

Data and Network Analysis of 12 Tragedies of William Shakespeare

Khoula Al Riyami

Project Summary

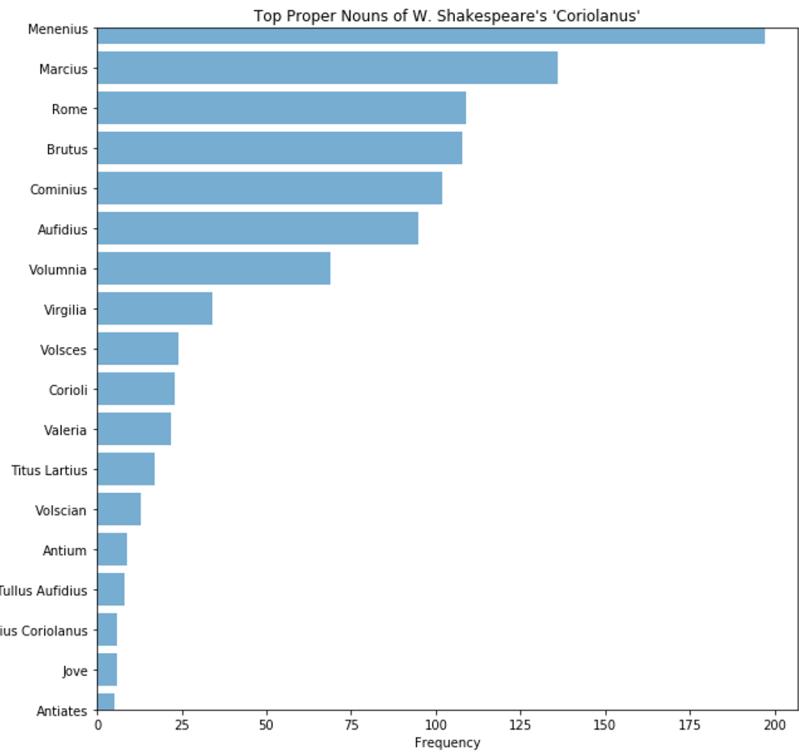
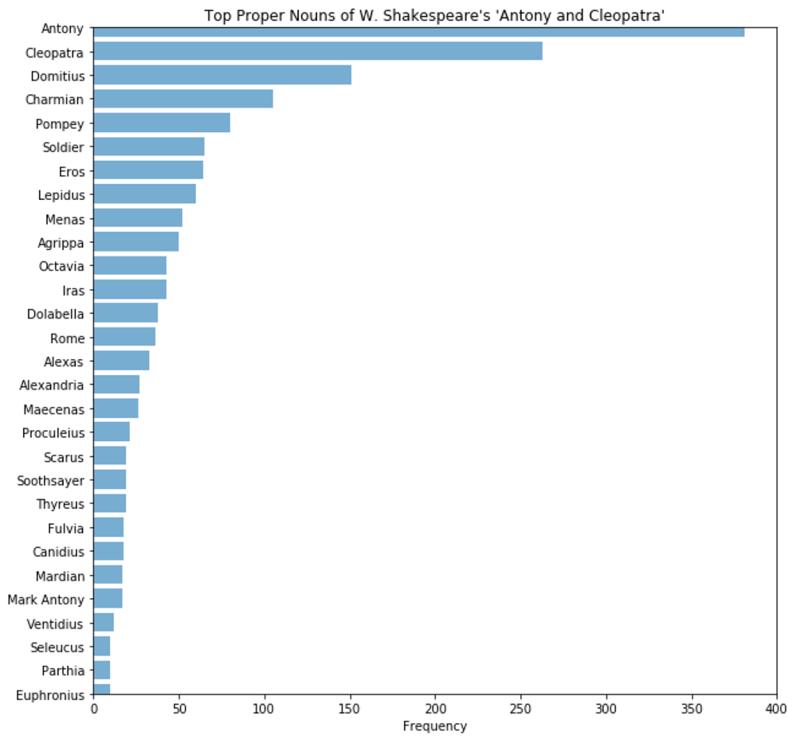
1. Extracting proper nouns in each tragedy.
 - Using NLP techniques for POS tagging.
 - Creating dictionaries of aliases among proper nouns.
 - Frequencies of occurrence of proper nouns in the text of each tragedy.
 - Bar plots of top proper nouns of tragedies.
 - Word clouds of proper nouns of tragedies.
1. Constructing the network of sentimentally co-occurring proper nouns in tragedies.
 - Computing sentential co-occurrences.
 - Measuring the sentiment score (polarity) of each co-occurrence.
 - Plotting the graph of sentimentally co-occurring proper nouns with co-occurrences labeled according to sentiment scores.

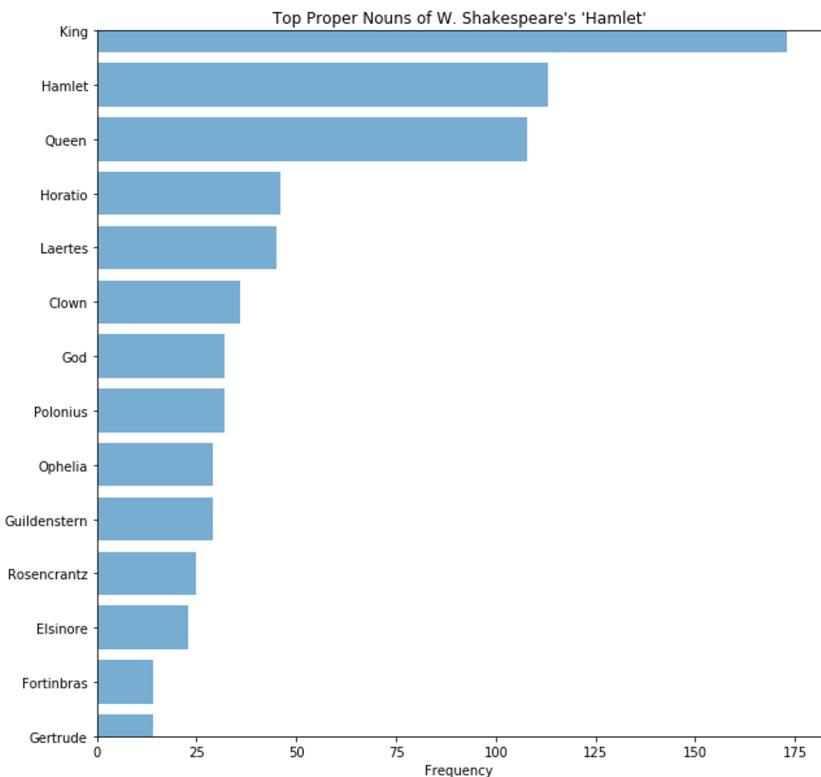
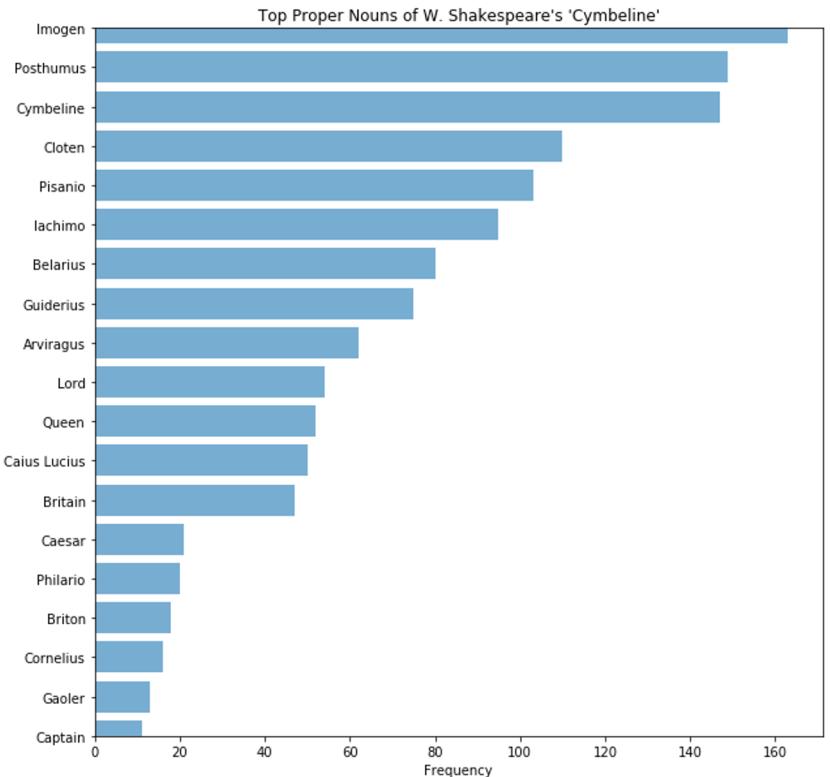
1. Extraction of Names and Descriptive Statistics

