

Retrieval exercise

In this exercise, you will implement the query likelihood model with Jelinek-Mercer smoothing. This assignment builds on the previous assignment and you will need to reuse some code from the Indexing notebook.

1. Build the index

Download the MS MARCO passage collection and build an index using [Pyserini](#).

In [1]:

```
pip install pyserini
```

```
Collecting pyserini
  Downloading pyserini-0.13.0-py3-none-any.whl (72.8 MB)
    |████████████████████| 72.8 MB 35 kB/s
Requirement already satisfied: tqdm in /usr/local/lib/python3.7/dist-packages (from pyserini) (4.62.3)
Collecting pyjnius>=1.2.1
  Downloading pyjnius-1.4.1-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux2_12_x86_64.manylinux2010_x86_64.whl (1.1 MB)
    |████████████████████| 1.1 MB 65.9 MB/s
Requirement already satisfied: scikit-learn>=0.22.1 in /usr/local/lib/python3.7/dist-packages (from pyserini) (0.22.2.post1)
Requirement already satisfied: numpy>=1.18.1 in /usr/local/lib/python3.7/dist-packages (from pyserini) (1.19.5)
Collecting sentencepiece>=0.1.95
  Downloading sentencepiece-0.1.96-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.2 MB)
    |████████████████████| 1.2 MB 69.8 MB/s
Collecting transformers>=4.6.0
  Downloading transformers-4.12.3-py3-none-any.whl (3.1 MB)
    |████████████████████| 3.1 MB 43.4 MB/s
Requirement already satisfied: scipy>=1.4.1 in /usr/local/lib/python3.7/dist-packages (from pyserini) (1.4.1)
Collecting nmslib
  Downloading nmslib-2.1.1-cp37-cp37m-manylinux2010_x86_64.whl (13.5 MB)
    |████████████████████| 13.5 MB 75 kB/s
Requirement already satisfied: pandas>=1.1.5 in /usr/local/lib/python3.7/dist-packages (from pyserini) (1.1.5)
Requirement already satisfied: Cython>=0.29.21 in /usr/local/lib/python3.7/dist-packages (from pyserini) (0.29.24)
Requirement already satisfied: python-dateutil>=2.7.3 in /usr/local/lib/python3.7/dist-packages (from pandas>=1.1.5->pyserini) (2.8.2)
Requirement already satisfied: pytz>=2017.2 in /usr/local/lib/python3.7/dist-packages (from pandas>=1.1.5->pyserini) (2018.9)
Requirement already satisfied: six>=1.7.0 in /usr/local/lib/python3.7/dist-packages (from pyjnius>=1.2.1->pyserini) (1.15.0)
Requirement already satisfied: joblib>=0.11 in /usr/local/lib/python3.7/dist-packages (from scikit-learn>=0.22.1->pyserini) (1.0.1)
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages (from transformers>=4.6.0->pyserini) (2.23.0)
Collecting sacremoses
  Downloading sacremoses-0.0.46-py3-none-any.whl (895 kB)
    |████████████████████| 895 kB 46.3 MB/s
Collecting tokenizers<0.11,>=0.10.1
  Downloading tokenizers-0.10.3-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux2_12_x86_64.manylinux2010_x86_64.whl (3.3 MB)
    |████████████████████| 3.3 MB 42.5 MB/s
Collecting huggingface-hub<1.0,>=0.1.0
  Downloading huggingface_hub-0.1.0-py3-none-any.whl (59 kB)
    |████████████████████| 59 kB 6.3 MB/s
Collecting pyyaml>=5.1
  Downloading PyYAML-6.0-cp37-cp37m-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux2_12_x86_64.manylinux2010_x86_64.whl (596 kB)
    |████████████████████| 596 kB 59.9 MB/s
```

```
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.7/dist-packages
s (from transformers>=4.6.0->pyserini) (2019.12.20)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-packages (from t
ransformers>=4.6.0->pyserini) (3.3.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.7/dist-packages
(from transformers>=4.6.0->pyserini) (21.0)
Requirement already satisfied: importlib-metadata in /usr/local/lib/python3.7/dist-packag
es (from transformers>=4.6.0->pyserini) (4.8.1)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.7/dist-package
s (from huggingface-hub<1.0,>=0.1.0->transformers>=4.6.0->pyserini) (3.7.4.3)
Requirement already satisfied: pyparsing>=2.0.2 in /usr/local/lib/python3.7/dist-packages
(from packaging>=20.0->transformers>=4.6.0->pyserini) (2.4.7)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from
importlib-metadata->transformers>=4.6.0->pyserini) (3.6.0)
Requirement already satisfied: psutil in /usr/local/lib/python3.7/dist-packages (from nms
lib->pyserini) (5.4.8)
Collecting pybind11<2.6.2
  Downloading pybind11-2.6.1-py2.py3-none-any.whl (188 kB)
    | 188 kB 75.5 MB/s
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-package
s (from requests->transformers>=4.6.0->pyserini) (3.0.4)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/
python3.7/dist-packages (from requests->transformers>=4.6.0->pyserini) (1.24.3)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packag
es (from requests->transformers>=4.6.0->pyserini) (2021.5.30)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (fr
om requests->transformers>=4.6.0->pyserini) (2.10)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages (from sacr
emoses->transformers>=4.6.0->pyserini) (7.1.2)
Installing collected packages: pyyaml, tokenizers, sacremoses, pybind11, huggingface-hub,
transformers, sentencepiece, pyjnius, nmslib, pyserini
  Attempting uninstall: pyyaml
    Found existing installation: PyYAML 3.13
    Uninstalling PyYAML-3.13:
      Successfully uninstalled PyYAML-3.13
Successfully installed huggingface-hub-0.1.0 nmslib-2.1.1 pybind11-2.6.1 pyjnius-1.4.1 py
serini-0.13.0 pyyaml-6.0 sacremoses-0.0.46 sentencepiece-0.1.96 tokenizers-0.10.3 transfo
rmers-4.12.3
```

In [2]:

```
!git clone https://github.com/castorini/anserini.git
!cd anserini && git checkout ad5balc76196436f8a0e28efdb69960d4873efe3
```

```
Cloning into 'anserini'...
remote: Enumerating objects: 17780, done.
remote: Counting objects: 100% (2251/2251), done.
remote: Compressing objects: 100% (1082/1082), done.
remote: Total 17780 (delta 1338), reused 1795 (delta 1012), pack-reused 15529
Receiving objects: 100% (17780/17780), 31.31 MiB | 18.17 MiB/s, done.
Resolving deltas: 100% (10256/10256), done.
Note: checking out 'ad5balc76196436f8a0e28efdb69960d4873efe3'.
```

You are in 'detached HEAD' state. You can look around, make experimental changes and commit them, and you can discard any commits you make in this state without impacting any branches by performing another checkout.

If you want to create a new branch to retain commits you create, you may do so (now or later) by using `-b` with the checkout command again. Example:

```
git checkout -b <new-branch-name>
```

HEAD is now at ad5balc7 Release notes for v0.9.2 (#1197)

In [3]:

```
!wget https://msmarco.blob.core.windows.net/msmarcoranking/collection.tar.gz -P data/msm
arco_passage/
```

```
--2021-11-04 12:45:10-- https://msmarco.blob.core.windows.net/msmarcoranking/collection.
tar.gz
```

```
Resolving msmarco.blob.core.windows.net (msmarco.blob.core.windows.net) 20 150 34 4
```

```
Receiving msmarco.blob.core.windows.net (msmarco.blob.core.windows.net)... 20.150.34.4  
Connecting to msmarco.blob.core.windows.net (msmarco.blob.core.windows.net) |20.150.34.4|:  
443... connected.  
HTTP request sent, awaiting response... 200 OK  
Length: 1035009698 (987M) [application/octet-stream]  
Saving to: 'data/msmarco_passage/collection.tar.gz'
```

```
collection.tar.gz 100%[=====>] 987.06M 20.6MB/s in 39s
```

```
2021-11-04 12:45:49 (25.1 MB/s) - 'data/msmarco_passage/collection.tar.gz' saved [1035009  
698/1035009698]
```

In [4]:

```
!!ls data/msmarco_passage/  
!!tar xvfz data/msmarco_passage/collection.tar.gz -C data/msmarco_passage  
  
collection.tar.gz  
collection.tsv
```

