Spoon River Anthology

Written By Edgar Lee Masters

A Project on Natural Language Processing

Spoon River?

1900's Masterpiece

Epitaphs of dead citizens of Spoon River residents

De-mystify rural & small town American life

212 characters

244 accounts of their lives, losses, and manner of death

Cross-references between characters

Content

- Histories of residents and turning points throughout their life
- Observations of life from outside
- Treatment of graves
- How they really died

Overall Goal

1. Determine similarities between poems based on cross-referenced terms and synonyms.
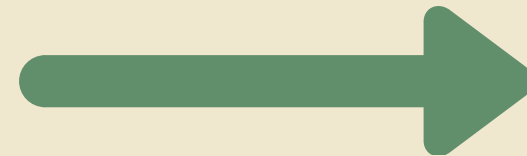2. Cluster poems based on context such as love, loss, death, guilt, etc..
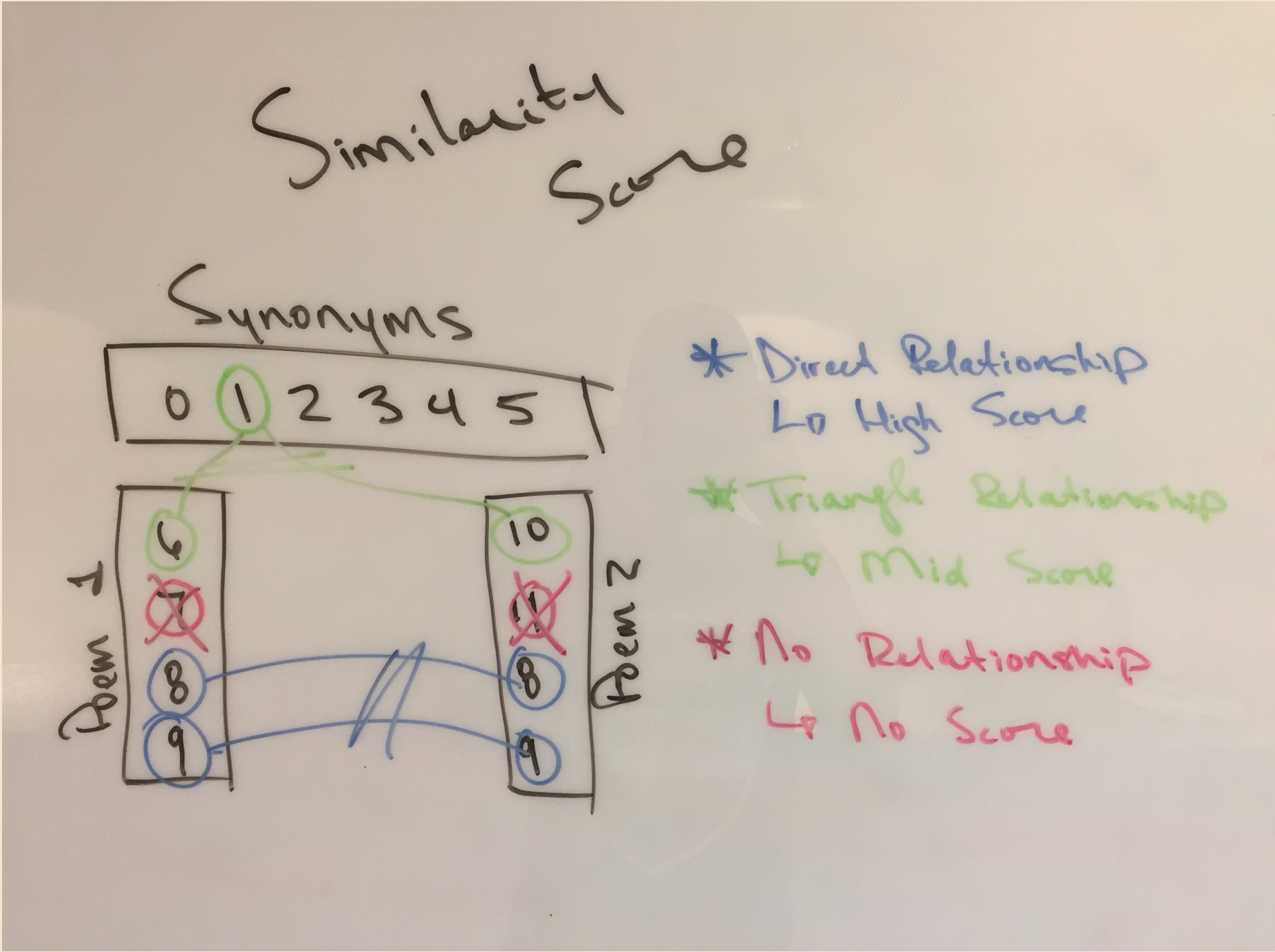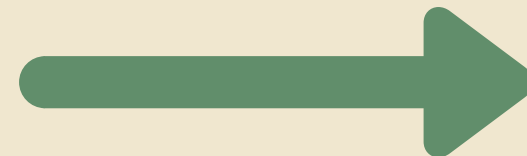
MVP
Final Outlook