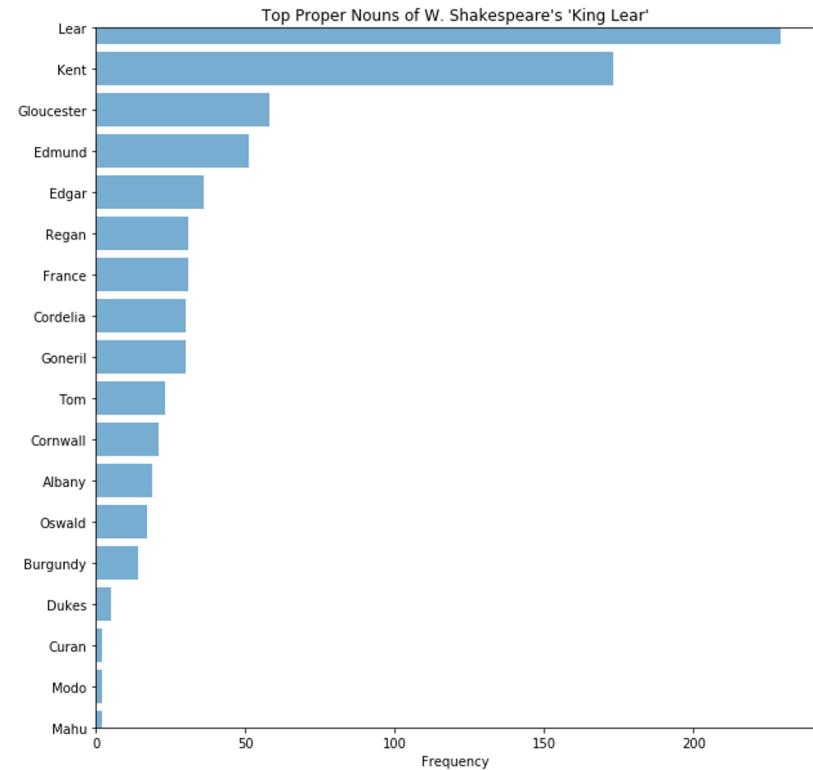
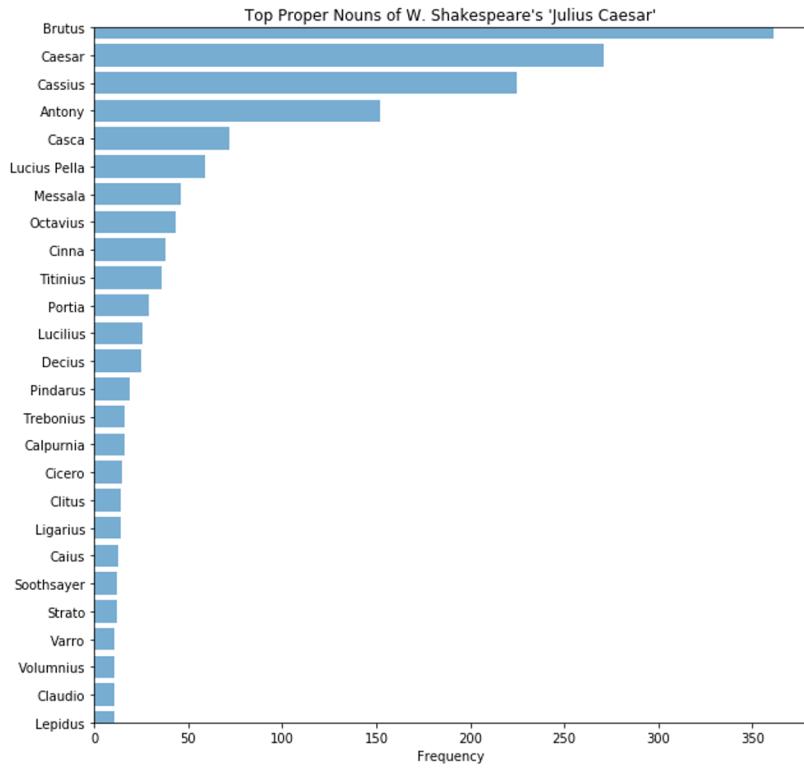
Number of words in each tragedy

