LOYOLA UNIVERSITY CHICAGO

CENTER FOR TEXTUAL STUDIES AND DIGITAL HUMANITIES

Introduction to Digital Humanities Research & Computing

Fall Semester 2015

Week 12

Final Presentation & Report

Any questions?

Conceptual Design Specification

- use your outline for a conceptual project as the basis
- this project focuses on the conceptual implementation rather than using a proposal model
- think about how to plan the construction of the programming or output of the project
- choose a 'software engineering model' to help plan and outline your project

eg: you might choose to base the process on the waterfall model using the steps for analysis, design, implementation, and testing

- detail each step of your model for your project
- implementation and testing phases can be conceptual
- explain in an introduction or overview why you chose your particular software model
- models can also be used, such as flowcharts or UML EXAMPLE

(Here Mr. Carmichael, who was reading Virgil, blew out his candle.)

3

But what after all is one night? A short space, especially when the darkness dims so soon, and so soon a bird sings, a cock crows loudly, or a faint green quickens, like a turning leaf, in the hollow of the wave.

It seemed now as if, touched by human penitence and all its tool, divine goodness had parted the curtain and displayed behind it, single, distinct, **the hare erect**; the wave falling; the boat rocking, which, did we deserve them, should be

Weekly Exercise - Example Solution

```
<?xml version="1.0" encoding="UTF-8"?>
<novel>
<page no=”198”>
<line no=”1”>
<title align=”centre” style=”underline” colour=”black”>To the Lighthouse</title>
</line>
<line no=”2” type=”empty”></line>
<para no=”1”>
<line no=”3”>(Here Mr. Carmichael, who was reading</line>
…
</para>
<line no=”6”>
<section align=”centre” style=”heading” colour=”blue” no=”3”>3</section>
</line>
...
<line no=”11”><rend style=”highlight” colour=”green”>green</rend>...</line>
...
</page>
</novel>
```

Project Presentation

4th December 2015 - ~10 mins per paper & questions from the class

- conference style short paper
- outline project proposal
- outline your design specification etc
- use template for guidance
- UML or flowchart outlines

Stylistic Analysis

Overview

- consideration of patterns in style
- influence of style on readers' perceptions
- disciplinary concerns of literary and linguistic interpretation
- patterns using computational stylistics

Stylistic Analysis

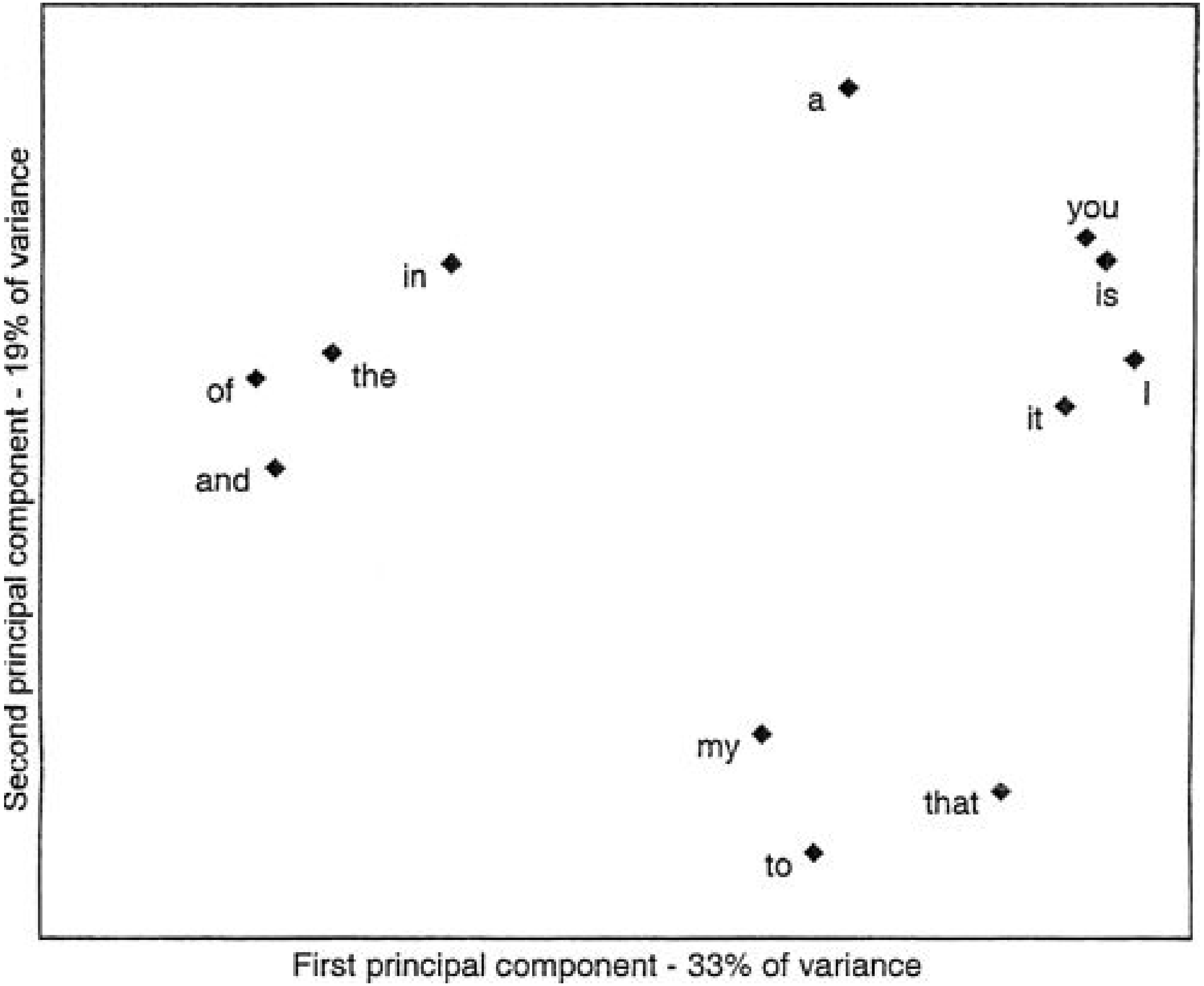
Analysing Shakespeare (Hugh Craig, 'A Companion to Digital Humanities')

