Distribution Agreement

Signature:

Hans-Kristof Nelson 04/18/2023

Exploring the Determinants of Internet Searches for Abortion-Related Information in the Midwest United States: A Google Trends Analysis

By

Hans-Kristof Nelson

MPH

Hubert Department of Global Health

Roger Rochat

Committee Chair

Exploring the Determinants of Internet Searches for Abortion-Related Information in the Midwest United States: A Google Trends Analysis

By

Hans-Kristof Nelson

-

Bachelor of Arts in Dietetics

Concordia College, Moorhead MN

2016

Thesis Committee Chair: Dr. Roger Rochat

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Global Health

2023

Abstract

Exploring the Determinants of Internet Searches for Abortion-Related Information in the Midwest United States: A Google Trends Analysis

By

Hans-Kristof Nelson

Background:
Google Trends (GT) analysis first became available in 2004 after petitioned improve collection of this data as well as to release this data to the public. GT analysis is used in various public health research applications including in areas of reproductive health access and justice.

Methods:
Interest_by_city is a variable that was drawn from aggregated Google Trends data using the package 'gtrendsR'. Data is returned as geolocated cities with a 'hits' value indicated that cities Google searching interest value (hits) relative to other cities in their state. We used this data as a measure of interest and examined this data against county level data from where the 'hit_city' exists. Examples of these county data are teen birth rate, metropolitan or non-metropolitan class, and percent of total poverty of each county. Simple t-tests, correlation tests, linear regression models and chi-squared tests are all methods of statistical analysis utilized. Table and spatial joins in ESRI ArcMap allowed for distance to nearest provider analysis.

Results:
For analysis the response variable (Y) was always 'hits' and the predictor variable (X) would change, there were significant as well as insignificant findings. Significant p-values are returned in the analysis of the county percent total poverty, and teen birth rate. Given the significant findings, there still exists a large limitation of the model. The project failed to control each hit city value by that cities population; giving a hit per capita variable is suggested to improve the balance of the analysis. Without controlling for population, a small, rural town measuring a 75 GT hits value is weighted the same as a large capital city measuring a 75 also. There may exist selection and sampling bias due to this fact.

Conclusions:
Even with the potential for selection bias, the model still was able to detect that of the hit cities that exist in counties with higher percent total poverty or teen birth rates, there could be increased Google search interest 'hits'. The findings highlight a few things; the abortion access landscape is changing week to week it seems, real time surveillance of Google traffic related to abortion could be more important now than ever. Not to assume much further than the 5 states in the Midwest but counties therein with initiatives combatting poverty or those with historically high teen birth rates should be evaluated for the equity of their health care access and financial support systems.

Exploring the Determinants of Internet Searches for Abortion-Related Information in the Midwest United States: A Google Trends Analysis

By

Hans-Kristof Nelson

-

Bachelor of Arts in Dietetics

Concordia College, Moorhead MN

2016

-

Thesis Committee Chair: Dr. Roger William Rochat

A thesis submitted to the faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Global Health

2023

Table of Contents

Exploring the Determinants of Internet Searches for Abortion-Related Information in the Midwest United States: A Google Trends Analysis

INTRODUCTION AND RATIONAL

Describe the public health problem – Additional to the very mobile state of abortion legality by state, a major public health problem we observe in the U.S is the lack of equal and reasonable access to reproductive health services.

Describe the specific problem being addressed – We seek to contribute to the knowledge pool on determinants of abortion care seeking via Google search, using Google trends data. We hope the efforts of this project can at least contribute to proof of concept that internet surveillance can be of particular use for those who care to increase abortion access in the upper mid-west.

The problem – There exists a need to understand the determinants and patterns surrounding abortion related search interest, as those factors contributing to abortion-related searches are not yet well known. We can use Google Trends search data as a proxy for estimating community interest in health-related topics, specifically reproductive health (Nuti et al., 2014, Seifter et al., 2010).

The background – The 1973 landmark Supreme court case Roe v. Wade established constitutional right to abortions for women. After that it became a battle ground of states implementing different restrictions like waiting periods, mandatory counseling, procedure and provider restrictions, and fetal heartbeat regulations. This type of legislation is often guised as health and safety regulations for the women and unborn child and varied greatly by state. Most recently in 2022, Roe v. Wade was overturned in the highly publicized Dobbs v. Jackson supreme court decision. This set-in motion a cascade of state legislation to change their abortion landscapes, many elected to restrict access and implement bans, other states elected to liberalize. Infodemiology is the study of the distribution and determinants of health-related information on the internet

Abortion care seeking surveillance using Google Trends data is a relatively new research approach gaining attention from reproductive justice researchers. Google Trends is a publicly available tool launched in 2006 that among other things, can facilitate investigation of the relative frequency of search queries related to specific search terms over specified time, and by geographic location. This feature makes it a useful tool for studying public interest on a range of topics, including those related to reproductive health. As this area of health legislation is changing rapidly in our country, to study Google Trends data may be more important than ever.

Its ramifications – Google Trends reproductive health surveillance and data collection will improve outreach efforts, with the assumption that those areas with high search interest represent persons seeking information and/or services. This research can visually display the overlapping layers of vulnerability, thus improving our knowledge about those populations and their communities, in addition to improving knowledge about Google Trends abortion related search patterns. We can answer some general questions like; who are these people that search for abortion information on Google? What situations of health or un-health do they find themselves in, or descriptive and geographic data from the communities where the searches originated.

Its theoretical and practical significance in the field of public health – If we would be able to predict where abortion services are needed based on where those who Google it the relative most are, we could deliver excellent care to highly specific places based on this surveillance. Additionally, if we also understand the socioeconomic and geographic factors associated with high internet interest in abortion related services there is an opportunity to anticipate need.

Problem Statement – Numerous populations in the upper mid-west U.S have little to no reasonable access to reproductive health services including abortions and face additional overlapping social, geographic, and economic barriers to accessing care. What should be – there should be geographically and financially accessible reproductive health services, including abortion services for all counties in the U.S.

Purpose Statement – Investigate the hypothesizes relationship between sampled Google Trends hit cities and socioeconomic status, metro and non-metro classification of county as well as distance to nearest provider.

The null hypothesis for my research would be that there is no association between Google Trends RSI hit value of a city and their counties metro and non-metro classification, teen birth rate, percent total poverty, (teen birth rate to percent total poverty ratio) or distance to nearest provider.

The alternative hypothesis is that there is an association between Google Trends RSI hit value of a city and their counties metro and non-metro classification, teen birth rate, percent total poverty, (teen birth rate to percent total poverty ratio) or distance to nearest provider.

Research Questions -

***My dependent variable (y) is Google RSI hit values for abortion related search terms. My independent variables (x) will be characteristics of the counties where these 'hit cities' exist, distance from those 'hit cities' to nearest providers. ***

1. What is the association between a city's Google Trends RSII value and their distance to nearest abortion provider?
2. What is the association between a city's Google Trends RSII value and their Census Bureau metro / non-metro classification?
3. What is the association between a city's Google Trends RSII value and their counties percent poverty for all ages?
4. What is the association between a city's Google Trends RSII value and their counties teen birth rate? Teen birth rate to percent poverty ratio?

Aim1: Geolocate cities (now referred to as 'hit cities') in Minnesota, Wisconsin, Iowa, North Dakota, and South Dakota returned with elevated Relative Search Interest Index (RSII) data from

Google Trends via package ‘gtrendsR’ for time of January 1, 2022, to December 31, 2022, to capture data surrounding the and including the Dobbs v. Jackson supreme court decision late June 2022.

Aim2: Geolocate abortion providers and map distances from each hit city to nearest provider. The provider list used was UpToDate as of the last day (Dec. 31) of 2021.

Aim3: Collect and assign values to all county layers for analysis. Values to be included for each county are, Census Bureau metro / non-metro classification, teen birth rate, percent poverty all ages.

Aim4: Conduct analysis and visual display of where these hit cities are and relation to the characteristics of their county, draw patterns for discussion and suggestions for improving health outcomes / health legislation.

Significance Statement -

The findings of this study are significant in shedding light on the factors that influence internet searches for abortion-related information in the Midwest region of the United States. Research done identifying key determinants contributes to our understanding of the social, political, and cultural factors that shape access to reproductive healthcare services and information regionally. The results of this study have important implications for public health professionals, policymakers, and advocates working to improve reproductive health outcomes and access to information in the Midwest and beyond.

Definition of Terms -

GT – Google Trends
RSI – Relative Search Interest Index – Referred to by Google simply as ‘hits’ or ‘hit value’ hyperlink to the FAQ from Google on their metric is here [https://support.google.com/trends/answer/4365533?hl=en]
hits – value associated with search interest 0-100 zero being lowest level of observed interest and one hundred being the highest level of search interest.
Hit City – any city/town with a non-0 RSI hit value returned from R package ‘gtrendsR’
RUCC – Rural urban continuum codes
hits - The project uses a synonym for ‘hits’, which is Relative Search Interest Index (RSI) value.

Thesis Title: Exploring the Determinants of Internet Searches for Abortion-Related Information in the Midwest United States: A Google Trends Analysis

REVIEW OF LITERATURE:

Introduction Paragraph -

Overview -

The literature review covers various stages of an iterative process of reviewing new literature as my project scope changed over time. An early yet formative piece of the literature review was reviewing national reproductive health organizations and advocates to help develop and frame the project idea. A primary source of information used is the Guttmacher institute and their dedicated work to reproductive health justice. Their investigative data collection returns quality reports to the public, a range from broad overviews of abortion trends and patterns to highly detailed reports regarding abortion in specific contexts. One example of these great resources is their latest worldwide report on the status of abortion from 2018 (Guttmacher Institute, 2018). The report specifically notes the disproportionate risk to marginalized populations that have limited access to abortion care. This is a finding that I am hoping to confirm and examine in greater detail in a specific area of the Midwest U.S. as there exist communities of poverty and limited access to health care services in general due to large distance and rurality.

Other institutions like the Institute for Womens Policy and Research, ANSIRH, and abortionfinder.org all contribute in large ways towards the common goal of making access to safe and legal abortions a fundamental human right. One way that these institutions push the needle on abortion access and justice is the visual display of abortion data and spatial analysis. Mapping abortion access and relevant state and national policies is not a new area of study. However, research institutions like the ones mentioned above are constantly working on creative ways to display different factors and layers of determinants of health related to abortion. Common displays of abortion access are calculated composite scores reflecting a state's reproductive health overall quality level of hostility (National Partnership for Women and Families, n.d). Most recently, a very strong display of abortion policies are time series maps of the changing abortion landscape in the U.S. prior to and directly following monumental abortion policy changes that have occurred (Abortion Finder, n.d., Center for Reproductive Rights, n.d.).

One semi-recent metric of trend analysis not originally created for reproductive health surveillance is the use of monitoring Google search queries. Google Trends was launched in 2004 and updated in 2006 with improved processing and data. Trends is a publicly available resource where users can track searches typed into Google search bars. Search data points such as terms used, geographic location, date and time are aggregated, synthesized and stored as a sample by Google for later use. This is a useful tool for researchers for many reasons. Understanding the trends and patterns of specific search terms over time can help with forecasting the need, interest, or satisfaction of a certain topic. Some of the earliest uses of the application were for economic forecasting. Another creative use of the application was to understand geographic racist sentiments following the election of former President Barack Obama (Stephens-Davidowitz, S, 2012). One final example of the application of the Google Trends analysis was from the same writer from the New York Times who investigated the actual number of homosexual American males. In conjunction with other data the same researcher pulled Google search queries to develop an estimate on the number of American males that identified as gay. His search of Google Trends data found many potential search queries to help build his estimate. One example that the researcher found was that most common search query starting with “is my husband” was, “is my husband gay” (Stephens-Davidowitz, S, 2013).

