

Lab 2

Name: Srivenkatesh Nair

Student ID: 378000329

Username: sn6711

ISTE.612

- Task 1:** Complete the constructor of the class that builds the positional index.

```
review      0:<75>  1:<5>
even        0:<77,302,387>  1:<161> 2:<207,254,359,414>  3:<90,297>  4:<301,514,622,672>
harder      0:<78>
write       0:<81>  4:<438>
since       0:<82,664>
i           0:<83,229,257,317,365,401,499>  2:<427> 3:<293>
generally   0:<84>
applaud     0:<85>
films       0:<86,114>  2:<403> 4:<375>
attempt     0:<88>  2:<455> 3:<13>
break       0:<90>  3:<215>
mold        0:<92>  3:<219>
mess        0:<93>  2:<262> 3:<435,474>
with        0:<94,142,370,511,646>  1:<153> 2:<125,166,275> 3:<52,183,207,283,290,407>  4:<20>
your        0:<95,397>  1:<228>
head        0:<96>  1:<13>
such        0:<98>
lost        0:<99,694>
highway     0:<100,695>
memento     0:<101,698>
there       0:<103,179,182,190,198,202,205,212,425,485>  1:<79>
are         0:<104,139,180,183,191,199,203,206,213,390,426,436,520>  1:<198> 2:<60,222,364> 3:<304,379,400,404>4
:<67,222,229,717>
good        0:<105,522>  1:<209> 3:<439> 4:<725>
ways        0:<108>
making      0:<110>
all         0:<111,244,338,361,458,606>  1:<90> 2:<264,484> 3:<331> 4:<220,361,444>
types       0:<112>
these       0:<116>  1:<150> 2:<4>  4:<235,357,402>
folks       0:<117>
just        0:<118,418,530,629>  1:<120,146>  3:<210> 4:<160,581>
didn't      0:<119,356>
snag        0:<120>
correctly   0:<123>
seem        0:<125>  2:<70>
have        0:<127,173>  2:<191> 4:<287>
taken       0:<128>
pretty      0:<130,489,521,591,600>  1:<232> 3:<85>
neat        0:<131>
concept     0:<132>
executed    0:<134>
terribly    0:<136>
```

- Task 2:** Complete implementing the intersect method that takes in two postings and output a merged postings list.

```

Part 2 => Testing the intersect Method:
this : [0:<74,121,129,163,268,275,323,373,471,619>, 1:<16,27>, 2:<260,437,481>, 3:<75,112,434,452,473>, 4:<248,267,373,394,428,715>]
is : [0:<71,149,224,247,267,312,374,384,417,620,637,643>, 1:<139,158,182>, 2:<1,113,118,137,179,203,237,291,310,322,342,348,384,432,439,445,450>, 3:<4,84,143,233,337,462>, 4:<50,146,159,185,249,261,268,300,388,421,432,505,618,626,640,675,682,706>]
Document 0 : [374, 620]
Document 4 : [249]

```

•**Task 3:** Implement a phraseQuery method that takes in a phrase query with multiple terms and return a list of DocId objects (use the DocId class from the Positional Index coding lecture).

```

Part 3 => Testing the phraseQuery Method:
Search Query :
that : [0:<55,150,218,313,343,362,375,381,504,539,610,633,640>, 1:<7,38,46,121>, 2:<5,111,131,176,288,332,381,398,426,436,447,473>, 3:<255,309,333,444>, 4:<92,157,184,219,247,259,445,686>]
is : [0:<71,149,224,247,267,312,374,384,417,620,637,643>, 1:<139,158,182>, 2:<1,113,118,137,179,203,237,291,310,322,342,348,384,432,439,445,450>, 3:<4,84,143,233,337,462>, 4:<50,146,159,185,249,261,268,300,388,421,432,505,618,626,640,675,682,706>]
what : [0:<72,138>, 1:<115>, 4:<186,209>]
it : [0:<64,135,155,223,245,281,287,330,354,386,424,443,479,483,506,580,585,589,613,649,667>, 1:<51>, 2:<0,69,100,107,112,167,172,286,330,444>, 3:<65,176,343>, 4:<187,285,313,418,462,557,600,674,681,687,705>]
sets : [4:<188>]

Result:
that : 4:<184>
is : 4:<185>
what : 4:<186>
it : 4:<187>
sets : 4:<188>