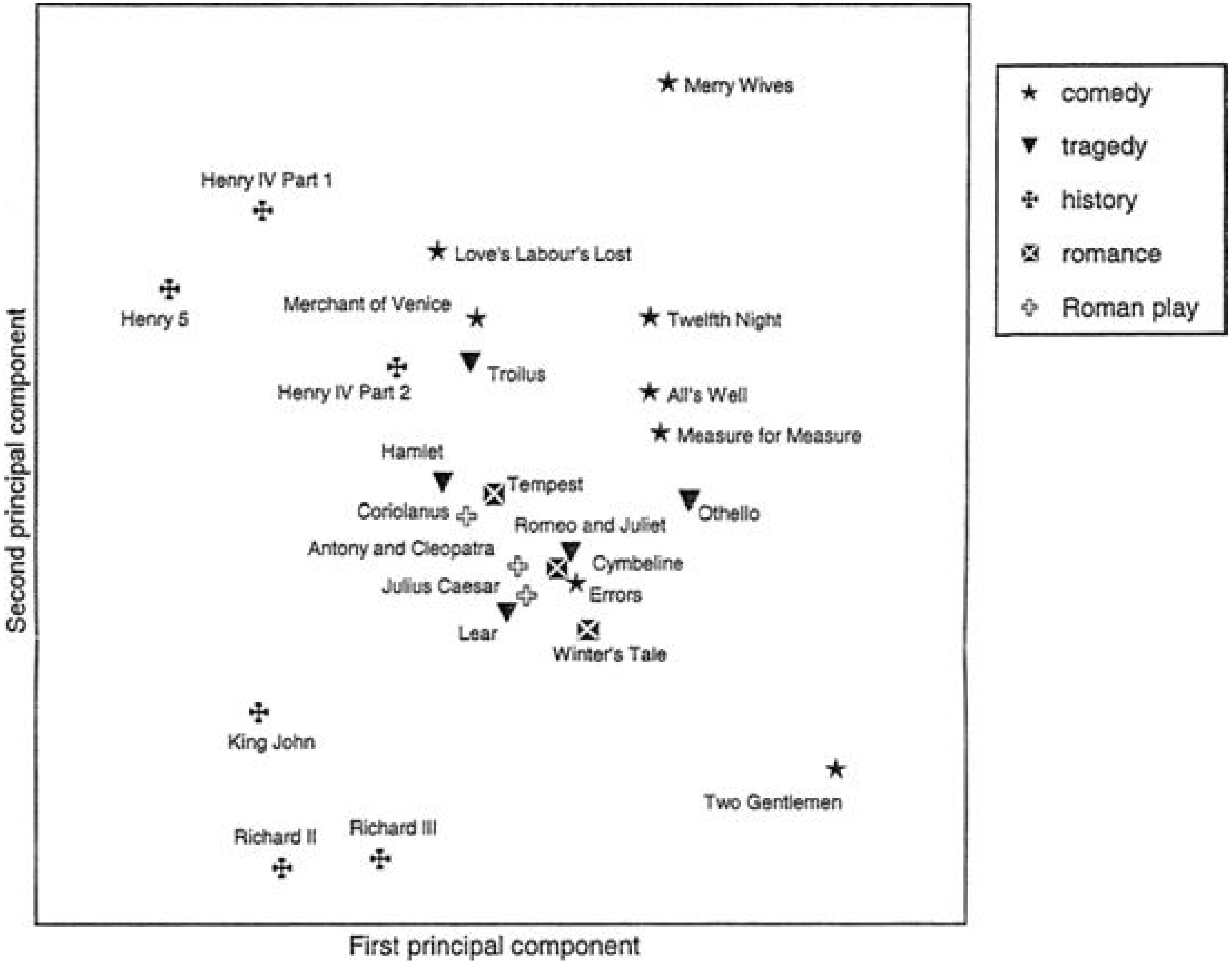
- analysis of 25 of 38 plays in standard Complete Shakespeare
- conduct analysis of the 12 most commonly used words
- Principal Components Analysis (PCA)
- PCA simplifies data by finding new variables that represent most relationships
- new composite variables still represent the variation in a data set
- PCA vectors are an extension of this principle

