**Design Plan Worksheet**

**MSDS 6110 Fall 2015**

Please fill out the following worksheet and submit it to the course website by September 20 at 11:59 p.m. in your local time zone.

**A.    Type the title of your project in the space below. The title can change as you learn more about your data. However, this will be the title for the purposes of grading.**

God’s Will: The Southern Media Influence on The Kim Davis Situation

**B. Please type the names of the group members below. By having your name on this form, you signify that you have contributed to and approve of this design plan.**

1.   Kevin Durham

2.   Adam Soto

3.   Cory Nichols

4.   Kareem Williams

**C. Type the research question below. The research question must be a hypothesis that is testable with data**

Our alternative hypothesis, based on anecdotal evidence, is that southern news outlets are more biased to support Kim Davis, a county clerk who refused to give out marriage licenses to gay couples. Therefore, we would expect to see a greater bias in the southern media in support of Kim’s religious position.

We intend to address this primary question:

Are southern news sources in Texas, Alabama, Kentucky and Georgia more biased in their coverage of the Kim Davis story than other news sources in California, Washington, Oregon and the Northeast?

Ho: Southern <= Other

Ha: Southern > Other

We will also answer the following **OTHER** questions:

In the presence of region, does the reach of the news outlet matter? Are regional outlets more biased than local outlets?

Is the tone of the article relevant? If it is, do biased articles tend to be more negative or positive in tonality?

Is the location of a legislative body or university relevant? This may or may not make the cut depending on the quality of articles we are able to gather based on random selection.

Using binary logistic regression, we will test for bias in regional and local news coverage of the Kim Davis story.

**D. Who or what is the target population for the study?**

Domestic, English speaking news outlets covering the Kim Davis incident that are regional or local in their coverage areas.

**E. Who or what are your subjects? How will you randomize items to different treatment groups (if appropriate)?**

Our subjects are individual news articles related to Kim Davis prior to her incarceration (prior to September 3rd, 2015). Our study doesn’t contain treatment groups as it is an observational study. We choose prior to incarceration to avoid assumed bias due to Kim Davis’ incarceration.

**F. What is (are) the response variable(s) for your study? Is each one categorical or quantitative? How do you plan to measure/gather this variable?**

Bias as a binary categorical value. We are only measuring bias as a binary value on purpose, not levels of bias, as we are only interested in a **presence** of bias. We plan to determine bias via a five-point method to determine if an article is biased (binary value of 1) or not.

1. Does headline predisposition reader?
2. Is the article unbalanced from an evidence standpoint?
3. Does the article directly slander or compliment the subject or important interests associated with subject matter? (e.g. one example was LGBT activists were called ‘militants’)
4. Are the visuals overly positive or negative?
5. Overuse of buzzwords and categorizations? E.g. LGBT, Christian Agenda, Homosexual Agenda

If the response to 3/5 of the question above are yes, we will consider the article biased. We will also attempt to acquire unbiased, journalistic sources to review the articles and score them appropriately.

**G. Please describe your sampling plan, including the mechanism you will use for random sampling of subjects from the population. Are you going to block for certain variables?**

We will randomly select at least 20 articles from southern news outlets that are local and regional covering the Kim Davis story prior to her incarceration. We will randomly select at least another 20 articles from western local and regional news outlets (ten for each geographic region level). We intend to balance each group. We will analyze these articles for bias based on the five-point procedure mentioned previously and avoid overly-biased sources such as catholicnews.com, baptistnews.com, gaystarnews.com, etc.. If an article meets 3/5 criteria, we will consider it biased.

**H. What explanatory variables will be measured/gathered to test the hypothesis? Please fill out the table below. The first two rows are given as examples. You can delete them when you type in your own variables. The point is to agree among group members about how you will measure each variable, particularly ones that defy measurement (e.g. motivation, likability, beauty, etc). The range is “expected” because it is possible that you might find a subject whose characteristics are outside the range of what you expected. That’s OK. Add as many rows as you need. Any blocking variables should be denoted with a \*.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable Name** | **Type** | **Units** | **Expected Range** | **Definition** |
| Geographic Region | Qualitative | N/A | West, South | Geographic Region of the news outlet |
| Media Coverage | Qualitative | N/A | Regional, Local | Level of media such as local newspaper vs. a regional paper such as the LA Times |
| Tone | Qualitative | N/A | Positive, Negative, Neutral | The overall tone of the article given the response result. If unbiased, neutral. If biased, was it positive (in favor) or negative? (not in favor). |
| Locality | Qualitative | N/A | University  Legislative  None | Is the news outlet part of a university paper or located in a legislative area? E.g. a state capital? |

**I.   What problems do you foresee in the data gathering process? How do you plan to address them?**

The major problem with our data gathering process is to effectively quantify bias without implementing personal biases of our reviewers into the process. We intend to obtain two journalists and randomly order the articles, have the journalist grade the article and retrieve the results to obtain our response. The question lies within our methodology: is this the right way to truly judge bias? Further, we know we will not be exhausting all avenues associated with bias, so our variables, such as locality and tone, may not show significance.

**J. What methods of analysis do you plan to use? Explain your reasoning.**

We plan to use binary logistic regression to analyze the explanatory variables and address our hypothesis directly. We are interested only in the binary result, of whether we could consider an article biased.

Our project will look at the Kim Davis story in Kentucky. Recall that Kim refused, for religious reasons, to award a marriage license to a gay couple. There has been a continuous string of articles, news releases and plenty of television time dedicated to Kim’s position about gay marriage.

In this project, we wish to look at the articles from large and small newspapers from large and small towns from sources known to be liberal and conservative to look for bias as well as correlation between the source and level of support or criticism for Kim.