math> (tweet)</li><li>• 3: <math>160 \leq \text{length}</math> (long)</li></ul>  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 7                         |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 8                         |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 9                         |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 10                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 11                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 12                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 13                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 14                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 15                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 16                        | params                  | 1  | datetime | DateTime  | For later microbatches, later datetime parameters are selected keep the variance low (<0.5%) |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 17                        |                         | <table><tr><td>1</td><td>year</td><td>32-bit Integer</td><td>R</td><td>year(message.creationDate)</td></tr><tr><td>2</td><td>isComment</td><td>Boolean</td><td>M</td><td>True for Comments, False for Posts</td></tr><tr><td>3</td><td>lengthCategory</td><td>32-bit Integer</td><td>C</td><td>0 for short, 1 for one-liner, 2 for tweet, 3 for long</td></tr><tr><td>4</td><td>messageCount</td><td>32-bit Integer</td><td>A</td><td>Total number of Messages in that group</td></tr><tr><td>5</td><td>averageMessageLength</td><td>32-bit Float</td><td>A</td><td>Average length of the Message content in that group</td></tr><tr><td>6</td><td>sumMessageLength</td><td>32-bit Integer</td><td>A</td><td>Sum of all Message content lengths</td></tr><tr><td>7</td><td>percentageOfMessages</td><td>32-bit Float</td><td>A</td><td>Number of Messages in group as a percentage of all messages created before the given date</td></tr></table> |          |   |  | 1                | year                      | 32-bit Integer  | R                                  | year(message.creationDate) | 2 | isComment | Boolean | M | True for Comments, False for Posts | 3 | lengthCategory | 32-bit Integer | C | 0 for short, 1 for one-liner, 2 for tweet, 3 for long | 4 | messageCount | 32-bit Integer | A | Total number of Messages in that group | 5 | averageMessageLength | 32-bit Float | A | Average length of the Message content in that group | 6 | sumMessageLength | 32-bit Integer | A | Sum of all Message content lengths | 7 | percentageOfMessages | 32-bit Float | A | Number of Messages in group as a percentage of all messages created before the given date |
| 1                            |                         |  |          |   |  | year             | 32-bit Integer            | R   | year(message.creationDate)         |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| 2                            |                         |  |          |   |  | isComment        | Boolean                   | M   | True for Comments, False for Posts |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| 3                            | lengthCategory          |  |          |   |  | 32-bit Integer   | C                         | 0 for short, 1 for one-liner, 2 for tweet, 3 for long |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| 4                            | messageCount            |  |          |   |  | 32-bit Integer   | A                         | Total number of Messages in that group                |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| 5                            | averageMessageLength    |  |          |   |  | 32-bit Float     | A                         | Average length of the Message content in that group   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| 6                            | sumMessageLength        |  |          |   |  | 32-bit Integer   | A                         | Sum of all Message content lengths                    |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| 7                            | percentageOfMessages    | 32-bit Float   | A        | Number of Messages in group as a percentage of all messages created before the given date |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 18                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 19                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| BI 20                        |                         |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| sort                         | 1                       | year   | ↓        |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
|                              | 2                       | isComment  | ↑        | False < True, i.e. Posts come first and Comments second                                   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
|                              | 3                       | lengthCategory   | ↑        | order based on the lengthCategory value   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| limit                        | n/a                     |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |
| CPs                          | 1.2, 3.2, 4.1, 4.2, 8.5 |  |          |   |  |                  |                           |   |                                    |                            |   |           |         |   |                                    |   |                |                |   |   |   |              |                |   |  |   |                      |              |   |   |   |                  |                |   |                                    |   |                      |              |   |   |

## BI / read / 2

|         |   |                |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
|---------|---|----------------|---|---|---|---|----------|--------------|---|---|--|---|--------------|----------------|---|---|---|------|----------------|---|--|--|--|--|
| query   | BI / read / 2   |                |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| title   | Tag evolution   |                |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| pattern | <pre>graph TD     TagClass -- hasType --&gt; tag     tag -- "«opt» hasTag" --&gt; M1     tag -- "«opt» hasTag" --&gt; M2     subgraph Window1 [countWindow1 = count(message)]         M1     end     subgraph Window2 [countWindow2 = count(message)]         M2     end     class TagClass fill:#800080,color:#fff     class tag fill:#ff69b4,color:#000     class M1 fill:#9370db,color:#000     class M2 fill:#9370db,color:#000</pre>   |                |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| desc.   | Find the Tags under a given TagClass that were used in Messages during in the 100-day period starting at date and compare it with the 100-day period that follows. For the Tags and for both months, compute the count of Messages.   |                |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| params  | <table><tr><td>1</td><td>date</td><td>Date</td><td></td></tr><tr><td>2</td><td>tagClass</td><td>Long String</td><td>TagClasses with a similar amount of Messages are selected</td></tr></table>   | 1              | date  | Date  |   | 2 | tagClass | Long String  | TagClasses with a similar amount of Messages are selected |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| 1       | date  | Date           |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| 2       | tagClass  | Long String    | TagClasses with a similar amount of Messages are selected |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| result  | <table><tr><td>1</td><td>tag.name</td><td>Long String</td><td>R</td><td></td></tr><tr><td>2</td><td>countWindow1</td><td>32-bit Integer</td><td>A</td><td>Occurrences of the tagClass during the first time window</td></tr><tr><td>3</td><td>countWindow2</td><td>32-bit Integer</td><td>A</td><td>Occurrences of the tagClass during the second time window</td></tr><tr><td>4</td><td>diff</td><td>32-bit Integer</td><td>A</td><td>Absolute difference of countWindow1 and countWindow2</td></tr></table> | 1              | tag.name  | Long String   | R |   | 2        | countWindow1 | 32-bit Integer  | A | Occurrences of the tagClass during the first time window | 3 | countWindow2 | 32-bit Integer | A | Occurrences of the tagClass during the second time window | 4 | diff | 32-bit Integer | A | Absolute difference of countWindow1 and countWindow2 |  |  |  |
| 1       | tag.name  | Long String    | R   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| 2       | countWindow1  | 32-bit Integer | A   | Occurrences of the tagClass during the first time window  |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| 3       | countWindow2  | 32-bit Integer | A   | Occurrences of the tagClass during the second time window |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| 4       | diff  | 32-bit Integer | A   | Absolute difference of countWindow1 and countWindow2      |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| sort    | <table><tr><td>1</td><td>diff</td><td>↓</td><td></td></tr><tr><td>2</td><td>tag.name</td><td>↑</td><td></td></tr></table>   | 1              | diff  | ↓   |   | 2 | tag.name | ↑            |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| 1       | diff  | ↓              |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| 2       | tag.name  | ↑              |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| limit   | 100   |                |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |
| CPs     | 2.4, 3.1, 3.2, 4.1, 4.2, 4.3, 5.3, 6.1, 8.2, 8.5  |                |   |   |   |   |          |              |   |   |  |   |              |                |   |   |   |      |                |   |  |  |  |  |