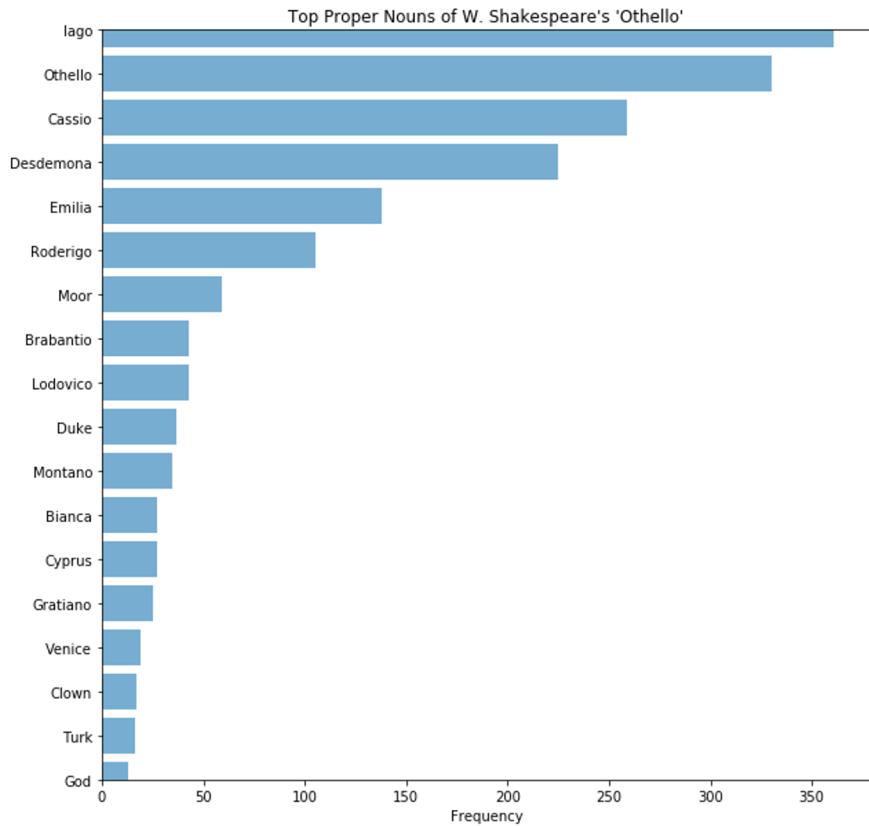
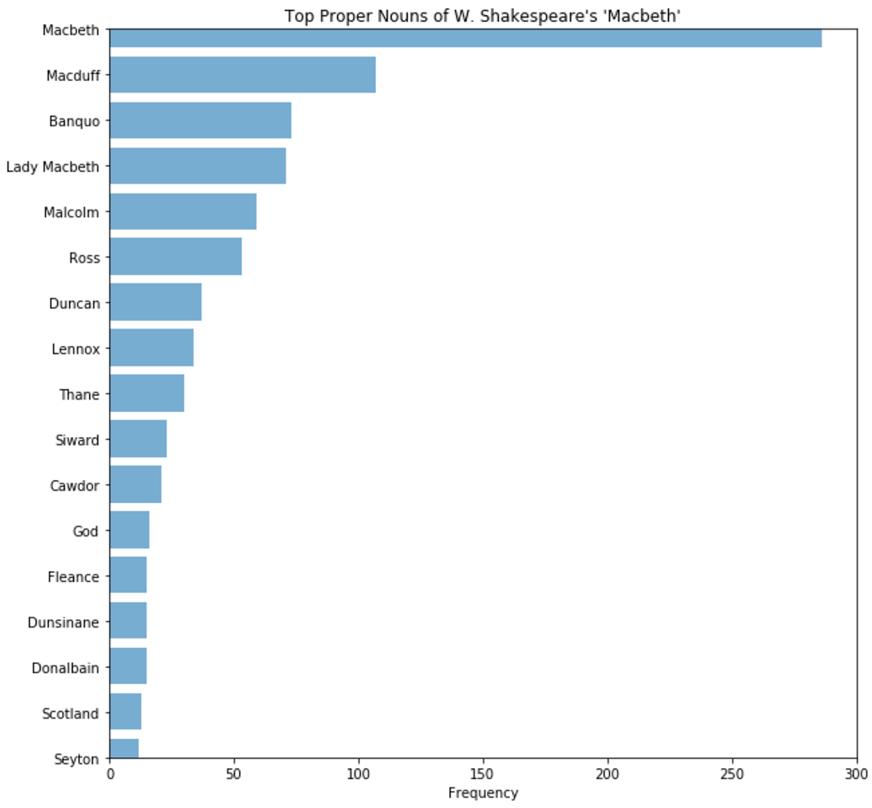
- W. Shakespeare's 'Antony and Cleopatra' contains 34842 nonunique and 4552 unique words
- W. Shakespeare's 'Coriolanus' contains 37539 nonunique and 4615 unique words
- W. Shakespeare's 'Cymbeline' contains 36103 nonunique and 4876 unique words
- W. Shakespeare's 'Hamlet' contains 40450 nonunique and 5378 unique words
- W. Shakespeare's 'Julius Caesar' contains 26161 nonunique and 3299 unique words
- W. Shakespeare's 'King Lear' contains 35463 nonunique and 4817 unique words
- W. Shakespeare's 'Macbeth' contains 22882 nonunique and 3816 unique words
- W. Shakespeare's 'Othello' contains 35613 nonunique and 4185 unique words
- W. Shakespeare's 'Romeo and Juliet' contains 32845 nonunique and 4224 unique words
- W. Shakespeare's 'Timon of Athens' contains 24749 nonunique and 3819 unique words
- W. Shakespeare's 'Titus Andronicus' contains 26791 nonunique and 3857 unique words
- W. Shakespeare's 'Troilus and Cressida' contains 35101 nonunique and 4821 unique words

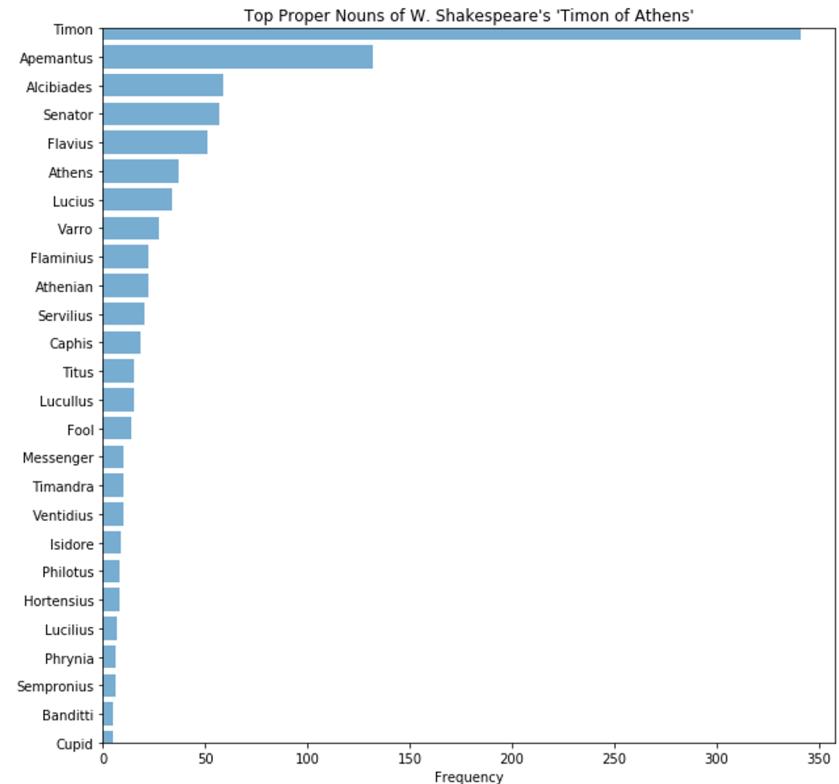
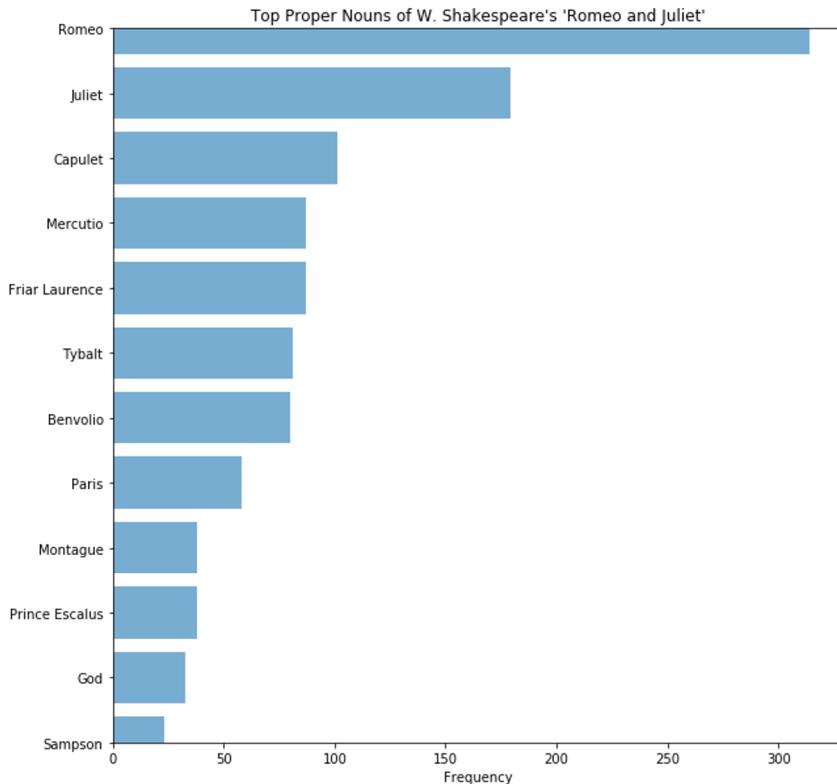
Bar Graphs

