

SENSE
CORP

alteryx

Alteryx Twin Cities User Group

April 2019

Hosted by Sense Corp



TWIN CITIES ALTERYX USER GROUP: MEET YOUR 2019 LEADERSHIP TEAM



Ben

TransPerfect



Garth

Daugherty



Nick

Slalom



Pedro

Haystack



Tessa

Quest Diagnostics

ALTERYX
INSPIRE
YOU • AMPLIFIED • 2019

REGISTRATION IS OPEN

NASHVILLE
JUNE 10-14

[ALTERYX.COM/INSPIRE-2019](https://www.alteryx.com/inspire-2019)



INDUSTRY/ DEPARTMENTAL USER GROUPS

Q2 USER GROUP MEETINGS

Wednesday (6/12) 7:00-8:00am

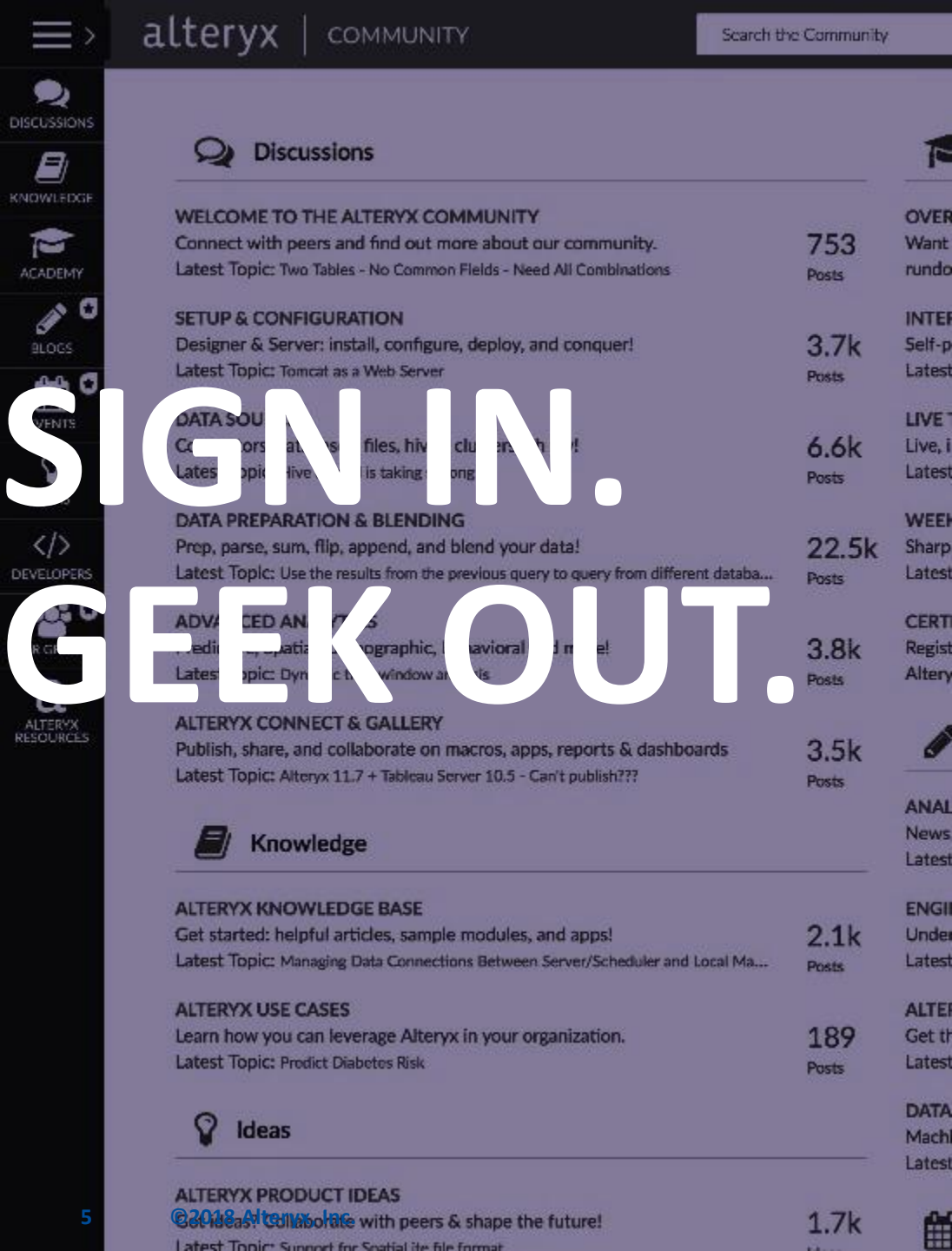
- Transportation
- Healthcare

Thursday (6/13) 7:00-8:00am

- Finance
- Marketing

**Check the inspire agenda for room information*

*Contact for more details:
cshernaman@alteryx.com*



SIGN IN.
GEEK OUT.



COMMUNITY.ALTERYX.COM

- Find answers quickly
- Share your knowledge
- Connect with peers
- Discover new Alteryx use cases
- Submit your bright ideas
- Read insightful blogs
- Develop tools, macros, and apps

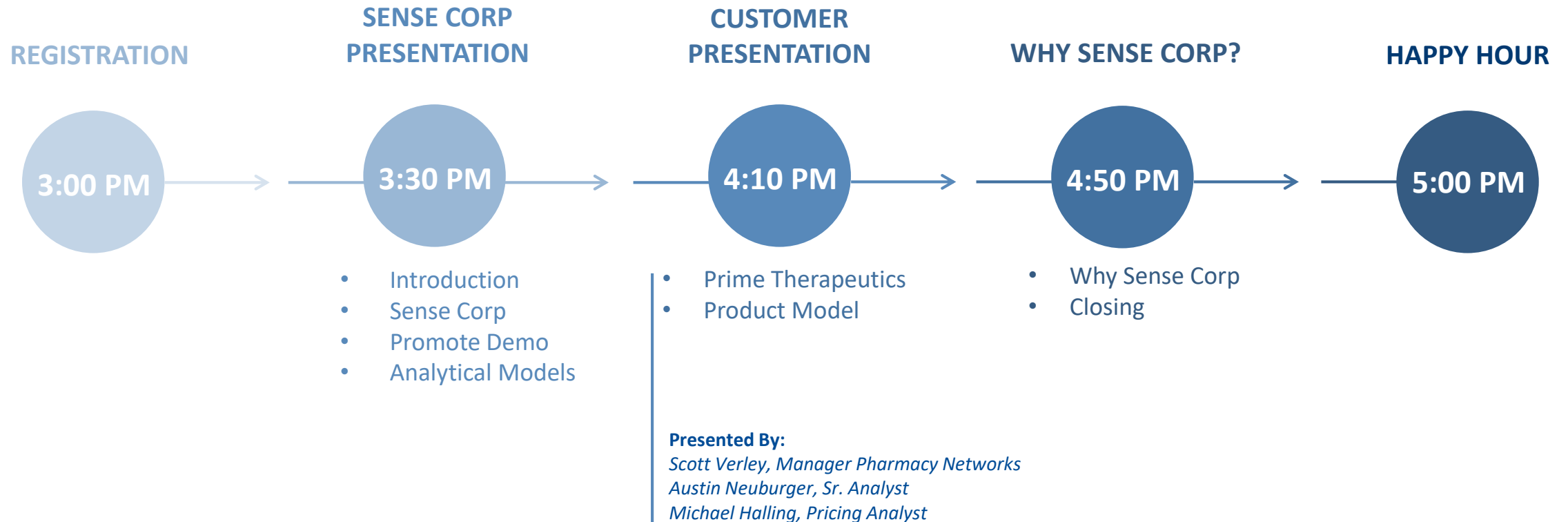
MEETING ROUND UP IS LIVE!



- Catch up on missed meetings!
- Share content with colleagues!
- Discover what other User Groups are up to!