## BI / read / 3

|         |  |                |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
|---------|--|----------------|---|-------------|---|---|----------|-------------|----------------------------|---|--|---|--------------------|----------|---|--|---|-----------|----|---|--|---|--------------|----------------|---|--|--|--|--|--|
| query   | BI / read / 3  |                |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| title   | Popular topics in a country  |                |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| pattern | <pre>graph TD     Country[Country<br/>name = \$country]     City[City]     Person[person: Person<br/>id]     Forum[forum: Forum<br/>id<br/>title<br/>creationDate]     TagClass[TagClass<br/>name = \$tagClass]     Tag[Tag]     Message[message: Message<br/>count(message)]     Post[Post]      Country -- isPartOf --&gt; City     City -- isLocatedIn --&gt; Person     Person -- hasModerator --&gt; Forum     Forum -- containerOf --&gt; Post     TagClass -- hasType --&gt; Tag     Tag -- hasTag --&gt; Message     Message -- count(message) --&gt; Tag     Message -- replyOf*0.. --&gt; Post</pre> |                |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| desc.   | <p>Given a TagClass and a Country, find all the Forums created in the given Country, containing at least one Message with Tags belonging directly to the given TagClass, and count the Messages by the Person who created it and by the Forum which contains them.</p> <p>The location of a Forum is identified by the location of the Forum’s moderator.</p>  |                |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| params  | <table><tr><td>1</td><td>tagClass</td><td>Long String</td><td>TagClasses with a similar amount of Messages are selected</td></tr><tr><td>2</td><td>country</td><td>Long String</td><td>Big Countries are selected</td></tr></table>  | 1              | tagClass  | Long String | TagClasses with a similar amount of Messages are selected | 2 | country  | Long String | Big Countries are selected |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| 1       | tagClass   | Long String    | TagClasses with a similar amount of Messages are selected |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| 2       | country  | Long String    | Big Countries are selected                                |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| result  | <table><tr><td>1</td><td>forum.id</td><td>ID</td><td>R</td><td></td></tr><tr><td>2</td><td>forum.title</td><td>Long String</td><td>R</td><td></td></tr><tr><td>3</td><td>forum.creationDate</td><td>DateTime</td><td>R</td><td></td></tr><tr><td>4</td><td>person.id</td><td>ID</td><td>R</td><td></td></tr><tr><td>5</td><td>messageCount</td><td>32-bit Integer</td><td>A</td><td></td></tr></table>   | 1              | forum.id  | ID          | R   |   | 2        | forum.title | Long String                | R |  | 3 | forum.creationDate | DateTime | R |  | 4 | person.id | ID | R |  | 5 | messageCount | 32-bit Integer | A |  |  |  |  |  |
| 1       | forum.id   | ID             | R   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| 2       | forum.title  | Long String    | R   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| 3       | forum.creationDate   | DateTime       | R   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| 4       | person.id  | ID             | R   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| 5       | messageCount   | 32-bit Integer | A   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| sort    | <table><tr><td>1</td><td>messageCount</td><td>↓</td><td></td></tr><tr><td>2</td><td>forum.id</td><td>↑</td><td></td></tr></table>  | 1              | messageCount  | ↓           |   | 2 | forum.id | ↑           |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| 1       | messageCount   | ↓              |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| 2       | forum.id   | ↑              |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| limit   | 20   |                |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |
| CPs     | 1.1, 1.2, 1.3, 2.1, 2.2, 2.4, 3.3, 8.2   |                |   |             |   |   |          |             |                            |   |  |   |                    |          |   |  |   |           |    |   |  |   |              |                |   |  |  |  |  |  |

## BI / read / 4

|         |   |                     |                |  |  |
|---------|---|---------------------|----------------|--|--|
| query   | BI / read / 4   |                     |                |  |  |
| title   | Top message creators by country   |                     |                |  |  |
| pattern | <div><div><div>1. select top 100 forums based on memberCount in country</div><div><div>Country</div><div>name</div></div><div>isPartOf</div><div><div>City</div><div>isLocatedIn</div><div><div>member: Person</div><div>memberCount = count(member)</div></div><div>hasMember</div><div><div>forum: Forum</div><div>creationDate &gt; \$date</div></div></div></div><div><div>2. for each country, for each of the top 100 forums (topForum), find the top posters</div><div><div><div>topForum: Forum</div><div>containerOf</div><div><div>Post</div><div>replyOf*0..</div><div><div>Message</div><div>messageCount = count(message)</div><div>creationDate &gt; \$date</div></div><div>hasCreator</div><div><div>person: Person</div><div>id</div><div>firstName</div><div>lastName</div><div>creationDate</div></div></div></div><div><div>Forum</div><div>is in top 100 forum</div><div>hasMember</div><div><div>person: Person</div><div>id</div><div>firstName</div><div>lastName</div><div>creationDate</div></div></div></div></div></div> |                     |                |  |  |
| desc.   | <p>Find the most popular Forums by Country, where the popularity of a Forum is measured by the number of members that Forum has from a given Country.</p> <p>Calculate the top 100 most popular Forums. If a Forum is popular in multiple countries, it should only be calculated once with its largest membership. In case of a tie, the Forum(s) with the smaller id value(s) should be selected.</p> <p>For each member Person of the 100 most popular Forums, count the number of Messages (messageCount) they made in any of those (most popular) Forums. Also include those member Persons who have not posted any Messages (have a messageCount of 0).</p>   |                     |                |  |  |
| params  | 1   | date                | Date           | Selected from the first 30 days of the network |  |
| result  | 1   | person.id           | ID             | R  |  |
|         | 2   | person.firstName    | String         | R  |  |
|         | 3   | person.lastName     | String         | R  |  |
|         | 4   | person.creationDate | DateTime       | R  |  |
|         | 5   | messageCount        | 32-bit Integer | A  |  |
| sort    | 1   | messageCount        | ↓              |  |  |
|         | 2   | person.id           | ↑              |  |  |
| limit   | 100   |                     |                |  |  |
| CPs     | 1.2, 1.3, 2.1, 2.2, 2.3, 2.4, 3.3, 5.3, 6.1, 8.2, 8.4   |                     |                |  |  |

