GREAT LAKES
DATA & ANALYTICS
SUMMIT

Thursday, March 12, 2020 | Metro Detroit

Welcome!

The presentation will begin shortly

@GLSummits
#AnalyticsSummit2020

Hosted by:

DATASCIENCE

Steelcase®

Starting a Data Science Practice in a non-digitally native organization:

5 Things you need to know

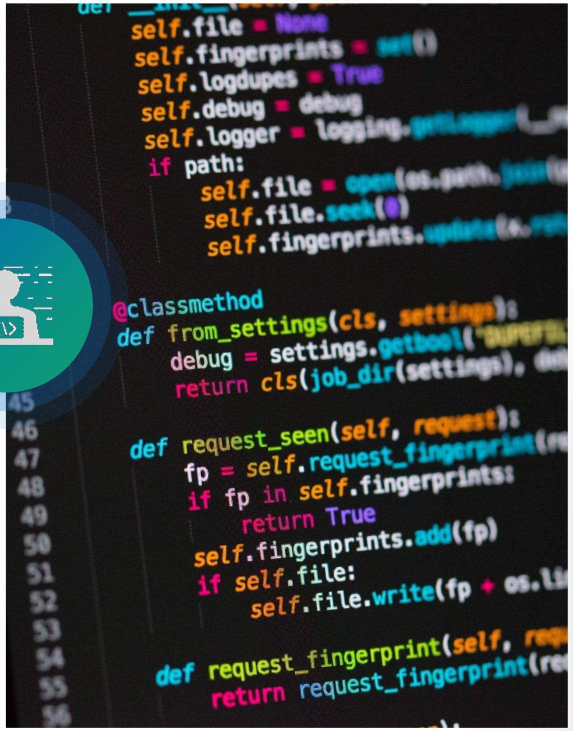
A little bit
About myself

Practicing Data Science for over 8 years

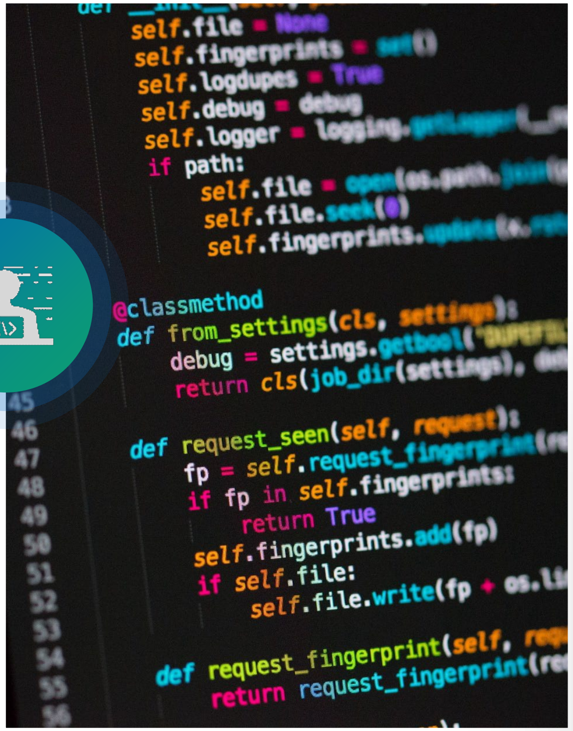
Since 2011, I have been involved in Data Science initiatives at Steelcase that have resulted in powerful transformations for the organization.

Experience

Advanced Analytics	Applied Data Science
Worked for a few years on our Advanced Analytics team at Steelcase	*Currently lead the Applied Data Science team.*

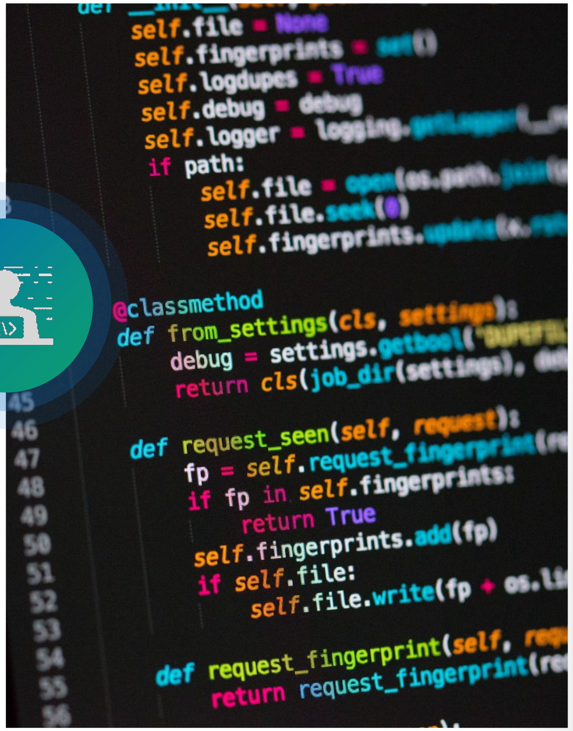


What does a **Data Scientist** do?