Plan of Attack

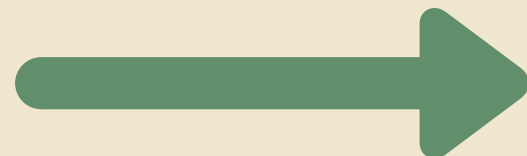*Compare two poems, resulting in a numeric score.*

$P_1 T_1 == P_2 T_1 \longrightarrow$ High

$P_1 T_1 * S_1 * P_2 T_1 \longrightarrow$ Mid

$P_1 T_1 !! P_2 T_1 \longrightarrow$ Low
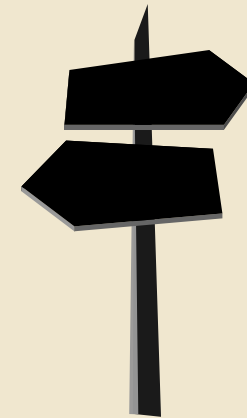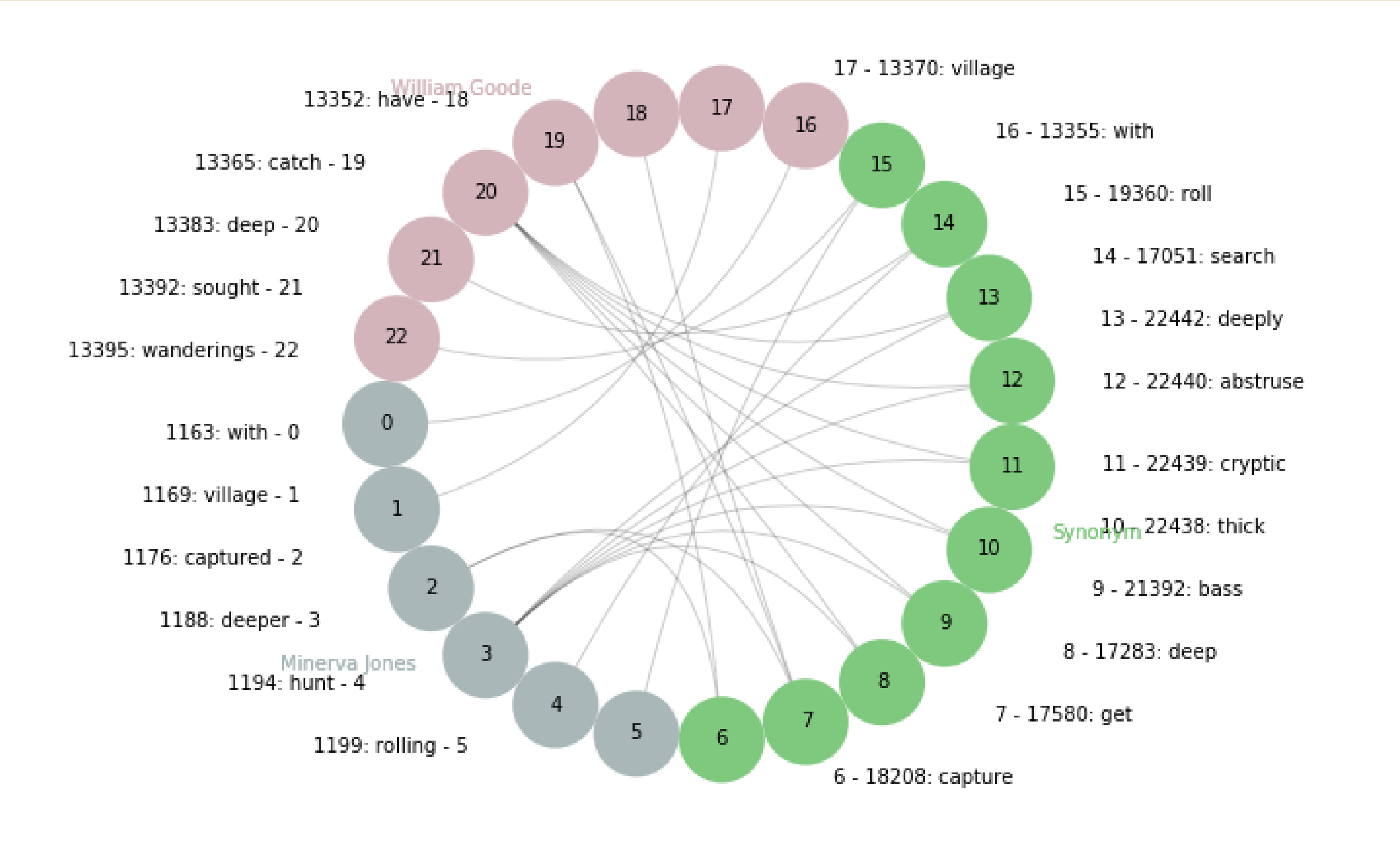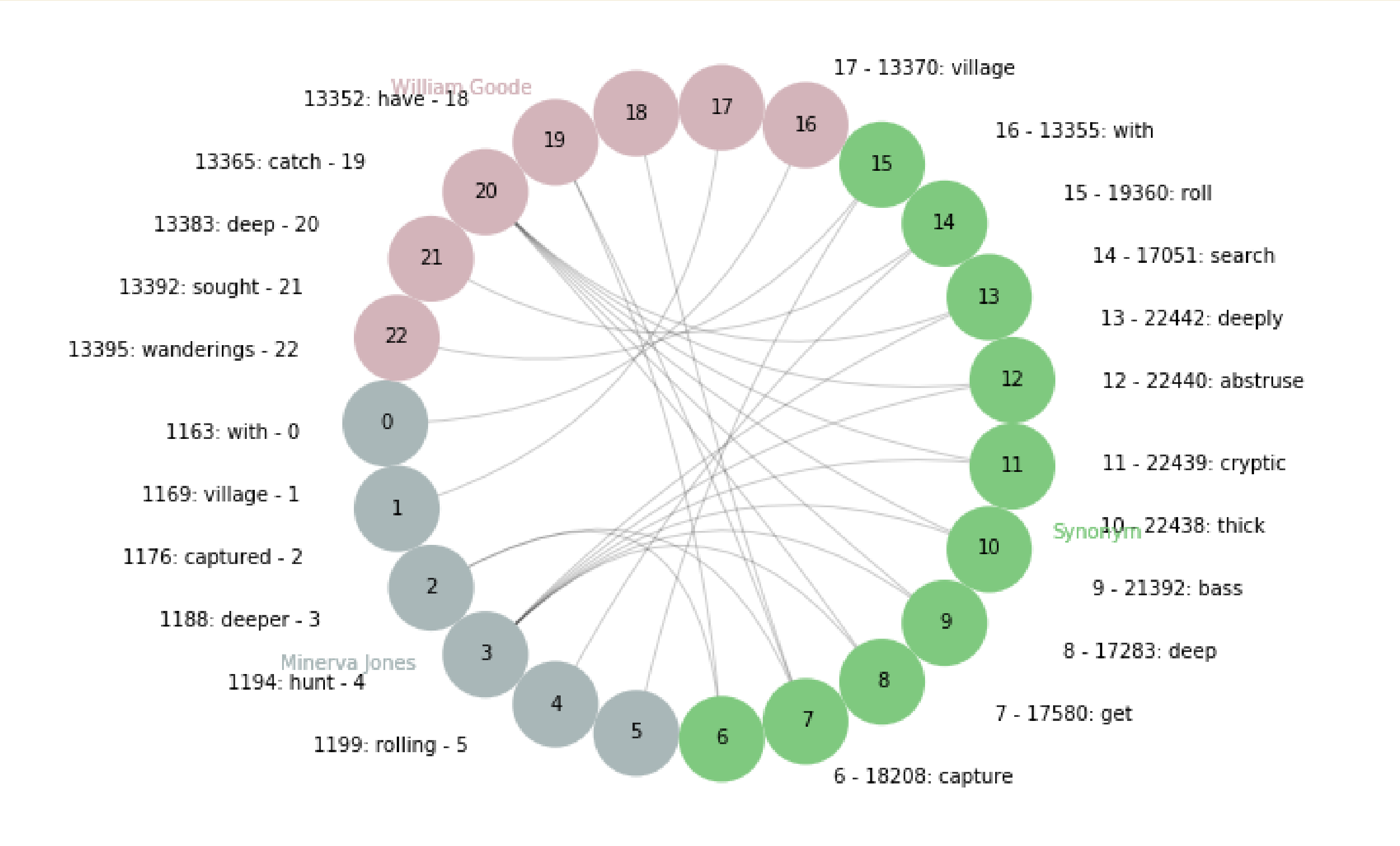
Where to?