Stylistic Analysis

Analysing Shakespeare (Hugh Craig, 'A Companion to Digital Humanities')

- relative weightings used to create vector which accounts for greatest proportion of variance
- then the second vector, and so on...
- strong associations between variables will lead to the first few composite variables accounting for most of the results


a
you
in
is
the
of
I
it
and
my
that
to
First principal component - 33% of variance
Second principal component - 19% of variance


Merry Wives
Henry IV Part 1
Love's Labour's Lost
Henry 5
Merchant of Venice
Twelfth Night
Troilus
Henry IV Part 2
All's Well
Measure for Measure
Hamlet
Tempest
Othello
Coriolanus
Romeo and Juliet
Antony and Cleopatra
Cymbeline
Julius Caesar
Errors
Lear
Winter's Tale
King John
Two Gentlemen
Richard II
Richard III
First principal component
Second principal component
comedy
tragedy
history
romance
Roman play

Stylistic Analysis

Considerations

- consider how constant such patterns are when we remove or add a play
- effect of modifying variables on the results
- why did we choose the given parameters for this particular study?

Stylistic Analysis

Computational Stylistics

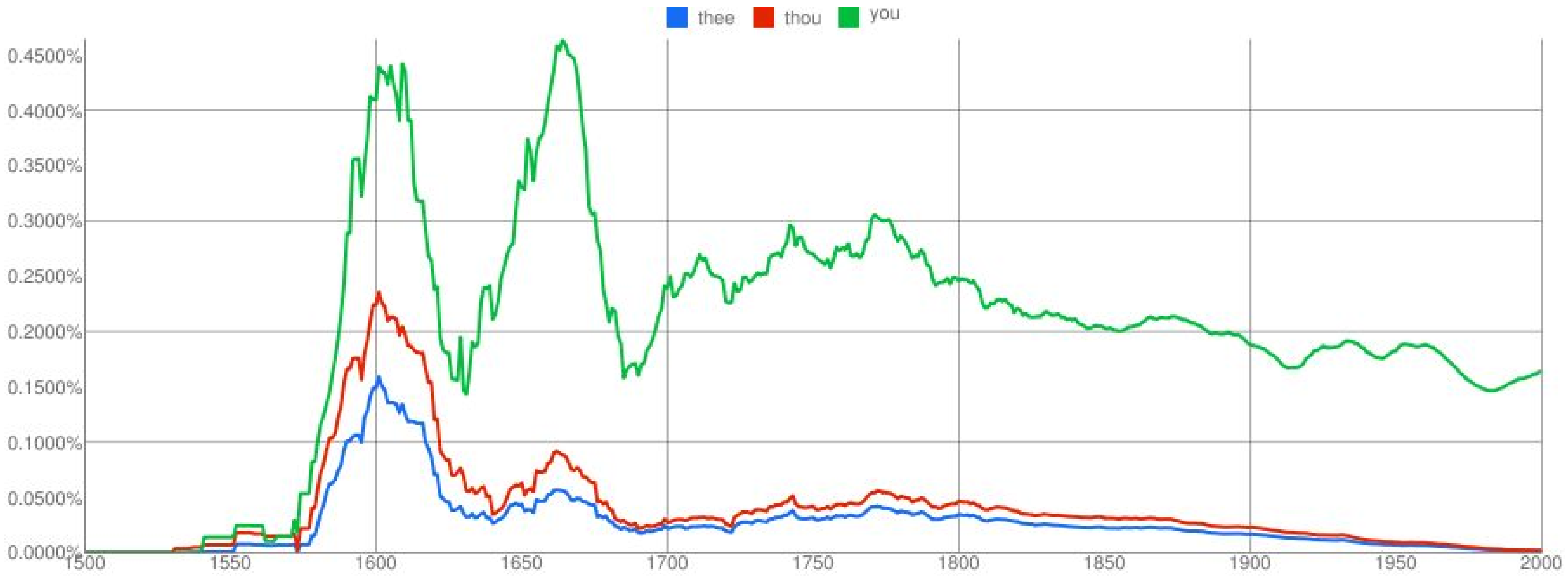
- extensive and perhaps best suited to large scale comparisons
- changes in language over time, eg: a writer's career, historical period...
- provision of a class of evidence not otherwise accessible
- not a solution by itself, requires knowledge of humanities and statistical techniques

