**Instruction for compiling and running Homework 3**

* Files: Zip file comes with following files
  1. *InvertedIndex.py:* python file implemented on python 3, which is creates the inverted index for given corpus.
  2. *tccorpus.*txt: the corpus file provided for indexing and ranking.
  3. *index.out*: output generated after running InvertedIndex.py and represents the inverted index for given corpus and it serves as an input to BM25.py.
  4. *BM25.py*: python file implemented on python 3, which implements the BM25 algorithm and score the documents based on inverted index and query.
  5. *results.eval*: the output file generated after running BM25.py as requested in assignment.
  6. *queries.txt*: the query file required to rank documents using BM25 algorithm.
* Steps to run HW\_3 Assignment:

1. Run InvertedIndex.py (it generates a file index.out which is input to BM25.py)

>python3 InvertedIndex.py “tccorpus.txt”

1. Run BM25.py

>python3 BM25.py “index.out” “queries.txt” 100

* Pre-requisite: python version 3.x (preferable 3.5.0)
* Outputs:

1. InvertedIndex.py generates *index.out* as output which represents the inverted index for given corpus and it serves as an input to BM25.py.
2. BM25.py generates *results.eval* as output which contain top 100 ranked documents based on the BM25 score computed for each query in required format.

***References***:

* Python Tutorials: <http://www.tutorialspoint.com/python/>
* Python Docs: <https://docs.python.org/>
* Stackoverflow: <http://stackoverflow.com/>

Note:

1. I have provided the output files generated by InvertedIndex.py and BM25.py in this zip file for your reference but in actual these output files are generated at runtime as per given corpus input.