Google Trends (GT) data is not stand-alone, and the best analysis is conducted when search query data is combined with other validated data. We know that abortion and reproductive health Google Trends data exist, and considering this limitation of independence we will think critically and continue to review literature with the intention of directing the project so that other validated data may be used in conjunction with this GT data.

Justification -

Abortion has and continues to be a common procedure in the US with over 850,000 abortions provided in clinical settings in 2017 (Guttmacher Institute, 2018), yet many people fight every day so the people around them have less access to that care (Kodjak, A, 2022).

Research exists that investigates the act of crossing state lines for safe and legal abortions, even more recently it has become possible to have medical abortion delivered by mail for self-use out of the clinic (Maddow-Zimet I and Kost K, 2022). In 2020 a study was published that quantified and discussed the demand for self-managed abortion in the U.S. The study found significantly higher demand for self-managed abortion services via telemedicine in states with more restrictive abortion policies (Aiken et al.,). This motivates the notion of surveying the internet to determine need, demand and other factors limiting abortion access.

Maternal Mortality rate is also a large motivator to investigate reproductive health and abortion access in the U.S especially in any area that actively restricts access to any type of health care for women, including reproductive health. Recently the U.S. recorded some rates of maternal mortality that are alarmingly bad to the tune of 32.9 deaths per 100,000 women when compared to other high income, developed countries with excellent health care like Spain, Japan, and Australia that have between 2 and 4 maternal deaths per 100,000 women (Hoyert, D.L. 2021).

As researchers and reproductive health allies and advocates we seek to untangle and address the dynamic root causes of these discrepancies. Research of underlying factors of abortion restrictive policies is needed, research on the populations disproportionately affected by such policies and barriers is also needed to form a response to the cascade of consequences these restrictive policies carry.

Body -

Much of the literature on the analysis of Google Trends data shows wide ranges of applications for this type of analysis. Broadly speaking researchers suggest high utility of data for predicting current trends in economic and social behaviors. The authors argue that Google Trends can provide real-time insights into consumer behavior, preferences, and sentiment, and demonstrate this with several examples, including predicting the volume of automobile sales and housing starts. The limitations noted by the authors are those issues of data quality, representativeness, and the inability to capture certain types of behaviors, one example given was perceived impulse buying (Choi, H., 2012). Ultimately the authors conferred with what numerous other resources conclude on, that being that the use of Google Trends data is best when analyzed in conjugation with other data. This resource is important in highlighting the potential use of Google Trends data analysis for behaviors, perceptions and outcomes associated with abortion, poverty, or unintended pregnancies just to give a few examples.

A systematic review of Google Trends data published in 2014 outlines the areas of public health research that could be monitored by Google Trends data analysis. "The Use of Google Trends in Health Care Research: A Systematic Review" by Nuti et al. (2014) provides a comprehensive review of the use of Google Trends in public health research. The authors conducted a systematic review of published articles that utilized Google Trends data in health care research and found that Google Trends has been used to track disease outbreaks, investigate health-related search behavior, and forecast health care demand. The authors also discuss the advantages and limitations of using Google Trends data, including issues related to data quality, representativeness, and privacy. They conclude that Google Trends can be a valuable tool for health care researchers, but it should be used in conjunction with other data sources and methods to ensure the validity and reliability of the findings (Nuti, S.V., 2014)

Tracking Disease Outbreaks

Literature reviewed also details the use of Google Trends data analysis in public health and outbreak response. "Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks" by Carneiro and Mylonakis (2009) discusses the potential of using Google Trends to monitor disease outbreaks in real-time. The authors used Google Flu Trends and compared them to CDC flu surveillance systems data and the Google Trends data detected the influenza outbreak 7-10 days prior. The authors argue that Google Trends data can provide early warning signals for disease outbreaks by tracking patterns of health-related searches and their changes in volume. They provide several examples of how Google Trends has been successfully used to monitor disease outbreaks, including influenza, dengue fever, and Chikungunya virus. The authors note that GT data analysis is probably best suited for developed countries where a large part of the population is using web search engines. Additionally, the authors note that GT data analysis results are difficult to interpret without other contextual information and data. This step can help to relate the findings to the public health issue of the specific region and population existing there. (Carneiro, H., 2009).

Epidemiological data of vector borne disease outbreaks have also been explored as a viable use of Google Trends data. "The Utility of 'Google Trends' for Epidemiological Research: Lyme Disease as an Example" by Seifter et al. (2010) discusses the potential of using Google Trends to monitor and track the incidence of Lyme disease. The authors argue that Google Trends can provide valuable information on the public's interest in Lyme disease and related topics, as well as track seasonal and geographic trends in the disease. The authors also concluded here that the use of Google Trends Data would be best analyzed in conjunction with other available data (Seifter, A., 2010).

Investigate abortion-related search behavior and motivators

One study looked to understand what people are looking for when they googled "self-abortion". Using a Google AdWords campaign, the researchers could send a survey to google searchers who entered 1 of 26 terms possibly indicating self-abortion seeking behavior, this survey investigated the reason for the initial search. The researchers sent out the survey 210,000 times and collected 1235 eligible completed surveys. 96% of respondents were women and 46% were minors. 73% of the respondents reported that they reverted to the internet because they were currently pregnant or didn't want to be pregnant. One third of the respondents reported being unfamiliar with their state laws regarding abortion. The authors drew the conclusion that much of the internet traffic for self-medicated abortion was driven by adolescents and young adults facing restrictive abortion policies and or an unintended pregnancy. (Jerman, J., 2018)

Another study published in 2011 produced survey results comparing patient and provider perceived safety, effectiveness, and acceptability. They found that the telemedicine option was highly acceptable to patients. The study found that providers saw telemedicine as safe and effective. Overall, the authors conclude that telemedicine can significantly increase access to abortion in areas with very limited access or where abortion is highly stigmatized. The results of this study are important to understanding some behaviors of those seeking abortion in these areas of limited resources, stigmatization, and limited reproductive health care or general health care (Grossman, D., 2011).

Maybe the most important resources that this literature review turned up were the works of the lead researcher Guendelman. The two studies use an infodemiologic approach to investigating Google searches related to abortion. The first investigated factors that motivate abortion-related searches and was published in 2018. The study used a Google search API to draw Google trends data related to their abortion related key terms. The study then used similar methodologies to tother references resources by mapping and comparing where these Google searches were relatively the highest and began to run analysis against other factors and determinants of health. The authors found results consistent with other research in that states with more restrictive abortion policies had a significantly higher relative Google search interest index. Interestingly, this study found that the abortion search volumes were positively associated with poor health outcomes, poor access to abortion facilities and non-rurality. These findings suggest that the highest search volumes generally came from non-rural areas, areas with limited access to abortion facilities. This leads one to infer that an individual in a rural zone, that has less access to abortion providers might be more interested in getting an abortion, or at least reverting to Google for information. According to this study, that might not be the case, and it makes some sense because in metropolitan areas we would assume to have much more people Googling at much higher volume. Overall, however, the authors conclude that those residing in the states with the most restrictive abortion policies are more likely than those in less restrictive states to Google abortion (Guendelman et al., 2020).

The second study (Guendelman et al., 2022) investigated abortion-related Google searches specifically for out of clinic medication abortion, referring to people searching for online pharmacies to access medical abortions but also abortion providers operating online and mailing medical abortions to states that restrict such procedures. This study used almost identical methodologies as the author's previous work. The results were conclusive in that states with more restrictions have higher search volume for abortion related queries. Another interesting finding was that searches for “home abortion” peaked in November 2020 of their research time frame. This coincided with the national allowance of telemedicine abortion provision following a federal district court overturning a previous FDA mandate that required providers have special license to dispense mifepristone. So, the researchers concluded that once it was legal everywhere to receive telemedicine abortions, this drove google search volume up, and especially in states where telemedicine abortion would be the only truly accessible option for abortions. A few strengths of these studies are the sampling methods used. In both studies, researchers drew samples of RSI 30 separate times and averaged them all together. This is something that my project will not be able to accomplish due to my recent exposure to this research method and newness to the statistical procedures and software.

In conclusion of the subpoint, we know that areas of limited access to abortion providers have a higher average RSI value of abortion related search terms

Forecasting abortion-related Healthcare demand

Other studies report that people might revert to searching the internet more frequently in states that have more restrictive abortion policies in place or incoming. These resources highlight incoming demand for abortion services

One primary example is an observational study of the relationship between abortion related search volume, local abortion rates and the local abortion policies. This study runs comparisons globally between countries in addition to comparison between all 50 states. The scale of the project and the data that was collected shows how different these variables are internationally when comparing restrictive countries to liberal countries when it comes to reproductive health. The results of the analysis show that both nationally and internationally, the areas with the most restrictive abortion policies had significantly higher abortion-related searches compared to areas with liberal legislation. Additionally, the authors found that in states with less than 10% of counties having an abortion provider, those states had significantly higher abortion related search volume than the rest. Finally, and maybe the most interesting policy and internet search relationship was that in states that had abortion restrictions ready for if Roe v. Wade were ever repealed had significantly higher abortion related search volumes than those states who made promises to protect abortion rights if Roe was ever repealed (Reis & Brownstein., 2010). Now having lived through the year 2022 and the Dobbs v. Jackson case, we could guess with some validity at which states would have been Googling for abortions after the Dobbs decision. Additional to new and impending state policies and restrictions, this study motivates more investigation into the factors surrounding abortion related web searches in local contexts where there may be highly specific context clues and overlapping determinants that affect searches.

One difference I believe that we will do that is different from the Reis & Brownstein., 2010 project is that we won't count the number of counties that have a provider because there are so few in our area of research, we will likely just use distance to the nearest provider.

Additionally, what we will do differently is have relative search interest index read back to us as interest by city, an option available in the package gtrendsR.

Most recently Poliak et al. 2022 completed research that investigated Google searches for medical abortions only hours after the SCOTUS Dobbs decision was leaked. Results of this investigation can provide validity to my research goals because Minnesota recorded the lowest relative search volume of itself and its 4 neighbors; Iowa, Wisconsin, North Dakota, and South Dakota (Poliak, S., 2022). This result is consistent with what has been discussed in this literature review that the states with the most restrictive abortion policies or have impending restrictions are those states that end up with the most relative search volume for abortion related searches. Limitations are known to exist as to the real reason for the search and who is making the search query. The study is also limited in the fact that it cannot estimate if any of these searches resulted in an abortion at all. We do have evidence to suggests that Google searches primarily come from adolescent or young adult females who are pregnant or do not wish to be pregnant. We deduce that most searches for self-abortion are in fact seeking abortive care for themselves or someone close to them (Jerman, J., 2018).

Conclusion -

In conclusion this literature review has identified the use of publicly available Google trends data as a promising tool for understanding the public engagement with reproductive health issues. By mapping the search areas based on various factors such as county designation, teenage birth rate, poverty rate, and proximity to abortion providers, this study aims to draw associations and relationships among all variables to gain insights into elevated abortion-related Google search interest.