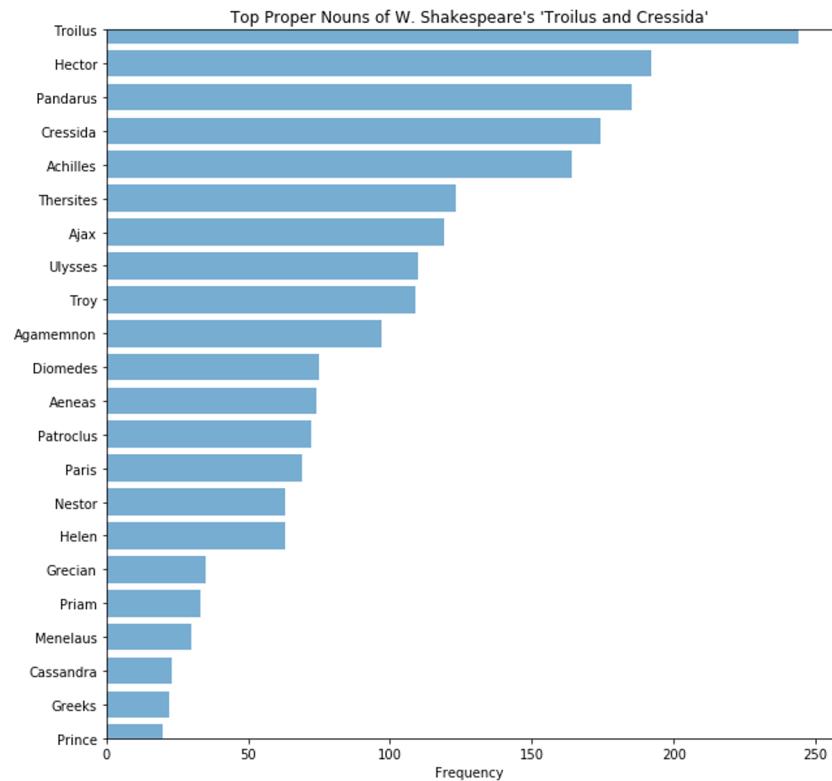
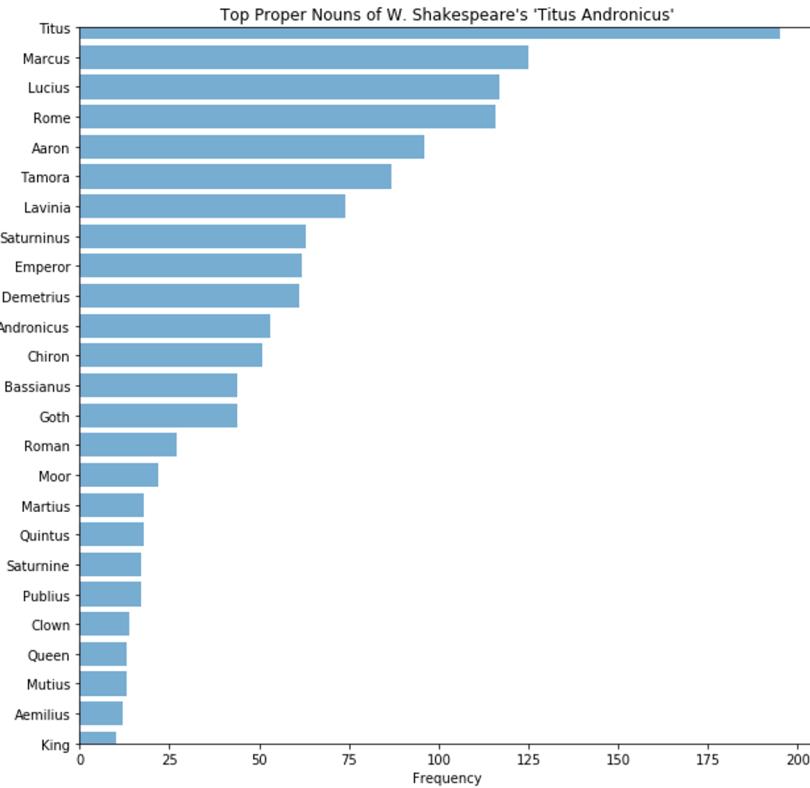












Word Clouds

Proper Nouns of W. Shakespeare's 'Antony and Cleopatra'

A word cloud visualization where the size and color of the words represent their frequency or importance. The names are categorized by color:

- Red:** Caesar, Mardian, Herod, Diomedes, Hipparchus, Bochus, Canidius, Maecenas, Marcus_Crassus, Demetrius, Antoniad, Bocchus.
- Blue:** Pompey, Antony, Menas, Parthia, Isis, Octavian, Proculeius, Silius, Alexandria, Lucius, Galen, Wistres, Thessaly.
- Green:** Domitius, Dolabella, Mark_Antony, Philo, Geron, Misenum, Cologene.
- Yellow:** Gallus, Ventiidius, Fulvia, Alcidies, Toryne, Philo, Geron, Brundusium, Guse, Lichas, Soothsayer, Silius, Juno, Iudea.
- Pink:** Eros, Sextus_Pompeius, Pompeius_Marcellus, Rome, Marcus_Octavius, Alexander, Brutus, Gallus, Messina, Alcidies, Alarum, Sicyon, Brundusium, Guse, Cleopatra, Agrippa, Iras.
- Orange:** Charmian, Seleucus, Scarus, Euphranius, Taurus, Thyreus, Menecrates, Marcus_Antonius, King_Manchus, Julius_Caesar, Dercetas, Sosias.
- Grey:** Clown, Cassius, Tyche, Aeneas, Cydnus, Octavian_Caeser.

Proper Nouns of W. Shakespeare's 'Cymbeline'

Imogen

Pisanio

Cymbeline

Posthumus

Iachimo

Cloten

Briton Lud Hercules Caius_Lucius Guiderius Caesar Arviragus Queen Cassibelan Augustus_Cesar Gaoler Captain Messenger Julius_Cesar_Cyulus Morgan Dian Simon Philomel Amnest Cytherea Lord Polydore Cornelius Phedrus Hecuba Belarius Senator Ladies Cleopatra Gordian Teres Jove Fidele Juno Soothsayer Philario Cadwal Leonati Diana Tenantius Sicilius Frenchman Britain

Proper Nouns of W. Shakespeare's 'Coriolanus'