## BI / read / 5

|         |   |              |                |   |
|---------|---|--------------|----------------|---|
| query   | BI / read / 5   |              |                |   |
| title   | Most active posters of a given topic  |              |                |   |
| pattern |   |              |                |   |
| desc.   | <p>Get each Person (person) who has created a Message (message) with a given Tag (direct relation, not transitive). Considering only these Messages, for each Person node:</p> <ul style="list-style-type: none"> <li>Count its messages (messageCount).</li> <li>Count likes (likeCount) to its messages.</li> <li>Count Comments (replyCount) in reply to it messages.</li> </ul> <p>The score is calculated according to the following formula: <math>1 \times \text{messageCount} + 2 \times \text{replyCount} + 10 \times \text{likeCount}</math>.</p> |              |                |   |
| params  | 1   | tag          | Long String    | Tags with a similar amount of Messages are selected |
| result  | 1   | person.id    | ID             | R   |
|         | 2   | replyCount   | 32-bit Integer | A   |
|         | 3   | likeCount    | 32-bit Integer | A   |
|         | 4   | messageCount | 32-bit Integer | A   |
|         | 5   | score        | 32-bit Integer | A   |
| sort    | 1   | score        | ↓              |   |
|         | 2   | person.id    | ↑              |   |
| limit   | 100   |              |                |   |
| CPs     | 1.2, 2.3, 8.2   |              |                |   |

## BI / read / 6

|       |           |   |                |                |   |
|-------|-----------|---|----------------|----------------|---|
| BI 1  | query     | BI / read / 6   |                |                |   |
| BI 2  | title     | Most authoritative users on a given topic   |                |                |   |
| BI 3  | pattern   |   |                |                |   |
| 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          |   |
|       | desc.     | <p>Given a Tag (tag), find all Persons (person) that ever created a Message with the Tag. For each of these Persons (person) compute their “authority score” as follows:</p> <ul style="list-style-type: none"> <li>The “authority score” is the sum of “popularity scores” of the Persons (p2) that liked any of that Person’s Messages with the given Tag (same criterion as for message1).</li> <li>A Person’s (p2) “popularity score” is defined as the total number of likes on all of their Messages (message2).</li> </ul> |                |                |   |
|       | params    | 1   | tag            | Long String    | Tags with a similar amount of Messages are selected |
|       | result    | 1   | person.id      | ID             | R   |
|       |           | 2   | authorityScore | 32-bit Integer | A   |
|       | sort      | 1   | authorityScore | ↓              |   |
|       |           | 2   | person1.id     | ↑              |   |
|       | limit     | 100   |                |                |   |
|       | CPs       | 1.2, 2.3, 3.3, 6.1, 8.2   |                |                |   |
|       | relevance | Computing the authority scores might involve computing the popularity score for the same Person multiple times. Implementations are advised to avoid such redundant computations.   |                |                |   |

## BI / read / 7

|       |         |  |                 |                |   |
|-------|---------|--|-----------------|----------------|---|
| BI 1  | query   | BI / read / 7  |                 |                |   |
| BI 2  | title   | Related topics   |                 |                |   |
| BI 3  | pattern |  |                 |                |   |
| BI 8  | desc.   | Find all Messages that have a given Tag. Find the related Tags attached to (direct) reply Comments of these Messages, but only of those reply Comments that do not have the given Tag. Group the Tags by name, and get the count of replies in each group. |                 |                |   |
| BI 12 | params  | 1  | tag             | Long String    | Tags with a similar amount of Messages are selected |
| BI 14 | result  | 1  | relatedTag.name | Long String    | R   |
| BI 15 |         | 2  | count           | 32-bit Integer | A   |
| BI 17 | sort    | 1  | count           | ↓              |   |
| BI 18 |         | 2  | relatedTag.name | ↑              |   |
| BI 19 | limit   | 100  |                 |                |   |
| BI 20 | CPs     | 1.4, 3.3, 5.2, 8.1   |                 |                |   |

## BI / read / 8

|           |   |                                      |  |   |  |
|-----------|---|--------------------------------------|--|---|--|
| query     | BI / read / 8   |                                      |  |   |  |
| title     | Central person for a tag  |                                      |  |   |  |
| pattern   | <div><p>For each person with a matching hasInterest and/or hasCreator edge, compute person.score = ((if hasInterest edge exists then 100 else 0) + count(message))</p><p>Calculate the sum of the friends' scores: friendsScore = sum(friend.score)</p><p>Detailed description: The diagram illustrates a graph query pattern. It starts with a <b>Tag</b> node (pink box) with a filter <code>name = \$tag</code>. A <b>person: Person</b> node (orange box) with filter <code>id</code> is connected to the Tag node via a dashed <code>hasInterest</code> edge (labeled <code>«opt»</code>). The person node is also connected to a <b>message: Message</b> node (purple box) via a dashed <code>hasCreator</code> edge (labeled <code>«opt»</code>). The message node has a filter <code>\$date &lt; creationDate</code> and is connected to the Tag node via a solid <code>hasTag</code> edge. A <code>count</code> aggregation is shown over the message node. Below this, a separate box shows the calculation of <code>friendsScore</code> as the sum of scores of friends. It features a <b>person: Person</b> node connected to a <b>friend: Person</b> node via a dashed <code>knows</code> edge (labeled <code>«opt»</code>).</p></div> |                                      |  |   |  |
| desc.     | <p>Given a Tag, find all Persons that are interested in the Tag and/or have written a Message (Post or Comment) with a creationDate after a given date and that has a given Tag. For each Person, compute the score as the sum of the following two aspects:</p> <ul style="list-style-type: none"><li>• 100, if the Person has this Tag as their interest, or 0 otherwise</li><li>• number of Messages by this Person with the given Tag</li></ul> <p>Also, for each Person, compute the sum of the score of the Person’s friends (friendsScore).</p>  |                                      |  |   |  |
| params    | <div><div>1</div><div>tag</div></div>   | <div><div>Long String</div></div>    | <div>Tags with a similar amount of Messages are selected</div>                                   |   |  |
|           | <div><div>2</div><div>date</div></div>  | <div><div>Date</div></div>           | <div>Dates from around the same day are selected. (TODO - how exactly? what distribution?)</div> |   |  |
| result    | <div><div>1</div><div>person.id</div></div>   | <div><div>ID</div></div>             | <div><div>R</div></div>  |   |  |
|           | <div><div>2</div><div>score</div></div>   | <div><div>32-bit Integer</div></div> | <div><div>A</div></div>  |   |  |
|           | <div><div>3</div><div>friendsScore</div></div>  | <div><div>32-bit Integer</div></div> | <div><div>A</div></div>  | <div>The sum of the score of the person’s friends</div> |  |
| sort      | <div><div>1</div><div>score + friendsScore</div></div>  | <div><div>↓</div></div>              |  |   |  |
|           | <div><div>2</div><div>person.id</div></div>   | <div><div>↑</div></div>              |  |   |  |
| limit     | 100   |                                      |  |   |  |
| CPs       | 1.2, 2.1, 2.3, 3.2, 5.3, 8.2, 8.4, 8.5  |                                      |  |   |  |
| relevance | <p>Similarly to BI 16, there are two major ways to compute this query: (1) creating an induced subgraph of the interested Persons and their friends and performing the scoring on this graph or (2) performing the scoring without creating an induced subgraph and scoring the friends of a Person on-the-fly. The first approach is more efficient as it avoids redundant computations, however, specifying it needs support for composable graph queries.</p>  |                                      |  |   |  |



