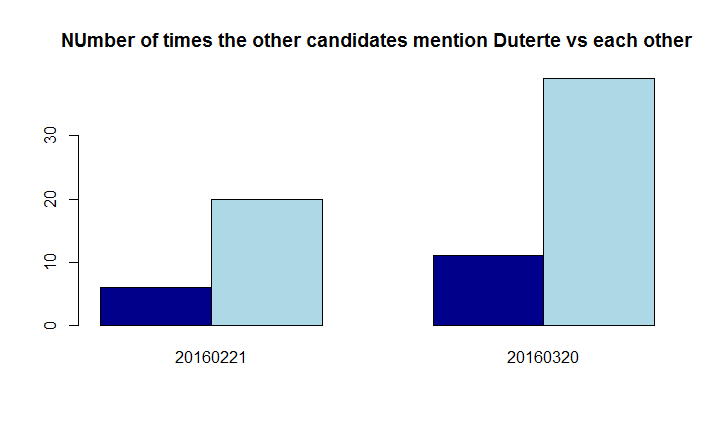
**The Candidates in Their Own Words:**

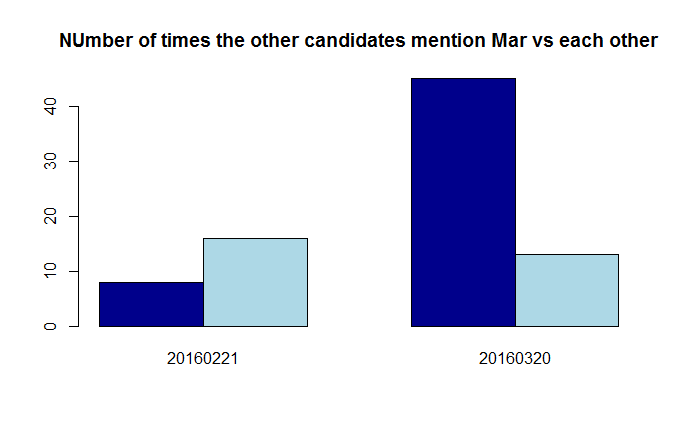
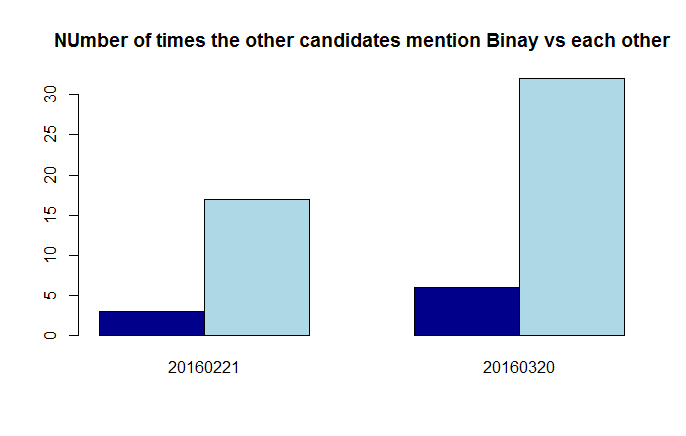
**A Textual Analysis of 2016 Philippines’ Presidential Debates**

The recent Philippines’ Presidential Debate (Pilipinas Debates 2016) drew unprecedented amounts of viewers and the usual lot of controversies and soundbites in the media. In our polarized media landscape, ensuing political analysis always suffer from political bias. Whether you trust the media, you will get a very different take on what was said.

So far, the candidates have spoken thousands of words in the two (2) debates held by a couple of TV networks. In this simple textual analysis, I turn all of the words into data and make inferences about the candidates’ characteristics that would otherwise be too difficult to discover.

1. **First, let us answer the following question: How often does a certain candidate get mentioned by name by the other candidates VS how often do the other candidates mention each other’s name.**





On the first debate, the candidates really did not refer to each other much at all. Things change around the second debate, where the candidates refer to each other significantly more. Three (3) candidates stood out that got mentioned by name by the other candidates.

