APPLICATIONS FOR MACHINE LEARNING IN SHOPPER MEASUREMENT AND TRACKING

Gary Angel
CEO
Digital Mortar

Machine Learning and Shopper Analytics

The experience inside is a mystery

The store has been a black-box

We measure how much we sold.

We count how many customers go in

New Collection Technologies have made the in-store journey MEASURABLE

Wi-Fi

Data gathered by in-store Wi-Fi access points track Wi-Fi emitting device movements within a networked space.

Phone

Mobile Apps and Bluetooth provide accurate, detailed journey measurement that can be integrated.

Video Camera

Multiple camera images, angles and lenses are used to stich together movement patterns, activity and demographics.

Where do customers go?

What engages and when does that sell?

Are areas of the store underperforming?

How much localization is necessary?

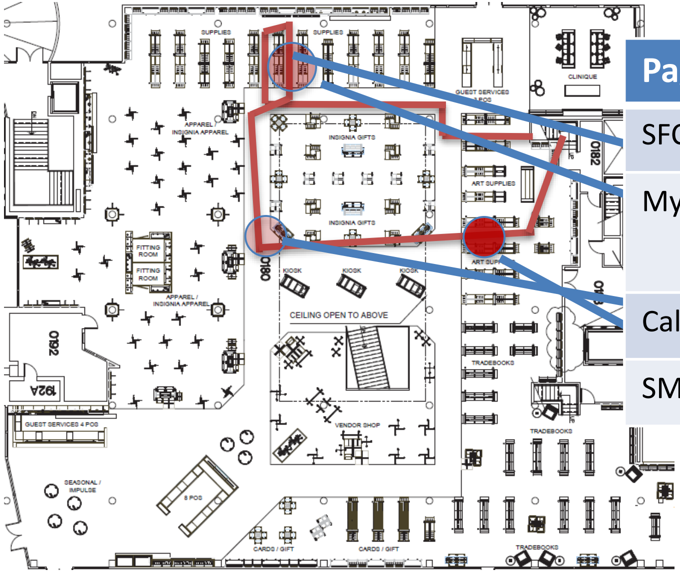
Is the staffing level appropriate?

Who are our best performers?

Measuring Physical Spaces

[illegible] where they spent their time…

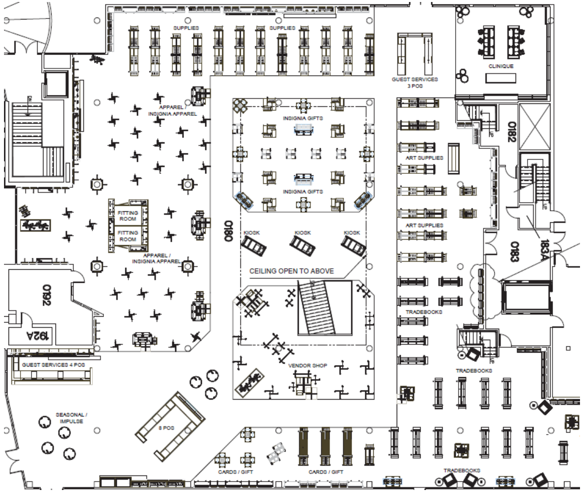
Imagine this was a Website



Page	Time
SFCat	5.35
MysteryCat	2.32
CalCat	1.59
SMagCat	6.12

Opening up a lot of Customer Experience and Omni-Channel Analysis

Web	Store
Path Analysis	Path Analysis
Functional Analysis	Draws, Corridors, Impulse
Funnel Abandonment	In-Store Funnel
Content Attribution	Staff Performance
Associated Content	View-based Market Basket



But there are some BIG Challenges

Data Quality	Analytics
Identifying Associates	Shopper Type Identification
Zone Stitching	Optimal Store Path