## BI / read / 9

|        |         |                       |        |        |   |
|--------|---------|-----------------------|--------|--------|---|
| BI 1   | query   | BI / read / 9         |        |        |   |
| BI 2   | title   | Top thread initiators |        |        |   |
| BI 3   | pattern |                       |        |        |   |
| BI 4   | BI 5    | BI 6                  | BI 7   | BI 8   | <p>For each Person, count the number of Posts they created in the time interval [startDate, endDate] (equivalent to the number of threads they initiated) and the number of Messages in each of their (transitive) reply trees, including the root Post of each tree. When calculating Message counts only consider Messages created within the given time interval.</p> <p>Return each Person, number of Posts they created, and the count of all Messages that appeared in the reply trees (including the Post at the root of tree) they created.</p> |
| BI 9   | BI 10   | BI 11                 | BI 12  | BI 13  |   |
| BI 14  | BI 15   | BI 16                 | BI 17  | BI 18  |   |
| BI 19  | BI 20   | BI 21                 | BI 22  | BI 23  |   |
| BI 24  | BI 25   | BI 26                 | BI 27  | BI 28  |   |
| BI 29  | BI 30   | BI 31                 | BI 32  | BI 33  | <p>1 startDate Date TODO</p> <p>2 endDate Date 8-10 days after the startDate</p>  |
| BI 34  | BI 35   | BI 36                 | BI 37  | BI 38  |   |
| BI 39  | BI 40   | BI 41                 | BI 42  | BI 43  |   |
| BI 44  | BI 45   | BI 46                 | BI 47  | BI 48  |   |
| BI 49  | BI 50   | BI 51                 | BI 52  | BI 53  |   |
| BI 54  | BI 55   | BI 56                 | BI 57  | BI 58  | <p>1 person.id ID R</p> <p>2 person.firstName String R</p> <p>3 person.lastName String R</p> <p>4 threadCount 32-bit Integer A The number of Posts created by that Person (the number of threads initiated)</p> <p>5 messageCount 32-bit Integer A The number of Messages created in all the threads this Person initiated</p>  |
| BI 59  | BI 60   | BI 61                 | BI 62  | BI 63  |   |
| BI 64  | BI 65   | BI 66                 | BI 67  | BI 68  |   |
| BI 69  | BI 70   | BI 71                 | BI 72  | BI 73  |   |
| BI 74  | BI 75   | BI 76                 | BI 77  | BI 78  |   |
| BI 79  | BI 80   | BI 81                 | BI 82  | BI 83  | <p>1 messageCount ↓</p> <p>2 person.id ↑</p>  |
| BI 84  | BI 85   | BI 86                 | BI 87  | BI 88  |   |
| BI 89  | BI 90   | BI 91                 | BI 92  | BI 93  |   |
| BI 94  | BI 95   | BI 96                 | BI 97  | BI 98  |   |
| BI 99  | BI 100  | BI 101                | BI 102 | BI 103 |   |
| BI 104 | BI 105  | BI 106                | BI 107 | BI 108 | limit   |
| BI 109 | BI 110  | BI 111                | BI 112 | BI 113 | 100   |
| BI 114 | BI 115  | BI 116                | BI 117 | BI 118 | CPs   |
| BI 119 | BI 120  | BI 121                | BI 122 | BI 123 | 1.2, 2.2, 2.3, 3.2, 7.2, 7.3, 7.4, 8.1, 8.5   |

## BI / read / 10

|         |   |                          |                |  |   |
|---------|---|--------------------------|----------------|--|---|
| query   | BI / read / 10  |                          |                |  |   |
| title   | Experts in social circle  |                          |                |  |   |
| pattern |   |                          |                |  |   |
| desc.   | <p>Given a Person (startPerson), find all other Persons (expertCandidatePerson) that live in a given Country and are connected to given Person by a <i>shortest path</i> with length in range [minPathDistance, maxPathDistance] through the knows relation.</p> <p>For each of these expertCandidatePerson nodes, retrieve all of their Messages that contain at least one Tag belonging to a given TagClass (direct relation not transitive). For each Message, retrieve all of its Tags.</p> <p>Group the results by Persons and Tags, then count the Messages by a certain Person having a certain Tag.</p> |                          |                |  |   |
| params  | 1   | personId                 | ID             | The ID of the startPerson. Persons with a similar degree of knows edges are selected |   |
|         | 2   | country                  | String         | Countries with a similar number of Persons are selected                              |   |
|         | 3   | tagClass                 | Long String    | TagClasses with a similar degree of hasType edges are selected                       |   |
|         | 4   | minPathDistance          | 32-bit Integer | 1 or 2   |   |
|         | 5   | maxPathDistance          | 32-bit Integer | 2 or 3   |   |
| result  | 1   | expertCandidatePerson.id | ID             | R  |   |
|         | 2   | tag.name                 | Long String    | R  |   |
|         | 3   | messageCount             | 32-bit Integer | A  | Number of Messages created by that Person containing that Tag |
| sort    | 1   | messageCount             | ↓              |  |   |
|         | 2   | tag.name                 | ↑              |  |   |
|         | 3   | expertCandidatePerson.id | ↑              |  |   |
| limit   | 100   |                          |                |  |   |
| CPs     | 1.2, 1.3, 2.3, 2.4, 3.3, 5.3, 7.1, 7.2, 7.3, 8.1, 8.6   |                          |                |  |   |

## BI / read / 11

|         |  |  |  |  |
|---------|--|--|--|--|
| query   | BI / read / 11   |  |  |  |
| title   | Friend triangles   |  |  |  |
| pattern |  |  |  |  |
| desc.   | <p>For a given country, count all the distinct triples of Persons such that:</p> <ul style="list-style-type: none"><li>• a is friend of b,</li><li>• b is friend of c,</li><li>• c is friend of a,</li></ul> <p>and these friendships were created after a given startDate.</p> <p>Distinct means that given a triple <math>t_1</math> in the result set <math>R</math> of all qualified triples, there is no triple <math>t_2</math> in <math>R</math> such that <math>t_1</math> and <math>t_2</math> have the same set of elements.</p> |  |  |  |
| params  | <div><div>1</div><div>country</div><div>Long String</div><div></div></div> <div><div>2</div><div>startDate</div><div>Date</div><div></div></div>   |  |  |  |
| result  | <div><div>1</div><div>count</div><div>32-bit Integer</div><div>A</div><div></div></div>  |  |  |  |
| limit   | n/a  |  |  |  |
| CPs     | 1.1, 2.3, 2.5  |  |  |  |