What does a **Data Scientist** do?

Truth is, depends on where you work!



What does a Data Scientist do?

Truth is, depends on where you work!

DIGITALLY NATIVE

NON DIGITALLY NATIVE

Everyone will have to be digital and Data Scientists play a crucial role in making that happen.

More than half of the Fortune 1000 companies have not achieved a full digital transformation.

DataScience

How is this profession different across companies??

Digitally Native

Non-Digitally Native

DataScience

How is this profession different across companies??

What you end-up doing as a Data Scientist will depend on:

DataScience

How is this profession different across companies??

What you end-up doing as a Data Scientist will depend on:

- How far an organization is on their digital transformation journey
- How fast an organization is moving on their digital journey
- Digital-dexterity across the organization

5 things you
should know

Data Science practices on a century-old organization presents certain challenges.

- People don’t know what to do with this skillset.
- Some teams are not ready to have a data science project
- Business doesn’t want to be disrupted.
- Legacy Data also requires transformation.

Data Science ✚ Non-Digital

5 things you should know

Know your customer

1

Who is your customer?

Who are you working for?

How do you work together?

DataScience

#1 Know your customer

Who is your customer?

Who are you working for?

How do you work together?

You need to define/segment areas of the business and build your network.

Get together and talk about their business objectives.

Start with the business problems, then allow those to help you build a backlog

5 things you should know

Keep it simple but keep it creative

2

The true value of a Data Scientist will come from the ability to translate a business problem into a data problem that they can solve.

My most successful projects have had very basic modeling components.

DataScience

#2 Keep it simple but keep it creative

$$z = \frac{x - \mu}{\sigma}$$

$\mu =$ Mean

$\sigma =$ Standard Deviation

DataScience

#2 Keep it simple but keep it creative

$$z = \frac{x - \mu}{\sigma}$$

μ = Mean

σ = Standard Deviation

$$\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x$$

DataScience

#2 Keep it simple but keep it creative

$$z = \frac{x - \mu}{\sigma}$$

$\mu =$ Mean

$\sigma =$ Standard Deviation

$$\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x$$

Percentiles!

DataScience

#2 Keep it simple but keep it creative

Predicting the Upcharge of a Custom Order

Predicting the Upcharge of a Custom Order

If I have an average of how much it has cost me to make in the past, how much should I upcharge this custom order?

Will the upcharge that I suggested be enough to reach my margin goals?

DataScience

#2 Keep it simple but keep it creative

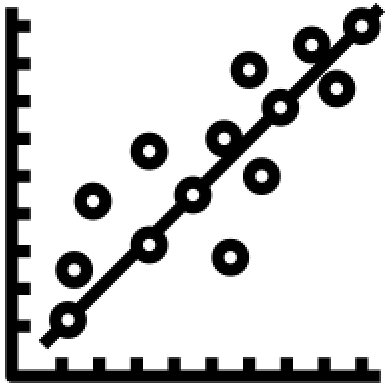
Predicting the Upcharge of a Custom Order

If I have an average of how much it has cost me to make in the past, how much should I upcharge this custom order?

Will the upcharge that I suggested be enough to reach my margin goals?

$$\left.\begin{array}{l} A \longrightarrow B \\ C \longrightarrow x \end{array}\right\} \quad x = \frac{B \cdot C}{A}$$

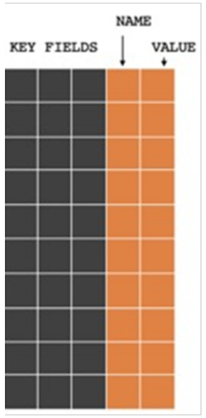
(Rule of 3's)

