PA3 Report

Page Rank and Indexing

Alex Huynh

Com S 435
Iowa State University

1.

```
Epsilon = 0.001, Beta = 0.85 -> Iterations: 14
[/wiki/Wayback_Machine, /wiki/United_States, /wiki/Association_football, /wiki/World_War_II, /wiki/France, /wiki/
National_Diet_Library, /wiki/Basketball, /wiki/Canada, /wiki/Japan, /wiki/The_New_York_Times]


Epsilon = 1.0E-5, Beta = 0.85 -> Iterations: 29
[/wiki/Wayback_Machine, /wiki/United_States, /wiki/Association_football, /wiki/World_War_II, /wiki/France, /wiki/
National_Diet_Library, /wiki/Basketball, /wiki/Canada, /wiki/Japan, /wiki/The_New_York_Times]


Epsilon = 0.001, Beta = 0.5 -> Iterations: 6
[/wiki/Wayback_Machine, /wiki/United_States, /wiki/Association_football, /wiki/Basketball, /wiki/American_footbal
l, /wiki/National_Football_League, /wiki/Ice_hockey, /wiki/World_War_II, /wiki/Baseball, /wiki/France]


Epsilon = 1.0E-5, Beta = 0.5 -> Iterations: 11
[/wiki/Wayback_Machine, /wiki/United_States, /wiki/Association_football, /wiki/Basketball, /wiki/American_footbal
l, /wiki/National_Football_League, /wiki/Ice_hockey, /wiki/World_War_II, /wiki/Baseball, /wiki/France]
```

Figure 1: Page Rank Algorithm Iterations via tuning Epsilon, Beta

2.

```
Query: left-handed

1. Handedness.txt: TPS Score: 0.0, VSS Score: 5.11649081556232, Relevance Score: 2.046596326224928

2. Left-handed_specialist.txt: TPS Score: 0.0, VSS Score: 4.812263980673482, Relevance Score: 1.9249055922693927

3. Platoon_system.txt: TPS Score: 0.0, VSS Score: 4.487459300499695, Relevance Score: 1.794983720199878

4. Lefty-righty_switch.txt: TPS Score: 0.0, VSS Score: 4.137233052124938, Relevance Score: 1.6548932208499754

5. Madison_Bumgarner.txt: TPS Score: 0.0, VSS Score: 4.0137056168723895, Relevance Score: 1.6054822467489558

6. Switch_hitter.txt: TPS Score: 0.0, VSS Score: 4.0137056168723895, Relevance Score: 1.6054822467489558

7. Comparison_of_cricket_and_baseball.txt: TPS Score: 0.0, VSS Score: 3.754477817426659, Relevance Score: 1.5017911269706636

8. Comparison_between_cricket_and_baseball.txt: TPS Score: 0.0, VSS Score: 3.754477817426659, Relevance Score: 1.5017911269706636

9. Comparison_of_baseball_and_cricket.txt: TPS Score: 0.0, VSS Score: 3.754477817426659, Relevance Score: 1.5017911269706636

10. Javier_L%C3%B3pez_(baseball).txt: TPS Score: 0.0, VSS Score: 3.4759710275237805, Relevance Score: 1.3903884110095124
```

Figure 2: Query: "left-handed"

```
Query: delicious food

1. Portland,_Oregon.txt: TPS Score: 0.2, VSS Score: 3.5195522296688786, Relevance Score: 1.5278208918675515

2. Korea.txt: TPS Score: 0.11764705882352941, VSS Score: 2.9063565357979364, Relevance Score: 1.2331308496132922

3. Salt_Lake_City.txt: TPS Score: 0.11764705882352941, VSS Score: 2.6015689999832023, Relevance Score: 1.1112158352873986

4. Concession_stand.txt: TPS Score: 0.11764705882352941, VSS Score: 2.6015689999832023, Relevance Score: 1.1112158352873986

5. Herbert_Hoover.txt: TPS Score: 0.11764705882352941, VSS Score: 2.2943681975101846, Relevance Score: 0.9883355142981916

6. United_States.txt: TPS Score: 0.11764705882352941, VSS Score: 2.2943681975101846, Relevance Score: 0.9883355142981916

7. Cuba.txt: TPS Score: 0.11764705882352941, VSS Score: 2.0337400396315486, Relevance Score: 0.884084251146737

8. Prisoner_of_war.txt: TPS Score: 0.11764705882352941, VSS Score: 2.0337400396315486, Relevance Score: 0.884084251146737

9. Hunting.txt: TPS Score: 0.11764705882352941, VSS Score: 1.93909504403953, Relevance Score: 0.8462262529099297

10. Minneapolis.txt: TPS Score: 0.11764705882352941, VSS Score: 1.8395870816128275, Relevance Score: 0.8064230679392487
```