Pretty Graphs and Results??

Or, In Depth Process..

Pretty Graphs It Is!!

William Goode (P1) - Minerva Jones (P2)

William Goode (P1) - Minerva Jones (P2)

Similarity Score - 0.25

P1 Contributing Nodes - 7
P2 Contributing Nodes - 6
Synonyms - 10

Direct Relationships - 2
Unique Triangles - 11

William Goode (P1) - Minerva Jones (P2)

P1T1 == P2T1 --> Direct Relationships

"with" - "village"
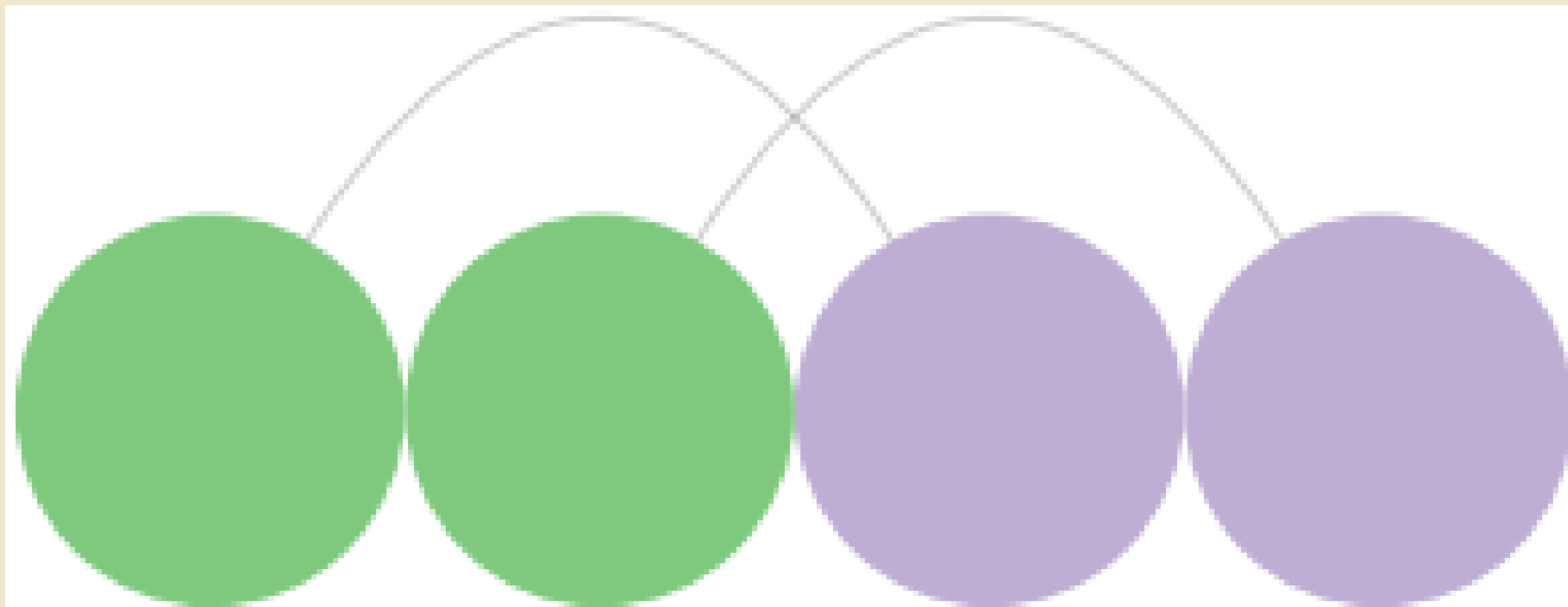
"TO all in the *village* I seemed, no doubt.."

"I AM Minerva, the *village* poetess.."

"Will some one go to the *village* newspaper.."

William Goode (P1) - Minerva Jones (P2)

P1T1 * S1 * P2T1 --> Synonym Relationship

"capture" - "get" - "catch"