The screenshot displays the Alteryx Community homepage. At the top, there's a navigation bar with the Alteryx logo, a search bar, and a user profile for 'LaurenU'. Below the navigation bar is a banner with the text 'ALTERYX YOUR POTENTIAL' and a sub-header 'Connect with a community of data science and analytic experts to discover new solutions to complex challenges, contribute fresh ideas, and gain valuable perspective on the rapidly developing industry.' An announcement banner follows, stating 'Announcement | We're thrilled to announce our newest Alteryx Designer Expert Certified users! Learn more.' The main content area is divided into several sections: 'DESIGNER' (15.3K Posts | 7.7K Solutions), 'SERVER' (937 Posts | 304 Solutions), 'CONNECT' (60 Posts | 11 Solutions), and 'PROMOTE' (23 Posts | 1 Solution). Below these are 'ACADEMY' (294 Posts) and 'BLOGS' (1.2K Posts). Further down are 'ALTER.NATION' (1.9K Posts) and 'CULTURE & EVENTS' (255 Posts). On the right side, there's a 'COMMUNITY ACTIVITY' box showing 118407 Posts, 68971 Stars, and 8757 Solutions. Below that are 'PARTNER COMMUNITY' and 'ALTERYX USE CASES'. At the bottom right, there's a 'LATEST PHOTOS' section with a grid of images.

Agenda



Introductions



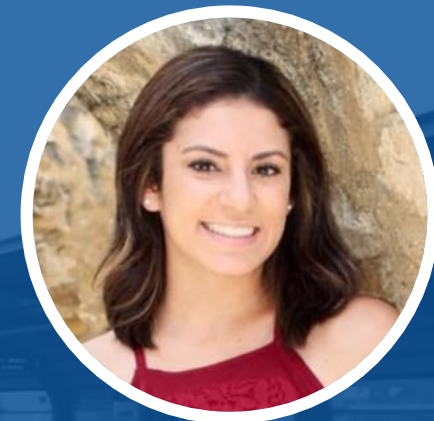
Tom Hymanson
Managing Director,
Minneapolis
THymanson@sensecorp.com



Tom Gonzalez
Analytics Director
TGonzalez@sensecorp.com



Kirk King
Alteryx Engineer
Kirk.King@sensecorp.com



Gaby Lio
Data Scientist
GLio@sensecorp.com

SENSE
CORP

Powering Insight-Driven Organizations

SENSE CORP

Powering Insight-Driven Organizations



Enterprise Data Strategy



Data Visualization



Data Governance



Advanced Analytics & Data Science



MDM



Big Data



Data Engineering



Data Security

... we turn data into actionable insights and transform organizations for the digital era.

Sense Corp has delivered exceptional results to hundreds of the world's largest global organizations and complex government agencies.

RETAIL	PUBLIC SECTOR	TELECOM/MEDIA	HEALTHCARE	FINANCIAL SERVICES	ENERGY	OTHER
   NORDSTROM    	      	  COMCAST    	     GEISINGER  	    MAKING SENSE OF INVESTING   	  bhpbilliton   HALLIBURTON  Shell  devon 	   TOYOTA    Panera BREAD 



Designer & Promote Demo

Gaby Lio and Kirk King

Some Basics

Level 1

Data Cleansing Tools



Level 2

Predictive Analytics Tools



Level 3

Developer Tools



Enterprise Solution

Level 1

Data Cleansing Tools



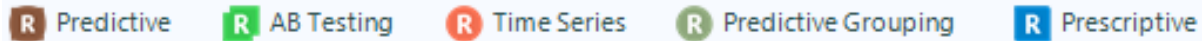
Analyst

Data Scientist

I.T.