In [5]:

```
!cd anserini && python ./src/main/python/msmarco/convert_collection_to_jsonl.py \  
--collection_path ../data/msmarco_passage/collection.tsv --output_folder ../data/msmarco  
_passage/collection_jsonl
```

```
Converting collection...  
Converted 0 docs in 1 files  
Converted 100000 docs in 1 files  
Converted 200000 docs in 1 files  
Converted 300000 docs in 1 files  
Converted 400000 docs in 1 files  
Converted 500000 docs in 1 files  
Converted 600000 docs in 1 files  
Converted 700000 docs in 1 files  
Converted 800000 docs in 1 files  
Converted 900000 docs in 1 files  
Converted 1000000 docs in 2 files  
Converted 1100000 docs in 2 files  
Converted 1200000 docs in 2 files  
Converted 1300000 docs in 2 files  
Converted 1400000 docs in 2 files  
Converted 1500000 docs in 2 files  
Converted 1600000 docs in 2 files  
Converted 1700000 docs in 2 files  
Converted 1800000 docs in 2 files  
Converted 1900000 docs in 2 files  
Converted 2000000 docs in 3 files  
Converted 2100000 docs in 3 files  
Converted 2200000 docs in 3 files  
Converted 2300000 docs in 3 files  
Converted 2400000 docs in 3 files  
Converted 2500000 docs in 3 files  
Converted 2600000 docs in 3 files  
Converted 2700000 docs in 3 files  
Converted 2800000 docs in 3 files  
Converted 2900000 docs in 3 files  
Converted 3000000 docs in 4 files  
Converted 3100000 docs in 4 files  
Converted 3200000 docs in 4 files  
Converted 3300000 docs in 4 files  
Converted 3400000 docs in 4 files  
Converted 3500000 docs in 4 files  
Converted 3600000 docs in 4 files  
Converted 3700000 docs in 4 files  
Converted 3800000 docs in 4 files  
Converted 3900000 docs in 4 files  
Converted 4000000 docs in 5 files  
Converted 4100000 docs in 5 files  
Converted 4200000 docs in 5 files  
Converted 4300000 docs in 5 files  
Converted 4400000 docs in 5 files
```

```
Converted 1100000 docs in 5 files
Converted 1200000 docs in 5 files
Converted 1300000 docs in 5 files
Converted 1400000 docs in 5 files
Converted 1500000 docs in 5 files
Converted 1600000 docs in 5 files
Converted 1700000 docs in 5 files
Converted 1800000 docs in 5 files
Converted 1900000 docs in 5 files
Converted 2000000 docs in 6 files
Converted 2100000 docs in 6 files
Converted 2200000 docs in 6 files
Converted 2300000 docs in 6 files
Converted 2400000 docs in 6 files
Converted 2500000 docs in 6 files
Converted 2600000 docs in 6 files
Converted 2700000 docs in 6 files
Converted 2800000 docs in 6 files
Converted 2900000 docs in 6 files
Converted 3000000 docs in 6 files
Converted 3100000 docs in 7 files
Converted 3200000 docs in 7 files
Converted 3300000 docs in 7 files
Converted 3400000 docs in 7 files
Converted 3500000 docs in 7 files
Converted 3600000 docs in 7 files
Converted 3700000 docs in 7 files
Converted 3800000 docs in 7 files
Converted 3900000 docs in 7 files
Converted 4000000 docs in 7 files
Converted 4100000 docs in 8 files
Converted 4200000 docs in 8 files
Converted 4300000 docs in 8 files
Converted 4400000 docs in 8 files
Converted 4500000 docs in 8 files
Converted 4600000 docs in 8 files
Converted 4700000 docs in 8 files
Converted 4800000 docs in 8 files
Converted 4900000 docs in 8 files
Converted 5000000 docs in 8 files
Converted 5100000 docs in 8 files
Converted 5200000 docs in 8 files
Converted 5300000 docs in 8 files
Converted 5400000 docs in 8 files
Converted 5500000 docs in 8 files
Converted 5600000 docs in 8 files
Converted 5700000 docs in 8 files
Converted 5800000 docs in 8 files
Converted 5900000 docs in 8 files
Converted 6000000 docs in 9 files
Converted 6100000 docs in 9 files
Converted 6200000 docs in 9 files
Converted 6300000 docs in 9 files
Converted 6400000 docs in 9 files
Converted 6500000 docs in 9 files
Converted 6600000 docs in 9 files
Converted 6700000 docs in 9 files
Converted 6800000 docs in 9 files
Converted 6900000 docs in 9 files
Converted 7000000 docs in 9 files
Converted 7100000 docs in 9 files
Converted 7200000 docs in 9 files
Converted 7300000 docs in 9 files
Converted 7400000 docs in 9 files
Converted 7500000 docs in 9 files
Converted 7600000 docs in 9 files
Converted 7700000 docs in 9 files
Converted 7800000 docs in 9 files
Converted 7900000 docs in 9 files
Converted 8000000 docs in 9 files
Converted 8100000 docs in 9 files
Converted 8200000 docs in 9 files
Converted 8300000 docs in 9 files
Converted 8400000 docs in 9 files
Converted 8500000 docs in 9 files
Converted 8600000 docs in 9 files
Converted 8700000 docs in 9 files
Converted 8800000 docs in 9 files
Done!
```

In [6]:

```
!ls data/msmarco_passage
!rm data/msmarco_passage/*.tsv
!ls data/msmarco_passage
!rm -rf sample_data
```

```
collection_jsonl collection.tar.gz collection.tsv
collection_jsonl collection.tar.gz
```

In [7]:

```
!python -m pyserini.index -collection JsonCollection -generator DefaultLuceneDocumentGen
erator -threads 9 \
-input data/msmarco_passage/collection_jsonl -index indexes/lucene-index-msmarco-passag
e -storePositions -storeDocvectors -storeRaw
```

WARNING: sun.reflect.Reflection.getCallerClass is not supported. This will impact perform

ance.
2021-11-04 12:48:02,065 INFO [main] index.IndexCollection (IndexCollection.java:643) - S

etting log level to INFO
2021-11-04 12:48:02,072 INFO [main] index.IndexCollection (IndexCollection.java:646) - S

tarting indexer...
2021-11-04 12:48:02,072 INFO [main] index.IndexCollection (IndexCollection.java:647) - =

=====Loading Parameters=====

2021-11-04 12:48:02,073 INFO [main] index.IndexCollection (IndexCollection.java:648) - D

```
2021-11-04 12:48:02,073 INFO [main] index.IndexCollection (IndexCollection.java:616) - DocumentCollection path: data/msmarco_passage/collection_jsonl
2021-11-04 12:48:02,074 INFO [main] index.IndexCollection (IndexCollection.java:649) - CollectionClass: JsonCollection
2021-11-04 12:48:02,074 INFO [main] index.IndexCollection (IndexCollection.java:650) - Generator: DefaultLuceneDocumentGenerator
2021-11-04 12:48:02,074 INFO [main] index.IndexCollection (IndexCollection.java:651) - Threads: 9
2021-11-04 12:48:02,075 INFO [main] index.IndexCollection (IndexCollection.java:652) - Stemmer: porter
2021-11-04 12:48:02,075 INFO [main] index.IndexCollection (IndexCollection.java:653) - Keep stopwords? false
2021-11-04 12:48:02,076 INFO [main] index.IndexCollection (IndexCollection.java:654) - Stopwords: null
2021-11-04 12:48:02,076 INFO [main] index.IndexCollection (IndexCollection.java:655) - Store positions? true
2021-11-04 12:48:02,077 INFO [main] index.IndexCollection (IndexCollection.java:656) - Store docvectors? true
2021-11-04 12:48:02,077 INFO [main] index.IndexCollection (IndexCollection.java:657) - Store document "contents" field? false
2021-11-04 12:48:02,077 INFO [main] index.IndexCollection (IndexCollection.java:658) - Store document "raw" field? true
2021-11-04 12:48:02,078 INFO [main] index.IndexCollection (IndexCollection.java:659) - Optimize (merge segments)? false
2021-11-04 12:48:02,082 INFO [main] index.IndexCollection (IndexCollection.java:660) - Whitelist: null
2021-11-04 12:48:02,084 INFO [main] index.IndexCollection (IndexCollection.java:661) - Pre-tokenized?: false
2021-11-04 12:48:02,084 INFO [main] index.IndexCollection (IndexCollection.java:681) - Directly building Lucene indexes...
2021-11-04 12:48:02,084 INFO [main] index.IndexCollection (IndexCollection.java:682) - Index path: indexes/lucene-index-msmarco-passage
2021-11-04 12:48:02,092 INFO [main] index.IndexCollection (IndexCollection.java:731) - ===== Indexing Collection =====
2021-11-04 12:48:02,377 INFO [main] index.IndexCollection (IndexCollection.java:829) - Thread pool with 9 threads initialized.
2021-11-04 12:48:02,380 INFO [main] index.IndexCollection (IndexCollection.java:831) - Initializing collection in data/msmarco_passage/collection_jsonl
2021-11-04 12:48:02,386 INFO [main] index.IndexCollection (IndexCollection.java:840) - 9 files found
2021-11-04 12:48:02,386 INFO [main] index.IndexCollection (IndexCollection.java:841) - Starting to index...
2021-11-04 12:49:02,422 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 530,000 documents indexed
2021-11-04 12:50:02,423 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 1,260,000 documents indexed
2021-11-04 12:51:02,427 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 2,020,000 documents indexed
2021-11-04 12:52:02,428 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 2,810,000 documents indexed
2021-11-04 12:53:02,429 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 3,610,000 documents indexed
2021-11-04 12:54:02,431 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 4,350,000 documents indexed
2021-11-04 12:55:02,435 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 5,020,000 documents indexed
2021-11-04 12:56:02,436 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 5,640,000 documents indexed
2021-11-04 12:57:02,437 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 6,240,000 documents indexed
2021-11-04 12:58:02,441 INFO [main] index.IndexCollection (IndexCollection.java:861) - 0.00% of files completed, 6,850,000 documents indexed
2021-11-04 12:58:53,367 DEBUG [pool-2-thread-2] index.IndexCollection$LocalIndexerThread (IndexCollection.java:248) - collection_jsonl/docs08.json: 841823 docs added.
2021-11-04 12:59:02,447 INFO [main] index.IndexCollection (IndexCollection.java:861) - 1.11% of files completed, 7,461,823 documents indexed
2021-11-04 13:00:02,448 INFO [main] index.IndexCollection (IndexCollection.java:861) - 1.11% of files completed, 8,071,823 documents indexed
2021-11-04 13:01:02,449 INFO [main] index.IndexCollection (IndexCollection.java:861) - 1.11% of files completed, 8,651,823 documents indexed
2021-11-04 13:01:03,015 DEBUG [pool-2-thread-8] index.IndexCollection$LocalIndexerThread (IndexCollection.java:248) - collection_jsonl/docs00.json: 1000000 docs added.
2021-11-04 13:01:04,593 DEBUG [pool-2-thread-11] index.IndexCollection$LocalIndexerThread
```

```

2021-11-04 13:01:01,935 DEBUG [pool-2-thread-1] index.IndexCollection$LocalIndexerThread
(IndexCollection.java:248) - collection_jsonl/docs07.json: 1000000 docs added.
2021-11-04 13:01:06,167 DEBUG [pool-2-thread-9] index.IndexCollection$LocalIndexerThread
(IndexCollection.java:248) - collection_jsonl/docs01.json: 1000000 docs added.
2021-11-04 13:01:09,858 DEBUG [pool-2-thread-3] index.IndexCollection$LocalIndexerThread
(IndexCollection.java:248) - collection_jsonl/docs02.json: 1000000 docs added.
2021-11-04 13:01:13,154 DEBUG [pool-2-thread-4] index.IndexCollection$LocalIndexerThread
(IndexCollection.java:248) - collection_jsonl/docs04.json: 1000000 docs added.
2021-11-04 13:01:13,393 DEBUG [pool-2-thread-6] index.IndexCollection$LocalIndexerThread
(IndexCollection.java:248) - collection_jsonl/docs03.json: 1000000 docs added.
2021-11-04 13:01:13,979 DEBUG [pool-2-thread-7] index.IndexCollection$LocalIndexerThread
(IndexCollection.java:248) - collection_jsonl/docs05.json: 1000000 docs added.
2021-11-04 13:01:16,520 DEBUG [pool-2-thread-5] index.IndexCollection$LocalIndexerThread
(IndexCollection.java:248) - collection_jsonl/docs06.json: 1000000 docs added.
2021-11-04 13:02:53,665 INFO [main] index.IndexCollection (IndexCollection.java:925) - I
ndexing Complete! 8,841,823 documents indexed
2021-11-04 13:02:53,665 INFO [main] index.IndexCollection (IndexCollection.java:926) - =
===== Final Counter Values =====
2021-11-04 13:02:53,666 INFO [main] index.IndexCollection (IndexCollection.java:927) - i
ndexed: 8,841,823
2021-11-04 13:02:53,667 INFO [main] index.IndexCollection (IndexCollection.java:928) - u
nindexable: 0
2021-11-04 13:02:53,668 INFO [main] index.IndexCollection (IndexCollection.java:929) - e
mpty: 0
2021-11-04 13:02:53,668 INFO [main] index.IndexCollection (IndexCollection.java:930) - s
kipped: 0
2021-11-04 13:02:53,668 INFO [main] index.IndexCollection (IndexCollection.java:931) - e
rrors: 0
2021-11-04 13:02:53,696 INFO [main] index.IndexCollection (IndexCollection.java:934) - T
otal 8,841,823 documents indexed in 00:14:51

