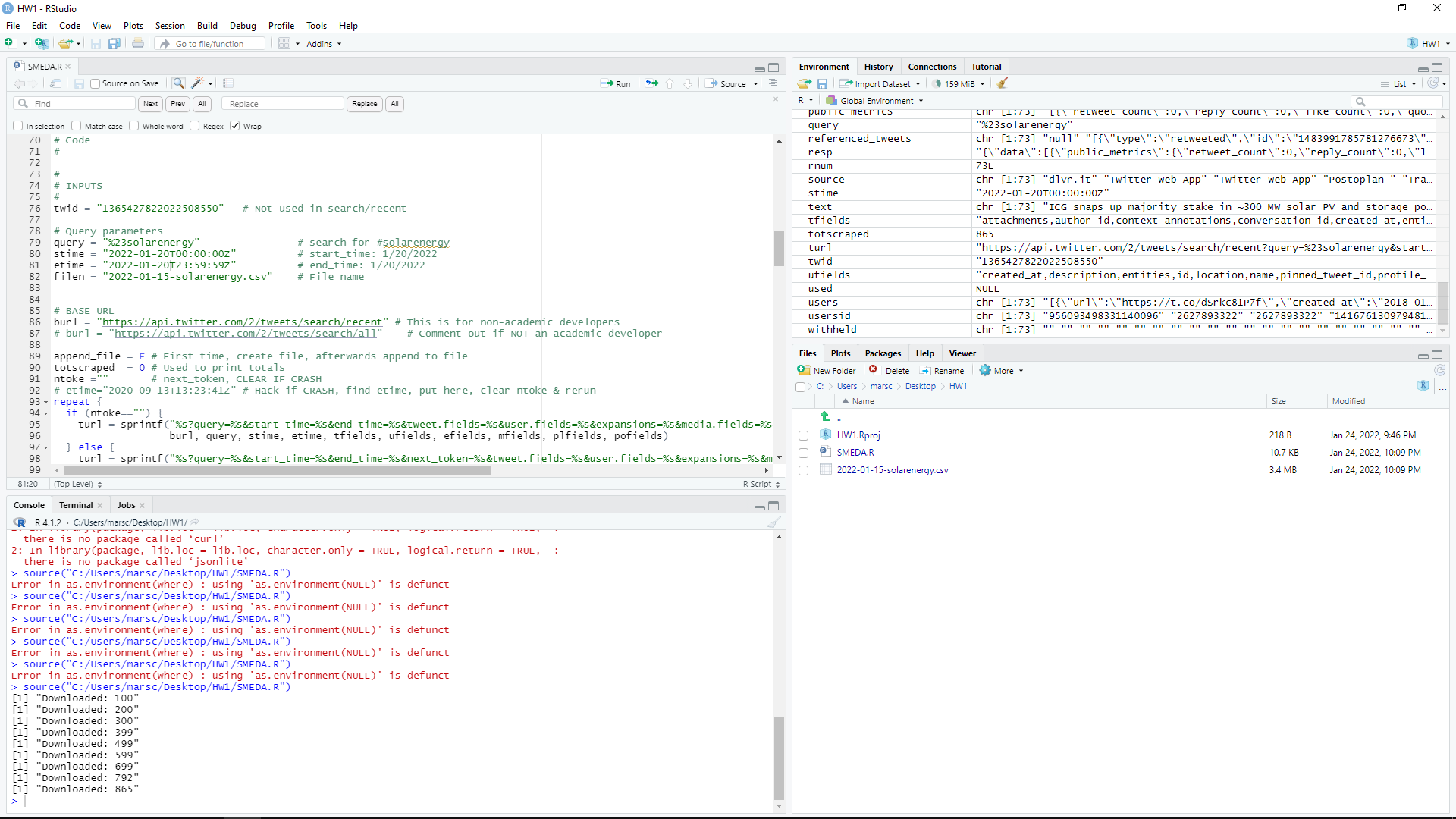
**Analytics Portfolio**

Andrew Marschall

1. **Social Media (Twitter) Data Scraping**

This data set is from twitter, scraping data from the hashtag #solarenergy.

The results show the amount of tweets scraped from 2022-01-20



1. **Data Cleaning**

This data set is from twitter and the tv show #ICanSeeYourVoice

The results show that we are removing data so we can analyze it more clearly. In this case we are removing words from tweets over the past seven days to transform the data. This makes the word clouds easier to read. We are using the gsub and grep functions to make this possible.

Graphical user interface, application

Description automatically generated

1. **Word Clouds**

This data set is still from #ICanSeeYourVoice on twitter.

The results show which words were used the most frequently within tweets over the past seven days.

Text

Description automatically generated

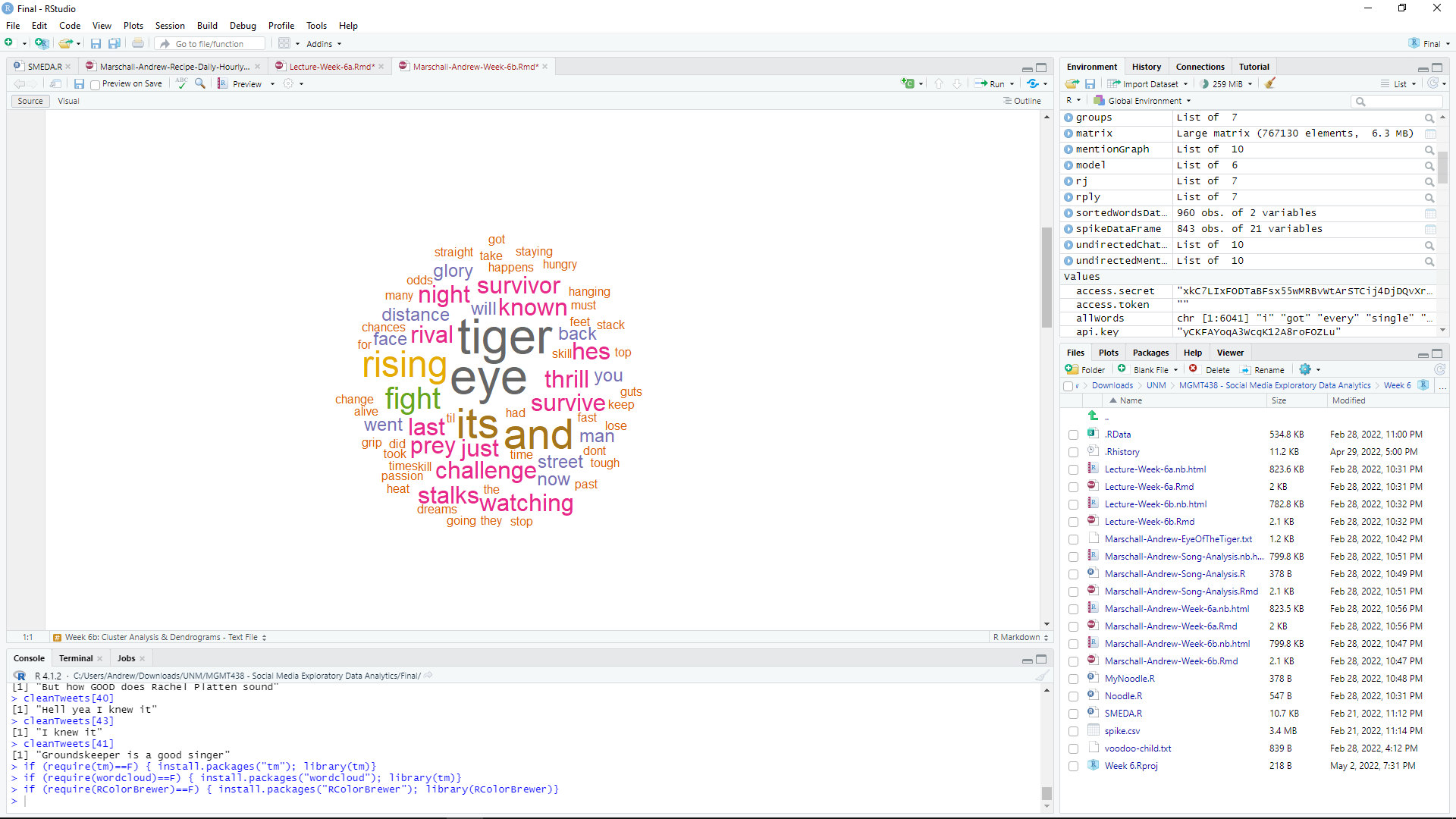
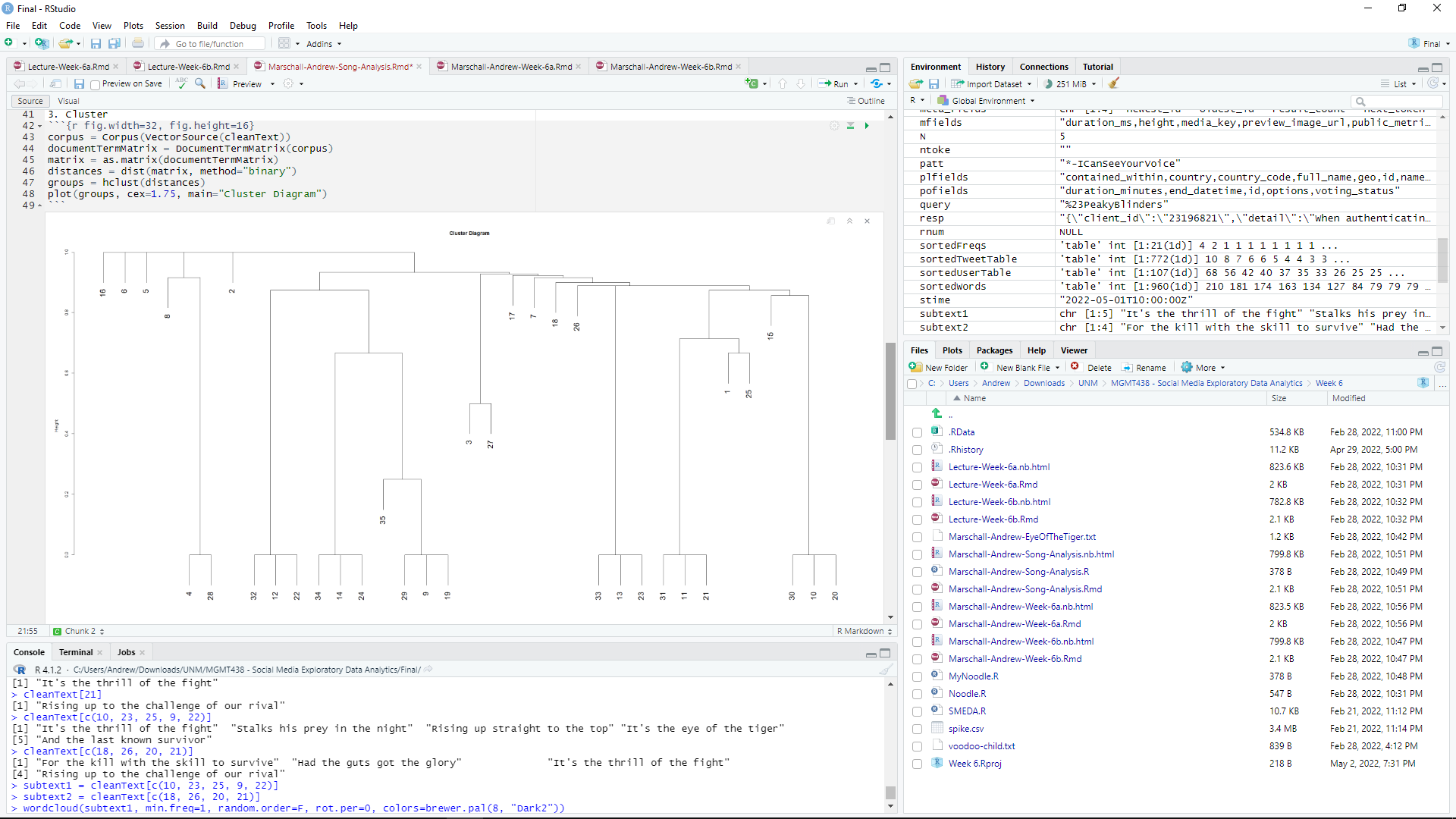
Graphical user interface, text

Description automatically generated

1. **Hierarchical Clustering & Dendrograms**

This data set is from a text file with the lyrics from Eye Of The Tiger.

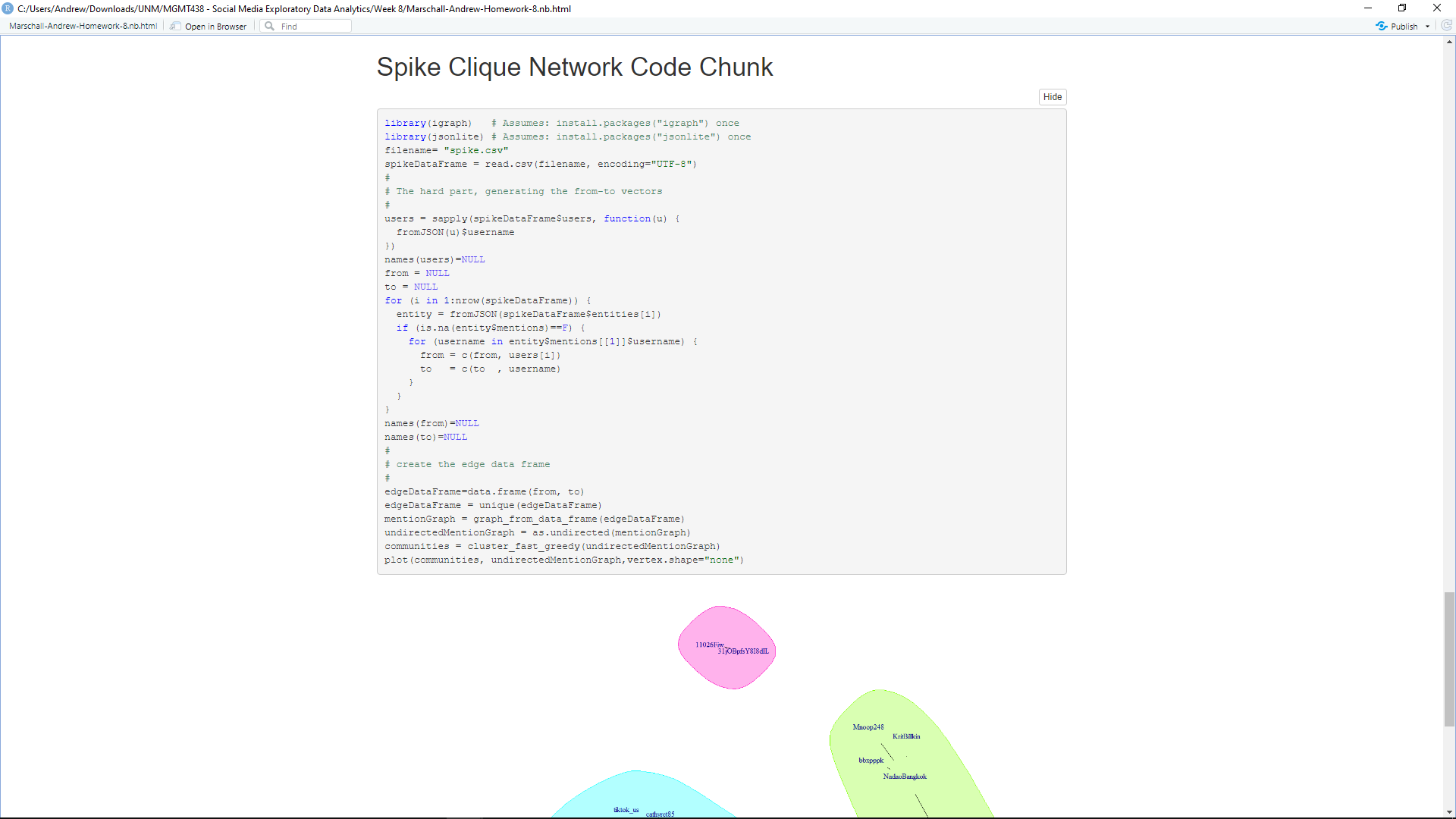
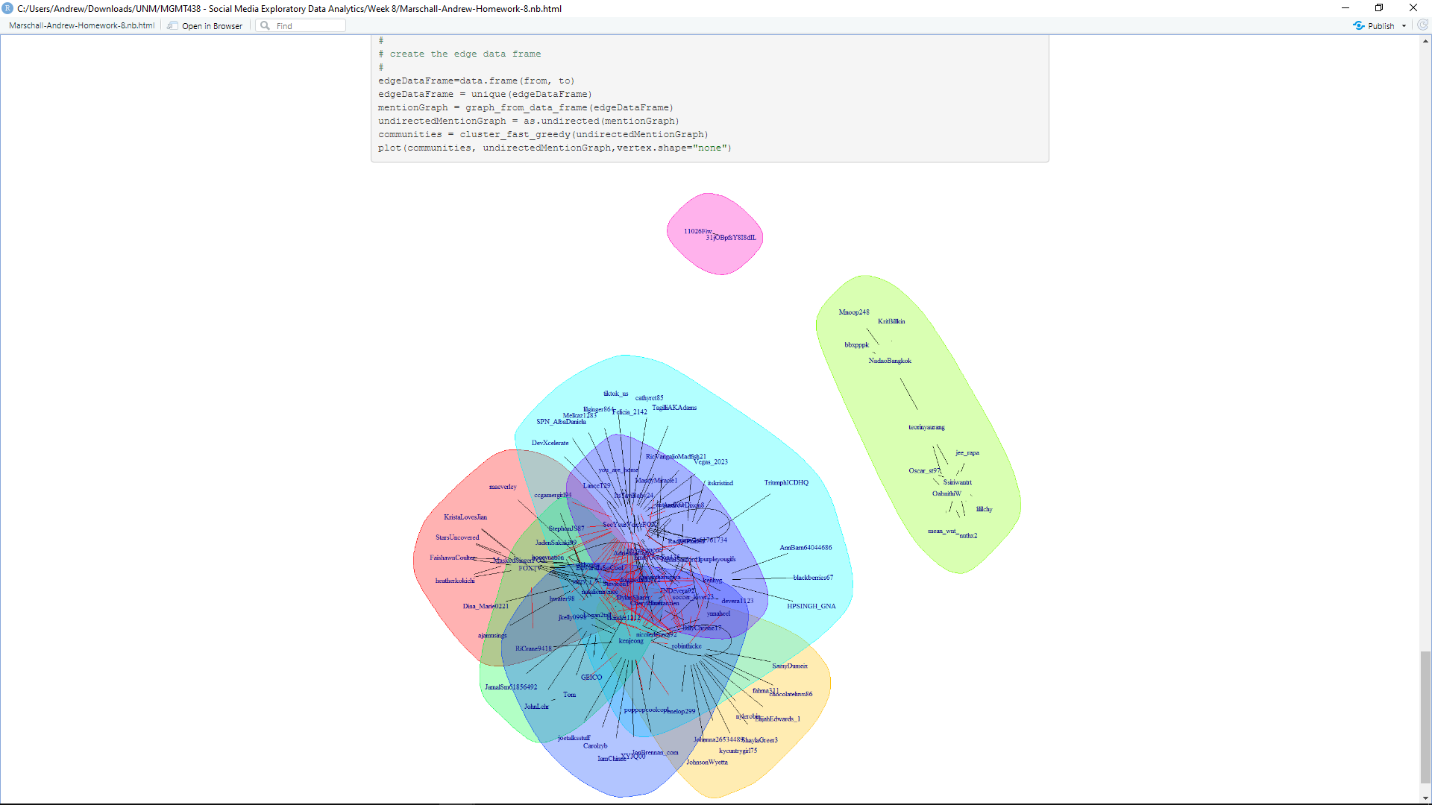
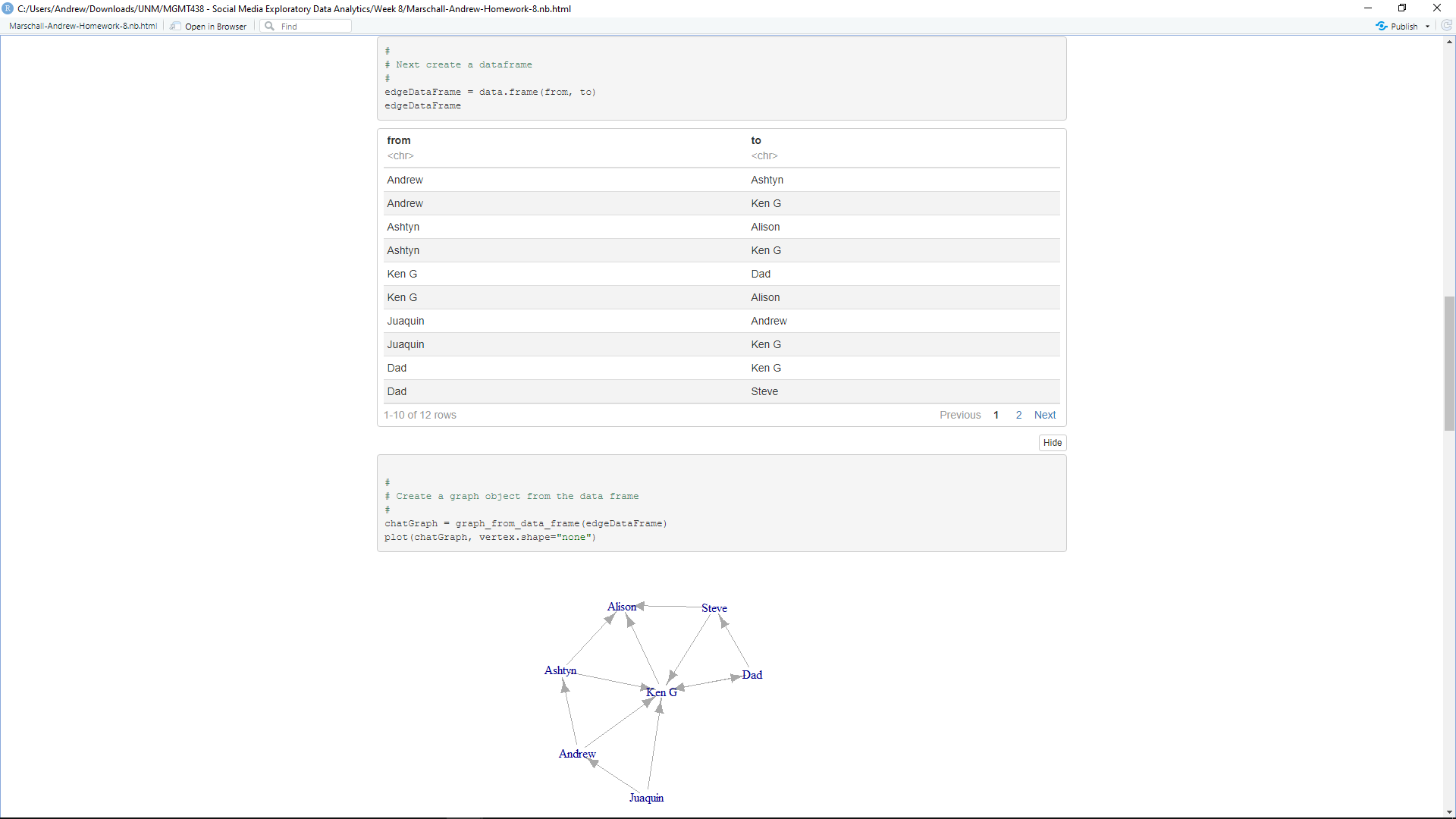
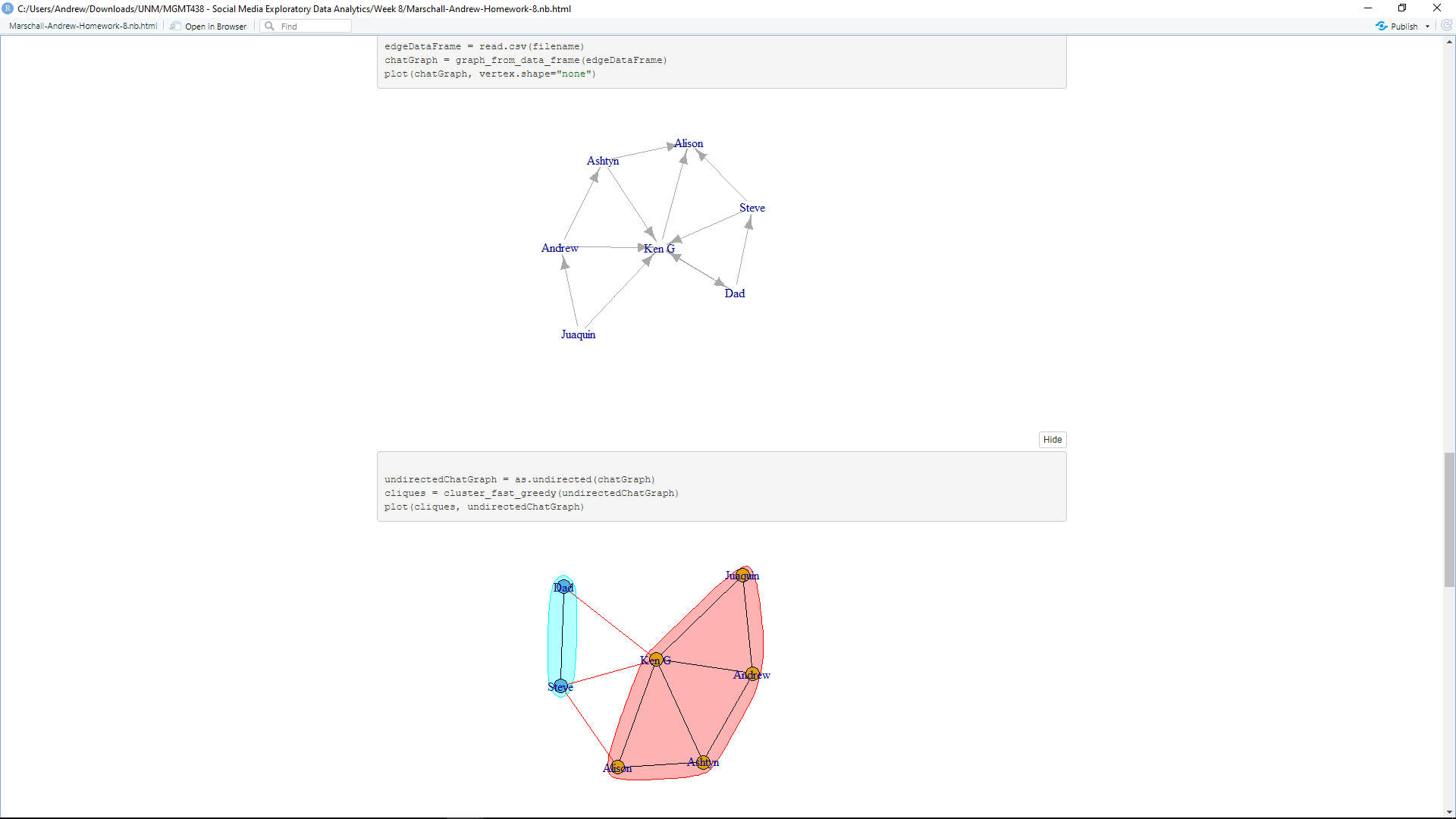
The results show the song lyrics converted into a word cloud as well as clustering the data as a dendrogram.



1. **Social Networking Analysis**

This data set is from a basic chat network of people I know, as well as from a twitter spike showing a more complex network of users.

The results show who is messaging who, on both a smaller and larger scale.



1. **Linear Regression**

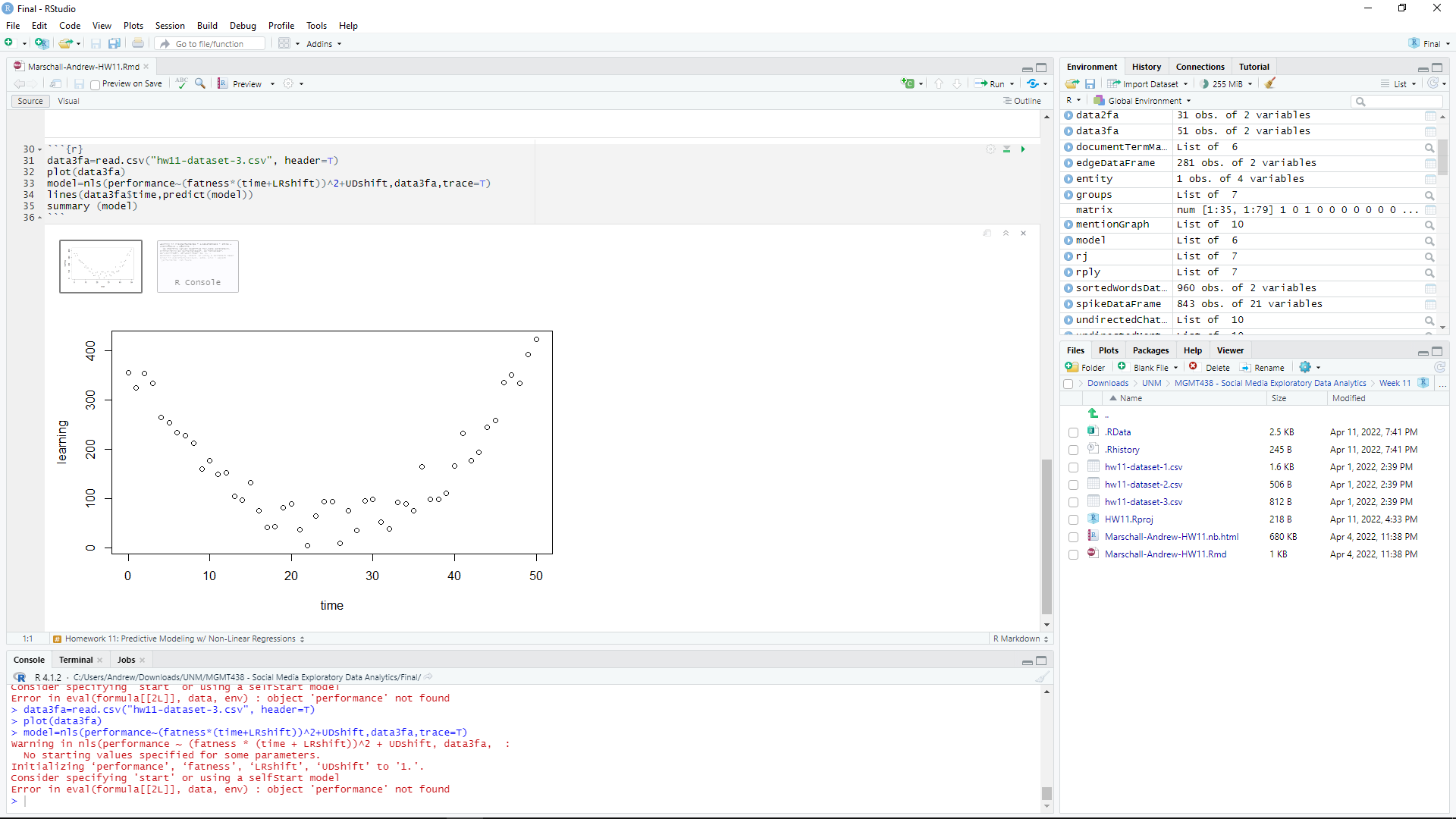
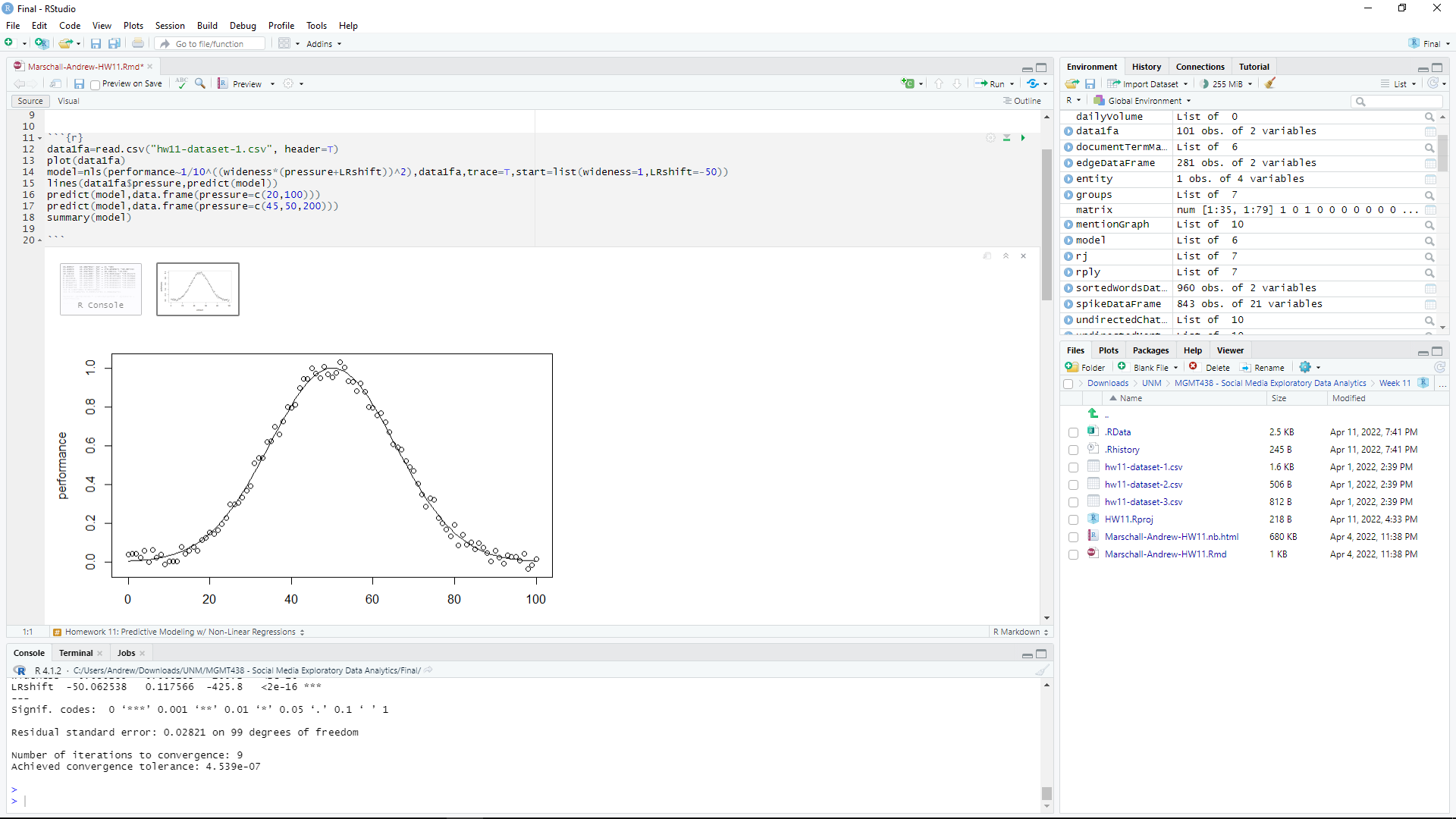
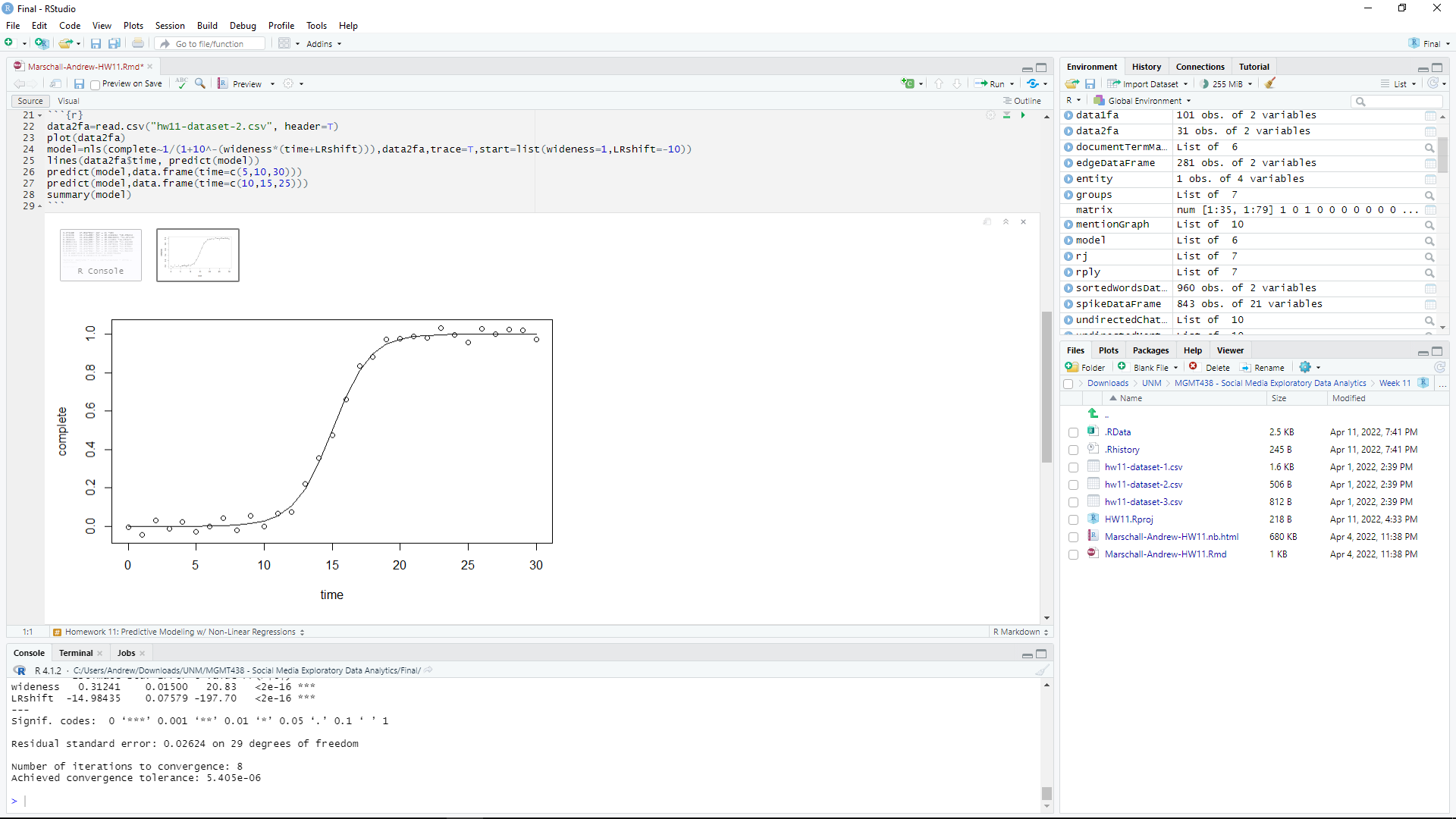
This data set is from…

The results show that…

1. **Non-Linear Regression**

This data set is from three csv files with data on variables such as pressure and performance, or time and completion / learning.

The first result shows that as pressure increases, performance will also increase up to a certain point.



1. **Neural Network**

1. **Bayesian Network**