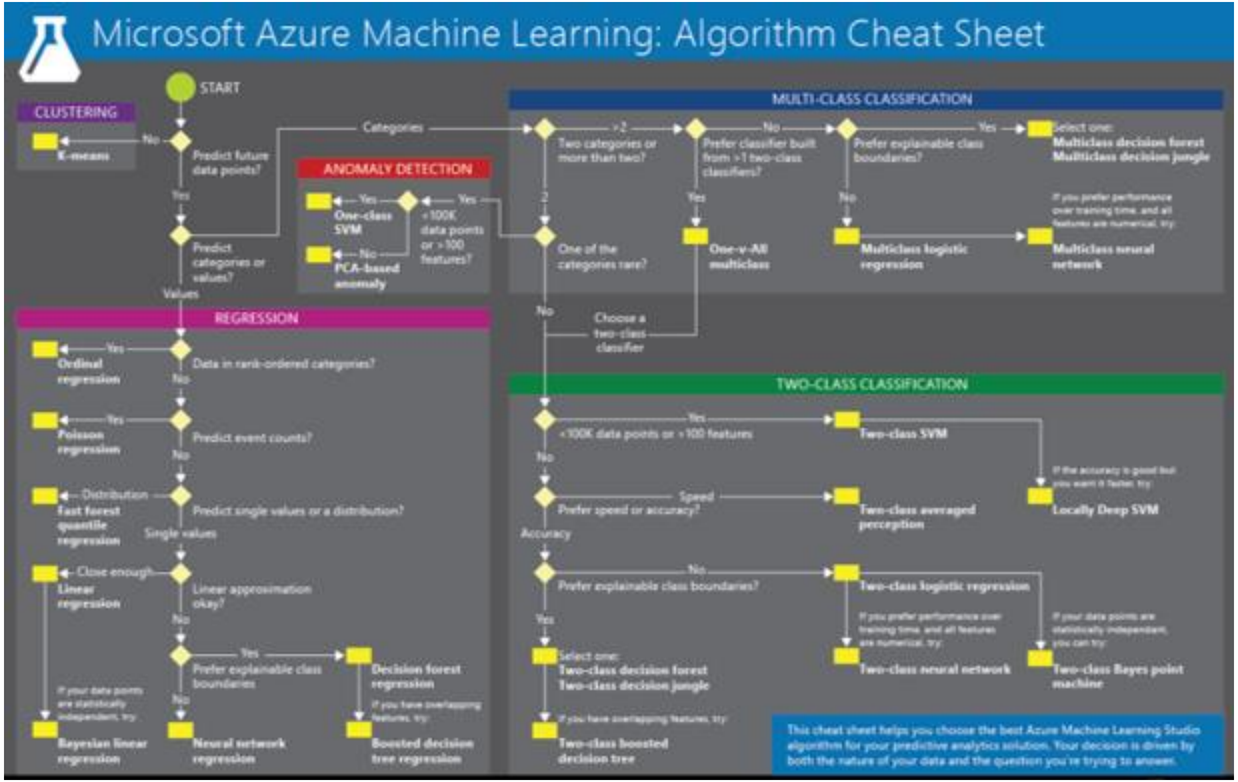
OK – there are tons of problems – but these are the ones we've targeted as ML appropriate…

What makes a good ML Problem

Lots of data : We have 100’s of thousands of events

Too hard to solve with simple rule-based processing : shopper paths are much too complex to if-then model

Oh yeah – and can be solved with ML



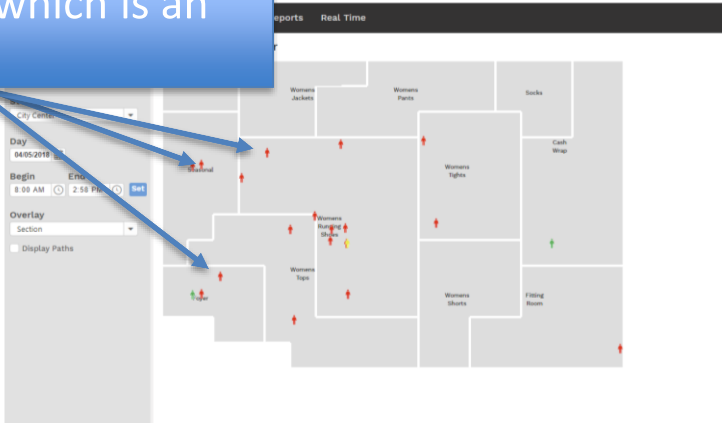
There's an appropriate ML technique that might work

Identifying Associates

Which is a shopper and which is an Associate?

Overview

We get data from electronics but we don't know which streams come from Associates and which come from shoppers. We want a way to use their in-store behavior to tell which is which.



Data

Avg. Time in Store

Visits this week

Staff Locations Fuzzy

Staff Locations Definite

Time at Cashwrap

Zone Stitching

Overview

Shoppers cross from one camera zone to another. Each zone overlaps for about 3 feet. We need to track shoppers from one zone to the other using the overlapping zone area and the movement vectors. This is simple in an uncrowded store but quite tricky when the zone is crowded.

Data

Shopper overlap zone x

Shopper Time Entered Zone Camera A

Shopper X Vector Camera A

Shopper Y Vector Camera A

Shopper Time Entered Zone Camera B

Shopper X Vector Camera B

Shopper Y Vector Camera B

Shopper Demographics Camera A

Shopper Demographics Camera B

Shopper Type Identification

Overview

There are a common set of store behavioral patterns that exist in most retail situations. These evince themselves in navigation patterns in the store – but those patterns are quite complex and varied. We want to be able to identify the patterns across many different types of stores. Sample patterns include Clearance Shopper, Dip-in, Right-Rail Shopper, Single Section Focused, Product Returner, etc.