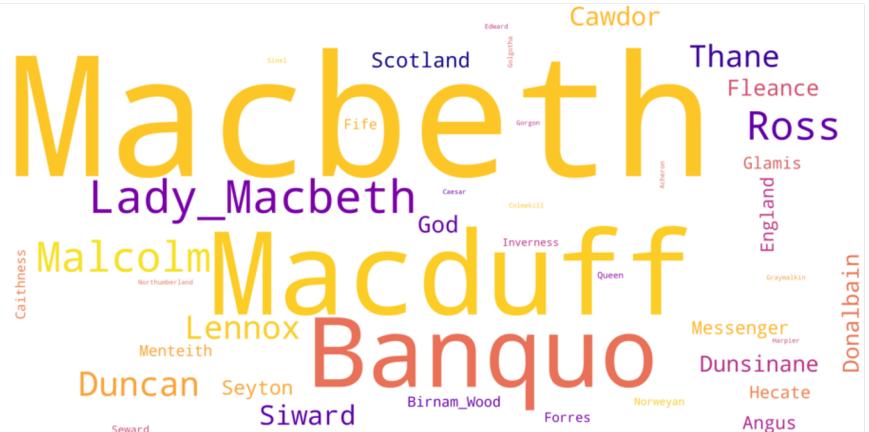
A word cloud diagram for the play 'Julius Caesar' by William Shakespeare. The words are arranged in a grid-like pattern, with larger words representing more prominent characters or themes. The words are color-coded: Brutus, Volumnia, Menenius, Marcus, Aufidius, Coriolanus, Galen, Tullus, Cato, Volscian, Titus, Lartius, Aediles, Sicinius, Velutus, Antiates, Hector, Ulysses, Afric, Junius, Juno, Lictors, Valeria, Cominius, Aufidius, Jove, Diane, Alexander, Adrian, Hercules, Gato, Tarquin, Penelope, Quintus, Publius, Menenius_Agrippa, and Nicobar.

Proper Nouns of W. Shakespeare's 'Hamlet'

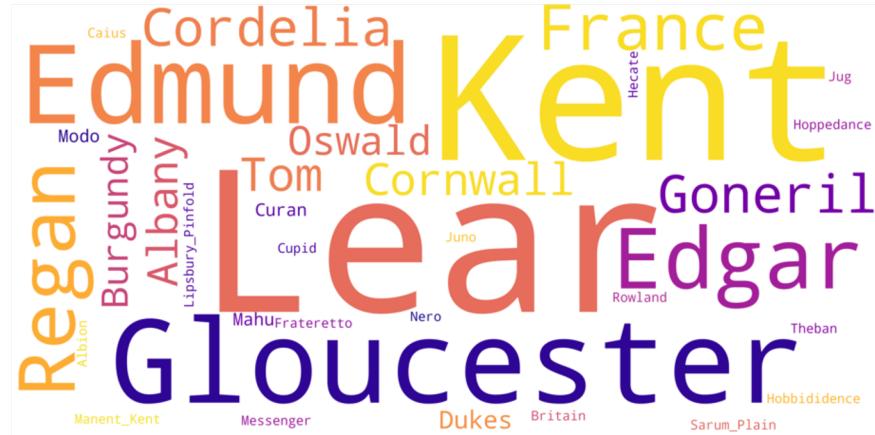
Proper Nouns of W. Shakespeare's 'Julius Caesar'



Proper Nouns of W. Shakespeare's 'Macbeth'



Proper Nouns of W. Shakespeare's 'King Lear'



Proper Nouns of W. Shakespeare's 'Othello'



Proper Nouns of W. Shakespeare's 'Romeo and Juliet'

A word cloud diagram for the play Romeo and Juliet. The words "Juliet" and "Romeo" are the largest and most central, both in purple. Other major characters like "Prince Escalus", "Paris", "Capulet", "Tybalt", "Mercutio", and "Benvolio" are also prominent in purple. Supporting characters and locations are shown in smaller, colored words scattered around the main names.

Proper Nouns of W. Shakespeare's 'Titus Andronicus'

Proper Nouns of W. Shakespeare's 'Timon of Athens'

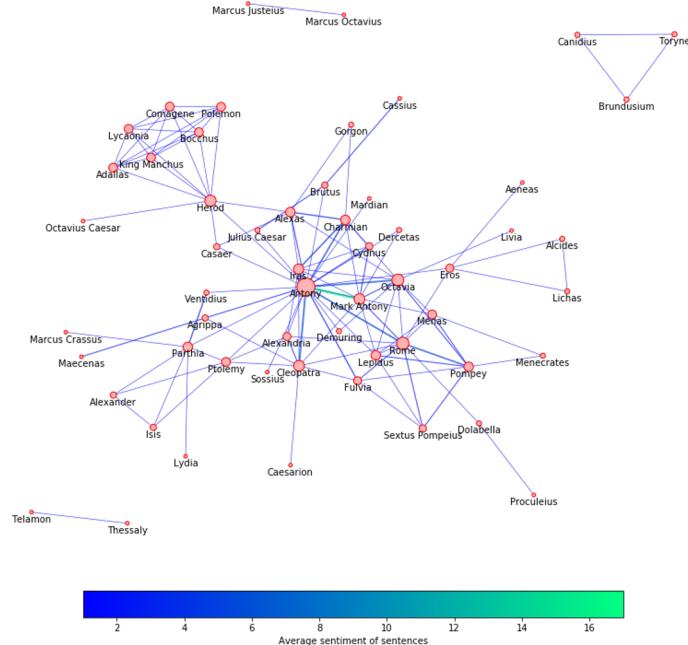
Proper Nouns of W. Shakespeare's 'Troilus and Cressida'

Networks of Co-Occurring Proper Nouns

Antony and Cleopatra

- W. Shakespeare's 'Antony and Cleopatra' contains 185 sentential co-occurrences among 86 aliased proper nouns
- The graph of sentimentally co-occurring proper nouns in W. Shakespeare's 'Antony and Cleopatra' is a weighted graph and it has 57 nodes and 122 edges
- The density of this graph is 0.076
- This graph is a disconnected graph and it has 4 connected components
- The largest connected component of this graph is a weighted graph with 50 nodes and 117 edges
- The density of the largest connected component of this graph is 0.096

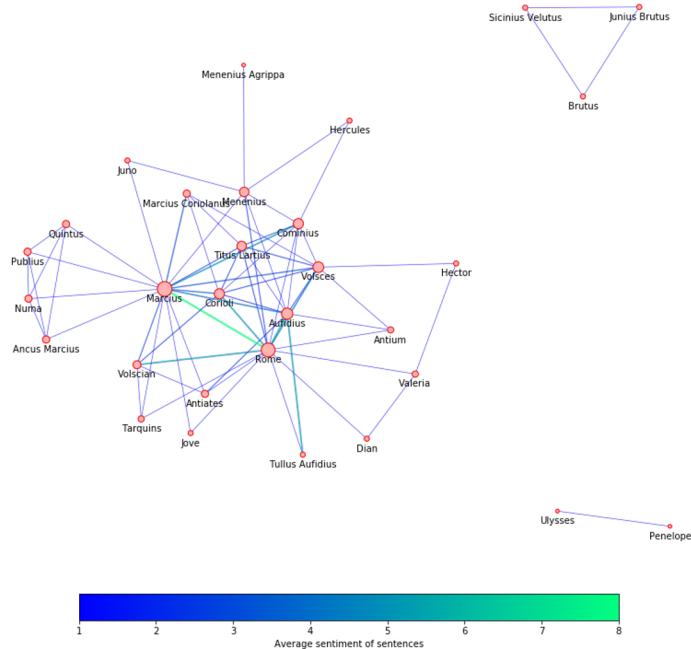
The graph of co-occurring proper nouns in W. Shakespeare's 'Antony and Cleopatra'
weighted over their average sentiment score