Level 2

Predictive Analytics Tools



Level 3

Developer Tools



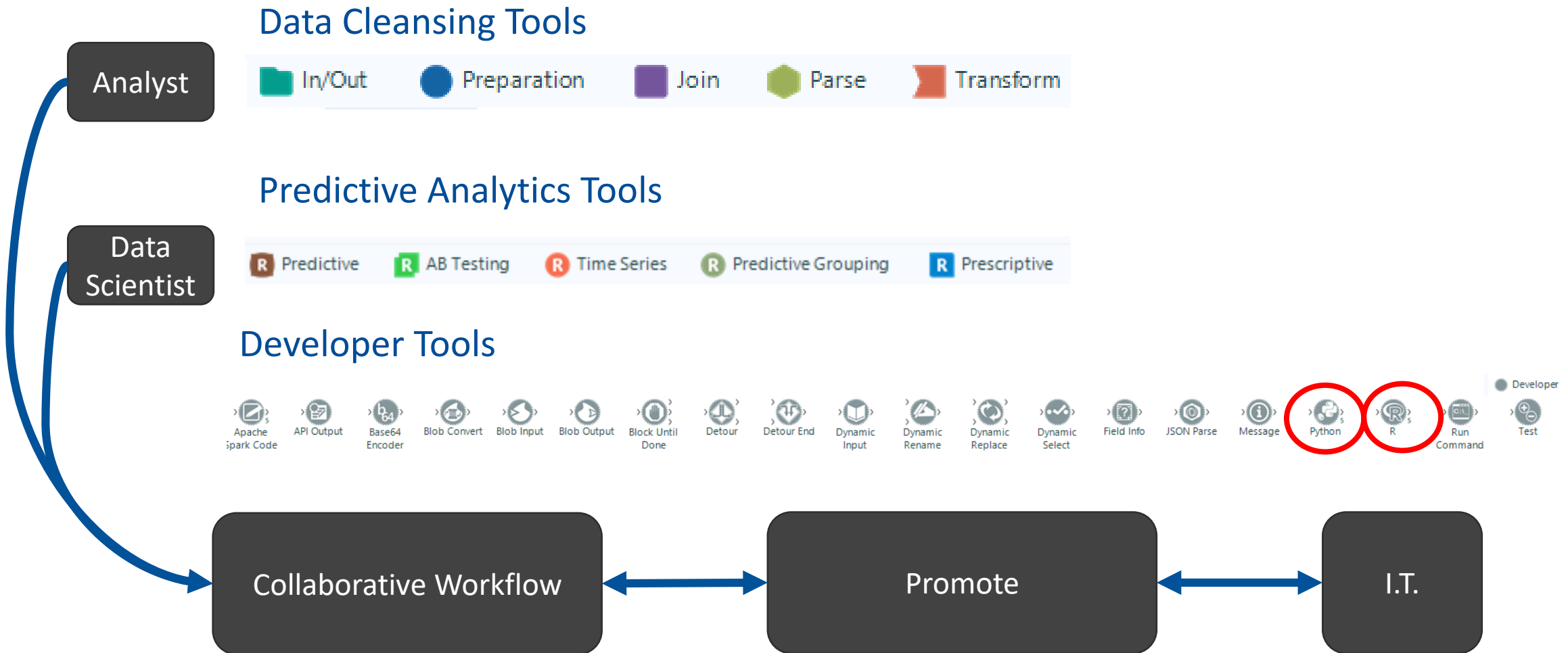
Gallery

Scheduling

Connect

Promote


Enterprise Solution





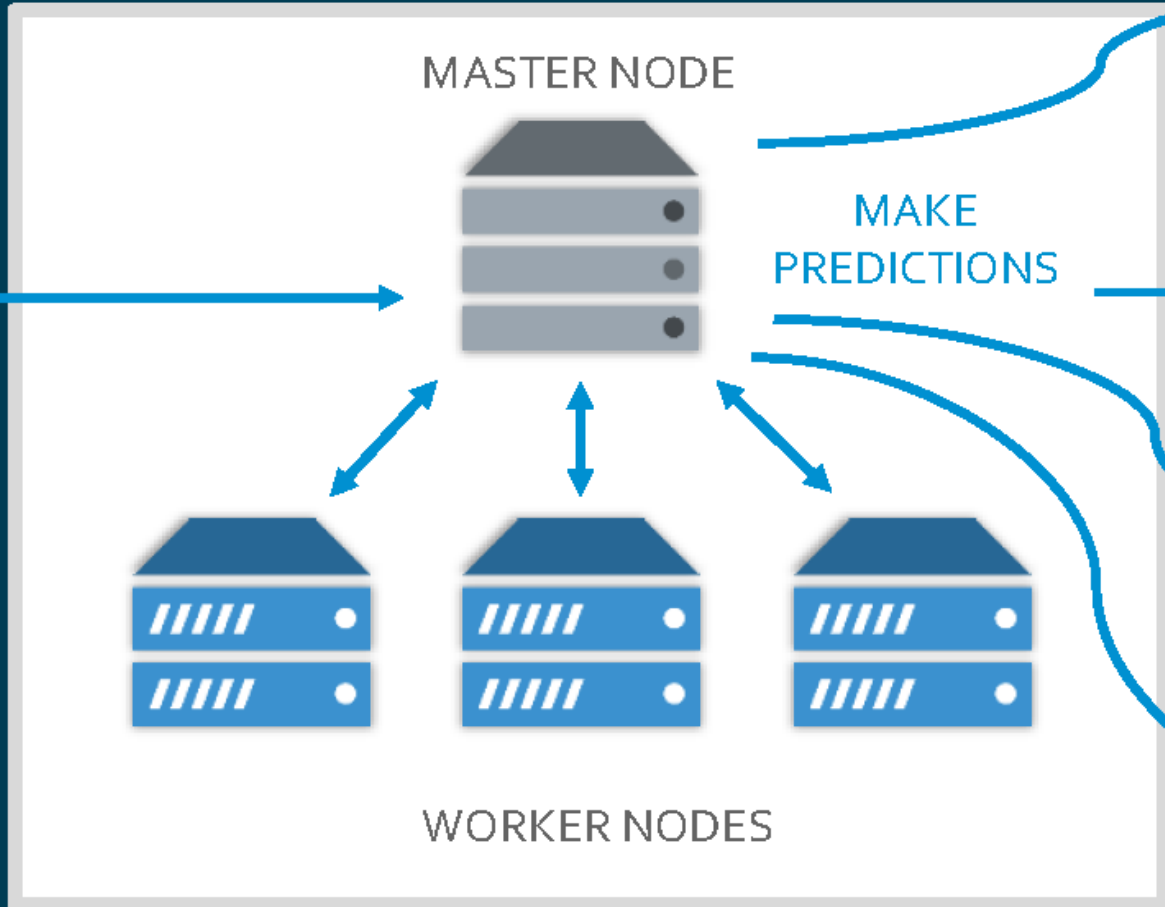
alteryx | PROMOTE

PREDICTIVE MODELS

DATA SCIENCE & ANALYTIC TEAMS





Code Free & Code Friendly


FRONT-LINE EMPLOYEES


MARKETERS


CUSTOMERS


APPLICATIONS

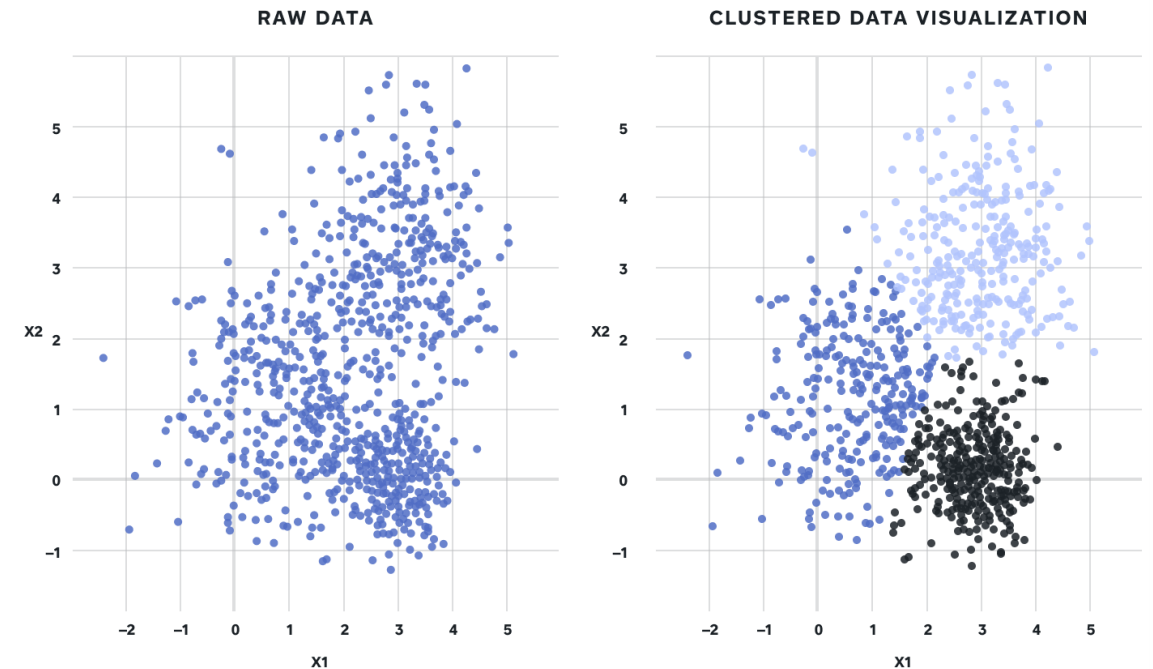
SENSE
CORP

Clustering Model

Retail Customer Segmentation

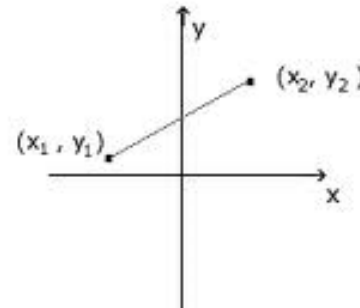
What is K-Means Clustering?

- **Set of observations with several available measurements**
 - We want to find out if the observations naturally group together in some predictable way
- **How do we cluster?**
 - Measuring the distance (or some other measure of similarity) between the different observations
- **K-means**
 - Starts first with a group of randomly selected centroids, then performs iterative calculations to optimize the positions of the centroids



Choices to Make

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$



- **Distance**

- Euclidean distance
- Manhattan distance

- **Standardization**

- Z-Scores: subtract their mean, and divide by their standard deviation
- Unit Interval: subtracting the minimum value of the field and then dividing by the range of the field

- **How many clusters do you pick?**

- Calinski-Harabasz Index- ‘how separated clusters are from each other and how compact within each cluster’
- Adjusted Rand Index – ‘how similar units are within a cluster’

The Data Set



Historical sales data for 45 Walmart stores located in different regions.

Each store contains several departments



Provided 3 Tables

Features Table

Sales Table

Store Table

Outcome



CLUSTER 1

- Large Stores
- Purchases are made when economy is doing good
- Brings in the most sales
- Even when CPI is high, consumers continue to shop at these stores

(11 stores out of 45)



CLUSTER 2

- Medium Sized Stores
- Most likely customers with less income in these stores as only shopping when CPI is low
- Average sales are made

(21 stores out of 45)



CLUSTER 3

- Small Stores
- Customers make purchases when economy is doing good (i.e. fuel prices are lower, unemployment is lower)
- Bring in the least sales

(13 stores out of 45)