## BI / read / 12

|         |   |                                      |  |  |  |
|---------|---|--------------------------------------|--|--|--|
| query   | BI / read / 12  |                                      |  |  |  |
| title   | How many persons have a given number of messages  |                                      |  |  |  |
| pattern | <div><div><div>2. personCount = count</div><div><div>Person</div></div><div>count Persons grouped by messageCount value</div></div><div><div>1. messageCount = count</div><div><div>Message</div><div>content not empty and length &lt; \$lengthThreshold and \$date &lt; creationDate</div></div></div><div><div>Post</div><div>language in \$languages</div></div><div><div>«opt» hasCreator</div><div>replyOf*0..</div></div></div>  |                                      |  |  |  |
| desc.   | <p>For each Person, count the number of Messages they made (messageCount). Only count Messages with the following attributes:</p> <ul style="list-style-type: none"><li>• Its content is not empty (and consequently, the <code>imageFile</code> attribute is empty for Posts).</li><li>• Its length is below the <code>lengthThreshold</code> (exclusive, equality is not allowed).</li><li>• Its <code>creationDate</code> is after date (exclusive, equality is not allowed).</li><li>• It is written in any of the given languages.</li></ul> <ul style="list-style-type: none"><li>– The language of a Post is defined by its <code>language</code> attribute.</li><li>– The language of a Comment is that of the Post that initiates the thread where the Comment replies to.</li></ul> <p>The Post and Comments in the reply tree's path (from the Message to the Post) do not have to satisfy the constraints for <code>content</code>, <code>length</code> and <code>creationDate</code>.</p> <p>For each <code>messageCount</code> value, count the number of Persons with exactly <code>messageCount</code> Messages (with the required attributes).</p> |                                      |  |  |  |
| params  | <div><div>1</div><div>date</div></div>  | <div><div>Date</div></div>           | <div></div>  |  |  |
|         | <div><div>2</div><div>lengthThreshold</div></div>   | <div><div>32-bit Integer</div></div> | <div>Selected as balanced against <code>date</code> to filter around 30% of the Messages within a language and keep the variance low</div> |  |  |
|         | <div><div>3</div><div>languages</div></div>   | <div><div>{String}</div></div>       | <div>Only the most frequently used languages are selected</div>  |  |  |
| result  | <div><div>1</div><div>messageCount</div></div>  | <div><div>32-bit Integer</div></div> | <div><div>A</div></div>  | <div>Number of Messages created</div>                                |  |
|         | <div><div>2</div><div>personCount</div></div>   | <div><div>32-bit Integer</div></div> | <div><div>A</div></div>  | <div>Number of Persons with <code>messageCount</code> Messages</div> |  |
| sort    | <div><div>1</div><div>personCount</div></div>   | <div><div>↓</div></div>              | <div></div>  |  |  |
|         | <div><div>2</div><div>messageCount</div></div>  | <div><div>↓</div></div>              | <div></div>  |  |  |
| limit   | n/a   |                                      |  |  |  |
| CPs     | 1.1, 1.2, 1.4, 3.2, 4.2, 4.3, 8.1, 8.2, 8.3, 8.4, 8.5   |                                      |  |  |  |

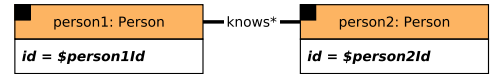
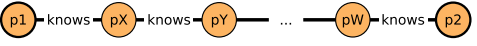
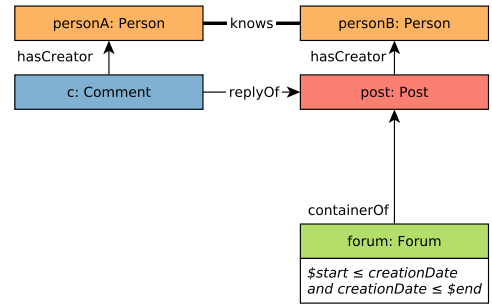
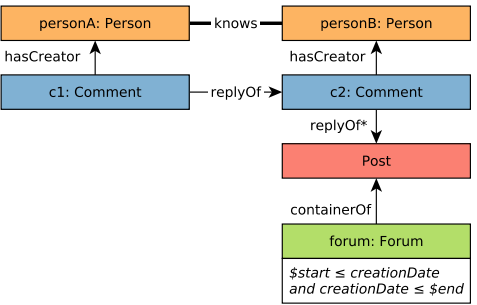
## BI / read / 13

|         |  |                 |                |   |  |
|---------|--|-----------------|----------------|---|--|
| query   | BI / read / 13   |                 |                |   |  |
| title   | Zombies in a country   |                 |                |   |  |
| pattern | <div><div>1. zombies = collect(zombie)</div><div><div><div>Country</div><div><b>name = \$country</b></div></div><div><div>City</div><div></div></div><div><div>zombie: Person</div><div>creationDate &lt; \$endDate<br/>and (messageCount / months) &lt; 1</div></div><div><div>message: Message</div><div>creationDate &lt; \$endDate</div></div><div>Country <math>\xrightarrow{\text{isPartOf}}</math> City</div><div>City <math>\xrightarrow{\text{isLocatedIn}}</math> zombie: Person</div><div>message: Message <math>\xrightarrow{\text{hasCreator}} \llcorner \text{opt} \llcorner</math> zombie: Person</div></div><div>messageCount = count(message)</div></div> <div><div>2. For each zombie IN zombies, calculate: zombieScore = zombieLikeCount / totalLikeCount</div><div><div><div>likerPerson: Person</div><div>creationDate &lt; \$endDate</div></div><div><div>Message</div><div></div></div><div><div>likerZombie: Person</div><div>creationDate &lt; \$endDate<br/>and likerZombie IN zombies</div></div><div>likerPerson: Person <math>\xrightarrow{\text{likes}} \llcorner \text{opt} \llcorner</math> Message</div><div>likerZombie: Person <math>\xrightarrow{\text{likes}} \llcorner \text{opt} \llcorner</math> Message</div><div>Message <math>\xrightarrow{\text{hasCreator}}</math> zombie: Person</div></div><div>totalLikeCount = count(likerPerson)</div><div>zombieLikeCount = count(likerZombie)</div></div> |                 |                |   |  |
| desc.   | <p>Find zombies within the given country, and return their zombie scores. A zombie is a Person created before the given endDate, which has created an average of <math>[0, 1)</math> Messages per month, during the time range between profile’s creationDate and the given endDate. The number of months spans the time range from the creationDate of the profile to the endDate with partial months on both end counting as one month (e.g. a creationDate of Jan 31 and an endDate of Mar 1 result in 3 months).</p> <p>For each zombie, calculate the following:</p> <ul style="list-style-type: none"><li>• zombieLikeCount: the number of likes received from other zombies.</li><li>• totalLikeCount: the total number of likes received.</li><li>• zombieScore: <math>\text{zombieLikeCount} / \text{totalLikeCount}</math>. If the value of totalLikeCount is 0, the zombieScore of the zombie should be 0.0.</li></ul> <p>For both zombieLikeCount and totalLikeCount, only consider likes received from profiles that were created before the given endDate.</p>   |                 |                |   |  |
| params  | 1  | country         | Long String    | Only the largest Countries are selected             |  |
|         | 2  | endDate         | Date           | Selected from the last days of the initial data set |  |
| result  | 1  | zombie.id       | ID             | R   |  |
|         | 2  | zombieLikeCount | 32-bit Integer | A   |  |
|         | 3  | totalLikeCount  | 32-bit Integer | A   |  |
|         | 4  | zombieScore     | 64-bit Float   | A   | Determined as $\text{zombieLikeCount} / \text{totalLikeCount}$ |
| sort    | 1  | zombieScore     | ↓              |   |  |
|         | 2  | zombie.id       | ↑              |   |  |
| limit   | 100  |                 |                |   |  |
| CPs     | 1.2, 2.1, 2.3, 2.4, 3.2, 3.3, 4.2, 5.1, 5.3, 8.2, 8.4, 8.5   |                 |                |   |  |