Multi-Section Buyers

Right Rail Shoppers

Customer Servicing

Omni-Channel

In-n-Out Shoppers

Section Focused

Data

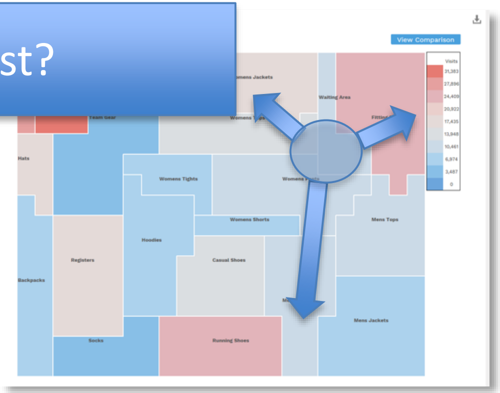
First Zone Type
Time at First Stop
Time Entering Store (HH)
Longest Stop
Time Inside Store
Clearance Zone Stop #
# of Linger Points
Clearance Zone % of Time
First Navigation Zone Direction

Optimal Store Path

What path works best?

Overview

We want a way to evaluate what paths and behaviors in the store are most associated with shopper success. This includes which sections are drivers of conversion, the role of associates, and general shopping behaviors like time and store navigation.



Data

Sections Visited

Time in Sections

CashWrap (Checkout)

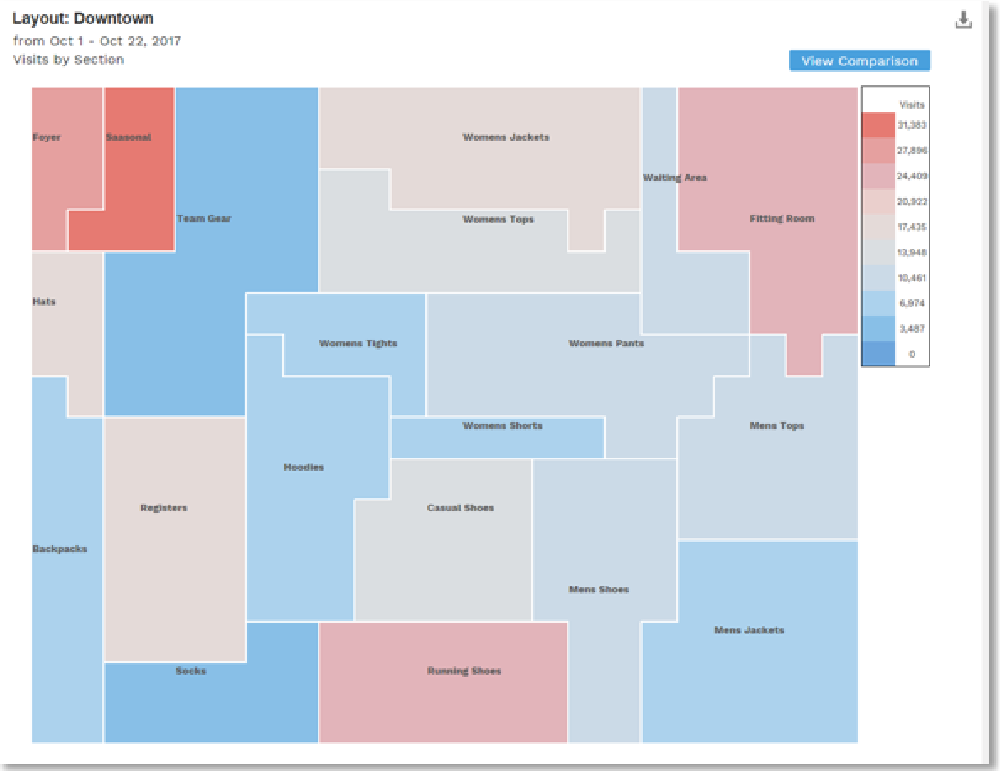
Total Time in Store

Day of Week

Time of Day

Associate Interaction Count

ML Case Study

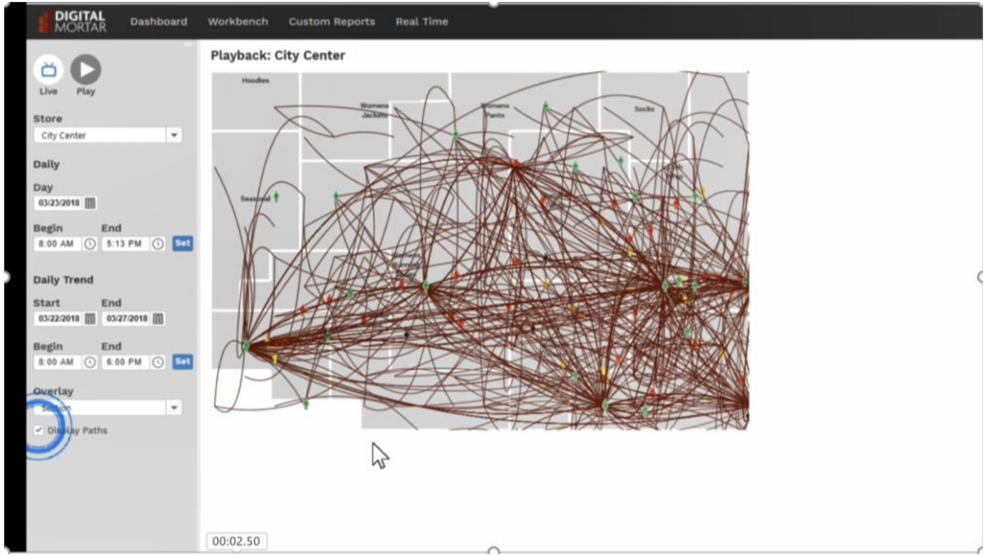


Use ML to determine whether the store layout and Associate placement strategies are optimal.

DXi is a SaaS based solution that enables digital marketers to increase Sales and ROI by leveraging the power of Machine Learning. DXi scores each customer and visitor on a scale of 0-100 and through this score it helps to identify the top customers and prospects as well as the factors that drive conversion.

Which Paths Worked Best

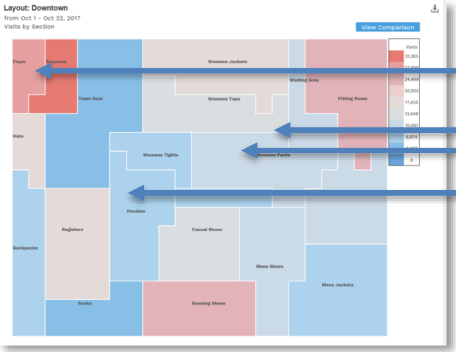
- Athletic Apparel Store with 25+ departments/sections
- Data consisting of over 25,000 unique visitors
- Data modeled to time of visit, sections visited and associate assistance
- Desired Outcome = “Conversion” = Visitor to Buyer



Analytics Process 1

The raw data is a stream of visit/location events:

VisitID	StoreID	PlanoID	VisitStartTime	MeasurementType	OffsetofPurchaseFrame	OffsetofEngageFrame	FrameTime	XGrid	YGrid	ZGrid	AssociateID	ConversionFrameFlag
4.2017E+34	2	42	20161201:09:10:0	storeTracking	-1	-1:02	20161201:09:10	8	13	0		0
4.2017E+34	2	47	20161201:09:09:1	storeTracking	-1	-1:17	20161201:09:09	17	10	010010		0
4.2017E+34	2	47	20161201:09:09:1	storeTracking	-1	-1:15	20161201:09:11	20	9	0		0
4.2017E+34	2	47	20161201:09:09:1	storeTracking	-1	-1:43	20161201:09:12	3	2	0		0

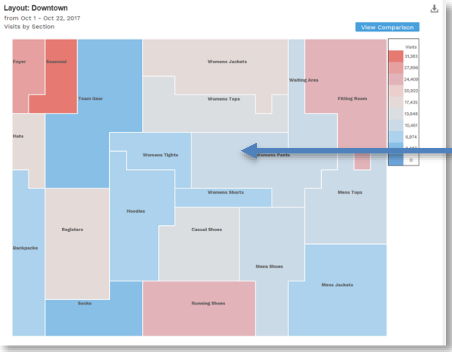


These get mapped to the store layout so we know where people shopped

Analytics Process 2

By tracking associates we know were interactions happen:

VisitID	StoreID	PlanoID	VisitStartTime	MeasurementType	OffsetofPurchaseFrame	OffsetofEngageFrame	FrameTime	XGrid	YGrid	ZGrid	AssociateID	ConversionFrameFlag
4.2017E+34	2	42	20161201:09:10:0	storeTracking	-1	-1	20161201:09:10:02	8	13	0		0
4.2017E+34	2	47	20161201:09:09:1	storeTracking	-1	-1	20161201:09:09:17	17	10	0	10010	0
4.2017E+34	2	47	20161201:09:09:1	storeTracking	-1	-1	20161201:09:11:15	20	9	0		0
4.2017E+34	2	47	20161201:09:09:1	storeTracking	-1	-1	20161201:09:12:43	3	2	0		0

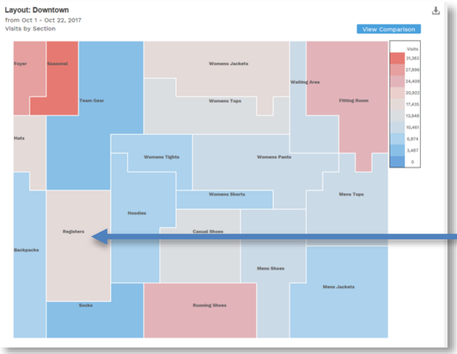


Those get mapped by Section and Time

Analytics Process 3

Time at CashWrap gives us the Purchase Flag:

VisitID	StoreID	PlanoID	VisitStartTime	Measurement Type	Purchase Frame	Engage Frame	FrameTime	XGrid	YGrid	ZGrid	Associate ID
4.2017E+34	2	4	20161201:09:21:11	storeTracking	5	-1	20161201:09:21:11	4	14	0	
4.2017E+34	2	4	20161201:09:21:11	storeTracking	5	-1	20161201:09:22:10	5	14	0	
4.2017E+34	2	4	20161201:09:21:11	storeTracking	5	-1	20161201:09:24:05	6	15	0	
4.2017E+34	2	4	20161201:09:21:11	storeTracking	5	-1	20161201:09:26:46	6	15	0	
4.2017E+34	2	4	20161201:09:21:11	storeTracking	5	-1	20161201:09:29:10	5	14	0	



This tells us when a shopper succeeded

Analytics Process 4

We aggregate this up to a single Row per visitor:

Interaction Data

Conversion Flag

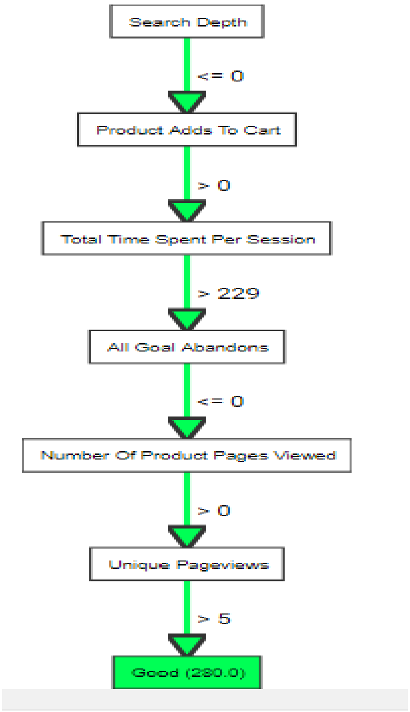
Time in Section

VisitID	4.2E+20	Mens Shirts	0
StartTime	20180101: 09:10:02	Mens Shoes	0
Visit Duration	95	Womens Running	0
InteractionCount	0	Womens Shoes	0
FirstInteractionID	0	Womens Tops	0
First InteractionTime	0	Foyer Time	0
Last Interaction ID	0	Backpacks Tim	0
Last Interaction Time	0	Casual Shoes Time	0
Purchase Flag	0	Fitting Room Time	0
Foyer	1	Hats Time	35
Backpacks	0	Hoodies Time	60
Casual Shoes	0	Men's Running Time	0
Fitting Room	0	Mens Shirts Time	0
Hats	1	Mens Shoes Time	0
Hoodies	1	Womens Running Time	0
Men's Running	0	Womens Shoes Time	0
		Womens Tops Time	0

Analytics Process 5

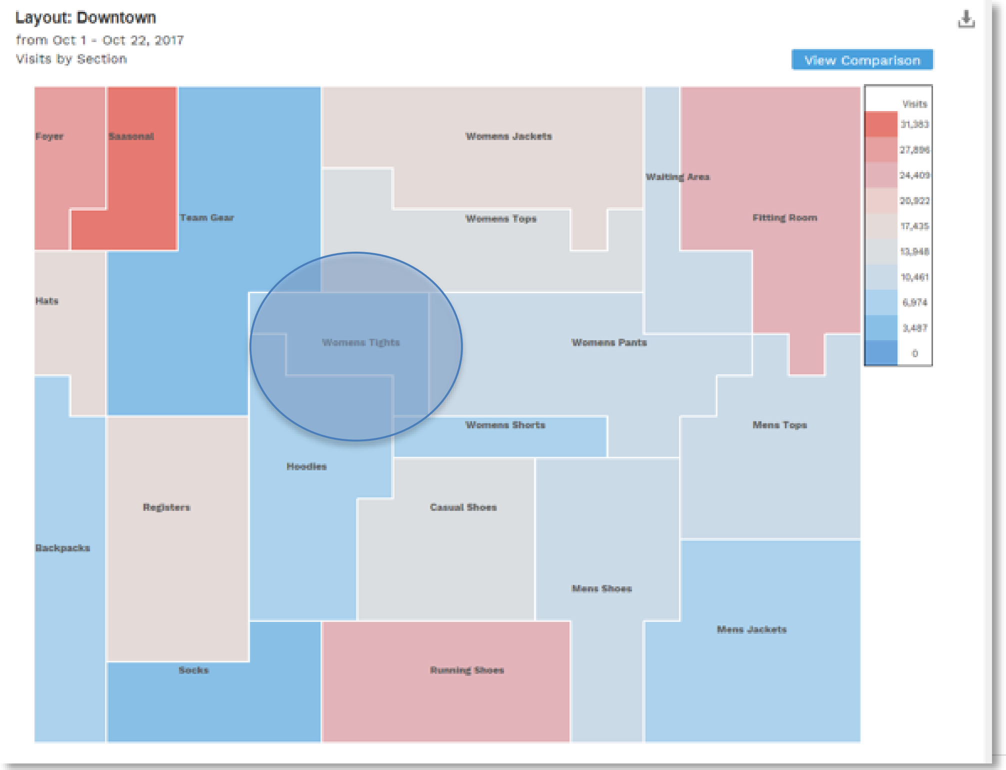
DXi then processes it to find the key behaviors that drive shopper conversion:

- Data Aggregated by unique visitor and department sections
- Data cleansed and outliers removed
- Data normalized to a common standard range
- Data fed simultaneously into 25 ML algorithms
- Results analyzed and inferred for actionable insights



Results 1

Women's Tights is a navigational problem:



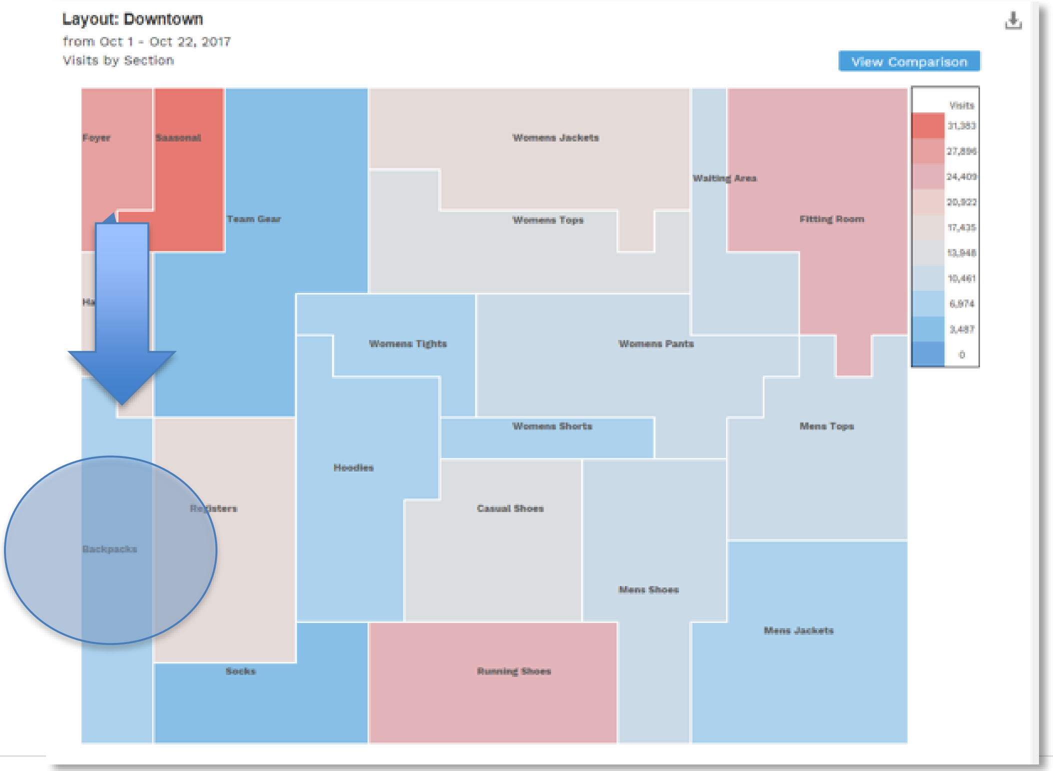
It is the poorest performing section in supporting or driving conversion.

Not a single DXi ML algorithm picked it as important.

Yet it has a central position in the store.

Results 2

This common Right Rail shopper pattern is very effective:

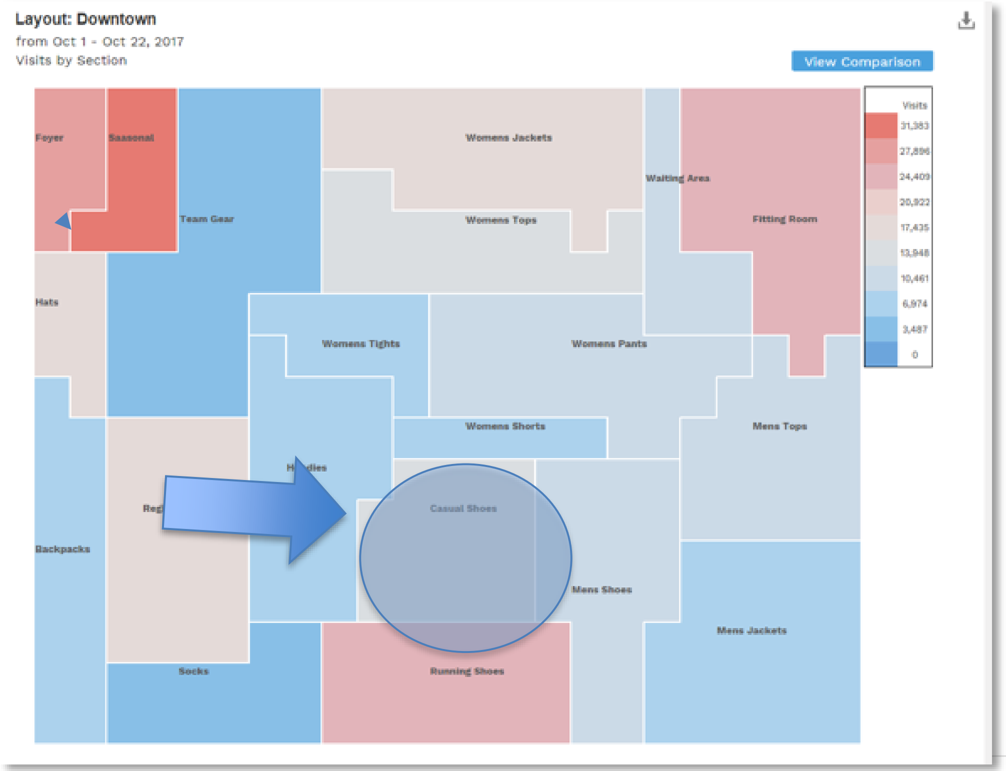


The path through Backpacks was one of the best performing store patterns in the DXi findings.

The "Right-Rail" behavioral pattern is common for shoppers without a specific destination.

Results 3

The Casual Shoes section is key to shopper performance:

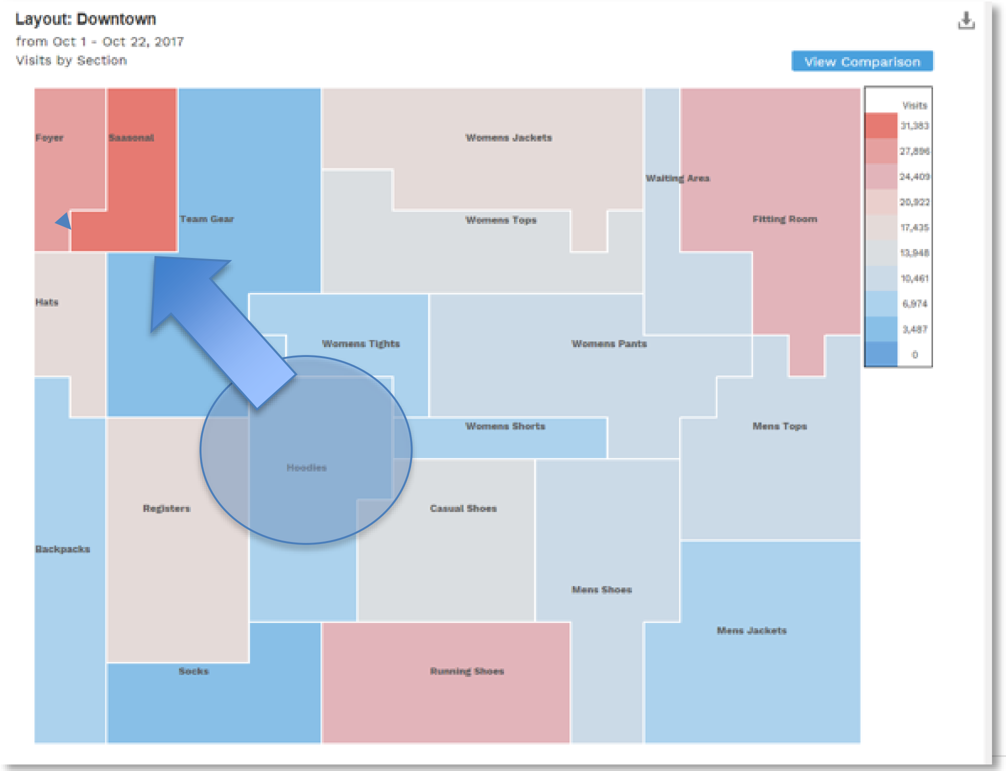


It was picked as important by every DXi ML algorithm.

It had the largest optimal time value (more time = better) of any section in the store.

Results 4

The Hoodies sections placement near Cash-Wrap is problematic:



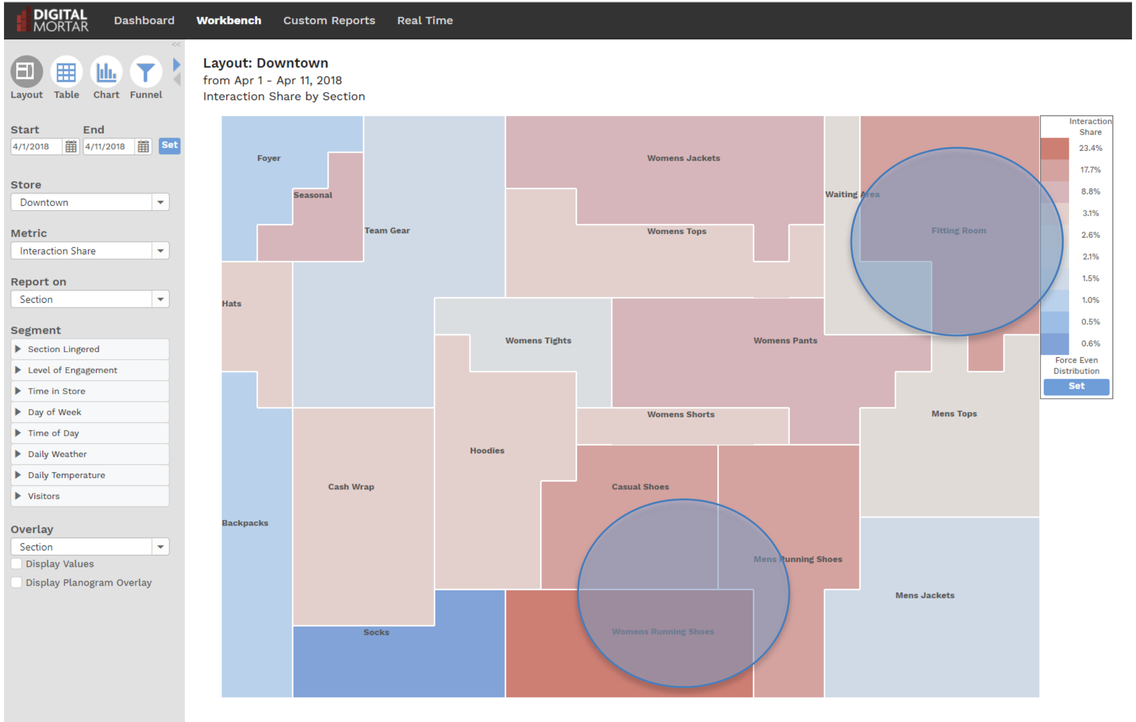
Shoppers who spent time here were less likely to convert.

Though it performed well in some DXi ML algos – we infer that this is because it's on the path to conversion.

That space near cash-wrap is super-valuable in stores and needs to be maximized.

Results 5

You can’t have too much interaction with Associates:



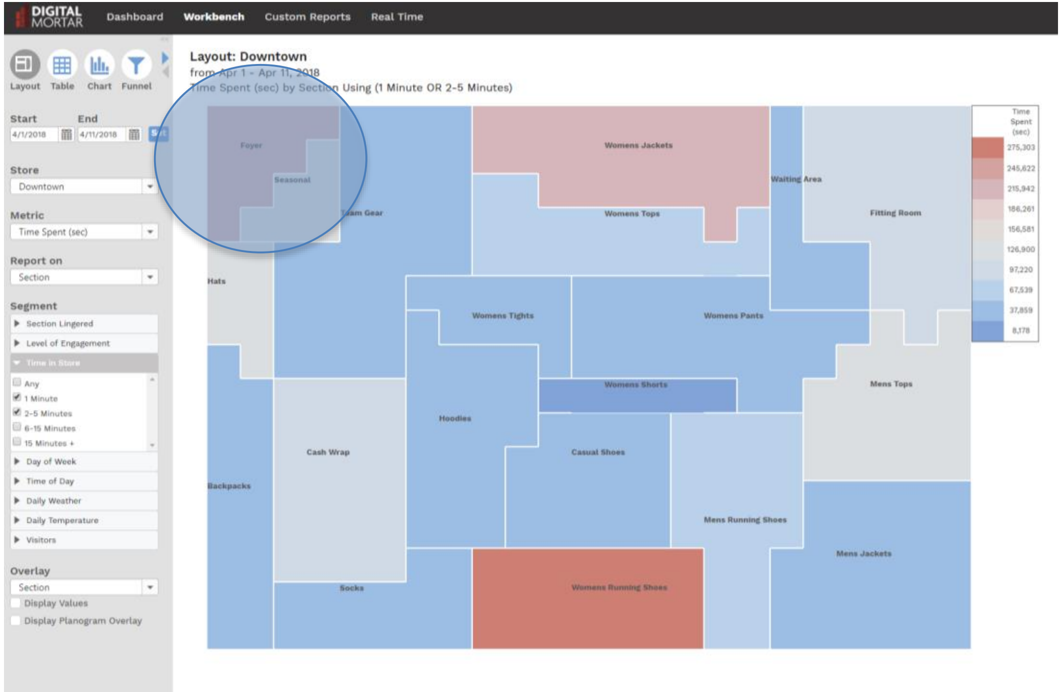
Every DXi ML Algorithm picked Associate Interaction as important.

More than 2 interactions was optimal.

Even 4-5 interactions had positive impact.

Results 6

And some general findings:



Long times at Cash Wrap reduce visitor success.

Linger times near the entrance are indicators of a low-qualified shopper.

Wrap Up

The Keys to the Car

- Measurement of the In-Store Journey is a significant opportunity for anyone with a physical space.
- Measurement technologies exist that can provide excellent shopper journey tracking down to a fairly small area of the store.
- Those technologies spin out data that is remarkably similar to what is produced in the digital world for digital analytics solutions to measure Web usage.
- That data opens up several areas where ML can be used to drive better data and better analytics.
- Those problems include Associate Identification, Zone Stitching, Shopper Type identification, and Optimal Store Path Analysis.
- We used DXi's multiple ML approaches to tackle the Optimal Store Path Analysis to understand whether the layout of the store is optimal and where opportunities to improve might exist.

Questions and Discussion

Gary.Angel@digitalmortar.com

Check out my presentation on ML Process Re-engineering on AnalyticsNexus
Check out my blog series on ML at http://digitalmortar.com/blog