**CS/SE 6301 Software Analysis and Comprehension**

**Spring 2016**

**Assignment #3 – March 8, 2016**

**Team 2**

**Ronaldo Goncalves Junior**

**Sungsoo Ahn**

**HADOOP**

**DERBY**

# Total number of bug reports

HADOOP: 7320

DERBY: 4017

# Average and median number of sentences, terms, nouns, verbs, adjectives, and adverbs that compose a bug report

Table 1: Bug report details for Title-Only input.

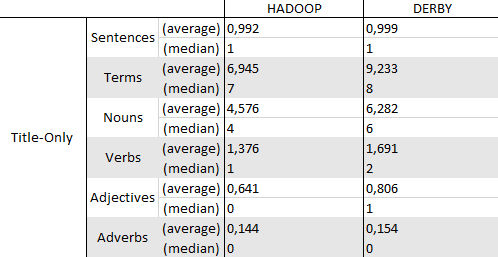


Table 2: Bug report details for Description-Only input.

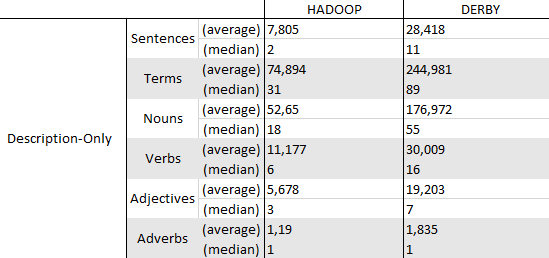
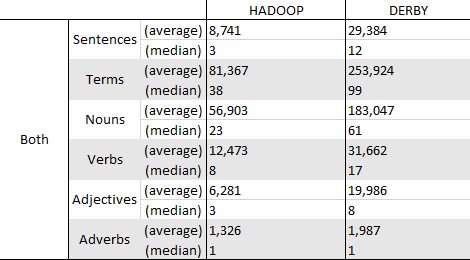


Table 3: Bug report details for both title and description.



# Top-10 most frequent nouns, verbs, adjectives and adverbs for each system and corpus

## DERBY Noun

1. derby=1590
2. org=1313
3. apache=1270
4. java-=948
5. java=913
6. impl=848
7. Statement=372
8. Connection=359
9. sql=328
10. Derby=271

## DERBY Verb

1. create=228,
2. monitor=225,
3. execute=224,
4. Set=148,
5. set=127,
6. updated=123,
7. inserted=113,
8. deleted=108,
9. Name=95,
10. connect=90

## DERBY Adjective

1. select=149,
2. main=120,
3. raw=92,
4. net=73,
5. Basic=61,
6. http=58,
7. embedded=54,
8. true=51,
9. Found=36,
10. Unknown-=35

## DERBY Adverb

1. org=24,
2. incorrectly=19,
3. correctly=18,
4. However=17,
5. close=17,
6. org-=16,
7. Also=16,
8. not=16,
9. Here=16,
10. currently-=12,

## HADOOP Noun

1. hadoop=16030,
2. org=15081,
3. apache=13434,
4. java-=11179,
5. java=9443,
6. INFO=3851,
7. Task=2922,
8. Job=2813,
9. Test=2795,
10. Method=2749,

## HADOOP Verb

1. mapred=3274,
2. run=1907,
3. Name=1431,
4. reflect=1366,
5. invoke=1311,
6. create=879,
7. fails=855,
8. build=764,
9. set=762,
10. read=754,

## HADOOP Adjective

1. Native=1639,
2. main=869,
3. hadoop=863,
4. common=742,
5. Local=623,
6. common-=566,
7. net=543,
8. local=455,
9. Unknown=429,
10. key=413,

## HADOOP Adverb

1. org=551,
2. n't=538,
3. n't-=442,
4. Not=334,
5. org-=197,
6. Currently=178,
7. close=173,
8. Here=168,
9. currently=154,
10. However-=152,