## BI / read / 14

|         |   |                |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
|---------|---|----------------|--|-------------|--|---|------------|-------------|----|---|------------|---|------------|-------------|---|--|---|-------|----------------|---|--|
| query   | BI / read / 14  |                |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| title   | International dialog  |                |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| pattern | <div><div>For each pair of countries, calculate the cost as a sum of cases #1-5. Cases that have a match add to the final score with the specified value. Each case only counts once, multiple matches do not increase to the score.</div><div><div><div><div>Country</div><div>name = \$country1</div></div><div><div>city1: City</div><div>name</div></div><div><div>person1: Person</div><div>id</div></div><div>isPartOf</div><div>isLocatedIn</div></div><div><div><div>Country</div><div>name = \$country2</div></div><div><div>City</div></div><div><div>person2: Person</div><div>id</div></div><div>isPartOf</div><div>isLocatedIn</div></div></div><div><div>Case 1: score += 4</div><div><div><div>person1: Person</div><div>hasCreator</div><div>Comment</div></div><div><div>person2: Person</div><div>hasCreator</div><div>Message</div></div><div>replyOf</div></div></div><div><div>Case 2: score += 1</div><div><div><div>person1: Person</div><div>hasCreator</div><div>Message</div></div><div><div>person2: Person</div><div>hasCreator</div><div>Comment</div></div><div>replyOf</div></div></div><div><div>Case 3: score += 15</div><div><div><div>person1: Person</div><div>knows</div><div>person2: Person</div></div></div></div><div><div>Case 4: score += 10</div><div><div><div>person1: Person</div><div>likes</div><div>Message</div></div><div><div>person2: Person</div><div>hasCreator</div><div>Message</div></div></div></div><div><div>Case 5: score += 1</div><div><div><div>person1: Person</div><div>hasCreator</div><div>Message</div></div><div><div>person2: Person</div><div>likes</div><div>Message</div></div></div></div></div> |                |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| desc.   | <p>Consider all pairs of people (person1, person2) such that one is located in a City of Country country1 and the other is located in a City of Country country2. For each City of Country country1, return the highest scoring pair. The score of a pair is defined as the sum of the subscores awarded for the following kinds of interaction. The initial value is score = 0.</p> <ol style="list-style-type: none"><li>1. person1 has created a reply Comment to at least one Message by person2: score += 4</li><li>2. person1 has created at least one Message that person2 has created a reply to: score += 1</li><li>3. person1 and person2 know each other: score += 15</li><li>4. person1 liked at least one Message by person2: score += 10</li><li>5. person1 has created at least one Message that was liked by person2: score += 1</li></ol> <p>Consequently, the maximum score a pair can obtain is: 4 + 1 + 15 + 10 + 1 = 31.</p> <p>This query has two variants based on whether the input parameters are selected as correlated (close countries) or uncorrelated (far countries).</p>  |                |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| params  | <table><tr><td>1</td><td>country1</td><td>Long String</td><td>A: correlated with parameter country2, i.e. the countries are close and there are many Persons visiting both Countries.<br/>B: uncorrelated with parameter country2, i.e. the countries are afar and there are few Persons visiting both Countries.</td></tr><tr><td>2</td><td>country2</td><td>Long String</td><td></td></tr></table>  | 1              | country1   | Long String | A: correlated with parameter country2, i.e. the countries are close and there are many Persons visiting both Countries.<br>B: uncorrelated with parameter country2, i.e. the countries are afar and there are few Persons visiting both Countries. | 2 | country2   | Long String |    |   |            |   |            |             |   |  |   |       |                |   |  |
| 1       | country1  | Long String    | A: correlated with parameter country2, i.e. the countries are close and there are many Persons visiting both Countries.<br>B: uncorrelated with parameter country2, i.e. the countries are afar and there are few Persons visiting both Countries. |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| 2       | country2  | Long String    |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| result  | <table><tr><td>1</td><td>person1.id</td><td>ID</td><td>R</td><td></td></tr><tr><td>2</td><td>person2.id</td><td>ID</td><td>R</td><td></td></tr><tr><td>3</td><td>city1.name</td><td>Long String</td><td>R</td><td></td></tr><tr><td>4</td><td>score</td><td>32-bit Integer</td><td>C</td><td></td></tr></table>   | 1              | person1.id   | ID          | R  |   | 2          | person2.id  | ID | R |            | 3 | city1.name | Long String | R |  | 4 | score | 32-bit Integer | C |  |
| 1       | person1.id  | ID             | R  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| 2       | person2.id  | ID             | R  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| 3       | city1.name  | Long String    | R  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| 4       | score   | 32-bit Integer | C  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| sort    | <table><tr><td>1</td><td>score</td><td>↓</td><td></td></tr><tr><td>2</td><td>person1.id</td><td>↑</td><td></td></tr><tr><td>3</td><td>person2.id</td><td>↑</td><td></td></tr></table>   | 1              | score  | ↓           |  | 2 | person1.id | ↑           |    | 3 | person2.id | ↑ |            |             |   |  |   |       |                |   |  |
| 1       | score   | ↓              |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| 2       | person1.id  | ↑              |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| 3       | person2.id  | ↑              |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| limit   | n/a   |                |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |
| CPs     | 1.3, 1.4, 2.1, 3.1, 3.3, 5.1, 5.2, 5.3, 8.3, 8.4  |                |  |             |  |   |            |             |    |   |            |   |            |             |   |  |   |       |                |   |  |