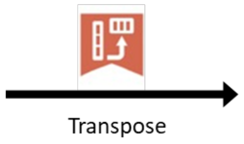


DataScience

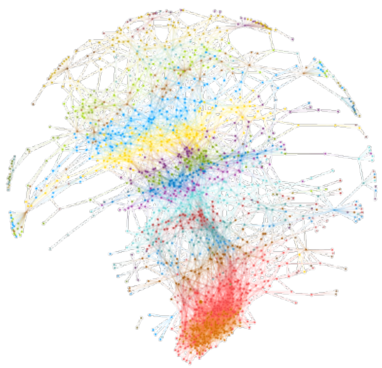
#2 Keep it simple but keep it creative

Extracting key features + configurations creates a data challenge





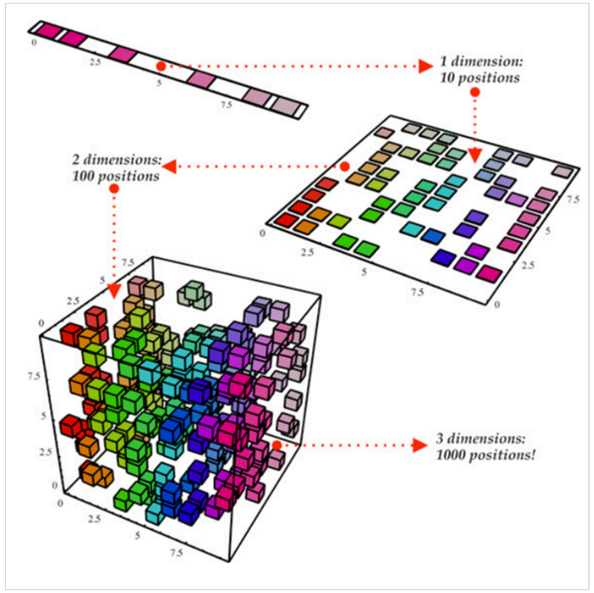
Almost impossible to do with regular tools like excel, Tableau, etc.



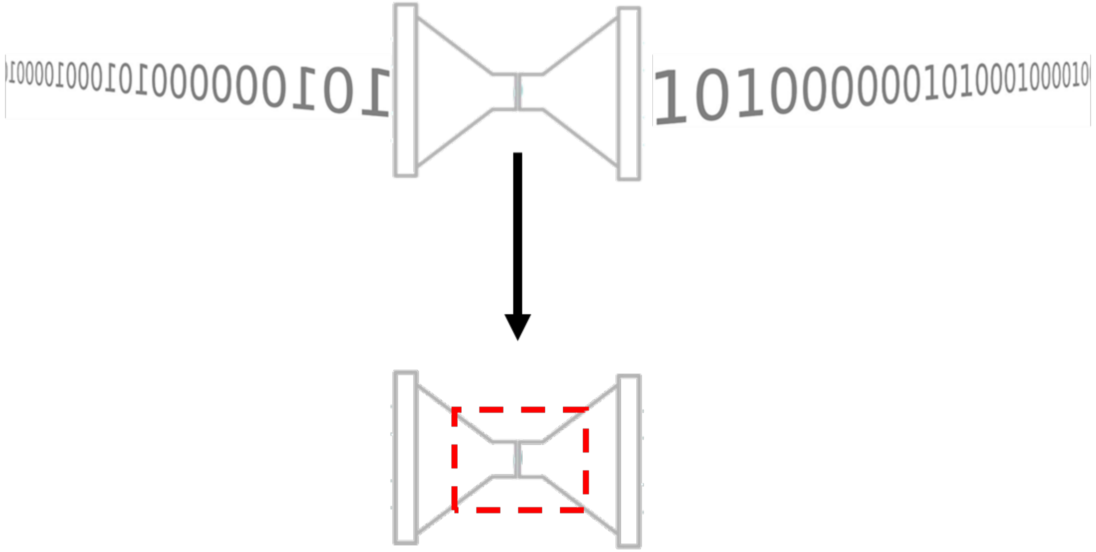
DataScience

#2 Keep it simple but keep it creative

We solve this by creating a lower dimensional representation of the data



(3 dimensional data in 2 dimensions)

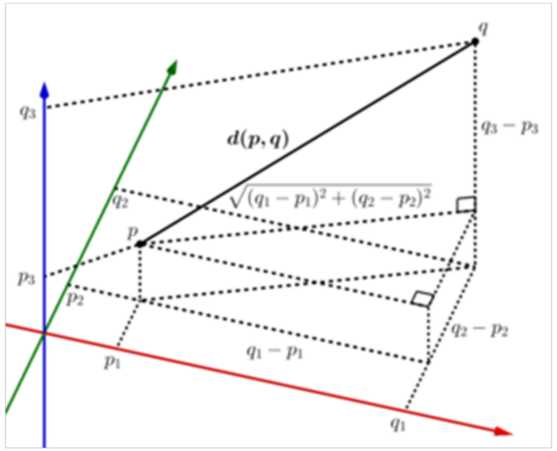


We create a funnel that forces the product configurations to be described in 10 columns (instead of 250)

