**Thesis Outline**

Methodology

* Data creation – use two word phrases only, see how that goes, justify because of stop list. Also, Excludes the speaker, president, and procedural hoohaa (Gentzkow). Also not using extensions

Methodology

* Data creation – use two word phrases only, see how that goes, justify because of stop list. Also, Excludes the speaker, president, and procedural hoohaa (Gentzkow). Also not using extensions

<http://www.mysqlfaqs.net/mysql-faqs/Indexes/What-are-best-practices-to-pick-columns-to-index>

I Had to go through an change all similar names in the excel file. Boehner (speakers) not counted