## BI / read / 15

|       |         |   |           |              |   |
|-------|---------|---|-----------|--------------|---|
| BI 1  | query   | BI / read / 15  |           |              |   |
| BI 2  | title   | Trusted connection paths through forums created in a given timeframe  |           |              |   |
| BI 3  | pattern | <p>Enumerate all unweighted shortest paths on knows edges between person1 to person2.</p>    |           |              |   |
| BI 4  |         | <p>For each knows edge in the path, calculate a weight based on interactions between the pair of Persons of the edge, calculated as a sum of cases #1 and #2 for the Persons (both ways), and the sum of these weights determine the total weight of each path.</p>   |           |              |   |
| BI 5  |         | <p>Case 1: Replies on Posts, weight += 1.0 × count(c)</p>    |           |              |   |
| BI 6  |         | <p>Case 2: Replies on Comments, weight += 0.5 × count(c1)</p>   |           |              |   |
| BI 7  |         |   |           |              |   |
| BI 8  |         |   |           |              |   |
| BI 9  |         |   |           |              |   |
| BI 10 |         |   |           |              |   |
| BI 11 | desc.   | <p>Given two Persons, find all (unweighted) shortest paths between these two Persons, in the subgraph induced by the knows relationship.</p> <p>Then, for each path calculate a weight. The nodes in the path are Persons, and the weight of a path is the sum of weights between every pair of consecutive Person nodes in the path.</p> <p>The weight for a pair of Persons is calculated based on their interactions:</p> <ul style="list-style-type: none"> <li>• Every direct reply (by one of the Persons) to a Post (by the other Person) contributes 1.0.</li> <li>• Every direct reply (by one of the Persons) to a Comment (by the other Person) contributes 0.5.</li> </ul> <p>Only consider Messages that were created in a Forum that was created within the timeframe (interval) [startDate, endDate]. Note that for Comments, the containing Forum is that of the Post that the comment (transitively) replies to. Also note that interactions are counted both ways.</p> <p>Return all paths with the Person IDs ordered by their weights descending.</p> |           |              |   |
| BI 12 |         |   |           |              |   |
| BI 13 |         |   |           |              |   |
| BI 14 |         |   |           |              |   |
| BI 15 |         |   |           |              |   |
| BI 16 |         |   |           |              |   |
| BI 17 |         |   |           |              |   |
| BI 18 |         |   |           |              |   |
| BI 19 | params  | 1   | person1Id | ID           |   |
| BI 20 |         | 2   | person2Id | ID           |   |
|       |         | 3   | startDate | Date         |   |
|       |         | 4   | endDate   | Date         |   |
|       | result  | 1   | person.id | [ID]         | C Ordered sequence of the Person IDs in the path          |
|       |         | 2   | weight    | 64-bit Float | C   |
|       | sort    | 1   | weight    | ↓            | The order of paths with the same weight is unspecified    |
|       |         | 2   | personIds | ↑            | The IDs in the paths are used for lexicographical sorting |
|       | limit   | n/a   |           |              |   |
|       | CPs     | 1.2, 2.1, 2.2, 2.4, 3.3, 5.1, 5.3, 7.2, 7.3, 7.5, 7.7, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6   |           |              |   |

## BI / read / 16

|           |   |  |  |  |  |
|-----------|---|--|--|--|--|
| query     | BI / read / 16  |  |  |  |  |
| title     | Fake news detection   |  |  |  |  |
| pattern   | <div><div>For \$tagX/\$dayX in [tagA/dateA, tagB/dateB], compute scoreX = count(messageX)</div><div><div>1. Create an induced subgraph of Persons who created a Message with Tag \$tagX on \$dateX</div><div><div><div>tag: Tag</div><div>name = \$tagX</div></div><div>hasTag</div><div><div>Message</div><div>day(creationDate) = \$dateX</div></div><div>hasCreator</div><div><div>person: Person</div></div></div></div><div><div>2. In the subgraph, count the Messages (using the same conditions) from People with ≤ \$maxKnowsLimit friends</div><div><div><div>tag: Tag</div><div>name = \$tagX</div></div><div>hasTag</div><div><div>count(messageX)</div><div><div>messageX: Message</div><div>day(creationDate) = \$dateX</div></div></div><div>hasCreator</div><div><div>person: Person</div><div>count ≤ \$maxKnowsLimit</div><div>«opt» knows</div><div>Person</div></div></div></div></div> |  |  |  |  |
| desc.     | <p>Given two Tag/date pairs (tagA/dateA and tagB/dateB), for each pair tagX/dateX:</p> <ul style="list-style-type: none"><li>• Create an induced subgraph between Persons where for each pair of Persons person1/person2, both have created a Message on the day of dateX with Tag tagX.</li><li>• In the induced subgraph, only keep pairs of Persons who have at most maxKnowsLimit friends (in the induced subgraph).</li><li>• For these Persons, count the number of Messages created on dateX with Tag tagX.</li></ul> <p>Return Persons who had at least one Messages for both tagA/dateA and tagB/dateB ranked by their total number of Messages (descending).</p>  |  |  |  |  |
| params    | <div><div>1</div><div>tagA</div><div>Long String</div><div></div></div> <div><div>2</div><div>dateA</div><div>Date</div><div></div></div> <div><div>3</div><div>tagB</div><div>Long String</div><div></div></div> <div><div>4</div><div>dateB</div><div>Date</div><div></div></div> <div><div>5</div><div>maxKnowsLimit</div><div>32-bit Integer</div><div>Selected between 3 and 6</div></div>   |  |  |  |  |
| result    | <div><div>1</div><div>person.id</div><div>ID</div><div>R</div><div></div></div> <div><div>2</div><div>messageCountA</div><div>32-bit Integer</div><div>A</div><div>Message count for tagA/dateA</div></div> <div><div>3</div><div>messageCountB</div><div>32-bit Integer</div><div>A</div><div>Message count for tagB/dateB</div></div>   |  |  |  |  |
| sort      | <div><div>1</div><div>messageCountA + messageCountB</div><div>↓</div><div></div></div> <div><div>2</div><div>person.id</div><div>↑</div><div></div></div>   |  |  |  |  |
| limit     | 20  |  |  |  |  |
| CPs       | 5.3, 8.4, 8.5   |  |  |  |  |
| relevance | <p>There are two major ways to compute this query: (1) create the induced subgraph as suggested by the specification (either as a view or in materialized form), or (2) skip creating the induced subgraph and perform on-the-fly check for the number of friends (who also posted at least one Message with the given Tag on the given date). The latter approach is easier to express in systems which do not provide graph views but might result in redundant computations (the query engine will might repeatedly check whether a Person has at least one Message that satisfies the conditions).</p>  |  |  |  |  |



## BI / read / 17

|         |  |              |                |   |
|---------|--|--------------|----------------|---|
| query   | BI / read / 17   |              |                |   |
| title   | Information propagation analysis   |              |                |   |
| pattern |  |              |                |   |
| desc.   | <p>This query aims to identify instances of “information propagation” when a Person (person1) submits a Message (message1) with a given Tag (tag) to a Forum (forum1). This is read by other members of forum1, Persons person2 and person3. Some time later (specified by the delta parameter), these persons have a discussion with the same tag in a different Forum (forum2) where person1 is not a member. The discussion consists of a Message (message2) by person2 and a direct reply Comment (comment) by person3.</p> <p>Return IDs of person1 with the number of interactions their Messages (might have) caused.</p> |              |                |   |
| params  | 1  | tag          | Long String    | Tags with a similar amount of Messages are selected       |
|         | 2  | delta        | 32-bit Integer | Measured in hours, selected to be between 8 and 16 hours. |
| result  | 1  | person1.id   | ID             | R   |
|         | 2  | messageCount | 32-bit Integer | A   |
| sort    | 1  | messageCount | ↓              |   |
|         | 2  | person1.id   | ↑              |   |
| limit   | 10   |              |                |   |
| CPs     | 2.1, 2.3, 8.1  |              |                |   |

