# Health Economics Query Requirements

**Note**: There are still issues with the data, the query result will make better sense when issues are fixed.

**Color Legend**: comment parameter todo-item

1) Sentences that contain at least two different aspects that can be linked to a specific drug, along with the aspects, associated opinions/polarities, forum ID, post ID, timestamp, and user demographics/characteristics if available.

Todo: need better method for linking drugs, currently a drug is “linked” if it appears in the same sentence when an aspect/opinion pair is extracted

select d1.describes\_id, d2.describes\_id, d1.describes\_aspect\_name as asp\_1, o1.opinion\_word\_polarity as opi\_1, d2.describes\_aspect\_name as asp\_2, o2.opinion\_word\_polarity as opi\_2, sentence\_body, forum\_name, post\_id, post\_timestamp, author\_age, author\_gender

from describes as d1, opinion\_word as o1, describes as d2, opinion\_word as o2, sentence, author, post, forum

where d1.describes\_sentence\_id = sentence.sentence\_id

and d2.describes\_sentence\_id = sentence.sentence\_id

and d1.describes\_id > d2.describes\_id

and d1.describes\_opinion\_word\_name = o1.opinion\_word\_name

and d2.describes\_opinion\_word\_name = o2.opinion\_word\_name

and o1.opinion\_word\_name > o2.opinion\_word\_name

and sentence.sentence\_post\_id = post.post\_id

and author.author\_username = post.post\_author\_username

and post.post\_forum\_name = forum.forum\_name

and sentence.sentence\_body like "%Ocrevus%"

A screenshot of a computer

Description automatically generated

2) Sentiment for DrugA associated with **mentions** by female patients over age 55 (presumably this would return a table with each row being an aspect-opinion dyad extracted from the text, right?), along with timestamp, forum ID, and user demographics/characteristics if available.

Question: “mention” means: the treatment occurs in the sentence, or extracted aspect is the treatment?

select describes\_aspect\_name, describes\_opinion\_word\_name, sentence\_body, post\_timestamp, author\_flair, author\_age, author\_gender, forum\_name

from describes, sentence, post, author, forum

where describes.describes\_sentence\_id = sentence.sentence\_id

and post.post\_id = sentence.sentence\_post\_id

and author.author\_username = post.post\_author\_username

and forum\_name = post.post\_forum\_name

and author.author\_gender = "F"

and author.author\_age > 40

and sentence.sentence\_body like "%treatment%"

A screenshot of a social media post

Description automatically generated

in this case the drug name is "treatment"

3) Volume of posts with at least one mention of DrugB over time in ForumX, independently of whether the drug or specific reference was extracted as an aspect by ABSApp

Todo: Populate database with drug aliases

SELECT count(\*), extract(day from post.post\_timestamp) as day, extract(month from post.post\_timestamp) as month, extract(year from post.post\_timestamp) as year

FROM sentence, post, forum

WHERE sentence.sentence\_post\_id = post.post\_id

AND forum.forum\_name = "r/MultipleSclerosis"

AND sentence\_body like "%like%"

GROUP BY extract(day from post.post\_timestamp), extract(month from post.post\_timestamp), extract(year from post.post\_timestamp)

A screenshot of a social media post

Description automatically generated

4) RA forum users who mentioned at least one drug that was extracted as an aspect by ABSApp in at least one of their posts, along with a list of all the drugs each user mentioned and **associated sentiments** (if available). Question: sould just other drug’s average overall sentiment? Alias?

Todo: in the process of working, currently “Ocrevus” is the only treatment in the database, so cannot extract other drugs yet

In progress of development