DataScience

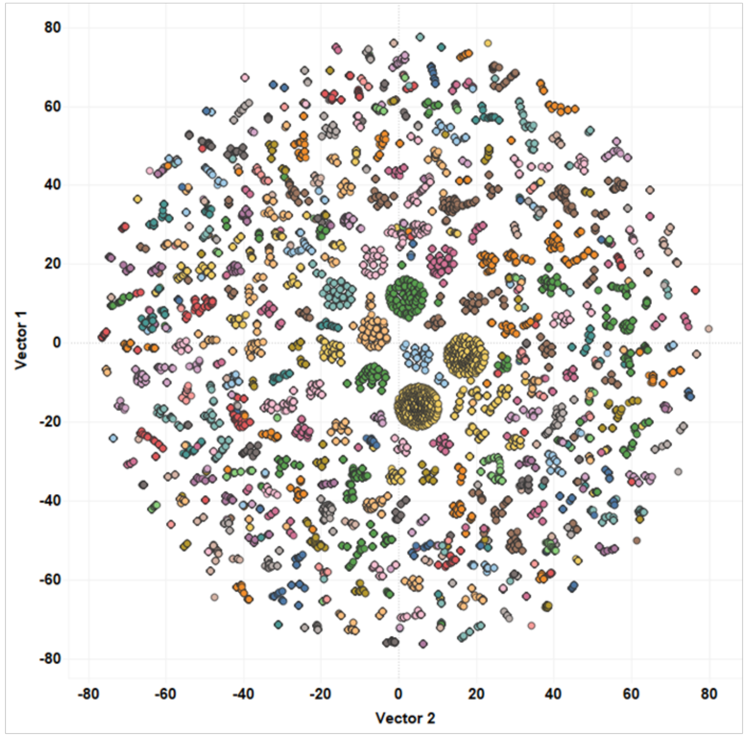
#2 Keep it simple but keep it creative

We obtain coordinates that are useful to measure distance (i.e. similarities)



The closer the point are together from each other, the more similar they will be.

***Actual 2-dimensional representation of the data.**

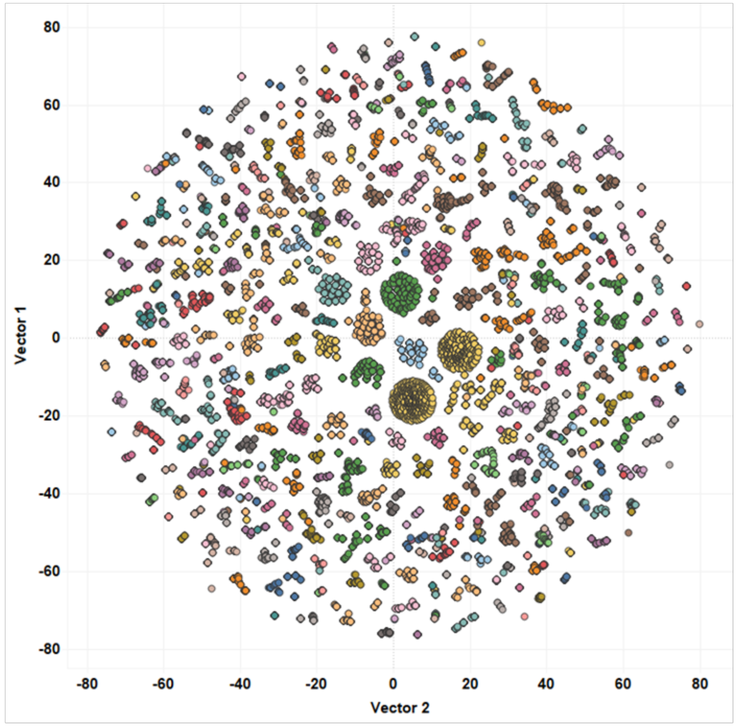


Each dot is one particular product configuration.

