Paper 3352-2015

Tactical Marketing with SAS® Visual Analytics – Aligning a Customer's Online Journey with In-Store Purchases

Johann Pasion, 89 Degrees; Tricia Aanderud, Zencos

ABSTRACT

Marketers often face a cross-channel challenge in making sense of the behavior of web visitors who spend considerable time researching an item online, even putting the item in a wish list or checkout basket, but failing to follow up with an actual purchase online, instead opting to purchase the item in the store. This research shows the use of SAS® Visual Analytics to address this challenge.

This research uses a large data set of simulated web transactional data, combines it with common IDs to attach the data to in-store retail data, and studies it in SAS Visual Analytics. In this presentation, we go over tips and tricks for using SAS Visual Analytics on a non-distributed server. The loaded data set is analyzed step-by-step to show how to draw correlations in the web browsing behavior of customers and how to link the data to their subsequent in-store behavior. It shows how we can draw inferences between web visits and in-store visits by department. You'll change your marketing strategy as a result of the research.

INTRODUCTION

In today's highly competitive retail environment, marketing professionals face the challenge of understanding and promoting to customers in a growing online channel. Analytically integrating online behavior with in-store transactional data gives organizations a significant competitive advantage. Imagine a retailer that could:

- Accurately classify customers by their purchase tendencies: in-store only, online only, or both.
- Know if a customer has a propensity to browse items online but abandon cart and visit the store.
- Identify departments that have a particularly strong channel behavior, such as 90% online only sales.
- Have the ability to predict customer behavior quickly enough to translate into effective marketing actions, such as trigger emails to cart abandoners likely to shop in-store.

This research is inspired from our marketing consulting work with retail clients facing these exact issues. We start by describing the motives behind a client's decision to implement SAS Visual Analytics, and then demonstrate, through simulated marketing data, how it serves as an effective analytical solution to generate the desired marketing insights found in loyalty-card member data.

ISSUES WITH THE RESEARCH CYCLE

Working with transactional data from large retailers introduces many challenges. Large datasets require special skill-sets and tools to prepare data for customer viewing, and reaction time is important. According to McKinsey, "The use of big data is becoming a key basis of competition and growth for individual firms. We estimate that a retailer embracing big data has the potential to increase its operating margin by more than 60 percent." [1] Store transaction data alone can consist of hundreds of thousands of observations on a daily basis for large retailers, so the ability to access valuable parts of the data quickly enough to influence the marketing strategy grants a significant competitive advantage.

In many organizations, there is a divide between the marketing research team and the strategy/brand management team. Part of this split is the specialized nature of marketing research practitioners requiring a skill-set in coding and database manipulation.

The market research team can spend hours coding visualizations. Often datasets are sliced and diced into multiple smaller datasets so the end users can find patterns more easily and determine the next steps. The analyst ports the results into Microsoft Excel or PowerPoint to create results that are easier for the business user or management team to understand. This creates a cumbersome process when working with a large number of datasets on a daily basis.

The second issue of timeliness is highlighted more when the business user relies on a research team to produce actionable results, even for small ad-hoc requests such as looking at purchases of a particular SKU in a store. By the time the research is completed, the window of opportunity to act on the information may have passed, which is a situation unacceptable in a high-stakes retail-marketing environment.

Our research cycle was taking longer because of visualization issues, simple request used people with advanced programming skillsets, and many times marketing opportunities were missed.

IMPLEMENTING SAS VISUAL ANALYTICS

As long time SAS users, we decided SAS Visual Analytics could address our requirements, which were:

- Ability to handle large volumes of marketing data
- Visualize data results without multiple small datasets thereby shortening our review cycle
- Quick implementation and training cycle
- Provide self service capability to the market analysts and program managers
- No additional hardware purchases until we knew it would work

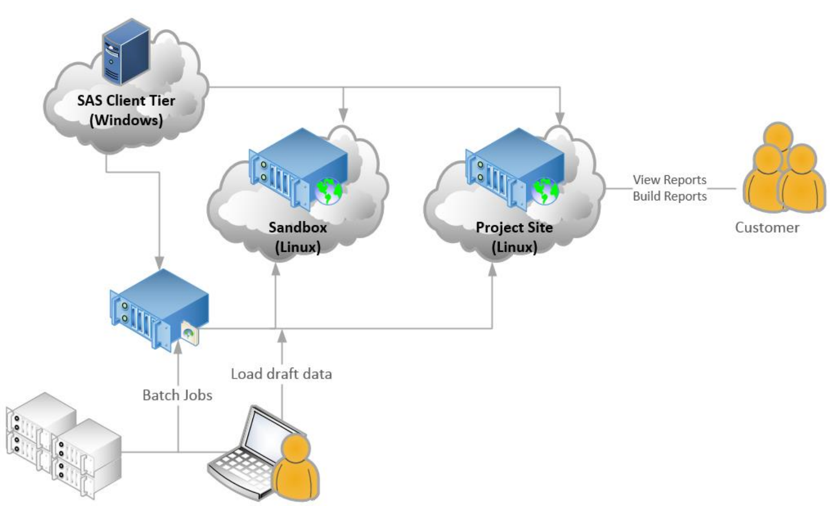
Choosing Distributed Versus Non Distributed SAS Visual Analytics

We implemented the non-distributed version of SAS Visual Analytics 6.3. The non-distributed version runs on a single server and is intended for limited users with smaller data. It does have size limits to the datasets with the largest dataset being around 18GB based on the overall number of users. Note that starting with SAS Visual Analytics 7.1 there are many different sizing options based on the machine core count. The distributed version of SAS Visual Analytics handles larger datasets.

Proof of Concept – In the Cloud!

Our first step was to complete a proof of concept for a smaller project. We chose to implement SAS Visual Analytics in a cloud-based environment. Running SAS Visual Analytics in the cloud-based environment solved many issues for us. Amazon Web Services provide a way to quickly implement and evaluate SAS Visual Analytics before making a long-term commitment. For a reasonable monthly fee, we were able to rent space on their servers and avoid any hardware or administration costs. Our cloud-based environment is shown in Figure 1. There may be some security concerns with having data in the cloud so be sure to discuss this topic with your IT department.

Figure 1 Cloud Based Environment



We had three Amazon servers:

- Linux environment with non-distributed SAS Visual Analytics installed. There were two of these servers, which were used as follows:
  - Sandbox: Area where developers could play with data for learning or create demo reports for new customers. Developers could load data through the user interface or from batch programs that ran nightly.
  - Project Site: Area where customers could view or create reports
- Windows environment with the SAS client tier tools, such as SAS Management Console and SAS Enterprise Guide.

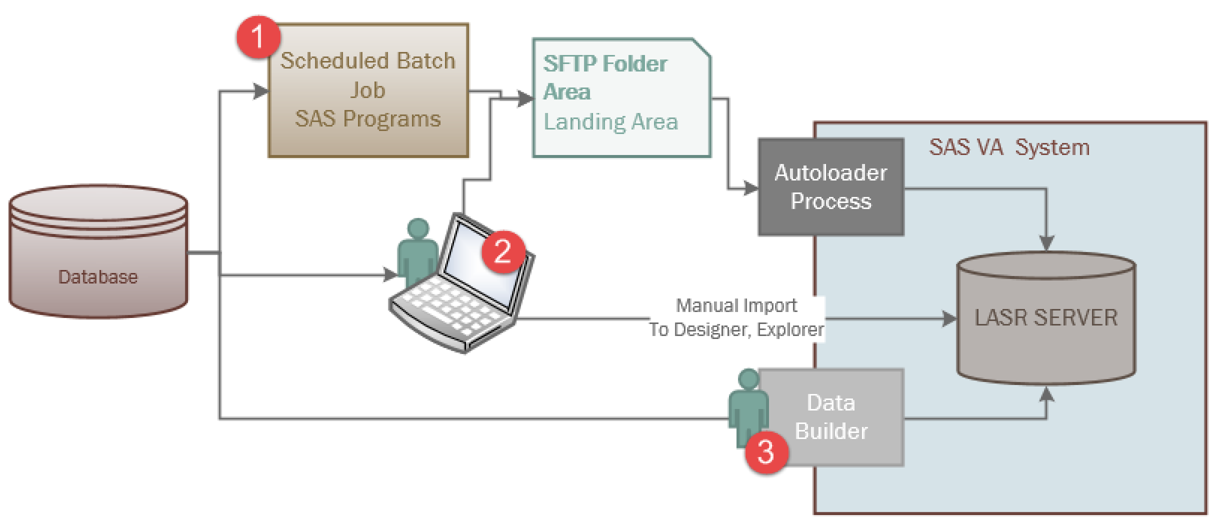
Moving Larger Datasets Around

Within a few weeks' time, we were able to install the software, train the team, and have live data for reports. Since most of our data was already in SAS datasets, there was little extra preparation required. The training was quick because the tool was very intuitive. The biggest issue was how to get the data into the servers.

As shown in the following figure, there are three methods for moving data into the system.

1. Schedule a batch job that runs nightly. This job essentially serves an ETL process since it moves cleaned data to the SFTP server. From there the SAS Visual Analytics Autoloader process copies the data to the LASR Server. The Autoloader process runs every 15 minutes.
2. Users can import data into the SAS Visual Analytics LASR server through the user interfaces. The user can also copy the dataset to the SFTP server and let the Autoloader process run it.
3. SAS Visual Analytics Data Builder can be set up to extract data from a database or even a dataset. The tool has an option to schedule the job so data is reload as desired.

Figure 2 Loading Data into the System

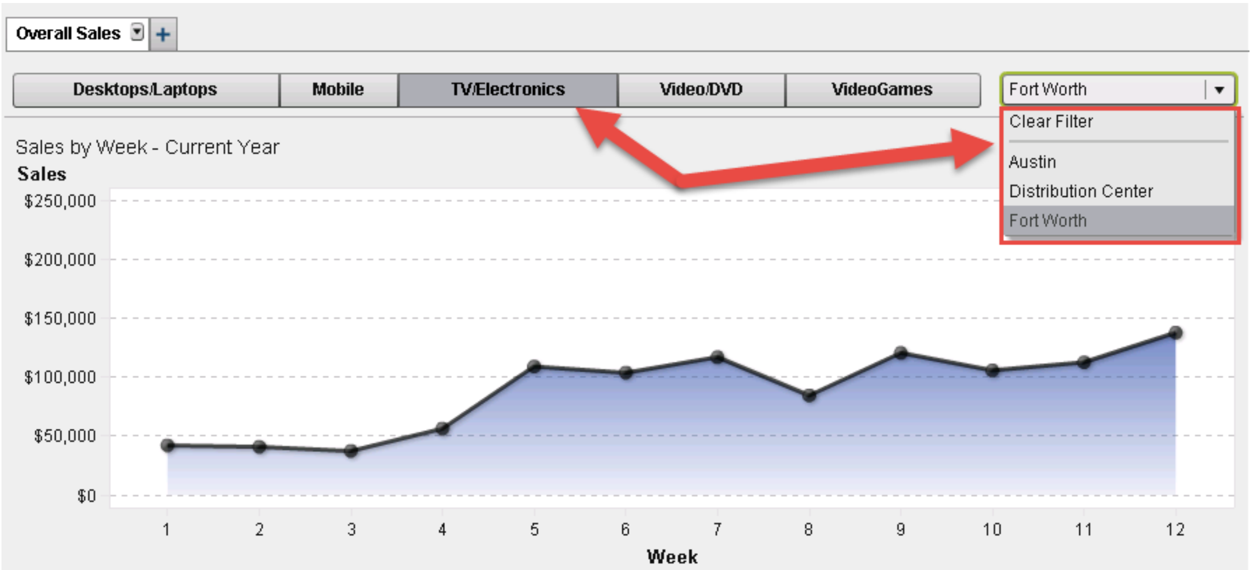


BENEFITS OF SAS VISUAL ANALYTICS

The ease of report creation and modification for both analysts and end-users allows simple research requests to be handled quickly by junior or non-technical personnel. Filters in particular can be placed in reports to allow marketers to drill down to specific stores or regions quickly.

As an example, assume a Marketing Manager wants to determine the weekly trend of sales across different stores or from sales generated online shipped through the distribution center. A pre-built bar chart showing sales by week is available, reporting across the entire organization, but a drop-down filter is available to cut down quickly to specific stores and even departments.

Figure 3 Marketing Managers can filter data multiple ways



CASE STUDY: SEGMENTING MULTI-CHANNEL CUSTOMERS

In this following case study, we used simulated fictional data to demonstrate how SAS Visual Analytics can be employed to effectively visualize data and obtain insights on customer loyalty behavior.

Background

The current retail environment is saturated with loyalty card use, but retailers need to ensure their programs are engaging by understanding customer behavior patterns. Nearly 90% of Americans participate in some type of rewards program, and most are enrolled in more than one. [2] According to a study by RSR Research, 61% of their surveyed retailers felt that "Customer retention has become more difficult and building customer loyalty is challenging"[3]

Customers now shop in a multi-channel environment, with most major retailers having an online store presence. Integrating this data allows the team to understand diverse customer behavioral trends and provide richer segmentation.

According to an article from Bloomberg Research, “Physical retailers are also using big data, storing massive amounts of information on servers and using software to search for trends, to drive more people into their stores." In particular, they noted, "Target creates customer profiles containing a person’s buying history, age, marital status, estimated salary and Web history."[4]

Loyalty cards offer additional informational advantages to retailers, allowing for the tracking of customer behavior in detail throughout different channels, significantly enhancing their promotional effectiveness.

THE RESEARCH OBJECTIVE

The research makes it self-evident that uncovering individual customer patterns with available data is crucially important. In our own experience with various retailer clients, we had found that many customers would browse items online to get product descriptions, prices, and so on. Cart abandonment is of particular interest as customers can use the website to gain product knowledge, but make the actual purchase in-store.

Situation

A marketing manager for a small Texas based chain store has the following general questions about loyalty card customers:

- Online Session Behavior: If a customer session results in having items placed in the online shopping cart, what is the session length and the breakout of pages viewed?
- Channel Purchases: What are the percentage of purchases online versus in-store for individuals that have trackable web browsing activity?
- Customer Location: Where do the majority of our in-store customers reside?
- Abandoned Carts:
  - If a customer abandons their cart, do they come into the store to buy the same products that they have abandoned?
  - If they buy in store, are they doing it with a specific cadence?

DESCRIBING THE DATA

To complete the research, we are focusing on loyalty card members since we have access to their online and in store behavior. The in-store behavior can be determined from the cash register data when the customer's use their loyalty card. The customer also uses the loyalty card ID when they log into the store website. Thus, we can link the datasets by the card ID.

We have an online behavioral dataset. This could come from a service such as Omniture SiteCatalyst, which allows digital analysts to tag certain pages of interest to allow for differentiation of page views across departments as well as cart information. Positive Identification through cookies or logins allows us to track this behavior for a given recognized Card ID.

For simplification, we assume this data has already gone through intermediary preparation by aggregating department-specific page views, overall session length, and creating an abandon cart flag. We also are not considering browsers not linking back by Card ID, focusing on Loyalty Card members.

We have a dataset showing the Session ID, Card ID, and the products a customer had loaded into a card for a given session. This is of importance as we can use this information to see if the customer had subsequently purchased these abandoned items in store.

PROVIDING ACTIONABLE INSIGHTS IN SAS VISUAL ANALYTICS

Ultimately, the aim is to achieve a deeper understanding of customer behavior. The integrated analytic data in SAS Visual Analytics reveals relationships between the online and in-store channels. The following shows how the data can be visualized to answer the marketing manager's research needs, and how the information can be used to adapt the marketing strategy for more effective promotions.

ONLINE SESSION BEHAVIOR

This bar chart shows session duration across the customer base. An interaction has been set up in SAS Visual Analytics that allows for the drilling down of session classification to report the total page views for people that had spent that range of minutes browsing the website. Moreover, with the brushing technique, we spot the corresponding page view percentage of total.

We can use this data visualization to find:

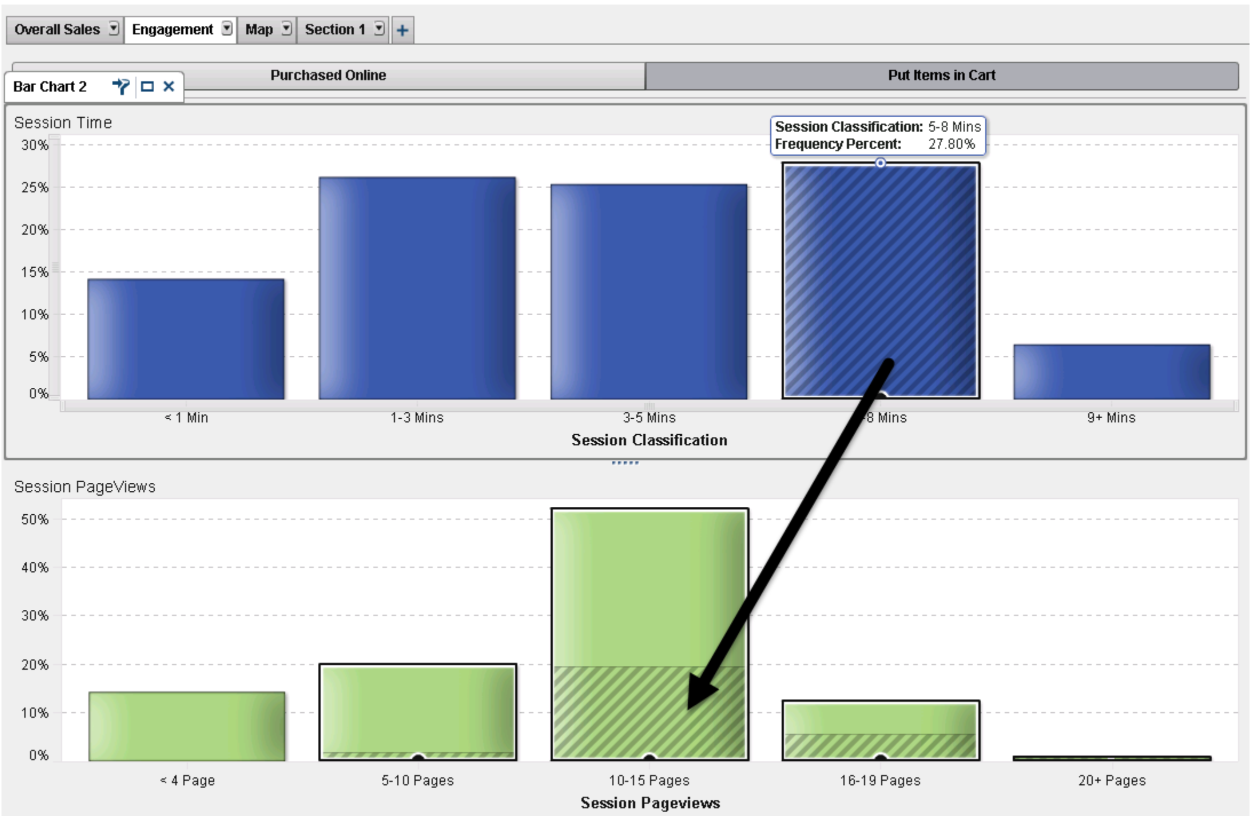
- The majority of loyalty customers that had placed an item in the cart spends 1 to 8 minutes on the site.
- Within each Session Time Classification, we can further drill down to see the total page view breakout.
- A selection bar on the top of the screen allows for quickly examining how this trend may differ from actual online purchasers versus those that had just put the items in the shopping cart.

The information does generally show significant website engagement across individuals who are putting items in the cart. It also reveals some potential marketing opportunities:

Around 14% of sessions end before 1 minute has elapsed. This might suggest that particular subgroup is using the cart as a bookmark or that the customer abandons after seeing the final price.

Potential Action: Customers with sessions that last less than a minute could be considered later for trigger emails or discounts to attempt to increase conversion.

Figure 4 Engagement Example



CHANNEL PURCHASES

This visualization shows the split between loyalty customers that are Brick and Mortar only versus the Multi-Channel Shoppers who have trackable online activity. The Multi-Channel Shoppers have been explored further by showing their channel purchase percentage of breakout across different departments.

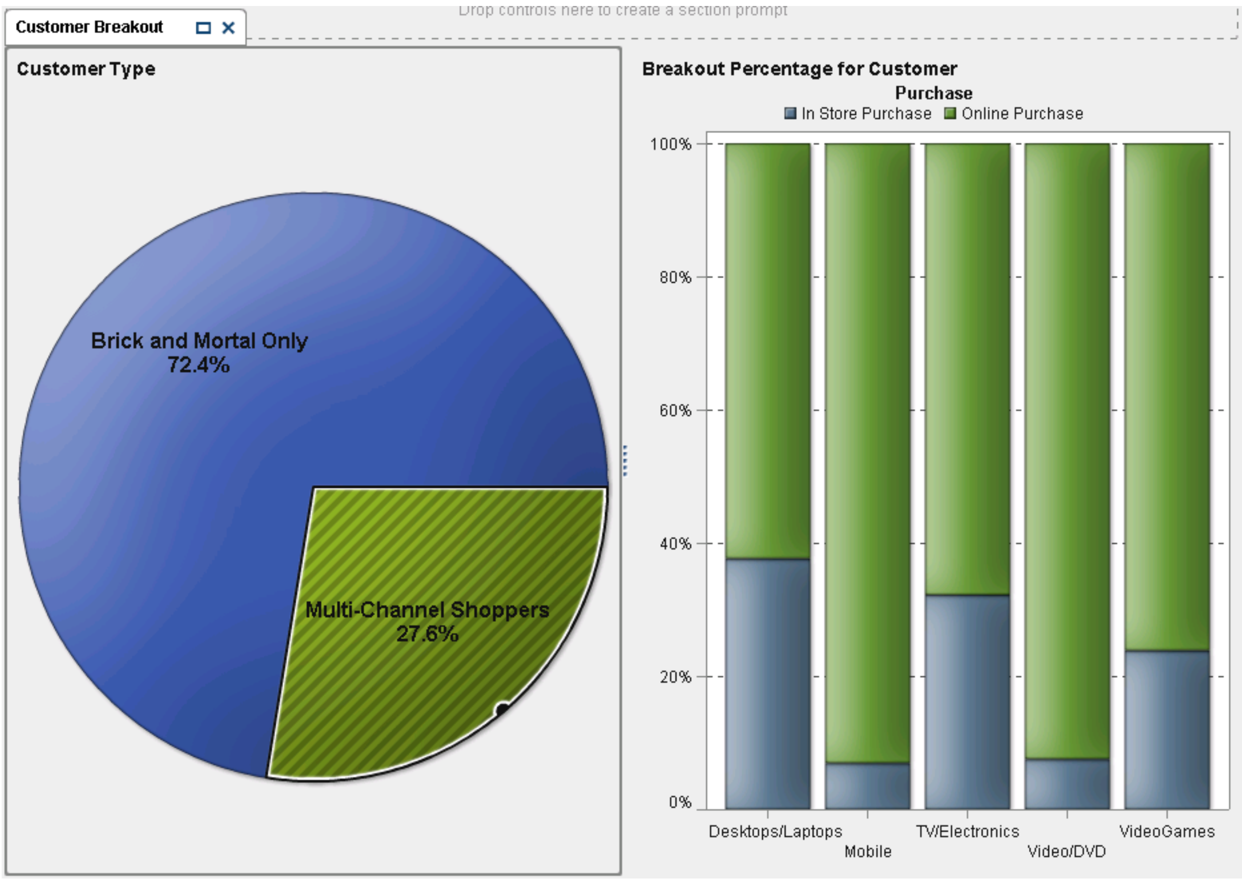
- Brick and Mortar customers only shop in the stores and consist of the majority of customers.
- Customers that have both in-store and online purchase behavior exhibit some variance across departments.
- Multi-Channel Shoppers tend to purchase the majority of their products through the online channel.

Many customers have no activity in the online channel and this could potentially be due to not requiring email information at registration. It could also be based on customer channel preference, as it does seem that customers that do exhibit online behavior tend to have the majority of their purchases be through the online channel.

Potential Actions:

- *Promote online registration for the significant 72% of Brick and Mortar customers to increase online channel awareness.*
- *For* Multi-Channel Shoppers*, a good number come in-store for higher ticket items like Desktops and TVs, perhaps to save on shipping costs. Customers with significant page views in these departments could be educated by the site on the store location and current in-store specials in order to drive further incremental in-store spend.*

Figure 5 Customer Type Channel Split Example



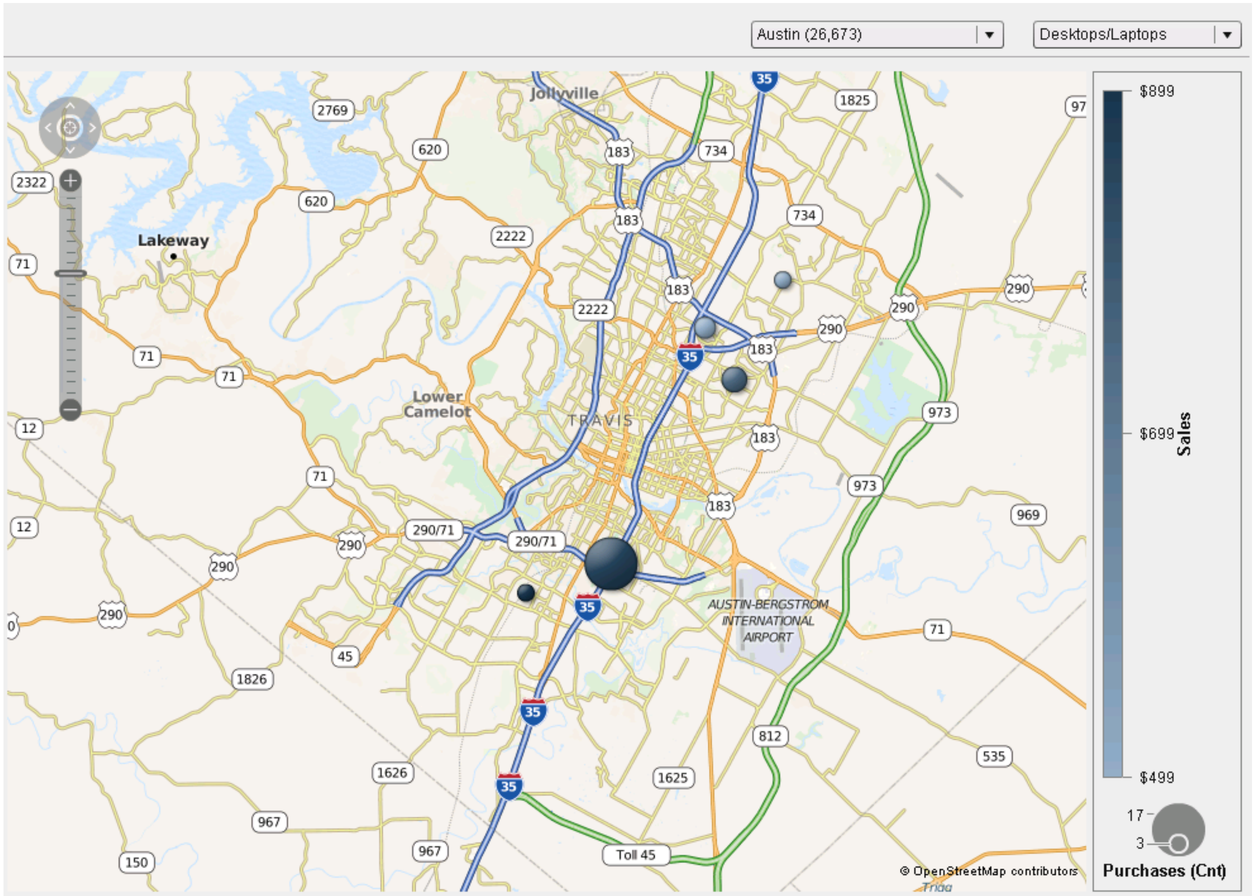
CUSTOMER LOCATION

- An interesting capability of SAS Visual Analytics is geospatial analysis. SAS Visual Analytics 7.1 introduced the capability to do postal code shading. The geographic location of customers along with their relative spend is reported below. The drop-down filters are used to switch between particular store numbers, as well as examine sales in a specific department from customers in the surrounding area.
- The mouse can hover over a bubble to obtain more details such as customer count and postal code. In addition, the color shading indicates the relative sales coming from a region's customers.
- The size of the bubble indicates the level of customers in that area based on their loyalty card registration.

High-value regions with significant population can be determined through a visual glance at this map. We can see there is a specific concentration of customers for Desktops/Laptops in the Austin area in the south portion.

Potential Actions: Having identified high-value regions, marketing efforts can be concentrated in those locations if those postal codes are close enough to physical store locations. In addition, smaller regions with high value customers can have sales expansion either through potential new store openings or through online promotions.

Figure 6 Geospatial Analysis



ABANDONED CARTS

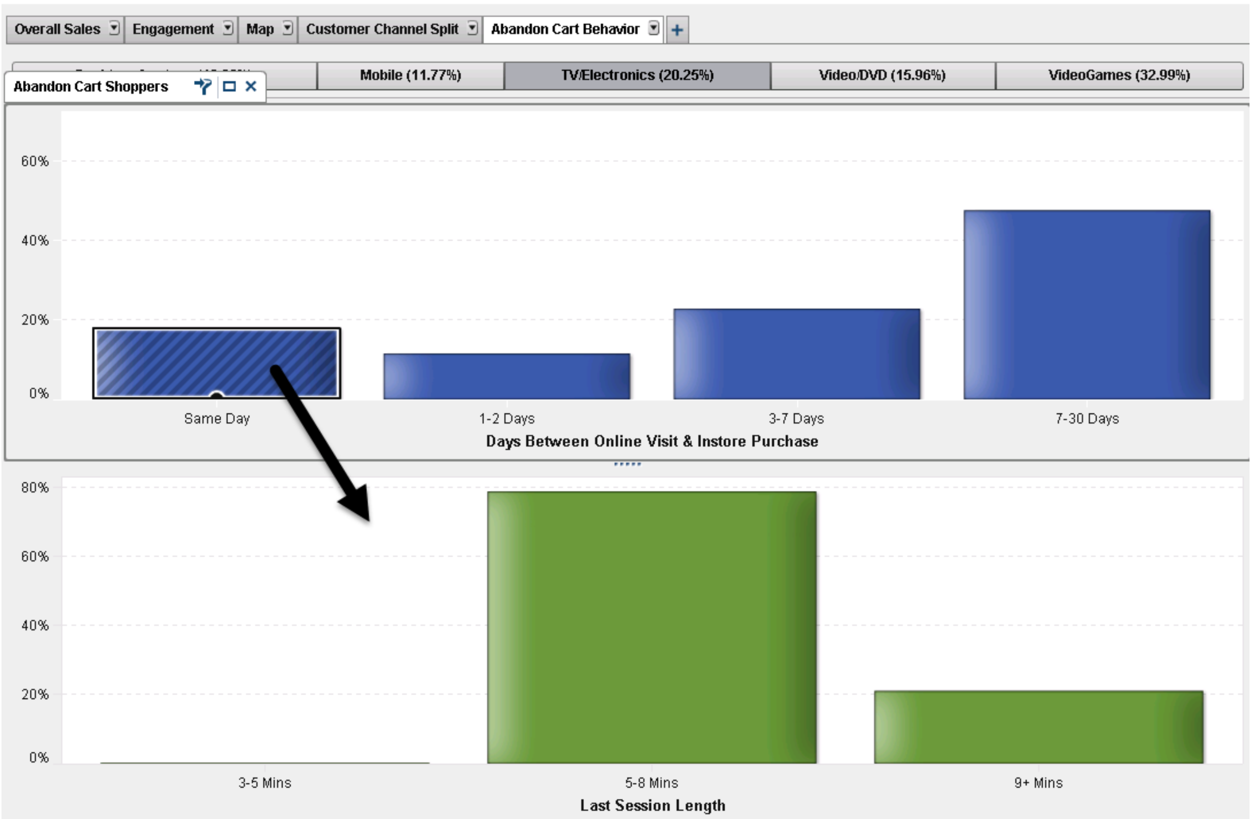
The web customers that abandon their carts but subsequently proceed to purchase in-store are shown in this visualization. An interaction links the days between the last online visit and the in-store purchase with the last session length. We find from this report that:

- Significant time is spent in the last session.
- There is a time lapse between the last visit and the purchase date, which can be for many days.
- The department selection bar allows for comparison to see if the abandonment patterns have significant differences cross-department.

Note that for the TV department, the customers spend significant amounts of time before abandonment. There is generally a major time delay before going in-store.

The purchase decision actually may have been finalized at this last visit, or the customer may have waited to do price comparisons with competitors or to save up enough income to make the purchase. As there is a considerable amount of time between the cart abandonment and the in-store purchase, *Potential Action: Email reminders and special offers can be used to drive customer's in-store sooner and protect from competitors.*

Figure 7 Abandon Cart Analysis



CONCLUSION

The integration of cross-channel data is a challenge, but understanding, classifying, and adapting to customer behavior is vital to the future of retail marketing. SAS Visual Analytics provides the tools needed to gain insights into marketing patterns quickly enough to have a viable impact on promotional decision-making.

REFERENCES

1 McKinsey Global Institute, " Big data: The next frontier for innovation, competition, and productivity" http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation#sthash.r97h78Ll.dpuf

2 BusinessWeek, "The Problem with Loyalty Programs" http://www.bloomberg.com/bw/stories/2007-05-10/the-problem-with-loyalty-programsbusinessweek-business-news-stock-market-and-financial-advice

3 "New Marketing Data: Back to the Future" https://www.rsrresearch.com/2013/09/10/new-marketing-data-back-to-the-future/

4 Bloomberg Business "Retailers Use Big Data to Turn You into a Big Spender" http://www.bloomberg.com/news/articles/2014-06-03/retailers-use-big-data-to-turn-you-into-a-big-spender

ACKNOWLEDGMENTS

The authors would like to acknowledge the following individual for assistance with preparing this paper:
Rosie Poultney, Ben Zenick and Bill Pena

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Johann Pasion
89 Degrees, Inc
pasionh@89degrees.com
http:/www.89degrees.com

Tricia Aanderud
Zencos Consulting, LLC
taanderud@zencos.com
http://www.zencos.com