Google Books Ngram Viewer

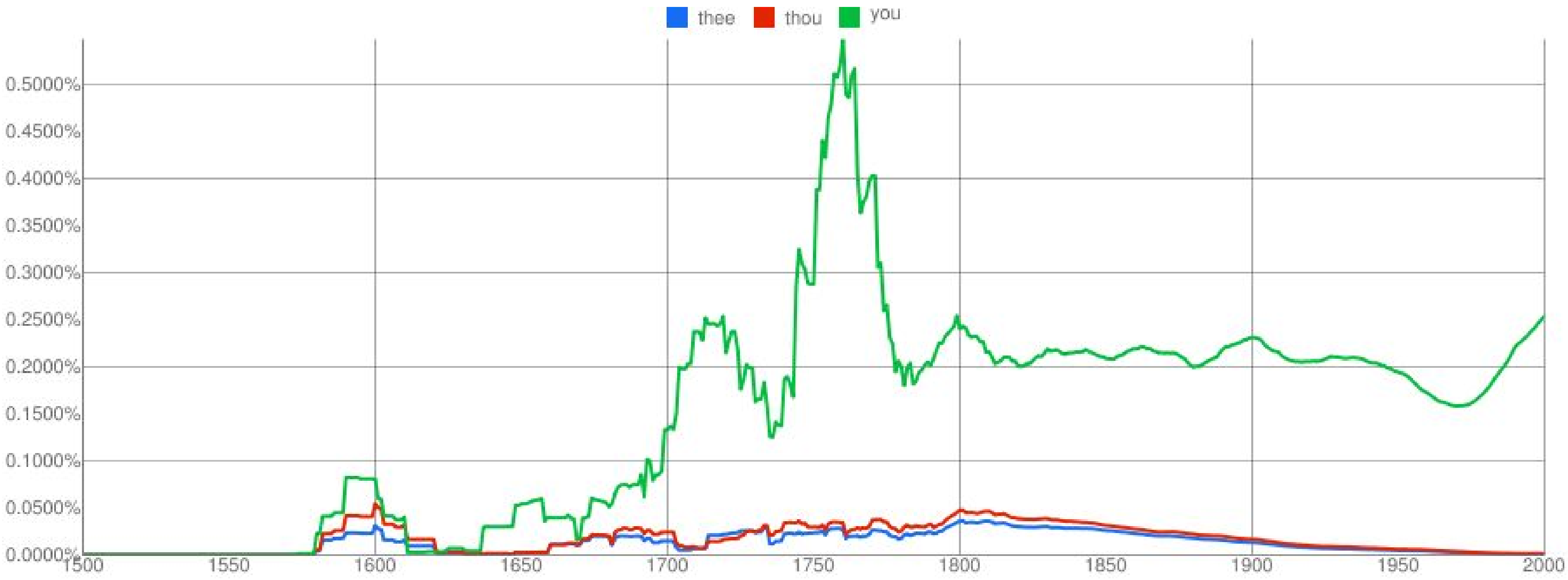
What does it actually do?

- searches a selected corpus of books for a user selected set of phrases
- select years for search
- apply smoothing to specify moving average of results returned
- search Google Books from returned set of results
- use the returned raw data to create your own visualisations, tests...

