



## HealthWise Coffee Market Research Report

**By:** Greg Peckham and Cody Frehr

**For:** Prof. Kaibo Liu and Abdallah Chehade

May 2, 2016

## **1 Introduction**

As an emerging online retailer in the niche market segment of low-acid coffee products, HealthWise stands to gain a significant competitive advantage with the assistance of data analytics. The recent integration of Google Analytics into their online marketplace has opened the door for large-scale data collection and analysis of business performance. In this report, we will present the findings of our research into HealthWise's product sales and customer base under the broader goal of extracting marketing insight.

## **2 Research Goals**

Management at HealthWise strongly desires to stimulate sales among existing customers and increase brand awareness abroad. In the early phases of this project, we worked closely with management to define the following research questions:

- Which day of the week, and times of day should specific email marketing campaigns be initiated to get the greatest amount of sales?
- Which products sell best in which regions of the country?
- What products are men versus women more likely to buy?
- What patterns exist that can identify a new visitor's coffee preferences in order to target them with a marketing offer?

## **3 Analysis and Findings**

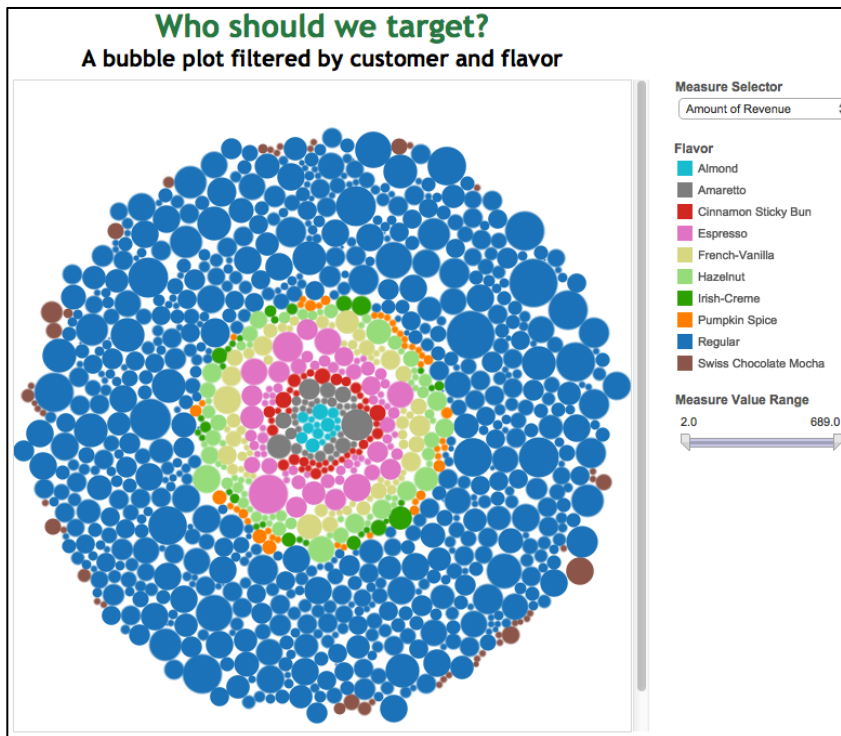
Although we set forth with a specific set of research questions to lead our investigation, we wanted to provide HealthWise with more than just answers to a very specific set of questions. Leaping into the world of data analytics for the first time can be a revitalizing experience for a business, and we think this is an opportunity to provide eye-opening visualizations of their marketing operations to hit our message home. In the following sections, we frame our analysis around four of the five Ws (Who, What, Where, and When) to add structure to our findings and to convey our message in a more coherent and comprehensible manner.

### 3.1 Who?

Who are HealthWise's customers and what are their product preferences? The bubble plot below is a multifaceted approach to understanding the customer base.

#### Key Figure Details:

- Each bubble represents a specific customer.
- Bubble size represents total revenue from a customer.
- Bubble color represents the flavor of coffee.