"hunt" - "search" - "wanderings"
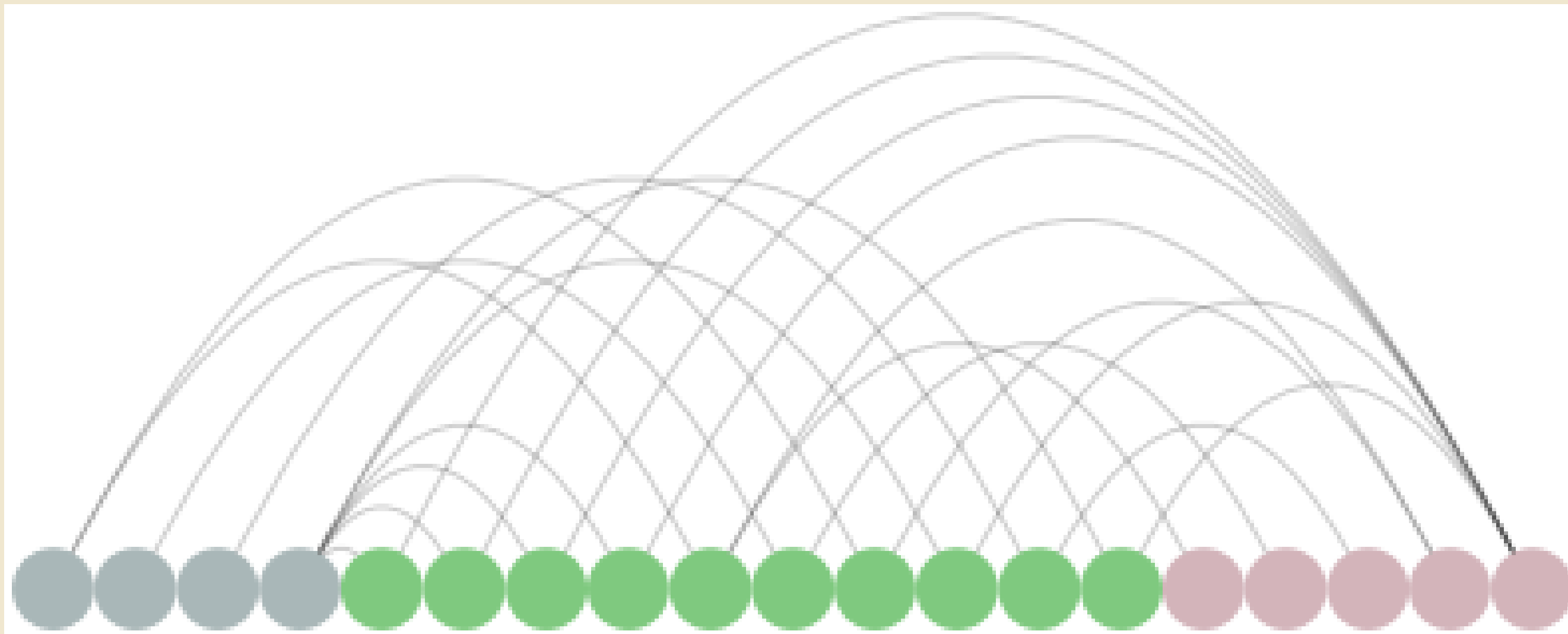
William Goode (P1) - Minerva Jones (P2)

P1T1 !! P2T1 --> Nodes that did not contribute

While these words do provide overall value to the text.. None of these words are either shared across poems, or connected through synonyms. So, they do not provide any benefit to this analysis.