However, the use of Google Trends data for abortion surveillance also presents limitations and potential biases, which need to be taken into account. These include the inability to differentiate between searches by individuals seeking abortion services and those seeking information for other reasons. To fully understand the potential and limitations of using Google Trends data for abortion surveillance, further research is necessary.

Overall, the findings of this study can provide valuable information for public health policymakers and practitioners to improve access to reproductive healthcare services and address the underlying factors that contribute to elevated abortion-related search interest.

-END LIT REVIEW-

METHODS:

IRB approval was not needed for this project as it was not considered human subject research and only uses aggregated Google Trends data. No addresses or names of abortion providers are displayed, and cities are not named on maps to protect those locations. A confidentiality agreement between the primary researcher and ANSIRH representative exists on file.

Returning Hit Cities in R

The two primary software used on the project are ESRI product ArcMap 10.8.2 (ESRI, 2020) and RStudio 1.5.679 (RStudio Team 2023). While throughout the project it was discovered that most if not all the steps could have been done in R, I was able to practice two skillsets. The first steps in the project were in R, using the package called 'gtrendsR' (Rouder, 2021). What this package allows users to do is search global Google search queries filtered for geography, time as far back as 2004, and up to 5 key words. The key terms selected for the project are "Abortion", Abortion near me", "Abortion cost", "Planned Parenthood", "Abortion pill". These key words were generated by researchers and associates as some of the more general terms that would capture what we thought would be the most amount of data. Other studies used similar terms like "planned parenthood" (Guendelman et al.,2020, Guendelman et al., 2022). The geography we chose to investigate is the region of the northern mid-western states of Minnesota and its neighbors Wisconsin (WI), Iowa (IA), South Dakota (SD), and North Dakota (ND). The study population will be every city recognized by google that is able to register Google Trends data. Data points and metrics are aggregated and returned in a variety of forms. Subsequently, we had 'gtrendsR' return the data to us sorted by city in the 5-state area that registered an elevated value, and that 'hits' value. There were 164 cities total returned from the completed code run on the draw date of February 1, 2023. The project uses a synonym for 'hits', which is Relative Search Interest Index (RSI) value. 'Hits' is the alias for the RSI value and is represented as 0-100 scale, and it represents Googles interpreted level of interest in a particular search query (Google, n.d). For example, city X that registers a 0 for the search term 'Abortion' has no observed change in googled interest over the time-period normalized by all google searches of the same time and geography. City Y that registers an RSI value of 100, is very much 'interested' in the search term 'Abortion' still being normalized for total Google searches over time and geography. The program returned one hundred and sixty-five (165) cities in all the 5-state area that registered an RSI value for the 5 search terms for the year of 2021. For additional aggregating and ease of use hit volume was categorized into low, medium and high categories and that new variable named 'hits_cat'.

State and County Shapefiles

All the subsequent steps happened in ArcMap. We imported the state and county shapefiles from the publicly available database on the census bureau's website.

State and County Shapefiles and County Rural Urban Continuum Codes (RUCC)

The map data set was created in several steps. Initially, the entire country was represented on the map at a high level of zoom. The next step was to zoom in to the state and county level. Drawing the boundaries of each state and county was the first task in the GIS portion of the project. The second task involved mapping each county to its respective Rural Urban Continuum Code (RUCC) classification using data from the most recent census in 2013.

To achieve this, the project team obtained public records from the United States Department of Agriculture (USDA) website. These records were available as an Excel spreadsheet and contained the RUCC codes for all US counties. The spreadsheet was added to the map layout, and a join procedure was performed on each county's FIPS code to link the data from the table to each county's polygon on the map. The result was a layer of county polygons, each of which was paired with its respective RUCC value, indicating whether it was classified as metro or non-metro.

The RUCC codes divide counties into nine categories, with codes 1-3 assigned to metropolitan areas and codes 4-9 assigned to non-metropolitan areas. Based on this classification, a new variable called "metro" was created. The "metro" variable categorized counties into two groups, with a value of 1 assigned to metro counties and a value of 0 assigned to non-metro counties. This allowed the map to display a clear distinction between urban and rural areas, providing valuable information for various applications, including economic research, urban planning, and public policy.

Mapping Hit Cities

This element of the project required geo-locating all the 165 hit cities returned from gtrendsR. This process is important to accomplish because the exact geographic coordinate we will use for analysis is a reference point to nearest a nearest provider, to demographics of its corresponding county, as well as the RSI hit value representing relative interest in abortions. We exported the list of hit cities to Microsoft Excell. At this point the data was a name of a city, its state and the RSI hits value. In combination with the search function on Google Maps, we documented each individual city/town centroid's latitude and longitude to that sheet. The 165 cities all came from either Minnesota (MN), North Dakota (ND), South Dakota (SD), Iowa (IA), or Wisconsin (WI). Then, with ArcMap we imported the dataset and imported the XY data as point data on the default trans Mercader projection.

Mapping 2021 Provider Location

The process of creating a map dataset of abortion providers started by obtaining the geographic data of the providers. To achieve this, the project requested access to a database created by a program at the University of California San Francisco called Advancing New Standards in Reproductive Health (ANSIRH). ANSIRH is a highly reputable and accurate database that is updated annually and has been utilized in various research projects over the years. Information on ANSIRH's policies and procedures can be found on their website. The database provided each provider's geographic location in terms of longitude and latitude, along with many other variables. No identifiable data are displayed in any maps or tables.

The project only displayed providers from this national data that existed, were open, reported providing at least one abortion in the year 2021, and were located within the 11-state area. This 11-state area includes Minnesota (MN), North Dakota (ND), South Dakota (SD), Iowa (IA), and Wisconsin (WI), along with six other states. Once the necessary data was filtered, it was then uploaded to the mapping software for further processing.

Percent Total Poverty

The percentage of the total population living in poverty was assigned to each county by a table join in ArcMap by county FIPS code. This data was also retrieved from the (United States Census Bureau, n.d.)

Teen Birth Rate

The teen birth rate is also joined to each county polygon by table join in ArcMap by county FIPS code. This data was also retrieved from the National Center for Health Statistics National Vital Statistics System (NCHS, n.d.).

Teen Birth Rate to Percent Total Poverty Ratio

By joining the two data layers together, percent total poverty and teen birth rate, we can observe those counties with a higher ratio have a stronger association between the teen birth rates and the percent of total poverty in that county.

Joining them all together

The following steps are some of the final ones before the data set is ready for analysis. Each city carries an RSI hits value and through spatial and table joins of the feature layers described here above we can determine if each city is in a metro or non-metro county and their distance in meters to their nearest provider that did provide an abortion in 2021. The following steps are then to categorize the layers differently. The hit cities values can be quantified into low, med-low, medium, med-high and high for deeper analysis to know if there is an association between the RSII scores and distance to nearest provider. Additionally, if there is an association between RSI scores and if the city is in a county that is metro or non-metro.

Continued Data Creation

The next step to take the map to another level is to create a map layout with an updated abortion provider landscape, as of the final days of 2022. This is possible through diligent work done by all ANSIRH, the Guttmacher Institute and abortionfinder.org. With their public and highly reliable data it is possible to re-run the distance analysis. This is an important step because as of the end of 2022 three states in the original 5-state area no longer had any abortion providers due to changing and impending state laws that would make abortions illegal. This map layout would illustrate the people who were furthest away from a provider and how their options may become further away, or populations that previously had good access but now are forced to travel much longer distances and even cross state lines in search of an abortion provider. One example includes Wisconsin and the large number of clinics that have closed leaving large population centers like Milwaukee to cross state lines into Illinois or Minnesota seeking abortion care.

Limitations of the methodology

There are several factors beyond the control of a researcher that follows this or a similar methodology. First is that there is little control surrounding the variety of socio-political events that happen throughout the study time and may influence the results in certain areas more than in others. Other limitations that exist with the methodology detailed here is the single sample draw of 'hit_cities'. The primary references for this project Guendelman et al., 2023 & 2021 draw the same type of data through a Google API method that allows them to draw larger number of samples before running analysis on all of those. I did not have the ability to run the number of samples and analyze them all with the knowledge and human hours I had available as a student. One convenience sample decreases the generalizability of the results.

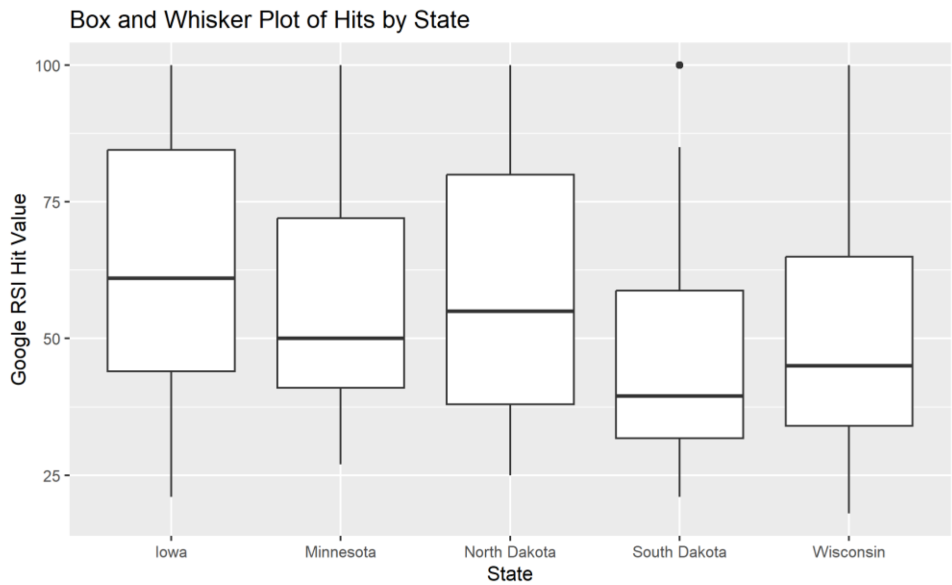
Delimitations of the methodology:
There are a number of factors that the research team controlled in order to narrow the scope of the project. Examples of such delimitations are focusing the area of research on 5 states in a specific region and examining the counties determinants to abortion access. The number of search terms that could be used by gtrendR was limited to only 5 unique terms, so the research team chose what we thought to be the most common terms used by someone seeking information about getting an abortion.

ANALYSIS & RESULTS:

I, as the primary researcher, had a hunch that given the change in the abortion landscape with the Dobbs v. Jackson decision (date), I would be able to capture useful information about abortion related information seeking via Google Trends data. Additional to this hunch, I thought that there would be significant results from my analysis indicating that rural counties would have cities with higher Google RSI hit values. I also thought that there would exist a relationship between hit values and a poverty metric as well as a teen pregnancy metric.

First, we will visualize the layout of ‘hits’ by state

Graphic 1




```
ggplot(data, aes(x = State1, y = hits)) +
geom_boxplot() +
labs(title = "Box and Whisker Plot of Hits by State",
x = "State", y = "Google RSI Hit Value")
```

To answer research question 1 ## What is the association between the city's Google Trends RSI 'hits' and the metro / non-metro classification of the county where that city exists? Tests that were conducted are bolded below.

t.test(data$hits ~ data$metro)

The Welch Two Sample t-test is a statistical test that compares the means of two independent groups. The null hypothesis of the test is that there is no difference in means between the two groups, while the alternative hypothesis is that there is a difference in means between the two groups.

