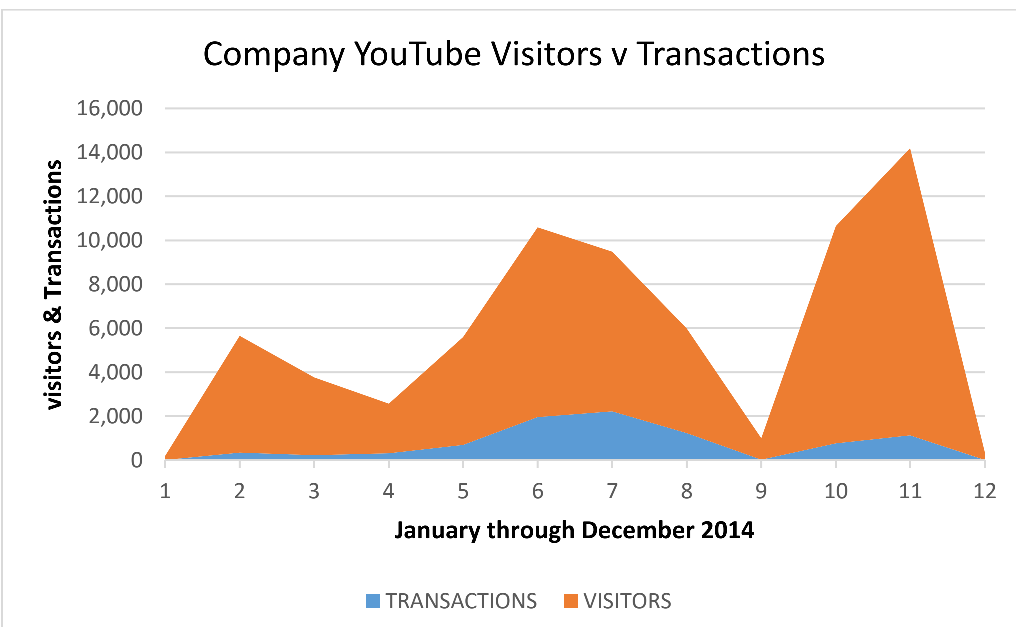
Beyond Traditional SERPs
~
Optimizing for YouTube

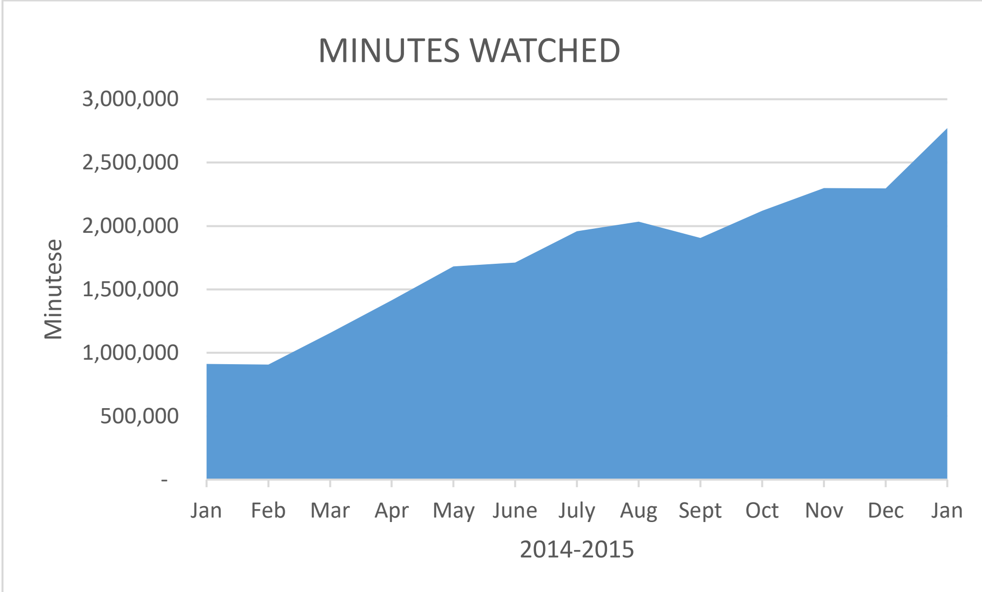
The SEO practitioner is currently faced with an interesting dilemma. The process of optimization has become to a large extent commoditized. In this increasing competitive environment, not only do practitioners need to differentiate their practices but they need to expand their capabilities with respect to driving views and ultimately positively impacting on customer conversions. How can we evaluate the customer journey in order to develop strategies aimed at influencing and determining the optimal channel mix? The basic question that comes from this bit of introspection is what does the long term evolution of SEO look like?

There is no question that we are in the era of big data. However, in the context of SEO, we are just entering the era of rigorous analysis. We need to start small and build upon a solid foundation of good data and sound technique. The area of social media presents an excellent starting point for this type of research. With respect to our client Company, their foray into YouTube has enabled the collection and analysis of customer journey data with a direct link to conversion. Further, a parallel data set that did not include YouTube in the channel mix was examined so as to compare the conversion rates between groups. The focus was on the grail of determining attribution. Does YouTube optimization make a difference in contrast to conventional web optimization?

Data was collected from YouTube analytics tracking information for the period of January 1, 2014 through December 31, 2014. This data included not only site visitors but a wide range of social media related variables. Variables included video views, minutes viewed, shares, likes, dislikes, subscribers, and most importantly, transactions directly generated from the YouTube channel.

The strategy for the analysis began as we worked to optimize Company's videos/content for high-value keywords. Over a period of time and through various optimization efforts (content, comments, copy, back-linking, etc.), we noted tremendous growth in the number of minutes watched (300% over the one year period). Month-on-month the minutes viewed grew at an average rate of 10% per month. Furthermore, subscribers more than doubled over the one year period. Overall engagement also increased as viewers began asking questions, suggesting content, sharing stories, and leaving more comments. Basic descriptive statistics are presented in Table 1 in the Appendix.





One of the pitfalls of statistical analysis is to establish a research hypothesis that is beyond the capabilities of the data to confirm or reject. In order to establish a basic understanding of the association of the different variables in the data set, a correlation matrix was calculated for all the variables. Considerations relative to the statistical method used finally fell in favor of Pearson's r. The large sample size was the real determining factor assuming normally distributed data. The results are presented as Table 2 in the Appendix. The most interesting result from this matrix was the high correlation between visitors and transactions. Although an obvious result, it lead to further exploration into the association between these variables.

Consequently, although many different variables were available for analysis, the simple approach was adopted. We looked at the correlation of YouTube visits with transactions in greater detail. As has already been said, there was a significant positive relationship. Although causality cannot be established, we can safely say there is a strong association between YouTube visits and transactions. The elimination of the other broad based variables from the correlation analysis alters the results slightly. However, the difference is small and for our purposes insignificant.

There are a few points that are worth exploring beyond the month on month type of analysis. The question arises as to the delayed impact of YouTube viewership. To answer this question, a number of "lag" variables were created to formally look at this phenomenon. Four lag variables for visitors were created. The question we were asking was what impact does the number of visitors in the previous month have on current transactions? What was the impact of visitors two months prior, and three months prior, and finally, four months prior?

A correlation matrix was calculated for these four lag visitor variables as well as the concurrent visitor variable with current transactions. What we find is the impact is immediate as we see the highest correlation coefficient. In addition, we see that the impact is distributed over a four month period. Although the correlation coefficient declines over this period, there remains a significant positive correlation between visitors and transactions over a four month time frame. The value of the correlation coefficients may be considered as proxies for "weights" attributed to the specific monthly lag variable.

This result is a reflection of a distributed lag model. Distributed lags are not uncommon as we see this kind activity in a wide range of business sectors. Effects are not always immediate. In this instance you can imagine that customers don't view a video and immediately book a vacation. They may book in the same month or come back to it in a future month when circumstances are right for them. The main point here is to identify this type of purchase behavior. Where the traditional SEO practitioner is looking for immediate results from the SERPs, we have revealed a different but not unexpected customer behavior pattern.

In looking at the non-YouTube customer journey data, we see a much lower conversion rates as opposed to the customer journey including YouTube in the channel mix. There may be any number of explanations for this result. However, it seems clear that driving YouTube as part of the customer journey has a significant influence on transactions. Optimizing for YouTube is a tactic that has important consequences for not only Company but for any enterprise with either a strong YouTube presence or the inclination to explore video as a marketing tool. Clearly, companies without a YouTube position are clearly missing an opportunity to drive business activity.

A matching distributed lag correlation matrix was calculated for the non-YouTube channel mix data. The results are clearly not as interesting as with the YouTube channel mix. When YouTube is absent from the customer journey, the effect is more immediate but less positively correlated. Conversion rates related to visitors remains positively correlated, but the lag impact is far less. We see a one month lag as positively correlated but beyond that point the effect is negligible.

Looking at the graphical data presented below, it is clear there is a seasonal variation in the data. In statistical terms, we may call this a cyclostationary process. To state this in a different way, we have a seasonal variation that also has regular intervals year on year. In addition, we may hypothesize that each repeating season may exhibit the same values as far as conversion rates. But driving the viewership of the YouTube channel may alter this interpretation. We may argue that this is a nonstationary process in which case each repeating season may exhibit a different conversion rate from the previous season. In this example, we may see in future analysis that the conversion rates from subsequent matching seasons are increasing. If the current statistics are a harbinger of future YouTube efficacy, this in fact may be the case.

The results of this study illustrate a few major points of relevance. Optimizing for YouTube is an important if not critical aspect of any SEO campaign driving viewership and conversions. Social media cannot be ignored as a viable optimization channel. This is a clear direction for SEO practitioners as they move from nano-analytics to big analytics. The sector needs to think wider and expand into new directions as big data provides the guidance. As data gives us a path, we need to match it with more robust analytics in order to be effective as practitioners.

It’s not just Company who can benefit from these findings!

APPENDIX

Table 1

Month	TXNs	VISITORS	VIEWS	SUB SCRIBERS	MIN. WATCHED	LIKES	DISLIKES	Shares
Jan 14	10	185	285,831	519	911,409	1,158	41	540
2/14	341	5,310	282,519	514	906,528	1,390	58	719
3/14	228	3,527	349,930	703	1,156,741	1,488	51	646
4/14	306	2,257	432,660	887	1,414,967	1,745	66	754
5/14	696	4,901	509,396	952	1,681,681	2,230	62	1,098
6/14	1,954	8,629	530,053	1,073	1,711,202	2,299	66	1,116
7/14	2,219	7,270	602,422	1,106	1,959,433	2,676	90	1,162
8/14	1,230	4,759	631,998	1,241	2,034,594	2,634	87	1,228
9/14	21	972	605,189	1,105	1,906,018	2,222	85	1,341
10/14	770	9,884	661,769	1,141	2,118,879	2,626	71	1,018

11/14	1,122	13,068	716,955	1,151	2,299,729	2,672	84	1,130
12/14	10	353	709,712	1,127	2,295,989	2,738	98	1,713

Table 2

Correlation Matrix

Transactions	1
VISITORS	0.70353168
VIEWS	0.15210004
SUBSCRIBERS	0.34752583
MINUTES WATCHED	0.1539397
LIKES	0.27056292
DISLIKES	0.19071109
Shares	0.08870218
YouTube Search Mins Viewed	0.13946974
YouTube Suggested Video	0.2651075
Unknown Direct	0.00044853
Embedded Player	-0.25625239
Advertising	-0.420045
YouTube Search Views	0.13559038
YouTube Suggested Video2	0.2446942
Unknown Direct3	-0.05976851
Embedded Player4	-0.2637169
Advertising5	-0.31786928

Cohen's standard was used to evaluate the correlation coefficient to determine the strength of the relationship, or the effect size, where coefficients between .10 and .29 represent a small association; coefficients between .30 and .49 represent a medium association; and coefficients above .50 represent a large associate or relationship.

These increases revealed some surprising findings:

- YouTube viewers who click-through to site have a higher conversion rate. [%?]
- YouTube viewers who click-through and purchase spend [%?] more than purchasers who come through other channels.
- [%?] of people who watched the videos clicked through to the Company site.