KENDRA OSBURN | 2-5-19 | IST707 | HW4 | CLUSTERING THE DISPUTED PAPERS HAMILTON VS MADISON

VI CONTENTS.	
NUMBER	ter warm ar she ar a standard and a
XVII. The Subject continued, and illui-	and the second to be an a second to be a
trated by Examples to show the	ТНЕ
Tendency of Federal Govern-	
ments, rather to Anarchy among the Members than Tyranny in	FEDERALIST:
the Head. 102	TEDERALIDI:
XVIII. The Subject continued, with farther	1. 「「「「「「「「」」」」」「「「「」」」」「「」」」」」「「」」」」」」」」
Examples. 107	ADDRESSED TO THE
XIX. The Subject continued, with farther	
Examples. 114 XX. The Subject continued, with farther	PEOPLE OF THE STATE OF
Examples. 121	NEW-YORK.
XXI. Further Defects of the prefent Con-	
, ftitution, 125	and the second sec
XXII. The fame Subject continued and	
concluded. 132 XXIII. The necessity of a Government, at	NUMBER I.
leaft equally energetic with the	Thus a setting and the setting of the set of the set of the set
one propoled. 143	. Introduction.
VYIV. The Subject continued, with an	W RTDD na namericanal
Antiver to an Objection con- cerning flanding Armies. 149	A FTER an unequivocal experience of the ineffi- cacy of the fublifting federal government, you
XXV. The Subject continued with the	"are called upon to deliberate on a new conflictation for
fame View. 155	the United States of America. The fubicit fneaks its
XXVI. The Subject continued with the	own importance; comprehending in its confequences,
fame View. 101 XXVII. The Subject continued with the	nothing lefs than the exifience of the UNION, the fafety and welfare of the parts of which it is com-
fame View. 168	poled, the fate of an empire, in many respects, the
XXVIII. The fame Subject concluded. 173	molt interesting in the world. It has been frequently
XXIX. Concerning the Militia. 170	remarked, that it leems to have been referved to the
XXX Concerning Taxation. 104	people of this country, by their conduct and example, "to decide the important queffion, whether focieties of
XXXI. The fame Subject continued. 190 XXXII. 'The fame Subject continued. 196	men are really capable or not, of establishing good
XXXIII. The fame Subject continued. 200	government from reflection and choice, or whether
XXXIV. The fame Subject continued. 20)	they are forever defined to depend, for their political
XXXV The fame Subject continued.	conditions, on accident and force. If there he any
XXXVI. The fame Subject concluded. 219	trath in the remark, the crifis, at which we are arrived, may with propriety be regarded as the era in which
THE	A that

Introduction

The Federalist Papers are a collection of essays published between October 1787 and April 1788. They were intended to promote ratification of the then-new U.S. Constitution.

Following a series of widely disseminated letters criticizing this document, three of its architects set out to defend the Constitution. Alexander Hamilton, James Madison, and John Jay decided to produce anonymously-written articles of their own, responding to popular criticisms and explaining the merits of the new framework. Originally intended to consist of around 25 documents, The Federalist Papers eventually totaled 85. Today, The Federalist Papers remain hugely influential in the field of political scholarship.

While the public eventually uncovered the identities of the three men, the authorship of the individual essays in The Federalist Papers remains a matter of some controversy. Following Hamilton's death, reports suggested that he had written the vast majority of the essays, sparking a debate about each writer's contributions. Years of scholarship led to a general consensus about most of The Federalist Papers, but the authorship of twelve essays remains disputed. Was Hamilton, in fact, the sole author of these essays? Was there another sole author? Or were they the result of collaboration between the three men?

Analysis and Models

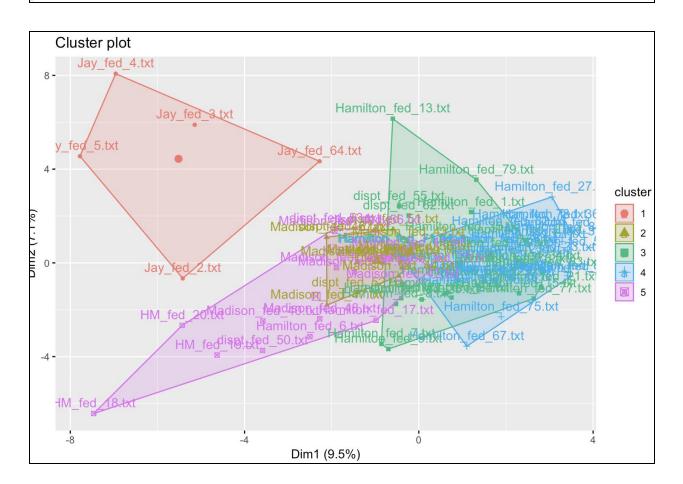
CORPUS METHOD

First, the documents were downloaded from the official website. Each document was a txt file and included the author's name or the word "disputed." Since John Jay isn't of interest to the current study, his papers were removed before the analysis began. Additionally, the co-authored papers were set aside as there is no clear indication who wrote which sections rendering the entire paper unhelpful for this exercise. Shortly after importing the documents this way, many roadblocks arose and the author of this paper decided that she, attempting to balance everything from work to school to a very ill kitten, was ok with the "done instead of perfect" approach so she could finally get some sleep.

CSV METHOD

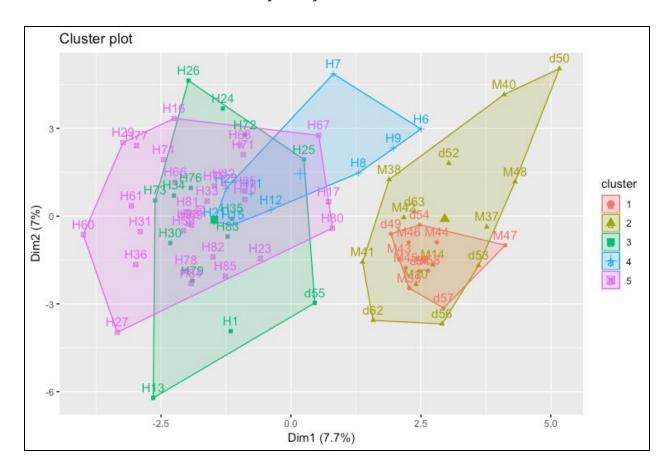
The csv was imported and the following R script was run. This is where we began.

```
fpcsv <- read.csv("/Users/kosburn/syracuse/IST707/fedPapers85_withHM.csv", header = TRUE, sep=",")
fpcsv <- na.omit(fpcsv)
head(fpcsv)
#saving author name in vector called names
author <- fpcsv$author
#changing row names to animal name
row.names(fpcsv) <- author
#removing author column
fpcsv <- fpcsv[,-1]
#removing filename column
fpcsv <- fpcsv[,-1]
#Creating 5 clusters
fpcsv5 <- kmeans(fpcsv, centers = 8, nstart = 30)
fviz_cluster(fpcsv5, data = fpcsv)</pre>
```



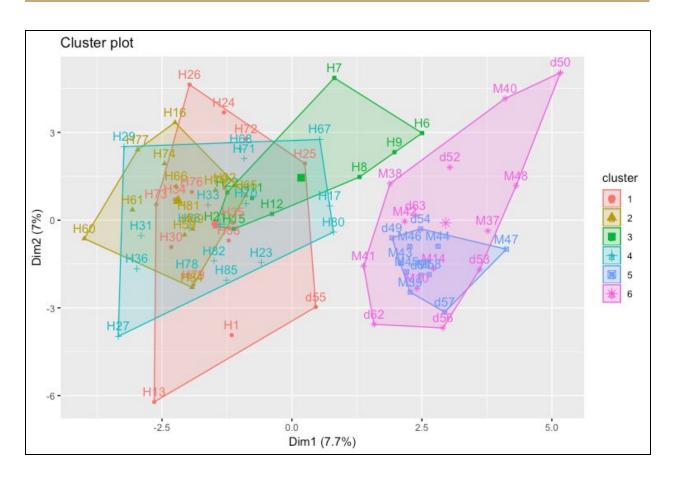
Clearly, this is too much information and more cleaning is needed. Since John Jay isn't of interest to the current study, his papers were removed. Additionally, the co-authored

papers were set aside for the initial analysis. Further, the document names were condensed to be more readable while still being unique. The letter "H" denotes Hamilton, "M" is Madison, and "d" is disputed.



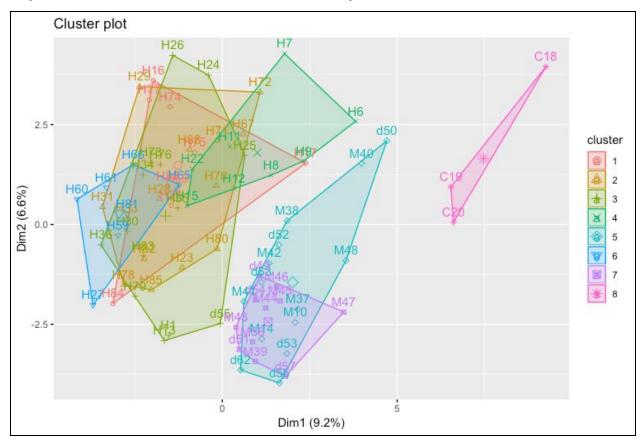
The first "Hamilton and Madison only" analysis included five clusters.

The second "Hamilton and Madison only" analysis included six clusters.



The information differences between the two clusters is minimal, meaning the elbow, or an appropriate number of clusters, was found. The author of this paper would like to note that she considered many other ways of finding the elbow, but again, despite the additional day given, ran out of time. The methods can be seen in the additional R code. She has a spreadsheet of her time spent (she tracks her time for two reasons -- a personal data project that has yet to be actualized and to see where she misjudged the amount of time something would take vs, the amount of time it actually took).

Returning to the papers set aside earlier -- the collaboration documents were re-introduced to see if any information could be gained from including them. This plot tells an interesting story of the collaboration between the two authors. When writing together, their style is



very different from both of their own, individual styles.

Results

From the clusters above, even as imperfect and unsatisfying they are to the author (who ran this model with both euclidean and cosine distance matrices, and did a bunch of other analysis not shown here) it is obvious that the two writers have a different enough style that authorship of the disputed papers can be determined. The plot above shows Hamilton in the leftmost corner occupying clusters 1 - 4 and also 6. Madison occupies plot 5 and 7 and the two in collaboration are noted by cluster 8. There is a single instance of one of the disputed papers appearing in a Hamilton cluster (d55), however, this is the result of the imperfect clustering and randomly generated centeroids, as there is still a relatively clear demarcation line that could have included d55.

Conclusion

After analyzing the results gathered, the data suggests that Madison was, in fact, the author of the twelve disputed papers. This brings Madison's total to 29 essays. Hamilton, in turn, wrote 51, with Jay providing the remaining five. Though Hamilton remains the most prolific of the three authors, this conclusion greatly expands Madison's contribution to the documents and, therefore, this important moment in history.

This is not the first time The Federalist Papers have been analyzed in the hopes of uncovering this mystery. It's a popular subject in data analysis for good reason: The essays are a seminal part of American culture, both in the past and today. By identifying the authors of each of these influential works, we can reveal important truths about the beliefs and accomplishments of our founding fathers. In addition, the world can rest easy knowing Lin-Manuel Miranda was correct in his representation of Alexander Hamilton in his musical of the same name.