Figure 3: Query: "delicious food"

```
Query: many points scored

1. BMX_racing.txt: TPS Score: 0.16666666666666666, VSS Score: 3.6460258235354073, Relevance Score: 1.5584103294141631

2. New_York_Islanders.txt: TPS Score: 0.11538461538461539, VSS Score: 2.9713370788871583, Relevance Score: 1.2577656007856326

3. Wayne_Gretzky.txt: TPS Score: 0.15, VSS Score: 2.86093776442955, Relevance Score: 1.2343751057718202

4. Michael_Jordan.txt: TPS Score: 0.08823529411764706, VSS Score: 2.7559941992303045, Relevance Score: 1.15533885616271

5. Los_Angeles_Lakers.txt: TPS Score: 0.16666666666666666, VSS Score: 2.5931545137287224, Relevance Score: 1.1372618054914891

6. Playoffs.txt: TPS Score: 0.13636363636363635, VSS Score: 2.4516021590492403, Relevance Score: 1.062459045437878

7. New_Jersey_Devils.txt: TPS Score: 0.08823529411764706, VSS Score: 2.440184643209675, Relevance Score: 1.0290150337544584

8. Curling.txt: TPS Score: 0.16666666666666666, VSS Score: 2.3203632120695077, Relevance Score: 1.0281452848278032

9. Anaheim_Ducks.txt: TPS Score: 0.10714285714285714, VSS Score: 2.341176522791557, Relevance Score: 1.0007563234023371

10. American_football.txt: TPS Score: 0.16666666666666666, VSS Score: 2.2362778231704006, Relevance Score: 0.9945111292681602
```

Figure 4: Query: "many points scored"

```
Query: rain and windy weather

1. Candlestick_Park.txt: TPS Score: 0.11428571428571428, VSS Score: 2.2096146296033927, Relevance Score: 0.9524172804127857

2. Cricket_pitch.txt: TPS Score: 0.21052631578947367, VSS Score: 1.9008549170331546, Relevance Score: 0.886657756286946

3. Montreal.txt: TPS Score: 0.36363636363636365, VSS Score: 1.6502035578345848, Relevance Score: 0.8782632413156521

4. Seattle.txt: TPS Score: 0.11428571428571428, VSS Score: 1.9147319442802841, Relevance Score: 0.8344642062835423

5. WLS-TV.txt: TPS Score: 0.1, VSS Score: 1.6619028270943454, Relevance Score: 0.7247611308377382

6. 2008_World_Series.txt: TPS Score: 0.09302325581395349, VSS Score: 1.6203189746192925, Relevance Score: 0.7039415433360892

7. WGN-TV.txt: TPS Score: 0.0784313725490196, VSS Score: 1.4909249573929024, Relevance Score: 0.6434288064865727

8. San_Francisco.txt: TPS Score: 0.11428571428571428, VSS Score: 1.4258024976367585, Relevance Score: 0.6388924276261321

9. 1886_World_Series.txt: TPS Score: 0.0975609756097561, VSS Score: 1.450263620499712, Relevance Score: 0.6386420335657383

10. 2009_American_League_Championship_Series.txt: TPS Score: 0.11428571428571428, VSS Score: 1.4179682898232957, Relevance Score: 0.6357587445007469
```

Figure 5: Query: "rain and windy weather"

```
Query: most boring match of all time

1. David_Beckham.txt: TPS Score: 0.13953488372093023, VSS Score: 2.061238579830374, Relevance Score: 0.9082163621647077

2. Cricket.txt: TPS Score: 0.16216216216216217, VSS Score: 1.8080767258154404, Relevance Score: 0.8205279876234736

3. Melbourne_Cricket_Ground.txt: TPS Score: 0.10909090909090909, VSS Score: 1.733787698224698, Relevance Score: 0.7589696247444246

4. Match_fixing.txt: TPS Score: 0.10714285714285714, VSS Score: 1.661839144990499, Relevance Score: 0.7290213722819139

5. Pel%C3%A9.txt: TPS Score: 0.16216216216216217, VSS Score: 1.5160145165158776, Relevance Score: 0.7037031039036484

6. Tom_Wills.txt: TPS Score: 0.11320754716981132, VSS Score: 1.5530946630124778, Relevance Score: 0.6891623935068779

7. Curling.txt: TPS Score: 0.15789473684210525, VSS Score: 1.482562485967686, Relevance Score: 0.6877618364923376

8. Don_Bradman.txt: TPS Score: 0.16216216216216217, VSS Score: 1.4668559652442656, Relevance Score: 0.6840396833950035

9. Playoff_format.txt: TPS Score: 0.1016949152542373, VSS Score: 1.5422547404464453, Relevance Score: 0.6779188453311206

10. Sudden_death_(sport).txt: TPS Score: 0.12, VSS Score: 1.45367831505556, Relevance Score: 0.653471326022224
```

Figure 6: Query: "most boring match of all time"

Overall, these results are reasonable. For "left-handed", we got results related to the query. For "delicious food", we got some places with good food (I can testify for Korean food). For "many points scored", we got players and teams that scored lots of points. For "rain and windy weather", we got places or events that had that type of weather. For "most boring match of all time", this was more of a struggle because the preposition "of" makes it difficult. Specifically, "of all time" made David_Beckham.txt rank first. Boring is the most significant part of this query, but we do not have a weighing process, hence the results returned are poor.