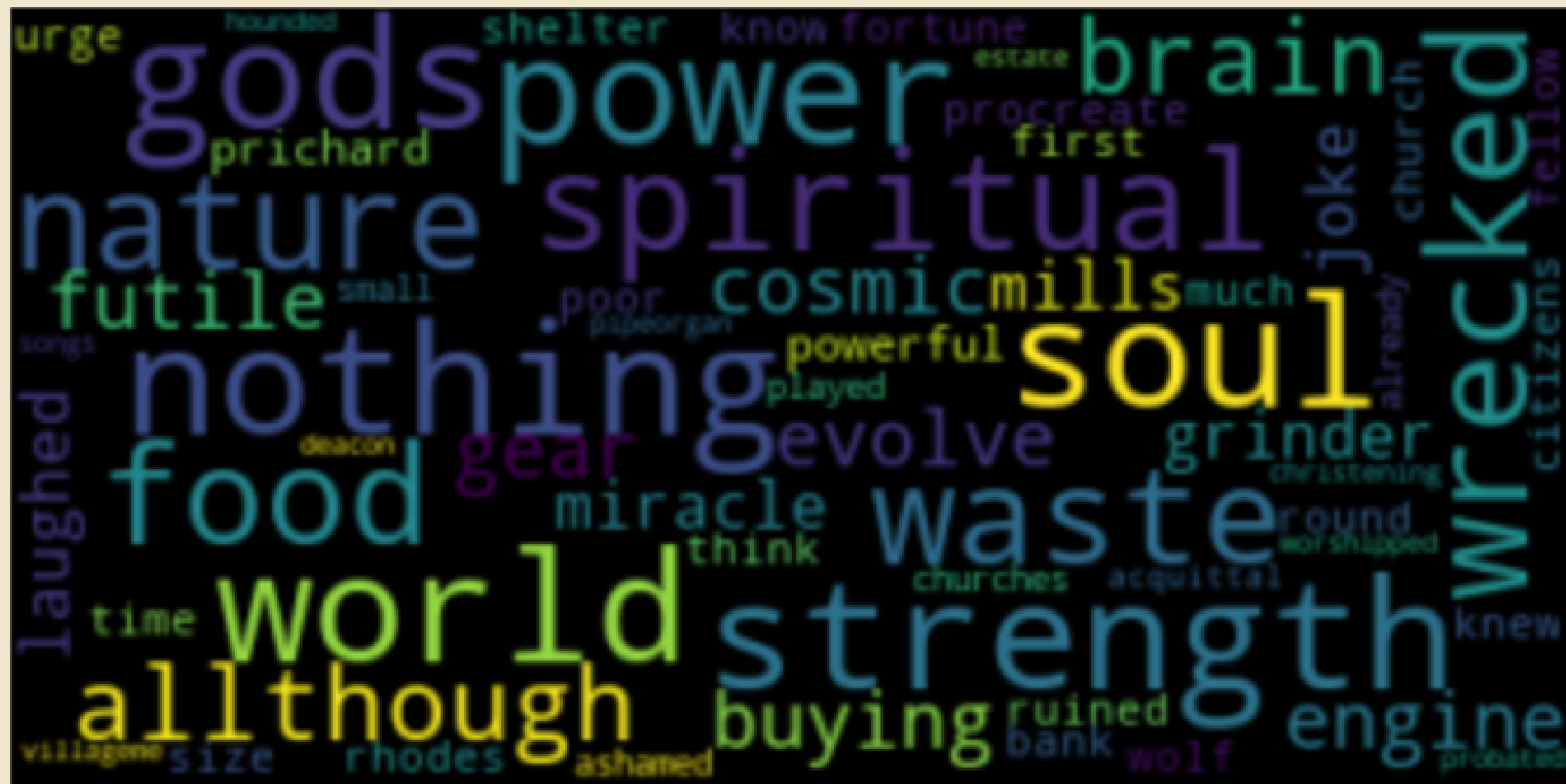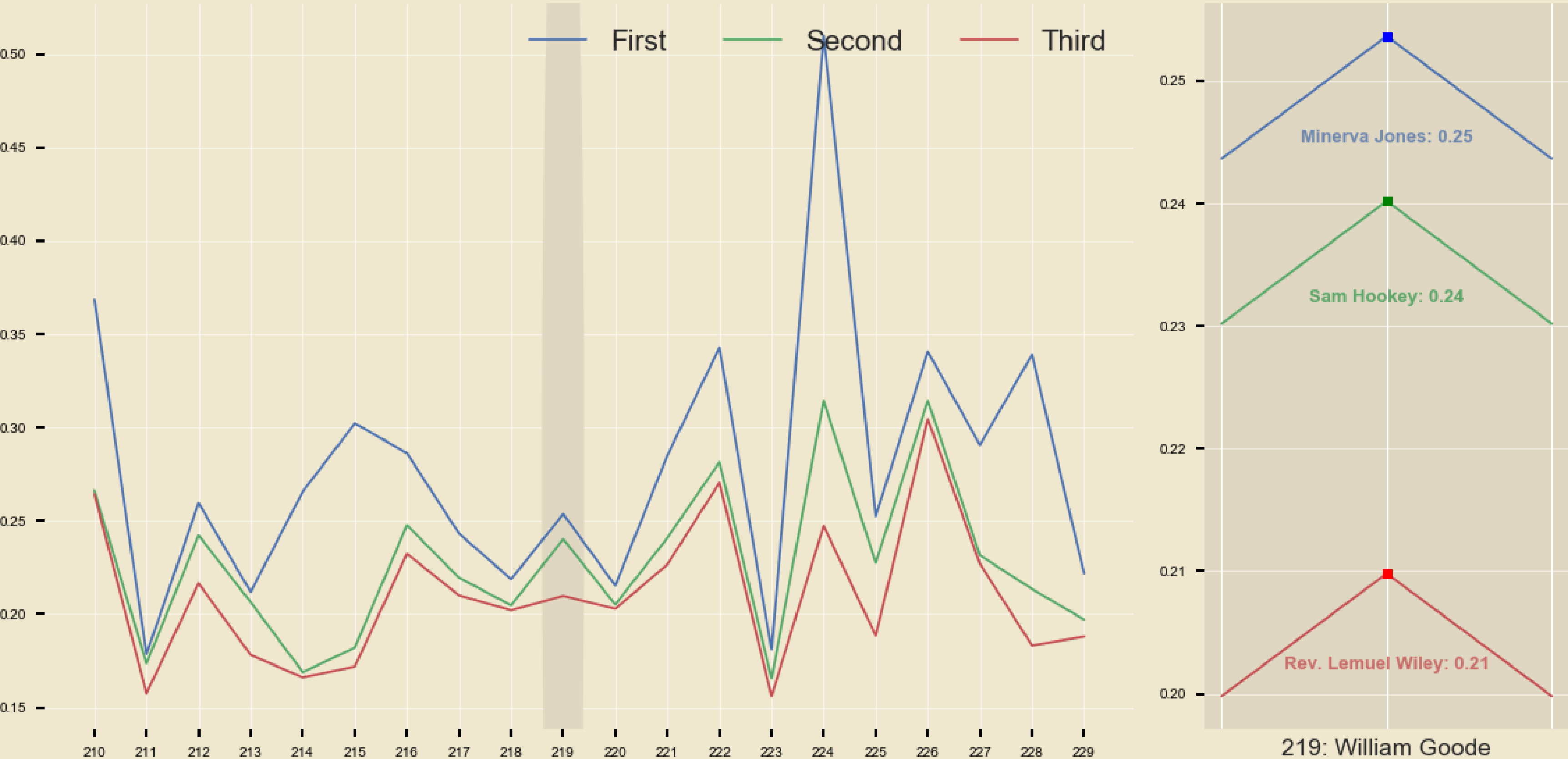*Imagine that process with every poem.. Vs every poem..*

