Third (or Fourth) and Goal: A Comparison of Google Search Interest in 'CFL' vs. 'NFL' Related Keywords from 2004-2014

By

Jason Buzzell

Submitted to the Faculty of Extension

University of Alberta

in partial fulfillment of the requirements for the degree of

Master of Arts in Communications and Technology

July 1, 2015

Acknowledgments

I would like to thank a lot of people but I only have one page to acknowledge so many of you, who helped me not only in my academic life, but personally and otherwise. This is merely a snapshot.

First, I'd like to thank my dad. You pushed me to pursue my Master's, and I finally did it. You never stopped encouraging me to get started. You have always been, and continue to be, an advocate for lifelong learning. To my mom and my late grandfather, thank you for bringing me to Canada, my home it will always be. Mom, you gave me the drive and work ethic to make this all possible through your early career days and your many miles driven to help me succeed in hockey and life. To Maynard (Grampy Banks), I wish you were here to hang out with us.

To UAlberta, Prof. Gordon Gow and all of the professors who helped make the MACT program excellent, engaging and personal. A special thanks to Prof. Rob Shields for his patience, and his encouraging words in March when I felt like I was stuck beyond belief in lit review hell. Additionally, many thanks to those in University Relations and all across campus who were a pleasure to work with from 2012 to 2015. To all my Edmonton family: I can't name one without naming you all. My classmates, co-workers, billets, friends, teammates and neighbours. You all helped me and my family in some way or another to make this possible. I will, until the day I die, be an Edmontonian through and through because of all of you.

Finally, my wife. I owe her more than anything money can buy. For watching the kids on Sundays while I toiled in six hours of homework, or the three weeks of solitude during each of the spring institutes. I love you. Without you, this work and my life mean nothing.

Table of Contents

Abstract

Google Search Trends are a growing area of interest for organizations and companies. Marketing and Internet news teams monitor trends to gauge interest in a variety of topics both for content ideas and how much a particular topic resonates with the public. This research project involves content analysis to identify trends related to Canadian Football League (CFL) and U.S. National Football League (NFL) keyword use in Canada from 2004 to 2015.

Using the foundations of issue salience, or the prominence of a particular topic in the public, this research focuses on current and past search volume trends for similar queries since 2004, when data began being collected, for NFL and CFL Past research in other topical areas has determined that issue salience can be judged from using Google search interest over time to identify trends. By applying similar methodologies, and shedding light on the comparisons in Canada of interest relating to CFL and NFL, this project could bring about more research into the effects on the country, and the business of sports or media in Canada.

Introduction

The purpose of this study is to examine interest in professional football in Canada by comparing keyword search data used by Canadians back to 2004. Through comparison of keyword data between Canadian Football League (CFL) and National Football League (NFL), the research seeks to show the differences in salience between the two leagues and provide further context to an overall pattern of rising salience.

Understanding the trends and measuring the salience of both leagues will support further

research that cites agenda setting theory as an explanation for these increases. Agenda setting theory is the idea that more media and publicity through traditional broadcast mediums increases public interest, and thus salience, in a particular topic. By using another method of measuring salience of both leagues, researchers can explore differences and similarities in future studies to explain why a trend may be occurring.

Background

CFL and NFL in Canada

The CFL is a nine-team professional football league in Canada, consisting of teams in Vancouver, Edmonton, Calgary, Regina, Winnipeg, Toronto, Hamilton, Montreal and Ottawa (CFL.ca, 2014). The league plays an 18-game regular season, culminating in a playoff in which the championship team is awarded the Grey Cup, a tradition dating back to 1909 (CFL.ca, 2014). The NFL, on the other hand, is a 32-team league with teams located only in cities of the United States as of 2014. Teams play for the Vince Lombardi Trophy at a yearly event called the Super Bowl, an event that dates back to 1967. More than 100,000 Canadians attend NFL games and 700,000 watch games on television each year (Canada.NFL.com, 2015). Though the leagues have gone through ebbs and flows of being in direct or indirect competition, they now work together to grow the game and "strengthen interest, awareness and participation in football among Canadians" (Canada.NFL.com, 2015).

According to NFL Canada, since 1997, "[the Canadian branch] handles all of the National Football League's business interests in Canada. One of the National Football League's

goals, on a multinational basis, is to support football wherever it is played. NFL Canada, formed in 1997 as a division of NFL International, has also serviced this mission. Canadians enjoy a rich history of watching and playing professional football, with both the National Football League and the Canadian Football League. The National Football League, through NFL Canada, makes complementing and working in partnership with the Canadian Football League a priority. Both leagues share the goal of strengthening the interest, awareness and participation in football among Canadians." Although there are no Canadian teams in the NFL, a handful of Canadian players play in the league. Also, from 2008, a game was played in Toronto, Ontario, once per year between the Buffalo Bills and an NFL opponent, coined the Bills Toronto Series. Six regular-season games and two preseason games were played as part of the series until the agreement was ended in 2014 (BuffaloBills.com, 2015).

Google Search and Google Trends

Google, a technical product company founded in 1998 with a primary focus on Internet search, generates a majority of its revenue from a service associated with search called Adwords (Financial Tables, 2015). In 1996, Google didn't exist. The Internet as we know it barely existed. Startups and companies rose and fell, many without any viable economic means, including a brand-new search engine called AltaVista that garnered 300,000 search queries on its first day with no marketing or formal announcement (Battelle, 2005). AltaVista stayed focused on search until its overarching company, Digital Equipment (DEC), was bought by Compaq in January 1998 (Battelle, 2005).

For a variety of reasons, AltaVista never again hit the peak it had in its early days, and other engines such as Excite, Lycos and Yahoo!, among many others, came and went, some managing to hang around for many more years (Battelle, 2005). But it wasn't until Sergey Brin and Larry Page met in school, and founded what is today known as Google in 1998, that search became something bigger—although the duo was not a match from the start, neither in terms of their personalities nor their passions in the classroom (Battelle, 2005).

> The World Wide Web, Page theorized, may have been the largest graph ever created, and it was growing at a breakneck pace. One could reasonably argue that many useful insights lurked in its vertices, awaiting discovery by inquiring graduate students. Winograd agreed, and Page set about pondering the link structure of the Web. It proved a fruitful study. Page noticed that while it was trivial to follow links from one page to another, it was non-trivial to discover links back. In other words, when you looked at a given Web page, you had no idea what pages were linking back to it. This bothered Page. He thought it would be very useful to know who was linking to whom. After all, very important people might be linking to you—if so, wouldn't you want to know that? (Battelle, 2005, pp. 68–69)

Using this insight as a basis, Page and Brin set out to discover links and store them for analysis. Using a crawler, as well as a similar foundation to how academics cite sources, Google began crawling as best it could and using a logarithm to display results (Battelle, 2005). Over time, these results would become more and more instantaneous, and information was not gated to

users, meaning content was delivered unbiasedly to whomever needed it, and whenever the person needed it (Hillis, Petit, & Jarrett, 2012). As Hillis et al. note:

> Online search displays information within moments of being sought and that, by the same logic of immediacy, also can disappear instantly back into the index, or cloud, from whence it came. What Google's model of search proposes is, then, not accretion of knowledge but the immediacy and ephemerality of information retrieval. (2012, p. 1733)

Google also focused on technical insights, such as applying a new design to decrease cost or increase usability (Schmidt, 2014). Schmidt notes this is one of the reasons Google began to create products that were better than those offered by the competition, reasoning that:

> The best products had achieved their success based on technical insights, not business ones, whereas the less stellar ones lacked technical distinction. Our brand had gotten strong enough that any product we launched would gain a certain amount of market momentum just by virtue of it coming from Google. If we measured success by number of users, we could (and did) trick ourselves into believing the products were successful. (2014, p. 73)

The theory of crawling and building off of technical insights is what ended up setting Google apart from other competitors in the market (Schmidt, 2014). It also used the technical insights of geography and location to increasingly tailor the results to the individual based on the searcher's IP address (Hillis et al., 2012). Hillis et al. argue that the act of search is increasingly easier and costs far less with little friction, enabling an experience between the searcher's

questions and the immediate answers to whatever they are interested in knowing (Hillis et al., 2012). Hillis et al. continue to focus on personalization as a key to the Google advantage:

> Its search results gain validity from the performative power of their own "findability" and immediate utility to a specific searcher and not from being based on access to any coordinated sets of knowledge per se. (2012, p. 1763)

The combination of looking for ways to find a competitive advantage, such as speed and personalization, and using its core product as a catalyst led to other advancements by using the same principles. Google made strategic decisions that in hindsight separated it from competitors, but the link structure was critical, as Schmidt notes:

> There were a lot of other factors that made Google Search so much better than the competition when it launched—for example, it placed more faith in results found on academic websites—but the heart of the product's advantage consisted of this single technical insight about using the web's link structure as a roadmap to the best answer. Since then, most of Google's successful products have been based on strong technical insights, while most of the less successful ones lacked them. (2014, p. 70)

This idea of technical insights is one of the driving factors that eventually led to Google Trends in the mid-2000s, born out of Google's idea to have 20 per cent of the work week dedicated to a project of its employees' choosing—the same idea that led to other products such as Google Maps, Google Transit and Google Reader (Schmidt, 2014). Schmidt notes that this "20 per cent time" was a bit misunderstood but did lead to more insights and better products:

> The program doesn't mean that the campus turns into a summer camp every Friday, with all the engineers going off in (hopefully) creative ways. In fact, 20 per cent time is more like 120 per cent time, since it often occurs on nights and weekends. But it can also be stored up and used all at once. (2014, p. 226–227)

With the 20 per cent time rule in place, it wasn't until 2006, after Google had grown and monetized its business model, that it released Google Trends (Google.com, 2015). One of many public tools Google offers, Trends analyzes the amount of search on words called "keywords" that are typed in by users to query for webpages that could help answer their questions.

> Google Trends analyzes a percentage of Google web searches to determine how many searches have been done for the terms you've entered compared to the total number of Google searches done during that time. For example, if you search for tea in Scotland in March of 2007, Trends analyzes a percentage of all searches for tea within the same time and location parameters. (Trends Help, 2015)

The data is normalized to make it easier to compare (Trends Help, 2015). "Normalized means that sets of search data are divided by a common variable, like total searches, to cancel out the variable's effect on the data. Just because two regions show the same number for a particular search term doesn't mean that their absolute, or total, search volumes are the same. Data from two regions with significant differences in search volumes can be compared equally because the data has been normalized by the total searches from each region," according to Google Trends Help.

This normalization of the data is key to the year-over-year comparisons necessary for this research. Although there are potential issues concerning sampling, decreased usage in Google or increased usage in other social media categories, preliminary research in this area, and our study, did not see this occurring, but did note it as a potential issue with this kind of research (Carrière-Swallow & Labbé, 2010). Carrière-Swallow and Labbé note the advantages and how normalized data works using Google Trends:

> The raw data undergoes two transformations prior to public release. First, the data are normalized by the total number of search queries in the geographical region of interest. As such, any trends from growth in the total number of Internet users or from a change in the relative popularity of Google as a search engine are removed from the data. Second, the normalized data are rescaled to an index with a maximum value of 100. This means that magnitudes are not directly comparable across series as a measure of relative popularity (2010, p. 2).

Increased Interest and Why It Matters

Traditional public opinion surveys have indicated an increase of interest in professional football, including the CFL and NFL. In 2013, interest increased by about 10 per cent for both CFL and NFL since 2005 (Bibby, 2013). The increased interest comes off the back of several initiatives and increased media exposure for both leagues in Canada. In 2008, Bell Media's TSN became the exclusive rights holder for all CFL games and the league's Grey Cup Championship (TSN.ca, 2014). TSN recently re-upped its deal and offered the league a reported $40 million per year (TheGlobeandMail.com, 2013). The NFL has signed TV deals with both Bell and Rogers to

showcase the game. They also have hosted a regular-season NFL game in Toronto in 2008 (CBC.ca, 2014).