```

In [8]:

```

from pyserini.index import IndexReader

index_reader = IndexReader('indexes/lucene-index-msmarco-passage')

```

2. Download and read the query file

You will rank MSMARCO passages for this set of queries.

In [9]:

```

[!]wget http://gem.cs.ru.nl/IR-Course-2021-2022/queries.txt

queries = dict()
with open("queries.txt", "r") as f:
    for line in f:
        cols = line.split("\t")
        queries[cols[0].strip()] = cols[1].strip()

# queries

--2021-11-04 13:02:56-- http://gem.cs.ru.nl/IR-Course-2021-2022/queries.txt
Resolving gem.cs.ru.nl (gem.cs.ru.nl)... 131.174.31.31
Connecting to gem.cs.ru.nl (gem.cs.ru.nl)|131.174.31.31|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 2275 (2.2K) [text/plain]
Saving to: 'queries.txt'

queries.txt      100%[=====>]    2.22K  --.-KB/s    in 0s

2021-11-04 13:02:57 (189 MB/s) - 'queries.txt' saved [2275/2275]

```

3. Implement the retrieval model

You will implement language model with Jelinek-Mercer (JM) smoothing:

$$score(q, d) =$$

$$\sum_{t \in q} \log \left(\left(1 - \lambda \right) \frac{c(t, d)}{|d|} + \lambda \frac{c(t, C)}{|C|} \right),$$

where $c(t, d)$ and $c(t, C)$ represent frequency of a term in a document and collection, respectively.

Notes about your implementation:

- Skip a term if it does not exist in the whole collection. This avoids $\log(0)$.
- Make sure to use the right form of a query (analyzed vs. not analyzed)
- Use natural logarithm

3.1. Obtain collection length

In this code, the global variable `len_C` denotes collection length.

In [10]:

```
global len_C
len_C = index_reader.stats()['total_terms']

print("Collection length =", len_C)
```

Collection length = 352316036

3.2. Obtain document length

In [11]:

```
def len_doc(d):
    len_d = sum(index_reader.get_document_vector(d).values())
    return len_d

doc = "2674124" # this is an example document
print("Length of document \""+doc+"\":", len_doc(doc))
```

Length of document "2674124": 31

3.3. Obtain collection frequency of a term

In [35]:

```
def coll_freq(term):
    df, cf = index_reader.get_term_counts(term)
    return cf

term = "record"
print("Frequency of term \""+term+"\" in the collection:", coll_freq(term))
```

Frequency of term "record" in the collection: 226439

3.4. Obtain term frequency

In [36]:

```
def term_freq(t, d):
    dv = index_reader.get_document_vector(d)
    if t in dv.keys():
        return dv[t]
    return 0
```

```
term = "record"
doc = "2674124"
print("Frequency of term \""+term+"\" in document \""+doc+"\":", term_freq(term, doc))
```

Frequency of term "record" in document "2674124": 2

3. Implement the retrieval model

You will implement language model with Jelinek-Mercer (JM) smoothing:

$$\text{score}(q, d) = \sum_{t \in q} \log \left((1 - \lambda) \frac{c(t, d)}{|d|} + \lambda \frac{c(t, C)}{|C|} \right),$$

where $c(t, d)$ and $c(t, C)$ represent frequency of a term in a document and collection, respectively.

Notes about your implementation:

- Skip a term if it does not exist in the whole collection. This avoids $\log(0)$.
- Make sure to use the right form of a query (analyzed vs. not analyzed)
- Use natural logarithm

3.5. Compute JM-smoothed probability for a single term

In [37]:

```
import math

def prob_t_Md(t, d, lambd):
    cf = coll_freq(t)
    if cf <= 0:
        return 0
    p_t_Md = ((1 - lambd) * term_freq(t, d) / len_doc(d) +
              lambd * cf / len_C)
    return p_t_Md

term = "record"
doc = "2674124"
print("p(t|Md) =", prob_t_Md(term, doc, 0.1))
```

p(t|Md) = 0.05812878768549357

3.6. Compute JM-smoothed probability for a query

In [38]:

```
import math

def score_doc(q, d, lambd):
    p_q_Md = 0
    for term in q:
        p_q_Md += math.log(prob_t_Md(term, d, lambd))
    return p_q_Md

query = index_reader.analyze(queries["23849"])
doc = "2674124"
print("p(q|Md) =", score_doc(query, doc, 0.1))
```

p(q|Md) = -12.763186444333824

4. Rank documents for the given queries

Ranking is done in two steps:

1. **First pass retrieval:** Use a fast ranker (i.e., Anserini SimpleSearcher) to rank all documents for a given query.
2. **Second pass retrieval:** Re-rank top-100 documents from the 1st pass retrieval using your retrieval model.
This is to make the ranking process efficient.

Notes:

- You need to change the default values of SimpleSearcher functions to obtain top-100 documents
- Set the value of lambda to 0.1
- Store your final ranking results in the `results` variable. Every item in the `results` list is a list containing queryID, documentID, and score. This is an example how the content of results should look like:

```
[[ '23849', '4348282', -10.65],  
 [ '23849', '7119957', -12.63],  
 [ '23849', '', -17.687729001682484],  
 ...]
```

In [39]:

```
from pyserini.search import SimpleSearcher  
  
results = []  
searcher = SimpleSearcher("indexes/lucene-index-msmarco-passage")  
lambda = 0.1  
  
for qID, q in zip(queries.keys(), queries.values()):  
    for d in searcher.search(q, 100):  
        results.append([qID, d.docid, score_doc(index_reader.analyze(queries[qID]), str(  
d.docid), lambda)])  
results
```

Out[39]:

```
[[ '1030303', '8726436', -13.36961560248476],  
 [ '1030303', '8726435', -15.502391047225595],  
 [ '1030303', '8726429', -15.604977448038678],  
 [ '1030303', '8726437', -15.851837894324593],  
 [ '1030303', '7156982', -15.91338110569287],  
 [ '1030303', '8726433', -16.0310618938083],  
 [ '1030303', '8726434', -16.08740354177648],  
 [ '1030303', '8726430', -16.395704265639907],  
 [ '1030303', '1305521', -27.64988446875305],  
 [ '1030303', '1305520', -27.67929834596989],  
 [ '1030303', '6222298', -27.945001427813594],  
 [ '1030303', '7284047', -24.083997913808915],  
 [ '1030303', '5285789', -23.558285653614412],  
 [ '1030303', '3302257', -23.341453320667405],  
 [ '1030303', '1305528', -27.664699550543464],  
 [ '1030303', '8230659', -24.129391038838868],  
 [ '1030303', '336399', -28.03015920419607],  
 [ '1030303', '309441', -27.377017546002076],  
 [ '1030303', '5824363', -24.53991606373053],  
 [ '1030303', '7508058', -24.7620958889916],  
 [ '1030303', '3302249', -24.818348396489505],  
 [ '1030303', '8726432', -22.98438075592842],  
 [ '1030303', '6764572', -28.80227394548036],  
 [ '1030303', '6271629', -28.415004841344015],  
 [ '1030303', '1905988', -29.218433905387936],  
 [ '1030303', '8441139', -29.218433905387936],  
 [ '1030303', '1492373', -28.532787805095303],  
 [ '1030303', '2699097', -29.33857805323605],  
 [ '1030303', '7168318', -28.54634076266955],  
 [ '1030303', '7284053', -23.74801135349862],  
 [ '1030303', '7970461', -28.60696534348747],  
 [ '1030303', '705402', -23.850415195198945],  
 [ '1030303', '1451846', -29.64495173938598],  
 [ '1030303', '3738183', -28.864794247797015],
```