Coriolanus

- W. Shakespeare's 'Coriolanus' contains 122 sentential co-occurrences among 44 aliased proper nouns
- The graph of sentimentally co-occurring proper nouns in W. Shakespeare's 'Coriolanus' is a weighted graph and it has 30 nodes and 70 edges
- The density of this graph is 0.161
- This graph is a disconnected graph and it has 3 connected components
- The largest connected component of this graph is a weighted graph with 25 nodes and 66 edges
- The density of the largest connected component of this graph is 0.220

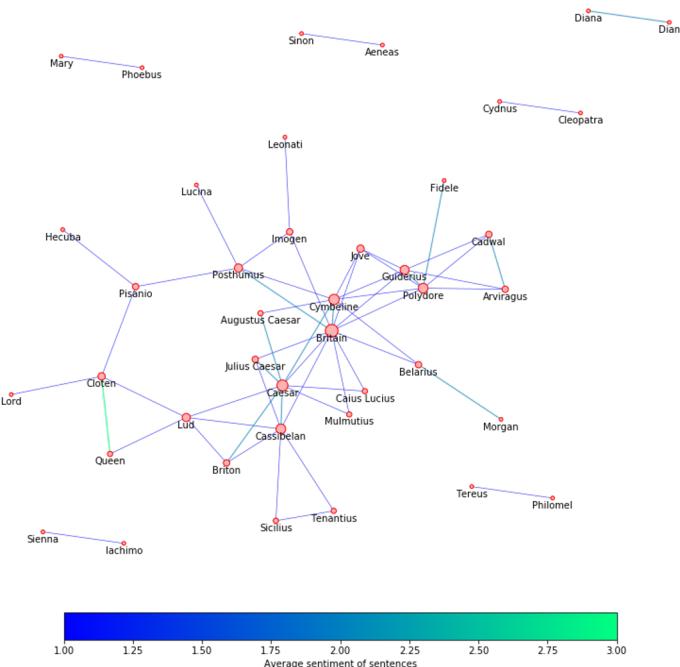
The graph of co-occurring proper nouns in W. Shakespeare's 'Coriolanus'
weighted over their average sentiment score



Cymbeline

- W. Shakespeare's 'Cymbeline' contains 73 sentential co-occurrences among 62 aliased proper nouns
- The graph of sentimentally co-occurred proper nouns in W. Shakespeare's 'Cymbeline' is a weighted graph and it has 41 nodes and 59 edges
- The density of this graph is 0.072
- This graph is a disconnected graph and it has 7 connected components
- The largest connected component of this graph is a weighted graph with 29 nodes and 53 edges
- The density of the largest connected component of this graph is 0.131

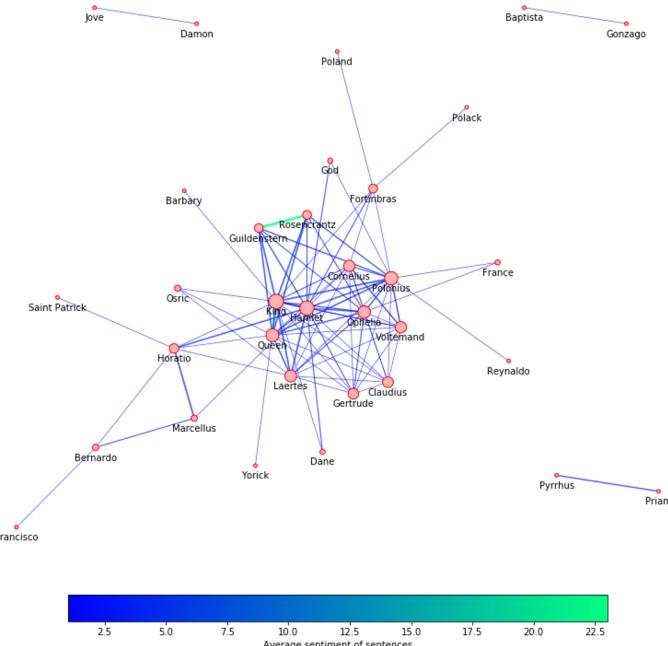
The graph of co-occurred proper nouns in W. Shakespeare's 'Cymbeline' weighted over their average sentiment score



Hamlet

- W. Shakespeare's 'Hamlet' contains 205 sentential co-occurrences among 54 aliased proper nouns
- The graph of sentimentally co-occurring proper nouns in W. Shakespeare's 'Hamlet' is a weighted graph and it has 33 nodes and 87 edges
- The density of this graph is 0.165
- This graph is a disconnected graph and it has 4 connected components
- The largest connected component of this graph is a weighted graph with 27 nodes and 84 edges
- The density of the largest connected component of this graph is 0.239

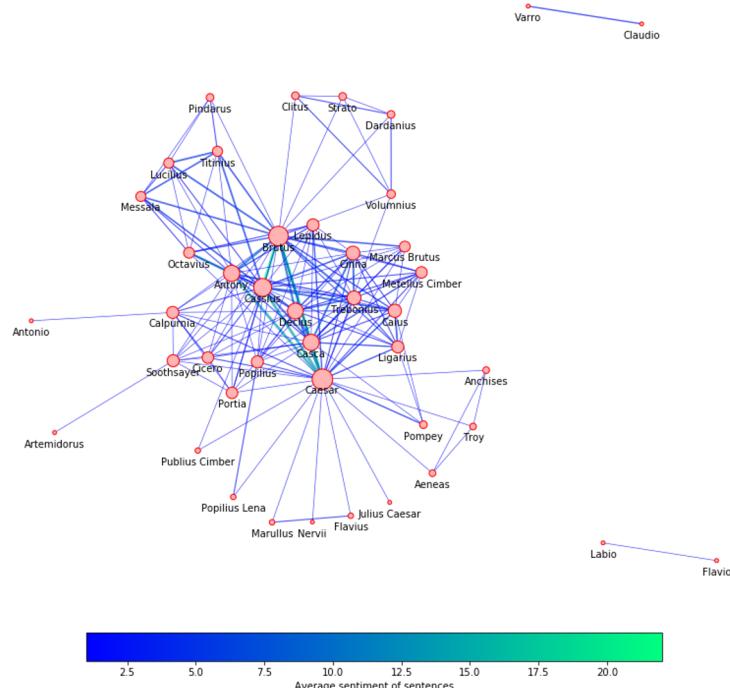
The graph of co-occurring proper nouns in W. Shakespeare's 'Hamlet'
weighted over their average sentiment score



Julius Caesar

- W. Shakespeare's 'Julius Caesar' contains 428 sentential co-occurrences among 48 aliased proper nouns
- The graph of sentimentally co-occurred proper nouns in W. Shakespeare's 'Julius Caesar' is a weighted graph and it has 43 nodes and 166 edges
- The density of this graph is 0.184
- This graph is a disconnected graph and it has 3 connected components
- The largest connected component of this graph is a weighted graph with 39 nodes and 164 edges
- The density of the largest connected component of this graph is 0.221