Out of these trio, the other candidates collectively referred to *Mar* more often than they refer to each other. In particular, they throw jabs to each other 13 times in total (light blue bar), where as they ganged-up on *Mar* 45 times! (dark blue bar)

1. **Let’s now turn our attention to the words themselves.**

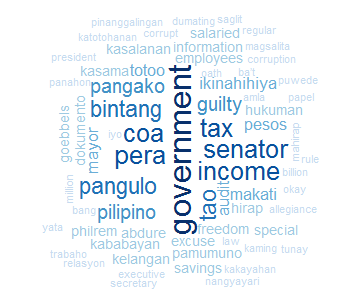
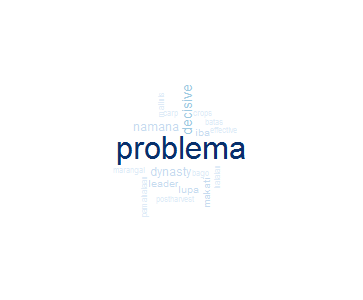
**For each debate per candidate, what are the top 100 most words? (in no particularly order)**

The candidates’ vocabulary are further simplified by removing common words that convey little meaning (e.g. prepositions, pronouns) and stemming the remaining, more substantive words.

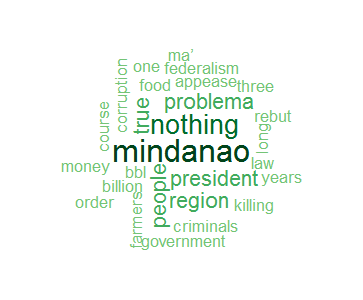
*Poe* and *Mar* were the most active during the first and second debates (60% of the total words thrown). They also both used longer words (around 7 letters per word). With *Binay*, these three candidates used longer words than the rest of the candidates.

In terms of the diversity of words used (i.e. unique words per 1000 words), this same set of candidates + *Duterte* used fewer words except *Miriam*. She was the least active and have the most diversified vocabulary among the five candidates. It does not came as a surprise, considering that she was not around during the 2nd run of the debates.

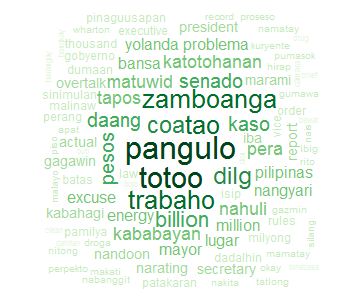
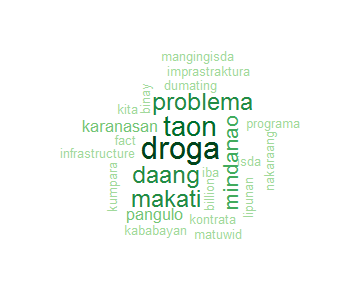
**BINAY’s WORDCLOUD (1st and 2nd debates)**

****

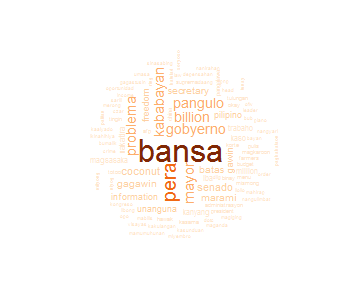
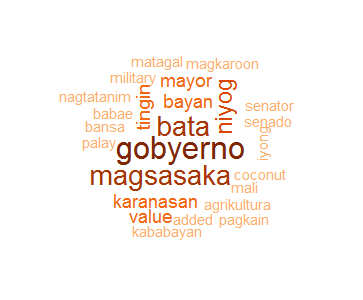
**DUTERTE’s WORDCLOUD (1st and 2nd debates)**

****

**MAR’s WORDCLOUD (1st and 2nd debates)**



**POE’s WORDCLOUD (1st and 2nd debates)**



**MIRIAM’s WORDCLOUD (1st debate only)**