['1030303', '2488127', -28.909245969369344],
['1030303', '6939422', -29.587793434875273],
['1030303', '6939426', -29.63096552474355],
['1030303', '3234728', -24.295789062616574],
['1030303', '3302258', -29.03184816846617],
['1030303', '4358004', -28.83426297197999],
['1030303', '4567726', -24.447584315454787],
['1030303', '3302252', -24.724917038351006],
['1030303', '4930087', -24.724917038351006],
['1030303', '7284050', -24.724917038351006],
['1030303', '7284048', -29.38466909210551],
['1030303', '3327029', -29.41133729818919],
['1030303', '4179453', -29.41133729818919],
['1030303', '6054027', -29.41133729818919],
['1030303', '7168317', -29.41133729818919],
['1030303', '5681381', -29.43731274359397],
['1030303', '2620119', -29.462630510579785],
['1030303', '3302255', -24.966909561390956],
['1030303', '2563266', -29.534951049163986],
['1030303', '3302251', -25.023069100558423],
['1030303', '848943', -25.023069100558423],
['1030303', '6939423', -29.580413341243805],
['1030303', '7168323', -29.55794052639022],
['1030303', '2563269', -29.60239220696411],
['1030303', '3302254', -29.60239220696411],
['1030303', '7168325', -29.60239220696411],
['1030303', '6108837', -25.233061604352756],
['1030303', '7284052', -25.282264780478428],
['1030303', '6939428', -29.744042437037916],
['1030303', '2615250', -25.377156879652734],
['1030303', '5086539', -25.330277867687474],
['1030303', '2444312', -29.762734529051613],
['1030303', '1815707', -29.834193329039962],
['1030303', '4966548', -29.834193329039962],
['1030303', '6939430', -29.851287721400816],
['1030303', '2001450', -28.372445250893584],
['1030303', '6989780', -29.884624059671523],
['1030303', '5681379', -25.596249251793903],
['1030303', '176183', -29.948137301400095],
['1030303', '3290826', -29.93263315586257],
['1030303', '4760232', -28.49504750108082],
['1030303', '8126314', -28.49504750108082],
['1030303', '7616223', -28.532787805095303],
['1030303', '2597061', -28.604246721140722],
['1030303', '4130105', -28.604246721140722],
['1030303', '5156631', -28.63814824884804],
['1030303', '8683568', -25.86724180142196],
['1030303', '2699093', -25.938314892098514],
['1030303', '3185530', -28.763311295928602],
['1030303', '3340457', -30.180468581795566],
['1030303', '1292823', -28.792298808833493],
['1030303', '1451848', -28.864794247797015],
['1030303', '2563261', -28.951805542789636],
['1030303', '4658938', -30.204566051377817],
['1030303', '1292822', -28.900512297600294],
['1030303', '2563265', -26.168708798469133],
['1030303', '1292821', -28.925830081616226],
['1030303', '3784253', -28.950522670238243],
['1030303', '8777713', -28.950522670238243],
['1030303', '7168316', -29.174948848112855],
['1030303', '2782702', -28.998150671290787],
['1030303', '5929352', -26.375034269855163],
['1030303', '4846462', -29.3001118270729],
['1030303', '5391149', -29.357270158915878],
['1030303', '1292819', -29.02114016554713],
['1030303', '5120735', -29.043612997430834],
['1037496', '4760912', -15.337810777776047],
['1037496', '7766587', -15.448944329257781],
['1037496', '5927420', -15.654240335810588],
['1037496', '7766585', -16.129549810862436],
['1037496', '4174377', -15.009341598249348],
['1037496', '3725936', -18.773960750746635],

['1037496', '4905511', -12.872790736790908],
['1037496', '3725937', -16.087282632674114],
['1037496', '4760914', -16.395563206005185],
['1037496', '4174376', -23.857785756860633],
['1037496', '4174378', -23.98116393516517],
['1037496', '2725825', -24.23247628872673],
['1037496', '7287404', -18.885587207071964],
['1037496', '3725935', -24.14546568975635],
['1037496', '7766583', -24.807833172146466],
['1037496', '4187304', -24.189917063317484],
['1037496', '3725940', -24.3866254125173],
['1037496', '4912163', -24.421716343319776],
['1037496', '7766584', -19.494990047687278],
['1037496', '3725932', -24.455617505986815],
['1037496', '4174380', -24.455617505986815],
['1037496', '3725931', -24.58077909290777],
['1037496', '4912166', -24.58077909290777],
['1037496', '8196479', -24.665335313912628],
['1037496', '4187303', -24.69200317198736],
['1037496', '4912160', -24.767987911943894],
['1037496', '553180', -20.113825304773886],
['1037496', '4912167', -24.83860431213782],
['1037496', '3356592', -20.130249220032518],
['1037496', '2725828', -20.189772169188],
['1037496', '4760906', -24.904561112910958],
['1037496', '3725938', -20.24756429898171],
['1037496', '3725934', -24.94623303129984],
['1037496', '3356595', -20.358338212910642],
['1037496', '3725933', -20.513712861910125],
['1037496', '2540328', -20.156728735383606],
['1037496', '2222120', -20.69869762259915],
['1037496', '365861', -21.13586719156924],
['1037496', '4760915', -25.213289981480543],
['1037496', '1297032', -21.046440196031114],
['1037496', '3275556', -25.302899806152425],
['1037496', '7286317', -19.69934037947567],
['1037496', '6914894', -20.645147429434743],
['1037496', '5033381', -21.23842433764468],
['1037496', '3562314', -20.209568989059058],
['1037496', '5786130', -20.209568989059058],
['1037496', '5507719', -20.270929178522742],
['1037496', '7021265', -20.55217123350225],
['1037496', '5788004', -21.041801187647557],
['1037496', '6595860', -21.109637163976572],
['1037496', '5540750', -20.37033832701762],
['1037496', '57410', -21.386209961576277],
['1037496', '2396654', -21.451889856405725],
['1037496', '365859', -21.48048564782412],
['1037496', '6292809', -21.508673267082045],
['1037496', '7407577', -20.514701685332636],
['1037496', '2124626', -20.60544116527926],
['1037496', '3852733', -21.4801697713338],
['1037496', '4905509', -21.54041666114528],
['1037496', '8131680', -21.176928358109308],
['1037496', '8831603', -21.252044114851397],
['1037496', '4187305', -24.69200317198736],
['1037496', '197695', -20.418690549360186],
['1037496', '519751', -20.418690549360186],
['1037496', '7779272', -20.615019768794554],
['1037496', '6101816', -21.005856714984],
['1037496', '3004783', -20.67637951051313],
['1037496', '5703441', -22.00140637286335],
['1037496', '5937726', -20.73590189440194],
['1037496', '929044', -20.73590189440194],
['1037496', '2902707', -24.883055296696064],
['1037496', '4187309', -24.883055296696064],
['1037496', '4760910', -24.883055296696064],
['1037496', '5937727', -21.590019667456534],
['1037496', '1896330', -20.79369345893018],
['1037496', '2221971', -20.79369345893018],
['1037496', '4881404', -20.79369345893018],
['1037496', '293184', -20.849851736825496],

['1037496', '4625262', -20.849851736825496],
['1037496', '6354370', -20.849851736825496],
['1037496', '6499341', -20.849851736825496],
['1037496', '2442404', -20.904466242331758],
['1037496', '8717869', -20.904466242331758],
['1037496', '1812968', -21.21238296754847],
['1037496', '7346511', -21.17941686041504],
['1037496', '880484', -21.276719723375464],
['1037496', '2646095', -20.957619326782147],
['1037496', '2716720', -20.957619326782147],
['1037496', '3157574', -20.957619326782147],
['1037496', '3562319', -20.957619326782147],
['1037496', '4150415', -20.957619326782147],
['1037496', '7392264', -20.957619326782147],
['1037496', '7948039', -20.957619326782147],
['1037496', '8805888', -20.957619326782147],
['1037496', '53465', -21.0093869230187],
['1037496', '2672141', -25.043394834726925],
['1037496', '1214193', -21.059839195549205],
['1037496', '1481537', -21.109041110411717],
['1037496', '2995695', -21.059839195549205],
['1037496', '5035625', -21.059839195549205],
['1043135', '650642', -17.18687410714317],
['1043135', '8696961', -22.00099448526447],
['1043135', '8696952', -21.778500753910663],
['1043135', '650641', -29.42977223944405],
['1043135', '8694703', -24.246231918646608],
['1043135', '2664994', -24.557367112422934],
['1043135', '8694697', -24.989918407033453],
['1043135', '8224813', -23.268164194917382],
['1043135', '8694701', -25.188286869552538],
['1043135', '7022077', -25.198691067516506],
['1043135', '4994428', -22.17286845298251],
['1043135', '2443317', -25.46779598193349],
['1043135', '2664995', -29.138273187975773],
['1043135', '650644', -28.802658369170462],
['1043135', '3378240', -23.835421071334757],
['1043135', '1845638', -28.227365878190206],
['1043135', '8696955', -27.048472501692185],
['1043135', '4721484', -29.32764795944131],
['1043135', '4355523', -22.599021794390797],
['1043135', '8696958', -22.734584837860147],
['1043135', '2822998', -24.101326072112478],
['1043135', '5262282', -29.303819811771778],
['1043135', '2664986', -29.631202631509606],
['1043135', '7945768', -28.975773778673254],
['1043135', '7022079', -29.405918105315546],
['1043135', '1845645', -28.082973358121595],
['1043135', '8696959', -28.082973358121595],
['1043135', '2664991', -29.530776372224523],
['1043135', '2664987', -29.209760348131447],
['1043135', '4541692', -29.209760348131447],
['1043135', '8787047', -28.545367394346382],
['1043135', '577276', -30.064034136046534],
['1043135', '8224806', -26.243186406343238],
['1043135', '7022076', -26.629633315223693],
['1043135', '1845636', -28.94590357054928],
['1043135', '1727793', -29.381201065889694],
['1043135', '5262285', -29.056097270026058],
['1043135', '1845643', -28.07519980380973],
['1043135', '3378235', -28.492999946335264],
['1043135', '7022075', -27.09235780168983],
['1043135', '4994429', -27.913065903295443],
['1043135', '7945773', -29.832136712815164],
['1043135', '2514458', -24.77702421430996],
['1043135', '3581772', -31.922854675996877],
['1043135', '650639', -25.0607394639633],
['1043135', '4541694', -29.826663294169144],
['1043135', '917852', -25.187529127516758],
['1043135', '1845642', -29.129214869951984],
['1043135', '1845644', -29.129214869951984],
['1043135', '3349026', -26.508468489030044],