In this case, the test was conducted on the variable "hits" (represents the Google Trends data Relative Search Interest Index, RSI) for two groups based on the variable "metro". The "metro" variable is a binary variable indicating whether the data point comes from a metropolitan area or not.

The output shows that the t-value of the test is 2.4188, with a degree of freedom of 80.226. The p-value of the test is 0.01784, which indicates that there is evidence to reject the null hypothesis at the 0.05 significance level. This might indicate that there is more than a random chance that the result of the sample contains a significant difference of mean hit value between metro and non-metro counties

Additionally, the 95% confidence interval for the difference in means between the two groups is (1.765649, 18.154048). This means that we are 95% confident that the true difference in means between the two groups lies within this interval.

Finally, the sample mean for the group in the metropolitan area is 58.14167 and the sample mean for the group not in the metropolitan area is 48.18182. This means that, on average, the group in the metropolitan area has a higher "hits" score than the group not in the metropolitan area.

cor.test(data$hits, data$RUCC_2013)

A correlation test is run to determine the relationship between cities in their hits category and each individual rural urban code. The output indicates that there is a significant negative correlation between hits and RUCC_2013 variables ($r = -0.182$, $p = 0.019$). This means that there is a weak negative relationship between the two variables - as one variable increases, the other tends to decrease. The correlation coefficient of -0.182 suggests that the relationship is not very strong. Based on the results of the correlation test between the hits variable (which represents search interest in abortion-related topics) and the RUCC_2013 variable (which represents the rural-urban classification of the city), there is a significant weak negative correlation between the two variables.

chisq.test(table(data$metro, data$hits_cat))

The 0-1 binary variable ‘metro’ was then used in analysis with a new variable called ‘hits_cat’. ‘hits_cat’ was created more or less arbitrarily starting from value 10 that being below the lowest recorded value in the data set and increasing by 3 increments of 30 for low, medium and high breaks at <40, <70 and <100.

The output shows that the chi-squared value is 3.4839 and the degrees of freedom are 2. The p-value of the test is 0.1752, which means that there is not enough evidence to reject the null hypothesis of the test.

Therefore, we can conclude that there is no significant association between the "metro" and "hits_cat" variables at the 0.05 significance level. This means that the proportion of data points in each category of "hits_cat" is not significantly different between the "metro" and "non-metro" groups.

This result is not significant but also worth reporting for future research because due to the amalgamation of variables into categories, the analysis loses granularity which may explain results showing no association. A limitation of this project would also be the small sample size and with more hit cities there might emerge more significant results.

model.1 <- lm(hits ~ metro+state, data = data)

summary(model.1)

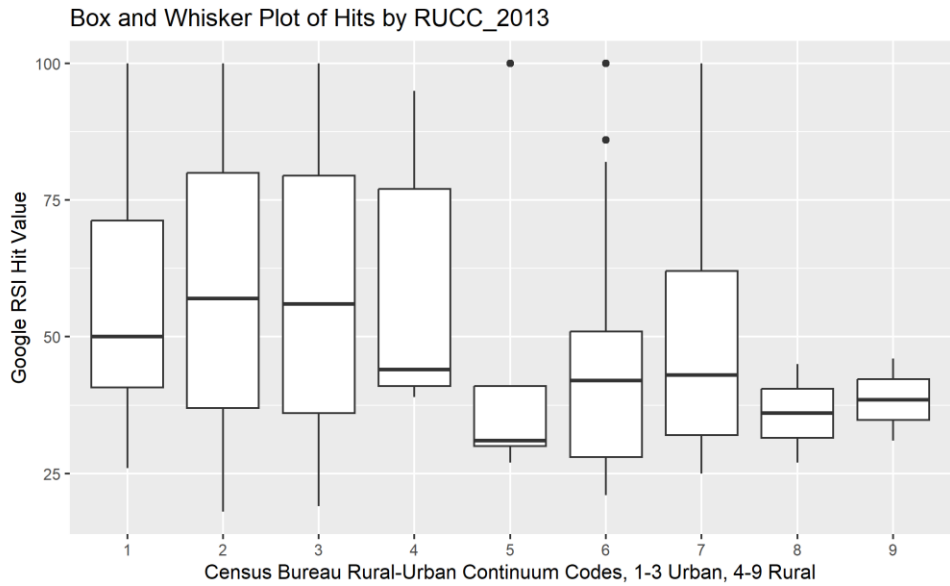
The "metro non-metro" coefficient is -10.802, with a significant p-value of 0.0127, which indicates that there is a significant difference in "hits" between the "metro" and "non-metro" groups. Specifically, on average, the "non-metro" group has 10.802 fewer hits than the "metro" group after controlling for the effect of the "state" variable.

The coefficients for the "state" variables (MN, ND, SD, and WI) show the expected change in "hits" associated with each state compared to the baseline state, IA. However, none of these coefficients have a statistically significant p-value, indicating that there is no significant difference in "hits" between the states after controlling for the effect of the "metro" variable.

The adjusted R-squared value of 0.04246 indicates that only 4.2% of the variation in "hits" can be explained by the "metro" and "state" variables in the model. The F-statistic of 2.446 on 5 and 158 degrees of freedom, with a p-value of 0.03634, indicates that the overall model is statistically significant at the 0.05 significance level, meaning that at least one of the independent variables is associated with the dependent variable.

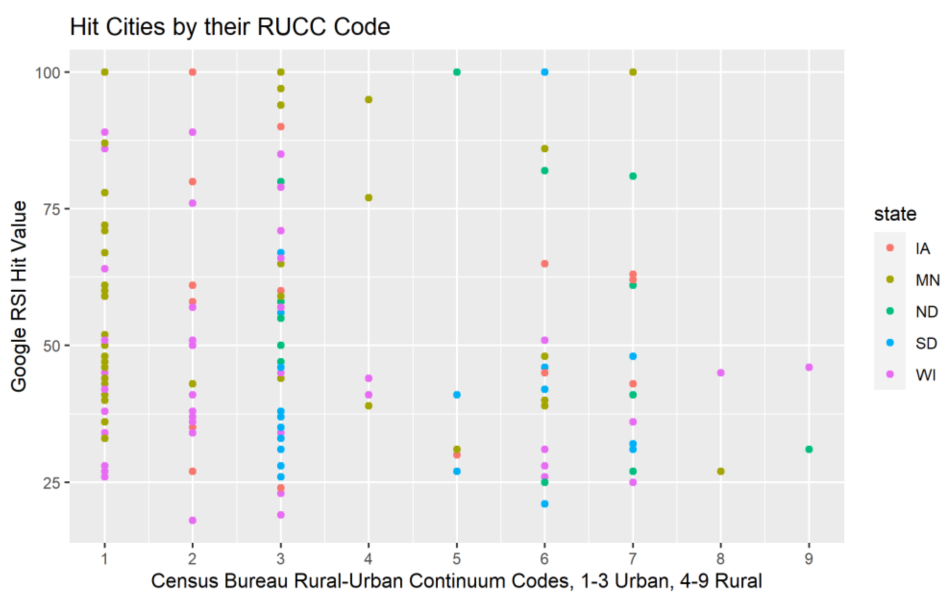
This data is visualized in figure X and code displayed below each figure:
Reminder that each point is a ‘hit city’ from the data set. Hits represent each city's Google RSI value.

Graphic 2



ggplot(data, aes(x = factor(RUCC_2013), y = hits)) +
geom_boxplot() +
labs(title = "Box and Whisker Plot of Hits by RUCC_2013",
x = "Census Bureau Rural-Urban Continuum Codes, 1-3 Urban, 4-9 Rural",
y = "Google RSI Hit Value")

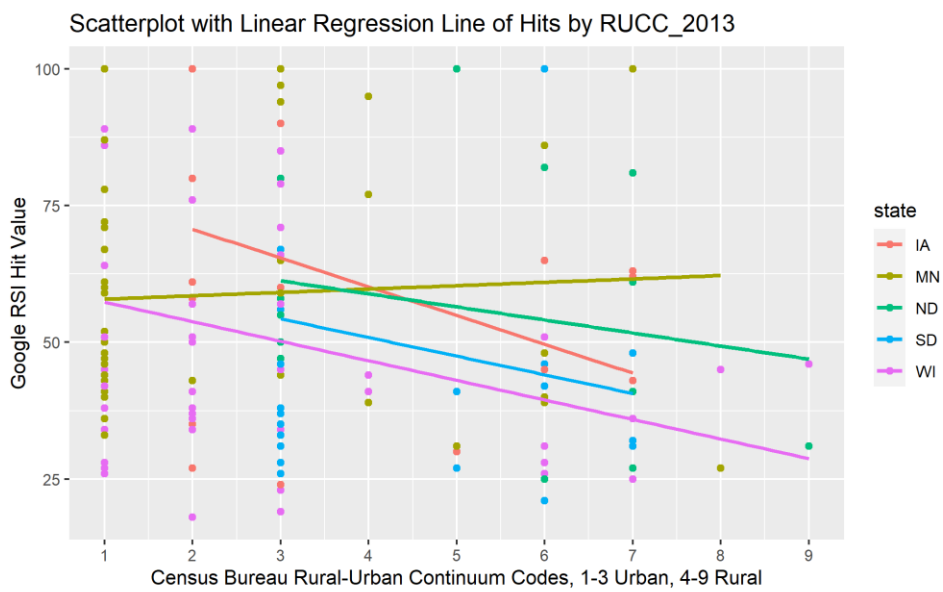
Graphic 3



ggplot(data, aes(x = RUCC_2013, y = hits)) +
geom_point(aes(color = state)) +

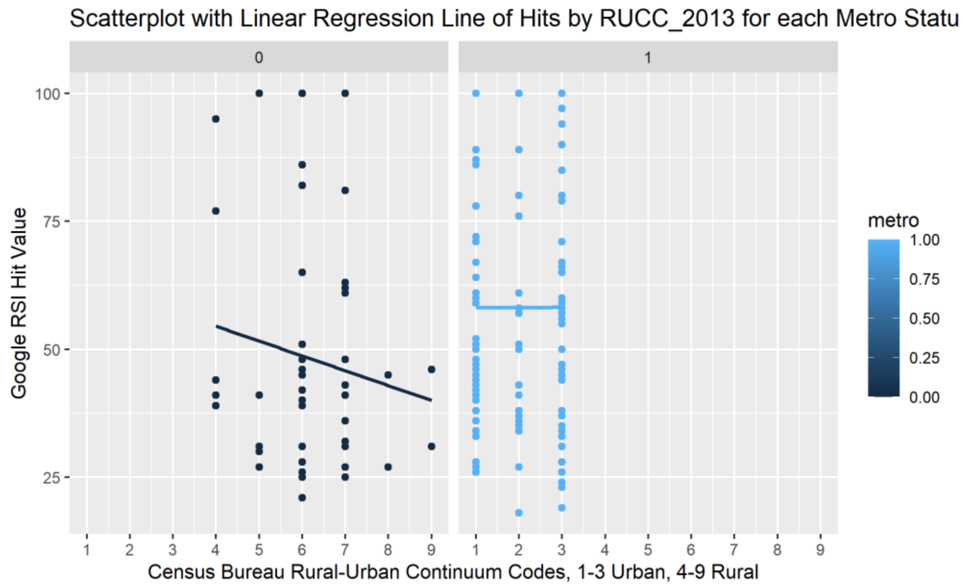
scale_x_continuous(breaks = seq(1,9,1), name = "Census Bureau Rural-Urban Continuum Codes, 1-3 Urban, 4-9 Rural") +
scale_y_continuous(name = "Google RSI Hit Value") +
labs(title = "Hit Cities by their RUCC Code")

Graphic 4



ggplot(data, aes(x = RUCC_2013, y = hits, color = state)) +
geom_point() +
geom_smooth(method = "lm", se = FALSE) +
scale_x_continuous(breaks = seq(1,9,1), name = "Census Bureau Rural-Urban Continuum Codes, 1-3 Urban, 4-9 Rural") +
scale_y_continuous(name = "Google RSI Hit Value") +
labs(title = "Scatterplot with Linear Regression Line of Hits by RUCC_2013")

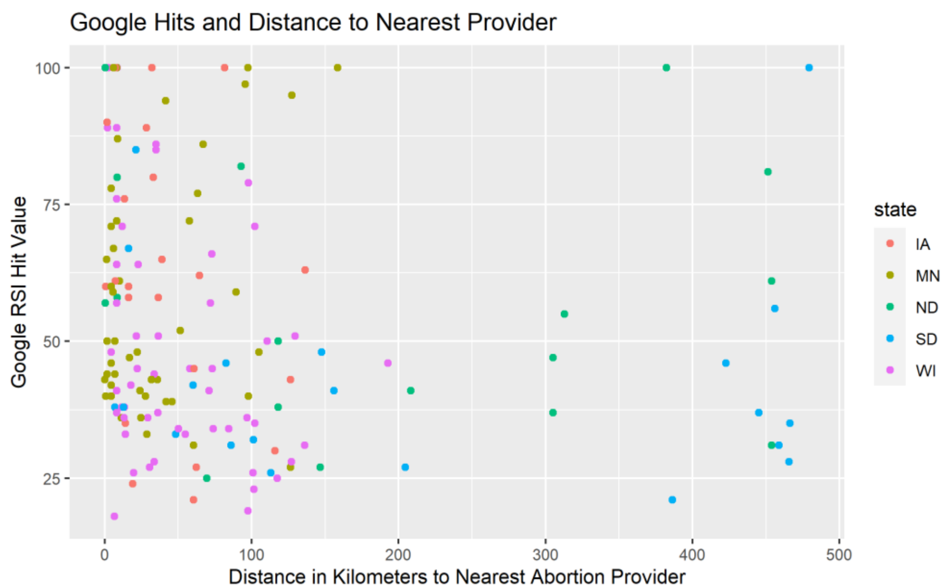
Graphic 5



ggplot(data, aes(x = RUCC_2013, y = hits, color = metro, group = metro)) +
geom_point() +
geom_smooth(method = "lm", se = FALSE) +
scale_x_continuous(breaks = seq(1,9,1), name = "Census Bureau Rural-Urban Continuum Codes, 1-3 Urban, 4-9 Rural") +
scale_y_continuous(name = "Google RSI Hit Value") +
facet_wrap(~metro) +
labs(title = "Scatterplot with Linear Regression Line of Hits by RUCC_2013 for each Metro Status")

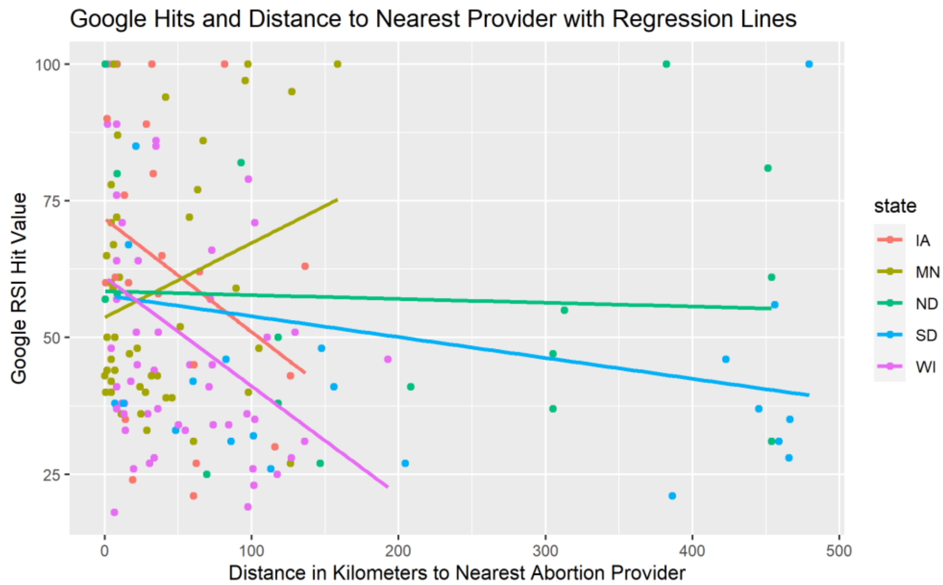
To answer research question 2: What is the association between a city's Google Trends RSI value and their distance to nearest abortion provider?

Graphic 6



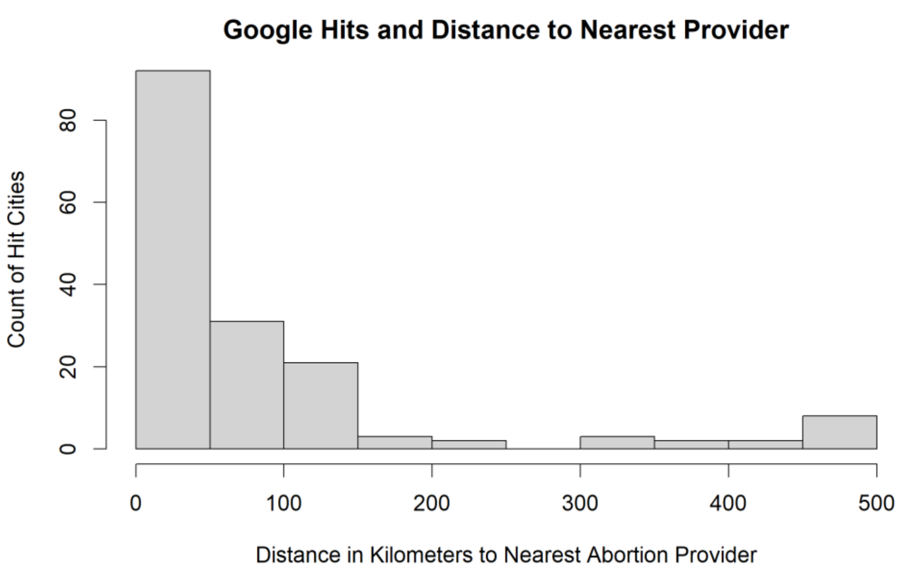

```
ggplot(data, aes(x = distance_km.1, y = hits, color = state)) +
geom_point() +
xlab("Distance in Kilometers to Nearest Abortion Provider") +
ylab("Google RSI Hit Value") +
ggtitle("Google Hits and Distance to Nearest Provider")
```

Graphic 7




```
ggplot(data, aes(x = distance_km.1, y = hits, color = state)) +
geom_point() +
geom_smooth(method = "lm", se = FALSE)+
xlab("Distance in Kilometers to Nearest Abortion Provider") +
ylab("Google RSI Hit Value") +
ggtitle("Google Hits and Distance to Nearest Provider with Regression Lines")
```

Graphic 8



hist(data$distance_km.1, ylab = "Count of Hit Cities", xlab = "Distance in Kilometers to Nearest Abortion Provider", main = "Google Hits and Distance to Nearest Provider")

Reserarch question 3: What is the association between a city's Google Trends RSI value and their counties percent poverty for all ages?

```
Call:
lm(formula = hits ~ PercentagePopPovertyAllAges, data = data)
Residuals:
    Min      1Q  Median      3Q     Max
-40.935 -17.569  -5.461  17.037  51.358

Coefficients:
                            Estimate Std. Error t value       Pr(>|t|)
(Intercept)                  41.7888     6.4140   6.515 0.000000000874 ***
PercentagePopPovertyAllAges   1.2931     0.5799   2.230         0.0271 *
---
Residual standard error: 23.97 on 162 degrees of freedom
Multiple R-squared:  0.02978,	Adjusted R-squared:  0.02379
F-statistic: 4.973 on 1 and 162 DF,  p-value: 0.02712
```

The coefficient estimate for the predictor variable "PercentagePopPovertyAllAges" is 1.2931, which means that for every one unit increase in the percentage of the population living in poverty, there is an expected increase of 1.2931 hits.

The p-value associated with the coefficient estimate for the predictor variable (percentage total poverty) is 0.0271, which is less than the commonly used alpha level of 0.05, indicating that the coefficient is statistically significant at the 0.05 level.

The R-squared value of 0.02978 indicates that approximately 2.978% of the variance in hits value can be explained by the percentage persons living in poverty in that county. The adjusted R-squared value of 0.02379 indicates that the model with the percentage persons living in poverty as a predictor variable is not a very good fit to the data.

The F-statistic value of 4.973 on 1 and 162 degrees of freedom corresponds to a p-value of 0.02712, indicating that the model as a whole is statistically significant at the 0.05 level.

Research Question 4: What is the association between a city's Google Trends RSII value and their counties teen birth rate? Teen birth rate to percent poverty ratio?

```
#Fit a linear regression model
model <- lm(hits ~ Teen_Birth_Rate, data = data)

#Print the model summary
summary(model)
```

The results:

```
Call:

lm(formula = hits ~ Teen_Birth_Rate, data = data)

Residuals:

   Min      1Q  Median      3Q     Max

-40.188 -18.553  -6.263  17.715  47.782

Coefficients:

                Estimate Std. Error t value            Pr(>|t|)

(Intercept)      49.2616     3.6142   13.63 <0.0000000000000002 ***

Teen_Birth_Rate   0.5097     0.2536    2.01              0.0461 *

---

Residual standard error: 24.04 on 162 degrees of freedom

Multiple R-squared:  0.02433,	Adjusted R-squared:  0.01831

F-statistic:  4.04 on 1 and 162 DF, p-value: 0.0461
```

The coefficient of 0.5097 for Teen_Birth_Rate suggests that there is a positive relationship between the Teen_Birth_Rate and hits. Specifically, for every one-unit increase in Teen_Birth_Rate, there is an expected increase of 0.5097 in hits, all other factors held constant.

The p-value of 0.0461 for the coefficient of Teen_Birth_Rate indicates that this relationship is statistically significant at a significance level of 0.05. However, the low R-squared value indicates that there may be other factors that also influence the level of interest in abortion-related searches that are not captured in the model.

In context of the data, this is important for the team because the hit cities are still not controlled for population and there were significant findings, this in meaningful contributions for the research team. The positive relationship between the Teen_Birth_Rate and hits suggests that there may be a link between the two variables. Specifically, areas with higher teen birth rates may have a greater interest in abortion-related searches on Google.

DISSCUSSION AND MAPS OF DETERMINANTS OF ANALYSIS:

Question #1 What is the association between a city's Google Trends Relative Search Interest 'hits' (RSI) value and their metro / non-metro classification?

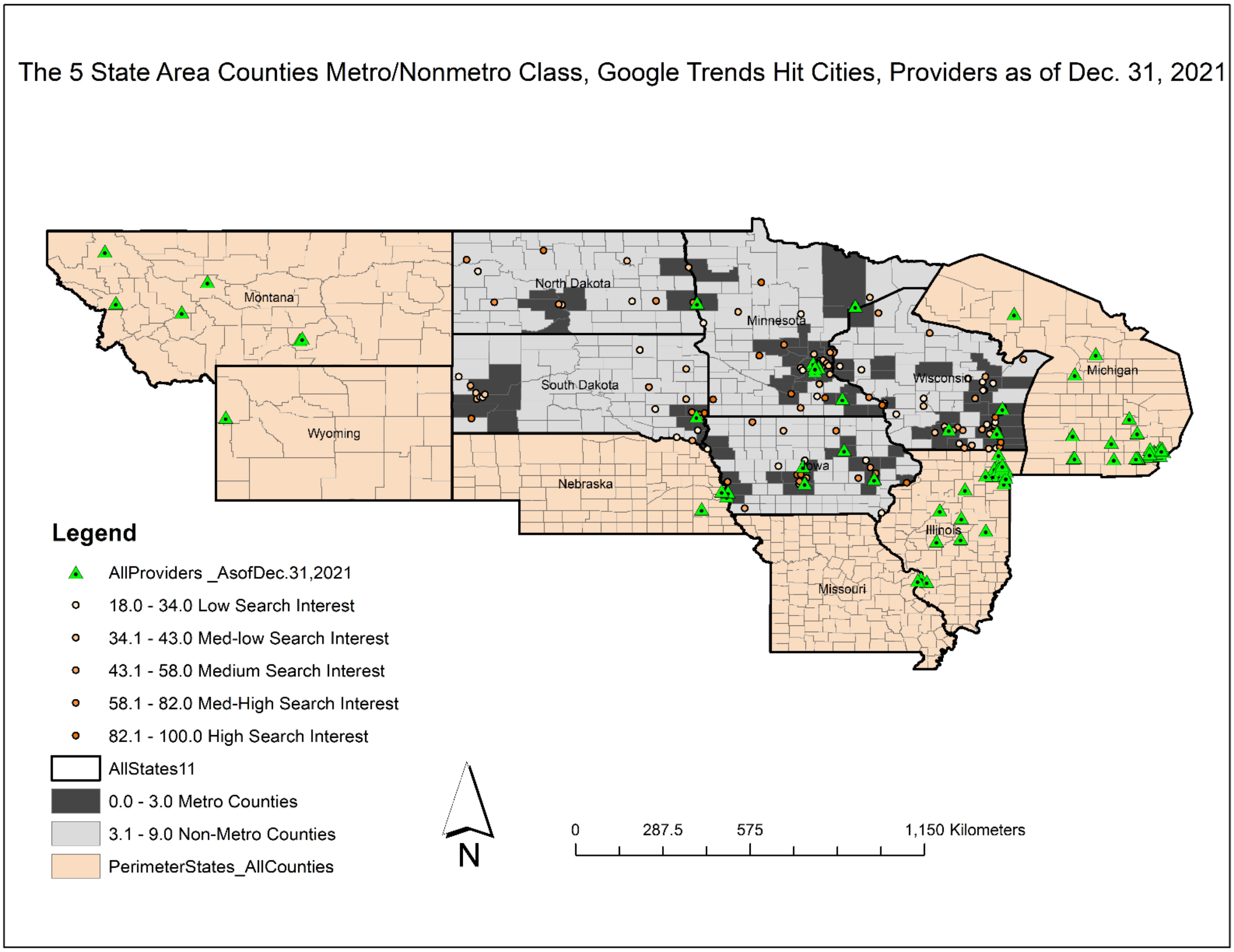
First and foremost, the mean 'hits' value of cities in metro counties had a higher mean than the mean hit values of cities in non-metro cities. My theory at the beginning of the project was that the opposite would be true and that those cities in non-metro counties would have higher average relative search interest for abortion related searches. This theory was developed because I thought that people who live in more rural areas would overall be searching the internet more for abortion and abortion related services. This was not the result that came from this portion of the analysis.

We are not surprised by these results as the metro counties have larger populations and more people who would search the internet for abortion information. These results suggest that there may exist a need to control for population within the areas of study. In the case of this project, it would be controlling hits values by population of the city to give a hits per capita value. This is a clear limitation of my project that I was not able to complete but should be considered in future research in this area. This finding is also supported by the metro and non-metro coefficient of -0.182 suggests that the relationship, although significant, is not very strong. This means that as the rural-urban classification becomes more rural, the search interest in abortion-related topics tends to decrease slightly. These findings are also supported by later linear regression models that show this negative correlation at the state level. Controlling for population by hit cities continues to be a prevalent limitation of this project.

These findings are supported by a significant p value from the correlation test that compared every RUCC value to each other. This test resulted in a weak negative correlation indicating that as the cities got more rural, hit volume went down. As mentioned, controlling each city's hit value for that city population would give a hits per capita value and it is possible that would give a more accurate reflection of the interest per person. Smaller cities/towns with the same relative search interest value as a capital city would then have more 'fair' impact in the analysis.

Overall, the analysis of this question shows the need for a standardization procedure that allows us to see hits per capita and allows us to compare large cities to small cities, and maybe this would impact the analysis of if cities in rural vs. Urban counties had different overall search interest in abortion related searches.

Map1



Question #2 What is the association between a city's Google Trends RSI value and their distance to nearest abortion provider?

One important note about this bit of analysis is that the Google Trends search data comes from the entirety of the year 2022. The abortion provider data (symbolized as green triangles) is dated as of the last day in the calendar year of 2021. The assumption I will make in this analysis is that all these distances are maintained the same as if nothing geographically changed over the year 2022. We now know that access in parts of the area was limited further. Some providers have since ceased to provide abortions, some clinics even closing due to local laws prohibiting abortion procedures. Knowing this, we understand that any estimates and results drawn here might very well be underestimates of true distances to nearest provider.

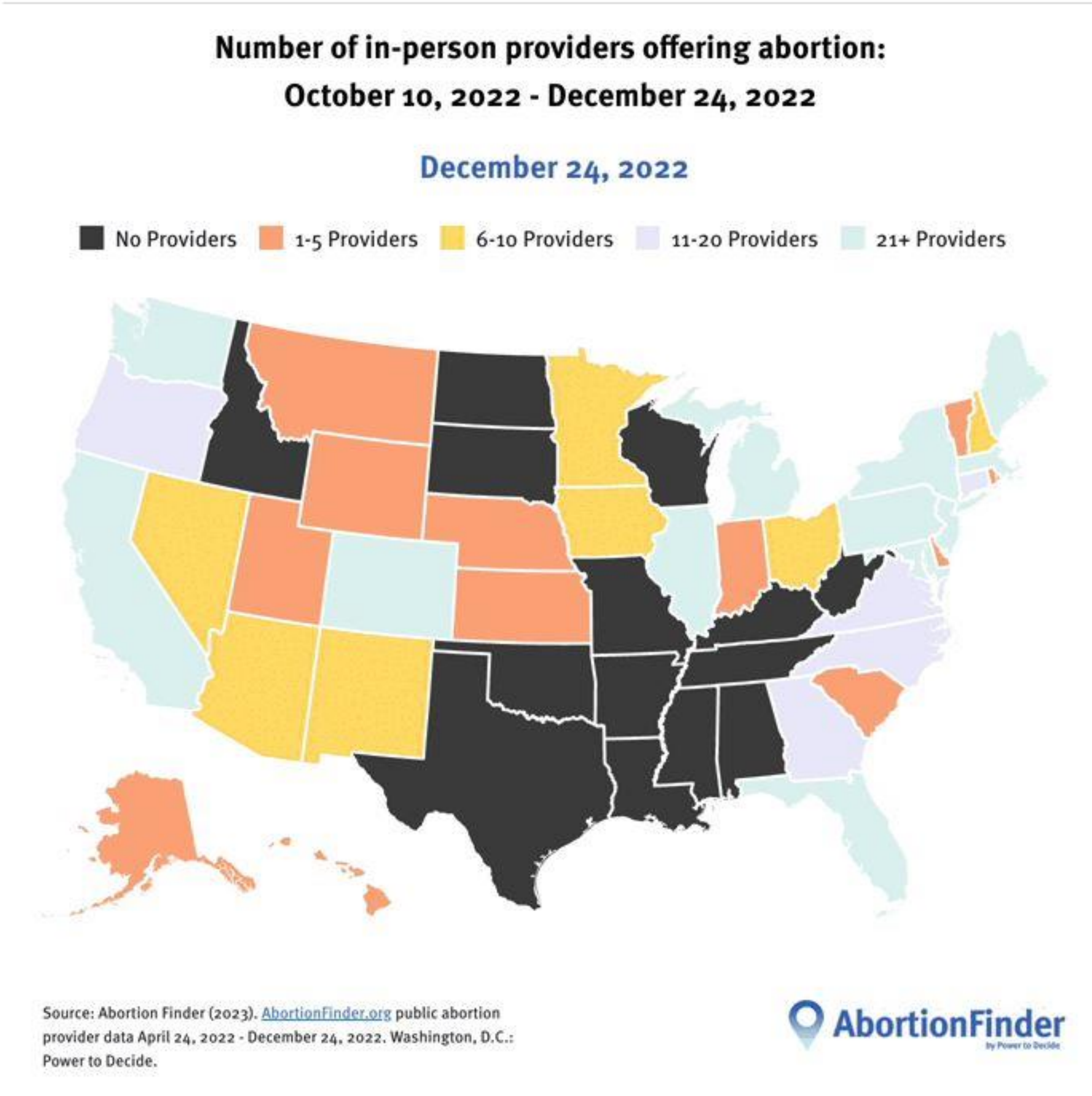
Graphics 6-8 show the relationship between the cities Google RSI hit value and that cities distance in kilometers to their nearest in-person abortion provider. Graphic 8 is a histogram of this data distribution and shows the right skewedness of the data. When used in conjunction with graphics 6 and 7 we can see that those hit cities furthest from nearest providers are the green and blue dots, indicating cities in SD and ND. More specifically these cities are in west central ND and SD where the nearest in person abortion provider would have been crossing all the way into Montana or all the way back east into Minnesota.

Interactive maps from the Guttmacher Institute data on abortionfinder.org show the changes in the abortion provider landscape over the course of the year 2022. This is important for the purpose of comparing my results to the real time context of abortion provision in the 5-state area. We can see in graphic 7 that there is a strong negative correlation between hits and distance to nearest provider in Wisconsin. Wisconsin however is surrounded by states that have maintained and protected abortion rights in the year 2022 including Minnesota, Illinois, and Michigan (Guttmacher, 2023, Abortion Finder, 2023).

Graphic 7 also shows us an interesting result of MN as the only positive regression line. This line indicates that in the sample of hit cities in MN, as the cities became farther from their nearest provider, the RSI value of those recorded hit cities increased as well. This is an interesting result but creates more questions than it provides answers. This may be the effect of not controlling city population but also maybe a bias introduced from the Google sample where the highest hit cities were recorded from MN.

Many of the MN providers exist in the south of the state near the capital leaving the large northern zone of MN with great distances to travel. Additionally, to the northeast is ND and SD, which do not offer any closer providers.

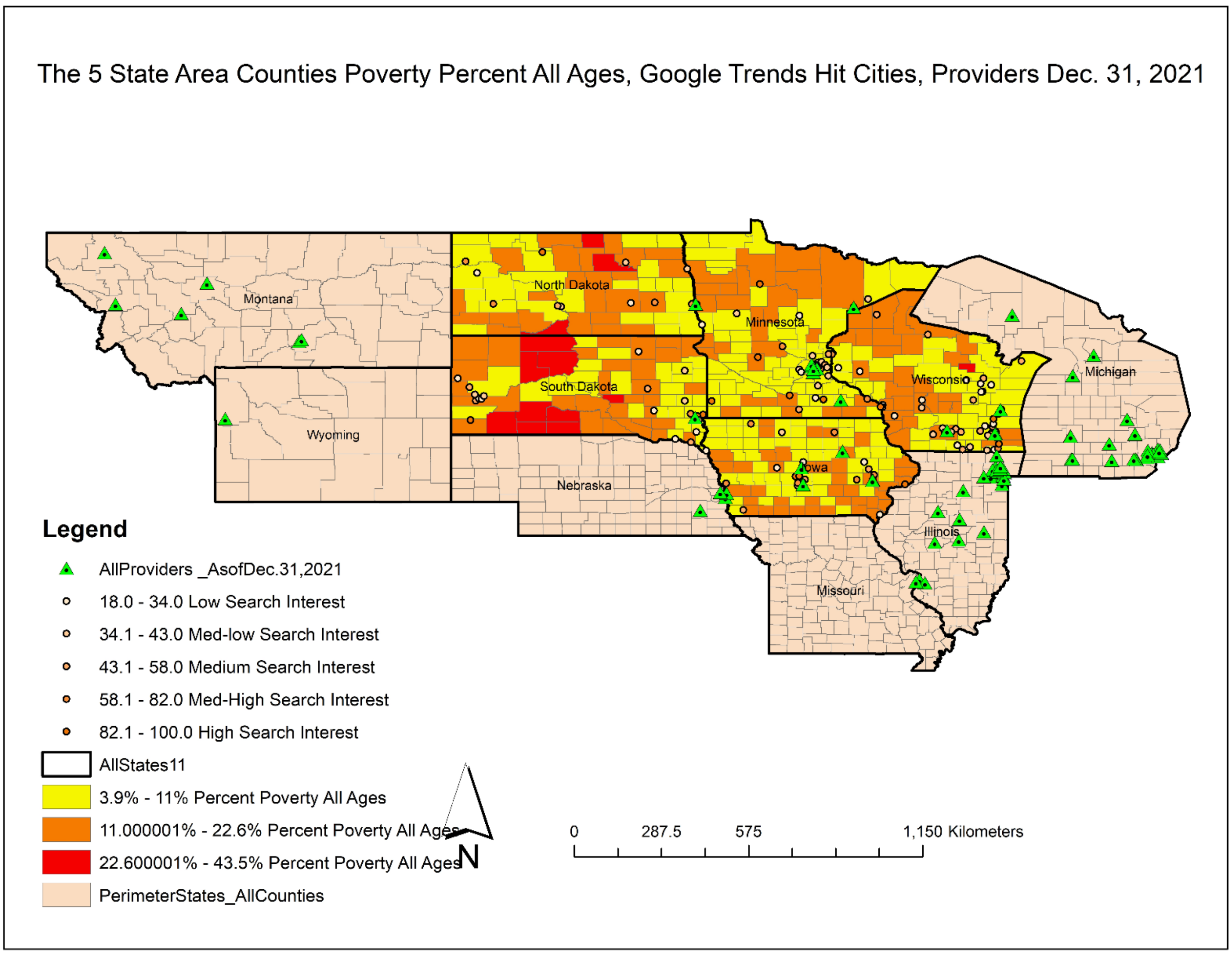
Map2



Question #3 What is the association between a city's Google Trends RSI value and their counties percent poverty for all ages?

The results of these analysis showed a weak positive but significant association between the 'hits' variable, and those cities corresponding percentages of the population living in poverty variable. This tells us that the sample is likely not random that as percentage of poverty increase, we can expect to see increase in ‘hits’ value searching Google for abortion or abortion services. This is also important to understand and continue to research given the political weaponization of abortion laws that occurred with the overturning of Roe v Wade during the time these data were being collected. Map3 shows the percentage of total poverty reported by the last census. Darker oranges and reds indicate a higher percentage of total poverty. Observed hot spots appear in the northern and central zones of MN, WI, ND, and SD. One important area for continued research would be to understand how these findings may impact different ethnic and indigenous groups? These results raise questions about the health equity of the national health system and could promote use and coordination with tribal health entities to improve health outcomes. We could begin the conversation of how national policy changes are affecting those populations living in various levels of poverty with relative access to health services?

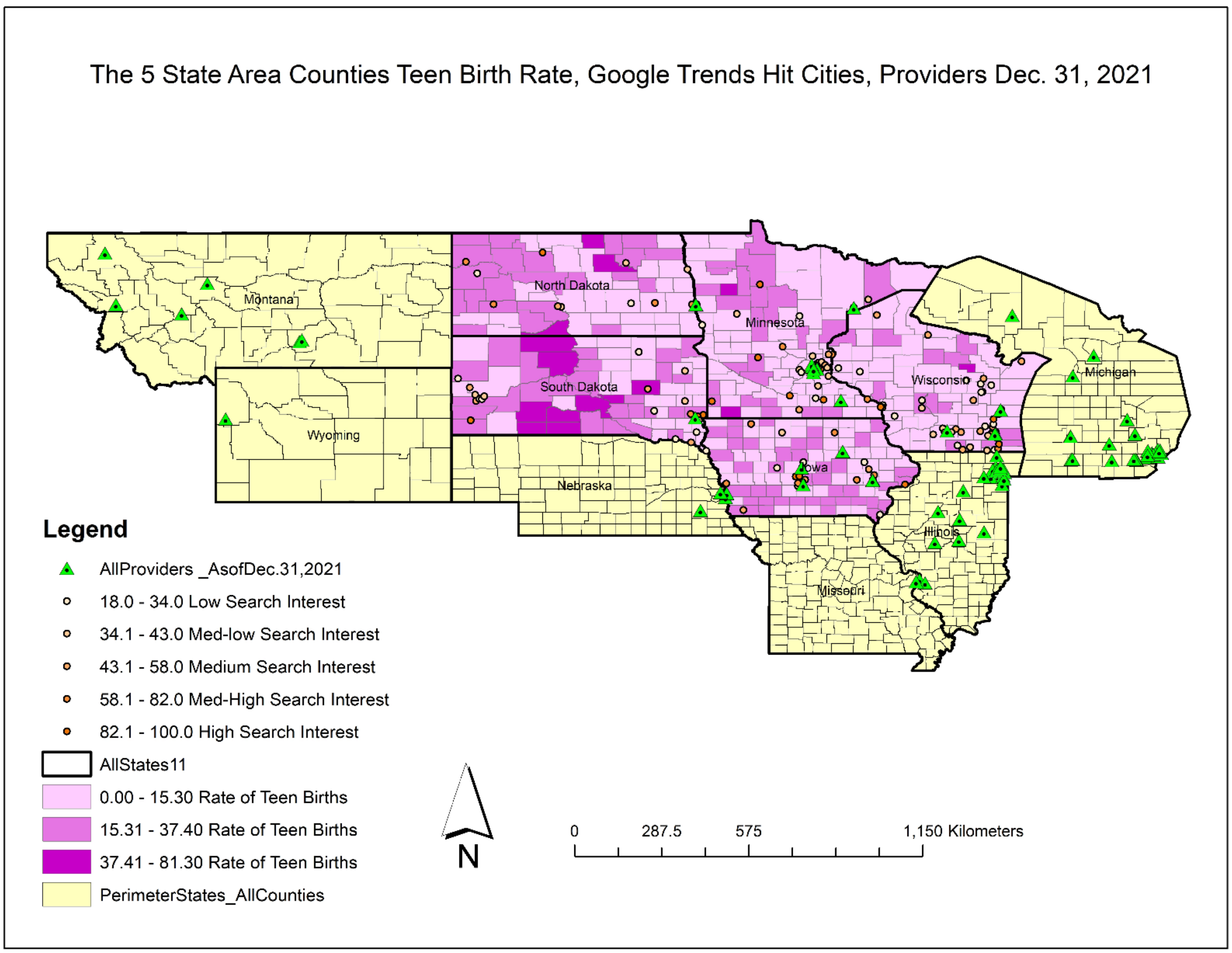
Map3



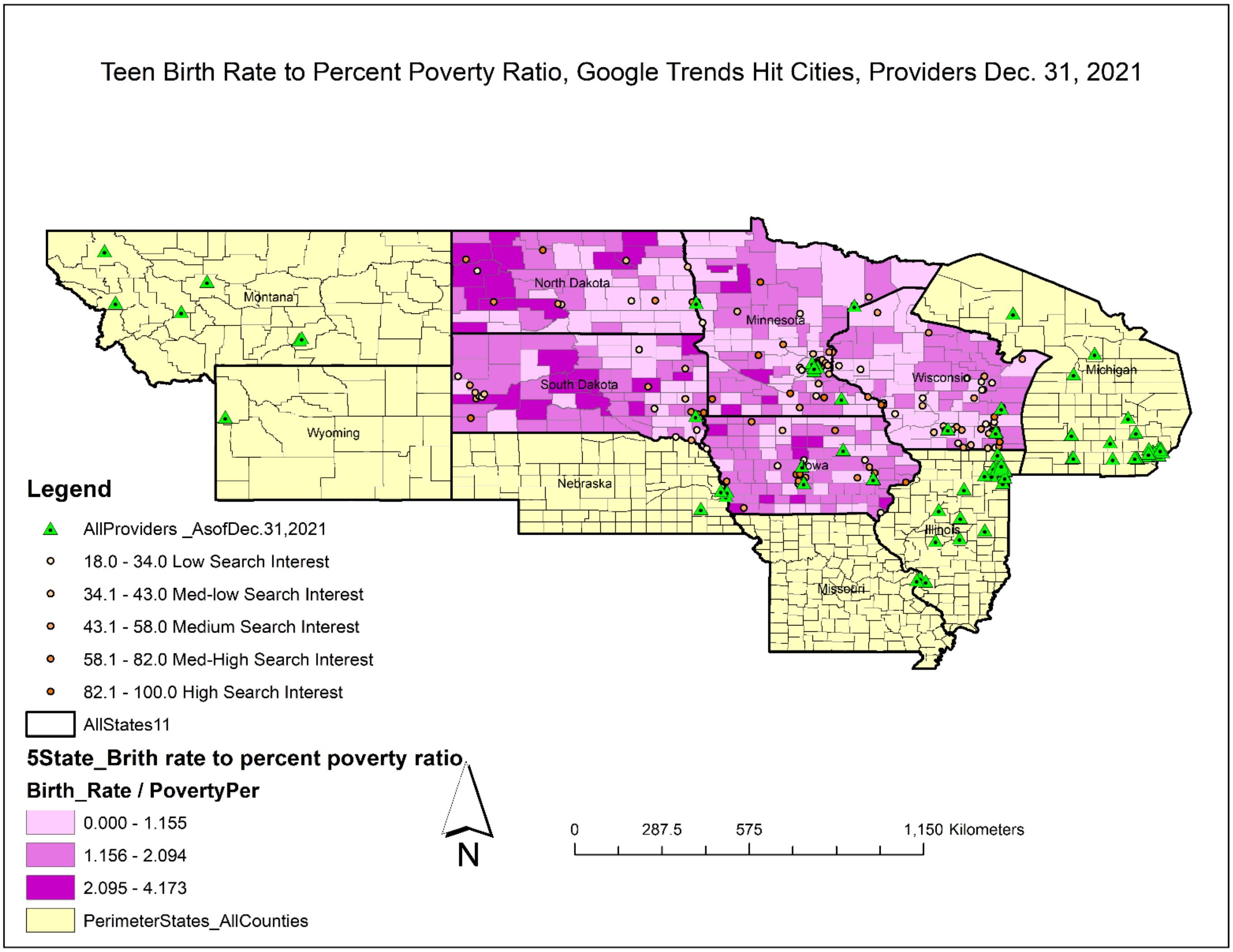
Question #4 What is the association between a city's Google Trends RSI value and their counties teen birth rate?

The final significant finding that resulted from the analysis of this project was the weak positive association recorded between the RSI hits variable and the teen birth rate of the county where that city exists. This suggests that we could consider a counties teen birth rate a determinant of the amount of perceived google search interest we expected from cites/towns like that one in that same county. This result makes some sense given that maybe teen pregnancies are not a health goal of the county and therefore more relative google searching going on for abortions and services. The finding is significant at the .05 level, but it is weak (explain) and so there exists a need to improve upon this monitoring technique to improve the quality of results. This is an important finding even considering the main limitation and issue with the model used in this project. Without controlling for hit city population, the model was still able to detect a significant positive association between counties teen birth rate and the Google search interest metric (RSI) of sample cities within that county.

Map4



Map5



STRENGTHS

There are some notable strengths to this project and overall study design. It is an observational study using majorly publicly available data from easy-to-use sources. The project could be conducted in RStudio which is free and open source which would allow similar analysis to be easily run, samples could be drawn, and results replicated mimicking methodology, likely improving on them. Another strength being that the data of RSI hit that was drawn is from recent years (2022) with important political implications taking place during the data draw period. Additionally, Google is one of the US most used search engine which makes the argument that the sample is a representative one.

LIMITATIONS

A limitation of Google trends analysis is that the reason for the search is unknown. We speculate that it is persons seeking information or services such as abortions, but exact reasons are assumed and supported by literature, but true reasons are difficult to confirm.

The most recent and complete data set on abortion providers we were able to obtain is that of the entire year (2021) prior to the year that the Google Trends data comes from (2022).

It is possible to have non-representative sample bias, since our data from Google Trends is just a sample. Other studies that informed the work done here had more comprehensive and ample sampling methods, one such study drew 30 independents samples and averaged the values. This would have been more work than I would have been able to complete alone.

FUTURE RESEARCH IN THIS AREA

After completing this project, it is my belief that well over half of the TRUE determinants of abortion related information seeking are still unknown to me. I can however speculate at some new variables that would inform future research and contribute to the "whirling vortex" of questions about mapping these determinants to reproductive health access.

A primary consideration is that when returning *interest_by_city* from 'gtrendsR' is that; the Google RSI hit value be divided by the population of that city, giving a 'hits_per_capita' variable. Without this step we are left with cities as small as 5000 people measuring an equal 75 in the 'hits' variable as a large city of 450,000 people. This is an obvious point of concern in replicating this type of analysis.

CONCLUSIONS

In conclusion, the project was a mixed bag of significant and nonsignificant findings based on the model used. Upon reflection there was meaning drawn from both results, we found significant determinants to people seeking abortion information on Google in 2022, as well as considerations to improve the metric used for measuring population interest. The significant results add to the knowledge base on the associations of abortion information seeking and poverty, as well as teen birth rates. Areas with high rates of poverty and higher rates of teen births are shown to be areas where there is significantly higher google search interest for abortions and abortion services. Health departments and health entities that exist in these areas are suggested to adopt creative ways in which to monitor and capture their populations interest in abortions or abortion related services online. These health enterprises could also engage with these populations online with creative outreach events, offering abortion medication online and over the phone where the landscape is in jeopardy with some politicians seeking to ban mailing abortion medications across state lines which would further limit access to specific populations. Political influences like these are the basis of our investigation, when a monumental decision happens like what happened with overturning Roe v. Wade, reproductive health researchers like ourselves asked the question, what are people googling as a response, and with what frequency, etc. Those who choose to replicate this type of model are encouraged to consider normalizing the ‘hits’ value from drawing results as ‘interest_by_city’ by that city’s population for the time of the sample draw. The implications of this are detailed in the limitations but briefly stated here, a town of 5000 people can score the same metric (RSI – hits) of interest as a city of 450,000 which casts these cities as more similar than different during analysis. A much more detailed picture is possible by dividing that RSI score by the total city population giving a ‘hits_per_capita’ variable allowing better comparison.

References - APA

1. Abortion Finder. (n.d.). Abortion Access Map. Retrieved March 8, 2023, from https://www.abortionfinder.org/abortion-guides-by-state/abortion-in-the-united-states/abortion-access-map#timeline
2. Aiken, A. R. A., Starling, J. E., van der Wal, A., van der Vliet, S., Broussard, K., Johnson, D. M., Padron, E., Gomperts, R., & Scott, J. G. (2020). Demand for self-managed medication abortion through an online telemedicine service in the United States. American Journal of Public Health, 110(1), 90–97. https://doi.org/10.2105/AJPH.2019.305369
3. ANSIRH. (n.d.). Advancing new standards in reproductive health. Retrieved April 3, 2023, from https://www.ansirh.org/.
4. Carneiro, H. A., & Mylonakis, E. (2009). Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clinical infectious diseases, 49(10), 1557-1564. https://doi.org/10.1086/630200.
5. Center for Reproductive Rights. (n.d.). Abortion Laws by State. Retrieved March 8, 2023, from https://reproductiverights.org/maps/abortion-laws-by-state/
6. Choi, H., Varian, H. (2012). Predicting the present with Google Trends. Economic Record, 88, 2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x.
7. ESRI. (2020). ArcMap (Version 10.8.2) [Software]. Redlands, CA: Environmental Systems Research Institute.
8. Google. (n.d.). Explore data with Google Trends. Retrieved April 3, 2023, from https://support.google.com/trends/answer/4365533?hl=en
9. Grossman, D., Grindlay, K., Buchacker, T. R., & Lane, K. (2011). Effectiveness and acceptability of medical abortion provided through telemedicine. Obstetrics and Gynecology, 118(2 Pt 1), 296-303. https://doi.org/10.1097/AOG.0b013e31822598ba
10. Guendelman, S., Pleasants, E., Cheshire, C., & Kong, A. (2022). Exploring Google Searches for Out-of-Clinic Medication Abortion in the United States During 2020: Infodemiology Approach Using Multiple Samples. JMIR Infodemiology, 2(1), e33184. https://doi.org/10.2196/33184
11. Guendelman, S., Yon, E., Pleasants, E., Hubbard, A., & Prata, N. (2020). Shining the light on abortion: Drivers of online abortion searches across the United States in 2018. PloS one, 15(5), e0231672. https://doi.org/10.1371/journal.pone.0231672
12. Guttmacher Institute. (2018). Abortion worldwide 2017: Uneven progress and unequal access. https://www.guttmacher.org/report/abortion-worldwide-2017-uneven-progress-and-unequal-access.
13. Guttmacher Institute. (n.d.). State Policy Landscape: Abortion Policies in Minnesota. Retrieved March 8, 2023, from https://states.guttmacher.org/policies/minnesota/abortion-policies
14. Hoyert, D. L. (2021). Maternal mortality rates in the United States. Division of Vital Statistics. National Vital Statistics Reports, 70(2), 1-33.

15. Jerman, J., Onda, T., & Jones, R. K. (2018). What are people looking for when they Google "self-abortion"? Contraception, 97(6), 510-514. https://doi.org/10.1016/j.contraception.2018.02.006.
16. Kodjak, A. (2022, June 24). Supreme Court Won't Overturn Roe v. Wade, But Upholds Mississippi's Abortion Ban. NPR. Retrieved March 8, 2023, from https://www.npr.org/sections/health-shots/2022/06/24/1107126432/abortion-bans-supreme-court-roe-v-wade
17. Maddow-Zimet I and Kost K, 2022. Even before Roe was overturned, nearly one in 10 people obtaining an abortion traveled across state lines for care, New York: Guttmacher Institute, 2022, https://www.guttmacher.org/article/2022/07/even-roe-was-overturned-near
18. National Center for Health Statistics. (n.d.). National Vital Statistics System (NVSS). Retrieved April 3, 2023, from https://www.cdc.gov/nchs/nvss/index.htm
19. National Partnership for Women & Families. (n.d.). Status of Women in the States: State Data. Retrieved from https://statusofwomendata.org/state-data/
20. Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I., ... & Ross, J. S. (2014). The use of Google Trends in health care research: a systematic review. PloS one, 9(10), e109583. https://doi.org/10.1371/journal.pone.0109583.
21. Poliak, Satybaldiyeva, N., Strathdee, S. A., Leas, E. C., Rao, R., Smith, D., & Ayers, J. W. (2022). Internet Searches for Abortion Medications Following the Leaked Supreme Court of the United States Draft Ruling. Archives of Internal Medicine (1960), 182(9), 1002–1004. https://doi.org/10.1001/jamainternmed.2022.2998.
22. Reis, B. Y., & Brownstein, J. S. (2010). Measuring the impact of health policies using internet search patterns: The case of abortion. BMC Public Health, 10(1), 514-514. https://doi.org/10.1186/1471-2458-10-514.
23. Rouder, J. N. (2021). gtrendsR: R functions to perform and display Google Trends queries. R package version 1.4.5. https://CRAN.R-project.org/package=gtrendsR
24. RStudio Team. (2023). RStudio (Version 1.5.679) [Integrated Development Environment]. Boston, MA: RStudio, PBC.
25. Seifter, A., Schwarzwalder, A., Geis, K., & Aucott, J. (2010). The utility of "Google Trends" for epidemiological research: Lyme disease as an example. Geospatial health, 4(2), 135-137. https://doi.org/10.4081/gh.2010.213.
26. Stephens-Davidowitz, S. (2012, June 9). How Racist Are We? Ask Google. The New York Times. https://archive.nytimes.com/campaignstops.blogs.nytimes.com/2012/06/09/how-racist-are-we-ask-google/
27. Stephens-Davidowitz, S. (2013, December 7). How Many American Men Are Gay? The New York Times. https://www.nytimes.com/2013/12/08/opinion/sunday/how-many-american-men-are-gay.html.
28. United States Census Bureau. (n.d.). Small Area Income and Poverty Estimates (SAIPE). Retrieved April 3, 2023, from https://www.census.gov/programs-surveys/saipe.html

29. United States Department of Agriculture. (n.d.). Rural-Urban Continuum Codes. Retrieved April 3, 2023, from https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx