"Did Shakespeare write the plays attributed to him?"

(from Amir Andwari, Stephanie Collins, Ellen Roberts, and Andressa Gomide)

Narrowing the question, and research aim

- HUGE question

- Shakespeare in relation to his contemporaries?
- Test some existing collaborations?
- Effects of genre?

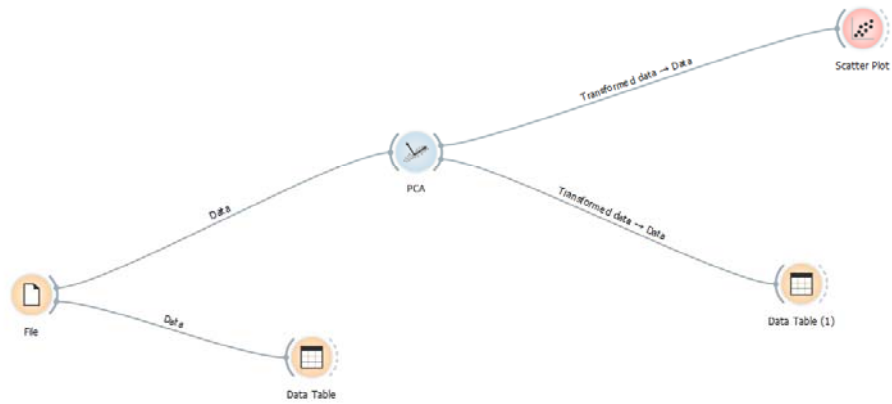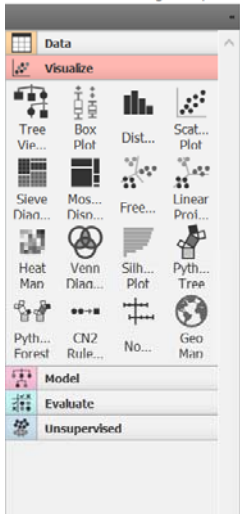- Relationship between different EME plays based on 50 most common bigrams

Method (1)

- Unsupervised method
  - Considered training texts
  - Time restraint
  - Principal Component Analysis (PCA) based on top 50 frequent Bigrams

- Corpus build
  - Size?
    - **Visualizing English Print (VEP), Expanded Drama 1660 (1,292 texts)**
    - **English Verse Drama (556 texts)**
  - Texts?
    - All of EME drama?
    - Professional plays?
  - Issues!

Method (2)

- Very experimental!
  - Use of R
    - Coding and software issues

  - Use of AntConc
    - Too much data

  - Use of LancsBox
    - Lack of N-Grams?

  - Use of Orange3
    - Easy to use!
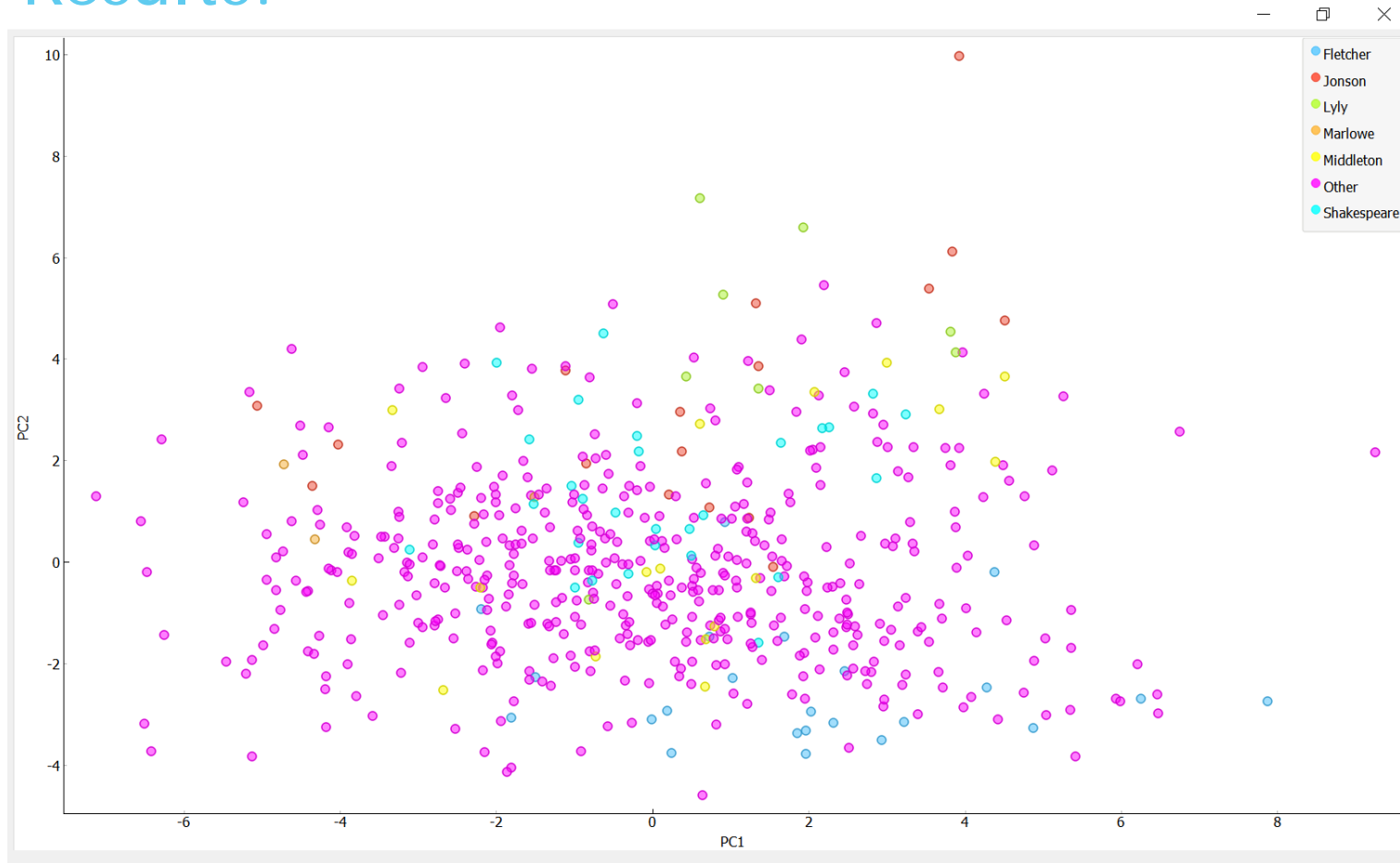
▶ Work flow in Orange 3

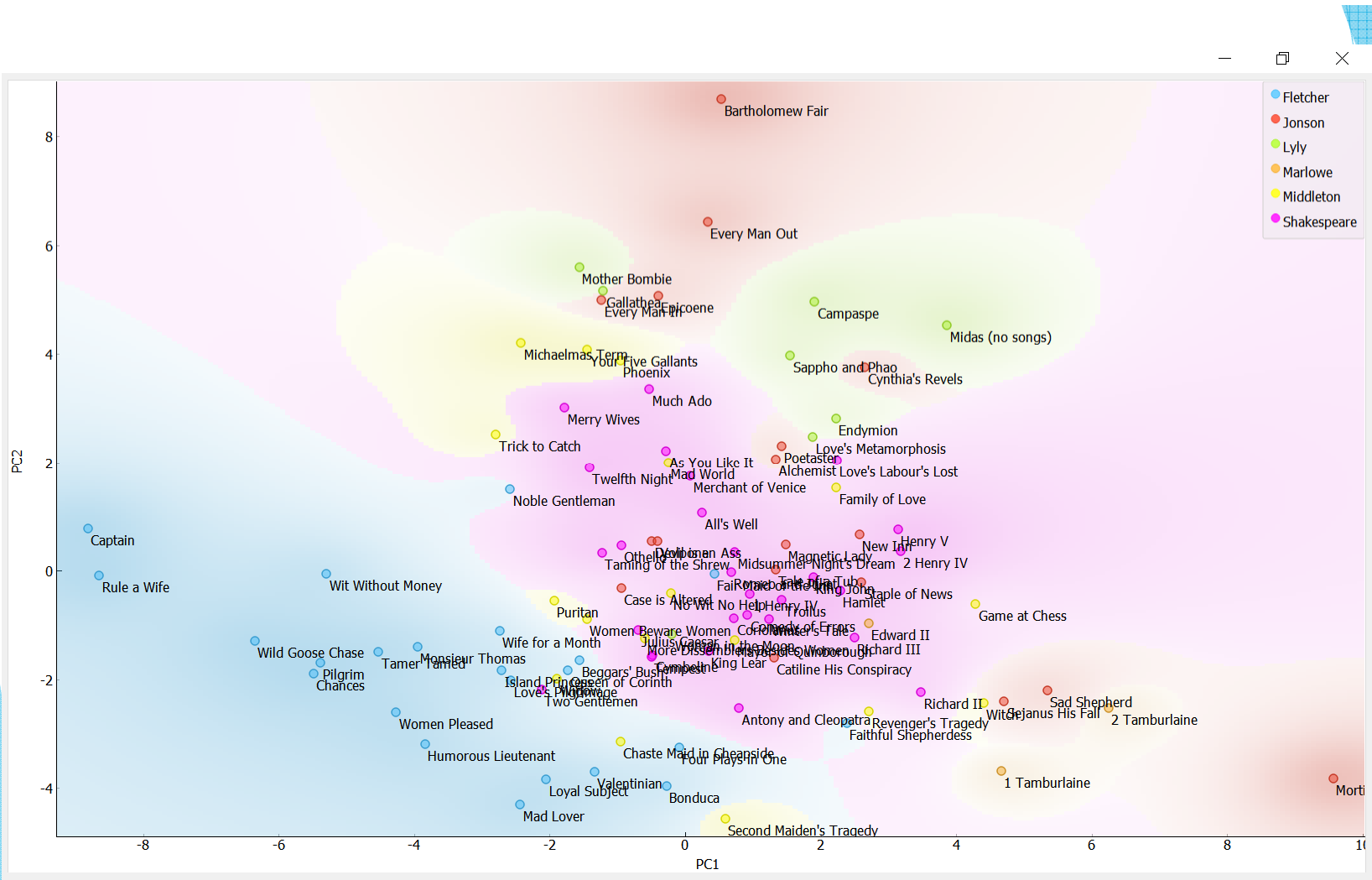# 50 Bigrams used:

- i.am
- i.will
- in.the
- of.the
- to.the
- to.be
- my.lord
- i.have
- that.i
- you.are
- it.is
- of.my
- and.i

- i.haue
- all.the
- of.a
- i.know
- i.would
- if.you
- of.your
- i.must
- in.my
- and.the
- i.shall
- n.a
- is.a

- you.shall
- for.the
- as.i
- is.the
- if.i
- shall.be
- to.my
- with.a
- such.a
- let.me
- of.his
- by.the
- with.the

- will.not
- this.is
- no.more
- to.me
- to.your
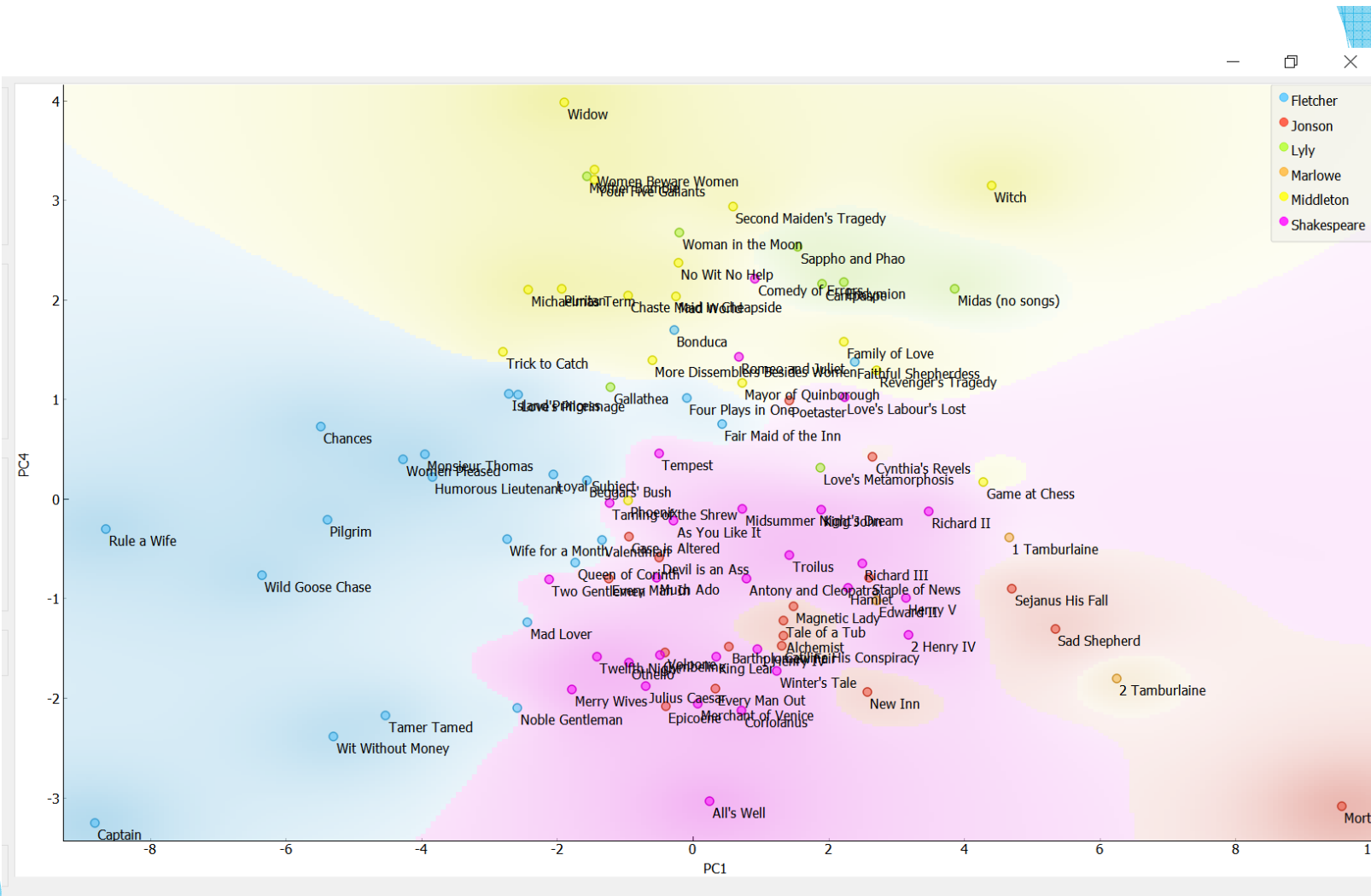- he.is
- out.of
- me.to
- my.selfe
- to.make
- for.a

Results:

- All EME plays in the corpus plotted using PC1 and PC2

Sole authorship plotted using PC1 and PC2

▶ Sole-authored plays plotted using PC1 and PC4

Conclusions:

▶ Very experimental study (based on methodology)

▶ Learnt a lot about different methods

▶ Clustering of authors seems to occur, even in this simple study