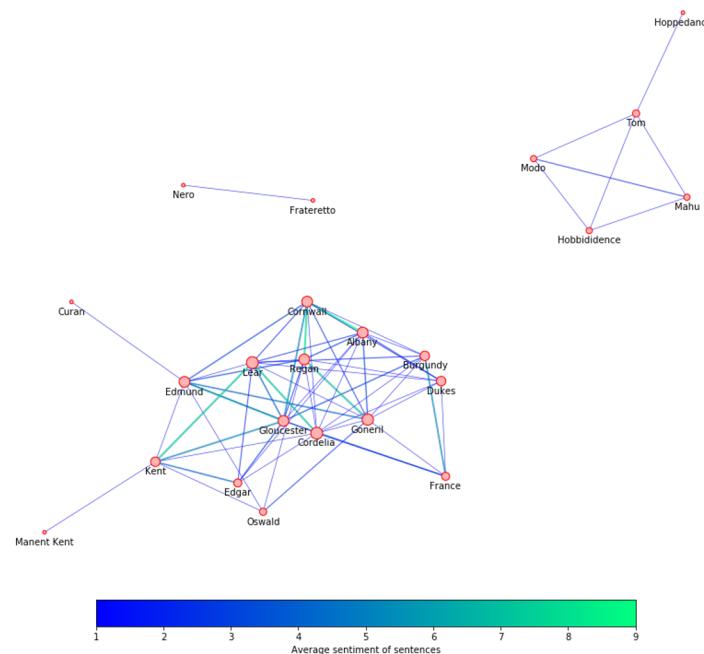
The graph of co-occurring proper nouns in W. Shakespeare's 'Julius Caesar'
weighted over their average sentiment score



King Lear

- W. Shakespeare's 'King Lear' contains 150 sentential co-occurrences among 35 aliased proper nouns
- The graph of sentimentally co-occurrent proper nouns in W. Shakespeare's 'King Lear' is a weighted graph and it has 23 nodes and 65 edges
- The density of this graph is 0.257
- This graph is a disconnected graph and it has 3 connected components
- The largest connected component of this graph is a weighted graph with 16 nodes and 57 edges
- The density of the largest connected component of this graph is 0.475

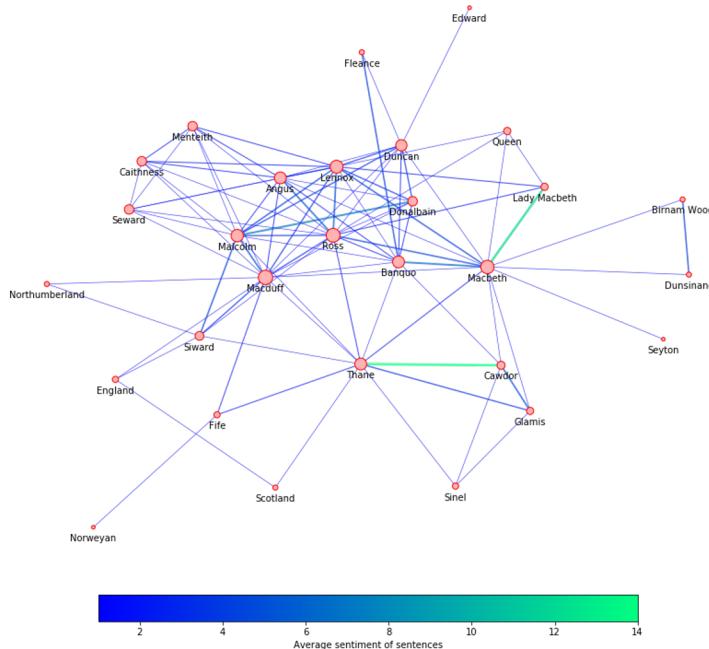
The graph of co-occurrent proper nouns in W. Shakespeare's 'King Lear'
weighted over their average sentiment score



Macbeth

- W. Shakespeare's 'Macbeth' contains 199 sentential co-occurrences among 41 aliased proper nouns
- The graph of sentimentally co-occurring proper nouns in W. Shakespeare's 'Macbeth' is a weighted graph and it has 29 nodes and 93 edges
- The density of this graph is 0.229
- This graph is a connected graph
- The largest connected component of this graph is a weighted graph with 29 nodes and 93 edges
- The density of the largest connected component of this graph is 0.229

The graph of co-occurring proper nouns in W. Shakespeare's 'Macbeth' weighted over their average sentiment score



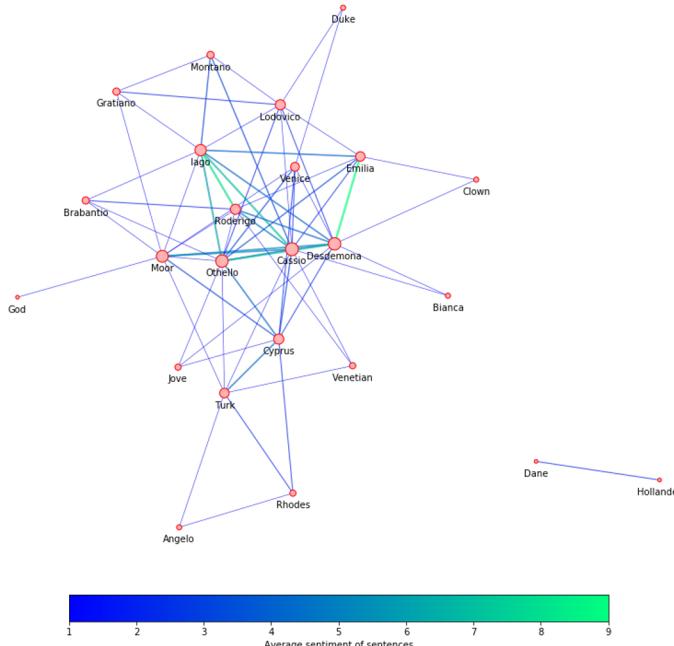
Othello

- W. Shakespeare's 'Othello' contains 154 sentential co-occurrences among 34 aliased proper nouns

The graph of sentimentally co-occurring proper nouns in W. Shakespeare's 'Othello' is a weighted graph and it has 24 nodes and 67 edges

- The density of this graph is 0.243
- This graph is a disconnected graph and it has 2 connected components
- The largest connected component of this graph is a weighted graph with 22 nodes and 66 edges
- The density of the largest connected component of this graph is 0.286

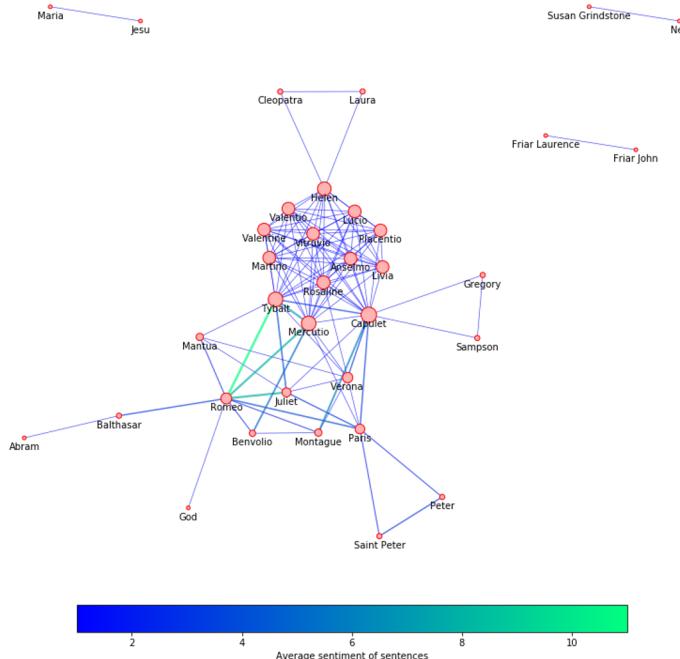
The graph of co-occurring proper nouns in W. Shakespeare's 'Othello'
weighted over their average sentiment score