BI / read / 18

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 / 18  |                   |                |   |
| title   | Friend recommendation   |                   |                |   |
| pattern | <div><div>For each person1 compute top-k(person2) based on mutualFriendCount</div><div><div><div>person1: Person</div><div>id = \$person1Id</div><div>«neg» knows</div></div><div>knows</div><div><div>Person</div><div>mutualFriendCount = count(*)</div></div><div>knows</div><div><div>person2: Person</div><div>≠ person1</div><div>id</div></div><div><div>tag: Tag</div><div>name = \$tag</div></div><div>hasInterest</div></div></div> |                   |                |   |
| desc.   | <div>For a given Person (person1) and a Tag (tag), recommend new friends (person2) who</div> <ul style="list-style-type: none"><li>do not yet know person1</li><li>have many mutual friends with person1</li><li>are interested in tag.</li></ul> <div>Rank Persons person2 based on the number of mutual friends.</div>  |                   |                |   |
| params  | 1   | person1Id         | ID             | Persons with a similar amount of friends are selected |
|         | 2   | tag               | Long String    | Tags with a similar amount of Messages are selected   |
| result  | 1   | person2.id        | ID             | R   |
|         | 2   | mutualFriendCount | 32-bit Integer | A   |
| sort    | 1   | mutualFriendCount | ↓              |   |
|         | 2   | person2.id        | ↑              |   |
| limit   | 20  |                   |                |   |
| CPs     | 2.5, 8.1  |                   |                |   |

## BI / read / 19

|       |           |  |             |              |   |
|-------|-----------|--|-------------|--------------|---|
| BI 1  | query     | BI / read / 19   |             |              |   |
| BI 2  | title     | Interaction path between cities  |             |              |   |
| BI 3  | pattern   | <p>Find the shortest paths between all pairs of Persons in city1 and city2</p> <p>city1: City<br/>id = \$city1id</p> <p>city2: City<br/>id = \$city2id</p> <p>person1: Person<br/>person2: Person</p> <p>isLocatedIn<br/>compute weighted shortest paths on knows.weight</p> <p>The weight of a knows edge is based on the number of interactions between its Persons:<br/>knows.weight = 1 / (count(i1)+count(i2))</p> <p>Case i1: Reply from personA to Person B's Message</p> <p>Case i2: Reply from personB to personA's Message</p>   |             |              |   |
| BI 14 | desc.     | <p>Given two Cities <i>city1</i>, <i>city2</i>, find Persons <i>person1</i>, <i>person2</i> living in these Cities (respectively) with the shortest <i>interaction path</i> between them. If there are multiple pairs of people with shortest paths having the same total weight, return all of them.</p> <p>The shortest path is computed using a weight between two Persons defined as the reciprocal of the number of interactions (direct reply Comments to a Message by the other Person). Therefore, more interactions imply a smaller weight.</p> <p><i>Note:</i> Interactions are counted both ways, i.e. if Alice writes 2 reply Comments to Bob's Messages and Bob writes 3 reply Comments to Alice's Messages, their total number of interactions is 5.</p> |             |              |   |
| BI 15 | params    | 1  | city1id     | ID           | Small Cities within the same Country are selected |
| BI 16 |           | 2  | city2id     | ID           |   |
| BI 17 | result    | 1  | person1.id  | ID           | R   |
| BI 18 |           | 2  | person2.id  | ID           | R   |
| BI 19 |           | 3  | totalWeight | 64-bit Float | C   |
| BI 20 | sort      | 1  | totalWeight | ↑            |   |
|       |           | 2  | person1.id  | ↑            |   |
|       |           | 3  | person2.id  | ↑            |   |
|       | limit     | 20   |             |              |   |
|       | CPs       | 3.3, 7.6, 7.7, 8.4, 8.6  |             |              |   |
|       | relevance | <p>Finding shortest paths between pairs of Persons in Cities can be implemented in theory with an <i>all-pairs shortest paths</i> algorithm. However, this needs to be executed on the whole Person-knows-Person graph (with edge weights derived from the number of interactions) so it is expected to be prohibitively expensive. A better approach is using multiple <i>single-source shortest path algorithms</i> (e.g. from the City with fewer inhabitants).</p>   |             |              |   |

## BI / read / 20

|       |           |   |             |                |   |
|-------|-----------|---|-------------|----------------|---|
| BI 1  | query     | BI / read / 20  |             |                |   |
| BI 2  | title     | Recruitment   |             |                |   |
| BI 3  | pattern   |  <p>The diagram illustrates the query pattern. It shows a <b>company: Company</b> node with a property <b>name = \$company</b>. A <b>person1: Person</b> node is connected to the company via a <b>workAt</b> edge. A <b>person2: Person</b> node is connected to <b>person1</b> via a <b>shortest path on knows.weight</b> edge, with a constraint <b>id = \$person2Id</b>. A separate sub-diagram shows the <b>knows</b> edge between <b>personA: Person</b> and <b>personB: Person</b>, with a weight calculation: <b>knows.weight: abs(saA.classYear - saB.classYear) + 1</b>. Both <b>personA</b> and <b>personB</b> are connected to a <b>University</b> node via <b>studyAt</b> edges, labeled <b>saA: studyAt</b> and <b>saB: studyAt</b> respectively.</p> |             |                |   |
| BI 10 | desc.     | <p>Given a Company <i>company</i> and a Person <i>person2</i> (who is known to be working at another Company), find a different Person (<i>person1</i>) who works in <i>company</i> and is reachable by from <i>person2</i> through people who have studied together. On this path, we only consider edges between Persons who know each other and attended the same university and set the weight of the edge to the absolute difference between the year of enrolment plus 1 (<i>studyAt.classYear</i> + 1). We return the 20 shortest paths.</p> <p>If there are multiple Person <i>person1</i> nodes with the same shortest path, return all of them.</p>   |             |                |   |
| BI 17 | params    | 1   | company     | Long String    | Companies with a similar number of employees (former or current) are selected |
| BI 18 |           | 2   | person2Id   | ID             |   |
| BI 19 | result    | 1   | person1.id  | ID             | R   |
| BI 20 |           | 2   | totalWeight | 64-bit Integer | C   |
|       | sort      | 1   | totalWeight | ↑              |   |
|       |           | 2   | person1.id  | ↑              |   |
|       | limit     | 20  |             |                |   |
|       | CPs       | 3.3, 7.6, 7.7, 8.4, 8.6   |             |                |   |
|       | relevance | Implementations can either pre-compute edge weights or compute them on-the-fly. To find a weighted shortest path efficiently, implementations can use e.g. a bidirectional Dijkstra algorithm.  |             |                |   |