['1043135', '7770351', -25.5453625741022],
['1043135', '5262290', -30.570100343280966],
['1043135', '1845639', -29.016081838740888],
['1043135', '7945770', -25.767232020602624],
['1043135', '8787048', -29.12928302322274],
['1043135', '7977270', -29.731093022436454],
['1043135', '4881713', -25.873700080034602],
['1043135', '8632895', -25.705925355861496],
['1043135', '2443321', -27.379727795603912],
['1043135', '2790624', -29.82062959055807],
['1043135', '2866084', -25.9773971301767],
['1043135', '2443319', -31.354161248475766],
['1043135', '7137851', -29.907569880133117],
['1043135', '8049995', -29.907569880133117],
['1043135', '6044059', -25.85117650204826],
['1043135', '8632897', -25.85117650204826],
['1043135', '6097574', -26.078463500056564],
['1043135', '1845641', -29.9920601902712],
['1043135', '2514951', -29.9920601902712],
['1043135', '8696956', -30.074234792356236],
['1043135', '1536028', -26.126709228754194],
['1043135', '3878608', -29.731176318173176],
['1043135', '7799671', -29.820715409943492],
['1043135', '8632894', -26.54857183317829],
['1043135', '2790617', -29.99215105687206],
['1043135', '5262288', -30.524947484564578],
['1043135', '8787049', -30.454378311117296],
['1043135', '1845640', -30.074328182523832],
['1043135', '3374872', -29.82087717915204],
['1043135', '650638', -30.59389372073802],
['1043135', '7799676', -30.23221978992106],
['1043135', '8696954', -30.308153417727446],
['1043135', '8696960', -30.308153417727446],
['1043135', '4881710', -27.047996933722697],
['1043135', '8694702', -27.047996933722697],
['1043135', '2443314', -31.98333432391911],
['1043135', '4708916', -30.916674125051152],
['1043135', '4875982', -30.525056015558917],
['1043135', '3209006', -29.34141092289755],
['1043135', '7770352', -27.201453166225747],
['1043135', '2280337', -29.46906689973508],
['1043135', '2996220', -27.349209495120338],
['1043135', '3378236', -31.118749092197728],
['1043135', '5751055', -31.21708102327872],
['1043135', '2514945', -29.82234740914968],
['1043135', '2955792', -29.82234740914968],
['1043135', '2514950', -29.9314274763781],
['1043135', '6781713', -29.9314274763781],
['1043135', '6781714', -29.9314274763781],
['1043135', '7770355', -29.9314274763781],
['1051399', '2845499', -23.739461957722252],
['1051399', '7848674', -23.758863312305245],
['1051399', '2866972', -24.386665681444175],
['1051399', '3446937', -23.782572355528416],
['1051399', '8712317', -25.253615554422982],
['1051399', '4681235', -24.194175662690093],
['1051399', '2768271', -24.48505464535203],
['1051399', '5041683', -24.716990822170242],
['1051399', '7848669', -25.05260909360851],
['1051399', '8818330', -26.0439440509407],
['1051399', '3773237', -28.05229139312238],
['1051399', '4328927', -26.42544034607142],
['1051399', '69813', -27.377022042091607],
['1051399', '69818', -27.558877472211247],
['1051399', '4840762', -26.126244865921226],
['1051399', '3434788', -25.585658985801256],
['1051399', '2953736', -26.01975525452439],
['1051399', '5284193', -26.622077130881834],
['1051399', '4622054', -26.062111903641007],
['1051399', '8229334', -25.437167562112446],
['1051399', '58138', -26.269435992841412],
['1051399', '5329142', -24.430296579059977],

['1051399', '4426188', -29.638005645266556],
['1051399', '4171290', -24.658723404792756],
['1051399', '5186393', -25.873383318356755],
['1051399', '3599718', -25.75638725750227],
['1051399', '3057965', -26.27869376494291],
['1051399', '2866978', -29.135940031234938],
['1051399', '4522483', -26.69954656405199],
['1051399', '3446936', -24.347382105695743],
['1051399', '4084314', -29.845022163579134],
['1051399', '8664506', -29.845022163579134],
['1051399', '2894408', -27.199373763151357],
['1051399', '5186392', -30.14523043963468],
['1051399', '3943228', -29.408853806904546],
['1051399', '4513312', -28.04592624071617],
['1051399', '8172956', -31.934791606018603],
['1051399', '3581849', -28.83890850110857],
['1051399', '69817', -27.427594932819588],
['1051399', '3378143', -31.87589846427605],
['1051399', '3773239', -32.51016832217201],
['1051399', '3969388', -30.785847495307472],
['1051399', '893262', -31.438785642443374],
['1051399', '4426187', -29.578647382120884],
['1051399', '893263', -31.596683971401877],
['1051399', '7813300', -32.844321356934515],
['1051399', '3969383', -26.98724212576725],
['1051399', '3588283', -31.31341425492955],
['1051399', '3107917', -27.11819547248219],
['1051399', '4729556', -26.501258688974705],
['1051399', '4513307', -29.04497286488948],
['1051399', '2764967', -27.24498435458184],
['1051399', '3599720', -27.24498435458184],
['1051399', '5056645', -33.40432799693404],
['1051399', '58139', -29.164798520395113],
['1051399', '2528523', -27.523217800202456],
['1051399', '4552288', -27.523217800202456],
['1051399', '2845500', -32.48778446694459],
['1051399', '2812921', -33.69106198215391],
['1051399', '3120981', -27.824683339018133],
['1051399', '7559979', -31.559732943825686],
['1051399', '5063127', -31.88956144476077],
['1051399', '4368659', -30.560056023527927],
['1051399', '2288256', -31.60071050550945],
['1051399', '7848670', -31.501024180610056],
['1051399', '8683249', -29.04574899665102],
['1051399', '4068420', -27.773178009742598],
['1051399', '5186386', -31.70111993520109],
['1051399', '4068418', -27.892379407371266],
['1051399', '2522098', -29.330762678188115],
['1051399', '8225042', -32.49706114434262],
['1051399', '2769054', -34.089563757294975],
['1051399', '7758785', -33.97923248514503],
['1051399', '2427011', -31.502794858567363],
['1051399', '2528525', -28.336073045069952],
['1051399', '4513313', -28.00811913338258],
['1051399', '3245659', -33.229976734710945],
['1051399', '3378142', -33.03500878906691],
['1051399', '7291530', -32.624697899758715],
['1051399', '8801997', -32.969719636827975],
['1051399', '5421904', -28.120592253243935],
['1051399', '3076345', -33.70420358871609],
['1051399', '2330719', -31.111180761133316],
['1051399', '8801999', -32.69684385981004],
['1051399', '5421901', -28.229977795455145],
['1051399', '5058868', -32.37941390655585],
['1051399', '6837220', -33.14793901621703],
['1051399', '4368654', -30.798837130860857],
['1051399', '4772458', -28.580022108552647],
['1051399', '2483093', -33.15325836373377],
['1051399', '4622485', -31.907797997224833],
['1051399', '7760091', -32.049602476219846],
['1051399', '3171080', -33.14219581490053],
['1051399', '1678171', -31.478915439200797],