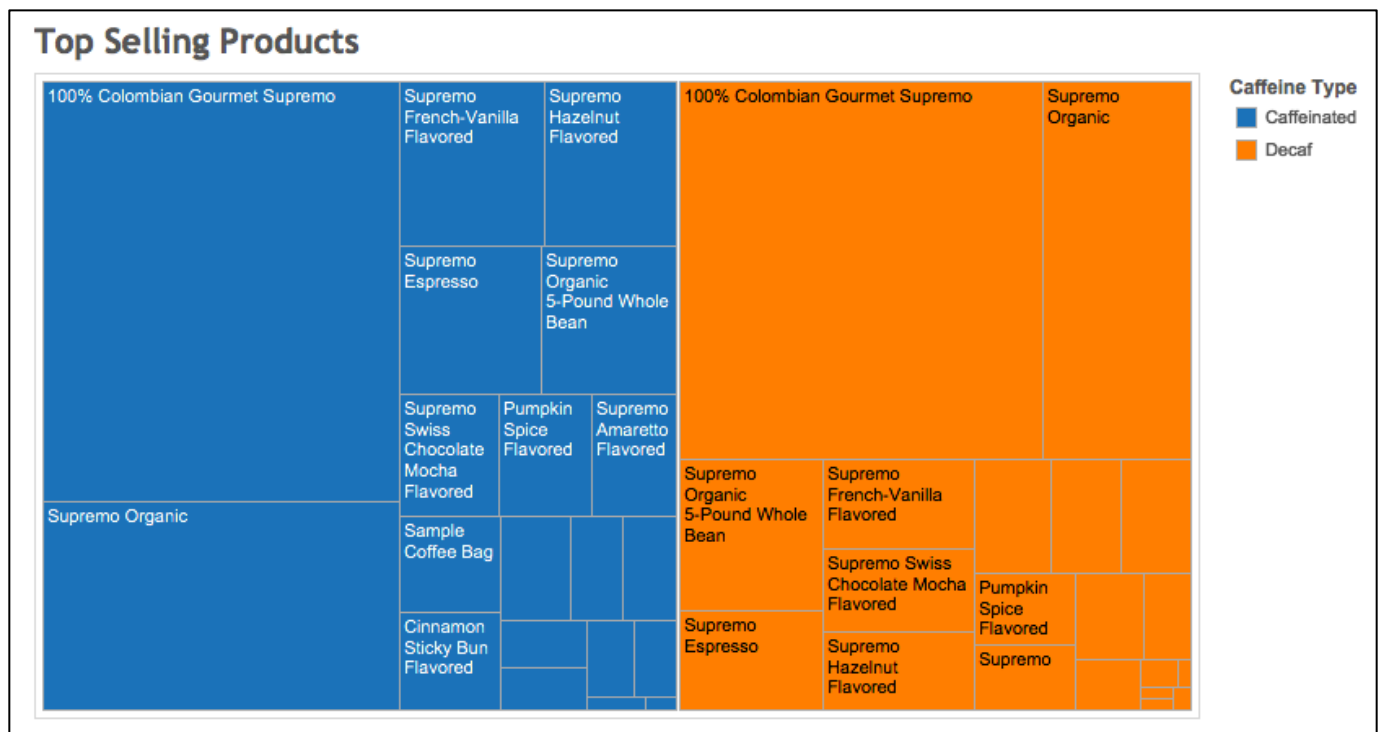
**Findings:** The great majority of HealthWise customers prefer regular flavored coffee, but there are a substantial number of customers who prefer the espresso flavor, as well; not only do more customers enjoy these flavors, but each customer tends to spend more money on these flavors. Customers do not enjoy the pumpkin spice flavored coffee. In addition to helping us understand the customer base as a whole, this bubble plot may serve as a useful tool for pinpointing HealthWise's most valuable customers. In selecting a measure range on the bubble plot, HealthWise can filter out the high spenders and target them with flavor-specific marketing offers to increase revenue. Conversely, HealthWise can target infrequent spenders with offers in an effort to strengthen customer loyalty and improve retention rates.

### 3.2 What?

What are HealthWise's best selling coffee products, and what are the distinguishing factors? The treemap figure below helped us to make insightful distinctions between products.

#### Key Figure Details:

- Each box represents a specific product.
- Box size represents total number of product sales.
- Color represents caffeine type.