Increased interest in professional football is creating revenue opportunities for Canadian media, and thus the leagues, that neither property has seen before. Although traditional public opinion survey methods show interest to be somewhat equal and increasing at the same rate (Bibby, 2013), search volumes may indicate a different story. The slower growth, even stagnation, in CFL search interest may indicate potential issues of stress for the Canadian league, whereas the recent rapid increase in NFL-related search terms points to a new phenomenon in Canadian society, as well as increased competition between the two leagues for viewers and advertisers. Or, it may reinforce the past strategy that both leagues can benefit from a shared partnership and using the media deals to increase interest and revenue to help increase owners' and players' bottom lines.

The basis of theory for this paper will start from using the concepts of content analysis to analyze keyword data from Google Search using Google Trends. It builds off the concepts of agenda-setting theory, "a robust and widespread effect of mass communication ... that results from specific content in the mass media" (McCombs, 2004). Also, the study will look at issue salience as it relates to information seeking, as we juxtapose how public opinion surveys done in the past, and search interest between the NFL and CFL in Canada during the past 10 years, have differed.

Objectives of the Research

The objective of this research is to begin to quantify the popularity or salience of CFL vs. NFL leagues in Canada and allow future researchers to juxtapose these numbers against previous research conducted using traditional polling methods such as surveys. By combining this quantitative analysis and past qualitative research results, future studies will afford a better grasp on the current and past state of salience regarding each of the leagues. The data presented in this study, and the conclusions made, pose interesting questions that require more research to dig into why NFL interest appears to be rising so rapidly in Canada. Popular belief and prior research often point to increased media exposure using agenda setting theories. Due to constraints in time and length, this research will not be able to identify specific media changes, both increases or decreases in content, or the rise in fantasy leagues to see whether these are contributing or key factors in the increased interest. The hope is that this study will provide further context and inspiration for such research.

Additionally, digital information seeking using search is becoming increasingly common among researchers (Kadali, 2015). People often note websites and digital content as the main source of information they are seeking (Percheski & Hargittai, 2011). The methodologies could also be applied to other topics of interest to gauge public sentiment or interest across Canada and in certain regions. This could play a powerful role in future studies analyzing, for example, the current trend of nationalism and globalization challenges that confront Canadian governments, news organizations and companies.

The Research Question

What is the difference in search volume trends between certain NFL- and CFL-related queries in Google search since 2004 in Canada?

Literature Review

Content Analysis

Content analysis is critical to this research. This analysis helps minimize interviewer biases and applies a broad framework of analyzing or collecting data that can help reveal potential trends (Krippendorff, 1980). Content related to a many topics is increasingly being published to the web through vehicles including online magazines and newspapers, social media and other sites, and with the Internet becoming more and more accessible, it has become useful to analyze this abundant content to answer a research question or problem (Oltra, Delicado, Prades, Pereira, & Schmidt, 2014). As noted earlier, by using a history of clicks and searches, Google essentially houses a real-time "database of intentions" (Battelle, 2005). Battelle notes that the digital world differs from print because news used to be consumed then talked about, whereas digital content is the actual conversation. Given this premise, it is clear that a content analysis—essentially in the case of a conversation, as Battelle notes— could make sense:

> The Database of Intentions is simply this: the aggregate results of every search ever entered, every result list ever tendered, and every path taken as a result. It lives in many places, but three or four in particular—AOL, Google, MSN, Yahoo—hold a massive

> amount of this data. Taken together this information represents a real-time history of post Web culture—a massive clickstream of database of desires, needs, wants, preferences that can be discovered, subpoenaed, archived, tracked and exploited for all sorts of ends. Nearly any question one might frame can be answered in one way or another by mining the implacable Database of Intentions that is building second by second across the Internet. (2005, p. 6)

For example, one piece of research analyzed forum content to determine a framework for assessing content (Marra, Moore, & Klimczak, 2009). Many other studies use content analysis of communication to identify trends and current issues. Some include social media such as Twitter hashtags. One study looked at #legalize on Twitter and used qualitative content analysis to provide meaning around each tweet and what category it fit into to evaluate the campaign's latency, which is the importance or prominence of a campaign (Komori, 2013). Another evaluated headlines as positive, negative or neutral using content analysis surrounding the use of the human papillomavirus vaccine (Habel, Liddon, & Stryker, 2009). Habel et al. note the dangers of doing qualitative content analysis during peak periods, because headlines could be more or less negative depending on the news. However, this research did not use Google Trends to actually perform a quantitative content analysis, nor was it used to scan headlines during seasonal peaks in the CFL and NFL when search interest is related to positive moments in the leagues, such as starts or endings of the season.

Some of the closest research to this study includes using Google Trends data to identify popularity or salience in different subject matter. Related to this specific research for

comparison, traditional opinion polls show relatively equal growth of interest in CFL and NFL, six per cent versus eight per cent growth respectively, over the past eight years, as well as slightly more interest in CFL than NFL, 26 per cent versus 21 per cent respectively (Bibby, 2013). By judging the content of communication—Internet search in this case—one can make "valid inferences" about the context of a communication trend regarding the "what" that is happening (Krippendorff, 1980). The volumes of search are thus an effective measure of issue salience or interest in a particular subject—in this case, sport (Wilde & Pope, 2013, pp. 211–222).

Contrasting traditional polls to salience using Google Trends, identification of topics and their popularity is considered a basic indicator of judging the popularity of events and then tailoring blog content appropriately (Duen-Ren, 2015). Many factors affect the salience of political topics, which have been the basis of much of the research using Google Trends (Mellon, 2014). Unlike politics, football seasons remain cyclical and predictable from beginning to end, especially given the availability of data over weeks and months for almost 10 years, whereas politics and issues can spike unpredictably, making it more difficult to calculate salience of a hot political topic or public concern at certain peak times (Mellon, 2014). Mellon reinforces this point by stating, "Any number of factors may affect the probability of an individual translating concern about an issue into Google searches: strength of partisanship, political interest, political knowledge, or strength of interest all could play a role in the link between individual concern and search behaviour." Worth noting is the difference between public interest and public opinion, as noted by Ripberger:

> Distinguishing between public attention and public opinion is necessary because there are times when the two variables are discernibly different. For instance, an individual might have exceptionally strong opinions about anthropogenic global warming, but after having formed these attitudes, they need not spend any time thinking about the issue. Rather than erroneously assuming perfect correspondence between potentially distinct concepts, the relationship between attention and opinion is an empirical question subject to verification (2011, p. 240).

Salience, concern and curiosity—or, as noted above, differences in opinion and casual interest from actual salience and engagement—are worth noting. Researchers have noted this issue and focused on more accessible ways to measure attention by media coverage (Ripberger, 2011). It is worthwhile to state that coverage and access have increased substantially for both the CFL and NFL in Canada since 2004, so simply monitoring the amount of media coverage, as opposed to measuring actual user queries, would bring much different results from those included in this paper.

Finally, the benefits of simply monitoring media coverage to measure public interest related to costs is also a great benefit of using Google Trends to perform the content analysis, but equally important is noting the deficiencies of traditional research using media mentions and polls (Ripberger, 2011). These deficiencies include not only that polls are costly, but also that people cannot recall or do not answer polls correctly, thus skewing results (Curtice & Sparrow, 1997). Ripberger states that using Google Trends, in what he refers to as a supply-based measuring system, has proven effective in past studies, also noted in this paper, and that media-

based coverage measuring systems can be challenged due to their subjectively and sample sizes. He notes:

> Given that the majority of questions require research over relatively lengthy periods of time, most researchers refrain from reading and coding every story within their frame of analysis. Rather, scholars generally employ theoretically guided (keyword) searches of extensive online archives such as the *New York Times Index* or *Readers' Guide* to identify the number of articles dedicated to a particular issue within a finite unit of time. (Ripberger, 2011, p. 241)

Despite these noted advantages over traditional polling and measurement systems, further research is required to discover why people have initiated the search. What is the practical context in everyday life from which the search springs? What is the value context that motivates the search? (Or, to put it another way, what if a given searcher has already decided to take the "against" position on a given issue and is collecting information to support that position, or is closer to a neutral position of merely informing themselves about a topic?) More research will also need to be done to evaluate the differences between topics. As noted earlier, it may be more difficult to use this supply-based measuring system for things such as political or entertainment interest, wherein certain topics encourage the public to seek out information out of "morbid curiosity" (Ripberger, 2011).

Scharkow and Vogelgesang also reference this, as the difference between program- and issue-related search. Google Trends has proven it is well known, for example, for providing a test of salience as an indicator of pragmatic need. Flu-related search terms correlated with

prevalence of influenza in regions of the United States is one example showing a high degree of salience (Araz, Bentley, & Muelleman, 2014). There is a high probability that people searching for information on the fly are in a situation of having to deal with influenza at that moment (Araz et al., 2014). Scharkow and Vogelgesang note the challenges with traditional polling methods and why, despite some of the different reasons people search, information seeking and issue salience can be connected:

> Methodologically speaking, salience represents an unobservable latent variable, whereas immediate audience responses such as salience-driven media use represent observable manifest variables. When immediate audience responses emerge from issue salience, there should be a relationship between the two concepts. Consequently, if there is a strong correlation between online information seeking and issue salience, the former could be used as a convenient proxy for the latter. (2011, p. 104)

The merits of traditional polling and measurement of salience have been debated in much research over time (Scharkow & Vogelgesang, 2011). Although using Google Trends specifically and content analysis to measure salience is new, researchers have defended the methodology. They argue the large aggregate numbers of search queries will balance out any variabilities over time to thus indicate a measure of the salience and public agenda appropriately (Scharkow & Vogelgesang, 2011). Wilde and Pope, as well as Scharkow and Vogelgesang, examine these challenges of traditional polling methods and how content analysis of Google Trends can be a more accurate measure of public interest in a particular subject. As noted, there are some topics, depending on the keywords and queries, that lend themselves better to this

methodology. Focusing on simple queries and words clearly identifiable to the issue or topic at hand is needed for this methodology to work well (Scharkow & Vogelgesang, 2011). On the other hand, no one methodology is necessarily always the best to choose. For example, Scharkow and Vogelgesang note a case in which traditional polling left the question open-ended so as not to lead the subject to answer, but as a result, the answers and questions became so "multi-dimensional," it was difficult to analyze salience:

> The open-ended question asked was: 'Can you recall any important issues recently covered in the news media that interest you?' and the responses were entered as free text by the interviewers. The wording of the question is multi-dimensional and does not clearly capture one of the processes described in the introduction. Rather unfortunately, this question measures issue awareness ('recall'), salience ('important issues'), and interest at the same time. We do not expect this to be an issue for correlational analyses but caution against inferring absolute levels of issue salience of the question. (2011)

As mentioned earlier, most of the research using content analysis uses a systematic qualitative categorization of press articles or content published either in print or digital mediums. One sports-related piece of research looked at a specific news source during a time when black segregation ended in Major League Baseball to analyze the sentiment of articles as the league changed (Carroll, 2008). This was deemed a proven methodology because it was reading objective news in prominent black publications during peak times, generally during the season (Carroll, 2008). This is similar to some aspects to the methodology used in this research paper.

The positives of such research were a time-tested approach, using a medium that had

essentially not changed much until the invention of the Internet. Content analysis focused on where articles ran, how often and in what fashion. Stories were then subjectively measured and categorized as negative, positive or neutral. One could evaluate the changes in black press coverage before, during and after integration (Carroll, 2008). What couldn't be judged was the salience or interest among the general public in certain black players or teams with black players. The linear path of media to consumer was what Scharkow and Vogelgesang have criticized as a limitation of such research, in that it assumed agenda setting as a process that occurs with audiences (2011). In today's world of search, consumers of news and content now communicate their interest into the search engines and find the content they want. They are not restricted or subjected to a finite amount of a certain type of publication in a certain region, as they may have been before and during Carroll's research.

Although the nature of positive or negative sentiment in the writer's or communicator's words has, in the past, been a significant focus in much of the research using content analysis, that is not the focus of this research. For the purpose of this paper, due to time limitations and the newness of the research methodology and topic, any interest in the leagues (both positive and negative) is perceived as interest or salience alone. Due to the nature of the content, the long history of the leagues and the cyclical nature of their seasons, other research referenced here is focused on interest (in fishing, for example, in Wilde and Pope). Also noted were specific political issues, which are easier to compare than traditional content analysis that is focused on the publisher and nature of the content, rather than whether the public as a whole is even interested in the topic. Until the past several years, this was impossible to do without traditional polling, hence the reason this research juxtaposes these traditional methodologies of public

opinion surveys and content analysis to form this new method of analyzing actual interest and salience from the reader or information seeker's own words in Google Trends.