['1051399', '4222649', -29.286728558877318],
['1051399', '8807183', -34.22200627706977],
['1051399', '8801994', -33.07687808583544],
['1051399', '6837221', -33.31008417044876],
['1051399', '58685', -33.88081957457164],
['1051399', '5150674', -33.51735314399898],
['1064670', '8680901', -15.901813370945721],
['1064670', '8680905', -15.733547697426363],
['1064670', '8680897', -16.614814926818884],
['1064670', '8680904', -15.505239156074884],
['1064670', '8680903', -28.67939664266608],
['1064670', '8680902', -23.682904260935885],
['1064670', '8337453', -31.707302354297514],
['1064670', '3807086', -36.50880047628972],
['1064670', '7734392', -37.472271138749655],
['1064670', '2998832', -34.155264046286646],
['1064670', '3277975', -35.528062850943435],
['1064670', '129462', -38.320162541363665],
['1064670', '8496831', -36.95673205393514],
['1064670', '402493', -37.35503638172448],
['1064670', '1946392', -36.507089567216134],
['1064670', '3077669', -36.57058340362874],
['1064670', '2977156', -38.31399594716337],
['1064670', '6878684', -36.80613523418614],
['1064670', '129468', -38.20327203928209],
['1064670', '6062648', -38.77259490644336],
['1064670', '129465', -39.89376281950868],
['1064670', '7556561', -38.59825458575035],
['1064670', '2183899', -38.53756553601821],
['1064670', '8734273', -38.75211403021596],
['1064670', '8030722', -36.03857104291345],
['1064670', '2389809', -36.17782169752546],
['1064670', '3555914', -37.42276132600323],
['1064670', '4853438', -36.642413580400756],
['1064670', '8533126', -34.501373377851365],
['1064670', '293152', -37.956763843672405],
['1064670', '7012470', -36.848467000776566],
['1064670', '8100020', -38.10189487501621],
['1064670', '6824284', -37.84463546095878],
['1064670', '8030718', -37.001140019331835],
['1064670', '736681', -38.50173659921724],
['1064670', '6062644', -41.10758500393567],
['1064670', '816346', -37.26372293457893],
['1064670', '8748215', -37.26372293457893],
['1064670', '110190', -35.975663873548704],
['1064670', '6824288', -38.347198457777516],
['1064670', '7522204', -35.07647485518207],
['1064670', '2977160', -38.40198817395747],
['1064670', '129459', -39.23027864693038],
['1064670', '2977159', -38.45531643598587],
['1064670', '7522208', -35.29835801733411],
['1064670', '6038320', -37.60740065308344],
['1064670', '2087457', -37.659347826704554],
['1064670', '4251134', -37.659347826704554],
['1064670', '6349303', -35.5578326729031],
['1064670', '55935', -37.709979643499],
['1064670', '6087393', -38.70250318814588],
['1064670', '6824287', -37.962313284676526],
['1064670', '3073057', -35.758549616741625],
['1064670', '2279163', -40.60783371819675],
['1064670', '2839843', -39.85070910736086],
['1064670', '5036872', -35.282586437460324],
['1064670', '4845469', -37.90058480652367],
['1064670', '6140058', -37.90058480652367],
['1064670', '1283724', -37.651586996628794],
['1064670', '5536005', -38.22932811409778],
['1064670', '4035283', -36.70192752402802],
['1064670', '4516313', -43.352927318988066],
['1064670', '4516312', -43.4582877549616],
['1064670', '3644829', -38.38938156009647],
['1064670', '5536007', -38.46484623174898],
['1064670', '5536001', -38.6077321918089],

['1064670', '2870990', -40.27863636327193],
['1064670', '3328187', -45.011152347672066],
['1064670', '8719108', -40.67529881904369],
['1064670', '1567821', -38.80456438738373],
['1064670', '3644827', -38.80456438738373],
['1064670', '5536002', -38.80456438738373],
['1064670', '8750942', -38.80456438738373],
['1064670', '378330', -38.3468301194749],
['1064670', '2902747', -37.915676593642864],
['1064670', '860041', -45.23886999250782],
['1064670', '6062642', -45.12868292938354],
['1064670', '976867', -40.340000512656545],
['1064670', '2759741', -38.03998685647696],
['1064670', '6986294', -37.80008978051497],
['1064670', '8680906', -39.04006652866305],
['1064670', '5089886', -37.07440269881104],
['1064670', '2905418', -39.09484849513244],
['1064670', '2560358', -40.146217748582316],
['1064670', '3413635', -38.586405926378404],
['1064670', '4660775', -40.04991728678403],
['1064670', '3328188', -45.144683202429405],
['1064670', '236964', -41.523795476181874],
['1064670', '5360417', -36.990109519066394],
['1064670', '5618788', -41.7734048903344],
['1064670', '7579792', -36.97791834749325],
['1064670', '3121310', -41.188952860582994],
['1064670', '3328192', -45.00158293251374],
['1064670', '429911', -39.20010399901137],
['1064670', '5159416', -37.057682537343965],
['1064670', '5874333', -39.30009287803109],
['1064670', '2049057', -41.42977986255955],
['1064670', '7899522', -41.63404206866498],
['1064670', '2310539', -41.997852347977044],
['1064670', '396429', -45.330085549237076],
['1071750', '3928913', -30.038930834977897],
['1071750', '8698274', -22.898580304750965],
['1071750', '8698268', -31.33233692470599],
['1071750', '8698276', -30.836836698859834],
['1071750', '8698269', -30.88370923302369],
['1071750', '3575934', -30.750381226611665],
['1071750', '8378390', -30.45315993930937],
['1071750', '8698273', -38.50744498790094],
['1071750', '8378385', -36.009944100286766],
['1071750', '8378389', -36.94565958712803],
['1071750', '8378386', -40.62627169520499],
['1071750', '27185', -39.24005824855831],
['1071750', '8378382', -43.16301794873607],
['1071750', '3016481', -39.80420052628254],
['1071750', '8378384', -43.634078635168194],
['1071750', '7018757', -43.954195875069026],
['1071750', '3016482', -43.85611201336018],
['1071750', '8698271', -43.85611201336018],
['1071750', '3086979', -44.2145020759563],
['1071750', '8378387', -43.22835126957563],
['1071750', '8698272', -41.46252251334863],
['1071750', '4133640', -44.74853002867994],
['1071750', '8378388', -43.86962344314892],
['1071750', '8378383', -46.782496417516086],
['1071750', '4717401', -47.18734802653364],
['1071750', '5510450', -44.42212637358192],
['1071750', '5510451', -44.907471381345644],
['1071750', '4133644', -48.07869721993279],
['1071750', '5510448', -45.236349961227624],
['1071750', '7018752', -45.44147393107864],
['1071750', '3719926', -51.232356134640995],
['1071750', '3928915', -46.915476451199254],
['1071750', '3928916', -46.35443983889974],
['1071750', '4094991', -51.39770095734484],
['1071750', '6508491', -48.487840467830736],
['1071750', '3928917', -46.07184451463229],
['1071750', '4094989', -48.00932028850734],
['1071750', '7018750', -46.71491982359704],

['1071750', '6508499', -52.750149426400625],
['1071750', '3928921', -52.18419977919603],
['1071750', '3016483', -55.12487106494463],
['1071750', '8698275', -53.817856505677504],
['1071750', '8698270', -48.30823449028813],
['1071750', '3795928', -49.92766067877611],
['1071750', '6335907', -51.26581529808824],
['1071750', '6330792', -52.64059075194065],
['1071750', '331170', -52.56514751278648],
['1071750', '5032111', -50.46870037762889],
['1071750', '2562017', -55.50004368582431],
['1071750', '8804338', -54.19439346313313],
['1071750', '6508493', -49.3825284664002],
['1071750', '4939145', -53.76967797057773],
['1071750', '4094990', -52.71918053578284],
['1071750', '7410371', -56.11310683945197],
['1071750', '7421125', -49.63526218729113],
['1071750', '3318540', -56.99725998200338],
['1071750', '962661', -56.98170110333541],
['1071750', '3214951', -49.73577395629657],
['1071750', '3993789', -55.51743370383557],
['1071750', '7778577', -50.02859864025358],
['1071750', '6983633', -49.719925876210745],
['1071750', '8017976', -59.37888394648701],
['1071750', '2012614', -53.72194434718866],
['1071750', '3928914', -52.61230262215064],
['1071750', '1628597', -49.946290614116876],
['1071750', '2850069', -52.54923790013609],
['1071750', '3930532', -55.485700771625424],
['1071750', '476679', -52.386461847613724],
['1071750', '90474', -54.62992925728229],
['1071750', '6733548', -53.73911154590763],
['1071750', '7850927', -55.963702600320055],
['1071750', '7141605', -53.95337419379093],
['1071750', '3850956', -50.62609300632718],
['1071750', '3042781', -54.0551373930807],
['1071750', '6138962', -53.127030615409396],
['1071750', '8837248', -54.24863073907461],
['1071750', '4360127', -59.317344337147816],
['1071750', '8017973', -59.37888394648701],
['1071750', '3335529', -55.28729675328502],
['1071750', '5047600', -54.43046215667206],
['1071750', '8017974', -59.496557530579025],
['1071750', '3719923', -59.96113400439165],
['1071750', '2840405', -57.05785265271382],
['1071750', '574320', -53.015522350649945],
['1071750', '8017972', -59.55289557655287],
['1071750', '3014192', -57.34777781191219],
['1071750', '8783990', -57.34777781191219],
['1071750', '7677637', -55.56658194989865],
['1071750', '4360128', -59.76360177614616],
['1071750', '3136366', -57.416000186720964],
['1071750', '8017968', -59.90823197563529],
['1071750', '6359062', -55.6388317614675],
['1071750', '7003298', -55.6388317614675],
['1071750', '208979', -57.67405510366912],
['1071750', '5377244', -57.67405510366912],
['1071750', '208982', -56.30966745241985],
['1071750', '7790571', -59.933337948648436],
['1071750', '8783991', -56.63155673105356],
['1071750', '3306573', -56.42657222178697],
['1071750', '6083597', -60.11527251445384],
['1105792', '8201611', -12.183648824591181],
['1105792', '8201618', -11.846094951628688],
['1105792', '8201616', -12.310400516301334],
['1105792', '8201612', -12.122348304856141],
['1105792', '8201613', -12.37939337943089],
['1105792', '8201615', -12.6670754100957],
['1105792', '6977017', -12.73161392009008],
['1105792', '8201610', -13.273211074376125],
['1105792', '8201614', -12.89021891962294],
['1105792', '8201617', -12.967179944044283],