```

•**Task 4:** Design and test phrase queries with 2, 3, 4, and 5 terms. Use the document collection from Lab 1 to design and test the queries.

```

Part 4 => Testing Phrase Queries:

Search Query 1 :
this : [0:<74,121,129,163,268,275,323,373,471,619>, 1:<16,27>, 2:<260,437,481>, 3:<75,112,434,452,473>, 4:<248,267,373,394,428,715>]
is : [0:<71,149,224,247,267,312,374,384,417,620,637,643>, 1:<139,158,182>, 2:<1,113,118,137,179,203,237,291,310,322,342,348,384,432,439,445,450>, 3:<4,84,143,233,337,462>, 4:<50,146,159,185,249,261,268,300,388,421,432,505,618,626,640,675,682,706>]

Result:
this : 0:<373>
is : 0:<374>
this : 0:<619>
is : 0:<620>
this : 4:<248>
is : 4:<249>

```

Search Query 2 :

middle : [1:<106>]

of : [0:<19,109,113,209,215,222,260,329,339,351,399,459,595,607,674,706>, 1:<35,107,114,131,165,175,192,237>, 2:<15,37,44,64,159,187,338,385,401,424,442,468,478>, 3:<73,103,119,131,186,250,302,346,352,433,466>, 4:<37,43,58,226,239,252,264,271,479,511,592,613,631,657,663>]

nowhere : [1:<108>]

Result:

middle : 1:<106>

of : 1:<107>

nowhere : 1:<108>

Search Query 3 :

signs : [3:<345>]

of : [0:<19,109,113,209,215,222,260,329,339,351,399,459,595,607,674,706>, 1:<35,107,114,131,165,175,192,237>, 2:<15,37,44,64,159,187,338,385,401,424,442,468,478>, 3:<73,103,119,131,186,250,302,346,352,433,466>, 4:<37,43,58,226,239,252,264,271,479,511,592,613,631,657,663>]

a : [0:<6,47,58,66,207,236,264,348,488,508,547,599,622,672>, 1:<12,30,33,36,41,48,73,86,122,172,190,234,238>, 2:<7,9,6,120,141,150,163,182,377,418,433>, 3:<132,178,195,225,313,328,347,376,385,412,476>, 4:<1,8,11,24,41,54,68,199,212,227,237,323,377,392,456,493,524,568,702,724>]

pulse : [3:<348>]

Result:

signs : 3:<345>

of : 3:<346>

a : 3:<347>

pulse : 3:<348>

Search Query 4 :

sounds : [2:<94>]

like : [0:<195,322,372>, 1:<85,189,212>, 2:<3,71,93,95>, 3:<414>, 4:<401>]

a : [0:<6,47,58,66,207,236,264,348,488,508,547,599,622,672>, 1:<12,30,33,36,41,48,73,86,122,172,190,234,238>, 2:<7,9,6,120,141,150,163,182,377,418,433>, 3:<132,178,195,225,313,328,347,376,385,412,476>, 4:<1,8,11,24,41,54,68,199,212,227,237,323,377,392,456,493,524,568,702,724>]

cool : [0:<60,601>, 2:<77,97,279>]

movie : [0:<41,50,144,416,466,494>, 1:<4,17,87,229,239>, 2:<9,98>, 4:<277,304>]

Result:

sounds : 2:<94>

like : 2:<95>

a : 2:<96>

cool : 2:<97>

movie : 2:<98>