Power Up Your Data by Bridging Language in Social Media

Ipsos' Colin Ho, Wendy Hunt and Claudia Meillat use language as a bridge between surveys and other data sources to achieve better results

Power Up Your Data by Bridging Language in Social Media

Our Point of View: Despite growing cynicism of survey research and its ability to give accurate information, our case study provides evidence that what people report from surveys can be accurate. But more importantly, connecting data from two sources, including unstructured text data, provides greater contextualization of the results.

Political leaders have long known that a common language unifies a country. A language facilitates the exchange of ideas, values and commerce. When people speak different languages, cultural and political differences are accentuated, and connections are not formed.

Parallels can be seen in the multiple data sources we have today. They all appear to be of different "languages." Survey data, for example, appears to be mostly about structured data (numbers) whereas social media data is mostly about unstructured data (text). Using survey and Twitter data, we illustrate how language can serve as a bridge. Why these two data sources? Survey is a primary data source used by market researchers, and Twitter is a leading exemplar of the available digital data sources. The process we use to synthesize findings from these two data sources, however, is applicable when bridging across all other data sources.

It's estimated that around 80-90% of usable business data is unstructured data (primarily text). This includes social media, online brand chatter, product reviews and emails. If true, a logical unifier across all these data sources is language. More specifically, concepts and ideas mined from text can provide linkages between data sources. A concept or idea extracted from survey data, for example, provides a connection to social media if the same concept is found there. We illustrate this by showing how survey and Twitter data, connected via consumers' language, provide a more holistic view of beer brands.

A Tale of Two Data Sources

Our data came from two sources: survey and Twitter. We fielded the survey among a representative sample of 900 beer drinkers in the last quarter of 2015. What was different in our survey was the usage of open-ended questions as the primary measuring instrument. This is a key departure from traditional surveys where closed-ended questions are the core instruments and open-ended questions are simply a "nice to have" or afterthought. We used open-ended questions to capture consumers' top of mind associations for three beer brands: Bud Light, Sam Adams and Stella Artois. Open-ended questions allowed us to capture consumers' spontaneous brand associations without the prompting or constraints of researcher generated attributes. The verbatim from the open-ended questions were text analyzed.
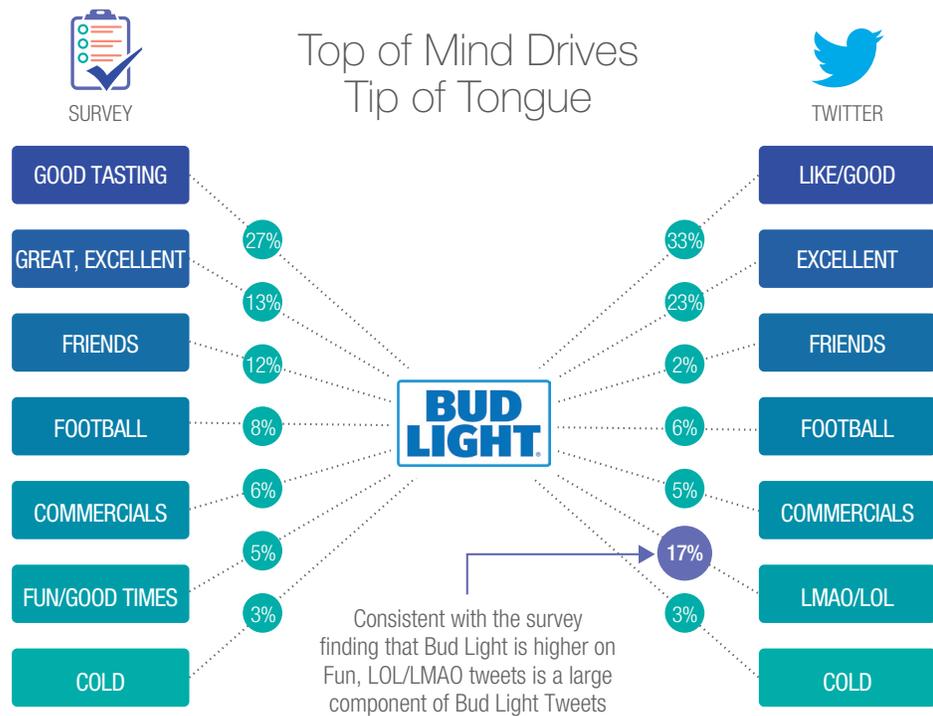
For our Twitter data source, we first identified tweets including mentions of any of the three brands in 2015. A random sample of 3,000 tweets each were downloaded for Bud Light and Sam Adams. We extracted 981 tweets for Stella Artois, which were all used. The extracted tweets were analyzed using the same text analytic software and process used to analyze the survey verbatim. This ensured that our comparison of the results between the survey and tweets was not influenced by the text analysis software or analytical process.

Bud Light: Hanging Out with Your Buddy

The survey revealed that Bud Light was associated with "Friends," "Football," "Commercials," "Fun/Good Times" and "Cold" (see left side of Figure 1). When these brand associations were compared with what was said about Bud Light on Twitter (see right side of Figure 1), we see a high degree of overlap. Put simply, top of mind leads to tip of tongue. For example, "Football" was a top of mind association from the survey as well as a topic tweeted for Bud Light. Besides validating the brand associations from the survey, however, Twitter also contributed incremental insights. For example, Bud Light's association with "Fun/Good Times" in the survey were substantiated by a relatively large proportion of "LMAO/LOL" tweets — most of which were related to Bud Light's humorous and entertaining commercials. Twitter, therefore, let us see how Bud Light manifested itself in consumers' lives. Consumers share Bud Light's commercials with their friends because it helps them feel part of a group and to be accepted. From a broader perspective, Bud Light's commercials allow people to build social currency via sharing. What is central to this paper, however, is how we have linked two apparently different data sources via the concepts extracted from text.

Figure 1. Bridging Survey and Twitter Data via Language, Bud Light

Top of Mind Drives Tip of Tongue

SURVEY

GOOD TASTING	27%
GREAT, EXCELLENT	13%
FRIENDS	12%
FOOTBALL	8%
COMMERCIALS	6%
FUN/GOOD TIMES	5%
COLD	3%

TWITTER

33%	LIKE/GOOD
23%	EXCELLENT
2%	FRIENDS
6%	FOOTBALL
5%	COMMERCIALS
17%	LMAO/LOL
3%	COLD

Consistent with the survey finding that Bud Light is higher on Fun, LOL/LMAO tweets is a large component of Bud Light Tweets
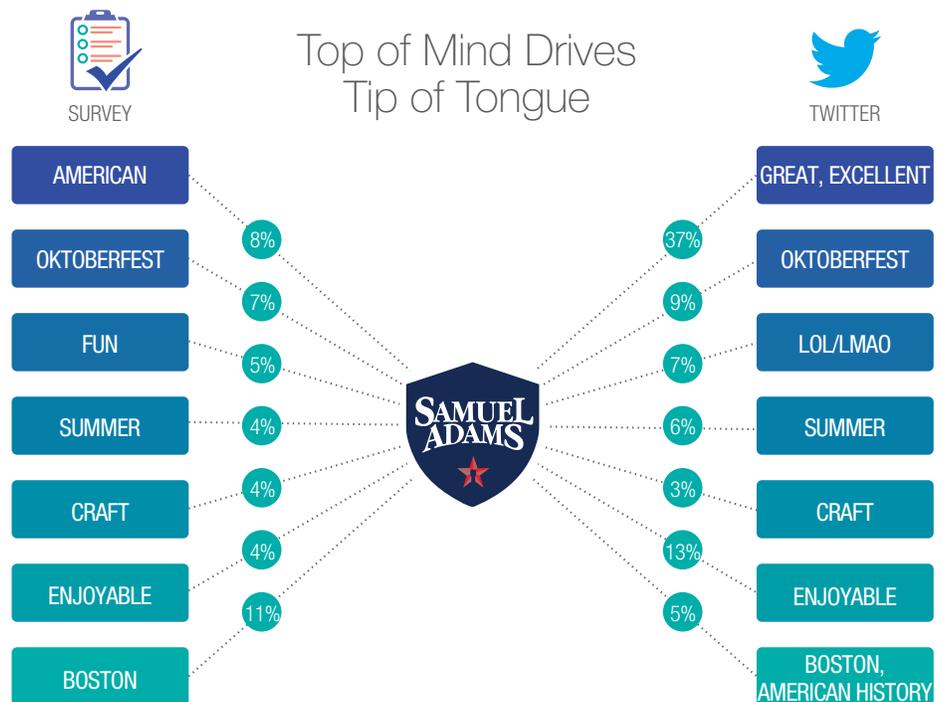
Sam Adams: The Land of the Free and Home of the Brave

We were able to create a similar bridge between survey and Twitter data for the Sam Adams brand. In the survey, Sam Adams was associated with "American," "Boston," "Fun," "Craft," "Summer," and "Octoberfest" (see left side of Figure 2). Similar to Bud Light, top of mind associations appeared to drive Sam Adams' tweets (see right side of Figure 2).

As with Bud Light, Twitter provided more than just validation of the survey results. It also provided a rich context for understanding the Sam Adams brand.

Looking at a sample of the tweets that related to Boston/American History (Figure 3), we are able to sense the underlying patriotism for this brand. This contextual information would be difficult to capture in a survey. Given the nature of Twitter, the tweets on Sam Adams provide a lens into consumers' free flowing thoughts and feelings about the brand without any researcher's prompting.

Figure 2. Bridging Survey and Twitter Data via Language, Sam Adams

Top of Mind Drives Tip of Tongue

SURVEY

AMERICAN		GREAT, EXCELLENT
OKTOBERFEST	8% — 37%	OKTOBERFEST
FUN	7% — 9%	LOL/LMAO
SUMMER	5% — 7%	SUMMER
CRAFT	4% — 6%	CRAFT
ENJOYABLE	4% — 3%	ENJOYABLE
BOSTON	4% — 13% / 11% — 5%	BOSTON, AMERICAN HISTORY

SAMUEL ADAMS

TWITTER

Figure 3: Sample tweets for Sam Adams relating to Boston/American History

tweets

RT @Swellcorrell Come home to pops knocking back a Sam Adams and blaring Pink Floyd through the house. It's good to be an American.

Thank you for great beer, Sam Adams! And Declaration of Independence, too. #Boston @ Granary Burying Ground – https://t.co/FkPuJmafot

Taking some time off enjoying the #Freedomtrail. Sam Adams a great patriot and wonderful beer maker. – http://t.co/eSaX3lqDoQ
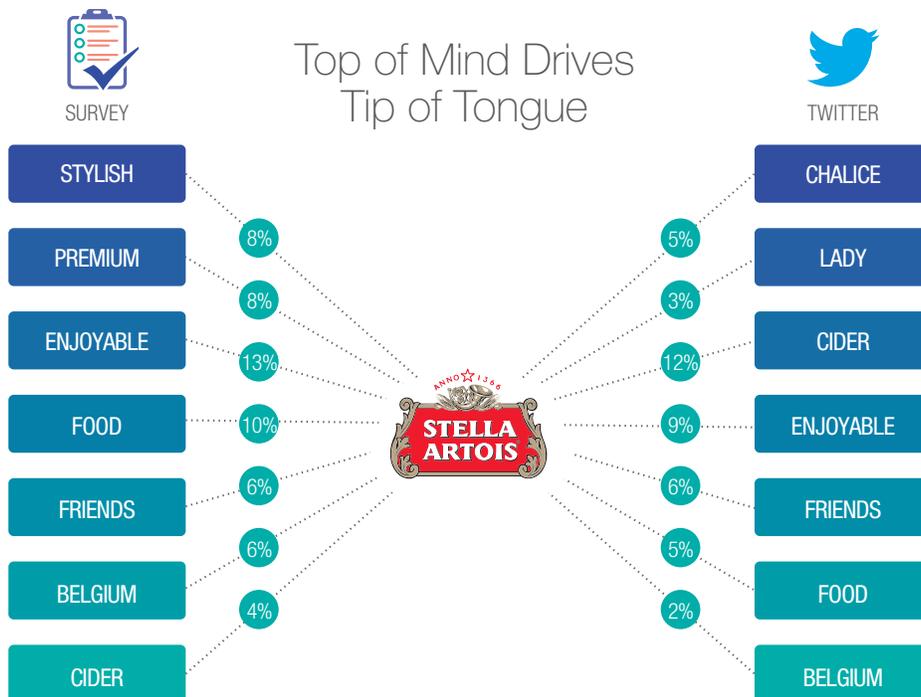
Sam Adams is a free tour where they give you 21 ounces of free beer in a free glass you get to keep. America is awesome.

RT @SirEviscerate Sam Adams is good. There should be more beers named after people mistakenly believed to be presidents, like Ben Franklin or George W. Bush.

I am in the land of Sam Adams, Dunkin' Donuts, the American Revolution and Chris Evans. Life is good. – https://t.co/L6kZB02UhQ

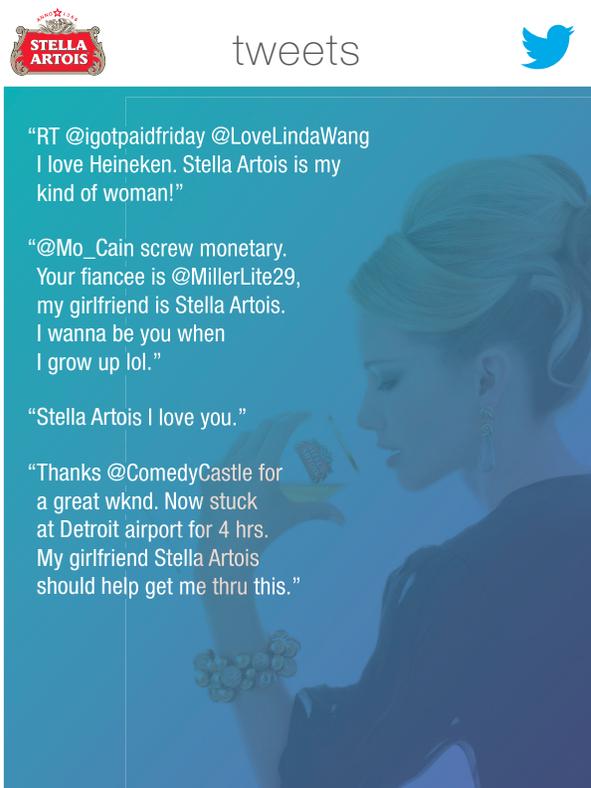# Stella Artois: A Drink with an Elegant Lady

Finally, we were also able to create a bridge from survey to Twitter data for the Stella Artois brand. Stella Artois' brand associations of "Premium," "Belgium," "Cider," "Stylish" and "Food" from the survey (see left side of Figure 4) all bridged to Twitter (see right side of Figure 4). Some of the concepts that bridged the two data sources were essentially the same exact concept (e.g., Belgium, Food). Other concepts that bridged the two data sources were different manifestation of a central idea. For example, "Stylish" and "Premium" were brand associations captured in the survey but manifested on Twitter as "Chalice." The "Chalice" comments on Twitter refer to the stylish and somewhat pricey cup the brand markets and sells to consumers.

Top of Mind Drives Tip of Tongue

SURVEY

| STYLISH |
PREMIUM	8%
ENJOYABLE	8%
FOOD	13%
	10%
FRIENDS	6%
BELGIUM	6%
CIDER	4%

TWITTER

| CHALICE |
LADY	5%
CIDER	3%
ENJOYABLE	12%
	9%
FRIENDS	6%
FOOD	5%
BELGIUM	2%

Figure 4. Bridging Survey and Twitter Data via Language, Stella Artois

Twitter also provided additional contextualization of the survey results. An additional theme captured in Twitter but not in the survey was "Lady." On reviewing a sample of these tweets, it is apparent that the Stella Artois brand is personified as a lady (see Figure 5).

Figure 5.
Sample "Lady" tweets of Stella Artois

STELLA ARTOIS   tweets

"RT @igotpaidfriday @LoveLindaWang
I love Heineken. Stella Artois is my
kind of woman!"

"@Mo_Cain screw monetary.
Your fiancee is @MillerLite29,
my girlfriend is Stella Artois.
I wanna be you when
I grow up lol."

"Stella Artois I love you."

"Thanks @ComedyCastle for
a great wknd. Now stuck
at Detroit airport for 4 hrs.
My girlfriend Stella Artois
should help get me thru this."

In essence, Stella Artois' premium and stylish associations manifest themselves as an emotional connection with a beer personified as an elegant woman. This is likely an outcome of the brand's 2010 campaign where Stella Artois is depicted as a "She" and as "one of the finer things in life."

We have seen how concepts extracted via text mining allowed us to bridge two seemingly different data sources together. The key is to understand that when it comes to consumer insights, it is about brand beliefs, motivations and goals. These are all abstract concepts that exist in our minds but can be captured with language. These concepts provide a simple but powerful way to link findings across data sources. Abstract consumer concepts are also best captured in context to eliminate any ambiguity in terms of what a concept means. In our case study, the tweets provided greater contextualization of the concepts found from the survey. Twitter, however, yielded more than just a greater contextualization of the survey results: it also allowed us to capture the salience of brand concepts over time. By so doing, we obtained insights into how the concepts that appear in both data sources changes in salience over time.
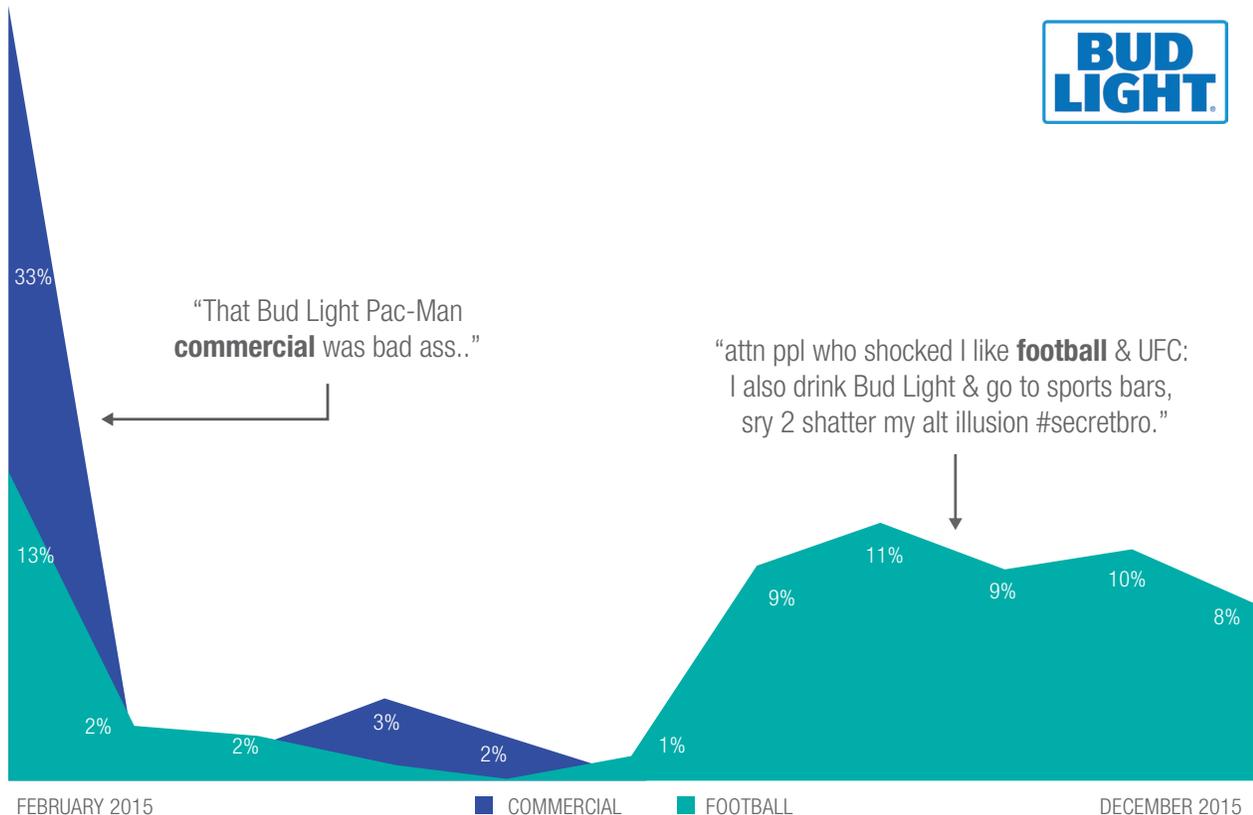
What's on Your Mind Now?

By examining the nature and volume of tweets over the course of 2015, we were able to determine what was being said about a brand, how much was said and when it was said. We look at two of the brands again and get a brief glimpse of how different each brand's "conversation" over the course of 2015 was. Stella Artois' sample size was too small to be analyzed by month. In the case of Bud Light, the connection with the brand appears to be strongest when football season is happening and peaking during the Super Bowl. Once football season is over, however, tweets of football or commercials practically disappear.

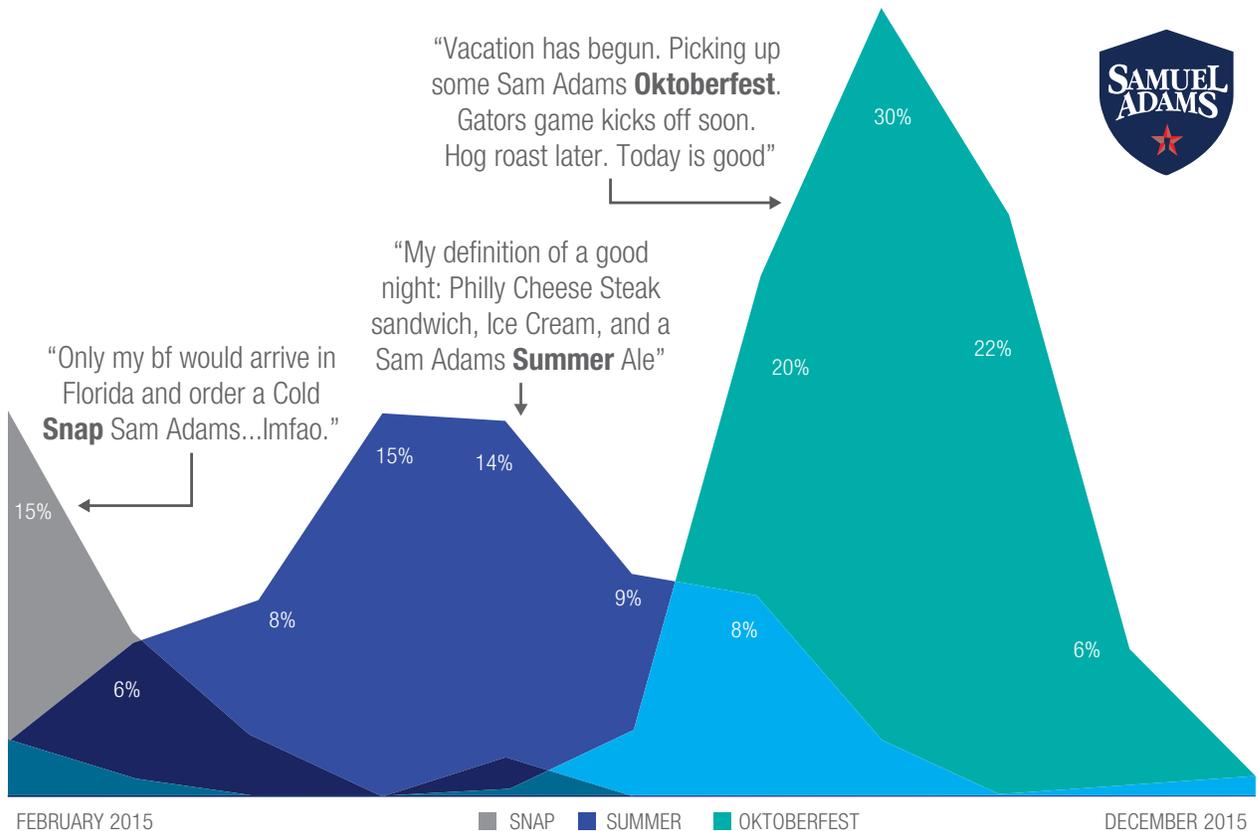> " ...the brand appears to be strongest when football season is happening... "

Figure 6. Tweets of football and commercial for Bud Light in 2015

"That Bud Light Pac-Man **commercial** was bad ass.."

"attn ppl who shocked I like **football** & UFC: I also drink Bud Light & go to sports bars, sry 2 shatter my alt illusion #secretbro."

33%

13%

2%

2%

3%

2%

1%

9%

11%

9%

10%

8%

FEBRUARY 2015   ■ COMMERCIAL   ■ FOOTBALL   DECEMBER 2015

In contrast, Sam Adams appears to generate conversations more constantly over the course of a year. The brand chatter for Sam Adams over the course of a year varied with its seasonal product offerings (see Figure 7). At the beginning of 2015, the chatter for Sam Adams was mostly about its winter products. In summer, Sam Adams chatter was mostly about its summer products. In fall, Sam Adams chatter was about its Oktoberfest line. From a marketing perspective, Sam Adams seasonal beer's strategy may well ensure that the brand-person relationship stays constant throughout the year. With each new season, tweets of Sam Adams are renewed with vigor. In short, Sam Adams' seasonal beers generate connections throughout the year and also re-invigorate the relationship on a quarterly basis. From these two brands, it is clear that the concepts that bridge the two data sources have different salience across time.

Figure 7.
Tweets of Sam Adams Seasonal Beer

"Vacation has begun. Picking up some Sam Adams **Oktoberfest**. Gators game kicks off soon. Hog roast later. Today is good"

"My definition of a good night: Philly Cheese Steak sandwich, Ice Cream, and a Sam Adams **Summer** Ale"

"Only my bf would arrive in Florida and order a Cold **Snap** Sam Adams...lmfao."

30% 22% 20% 15% 14% 9% 8% 8% 6% 6% 15%

FEBRUARY 2015 · SNAP · SUMMER · OKTOBERFEST · DECEMBER 2015

Final Thoughts

By using language as the common thread, we were able to connect the findings from two very different data sources. Concepts of three beer brands found in survey were also found in Twitter. At times, identical concepts were found in both data sources. At other times, the concepts found in survey took on a different manifestation in Twitter. In general, however, it was clear that the two data sources could be effectively linked via concepts. The replication of concepts from survey to Twitter also simultaneously validates both data sources. That is, the finding of similar concepts in both data sources reassures us that what was found in one data source was not just an aberration. Our case study is particularly noteworthy as it provides evidence that what people report from surveys can be accurate. In recent years, there has been a growing cynicism of survey research and its ability to give accurate information. While some of the criticism is justified (e.g., asking consumers to explain why they make a purchase decision), we believe that survey findings can be accurate when we ask people questions they can accurately respond to. In our survey, we simply asked consumers to report their top of mind/spontaneous thoughts of three alcoholic beverage brands.

While bridging across data sources is the key focus of this paper, we believe that there are additional benefits beyond a successful bridging. Multiple data sources can provide greater contextualization of the results. In the case of Twitter, it enabled us to better understand the concepts associated with the three beer brands. Each brand was brought to life when we were able to see the role each brand played in consumers' lives — Bud Light as a buddy to have fun with, Sam Adams as a subtle expression of patriotism, and Stella Artois personified as an elegant female drinking companion. Twitter brings a rich contextualization for understanding how a brand manifests itself in consumers' lives. Finally, as Twitter data are available over time, we were also able to view how brand concepts varied in salience over the course of a year.

History is filled with examples of how countries unified their country through the implementation of a common language. While we have illustrated how findings from survey and Twitter can be unified through language, this process can be easily applied to any unstructured text data. That would include data sources such as online communities, customer complaints via email, and online product reviews and blogs as well. Data unification through ideas and concepts is a simple but powerful approach. Through linkages via ideas and concepts, the potential to create a sum that's greater than the sum of its parts is achievable.

Team

Colin Ho
Senior Vice President
at Ipsos Marketing
colin.ho@ipsos.com

Wendy Hunt
Strategic Facilitator
at Ipsos UU
wendy.hunt@ipsos.com

Claudia Meillat
International Director
at Ipsos Connect
claudia.meillat@ipsos.com

About Ipsos

Ipsos is an independent market research company controlled and managed by research professionals. Founded in France in 1975, Ipsos has grown into a worldwide research group with a strong presence in all key markets. Ipsos ranks third in the global research industry.

At Ipsos we are passionately curious about people, markets, brands and society. We make our changing world easier and faster to navigate and inspire clients to make smarter decisions. We deliver with security, speed, simplicity and substance. We are Game Changers.

With offices in 88 countries, Ipsos delivers insightful expertise across six research specializations: advertising, customer loyalty, marketing, media, public affairs research, and survey management.

Ipsos researchers assess market potential and interpret market trends. We develop and build brands. We help clients build long-term relationships with their customers. We test advertising and study audience responses to various media and they measure public opinion around the globe.

Visit **www.ipsos.com/en-us** to learn more about Ipsos' offerings and capabilities.

GAME CHANGERS | Ipsos