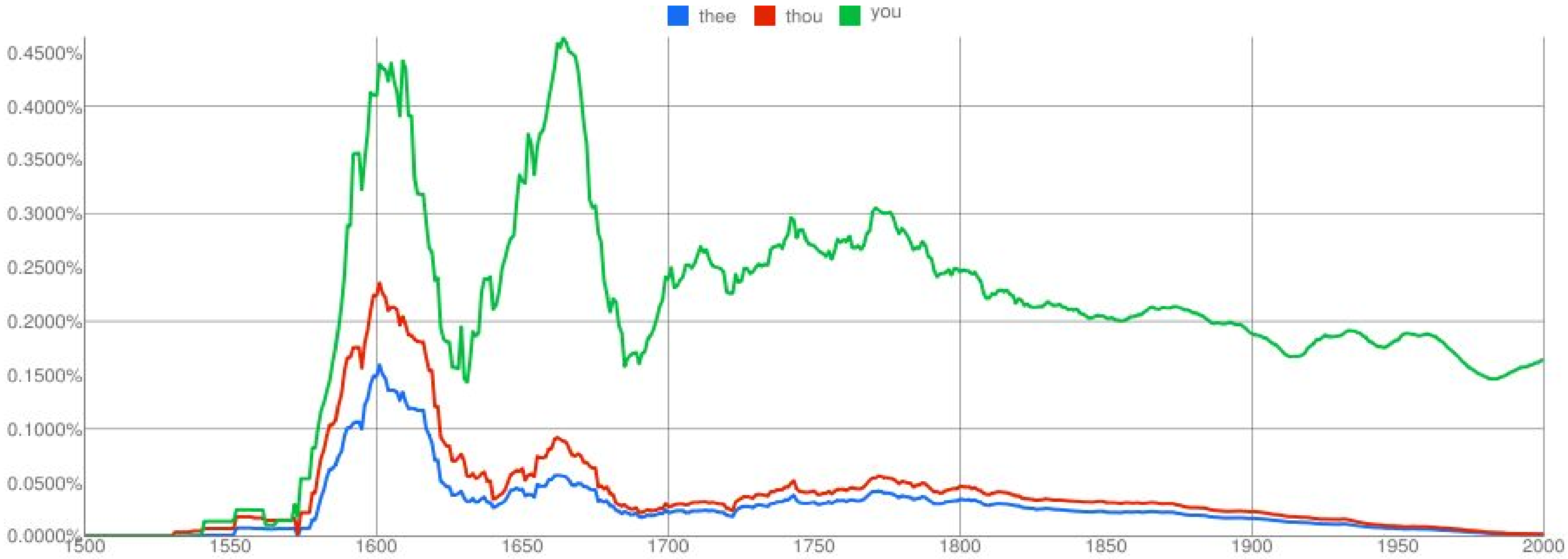
Try a few tests



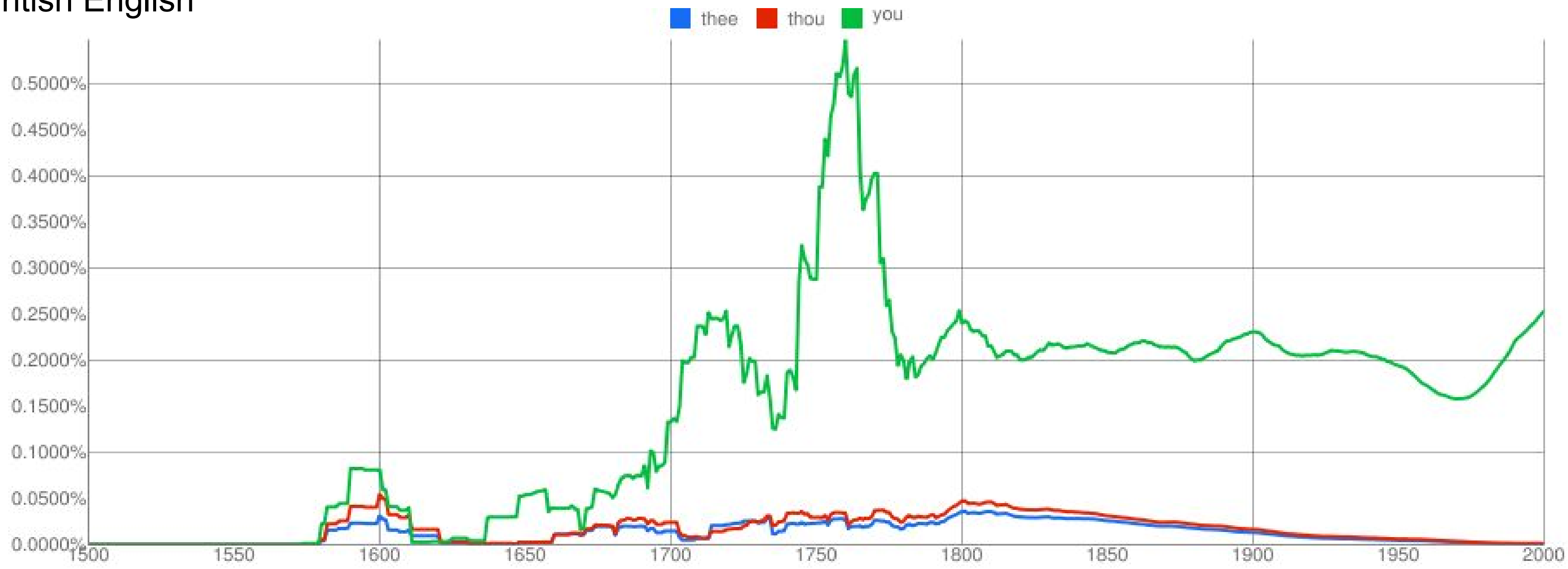
British English Corpus results for 1500 - 2000



American English Corpus results for 1500 - 2000



British English

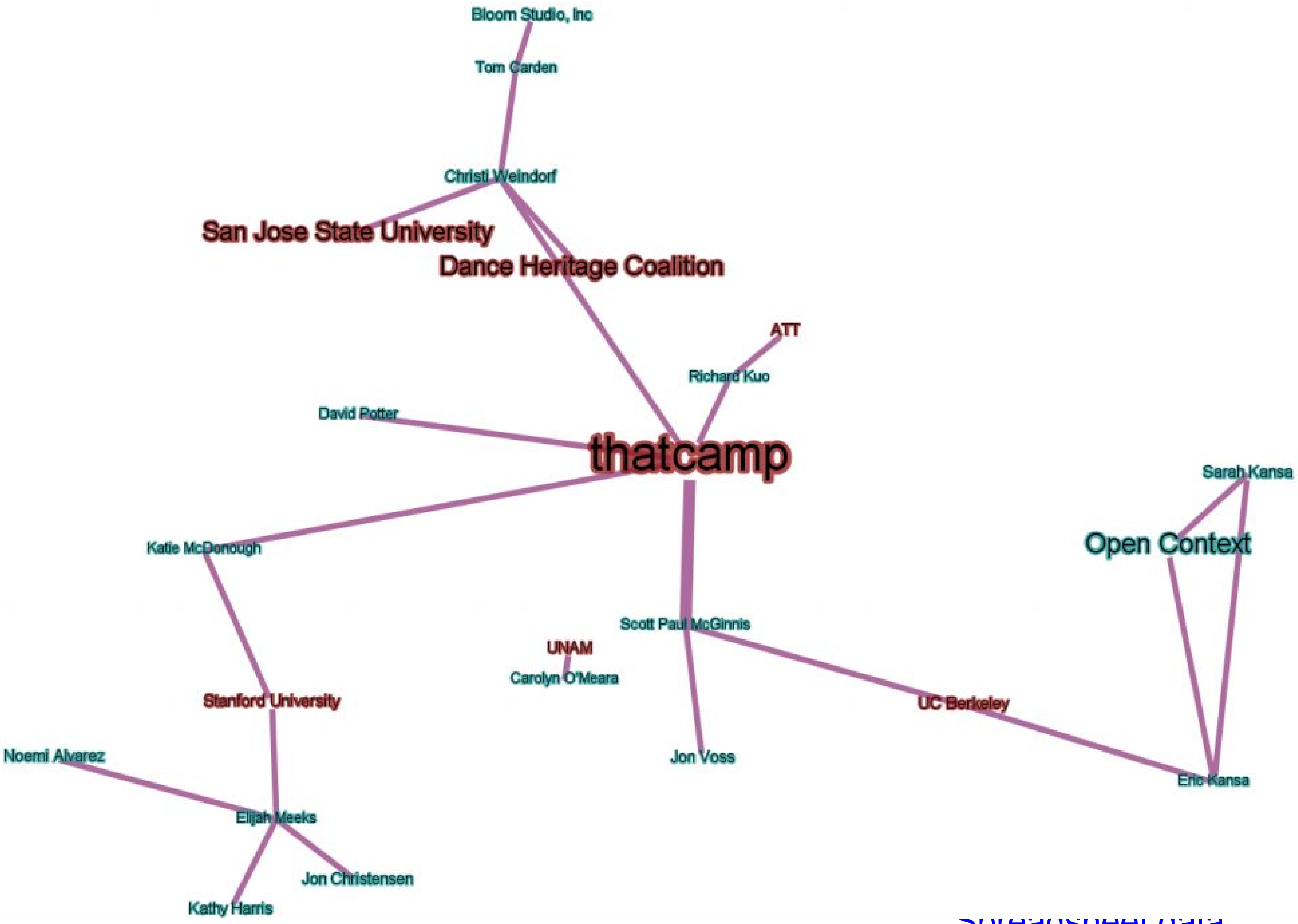


American English

Gephi

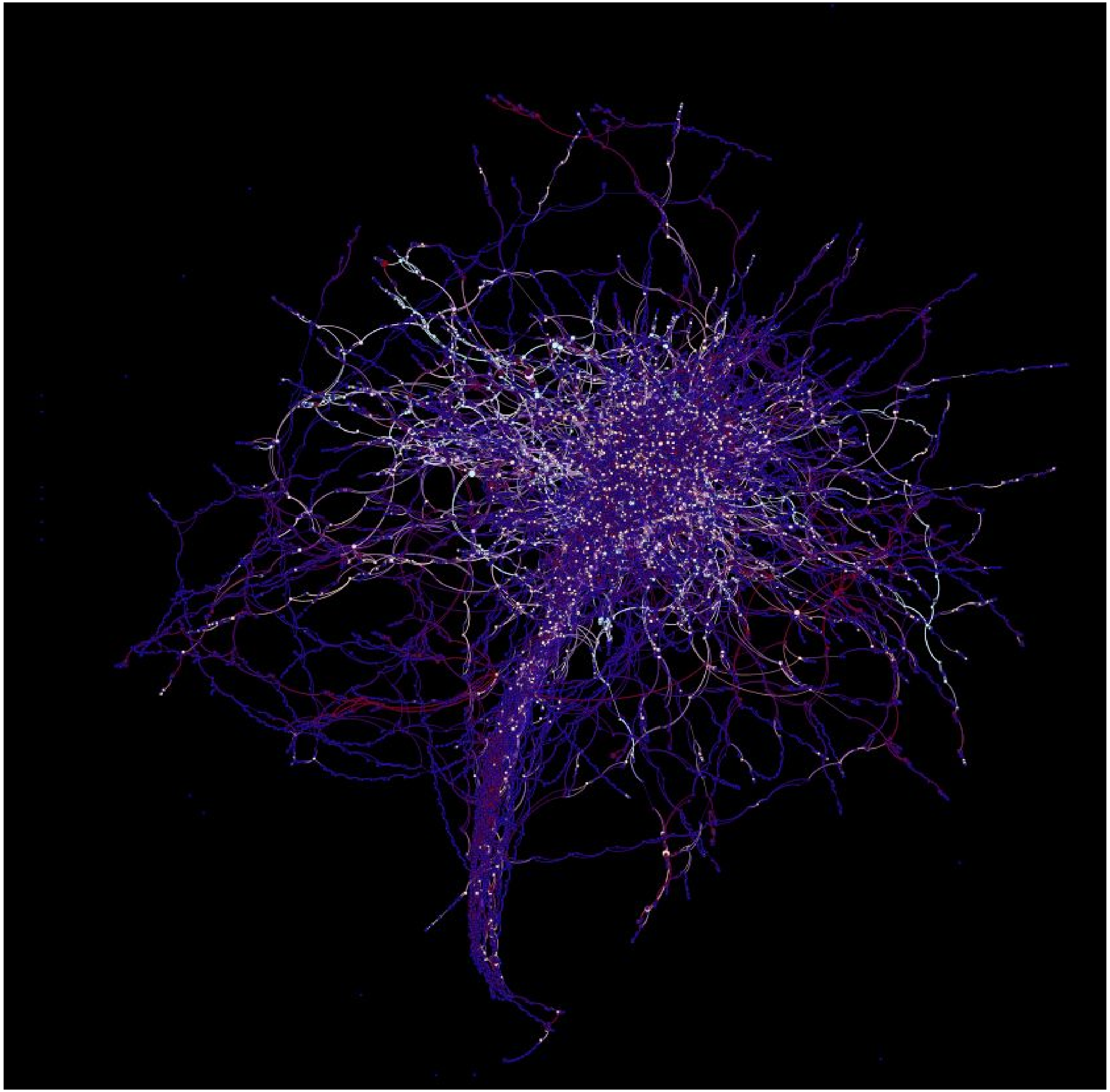
What does it actually do? (see https://gephi.org/about/)

- software for network visualisation and analysis
- helps data analysts to intuitively reveal patterns and trends...
- highlight outliers and tell stories with data
- displays large graphs in real time to speed up exploration
- built-in functionalities and flexible architecture for networks to
    - explore, analyse, spatialise, filter, cluster, manipulate & export



Spreadsheet data

Further Info



And now for something completely different...but somewhat relevant, and a lot of fun

Twitterology (?) - a bit of fun...

- used in diverse fields such as linguistics, sociology, and psychology
- analyse and examine language usage, patterns, location specific terms...
- immediacy and immensity
- University of Texas research into tweets and streams emanating from Libya and Egypt
- noticeable increase in usage following certain political events
- language patterns could also be discerned relative to such events

NYTimes Article on "Twitterology: A New Science"

Carnegie Mellon Article

Twitter API - what can we do?

- work with a user’s timeline
  - status updates
  - search tweets…
- manage your own account
  - post, delete, maintain tweets
  - check friends, followers, IDs…
- search Twitter
  - search feeds and hashtags
  - search friends, user IDs
  - get tweets by geo location

and lots more… Twitter Stream | Twitter API

Manipulating Data - Comparison Operators

- comparison of two data values to see which is bigger than the other
- comparison operator returns either true or false
- often used in conditional expressions

==	equal to	x == y
===	identical and of same type	5 === 5
<> or !=	not equal	x != y
!==	not identical	x !== y
<	less than	x < y
<=	less than or equal to	x <= y
>	greater than	x > y
>=	greater than or equal to	x >= y

Manipulating Data - Comparison Operators

String comparison

- comparing strings using this method can be slightly confusing
- computers use a number to represent a letter
- ASCII table contains numbers for characters

A > a

- compare multiple characters in a string

aA > aa

aA > a

Aa > a

- as soon as the computer finds values that differ, it will cease the comparison and return a boolean result

Manipulating Data - Using Boolean Operators

- comparison operators return Boolean values, either True or False
- as with numbers and strings, Boolean values can also be manipulated
- boolean operators also return either True or False results
    - Not, And, Or, Xor
- PHP known as 'Logical Operators'
    - And, Or, Xor, Not

```
x and y
x or y
x xor y
```

Manipulating Data - Using Boolean Operators

Java

!	NOT operator
&	AND operator
|	OR operator
^	XOR operator
||	short-circuit OR operator
&&	short-circuit AND operator
==	EQUAL TO operator
!=	NOT EQUAL TO operator

12 < 6 && 18 > 1
12 < 6 || 18 > 1

Manipulating Data - Using Boolean Operators

NOT operator

- takes a boolean value and converts it to the opposite
- True to False, and vice-versa

Not(3 > 2)

Manipulating Data - Using Boolean Operators

AND operator

- takes two boolean values and converts them into a single boolean value
- if boolean values are true, the And operator will return true
- if a False boolean value exists, the And operator will always return false

1.

True	True	
True	False	
False	True	
False	False	

2. (3 > 2) AND (3 >= 18)

Manipulating Data - Using Boolean Operators

OR operator

- takes two boolean values and converts them into a single boolean value
- both boolean values are false, it will return a false value
- otherwise it will return a true value

1.

True	True	
True	False	
False	True	
False	False	

2. (3 > 2) OR (3 >= 18)

Manipulating Data - Using Boolean Operators

XOR operator

- XOR is an exclusive OR operator
- converts two boolean values into a single boolean value
- both values True or False, XOR returns a False boolean value
- if one value is True and the other is false, XOR returns a True boolean value

1.

True	True	
True	False	
False	True	
False	False	

2. (3 > 2) XOR (3 >= 18)

Data

- every computer program needs to store data
- dump data into a variable
- need a separate variable for each chunk of data
- we can also store data together, for example
    - arrays
    - structures

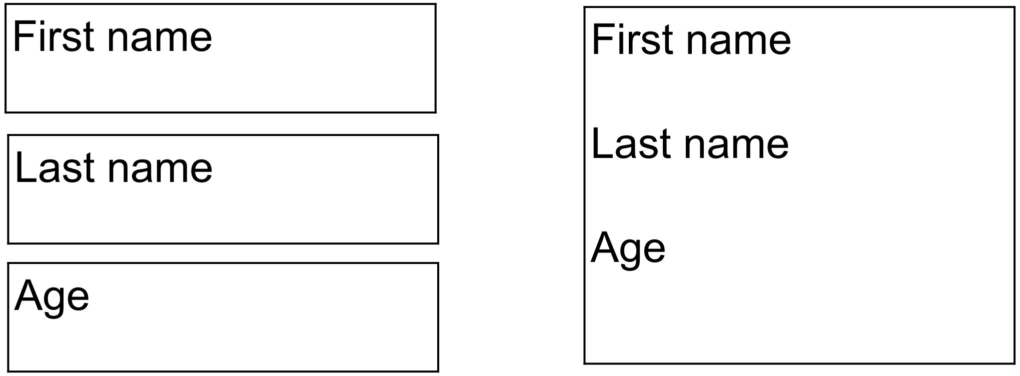
Data Structures

- group separate variables together
- store multiple variables within another variable
- user-defined data type
- cannot use a structure until you declare a variable to represent the structure

```
int[] anArray;
```

```
$students = array();
```

Data Structures



First name

Last name

Age

First name

Last name

Age

Data Structures - storing & retrieving data

- identify the variable that represents that structure
- identify the specific variable inside the structure to use

Create array: $students = array();

Insert a value: $students[0] = "Emma"

Output a value: echo $students[0];

Example

Data Structures - storing & retrieving data

```
<?php

$students = array();

$students[0] = 'Emma';

echo $students[0];

?>
```