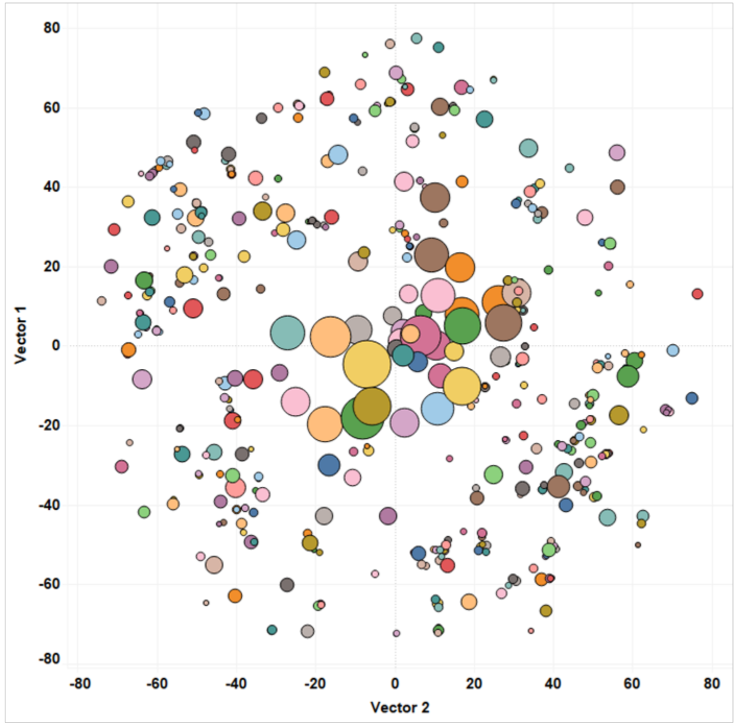
DataScience

#2 Keep it simple but keep it creative

We can now group products that are similar based on their distance



Each dot is one particular product configuration.

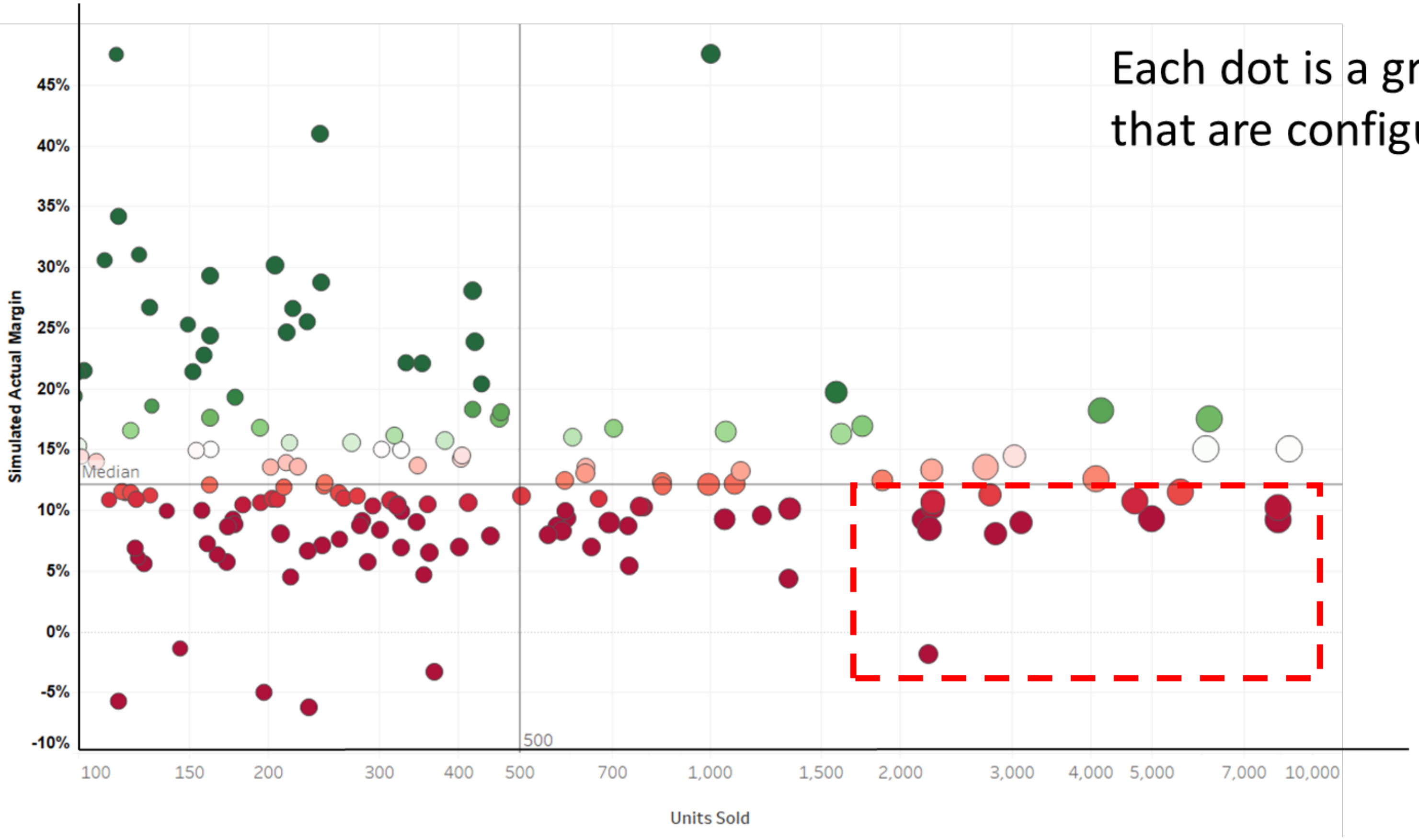


Each dot is one group of products that are similar enough that we can group them.