Romeo and Juliet

- W. Shakespeare's 'Romeo and Juliet' contains 181 sentential co-occurrences among 52 aliased proper nouns
- The graph of sentimentally co-occurred proper nouns in W. Shakespeare's 'Romeo and Juliet' is a weighted graph and it has 35 nodes and 118 edges
- The density of this graph is 0.198
- This graph is a disconnected graph and it has 4 connected components
- The largest connected component of this graph is a weighted graph with 29 nodes and 115 edges
- The density of the largest connected component of this graph is 0.283

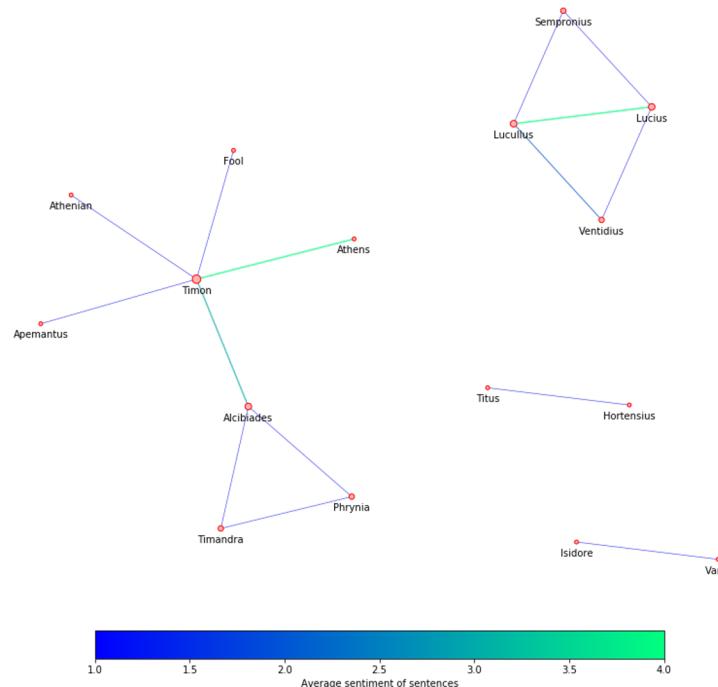
The graph of co-occurred proper nouns in W. Shakespeare's 'Romeo and Juliet'
weighted over their average sentiment score



Timon of Athens

- W. Shakespeare's 'Timon of Athens' contains 24 sentential co-occurrences among 30 aliased proper nouns
- The graph of sentimentally co-occurring proper nouns in W. Shakespeare's 'Timon of Athens' is a weighted graph and it has 16 nodes and 15 edges
- The density of this graph is 0.125
- This graph is a disconnected graph and it has 4 connected components
- The largest connected component of this graph is a weighted graph with 8 nodes and 8 edges
- The density of the largest connected component of this graph is 0.286

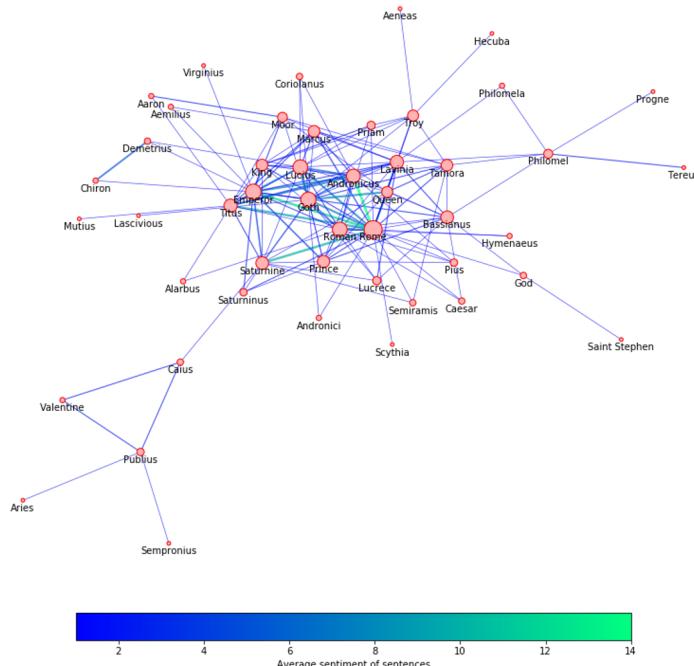
The graph of co-occurring proper nouns in W. Shakespeare's 'Timon of Athens'
weighted over their average sentiment score



Titus Andronicus

- W. Shakespeare's 'Titus Andronicus' contains 304 sentential co-occurrences among 64 aliased proper nouns
- The graph of sentimentally co-occurring proper nouns in W. Shakespeare's 'Titus Andronicus' is a weighted graph and it has 48 nodes and 147 edges
- The density of this graph is 0.130
- This graph is a connected graph
- The largest connected component of this graph is a weighted graph with 48 nodes and 147 edges
- The density of the largest connected component of this graph is 0.130

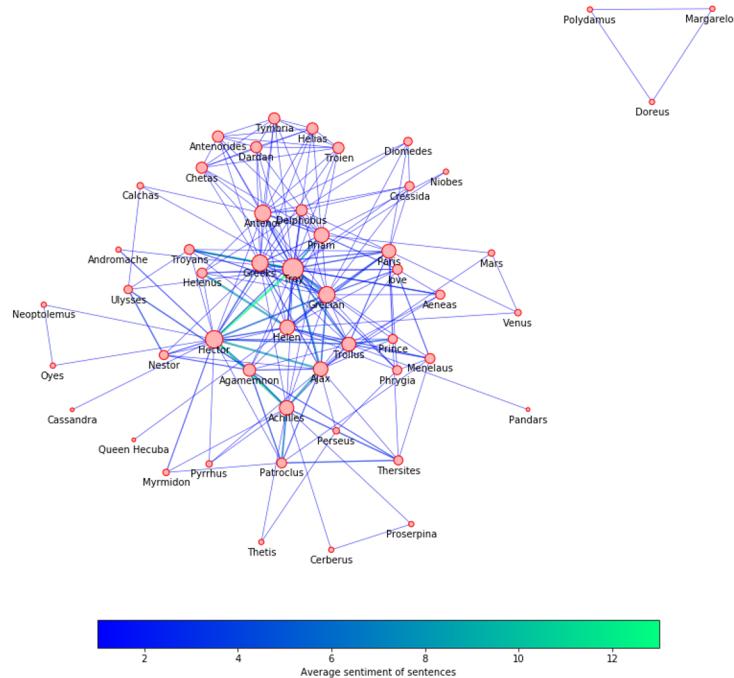
The graph of co-occurring proper nouns in W. Shakespeare's 'Titus Andronicus'
weighted over their average sentiment score



Troilus and Cressida

- W. Shakespeare's 'Troilus and Cressida' contains 357 sentential co-occurrences among 62 aliased proper nouns
- The graph of sententially co-occurring proper nouns in W. Shakespeare's 'Troilus and Cressida' is a weighted graph and it has 51 nodes and 198 edges
- The density of this graph is 0.155
- This graph is a disconnected graph and it has 2 connected components
- The largest connected component of this graph is a weighted graph with 48 nodes and 195 edges
- The density of the largest connected component of this graph is 0.173

The graph of co-occurring proper nouns in W. Shakespeare's 'Troilus and Cressida'
weighted over their average sentiment score



Future Plans

1. To study the common graph of sententially co-occurring proper nouns for the whole corpora of the 12 tragedies.
2. Topic Modeling of terms in the tragedies.
3. Graphs of co-occurring Topic Modeling terms in all tragedies.

THANK YOU