Part 2: Thinking

1. What variables would you extract for each ballot measure to study the information voters are exposed to about ballot measures? What might you be able to measure using them?

I would extract the numbering of ballot measures in order to see if it's local ballot measures, constitutional amendments, citizens' initiatives, or legislative referrals. This would allow us to measure what the ballot measure is focusing on, and this would serve as an ID on the specific ballot measure we are looking at.

I would extract the number of arguments for and against the ballot measure. This would allow me to measure the engagement of this ballot measure, with more arguments sent in, the more engagement, which likely means voters are more exposed.

Each argument for or against also includes who wrote that argument and where they are from in Arizona. By seeing which municipalities have the most people talking about a ballot measure, we would be able to measure if some municipalities are getting more exposure to ballot measures than others.

2. How would you structure that dataset you create from the voter guide and why?

I would structure the dataset by ballot measure, with each row containing a new ballot measure, with the variables ballot measure numbering, number of arguments for, number of arguments against, and then a count for the number of people in each city that are for and against the proposition. This would allow me to easily aggregate counts for and against arguments for each city in total, as well as for the type of ballot measure.