Content Analysis Using Google Trends

Although different from a content analysis in exploring present and past trends of salience, the methodology of using Google Trends to identify salience in a given time is similar. Although many other studies have looked into how web search and social phenomena are connected, the main question they seek to answer is whether it can be predictive (Vaughan & Romero-Frias, 2014). Using content analysis, the studies provide a non-biased, cheap and somewhat quick method of measuring and observing what is happening in communication during a certain point in time (Krippendorff, 1980). In the past, the methods of measuring interest and issue salience proved difficult because many actions were unobserved (Scharkow & Vogelgesang, 2011). This is why media use and search volumes using content analysis can be an effective way to contrast and compare traditional online polls, which require observed or triggered responses (Scharkow & Vogelgesang, 2011). Scharkow and Vogelgesang explain:

> A possible solution to the small sample problem of field studies employing participant observation lies in automated observation tools that work without human observers. However, these data are less well-suited for agenda-setting research because: (a) interest can only be measured at the granularity of programmes not issues; (b) the users' possibilities for salience-driven information seeking are limited in linear media like TV; and (c) information seeking may not be the only relevant motive for programme choice (2011, p. 106).

For example, content analysis can also be used in conjunction with other research methods to contrast and compare correlations of statistics, polls and traditional media consumption methods to make hypotheses on trends or on a current state (Krippendorff, 1980). This can be a powerful tool for research, but as Krippendorff points out, it also leaves room for systematic interpretation:

> For the content analyst, the systematic reading of a body of texts narrows the range of possible inferences concerning unobserved facts, intentions, mental states, effects, prejudices, planned actions, and antecedent or consequent conditions. Content analysts infer answers to particular research questions from their texts. Their inferences are merely more systematic, explicitly informed, and verifiable—ideally—than what ordinary readers do with texts. (1980, p. 30)

By using content analysis theories and Google Trends, one can collect and gather search patterns on a weekly basis back to 2004 in the country and province on a variety of search terms (Mellon, 2014). There are thousands of searches across Canada in relation to the two football leagues, so there is more than enough search data to quickly see and analyze trends. According to Melon:

> Google Trends provides information on search trends measured weekly. Second, there are many countries where surveys are only conducted sporadically, [but] Google search data are available anywhere in the world where sufficient numbers of people use its search engine. The Google Trends website allows researchers to download data for almost all countries at no cost and to download time series of any search term's

> popularity over time provided enough people have searched for it. For these reasons, Google Trends is an attractive data source for social scientists. (2014, p. 280)

Furthermore, researchers have questioned whether the public and commercial entities could use this ability in a concept known as "nowcasting," or predicting the present (Carrière-Swallow & Labbé, 2010). Using the tool, people can download data and view trends related to the keywords that are searched on to see where certain topics might be trending and where public interest lies. This idea of using data for prediction purposes is not a focus of this study, but it is worthy to note the same concepts of using content analysis to measure interest and salience in certain relevant topics is a key and emerging field of research. Carrière-Swallow and Labbé note:

> In 2009, Google began the public release of its users' search queries through a publicly accessible interface. The rapid expansion of the Internet into all aspects of modern life together with Google's dominance in the search engine market give the company a central role in the collection of market intelligence. The publication of user search queries offers researchers the tantalizing possibility to observe the interests of society in real time without carrying out costly surveys. (2010, p. 1)

This Google Trends data provides issue salience for the general public as it is related to keywords (Mellon, 2014). The difficulty at times, however, is finding the search terms to measure in order to identify levels of interest or engagement in the general public (Mellon, 2014). Some studies stop short, claiming that the correlation of their data to news coverage is evidence that information seeking theory and using Google Trends is a validation of predicting the present or the effectiveness of media or current events at influencing interest or salience

(Fazeli Dehkordy, Carlos, Hall, & Dalton, 2013). For example, Fazeli Dehkordy et al. noted when legislative action occurred on one particular topic, and then tracked search interest during those times, and correlated the peaks to the legislative decisions:

> Newsworthy events and legislative actions appear to correlate well with peaks in search volume of 'dense breast.' Geographic regions with the highest search volumes have passed, denied, or are currently considering the dense breast legislation. [The] study demonstrated that any legislative action and respective news coverage correlate with increase in information seeking for 'dense breast' on Google, suggesting that Google Trends has the potential to serve as a data source for policy-relevant research (2013, p. 1172).

This is very close to the research being reported in this paper; however, instead of noting the dates of specific media incidents, we look at comparing two queries over time using Google Trends. This quantitative content analysis of keyword queries using Google Trends is not focused on specific dates or incidents that may prompt the public to search more or less. I look at the trend over time, comparing the two leagues, during similar time periods of high interest. For example, the leagues' seasons, as noted previously, begin at different times. The CFL runs from June to November, whereas the NFL runs from September to February (Canada.NFL.com, 2015). Due to these seasonal patterns of when games are taking place and when repeatable events are happening, interest could hypothetically peak during each league's playoff or championship games and the opening month of the season.

The research being done still relies on the basic principle of content analysis, which is an

evaluation of which keywords or phrases are being used by the public and how often. Searching or seeking out information about an idea or product using certain keywords is a level of interaction being communicated through finger-taps on a screen or keyboard in hopes of receiving the information instantaneously. Hillis et al. describe action as a cybernetic feedback loop to give users what they want (Hillis et al., 2012). These millions and millions of micro-communications are logged and accessible in a database, meaning it was almost as if Krippendorff spoke in the present when he said:

> Content analysis has many commercial uses. For example, word-association databases (which collect huge numbers of pairs of words that consumers associate in their minds, as determined through word-association experiments) can serve as the context within which the advertising research can infer chains of associations for products, services or brand names (1980, p. 34).

Now, through advances in search technology and the "Database of Intentions," as Battelle said, one can actually use Google Trends to see which words the public is using to find associated content. The work today involves identifying those terms the public is using. There is a bit of trial and error to this process using Google Trends on its own. The researcher must have in mind terms or queries they know signify an issue or topic that the user would use. Google Trends does, however, offer up variations of queries using a word or series of words. It also lets searchers filter by topic, such as sports in this study.

Content Analysis as It Relates to Information Seeking

Information seeking and measuring search volumes is an excellent measure of issue

salience and interest, as there is limited bias or the context of timeliness as it relates to news and information that is in the public's mind compared with when a traditional poll is conducted (Wilde & Pope, 2013). Thus, search volume measurement and content analysis on an ongoing basis can be a useful means to measure interest and even predict future interest (Wilde & Pope, 2013). Public opinion surveys can also be difficult, as there are uncontrollable variables of what is in the news currently and when the poll is conducted that can affect responses (Scharkow & Vogelgesang, 2011). Using Google Trends, a researcher can quickly see the month and week when search interest rises to correlate to specific high-profile events (Trends Help, 2015). This makes the content analysis an effective complementary tool using both information-seeking and agenda-setting theories to formulate the necessary base for future research and solidification of issue salience.

The aggregate of Internet searches over time produces a strong correlation of information seeking leading to changes and future predictions of the public's agenda (Scharkow & Vogelgesang, 2011). As Scharkow and Vogelgesang note:

> From a measurement point of view, search queries have many desirable properties: Compared to survey questions, there is no interviewer bias or social desirability involved, the measurement is completely unobtrusive and happens in the field. Moreover, for many users, there is virtually no effort involved in using search engines, compared to buying a book or searching a paper encyclopedia. Of course, the demographic of Internet users is still different from the general population, so that we cannot take the complete public agenda, as measured by search queries, at face value. (2011, p. 105)

It is noteworthy that since Scharkow and Vogelgesang highlighted these demographic issues such as certain segments not using the Internet or Google, Internet usage has grown, with Google still holding 75 per cent of the search share amongst competing companies (USAToday.com, 2015). Internet use in Canada has now reached 90 per cent in 2013 (Internetworldstats.com, 2015). Many political scientists have used Internet search as a means to measure issue salience (Mellon, 2014). The ambiguity of search, and the context in which searchers perform and use certain keywords, still forces researchers to understand and validate search data (Mellon, 2013). Unusual news items or circumstances can cause keyword spikes unrelated to interest or a specific event, which heightens the importance of validated keyword data, trended over time (Ginsberg et al., 2008). There is now mounting support for the link between Internet search and human actions (Wilde & Pope, 2013).

It is suggested, using the theories of agenda setting, that the increase in coverage of the NFL and professional football as a whole is causing increased interest in the game (Bibby, 2013). Agenda setting has commonly been used to identify how the media, including newspaper, TV and radio affect the public agenda, which is then measured in traditional public opinion surveys (McComb, 2004). McComb notes:

> The agenda-setting role of the news media is their influence on the salience of an issue, an influence on whether any significant number of people regard it as worthwhile to hold an opinion about that issue. While many issues compete for public attention, only a few are successful in doing so, and the news media exert significant influence on our perceptions of what are the most important issues of the day. This is not a deliberate,

> premeditated influence—as in the expression 'to have an agenda'—but rather an inadvertent influence resulting from the necessity of the news media to select and highlight a few topics in their reports as the most salient news of the moment (2004, p. 2).

This suggestion and assumption built on the foundations of agenda setting needs further research to get at the root cause of the increased interest, though it is not argued that it does, in fact, play a significant role. One could argue that other factors—such as the success of certain teams, the increased quality of media coverage and increased access to digital content—are driving Canadians to become more interested in American content at a more rapid rate than Canadian content. Due to the limitations of time, the research in this study can add more context and will focus on Google Trends as a novel way to measure salience and the increases in interest in professional football, hopefully encouraging more research on reasons behind the increase and whether interest continues to rise in the future.

There is also a "dearth of research" on the effects that social media has on increasing interest and reach of football in the United States (Jensen, Ervin, & Dittmore, 2014). One cannot rule out the effect of this new medium as coaches, teams and even players use social media more and more to communicate their messages (Jensen et al., 2014). Likewise, most content analysis until recently had focused on the quantity of stories relating to a subject matter and its prominence, then coupling the technique with traditional questionnaires or polling (McCombs, 2005). By gathering data related to information-seeking behaviour on the Internet using Google Trends, one could reasonably identify trends in issue salience and interest in subjects (Scharkow & Vogelgesang, 2011).

Many studies have included varying degrees of content analysis using categorizations of keywords or topics used in information seeking to cover a particular area or topic of interest. This categorization leads to decisions of taxonomy and potential subjectivity. Such an approach requires the researcher to decide, based on the context of the text, how to classify types of content accordingly (Hider & Pymm, 2008).

This differs from the research presented in this paper, as it focuses on specific keyword usage and variations of the keywords. Although this helps focus the research on more definitive quantitative results, there are opportunities to question whether the data could include keywords that are not directly related to the keywords being researched. However, Google Trends says the tool detects misspellings and groups content into categories to avoid such mistakes (Trends Help, 2015). This in essence is doing a similar task to what Hider and Pymm were doing manually and allows for comparison of past searches year over year and study versus study.

Salience Using Non-biased Methods of Quantitative Measures

One downside to this research is the fact that the validity of issue salience measurements remains challenging (Scharkow & Vogelgesang, 2011). Early agenda-setting theories took aim at the accessibility of media in some countries as it relates to certain stories (McCombs, 2005). Today, access to the Internet is rarely limited. More than 80 per cent of Canadians had access to the Internet in 2010 (Canadian Internet Use Survey, 2011). However, the quality and quantity of content in certain subjects—NFL versus CFL, for example—could be in question. Due to length and time constraints, this research will not go into this area much, as it focuses on the content analysis and information seeking theories that identify how salient one league is versus the other

over time in Canada.

The challenge of this research stems from search data perceived as ambiguous compared with traditional surveys, since a web searcher is looking to perform a task (Mellon, 2013). For example, one could search for "CFL joke league" or "NFL criminals" and be included in such research as "interest." These searches don't necessarily indicate popularity, but do indicate some sort of interaction. It's also worth noting again that the aggregate of such searches over many months and years, and the nature of the content in terms of this study with leagues running during set periods, can offset this to a degree.

Finally, issues of translation could pose a difficulty. The CFL has a French version of its site and different words related to football for scores, schedules and standings. NFL fans in Canada are forced to use English to consume or find information, thus possibly skewing the results higher for NFL because Quebec has a large population of football fans. The changes mentioned previously in agenda setting theories align with Bibby's assumption that increased media coverage and access in Canada could make the NFL more viewable in the country. This is assumed to lead to more information seeking from fans, new and old, thus increasing the interest year after year.

A review of the literature shows strong correlations in methodology and the ability to use Google Trends to pull search data and compare keyword trends that contribute to issue salience in a specific topic. There are great possibilities of using search data to identify issue salience, but also drawbacks if the data are not accurately validated (Mellon, 2013). Past research is strong with regard to how media, teams and even fantasy sports affect American sports interest, but

specific research on the Canadian Football League, and the relevance of using the methodology of content analysis and Internet search trends for sports and other commercial interests, remains less explored.

Using traditional public opinion surveys, ratings and other basic metrics, leagues can measure interest and popularity as success measures and for setting advertising partnerships. However, accurate surveys are expensive and require a heavy investment of time. Also, broadcasts of games can be taped and watched by fans later. Measurements using keyword search data can be a quick and low-cost way to gauge issue salience of a particular topic (Martin, Pracheil, DeBoer, Wilde, & Pope, 2012). This low-cost method also quickly identifies patterns of content consumption by fans, and research continues to grow that points to the tie between search volumes of particular topics and people's actions (Wilde & Pope, 2013). This is to say online searching has become embedded in a wider present, pragmatic and future intentional context of everyday life. For example, search alone is the largest digital advertising spend at almost $50 billion and growing nearly 10 per cent per year, with Google having two-thirds of the searches (Quartz, 2015). In 2012, there were 1.2 trillion searches on Google and rising (Internetlivestats.com, 2015). It has become a natural part of life for most of us.

The concepts of information seeking, combined with a content analysis of search trends, can be used to measure interest in a particular subject matter. This theory and tying to content analysis can be used to identify trends and make inferences about the future, especially seasonal events or football seasons such as CFL versus NFL in this research (Scharkow & Vogelgesang, 2011). This theory has been widely used in political contexts and successfully applied to validate

or compare with public opinion polls (Mellon, 2014). This falls in line with a similar methodology and research questions posed in this research, just in a different context.

As noted earlier, much of the analysis in past research using similar methodologies of Google Trends data has been to seek out predictive measures to help explain or forecast (Vaughn, 2014). For example, one study showed how economic data in the present could be predicted using Google Trends (Choi & Varian, 2012).

Also noted previously, another project helped detect disease outbreaks, in particular the flu epidemics occurring globally, emphasizing the advantages of a content analysis of keywords or queries over media mentions (Santillana, Zhang, Althouse, & Ayers, 2014). This study did not guarantee results in predicting flu outbreaks, but it was fairly accurate, and also noted any miscalculations or predictions could be adjusted on a weekly basis instead of waiting for manual analysis after the season had ended (Santillana et al., 2014). Santillana et al.'s study included analysis on the most recent flu season, which the Centers for Disease Control listed as moderately severe (Centers for Disease Control and Prevention, 2015). Flu seasons, like football seasons, are predictable, with predictable peaks in flu seasons from December to January (Centers for Disease Control and Prevention, 2015). A related study of flu outbreaks earlier than 2012 notes the advantages of this newer form of content analysis as well in relation to information seeking as it relates to health information:

> Although news media represent an important adjunct to the public health infrastructure, the information they report pales in comparison to the potential collective intelligence that can be garnered from the public. An estimated 37 to 52 [per cent] of Americans seek

> health-related information on the Internet each year, generally using search engines to find advice on conditions, symptoms, and treatments (Brownstein, Freifeld, & Madoff, 2009, p. 2154).

Much of the similar research either using content analysis and measuring trends in social media or search is focused on these disease outbreaks and the predictive or alert-like qualities such data could produce. In Choi and Varian's research on various other topics besides economics, it was noted:

> One thing that we would like to investigate in future work is whether the Google Trends variables are helpful in predicting 'turning points' in the data. Simple autoregressive models [data that continues on the same trend related to its historic values] do remarkably well in extrapolating smooth trends; however, by their very nature, it is difficult for such models to describe cases where the direction changes. (2012, p. 18)

A group in Nebraska identified the trend in decreased search volumes around fishing worldwide, hinting at the prediction of decreased issue salience in the sport or livelihood (Wilde & Pope, 2013). As early as 2011, one study noted high success in predicting not only the future but also the present month using Google Trends data, again focused on economic data as it relates to private spending and consumer behaviour (Vosen & Schmidt, 2011). Products and online tools like Auction.com heralded the predictive nature of Google Trends and this content analysis to forecast home sales (Bloomberg.com, 2014).

Aside from predictive focuses of research, just being able to identify trends in a particular area has proven useful using content analysis and search trends in early research—for example,

to inform investors or tech companies of a particular rise in software companies or names (Rech, 2007). In the political arena, issue salience was researched using Google Trends to replace traditional polling methods (Mellon, 2014). Google itself has used the same data to map disease outbreaks and establish warning messaging for organizations and countries (Hillis et al., 2012, p. 43). Content analysis has been done looking at social media in American football teams and the incidences that contribute to popularity (Jensen et al., 2014). Also noted, search trends were analyzed to predict and measure one's academic fame (Vaughan & Romero-Frias, 2014). Prior to the creation of Google Trends, some studies examined search logs, looking at differences from country to country in search trends—for example, whether Americans searched for basketball more than residents of other countries (Silverstein, Marais, Henzinger, & Moricz, 1999). At that point, analyzing queries was costly, as Silverstein et al. note: "Determining query duplication is expensive because it requires storing each unique query. For our analysis, with over a hundred million unique queries, it is infeasible to store each query in memory" (1999). Scharkow and Vogelgesang, in relation to the extensive work needed to analyze log files in the past, point out that today's tools and storage abilities make the actual mining easier and enable research to focus on analysis versus theories of how to synthesize log file information:

> Fortunately, the world's largest search engine, Google, has recently begun to make aggregate log results available to the public. Google started their web service Insights for Search (http://google.com/insights/search) in August 2008 as a follow-up to their earlier Google Trends site. GIFS provides public access to Google's logged search queries and allows for filtering by search term, time frame, and region. Unlike older services, users cannot only see graphical presentations of these data, but also download actual data tables

of the search volume for a particular query. (Scharkow & Vogelgesang, 2011, p. 107)

Although it is common for media and Google itself to publish search trends by region and country related to sporting leagues or events, academic studies measuring league interest over time are limited. A study focused on trends in search and interest in football in Canada could not be found. Proceeding from this background, this research will use similar methodologies as noted earlier, on a different topic, "football in Canada," to begin to use aggregate data to measure salience on CFL versus NFL keywords both nationally and regionally.

Using the similar methodologies to the research noted above, and Google search interest analysis through Google Trends, the research outlined in this paper considers current trends in Canada of the salience of both professional football leagues and aims to spur future research to explain the past and present trend and why it might be occurring. Due to time and length constraints, the research in this paper will not explain why the trends in interest are or aren't occurring, but will provide a solid base for future research. When paired and analyzed with other research in the area using traditional polling methods, one could begin to examine further why or why not interest or salience in each of the leagues is trending and the underlying reasons or implications behind it.

Possible Issues with the Gathering and Methods

Issues related to language may exist in large portions of Quebec because the population is French-speaking. However, the number of searches for the French variations are not enough to affect the general trend across the country. The data are limited to Google, which accounts for a

nearly 70 per cent majority of Internet searches ("Google Fails to Gain Search Market Share, Bing Steals From Yahoo", 2014). However, audiences do vary and the search patterns may be different from search engine to search engine. Still, the aggregate of Google search data over the extended period of time on millions of well-known search queries continues to be a respected source for research related to monitoring keyword frequency and thus interest (Ginsberg et al., 2008).

The normalization of search data allows equal comparisons by geographic location, as the data show the percentage of all keyword searches at a given period of time, thus limiting the argument of simply higher-population locations or more computers as the reason for the increase in certain searches (Trends Help, 2015). The main issue related to data collection is the ambiguity of search. The NFL also has more teams, a more complex draft and standings, as well as the previously stated robust fantasy sports leagues and plethora of media options. This means there is more content to search. Every CFL game is also shown on one channel, whereas NFL games are carried on multiple broadcasters (TSN.ca, 2014). It is worth noting, and for future research to consider, that the reliance on digital content and information-seeking behaviour analyzed in this study below, may be more necessary and prevalent for NFL fans than CFL fans in most of Canada. On the other side, regional data in Saskatchewan do not correlate with this theory. More research will need to be done to examine these trends both nationally and provincially.

Methodology

This paper uses a content analysis with the help of a tool called Google Trends. Google Trends uses normalized trends to analyze keyword data and display it graphically over time to compare and contrast keywords worldwide and within a region. For example, in the graphs in Appendix A, time is represented on the horizontal axis and relative search frequency on the vertical axis. Relative search frequency is the search term volume, in this case NFL or CFL, and the normalized relative to the total search volume on Google during the time period. The search term frequency at each time point is presented as a percentage of the highest volume of searches, "NFL" compared with "CFL," during the period of interest, rather than absolute search volume. The peak volume within the period of interest represents 100 per cent, whereas the relative frequency at other time points is displayed as a proportion of this. If the total volume of searches for the term does not reach a required threshold, estimated at a minimum of 1,000 searches over the relevant period or geographic region of interest, Google Trends will report the search volume index as zero. This allows one to capture the temporal variations of information seeking over time, as reflected in search volume.

Google Trends also allows display of hotspots, the geographic regions with the highest search volumes, both as a heat map and a ranked list. Data are, however, limited in specific provinces due to small sample sizes as noted above. These provinces did not have enough searches during the time frame to reach the threshold of 1,000 searches during the time period. This study will focus on analyzing data in CFL hotbeds in Western Canada such as Manitoba, Saskatchewan and Alberta, as well as in high-population provinces such as Ontario and British

Columbia. Due to constraints of time and scope, this study will not go into the detail of searches in specific cities, though future research would be encouraged in this area.

Using Google Trends, an aggregator of searches on keywords within Google since 2004, one is allowed to discover the frequency of common words and phrases used in Google Search queries (Trends Help, 2015). The numbers for the keywords and phrases entered are scaled to the number of searches in Google over time, and duplicate searches by one person during a short period of time are excluded (Trends Help, 2015). Each observation is divided by the maximum value in a set of comparisons and is then multiplied by 100 (Wilde & Pope, 2013). For this study, the population was restricted to English searches in Canada and its provinces. Data were collected for the following queries:

- "NFL" vs. "CFL" in Canada and Saskatchewan as well as other provinces and derivatives of specific phrases that include "[league] + "draft," "schedule," "scores" and "standings"
- "Super Bowl" vs. "Grey Cup" in Canada as well as other provinces
- NFL star player "Tom Brady" vs. CFL star player "Ricky Ray" in Canada as well as others in other provinces
- NFL vs. CFL teams in Canada (Seattle Seahawks, Calgary Stampeders, Green Bay Packers, Saskatchewan Roughriders and Minnesota Vikings)

These searches produced data sets that were then charted and trended over the 10 years, compared with each other—including data from Saskatchewan, used as an abnormal province in their love for CFL football—and B.C. and Ontario as both are NFL hotbeds with the nation's largest populations. Top provinces for NFL and or CFL interest were noted to give more context

and validate the nationwide results, with the main hotbed of CFL football in the western Prairie provinces.

By contrasting and comparing the normalized Google Trends data between the selected keywords and phrases, we noted geographic contexts and differences. Given the seasonality and peaks during specific events such as championship games and drafts, the trends over time were important to note—seasonal highs and lows as well as the average interest. For example, most searches occur during the two seasons, which run from September to February for the NFL and June to November for the CFL (Canada.NFL.com, 2015). Peaks during the season occur rather consistently based on the queries. This will be discussed further in the results below.

By contrasting these trends of normalized search data, we can began to clarify the issue salience and to what extent NFL and CFL interest have potentially increased in Canada, and to some extent predict its growth (or lack thereof) moving forward. As the literature states, concerning the importance of validating normalized Google Trends data, the search data are cross-referenced on unpublished data concerning a poll in 2013 showing that interest in professional football has in fact increased by five to six per cent, but somewhat equally for NFL and CFL, and that the CFL is slightly more popular than the NFL (Bibby, 2013).

When it comes to looking at the specific data points, it's important to find comparative sets. Looking at the peaks in data between the two leagues allows us to compare the differences during similar times, because the data are so seasonal and predictable. As mentioned above, the NFL season starts in September and ends in February, whereas the CFL season starts in June and ends in November. This is why it's important to use peaks when comparing, because the high

points for one league are not the same for the other. Averages are also misleading, as they smooth out the results and underaccentuate interest in leagues that are generally focused during short periods of time. Peak interest and the differences at peak times are the focus of this research due to time constraints and research limitations.

In this specific research, due to these time constraints and the research goals, we focus on the peak months for each league and compare high search interest in each year. Because the leagues are running at different times, comparing weeks or months during the same time period to track issue salience would not be appropriate. By comparing peaks and lows during the seasons, we can compare the interest relative to each league year over year for nearly a decade with the data available using Google Trends.

Why the Method Was Chosen

The method was chosen because of its ability to identify objectively the issue salience and interest in specific subjects during certain points in time dating back more than eight years. The tool can quickly gather and export large amounts of data critical to producing the needed trends and matching of data validity, and the preliminary research shows significant trends pointing to some kinds of phenomena that cannot be easily answered. The methods and variations of using Google Trends continue to evolve due to its efficiency and low cost, as well as the interesting outcomes of the research. Comparing annual search data and the trends related to certain queries has an impact on economic, health, social and communication issues. The analysis can now be done in weeks instead of months.

Also, a national traditional polling survey was conducted in 2005 and 2013, which matches up well to trend the increased interest in the CFL and NFL (Bibby, 2013). By comparing related keywords and showing somewhat different sets of data, we could hypothesize why the increase is occurring in the NFL at a more rapid pace, and why there are differences between traditional polling and the methods of investigating search trends using content analysis and Google Trends. One of the strongest arguments for content analysis in this case will be the ability to draw strong inferences from a large data set that would not be readily available otherwise (Krippendorff, 1980). This could lead to future research on why the increase is happening, and what this means for Canada and the CFL as a whole. Due to time constraints and the nature of the research, those research questions will not be addressed in this paper.

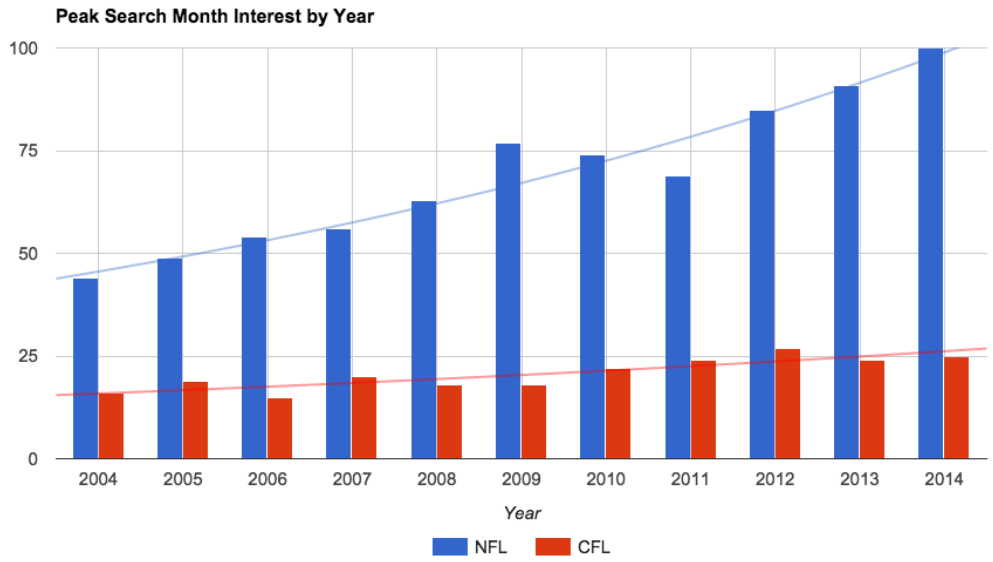
Outcomes of the Research

CFL vs. NFL

The NFL in Canada has almost always peaked for search interest in September of each season, which is when the season begins, with the exception of December 2007 when September and December were almost equal. Alternatively, CFL search interest peaks during the playoffs and Grey Cup, the league championship and end of the season, in November (see Appendix A). This difference between the two leagues is important for future researchers to note when trying to understand the motivations of Canadians interested in information about both leagues. In 2004, the NFL peaked at approximately twice the amount of search interest as the CFL. Each year, search interest in NFL-related terms shows gains of five to 10 per cent, with the only exception being in 2010. Contrary to that is search interest in the CFL, which has increased

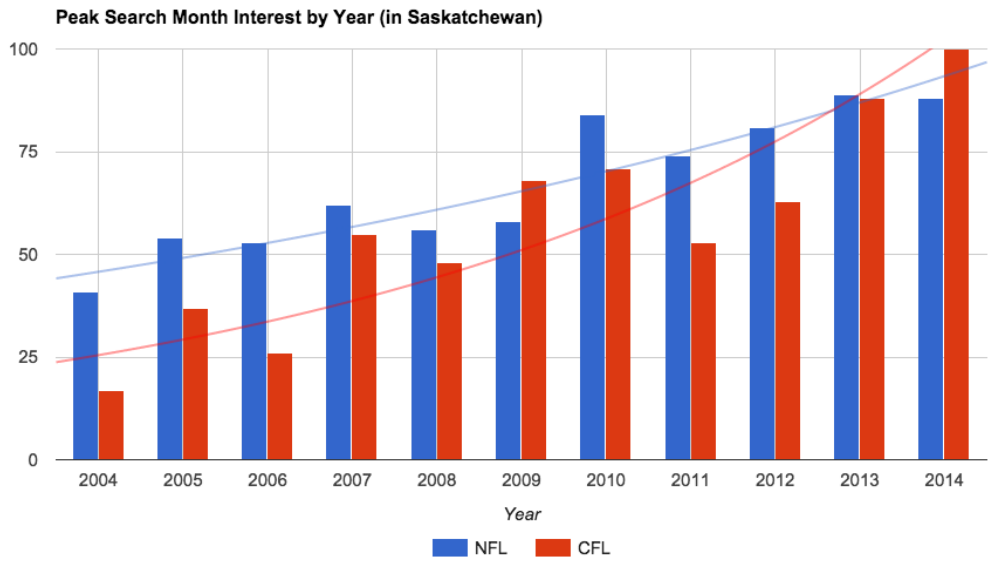
nearly 57 per cent in a pattern of slow but steady growth, with the exception of a spike in 2007 of nearly 20 per cent.

The NFL, in comparison, has increased nearly 220 per cent during its peak times, usually in September. It is clear, as a whole, that search interest in the NFL has been higher than the year prior since 2004, and that it has increased more rapidly than interest in the CFL.



As shown in the graph above, NFL search interest in that league's peak month of September has risen each year, whereas the CFL search interest in its peak month, November, is increasing, but more slowly in comparison. As noted in Appendix A, search interest during the highest times in CFL interest, in November, is still higher than the lowest times in the NFL, usually in July. All of these numbers look quite similar across the country in each province, until we look at Saskatchewan. In Alberta and Manitoba interest is stronger, but the pattern maps closely to the national peaks. Ontario has by far the biggest gap in search interest between the

two leagues. Saskatchewan, as noted in Appendix B, is the anomaly. Peak interest has traditionally been the closest between NFL and CFL in this province. Last year, CFL search interest overtook NFL for the first time ever. Saskatchewan also does not match the trend identified earlier, wherein it appears most CFL fans are more interested in the end of the season during playoffs. However, search interest still peaks at the end of the season.

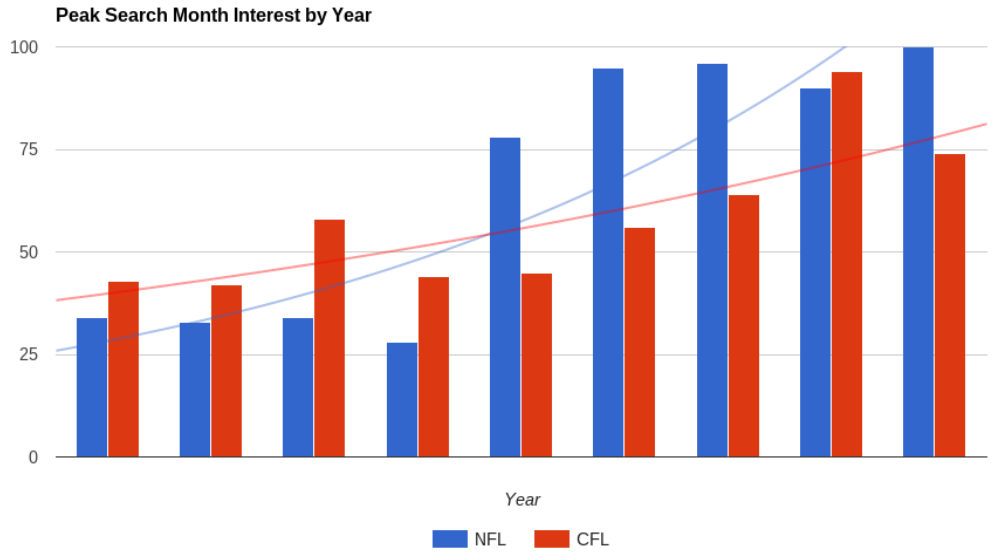


Other search variations all point to a wide gap in favour of NFL-related queries. Searches for NFL Draft, Scores, Schedule and Standings are all more searched on in Canada than their counterparts in the CFL, as noted in Appendix A - Other Queries. Interest is increasing in all queries that were gathered in this research, but the gap in increases between NFL- and CFL-related queries is widening, especially in the last three or four years since 2009 or 2010. It is also worth noting that 2014 was a banner year for CFL after relatively steady search interest from about 2008 through 2013. For example, in 2014 search interest in CFL standings and scores increased nearly 50 per cent year over year. This still was not as large as the increase the NFL

had as a whole during the same time period, although the gap between CFL- and NFL-related searches narrowed on very specific queries involving standings and scores, for example, in 2014. NFL-related queries were 300 to 500 per cent higher than CFL-related queries across Canada.

Grey Cup vs. Super Bowl

The two leagues' championship games are known as the Grey Cup in the CFL and the Super Bowl in the NFL. Search interest in the big game for the CFL outpaced the NFL until 2010. From that point forward, the NFL's Super Bowl has been equally or more searched on than the Grey Cup. The Super Bowl occurs in February, whereas the Grey Cup occurs in November. From 2004 to 2009, the Grey Cup generally garnered more search interest than the Super Bowl by as much as 20 per cent or more in some cases. Grey Cup search interest experienced steady growth while Super Bowl interest remained steady until 2010. It appears this is when a distinct shift in both NFL- and CFL-related search queries, noted above, and championship game queries took place. NFL-related searches for "Super Bowl" tripled and then went up another 25 per cent the year after and have maintained that mark for the last four years. The Grey Cup experienced steady growth but did see a maximum number of searches in 2013 when the popular Saskatchewan Roughriders, noted above, hosted, made and eventually won the game.



This shift for Canadians as a whole has remained constant in the following years. As noted in Appendix C and above, from 2004 until 2010, the Grey Cup was searched on more than the Super Bowl every year. Both games seem to be increasing in search interest at a more rapid rate than general queries involving the leagues "NFL" or "CFL." It wasn't until the turning point in 2010 that queries about schedule and standings for the two leagues increased dramatically for the NFL compared with the relative increases in the CFL.

Also worth noting are some of the regional differences in search trends. For example, according to Appendix C - Other Provinces, searches for the Grey Cup in Alberta and Saskatchewan and even British Columbia are not as high as Super Bowl searches, but are very close. This in contrast to Ontario, where, outside of small blips such as when the 2012 Grey Cup was hosted in Ontario, Super Bowl searches rose in 2010 and have stayed well above Grey Cup searches. It appears when a province has a city hosting the game, it causes an increase of searches in that province, but it does not necessarily affect the total searches nationwide.

Popular Player Comparisons

Since data were captured and could be analyzed, NFL marquee player names have traditionally outpaced CFL counterparts for the nation as a whole, according to Appendix D. Rarely, certain players regionally have equal or more search interest from time to time, according to Appendix D - Other Players in Other Provinces. Also interesting to note are spikes from when popular NFL players have come to the CFL. Ricky Williams, Chad Johnson and, more recently, Michael Sam all created a buzz when they entered the league. Historically, this passes, and NFL star players garner more search interest after an initial spike (see Appendix D). This is not a main focus of the research, and there aren't enough data to analyze in full, but it is worth noting that outside regional differences, popular NFL players are, as a whole, more searched on than CFL players.

Popular Team Comparisons

Again, although not a main focus of the research due to time constraints, it is worth noting specific teams and how search interest in popular CFL teams has compared with interest in popular NFL teams over the time period. Of late, according to Appendix F, the Seattle Seahawks of the NFL have performed well and garnered the highest search volume of any professional football team in Canada. The NFL's Green Bay Packers previously achieved similar results, and the Saskatchewan Roughriders of the CFL have had moments of high search interest when making or hosting the championship game. (The clear difference in search volume in Saskatchewan has been noted.) Despite this not being a main focus of the research, it is worth noting how individual team followings and successes may be affecting search trends over time in

Canada. Due to time constraints and the nature of the research, we are not able to go into this further, but it would be worthy for future research to look into the trends.

Conclusion

The NFL remains high and continues to grow in search interest across Canada. In 2004, search interest in NFL-related keywords was more than 2.5 times that of CFL-related keywords. Now, NFL interest during the peak month is more than four times that of CFL-related keywords. Canadians seem to search more for NFL-related keywords at the beginning of that league's season and search more for the CFL-related season at the end of its season. It is also worth noting that, despite a spike in NFL search interest in September, the CFL doesn't see a drop in search interest during that time, though it is usually much higher in November. It is generally considered in other research that the amount of NFL content, including broadcasts that far outnumber those for the CFL, contribute to these outcomes. It is noted that the NFL has 32 teams while the CFL has traditionally only had eight or nine during this time frame, meaning there is less content to actually search for, whether it's number of games, players or news media sites.

The data in Saskatchewan are noteworthy. Local and league CFL content cannot match that of the entire NFL in amount and variety. This research could be challenged by some who may assume that, because the NFL simply has more content out there, more people would search for it. This theory is bucked in Saskatchewan, where the league is searched on equally and even more of late and when the local team is good. These are league-related searches that include "CFL," not simply local team content. Due to time constraints and the focus of the research, we did not look at specific provincial and city trends between the two leagues for a variety of

queries. Data are also limited in specific provinces due to small sample sizes, so much of this study has focused on high-population areas or on regions where football is the most popular.

The reasons for such a rapid rate of increased interest and salience of the NFL for Canada as a whole could be attributable to several factors, including northern teams doing well of late, increased media coverage of the NFL and more digital content accessible to Canadians than ever before. More research will need to be done to examine why this trend is occurring and what it means to Canada. But it is also worth noting that NFL-related searches are not the only ones that have risen. CFL-related search interest has risen even in Ontario and British Columbia, albeit not at the rate that NFL-related queries have. It also appears—and we could infer, given the drop in Google search interest in NFL-related terms around playoffs and the Super Bowl—that searches turn from "NFL" to the championship game "Super Bowl." This is why it's worth it to look at multiple words and phrases to get a fuller picture.

More research will need to be done to examine what causes the interest or salience to rise when certain teams host or do particularly well in either league, and whether the recent increases related to the Seattle Seahawks are at all responsible for the gains from 2010 on. Only time and further analysis in the coming years would be able to justify such an inference. This might explain the increases out west and in the prairies for NFL, but does not explain Ontario. Many other factors are in play, and require more research. Due to the time constraints and nature of this research, we will not be able to examine this further; however, it is hoped that other studies will follow up on these trends and examine whether they continue, why they may be occurring and what it means—not only to football in Canada, but also to Canada as a whole. It is also hoped

that further research using search queries to predict salience and trending topics will be implemented in other industries.

Finally, it is hoped that this research will act as a wake-up call to Canadians. Something is happening in Canada when it comes to interest in and around American sports and entertainment. As interest increases in American-related leagues and sports, researchers and studies could focus on what this is doing to Canadian leagues and sport, and what these trends mean when it comes to the future of sports in our country, both in terms of amateur participation and of the sports entertainment industry as a whole. It will also be interesting to see when the trend of increased search interest in professional football ends or plateaus.

References

Araz, O. M., Bentley, D., & Muelleman, R. L. (2014). Using Google Flu Trends data in forecasting influenza-like-illness related ED visits in Omaha, Nebraska. *American Journal of Emergency Medicine, 32*(9), 1016–1023. doi:10.1016/j.ajem.2014.05.052

About NFL Canada. (2015). *Canada.NFL.com*. Retrieved May 23, 2014, from http://canada.nfl.com/about

Battelle, J. (2005). *The search: How Google and its rivals rewrote the rules of business and transformed our culture*. New York: Portfolio.

Bibby survey shows football thriving in Canada. (2013, November 22). *University of Lethbridge UNews*. Retrieved April 25, 2015, from www.uleth.ca/unews/article/bibby-survey-shows-football-thriving-canada

Bills and Rogers Communications agree to terminate Toronto series. (2014, December 3). Retrieved April 16, 2015, from http://www.buffalobills.com/news/article-1/Bills-and-Rogers-Communications- agree-to-terminate-Toronto-series/99855c5e-4373-4560-845a-0a31e6a9a18f

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. (2009). Digital disease detection—Harnessing the Web for public health surveillance. *New England Journal of Medicine*, *360*(21), 2153–2157. doi:10.1056/NEJMp0900702

Buffalo Bills won't play in Toronto next season. (2014, March 5). CBC Sports - Football - NFL. Retrieved June 7, 2014, from http://www.cbc.ca/sports/football/nfl/buffalo-bills-won-t-play-in-toronto-next-season-1.2560486

Carroll, B. (2006). From fraternity to fracture: Black press coverage of and involvement in Negro League baseball in the 1920s. *American Journalism, 23*(2), 69–95.

Carrière-Swallow, Y., & Labbé, F. (2010). Nowcasting with Google Trends in an emerging market. *Journal of Forecasting*. Retrieved April 29, 2015, from http://www.bcentral.cl/estudios/documentos-trabajo/pdf/dtbc588.pdf

Canada Internet stats, travel stats and telecommunications. (2013). *Internet world stats usage and populations*. Retrieved April 25, 2015, from http://www.internetworldstats.com/am/ca.htm

Canadian Internet Use Survey. (2011, May 25). *The Daily*. Retrieved June 7, 2014, from http://www.statcan.gc.ca/daily-quotidien/110525/dq110525b-eng.htm

Centers for Disease Control and Prevention. (n.d.) The 2012–2013 flu season. Retrieved April 25, 2015, from http://www.cdc.gov/flu/pastseasons/1213season.htm

CFL and pro football on a roll in year of 101st Grey Cup. (2013, November 21). *Research of Reginald Bibby*. Retrieved May 8, 2014, from http://www.reginaldbibby.com/images/PRO_SPORTS_Release_1_Thurs_Nov_21_2013.pdf

CFL, TSN and RDS reach new multi-year broadcast agreement. (n.d.). *RSS*. Retrieved May 23, 2014, from http://www.tsn.ca/cfl/story/?id=418717

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record, 88*(Suppl. s1), 2–9.

Curtice, J., & Sparrow, N. (1997). How accurate are traditional quota opinion polls? *Journal of the Market Research Society, 39*(3), 433–448. Retrieved from http://login.ezproxy.library.ualberta.ca/login?url=http://search.proquest.com/docview/214812845?accountid=14474

Fazeli Dehkordy, S., Carlos, R. C., Hall, K. S., & Dalton, V. K. (2014). Novel data

sources for women's health research: Mapping breast screening online information seeking through Google Trends. *Academic Radiology*,*21*(9), 1172–1176. doi:10.1016/j.acra.2014.05.005

Financial Tables. Google, Inc. Retrieved May 2, 2012.

Foster, D. (2009). Chasing the public: The CBC and the debate over factual entertainment on Canadian airwaves. *Canadian Journal of Communication, 34*(1), 61–77.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2008). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012–1014.

Google. (n.d.). Our history in depth. Retrieved from http://www.google.com/about/company/history/#2006

Google fails to gain search market share, Bing steals from Yahoo. (2013, November 14). *Search Engine Watch*. Retrieved June 7, 2014, from http://searchenginewatch.com/article/2307115/Google-Fails-to-Gain-Search-Market-Share-Bing-Steals-From-Yahoo

Google's big data to help Auction.com predict homebuying trends. (2014, October 30). *Mortgage Servicing News*.

Google has run away with the web search market and almost no one is chasing. (2014, July 25). *Quartz*. Retrieved April 26, 2015, from http://qz.com/239332/google-has-run-away-with-the-web-search-market-and-almost-no-one-is-chasing/

Google search slipping in U.S. market share. (2015, February 3). *USA Today.* Retrieved April 25, 2015, from http://www.usatoday.com/story/tech/2015/02/03/google-search/22787853/

Google Search Statistics. (n.d.) *Internet live stats*. Retrieved April 25, 2015, from

http://www.internetlivestats.com/google-search-statistics

Habel, M., Liddon, N., & Stryker, J. (2009). The HPV vaccine: A content analysis of online news stories. *Journal of Women's Health, 18*(3)*,* 401–407.

Hider, P., & Pymm, B. (2008). Empirical research methods reported in high-profile LIS journal literature. *Library & Information Science Research, 30*(2), 108–114. Retrieved April 16, 2015, from Library & Information Science Research.

Hillis, K., Petit, M., & Jarrett, K. (2012). *Google and the culture of search*. New York: Routledge.

Howell, C. D. (2001). *Blood, sweat and cheers: Sport and the making of modern Canada*. Toronto: University of Toronto Press.

Jensen, J. A., Ervin, S. M., & Dittmore, S. W. (2014). Exploring the factors affecting popularity in social media: A case study of football bowl subdivision head coaches. *International Journal of Sport Communication, 7*(2), 261–278.

Kadli, J. H., & Hanchinal, V. B. (2015). Information seeking behaviour of law students in the changing digital environment. *DESIDOC Journal of Library & Information Technology, 35*(1), 61–68.

Katz, E., Blumler, J. G., & Gurevitch, M. (1973). Uses and gratifications research. *Public Opinion Quarterly*, *37*(4), 509.

Koerber, D. (2009). Constructing the sports community: Canadian sports columnists, identity, and the business of sport in the 1940s. *Sport History Review*, *40*, 126–142.

Linda, K. (2013). We shouldn't have to smoke and hide: The legalize hashtag as a platform for collective identity and collective action framing. Retrieved January 1, 2015, from Education and Research Archive.

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.

Martin, D.R., Pracheil, B. M., DeBoer, J. A., Wilde G. R., & Pope, K. L. (2012) A look at today's angler: How technology reveals patterns in the Internet age. *Fisheries* 37, 000–000.

Marra, R., Moore, J., & Klimczak, A. (2004). Content analysis of online discussion forums: A comparative analysis of protocols. *Educational Technology Research and Development, 52*(2), 23–40.

McCombs, M. (2004). *Setting the agenda: The mass media and public opinion.* Cambridge: Polity.

McCombs, M. (2005). A look at agenda-setting: Past, present and future. *Journalism Studies, 6*(4), 543–557.

Mellon, J. (2013). Where and when can we use Google Trends to measure issue salience? *Political Science & Politics, 46*(02), 280–290.

Mellon, J. (2014). Internet search data and issue salience: The properties of Google Trends as a measure of issue salience. *Journal of Elections, Public Opinion & Parties*, *24*(1), 45–72.

NFL, Bell Media extend and expand broadcast agreement. (n.d.). *RSS*. Retrieved May 24, 2014, from http://www.tsn.ca/nfl/story/?id=438593

Oltra, C., Delicado, A., Prades, A., Pereira, S. & Schmidt, L. (2014) The Holy Grail of

energy? A content and thematic analysis of the presentation of nuclear fusion on the Internet. *Journal of Science Communication, 13*(4), 1–18.

Percheski, C., & Hargittai, E. (2011). Health information-seeking in the digital age. *Journal of American College Health, 59*(5), 379–386.

Phua, J. J. (2010). Sports fans and media use: Influence on sports fan identification and collective self-esteem. *International Journal of Sport Communication, 3*(2), 190–206.

Rech, J. (2007). Discovering trends in software engineering with google trend. *ACM SIGSOFT Software Engineering Notes, 32*(2), 1–2.

Report: TSN and CFL agree on deal worth $40-million a year. (2013, March 21). *The Globe and Mail*. Retrieved May 23, 2014, from http://www.theglobeandmail.com/sports/football/report-tsn-and-cfl-agree-on-deal-worth-40-million-a-year/article10053800/

Rex Yuxing, D., Ye, H., & Damangir, S. (2015). Leveraging trends in online searches for product features in market response modeling. *Journal of Marketing*, *79*(1), 29–43.

Ripberger, J. T. (2011). Capturing curiosity: Using Internet search trends to measure public attentiveness. *Policy Studies Journal, 39*(2), 239–259.

Santillana, M., Zhang, D. W., Althouse, B. M., & Ayers, J. W. (2014). Research and practice methods: What can digital disease detection learn from (an external revision to) Google Flu Trends? *American Journal of Preventive Medicine*, *47*(3), 341–347. doi:10.1016/j.amepre.2014.05.020

Scharkow, M., & Vogelgesang, J. (2011). Measuring the public agenda using search engine queries. *International Journal of Public Opinion Research*, *23*(1), 104–113.

Schmidt, E. (2014). *Google: How Google works*. New York: Grand Central Publishing.

Shank, M. D., & Lyberger, M. R. (2015). *Sports marketing: A strategic perspective*. Oxford: Routledge.

Silverstein, C., Marais, H., Henzinger, M., & Moricz, M. (1999, September). Analysis of a very large web search engine query log. *SIGIR Forum, 33*(1), 6–12.

Super Bowl 48 at NFL.com - Official Site of the National Football League. (n.d.). *NFL.com*. Retrieved May 22, 2014, from http://www.nfl.com/superbowl/history

Teghtmeyer, J. H. (2014). Media consumption of Alberta youth and voting participation.

Trends Help. (n.d.). *Trends Help*. Retrieved June 7, 2014, from https://support.google.com/trends

Vaughan, L., & Romero-Frias, E. (2014). Web search volume as a predictor of academic fame: An exploration of Google Trends (English). *Journal of the Association for Information Science and Technology*, *65*(4), 707–720.

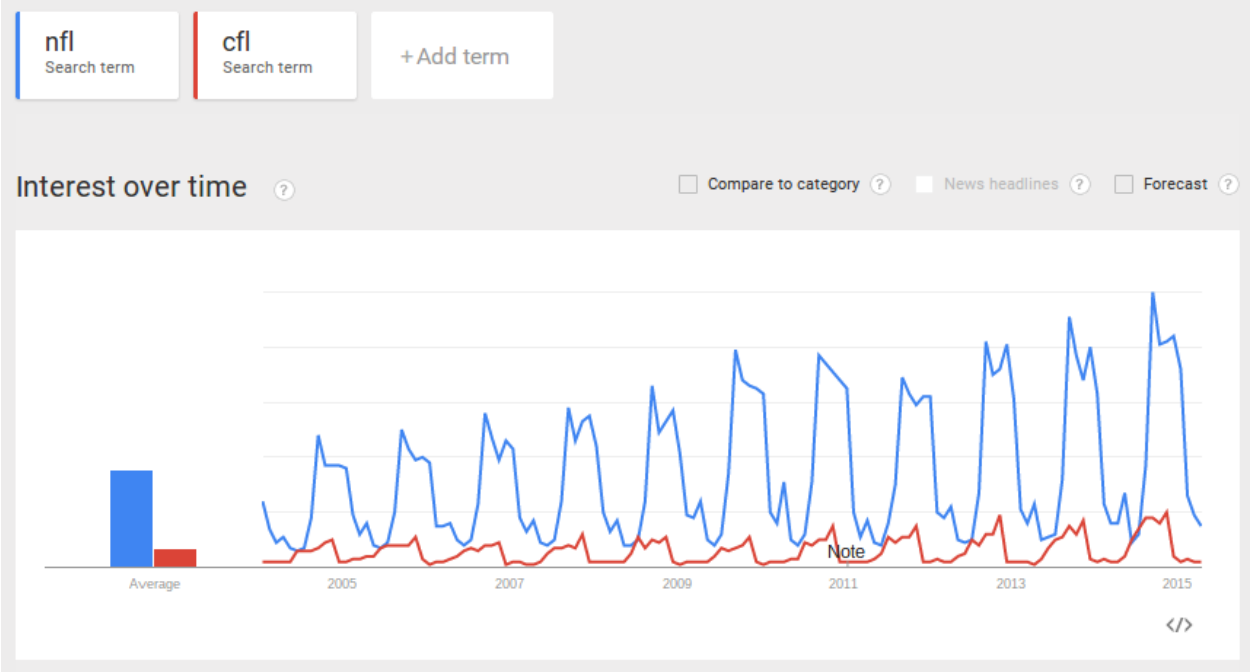
Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: Survey-based indicators vs. Google Trends. *Journal of Forecasting, 30*(6), 565–578.

Wilde, G. R., & Pope, K. L. (2013). Worldwide trends in fishing interest indicated by internet search volume. *Fisheries Management and Ecology*, *20*(2-3), 211–222.

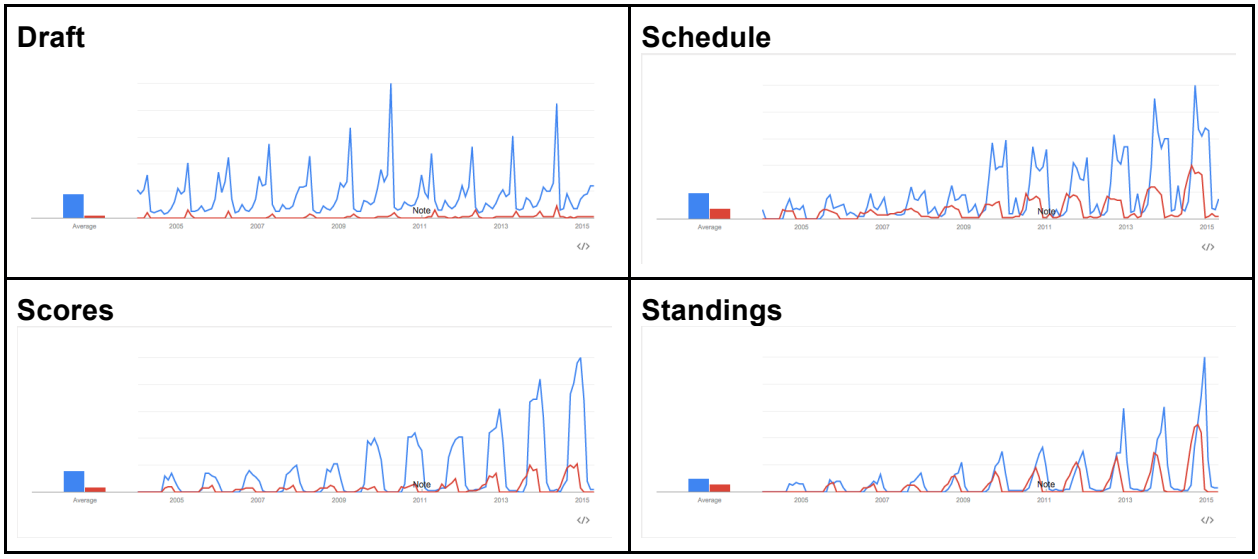
Won Jae, S., Green, B. C., Yong Jae, K., Seunghwan, L., & Schenewark, J. (2007). The effect of web cohesion, web commitment, and attitude toward the website on intentions to use NFL teams' websites. *Sport Management Review (Sport Management Association of Australia & New Zealand)*, *10*(3), 231–252.

Appendices

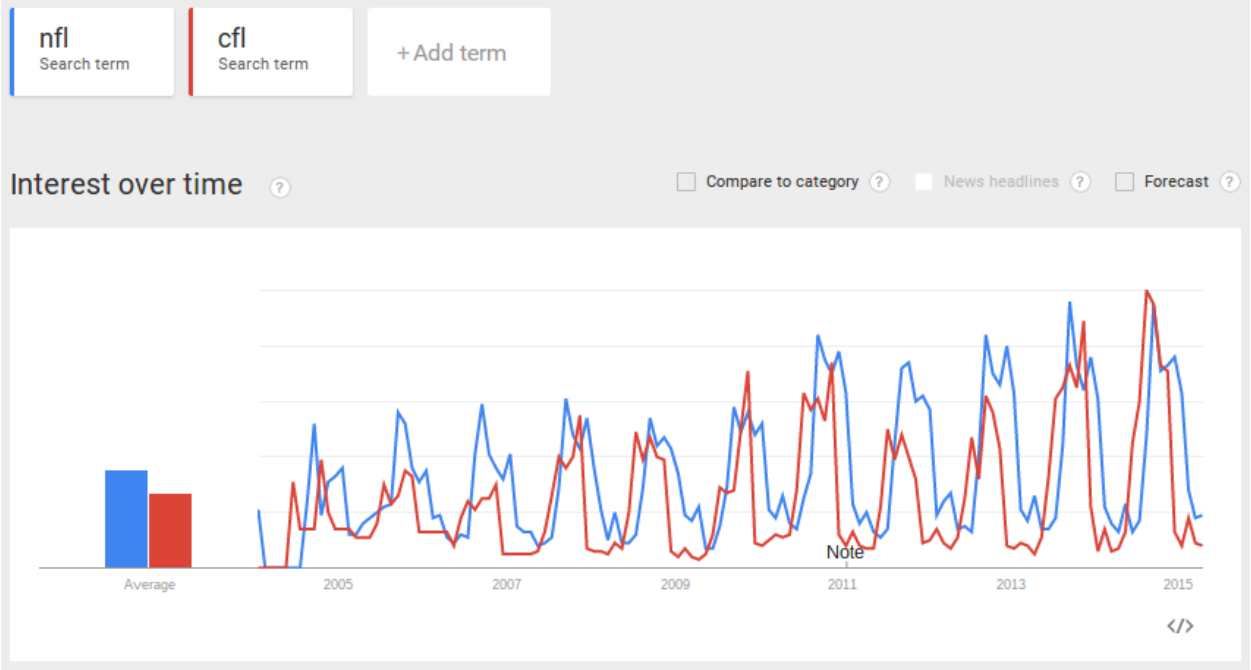
Appendix A. CFL vs. NFL (Canada), 2004-2015



Other Queries

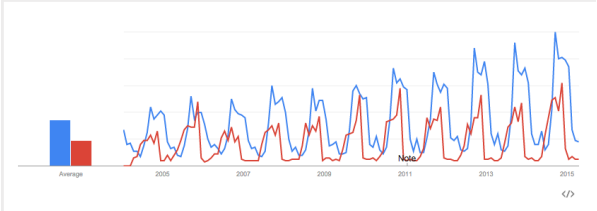


Appendix B. NFL vs. CFL (Saskatchewan), 2004-2015

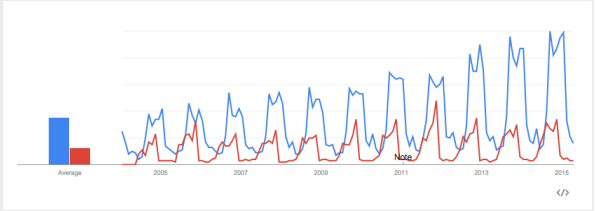


Other Provinces

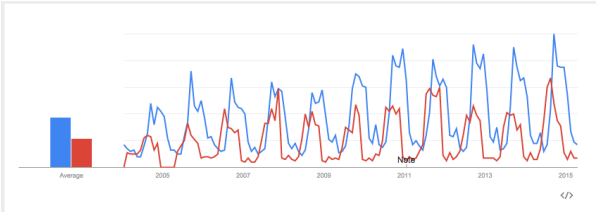
Alberta



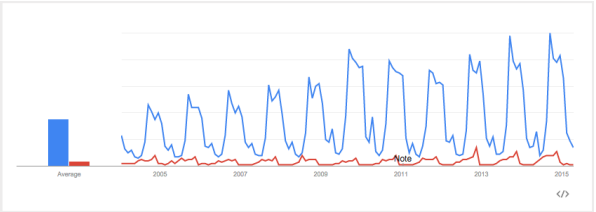
British Columbia



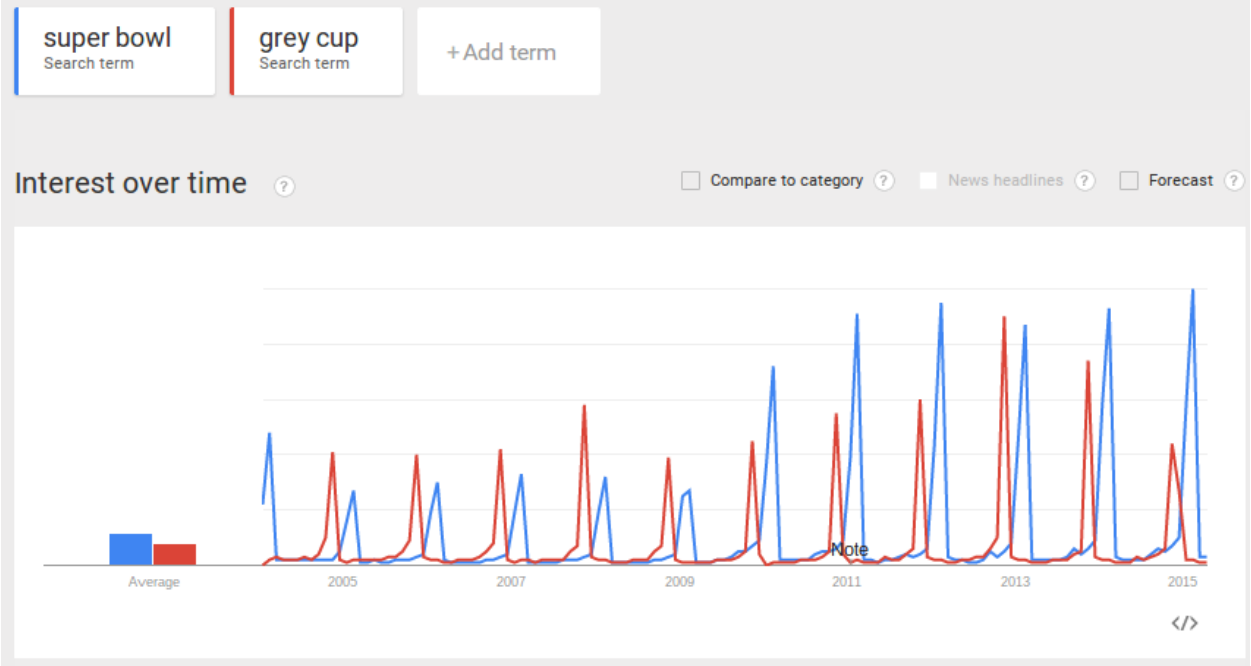
Manitoba



Ontario

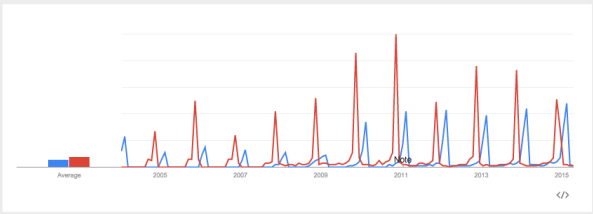


Appendix C. Super Bowl vs. Grey Cup (Canada), 2004-2015

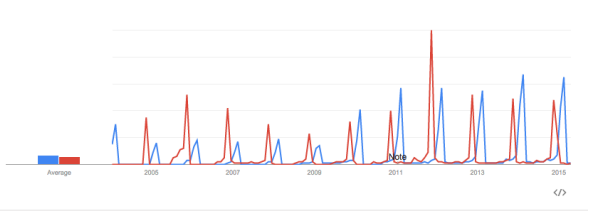


Other Provinces

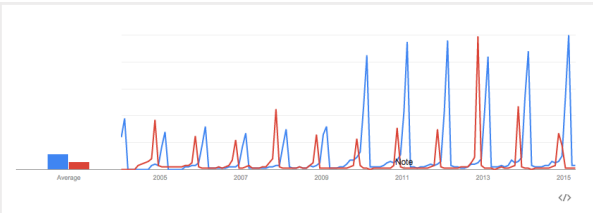
Alberta



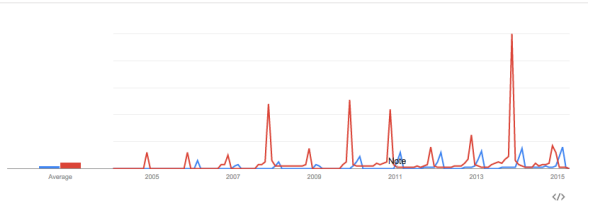
British Columbia



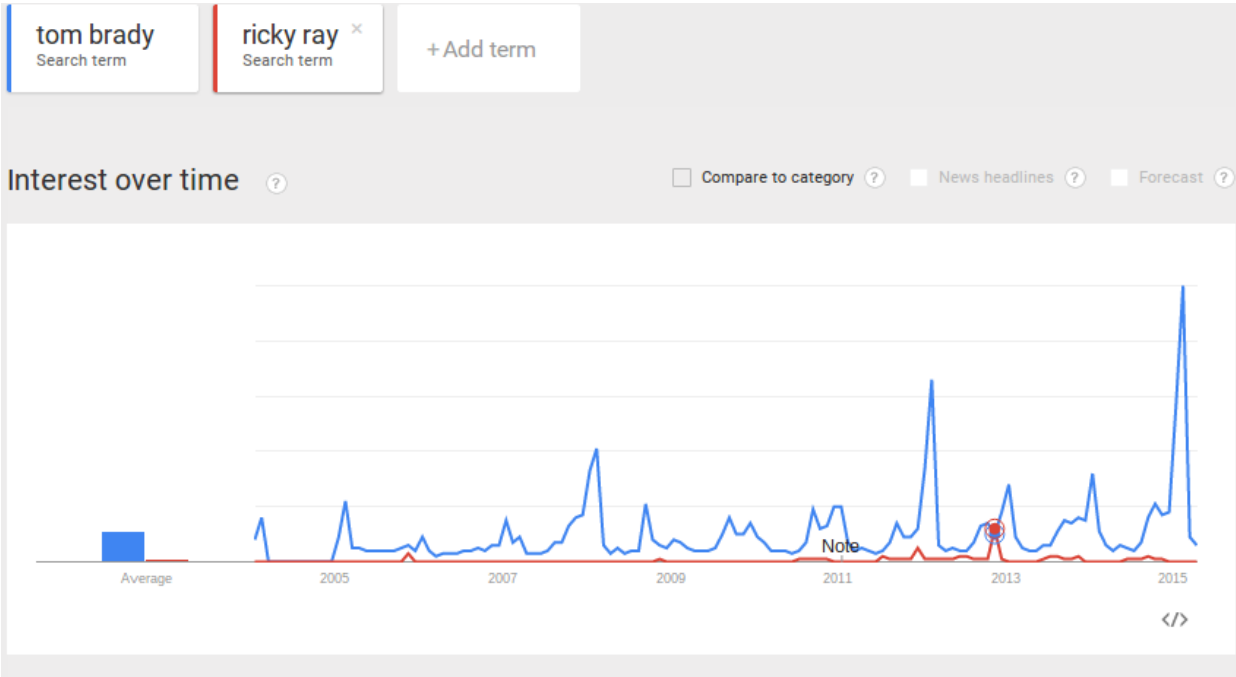
Ontario



Saskatchewan

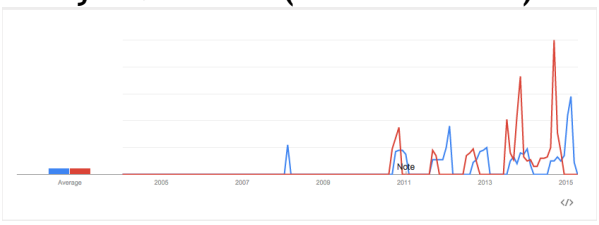


Appendix D. NFL Player Tom Brady vs. CFL Player Ricky Ray (Canada), 2004-2015

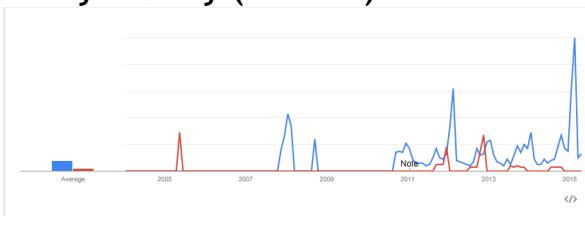


Other Players in Other Provinces

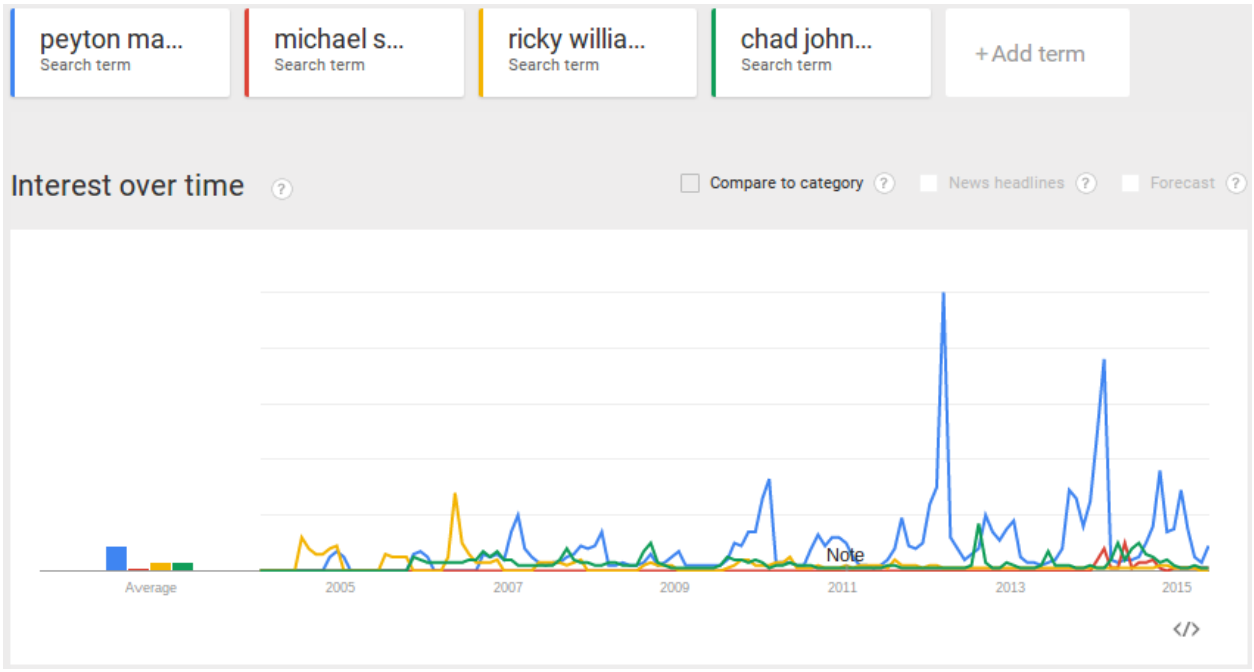
Brady vs. Durant (Saskatchewan)



Brady vs. Ray (Alberta)



Appendix E. Players Who Came to CFL from NFL in Canada, 2004-2015



Appendix F. NFL vs. CFL Teams in Canada, 2004-2015