['1105792', '7102720', -13.197703544154361],
['1105792', '771760', -13.455532569882532],
['1105792', '4545233', -13.499984315738581],
['1105792', '8201619', -13.137078939052712],
['1105792', '342876', -13.731785829507205],
['1105792', '7399820', -13.83022585217611],
['1105792', '2294057', -13.975407778446701],
['1105792', '5665835', -14.002076008814083],
['1105792', '2409768', -14.14867938271728],
['1105792', '8265458', -14.315733333662227],
['1105792', '8608603', -14.442026941983091],
['1105792', '6010402', -19.45254600263143],
['1105792', '3672136', -19.726936141457315],
['1105792', '5490903', -19.996208925659992],
['1105792', '3487762', -19.675652577489007],
['1105792', '4391107', -19.701623198670813],
['1105792', '6699898', -20.04894569253421],
['1105792', '1193877', -19.668106761921244],
['1105792', '4050698', -20.209241641368465],
['1105792', '6180434', -20.257428176590686],
['1105792', '775468', -20.280677255547996],
['1105792', '4651699', -19.726936141457315],
['1105792', '479601', -19.726936141457315],
['1105792', '1407489', -20.148828387830623],
['1105792', '475751', -19.759719477386746],
['1105792', '7615110', -19.791461688849846],
['1105792', '7389529', -20.06333053813293],
['1105792', '8658090', -20.196823013056683],
['1105792', '7459317', -19.852073337089557],
['1105792', '1193879', -19.90921877752109],
['1105792', '1376022', -20.178810067132815],
['1105792', '1376023', -20.160466488074338],
['1105792', '5490907', -20.178810067132815],
['1105792', '4391109', -19.96327302555126],
['1105792', '5971204', -20.45266558511505],
['1105792', '989490', -19.98924202539758],
['1105792', '7372393', -20.231903213803058],
['1105792', '1966229', -19.726936141457315],
['1105792', '8423372', -20.08685454915448],
['1105792', '3642904', -20.132303950580173],
['1105792', '8717896', -20.132303950580173],
['1105792', '1785018', -20.444571617671844],
['1105792', '5594852', -20.154276371036524],
['1105792', '6699897', -20.345808557398737],
['1105792', '2511437', -19.87532500987026],
['1105792', '4804671', -19.90921877752109],
['1105792', '5072015', -20.4686594409142],
['1105792', '2623416', -19.94200081637982],
['1105792', '7408031', -19.94200081637982],
['1105792', '4918035', -20.237632035346582],
['1105792', '1096440', -19.973741730790728],
['1105792', '4785106', -20.0045056056144],
['1105792', '3307879', -20.295881485440542],
['1105792', '2920212', -20.034350784981587],
['1105792', '6460339', -20.44804849227999],
['1105792', '6650998', -20.06333053813293],
['1105792', '2623408', -20.091493631438304],
['1105792', '7600283', -20.091493631438304],
['1105792', '8058698', -20.091493631438304],
['1105792', '2624570', -20.36861489627189],
['1105792', '2344062', -20.118884822025684],
['1105792', '4391103', -20.118884822025684],
['1105792', '919515', -20.118884822025684],
['1105792', '4391104', -19.621595086732654],
['1105792', '7549317', -20.145545285567692],
['1105792', '2148708', -20.17151298849138],
['1105792', '4935191', -20.17151298849138],
['1105792', '5460574', -20.17151298849138],
['1105792', '2335836', -20.22150784228865],
['1105792', '2669218', -20.196823013056683],
['1105792', '457729', -20.22150784228865],
['1105792', '7026606', -20.22150784228865],

['1105792', '7389527', -20.22150784228865],
['1105792', '6571367', -20.436411549832616],
['1105792', '1335682', -19.775716575302187],
['1105792', '766205', -19.775716575302187],
['1105792', '1869239', -20.26912032474283],
['1105792', '333470', -20.24559761056988],
['1105792', '5263288', -20.24559761056988],
['1105792', '5740463', -20.26912032474283],
['1105792', '5856095', -20.24559761056988],
['1105792', '6417724', -20.26912032474283],
['1105792', '7480463', -20.24559761056988],
['1105792', '7770851', -20.24559761056988],
['1105792', '7835708', -20.24559761056988],
['1105792', '5495682', -19.822226860707197],
['1105792', '1004084', -20.49989901344992],
['1105792', '2776428', -20.484401312377116],
['1105792', '4489529', -20.515160000159604],
['1105792', '4489534', -20.484401312377116],
['1106979', '8190137', -18.57368423836954],
['1106979', '8190140', -18.57368423836954],
['1106979', '7674419', -19.15596703005847],
['1106979', '8190134', -25.002891024352714],
['1106979', '348057', -25.361585969241155],
['1106979', '348054', -24.911851320643557],
['1106979', '7557632', -25.41071811255899],
['1106979', '799359', -25.73236440453072],
['1106979', '8190133', -26.49126725724565],
['1106979', '4219185', -20.46015280112794],
['1106979', '6221487', -26.7321642722356],
['1106979', '8190135', -19.98080348981657],
['1106979', '348055', -25.892643703486364],
['1106979', '3290646', -26.537963013222253],
['1106979', '5383176', -21.16440878492212],
['1106979', '3290651', -26.116111407029933],
['1106979', '1589841', -21.43113446002325],
['1106979', '2306620', -26.383145546085245],
['1106979', '7674423', -21.75214522276815],
['1106979', '8190136', -27.248749586585284],
['1106979', '6221486', -27.589397614374064],
['1106979', '6221490', -28.097656451604372],
['1106979', '3290649', -27.311573308963816],
['1106979', '3058146', -26.09336830096439],
['1106979', '2496608', -27.587797759278047],
['1106979', '8190138', -27.861419882078543],
['1106979', '6221491', -28.046794206638765],
['1106979', '2112598', -26.97917862196355],
['1106979', '357177', -27.367452180251796],
['1106979', '2884219', -28.45078315554236],
['1106979', '6221492', -27.425420842373526],
['1106979', '7674415', -27.5024036024845],
['1106979', '1207096', -27.454727987182565],
['1106979', '6221489', -27.97193698715582],
['1106979', '2375313', -33.80896493722357],
['1106979', '8190132', -27.941895495246676],
['1106979', '6221488', -28.047163798809734],
['1106979', '4085365', -34.26390992040396],
['1106979', '8177315', -27.212345768918638],
['1106979', '8190139', -28.14759930654678],
['1106979', '3996782', -28.53456820528139],
['1106979', '4624755', -27.65375203877834],
['1106979', '7920803', -34.634283485258926],
['1106979', '7920795', -34.866905591461475],
['1106979', '6221485', -28.52358922984085],
['1106979', '1089144', -28.57991189508855],
['1106979', '4779469', -28.57991189508855],
['1106979', '4624757', -29.006080026671903],
['1106979', '4085367', -34.55159182670661],
['1106979', '3417281', -34.585493355465175],
['1106979', '2789921', -34.650031830768505],
['1106979', '2375312', -35.27863984851179],
['1106979', '6221494', -28.79055657527885],
['1106979', '3293846', -28.862600857170854],

['1106979', '5784746', -28.862600857170854],
['1106979', '5920064', -26.18563158040202],
['1106979', '2375316', -34.82188195019247],
['1106979', '3996787', -28.870895753252388],
['1106979', '2496616', -28.888098730225288],
['1106979', '2375317', -34.99095811981661],
['1106979', '8196694', -29.180535495750398],
['1106979', '1308650', -29.294774433739498],
['1106979', '8585465', -31.077277492342414],
['1106979', '5488572', -27.013848957011398],
['1106979', '5920066', -27.115476084892975],
['1106979', '799365', -29.236652825410992],
['1106979', '7295948', -35.227346622875544],
['1106979', '4585827', -34.5342000954533],
['1106979', '8544395', -27.30893660079082],
['1106979', '4993951', -27.49065542775901],
['1106979', '4085368', -34.84785741367862],
['1106979', '2789920', -34.9570566025166],
['1106979', '947605', -29.69862821314426],
['1106979', '5733130', -32.01034668589431],
['1106979', '5908072', -32.01034668589431],
['1106979', '4254726', -32.078133732788444],
['1106979', '1039252', -35.08626819649394],
['1106979', '3553629', -32.8847132548497],
['1106979', '7706424', -28.123962645150403],
['1106979', '1655258', -33.52465645440882],
['1106979', '4624749', -32.730241246283654],
['1106979', '4624750', -32.85144678499773],
['1106979', '799360', -35.227346622875544],
['1106979', '6170459', -28.4610091290306],
['1106979', '6802961', -33.073810648590474],
['1106979', '7920802', -35.32742994392991],
['1106979', '1352235', -33.03843444399139],
['1106979', '8772719', -33.073819180273276],
['1106979', '4624754', -33.273890163356214],
['1106979', '8275110', -33.20939757256038],
['1106979', '8631020', -33.2419040474242],
['1106979', '2390743', -31.461998473540092],
['1106979', '7920797', -35.41840158463304],
['1106979', '6634295', -28.64640288742106],
['1106979', '6234047', -33.13905308010371],
['1106979', '5838049', -32.6125390940838],
['1106979', '6045982', -33.67457541416471],
['1106979', '4437575', -32.67406308509987],
['1106979', '6802958', -33.289998483930525],
['1106979', '3715369', -32.73374968526859],
['1108651', '8175412', -24.917853352791823],
['1108651', '8175407', -24.26545396176048],
['1108651', '6348741', -28.452046921874768],
['1108651', '6620938', -28.587078934192753],
['1108651', '1325612', -31.414489932568657],
['1108651', '6663873', -30.287468722358234],
['1108651', '1482137', -31.09019916151955],
['1108651', '2065792', -28.99220762318688],
['1108651', '1646383', -30.183532043779643],
['1108651', '6517914', -31.262766325849572],
['1108651', '3226775', -27.588176588429484],
['1108651', '3851277', -32.84392849930612],
['1108651', '6058576', -34.496353785209315],
['1108651', '4382120', -31.70379093150811],
['1108651', '1371583', -29.825198678318195],
['1108651', '7563462', -26.414383030711626],
['1108651', '4946657', -32.76476791568608],
['1108651', '6985088', -33.24641249931014],
['1108651', '4513249', -34.244108076965325],
['1108651', '488271', -34.16621330231418],
['1108651', '5027435', -29.02009707652473],
['1108651', '2475817', -32.13379913831775],
['1108651', '7038690', -30.923321043265037],
['1108651', '3459013', -35.02703013638022],
['1108651', '8175415', -35.17313497198818],
['1108651', '5621494', -31.677673566991196],

['1108651', '3727096', -33.33490016493763],
['1108651', '4125386', -32.3065045824921],
['1108651', '6782211', -31.80259652685507],
['1108651', '4617850', -33.5932273010712],
['1108651', '1400781', -32.61738053366643],
['1108651', '8266790', -33.57269245743241],
['1108651', '893423', -33.796592047917095],
['1108651', '8254975', -34.260180416885404],
['1108651', '3102791', -30.98629960472883],
['1108651', '6985085', -33.69262640040154],
['1108651', '8266789', -33.90951876964851],
['1108651', '5715055', -30.58734571776397],
['1108651', '8189032', -32.424466622603305],
['1108651', '8289561', -33.97261819589322],
['1108651', '1971162', -32.30435685571054],
['1108651', '6965', -33.575588459891485],
['1108651', '8576877', -36.611483039240305],
['1108651', '6410066', -31.88353132619484],
['1108651', '5997928', -34.04104517507877],
['1108651', '7321337', -32.662705620177505],
['1108651', '3136702', -33.078197625642304],
['1108651', '8512413', -33.419384895242146],
['1108651', '8205807', -33.573850386740126],
['1108651', '8512412', -33.631078269796475],
['1108651', '8006613', -32.09996435264465],
['1108651', '193384', -36.742876739873594],
['1108651', '4336299', -33.20941127630061],
['1108651', '4304301', -31.135640525643005],
['1108651', '4911228', -35.777478399262066],
['1108651', '8764446', -32.23020004513893],
['1108651', '8157105', -29.9108501042031],
['1108651', '918309', -29.9108501042031],
['1108651', '5397297', -31.486331684290185],
['1108651', '7910457', -32.588574906935584],
['1108651', '2512383', -33.58283918973201],
['1108651', '1457798', -37.220665519646815],
['1108651', '6266417', -33.0883713223739],
['1108651', '1325610', -34.09755439257215],
['1108651', '8202831', -33.61744692757866],
['1108651', '6623205', -36.573817657680735],
['1108651', '7729777', -31.283321446785514],
['1108651', '6410060', -31.754419846937154],
['1108651', '140506', -32.035087816129504],
['1108651', '6410062', -32.9974067519932],
['1108651', '1621557', -32.69807899022983],
['1108651', '3641700', -35.806459589275704],
['1108651', '8538781', -33.00125057288969],
['1108651', '7754981', -31.67643423576208],
['1108651', '140502', -31.285649648292928],
['1108651', '2745685', -36.152712588327475],
['1108651', '3013746', -32.92142203507791],
['1108651', '5141688', -31.624314359424055],
['1108651', '6410065', -32.01178681027552],
['1108651', '815597', -37.301185286540225],
['1108651', '5798501', -37.42814743201104],
['1108651', '6410061', -32.13447997277829],
['1108651', '2881395', -34.88713399665422],
['1108651', '3253130', -33.52055320049499],
['1108651', '3960291', -34.258211368422494],
['1108651', '1773494', -32.66009613843511],
['1108651', '7243391', -34.22290246697749],
['1108651', '7904315', -34.102345465171936],
['1108651', '2164085', -34.736117220920136],
['1108651', '3740104', -34.16705438044217],
['1108651', '1457501', -37.65476277387594],
['1108651', '5877822', -34.550890888437706],
['1108651', '4613064', -34.430161717524456],
['1108651', '3102792', -32.99553641014564],
['1108651', '2040434', -33.330882143205464],
['1108651', '1218618', -33.77467133891297],
['1108651', '8175408', -37.1482097946072],
['1108651', '5480469', -33.66628054600592],

['1108651', '6486048', -37.263018806416795],
['1108651', '3803874', -33.22224746630637],
['1109707', '4370753', -21.211338493065277],
['1109707', '1911564', -31.255064607303506],
['1109707', '8165491', -21.33175927463682],
['1109707', '771002', -28.86352716984195],
['1109707', '6872195', -26.90118231311617],
['1109707', '5947139', -31.048847923750508],
['1109707', '1143389', -26.504119267525592],
['1109707', '6238285', -26.504119267525592],
['1109707', '4122615', -31.20055464830697],
['1109707', '458599', -28.2671440885605],
['1109707', '4122618', -32.4515445313264],
['1109707', '1143386', -30.498939888475096],
['1109707', '2510185', -31.883980362197356],
['1109707', '2245437', -30.674284688424624],
['1109707', '3982977', -31.096781839735616],
['1109707', '1719374', -33.46548868592658],
['1109707', '4788762', -31.00246574201318],
['1109707', '1719371', -33.996941085405666],
['1109707', '6238290', -28.043855063687865],
['1109707', '5740304', -33.5385757511755],
['1109707', '5270487', -30.461937278568488],
['1109707', '4661916', -33.4056407489619],
['1109707', '5740302', -33.845059850048635],
['1109707', '1608567', -32.105863527278984],
['1109707', '2534438', -31.378103845091292],
['1109707', '6189800', -33.339298014848275],
['1109707', '6716572', -33.58234438698994],
['1109707', '771745', -30.816330135992633],
['1109707', '6733197', -34.971707774256025],
['1109707', '6716569', -34.584863945551376],
['1109707', '8008033', -32.63352522304887],
['1109707', '3037788', -30.811483791419604],
['1109707', '5121347', -32.855000922763274],
['1109707', '7241556', -31.32118870150879],
['1109707', '1340141', -38.7987988838373],
['1109707', '7126494', -33.63779149643954],
['1109707', '4236630', -33.77389226461635],
['1109707', '2247586', -36.086058540074326],
['1109707', '8441435', -39.03285796620938],
['1109707', '3037793', -38.67296865220359],
['1109707', '8746678', -35.71376837666553],
['1109707', '1789610', -38.17727366284258],
['1109707', '3325866', -34.231955198795795],
['1109707', '2031419', -33.677536001587086],
['1109707', '4236627', -33.739685265545674],
['1109707', '1823913', -38.12613730686325],
['1109707', '1143393', -38.70056346060897],
['1109707', '4867369', -37.698059639721954],
['1109707', '1667031', -39.31104820092822],
['1109707', '228686', -38.77932309630533],
['1109707', '3340791', -34.388522244572535],
['1109707', '3888607', -36.42480697782484],
['1109707', '7671120', -33.54744134305998],
['1109707', '1062367', -37.05229495186051],
['1109707', '2646538', -35.43375940416386],
['1109707', '7030351', -35.453301115482624],
['1109707', '5882871', -34.85190783096083],
['1109707', '7671117', -33.15821061279871],
['1109707', '459368', -34.84718681937199],
['1109707', '5826551', -35.764724477553784],
['1109707', '1951855', -36.418205772081194],
['1109707', '6939056', -38.50572914426136],
['1109707', '7159597', -38.93558598429734],
['1109707', '7566099', -34.71861798514549],
['1109707', '1117884', -38.810808396000105],
['1109707', '7566100', -34.83338281935637],
['1109707', '4285843', -39.232335534004925],
['1109707', '7671118', -33.80368694721679],
['1109707', '6752975', -38.77480748459492],
['1109707', '7241557', -33.935790257428984],

```
[ '1109707', '752145', -36.56547286193826],
[ '1109707', '353771', -40.45544680244038],
[ '1109707', '939951', -40.17984109146386],
[ '1109707', '4332695', -34.5061935532281],
[ '1109707', '1609738', -40.03435883662254],
[ '1109707', '3037794', -39.72738822417094],
[ '1109707', '1951853', -37.683763890225634],
[ '1109707', '3037786', -36.83250596618021],
[ '1109707', '8163716', -39.49066107370312],
[ '1109707', '1609739', -40.0145711450687],
[ '1109707', '5947140', -38.22738139572138],
[ '1109707', '2068094', -38.44739241735138],
[ '1109707', '5340019', -36.0364662382513],
[ '1109707', '4862297', -39.84569770538307],
[ '1109707', '1143388', -38.57159060776892],
[ '1109707', '458603', -36.71773701777966],
[ '1109707', '4059624', -38.67684001380163],
[ '1109707', '589762', -35.91789116059553],
[ '1109707', '2819919', -34.66068197883607],
[ '1109707', '1340134', -40.556056310062594],
[ '1109707', '7241559', -39.42681353118393],
[ '1109707', '8165492', -34.79381960069215],
[ '1109707', '2245436', -39.02897003691752],
[ '1109707', '8165494', -39.98404692052833],
[ '1109707', '1712222', -38.83153307510686],
[ '1109707', '4851004', -40.04892593103053],
[ '1109707', '4099609', -40.346070117267026],
[ '1109707', '1011113', -39.66810985043144],
[ '1109707', '2259393', -39.115862949679254],
[ '1109707', '4332698', -40.344220382065636],
...]
```

In [40]:

```
import numpy as np

fl_results = []
for result in np.array(results)[: ,2]:
    fl_results.append(float(result))
sum(fl_results)
```

Out[40]:

-160109.87527309445

Write your results into a file. Submit this file together with the completed notebook.

In [19]:

```
# check duplicates
check = set()
for res in results:
    if ((res[0], res[1])) in check:
        raise Exception("Error: Duplicate query-doc is found", res[0], res[1])
    check.add((res[0], res[1]))

# write results in a file
output_str = "\n".join([l[0] + "\tQ0\t" + l[1] + "\t0\t" + str(l[2]) + "\t1m_jm" for l
in results])
open("1m_jm.run", "w").write(output_str)
```

Out[19]:

246929

Handing in

Hand in both the result file and the filled-in notebook:

- The result file should be named `STUDENTNUMBER_FIRSTNAME_LASTNAME_1m_jm.run`

- **The notebook should be named** `STUDENTNUMBER_FIRSTNAME_LASTNAME_retrieval.ipynb`

In []: