Columns to be the variables.

How many when yes How Many when No

Use the R with this data. You want a probability metric whether the outcome will be yes or no.

Try and apply TTests to say with certainty whether we can say our data is good

If we are totally wrong, we should get about 15 right. Look over all of the vocabulary. How?

Logisitic Regression in the R and take the values and sum the weights

Oyez files are consistent

Find out who is fighting for what side.

Independent variables should also be the classifications of the case

We will be using xml (similar to html)

Write a fxn that takes the xml and puts out all of the speaker names

Then have something that goes through and changes all of the names

The flow

Expand to all 120 5-4 cases

Expands to all the words that he uses

Check significance

Expand to all cases

Check

Can also look at who he is talking to when he says these words