0.074188	0.289016	0.315098	0.433669	0.295456	0.000000	0.441061	0.436752	0.007746	0.431399
0.049329	0.403463	0.374097	0.270315	0.254198	0.441061	0.000000	0.582453	0.104436	0.228844
0.154552	0.231593	0.569522	0.213595	0.157407	0.436752	0.582453	0.000000	0.018722	0.308408
0.050322	0.007191	0.018275	0.038868	0.035323	0.007746	0.104436	0.018722	0.000000	0.033111
0.164652	0.365720	0.296071	0.259547	0.192703	0.431399	0.228844	0.308408	0.033111	0.000000
0.076548	0.319540	0.329047	0.426690	0.367510	0.225855	0.176896	0.252255	0.136090	0.204868
0.133129	0.331368	0.242928	0.191042	0.200507	0.325771	0.364600	0.295863	0.147962	0.404462
0.156201	0.215253	0.243915	0.244729	0.204826	0.213884	0.120251	0.050884	0.144356	0.179078
0.183861	0.310124	0.403089	0.252874	0.438133	0.347194	0.388937	0.296402	0.154910	0.348459
0.212736	0.345186	0.391636	0.357050	0.395082	0.329847	0.287086	0.411527	0.056319	0.234961
0.258121	0.187264	0.333896	0.398021	0.311479	0.365349	0.245467	0.341166	0.063777	0.290020

Hard to take in, huh?

Top 3 Scores Per Poem

First — Second — Third

219: William Goode

Minerva Jones: 0.25

Sam Hookey: 0.24

Rev. Lemuel Wiley: 0.21

Now Let's Compare!