DataScience

#2 Keep it simple but keep it creative

Using these created groups of products makes it easier to do margin analytics



Each dot is a group of products that are configured similarly.

This matrix helps us prioritize our areas of focus by looking at configurations that have high sales and low margins.

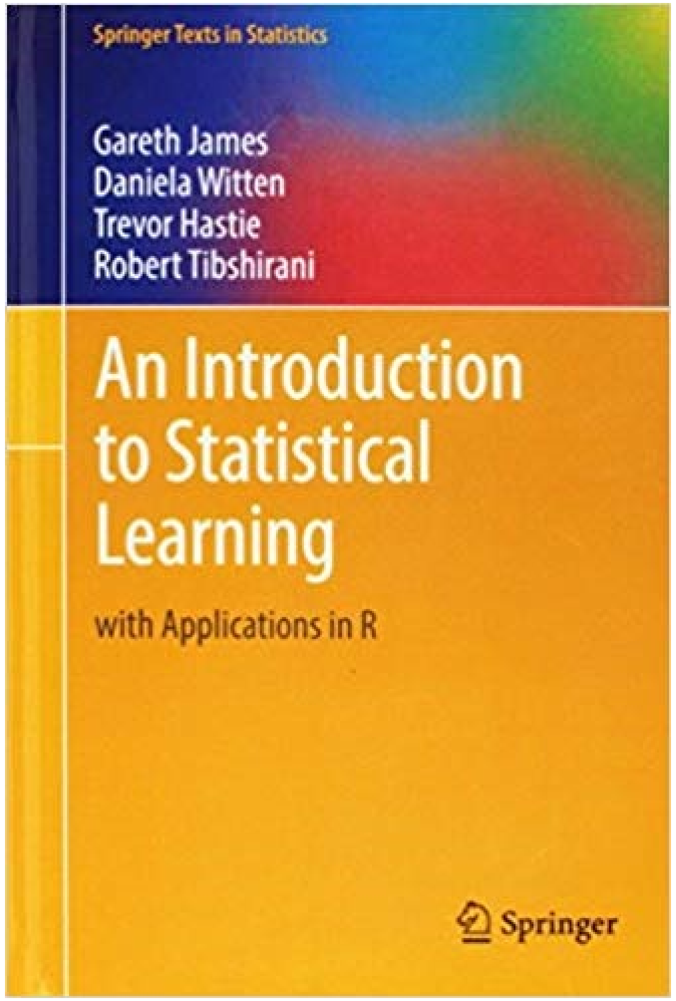
Confidence Intervals

Linear Models

K-Means Clustering

Basic Probability Distributions

(Normal, Log Normal, Beta, Weibull, Poisson, Binomial, Negative Binomial)


Springer Texts in Statistics
Gareth James
Daniela Witten
Trevor Hastie
Robert Tibshirani
An Introduction to Statistical Learning
with Applications in R
Springer

DataScience

#2 Keep it simple but keep it creative

It's not about building the most accurate or complex model…

…it's about creative solutions that enable new capabilities and ways to use the data.

5 things you should know

Those who tell the best stories, rule the world

3

People respond more to stories than to any other narrative.

People respond more to stories than to any other narrative.

- What was the business problem we were after?
- What was the main challenge? Why was it important?
- How are we thinking about the problem?
- What is the main insight?
- What are the decisions that may result from this insight?

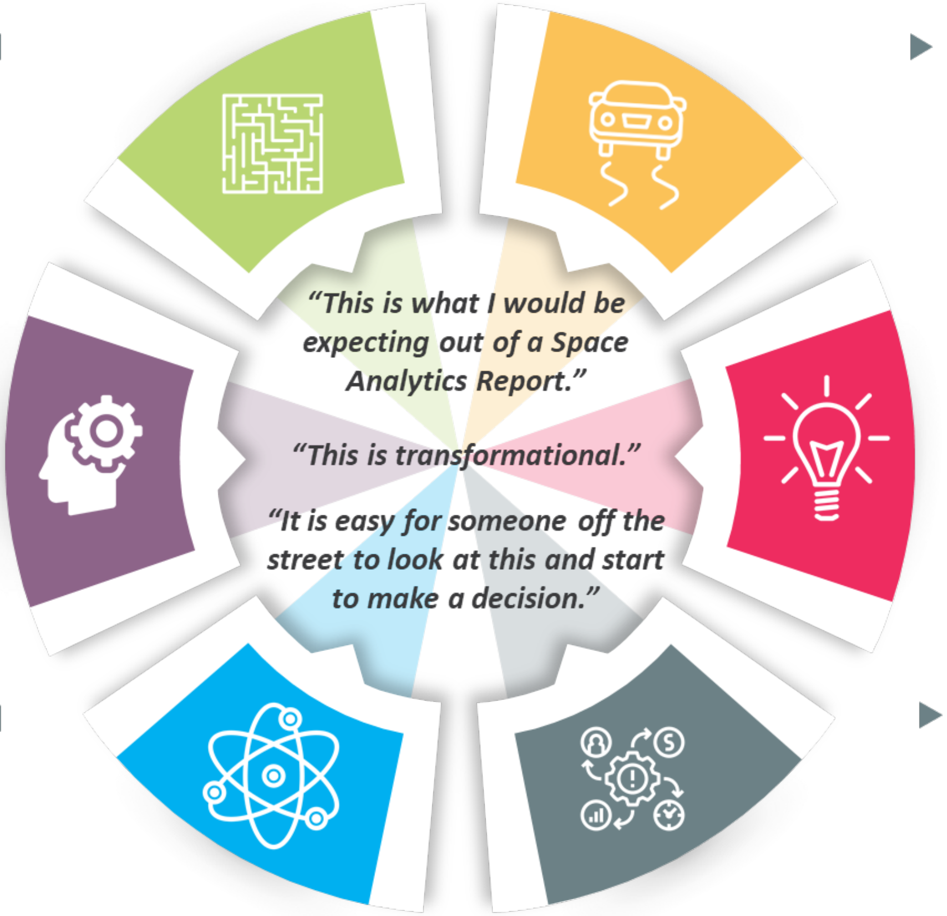
DataScience

Human-Centered Space Analytics

Rough start
Our initial focus was more on the technology than the insights

Gaining traction
Changing people's mindset

Busting the myth that the data has no value



Rebirth
Redefine and focus

Customer point of view

Data Lake

Building for growth
Creating data as a barrier to entry

Setting industry standards and digital IP

User-centered analytics
From data to insights to actions.

This is a differentiator
Expect a 2% increase in win rate → $200 M

5 things you should know

You will fail before you succeed… And that's ok

4

DataScience

#4 You will fail before you succeed. And that’s ok.

Most Data Science initiatives will fail.

DataScience

#4 You will fail before you succeed. And that’s ok.

Most Data Science initiatives will fail.

- Not enough data
- Lack of business engagement
- No efficient way to productionalize the work
- You tried to solve the wrong question

DataScience

#4 You will fail before you succeed. And that's ok.

Most Data Science initiatives will fail.

- You tried to solve the wrong question

DataScience

#4 You will fail before you succeed. And that's ok.

Customer Churn $\longrightarrow$ Predictive Leads

DataScience

#4 You will fail before you succeed. And that’s ok.

Customer Churn → Predictive Leads

ALIGNMENT
Business Understanding
Engaged Executive Leadership

DEVELOPMENT
Solid foundation for new models from prior data investments for pricing analytics

PILOT
Start small, Fail fast
Agile approach to Data Science

PLAN TO SCALE
Clear vision of the broader system. Maximize results by focusing on the user experience

DEPLOYMENT
Integration with CRM Systems
Training, Storytelling & Visualization

SUSTAINMENT
Self-learning Analytical Model
Transition to Sales Program Manager
Share best practices & success stories

Hedge your odds of failing by choosing a diverse backlog.

Hedge your odds of failing by choosing a diverse backlog.

	"Data Approach" Lot's of data available	"Strategy-derived Approach" Importance to the business
"Moon Shot Approach" (1 or 2 but very ambitious)	Lowest Technical Risky Highest risk for adoption	Highest Potential Value Will take time
"Low Hanging Fruit Approach" (Several 'easy wins')	Good for small pilots and POCs	Ideal to gain traction across the organization

5 things you should know

You don’t need a Data Scientist... You need a Data Science team!

5

DataScience

#5 You don’t need a Data Scientist. You need a Data Science team

The successful practice of Data Science in an organization is requires a mix of skills and personas.

Assuming that a single individual can handle Data Science initiatives from start to finish is risky.