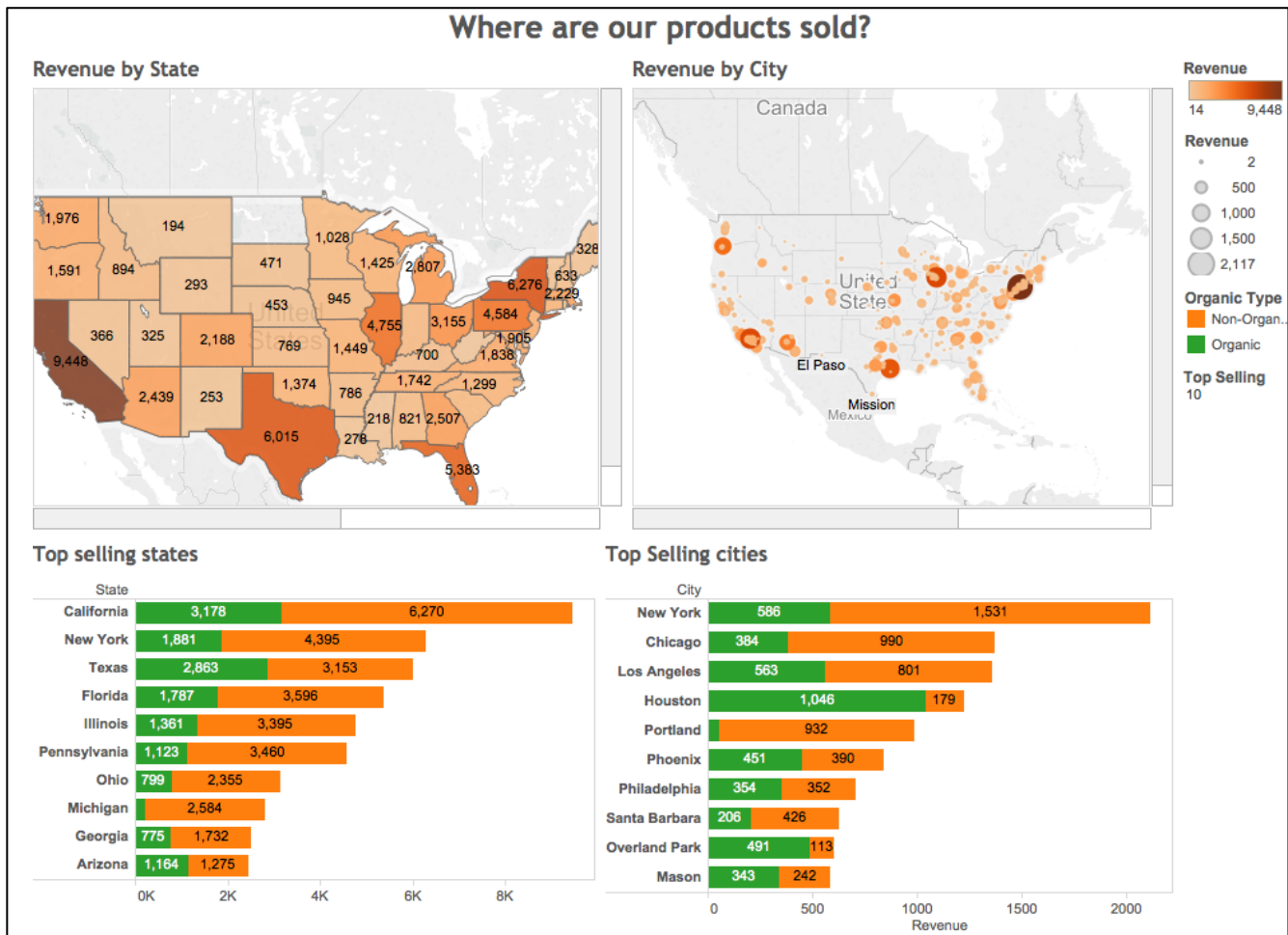
**Findings:** From this figure of top selling products by revenue, you can see that the 100% Colombian Supremo was the most popular product by far, and interestingly enough, the top 8 selling products all had the word 'Supremo' in it, although the 'Supremo' coffee itself is not significantly different in price or quality from other products. We think that this buzzword link could signify effective branding tactics by HealthWise, who may want to explore search optimization in future studies.

### 3.3 Where?

How do HealthWise's coffee sales vary regionally? Both geographic plots and bar charts are given below to illustrate sales across the county.

#### Key Figure Details:

- (Geographic Plots) Darker color indicates higher overall revenue from that source. Bubble size also represents total revenue.
- (Bar Charts) Length of bar represents total revenue. Color sorts by organic type.



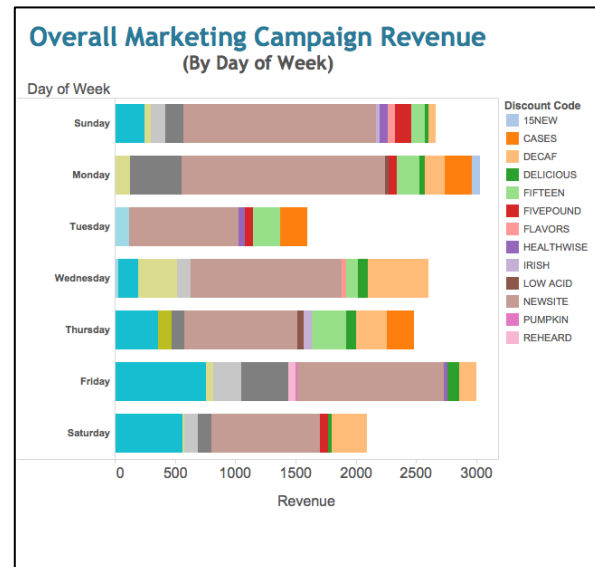
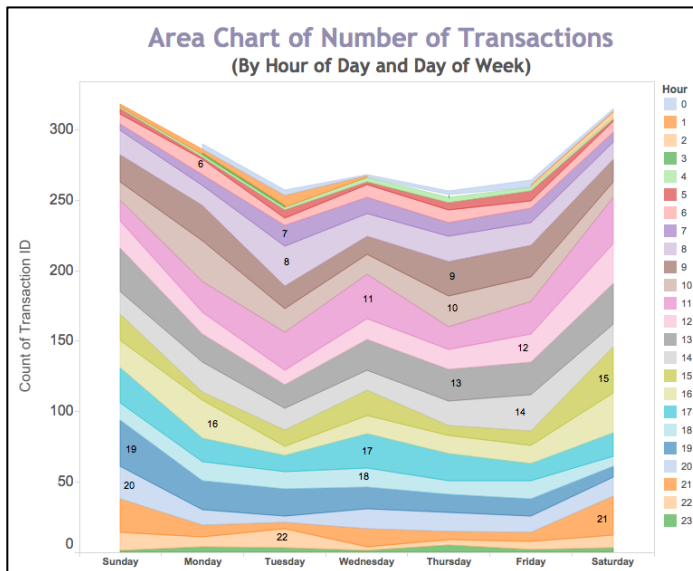
**Findings:** The dashboard above shows that HealthWise has sold coffee in 49-50 states in the US. They have been more successful in larger cities and states, while there does not seem to be a clear indication of organic preferences between states. Interestingly, the city of Portland has very low organic sales, which might go against one's intuition given the city's reputation. This is a strong example of why you cannot always follow your gut in marketing strategies.

### 3.4 When?

What time of day, and what day of week, does HealthWise get the most sales? What is the best day of week to launch a new marketing campaign?

#### Key Figure Details:

- (Area Chart) Days start on Sunday and proceed horizontally; hours start at midnight and proceed top to bottom.
- (Area Chart) Color represents an hourly range. The vertical thickness of an hourly range represents the total number of transactions occurring over that range.
- (Bar Chart) Length of bar represents total revenue. Color sorts by marketing campaign.



**Findings:** As shown by the total height of the hourly ranges in the area chart, Saturday, Sunday, and Monday have the greatest overall number of sales by a large margin.

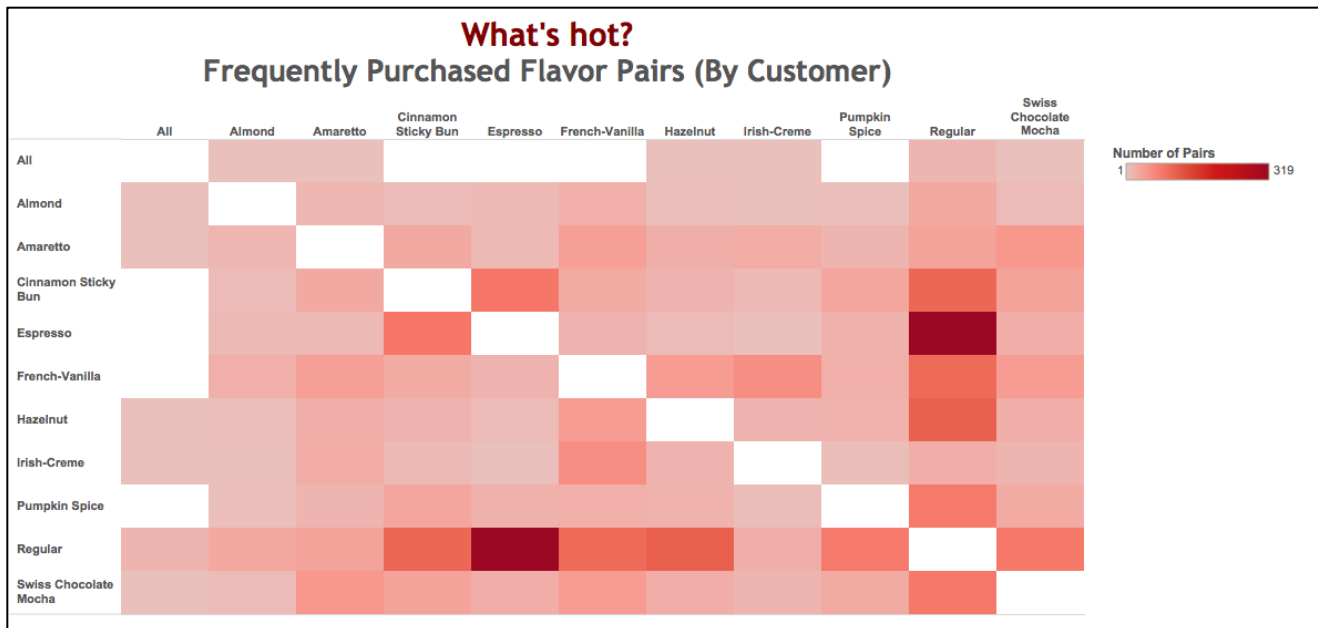
Interestingly, a large number of sales on these days occurs in the evening from between 7pm-10pm. Conversely, for weekdays Tuesday-Friday, peak sales occur between the hours of 8am-2pm; this seems somewhat expected as customers are likely tired out at work and have coffee on their minds. In the bar chart, we see that the greatest revenue from marketing campaigns comes on Mondays and Fridays. The best day to launch a campaign seems to be Friday; in comparison to the relative totals across the previous area chart, Friday seems to have a much greater proportion of its sales coming from marketing campaigns. Additionally, starting a campaign on Friday means that the three greatest days for overall sales will immediately follow, two of which are traditionally strong days for marketing revenue.

### 3.5 Purchasing Patterns

What patterns exist that can accurately predict a visitor's coffee preferences? Being able to predict a visitor's preferences can enable HealthWise to develop smarter marketing tactics.

#### Key Figure Details:

- (Heat Map) Darker color indicates a higher frequency of occurrence.

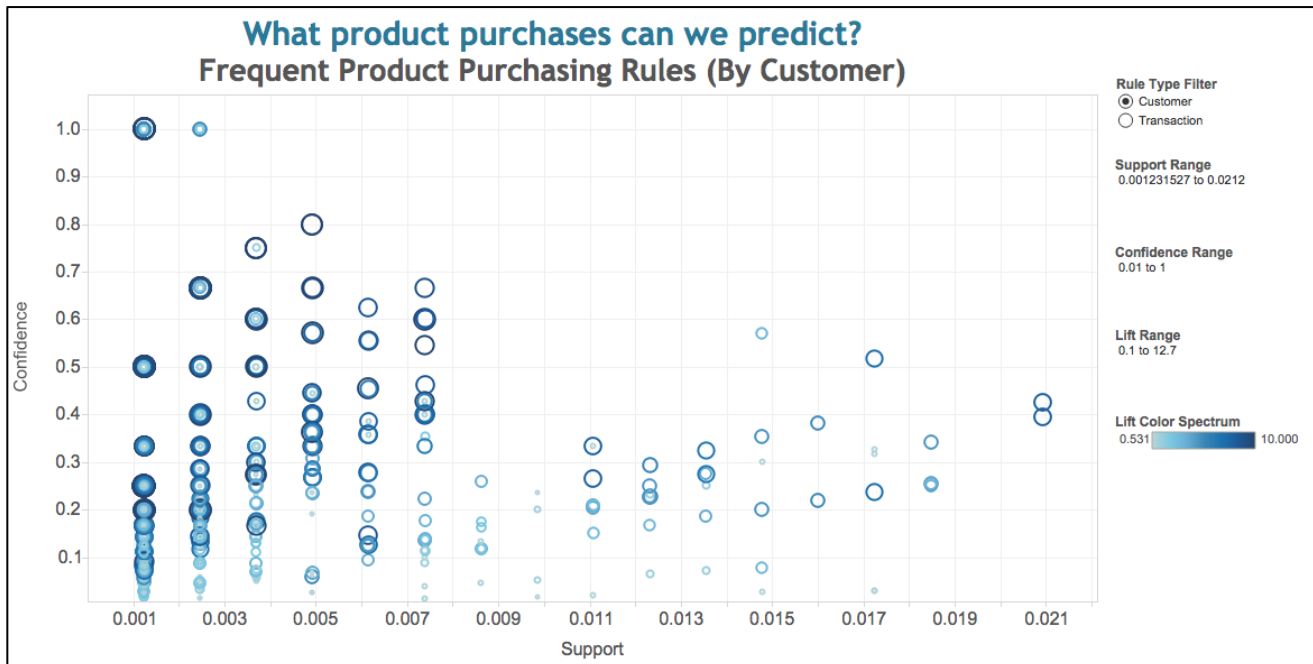


**Findings:** As part of our market basket analysis, we found that Regular coffee has a high frequency of occurrence with Espresso flavored coffee and several others. Additionally, French-Vanilla flavored coffee has a high occurrence rate with Irish-Crème flavored coffee. Frequent groupings, such as these, do not themselves imply causality in the purchasing of products. However, frequent sets are a prerequisite for developing those kinds of rules, which we explore on the next page. The top ten frequent pairs of flavors by customer are listed below:

#	Set	Support
8	{Espresso, Regular}	0.04802956
10	{Hazelnut, Regular}	0.04187192
12	{Regular, Swiss Chocolate Mocha}	0.03325123
13	{Pumpkin Spice, Regular}	0.03078818
14	{French-Vanilla, Regular}	0.03078818
15	{Cinnamon Sticky Bun, Regular}	0.02463054
16	{Cinnamon Sticky Bun, Pumpkin Spice}	0.02093596
18	{French-Vanilla, Swiss Chocolate Mocha}	0.01847291
19	{French-Vanilla, Hazelnut}	0.01847291
20	{French-Vanilla, Irish-Crème}	0.01724138

## Key Figure Details:

- (Scatter Plot) Each bubble represents an association rule between products with the specified support and confidence. A darker color indicates a higher measure of lift.



**Findings:** Although a discovered purchasing rule may have a high frequency and confidence level, the rule itself is not interesting unless the measure of lift is greater than or equal to 1. A higher lift reflects a higher dependency between the frequent items, and therefore demonstrates that the rule was not a consequence of pure chance. For this example, note that we chose to explore rules between specific items, rather than flavors; we picked several purchasing rules below that we found interesting due to their high support, confidence, and lift:

#	Association Rule	Support	Confidence	Lift
1292	{Supremo Irish-Creme Flavored} => {Supremo French-Vanilla Flavored}	0.017241379	0.51851852	7.13622097
1245	{Cinnamon Sticky Bun Flavored} => {Pumpkin Spice Flavored}	0.020935961	0.42500000	8.02558140
1572	{Organic Espresso} => {Supremo Organic}	0.014778325	0.57142857	3.03267974

There are many more rules we found that stuck out, but chose these few for their high lift, support, and confidence levels. Rule 1572, for example, states that Organic Espresso and Supremo Organic occur together in a customer's purchases approximately 1.5% of the time. HealthWise can then predict with 57% confidence that a customer who purchases Organic Espresso will also at some point purchase Supremo Organic. Say hello to predictive analytics!