**Improving transparency and accountability in New Mexico’s lobbying practices ti**

**Abstract**

The goal of this project was to use natural language processing to summarize bills and lobbyist information to help increase transparency and accountability in the legislative process. I summarized and analyze bills passed in New Mexico from 2013 to 2020 and analyzed lobbyist campaign donations and employers from the same time period. The basis of the analysis was data acquired two sources: 1) the NM [NM legislature website](https://www.nmlegis.gov/Legislation/BillFinder/Governor_Actions) and 2) [NM Secretary of State Office](https://www.cfis.state.nm.us/media/CFIS_Data_Download.aspx).

**Business Need**

My client is the public – though lawmakers and lobbyists could use it – increase public confidence in the integrity of legislative process.

New Mexico is one of fifteen states that has minimal disclosure rules for lobbyists. It only requires a lobbyist to register themselves and their clients, and report basic spending each session and candidate contributions. It does not require reporting of their compensation or reporting on which bills they had been hired to support, oppose, or otherwise worked For citizens as well as legislators, this type information would be useful in improving transparency and accountability in New Mexico’s lobbying practices

**Data**

I used four data sources for this analysis:

1. List of bills acted on by the governor from 2013 to 2020. The list contains four attributes: Bill ID, Bill Title, Action by Governor, Bill Sponsor. There were approximately 1480 bills in the list.
2. I downloaded approximately 1450 pdfs of bills enacted from 2013 to 2020. There were a number of corrupted pdfs and broken links. The average word length of the pdf was 5,428. The average page length was two pages.
3. List of registered lobbyists from 2013 to 2020. There are approximately 12,000 records. The list contains three main attribute type across numerous columns: name of registered lobbyists, lobbyist business information, and employers by year
4. Campaign donations from lobbyists to legislators. List of all donations downloaded from NM Secretary of State website. The spreadsheet contains 25,000 records and numerous fields. For this analysis, I used name of lobbyist, year of donation and name of legislator or campaign the donation was made to.

**Tools**

*Data Collection and Cleaning*

1. Selenium was used to scrape list of all bills
2. Pdfminer and pdf plumber were used to download pdf files of all bills and transform pdfs into text documents that could be processed
3. Re and pandas were used for data cleaning (formatting) of text

*Topic Modeling of NM Bills*

1. Pandas was used to create file of text
2. Sumy was used to created shortened summaries of bill text
3. Nltk and spacy was used to preprocess text.(stop words, lemmatization)
4. Gensim – LDA was used to perform topic modeling on bills

*Topic Modeling of Lobbyist Employers*

1. Nltk and TFIDF were used to preprocess data
2. NMF was used to model topics

*Communications*

1. Tableau was used to develop dashboards for presentation of lobbyist influence results

**Process**

Timeline

Description automatically generated

Figure 1: NLP process

Figure 1 illustrates the basic NLP process employed. The results topic modeling for the bill text was used as foundation for topic modeling of lobbyists.

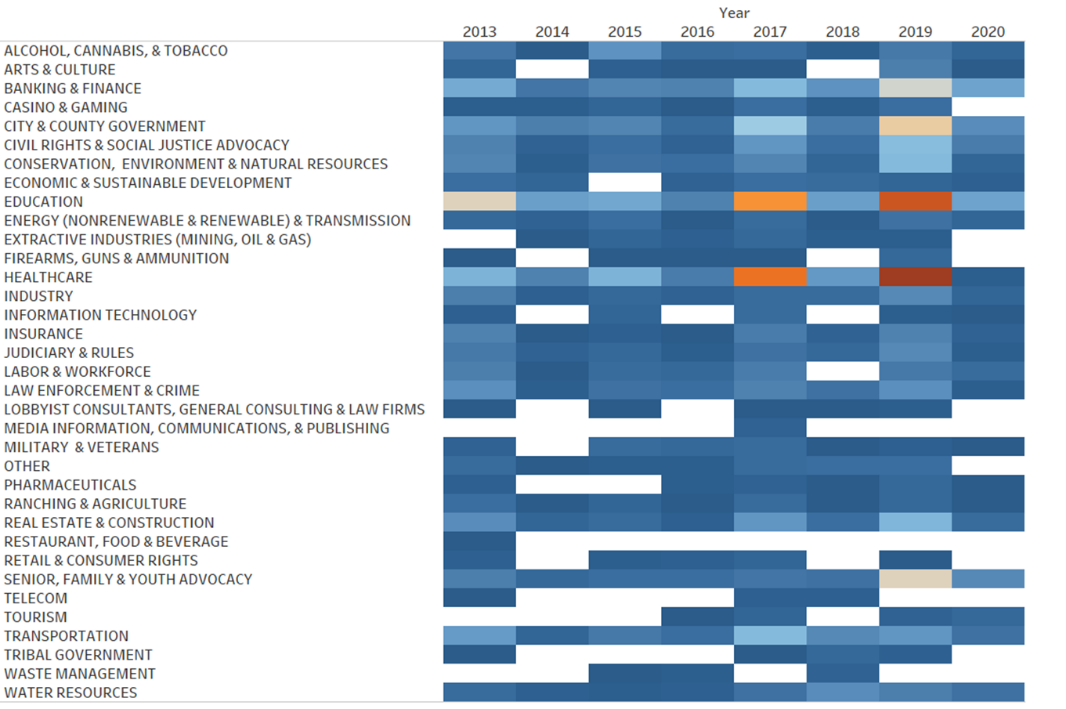
**Results**

Figure 2: Number of bills passed per year by topic

1. There were 38 topics derived from the bill analysis. Figure 2 depicts the topics along with the number of bills passed within those topics per year.
2. Healthcare and education are topics consistently have most bills acted on year to year.
3. Bill topics was combined with lobbyist employer data to show areas where there is higher potential for lobbyist influence. Lobbyist influence metric indicates areas where there is a large number of lobbyists employers for a topic than bills seen for that same area. Figure 3 shows results of this analysis developed in Tableau. Energy and extractive industries (mining and oil and gas) have considerably more lobbyist employers than bills being heard in the legislature.

Chart, bubble chart

Description automatically generated

Figure 3: Potential lobbyist influence

1. Tableau was also used to show lobbyist contributions to legislators per year alongside the lobbyist employers (Figure 4). This dashboard highlights who is contributing to legislators and their campaigns along with which industries are employing those lobbyists.

Graphical user interface, application

Description automatically generated

Figure 4: Lobbyist contributions to legislators and their campaigns