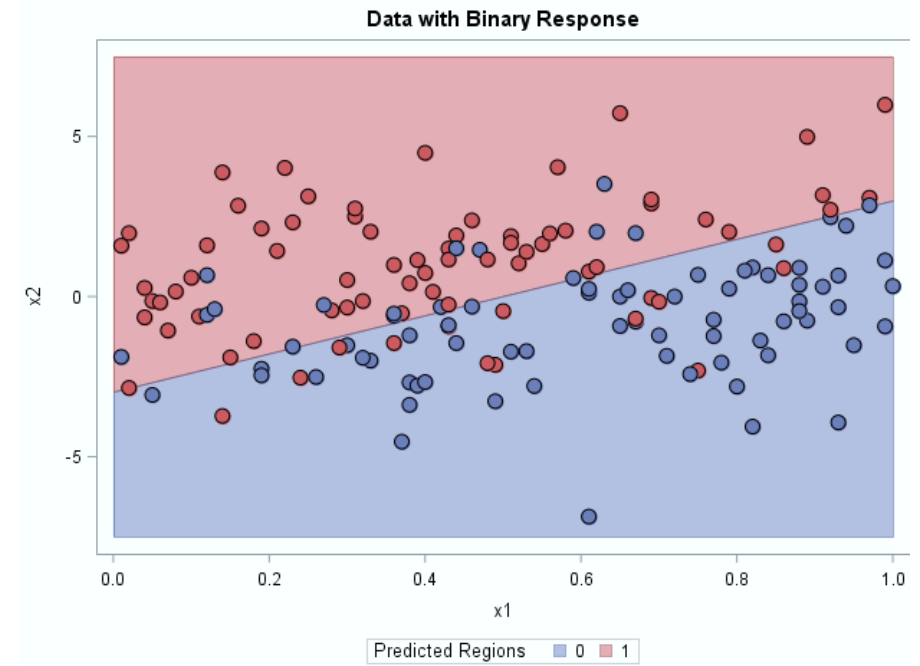
SENSE
CORP

Classification Model

Marketing Campaign Responses Success

What is Binary Classification?

- Falls into the domain of supervised learning since the training dataset is labelled. Simply a special case in which there are only two classes you are trying to predict.
 - Some typical examples include:
 - Credit Card Fraudulent Transaction detection
 - Medical Diagnosis
 - Spam Detection
- Various models can be used for binary classification which include:
 - Decision Trees
 - Neural Networks
 - Logistic Regression
 - Support Vector Machines

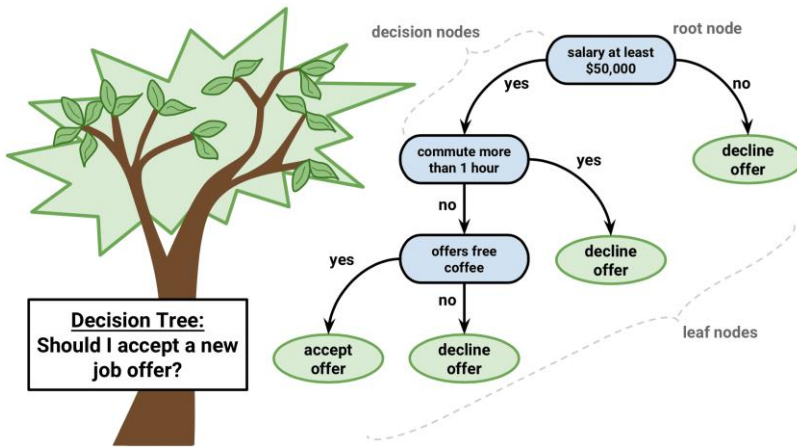


The Data Set

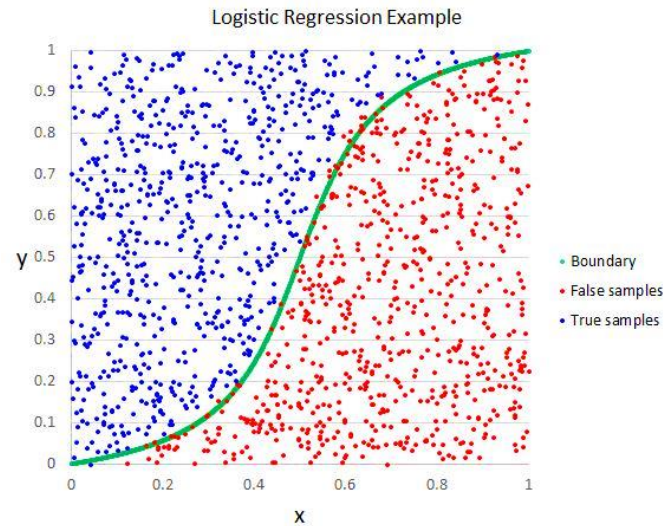
- **This is a bank telemarketing data set**
 - From UCI machine learning repository
- **Data set has information about customers and the campaigns done by bank towards term deposit.**
 - There is column named y in the data set which tells whether or not the customer opted term deposit
- **Some other predictor variables include:**
 - Age
 - Marital
 - Education
 - Job
 - Default – has credit in default?
 - Balance – average yearly balance
 - Pdays - # of days that passed after client was last contacted from a previous campaign
 - Duration – last contact duration (in seconds)

Models Used

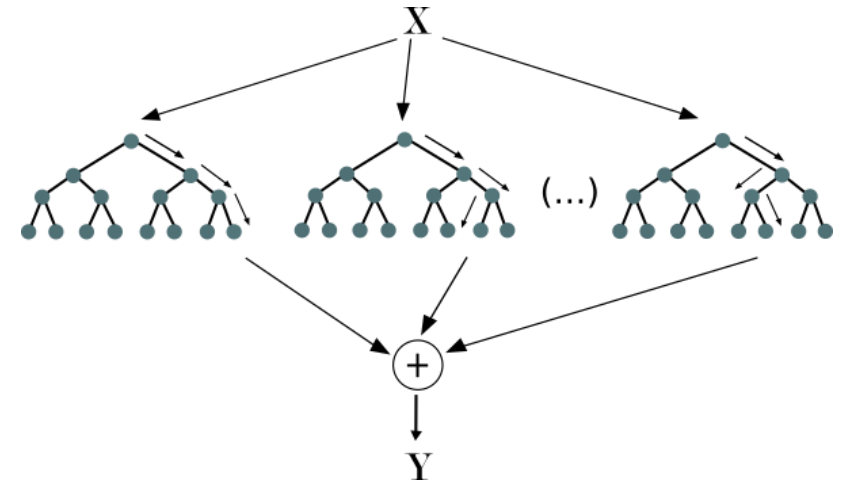
Decision Tree



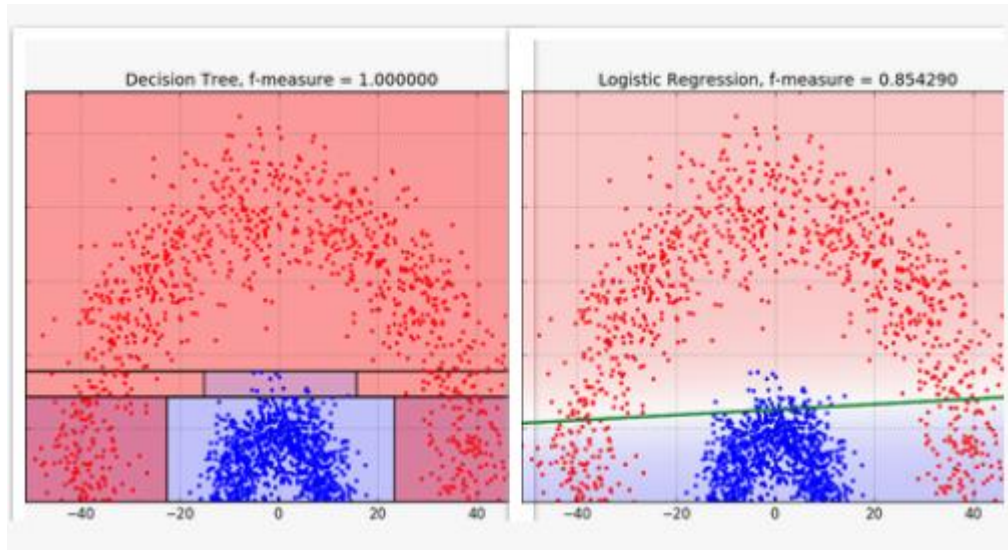
Logistic Regression



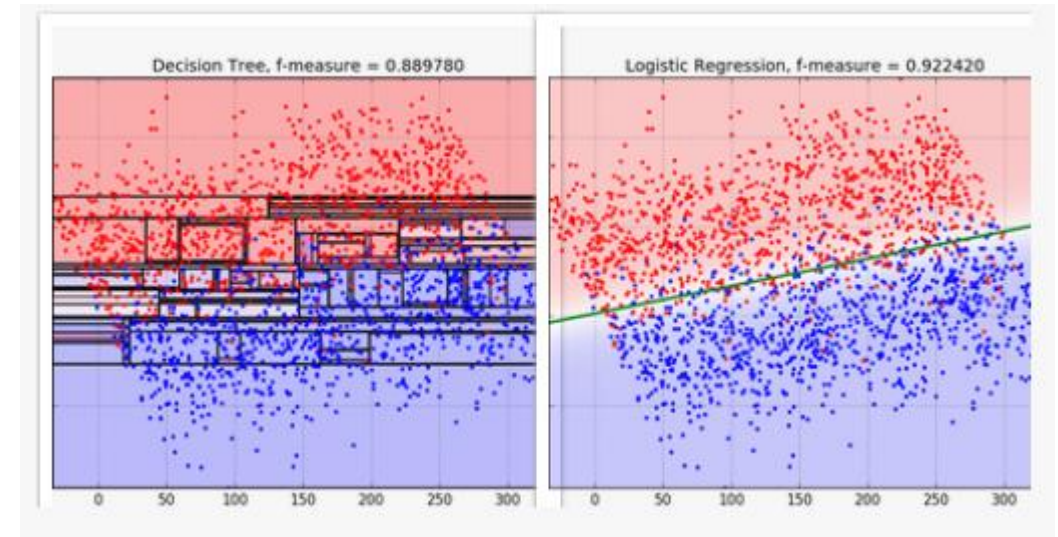
Forest Tree Model



Decision Trees vs. Logistic Regression



If this were the data, a decision tree would be the better model to use



If this were the data, a logistic regression would be the better model to use

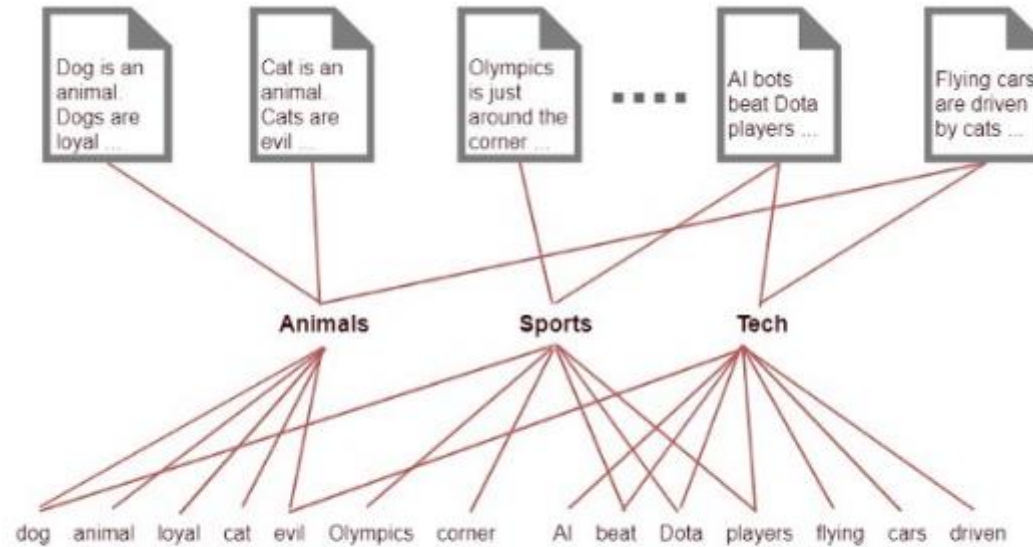
The background of the slide is a photograph of the Austin State Capitol building, a large, classical-style structure with a prominent dome. The building is surrounded by city streets with parked cars and other buildings. A semi-transparent blue rectangle is overlaid on the image, containing the text.

Topic Modeling

Austin Airport vs. MSP Airport Tweets

What is Topic Modeling?

- *“Each document can be described by a distribution of topics and each topic can be described by a distribution of words”*



Words are modeled by a set of topics and documents are modeled by a set of topics. The relationships are clearer than the first example because there's a fewer connections than the first example.

Layman's Explanation

- **Latent Dirichlet Allocation (LDA) is a type of topic modeling algorithm that learns the representation of a fixed number of topics and learns the topic distribution that each document in a collection of documents has as well as the distribution of words in each topic. For example, if we were given the following sentences:**
 - A: I spent the day watching Messi kick the ball around and play soccer.
 - B: I ate spaghetti with meatball sauce for dinner.
 - C: I love eating pasta with marinara sauce while watching Argentina play soccer in the World Cup.
- **LDA might say something like:**
 - Sentence A is 100% about Topic 1
 - Sentence B is 100% Topic 2
 - Sentence C is 50% Topic 1, 50% Topic 2
- **LDA will also discover that:**
 - Topic 1: Soccer 60%, Ball 20%, Play 10%, World Cup 10% ... (where we notice that topic 1 represents things related to Soccer)
 - Topic 2: Sauce 40%, Spaghetti 20%, Meatballs 20%, Pasta 20% , ... (where we notice that topic 2 represents things related to Italian Food.)

The Data Set

- **Used Twython (Python Package) to pull live data from twitter**
 - 100 pull limit per search term
 - Used @mspairport and @AUSTinairport twitter handles (official airport twitter handles) for the search
- **Used regex functions to pull out body of tweet and place all atx and msp tweets into each of their own data frames**



Past Topics

Record #	id	misp_topic
1	0	0.015*well + 0.015*trump + 0.010*time + 0.010*off
2	1	0.019*air + 0.012*today + 0.012*force + 0.012*it
3	2	0.020*mn + 0.016*paul + 0.016*-saint + 0.016*international

- Trump and Air Force One on 2019-04-16 run

Record #	id	misp_topic
1	0	0.009*hour + 0.009*also + 0.009*in + 0.007*way
2	1	0.015*today + 0.012*update + 0.010*power + 0.010*know
3	2	0.015*flight + 0.008*air + 0.008*make + 0.008*mn

- MSP Power Outage on 2019-04-17 run

Record #	id	atx_topic
1	0	0.020*bergstrom + 0.018*international + 0.011*flight + 0.010*time
2	1	0.025*parking + 0.009*full + 0.009*would + 0.009*update
3	2	0.011*flight + 0.008*today + 0.007*weather + 0.007*severe

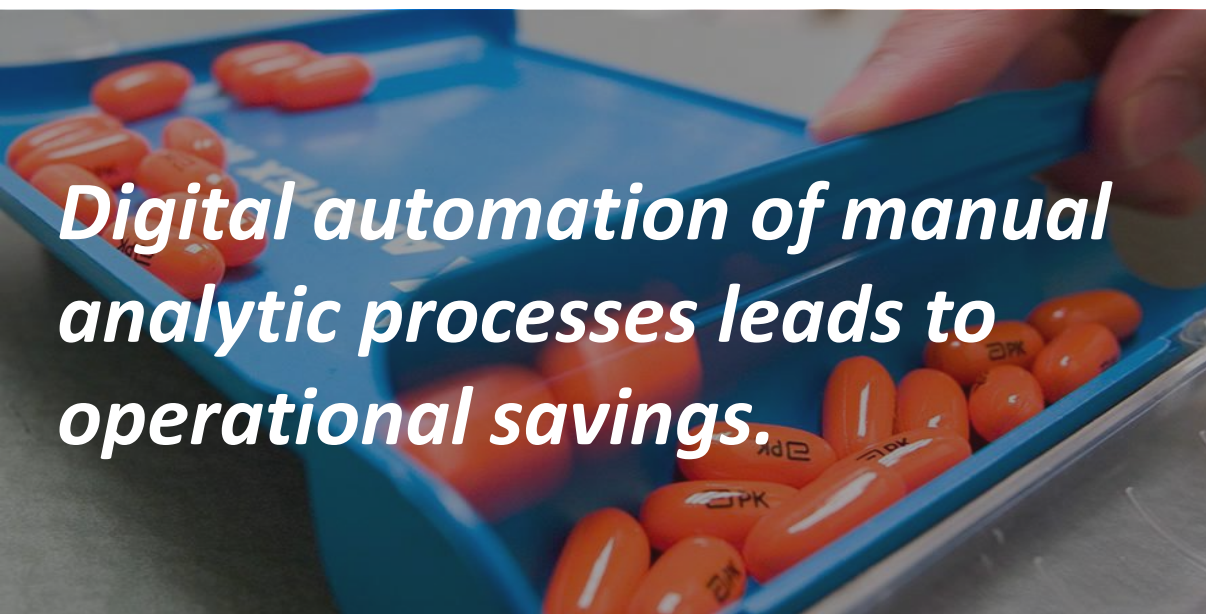
- Severe Weather (hail and tornado watch) and Parking lot full on 2019-04-17 run

Record #	id	atx_topic
1	0	0.010*parking + 0.010*update + 0.010*best + 0.010*robotic
2	1	0.014*terminal + 0.010*opening + 0.008*see + 0.008*flight
3	2	0.014*n + 0.012*international + 0.012*bergstrom + 0.012*stage

- New Terminal Opening where you can sit outside and see flights on 2019-04-23 run

Record #	id	atx_topic
1	0	0.013*fly + 0.013*new + 0.011*stage + 0.009*gate
2	1	0.013*n + 0.011*international + 0.011*bergstrom + 0.009*great
3	2	0.014*flight + 0.011*coffee + 0.011*robotic + 0.011*briggo

- 24/7 Briggo Robotic Coffee Machine on 2019-04-24 run



Digital automation of manual analytic processes leads to operational savings.

RESULTS

- ✓ Capability to monitor all pharmacy networks performance daily instead of quarterly.
- ✓ **Client earned the president's award** (highest level of recognition at the company).
- ✓ **Operational savings.** Avoided costly solution. Re-deployed analysts to higher-value work.
- ✓ Positioned to extend with Geo Spatial capability to correlate distance between customers getting prescriptions and their pharmacies in order to drive **new revenue opportunities with delivery service partners.**

Situation

The pharmacy network analytics team is responsible for the monitoring and negotiations of multiple pharmacy networks for their clients. This PBM has over 2,500 networks that they contract to manage drug cost for their clients.

Challenge

To achieve monitoring and performance management of their networks, many SAS scripts would consolidate data feeds from third party systems to be analyzed. Analysis could only be performed quarterly given technology solution limitations and the manual data engineering. This department was looking to spend \$1MM to have a system created and hire 11 additional analysts to support future growth.

Solution

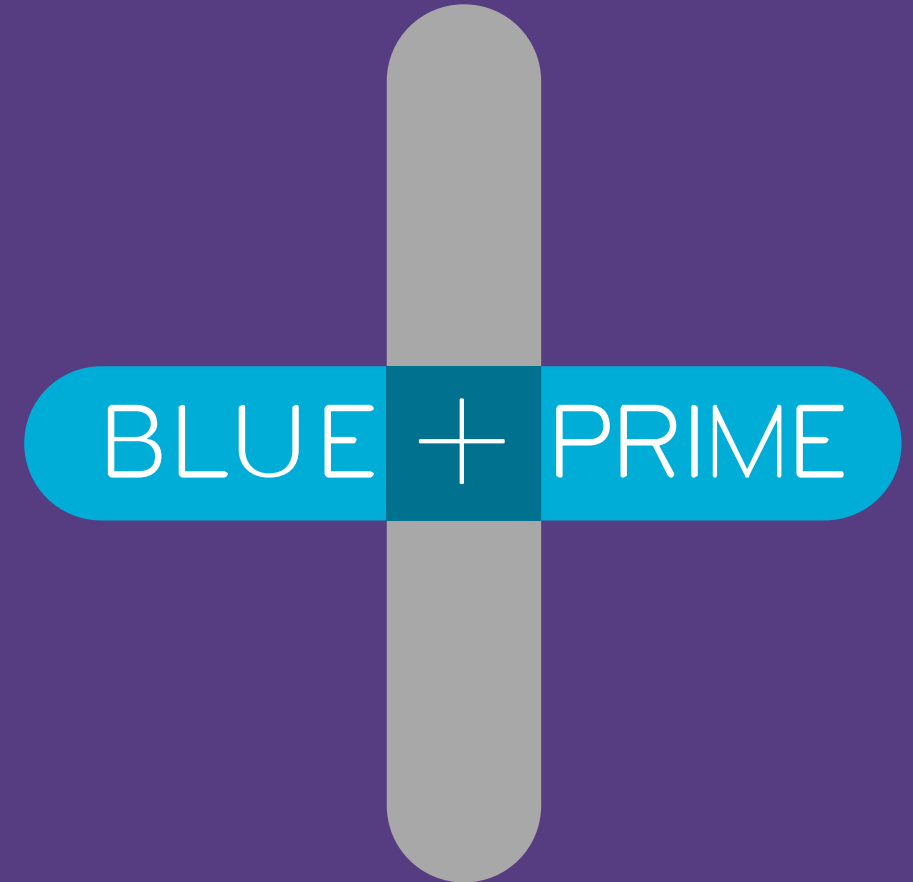
- ✓ Introduced Alteryx. In 90 days, delivered Alteryx solution to fully automate data engineering processes.
- ✓ Delivered Tableau visualizations of KPIs for analysis.

Prime Therapeutics - Scaling Business with Alteryx



Prime Therapeutics

- Processes claims and offers clinical services for people with complex medical conditions.
- Prime serves more than 27 million people.
- Collectively owned by 18 Blue Cross and Blue Shield Plans, subsidiaries or affiliates of those plans.



Business Expansion

- In 2018 the marketplace demanded an industry-wide increase in customization of pharmaceutical management.
- Scalability needed in all product spaces

- **Inability to scale any business based on current technologies, pre Alteryx.**

Needed the ability to customize to client with no barriers and quick TAT.

Data Collection and Blending

• OLD PROCESS

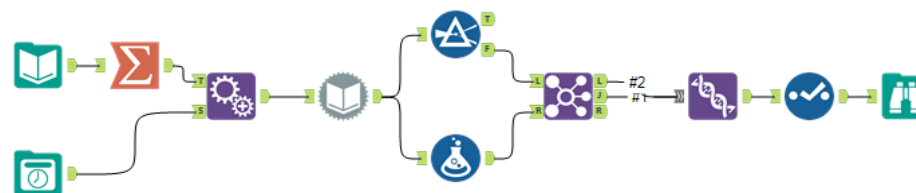
- Weekly Process
- Excel
- Access
- SPSS
- Transactional Data
- System Data
- Industry Data
- COPY/PASTING

• **15 Hours / Week**

NEW PROCESS

- **Dynamic Inputs**
- **Preparation Toolset**
- **Join Toolset**
- **Developer Toolset**
- **Transform Toolset**
- **Parse Toolset**

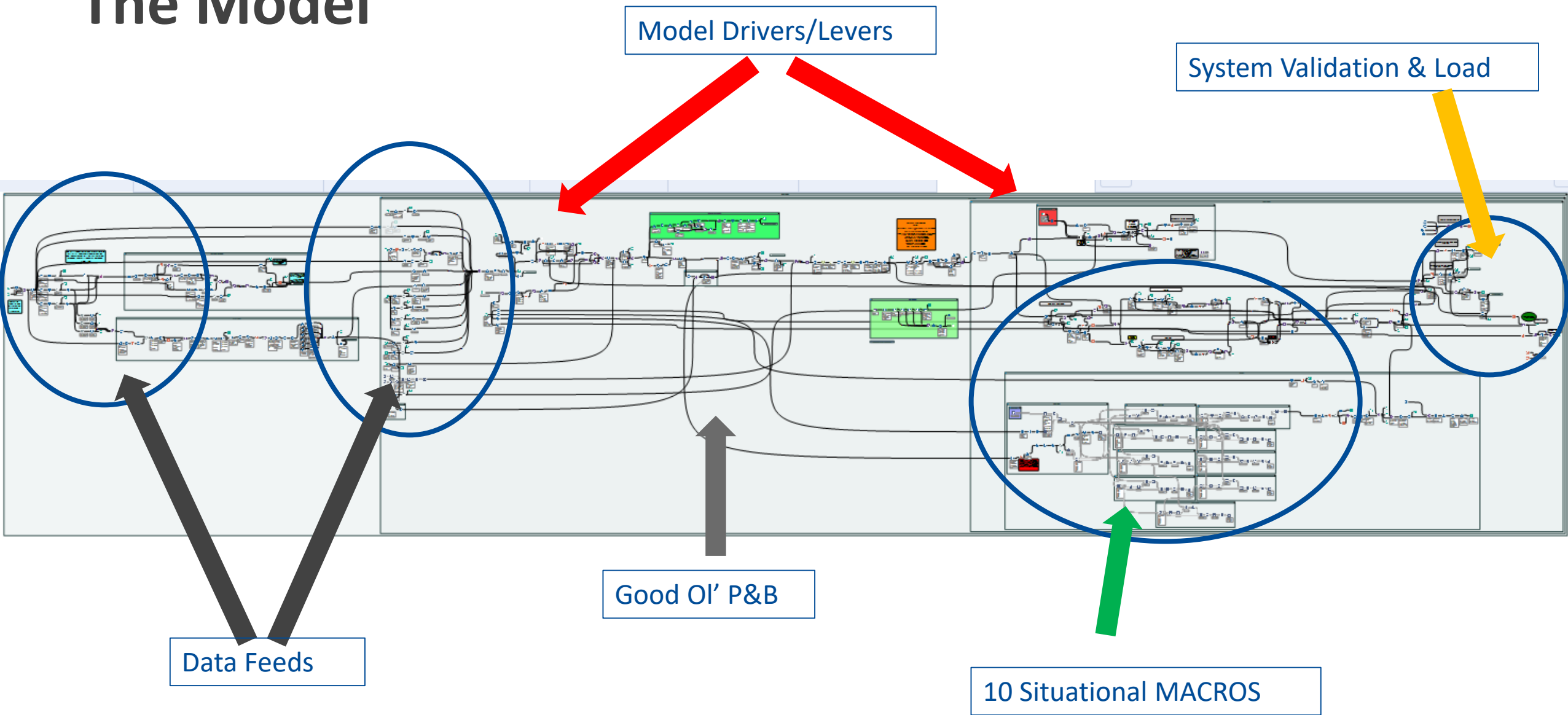
10 Minutes / Week



OLD PROCESS

UP	MAC Generic Name	Brand Name	Group	# Generics	# Manufacturers	# W/ TEE's	Current MAC	Old MAC	Last Changed Date	Intelligible	New Rate	Reason	Rate Change	Savings	ASP	Auth/Qty	Auth
11200040102105	DONORUBICIN HCL FOR INJ 10 MG	ADRIAMYC INJ 10MG	1	1	1	1											
10120020007210	GLUCOPHAGE ORALLY DISINTEGRATING TAB 12	FACADUO TAB 12 30MG	1	1	2	1									2.20	13	
44201040102005	ISOPROTERENOL HCL INJ 0.2 MG/ML	ISUPREL INJ 0.2MG/ML	1	1	4	1	405.1170	405.1170	6/30/2018		41.51%	0.0000	0.00% \$	-	770.28		
44200010000610	BISACODYL TAB DELAYED RELEASE 5 MG	DUALCOLAR TAB 5MG EC	1	10	34	1	0.2000	0.2000	5/24/2018		-278.11%	0.0000	0.00% \$	-	0.05	34	
799300102000	KCL 0.2% IN D5/D 45% NACL	KCL/D5W/NACL INJ 0.2/5/45	1	1	1	1	0.0097	0.0097	9/13/2018		-81.79%	0.0000	0.00% \$	-	0.01		
7993001020025	KCL 0.15% IN D5/D 45% NACL	KCL/D5W/NACL INJ 0.15/5/45	1	1	1	1	0.0081	0.0081	9/13/2018		-81.22%	0.0000	0.00% \$	-	0.01		
7993001020015	KCL 0.075% IN D5/D 45% NACL	KCL/D5W/NACL INJ 0.075/5/45	1	1	1	1	0.0036	0.0036	9/13/2018		3.80%	0.0000	0.00% \$	-	0.01		
13400010100005	NALTREXONE HCL TAB 50 MG	REXIA TAB 50MG	1	1	1	18	1.0509	1.0509	9/20/2018		75.01%	0.0000	0.00% \$	-	4.20	70,071	
10130010104120	HYDROCORTISONE BUTYRATE LOTION 0.2%	LOCOD LOT 0.2%	1	1	4	1	7.1380	7.1380	9/20/2018		11.80%	0.0000	0.00% \$	-	8.11	148	
10190002200010	AMOXICILLIN & 4 CLAVULANATE TAB 250-125	AUGMENTIN TAB 250MG	1	1	1	1	6.0687	6.0687	9/20/2018		-1.34%	0.0000	0.00% \$	-	5.82	35	
101340010000710	NETOCONAZOLE CREAM 2%	NIZORAL CRE 2%	1	4	4	15	1.1133	1.1133	9/20/2018		75.00%	0.0000	0.00% \$	-	4.46	90,613	
10130010104010	GLOBETASOL PROPIONATE GEL 0.05%	TEMOVATE GEL 0.05%	1	1	1	15	2.2820	2.2820	9/20/2018		84.82%	0.0000	0.00% \$	-	6.51	1,728	
10130010100705	HYDROCORTISONE VALERATE CREAM 0.2%	WESTCORT CRE 0.2%	1	1	1	12	3.1115	3.1115	9/20/2018		26.79%	0.0000	0.00% \$	-	4.25	4,283	
14000010000005	TRIMETHOPRIM TAB 200 MG	TRIMPRX TAB 200MG	1	4	4	7	0.3270	0.3270	9/20/2018		52.20%	0.0000	0.00% \$	-	0.88	14,580	
10130040004205	DESCHMETASONE OINT 0.25%	TOPICORT OIN 0.25%	1	1	1	18	1.2495	1.2495	9/20/2018		77.00%	0.0000	0.00% \$	-	5.45	1,813	
1401200102010	OLPATADINE HCL OPHTH SOLN 0.2% (BASE)	PATADAY SOL 0.2%	1	4	1	1	22.1230	22.1230	9/9/2018		88.83%	0.0000	0.00% \$	-	79.52	1,824	
11200010000100	DEXAMETHASONE TAB 1.5 MG	DECADRON TAB 1.5MG	1	2	2	1	0.2096	0.2096	9/30/2018		48.36%	0.0108	-11.80% \$	90.40	0.48	2,891	
10250010000170	TAZAROTENE CREAM 0.1%	TAZORAC CRE 0.1%	1	1	4	1	7.8055	7.8055	9/30/2018		39.50%	0.0000	0.00% \$	-	13.00	45,580	
10130002100710	MONOMETASONE FURIOATE CREAM 0.1%	EUCON CRE 0.1%	1	4	1	14	0.6130	0.6130	9/30/2018		50.10%	0.0000	0.00% \$	-	1.23	47,395	
11200040101820	MONOMETASONE FURIOATE NASAL SUSP 50 MG	NASONEX SPR 50MG/5ML	1	1	4	1	5.1353	5.1353	9/14/2018		66.78%	0.0075	-0.14% \$	529.20	15.44	71,494	
17200000002010	FLUDONAC OTHC SOLN 0.2%	FLUDIN OTHC DRG 0.2%	1	4	1	13	8.7034	8.9078	8/9/2018		74.07%	0.7208	-8.01% \$	12,881.21	31.24	17,751	
10019002000170	LIDOCAINE-PRILUCAINE CREAM 2.5-2.5%	EMLA CRE	1	1	1	13	0.7257	0.6540	5/31/2018		59.65%	0.0897	-9.63% \$	23,627.40	1.62	338,987	
10130002104210	MONOMETASONE FURIOATE OINT 0.1%	EUCON OIN 0.1%	1	4	1	12	0.3861	0.3861	9/14/2018		71.06%	0.0010	-0.27% \$	26.94	1.51	26,940	
449120430210	PEG 3350 HCL SOO BICARB-NACL FOR SOLN 4	NALTYEL SOL	1	4	1	1	0.0042	0.0041	7/11/2018		34.10%	0.0001	-1.98% \$	772.94	0.01	1,725,410	
11200010102003	METHOTREXATE SODIUM INJ 15 MG/2ML	METHOTREXATE INJ 25MG/5ML	1	1	1	1	2.4758	2.4582	8/9/2018		-4.99%	0.0178	-0.71% \$	-	2.31	340	
101210100002005	GENTAMICIN SULFATE OPHTH SOLN 0.3%	GENTACIDIN SOL 0.3% OP	1	1	1	22	1.4979	1.4450	7/5/2018		85.25%	0.0529	-3.53% \$	495.72	4.18	9,371	
10250010104410	TRIAMCINOLONE ACETONIDE IN CRABAGE 0	KENALOG/CRAB PST 0.1%	1	4	4	1	9.5805	9.5805	7/19/2018		40.92%	0.0000	0.00% \$	-	18.22	4,342	
1013000104210	TRIAMCINOLONE ACETONIDE OINT 0.1%	KENALOG OIN 0.1%	1	1	1	17	0.1889	0.1889	7/26/2018		-17.38%	0.0000	0.00% \$	-	0.14	90,877	
1001900102020	KETOROLAC TRIMETHAMINE OPHTH SOLN 0	ACULAR SOL 0.3% OP	1	4	1	15	2.5368	2.4797	7/5/2018		88.29%	0.0571	-2.20% \$	367.08	21.17	6,434	
10130000004205	FLUCINONIDE OINT 0.05%	LIDEX OIN 0.05%	1	1	1	17	1.3304	1.2041	7/26/2018		74.46%	0.1283	-9.49% \$	4,574.27	4.71	98,218	
1120001010710	TRAMADOL HCL TAB SR 24HR 200 MG	ULTRAM SR TAB 200MG	1	4	4	1	3.4191	2.8190	5/31/2018		57.06%	0.6001	-17.00% \$	9,517.29	6.48	13,860	
112000100001040	DEXAMETHASONE TAB 6 MG	DECADRON TAB 6MG	1	1	1	1	1.1987	1.1802	7/5/2018		10.83%	0.0185	-1.54% \$	12.75	1.78	688	
101300100001705	BETAMETHASONE DIPROPIONATE AUGMENT	DIPROLENE AF CRE 0.05%	1	4	1	12	0.4638	0.2617	6/28/2018		86.60%	0.2011	-40.57% \$	1,892.75	1.85	25,199	

The Model

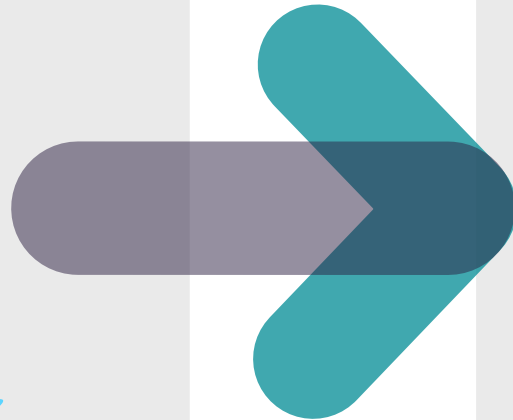


The Product Management Model

• OLD PROCESS

- Weekly Process
- Hundreds of Individual Excel Spreadsheets
- Multiple Access Databases
- VBA Macros - Loading
- 6 Employees

• **175 Hours / Week**



NEW PROCESS

- **1 Alteryx Workflow**
- **Teradata & In-DB Toolset**
- **Developer Toolset**
- **Interface Toolset**
- **Macros (Batch, Standard, Iterative, Crew)**
- **Robot Job (SFTP) - Loading**
- **1 Employee, maybe...**

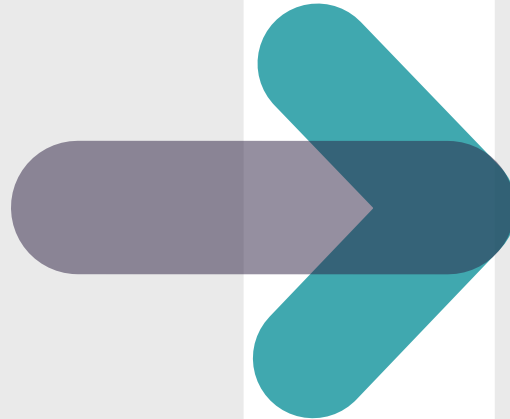
3 Hours / Week

Validation

• OLD PROCESS

- SPSS Queries
- Excel Vlookups

• **5 Hours / Week**



NEW PROCESS

- **Automated Alteryx Process**
- **Directory Tool**
- **Join Toolset**
- **Interface Toolset**
- **Standard Macros**

5 Minutes / Week

Accomplishments

Old Process

10,400

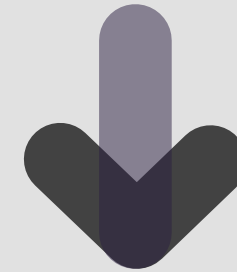
Annualized Hours
(Pre Business Scaling)

New Process

260

Annualized Hours (Post
Business Scaling)

+900% over 2
years.



10,140

Reduction
195 Hours / Wk



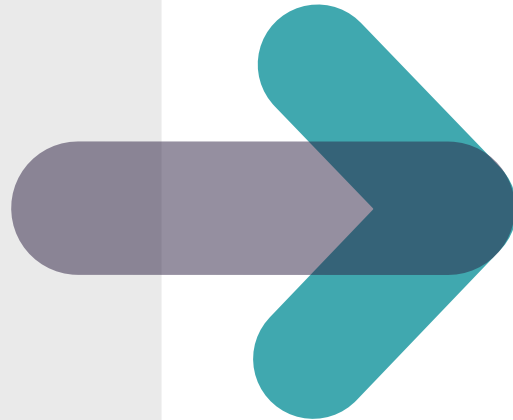
\$800K

~10 FTE Reduction

Client Reporting

Old Process:

- Minimal amounts of client reporting possible.
- Query/Excel based
- Hours to reproduce when needed



Current State:

- Fully Automated Process
- Creates ~500 Unique Reports Weekly and growing
- Reports Tailored to Each Client
- Customized Client Tableau Dashboards

Time Savings Benefits

- Alteryx Power Users:
 - Function as Consultants and Trainers
 - Add Efficiencies within Division
 - Project & TATs
 - Consistent Reporting / Documentation
 - Do what we want, not what we have to do
 - Grow and Evolve Business

What's Next??

- Alteryx Server
- Analytic Apps and Canned Reporting
- Further Tableau Integration
- Data Research and Gathering
 - Public API's
- Predictive Modeling
- Continued Scalability

WHY SENSE CORP?

- ✓ Alteryx partner.
- ✓ '20+ years of delivering enterprise-class data and analytic solutions. Strategy-through-Implementation.
- ✓ Deep data engineering expertise.
- ✓ Robust expertise at every step of the analytics journey towards advanced analytics & data science capabilities.
- ✓ Complimentary Transformation and Digital capabilities that help “close the last mile of analytics”.
- ✓ Local partner with ability to draw upon national strength of our firm.



Sense Corp will accelerate your client's analytics capabilities by streamlining the acquisition of data and supporting technologies with Alteryx.

SENSE
CORP

alteryx

Closing Announcements

- Thanks to the Training Haus
- Thanks to Prime Therapeutics
- Alteryx User Group Leadership

SENSE
CORP

alteryx

Happy Hour!

Hosted by Sense Corp