DataScience

#5 You don't need a Data Scientist. You need a Data Science team

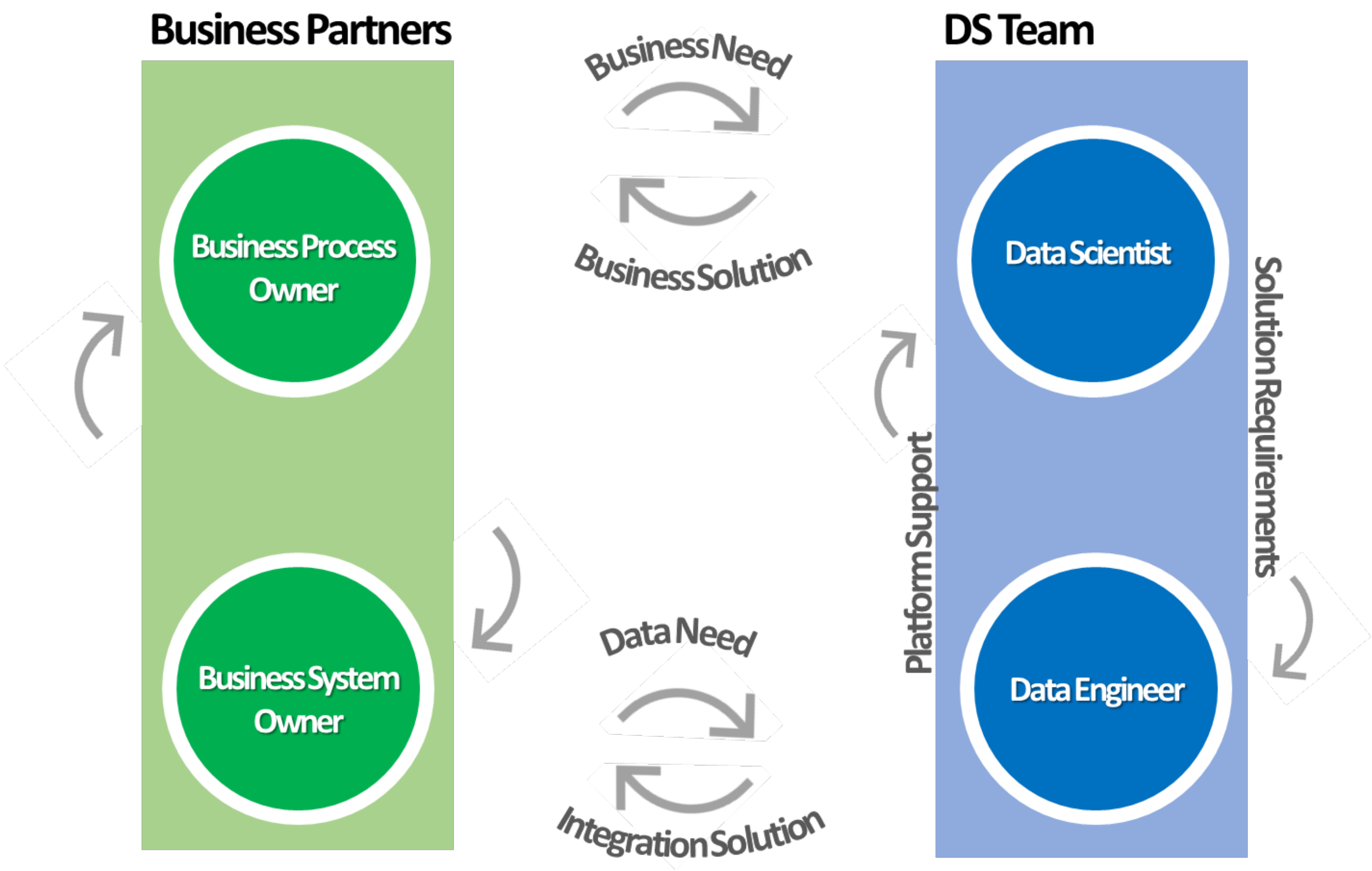
Analytics Roles

Analytics Translator	Data Engineer	Analytics Application & Interface Developer	Data Scientist
Define & Scope	Gather & Stage Data	Implement & Operationalize	Execute Analysis
*Analytics Portfolio Lead	*Data Engineer	*Machine Learning Engineer	*Lead Modeler

Source: Taken from "Fastest growing analytics + data science roles today" | Bill Franks, IIA (2019)

DataScience

#5 You don't need a Data Scientist. You need a Data Science team



- Data Scientist works with Business process owner to identify business problems and define solutions.
- Data Engineer works with business system owner and data scientist to develop data products

5 Things you should know

Quick recap

01 Know your customer

02 Keep it simple but keep it creative.

03 Those who tell the best stories, rule the world.

04 You will fail before you succeed. And that's ok.

05 You don't need a Data Scientist. You need a Data Science team.

Jorge Lozano

jlozano1@Steelcase.com

LinkedIn linkedin.com/in/jorge-lozano-flores-72141414/

APPLIED
DATASCIENCE
Steelcase®

GREAT LAKES DATA & ANALYTICS SUMMIT

Thursday, March 12, 2020 | Metro Detroit

Thank you for attending!

@GLSummits

#AnalyticsSummit2020

Hosted by: