Great Lakes Analytics Summit 2024

Outsmarting Big Data: When more data is not the answer

Chris Smith

"These views are my own and not the views of my employer"

April 18, 2024

Chris Smith

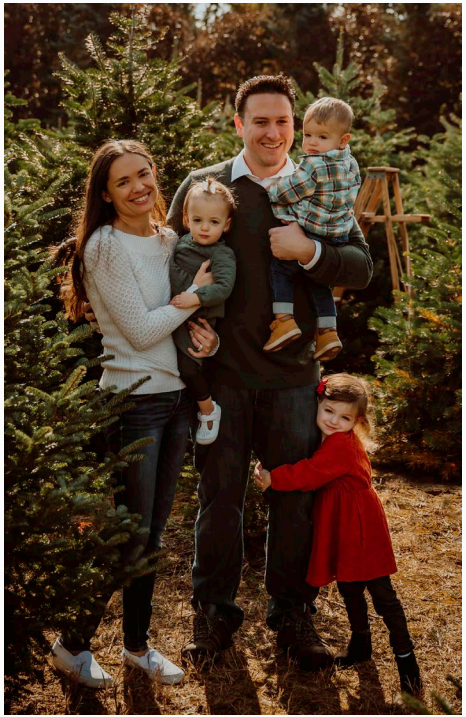
Education

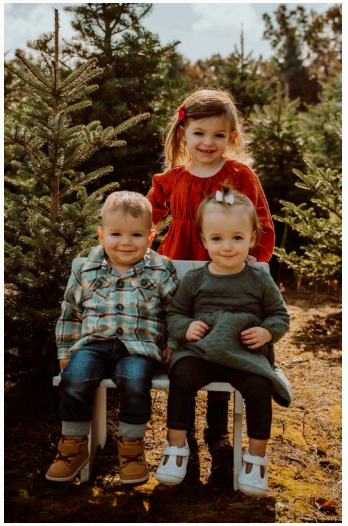
Lawrence Technological University

- Bachelor of Science in Computer Science, concentration in scientific software development
- Minor in Mathematics

Oakland University

- Masters in Business Administration (MBA), concentration in M.I.S.





Agenda

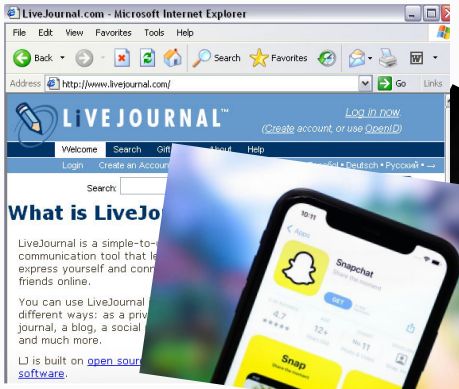
- **Outsmarting Big Data: When Shiny New Tech Is Not Always The Answer**
- **How did we get here**
  - A look at what led us here
- **Welcome to the now**
  - Bring that value
  - What is the new hype?
  - AI Overview
  - ML Overview
  - Analytics Overview
- **Where to go from here**
  - How to decide where to get started
  - Pick your battle and then pick your tech
  - Begin small, but with the end in mind

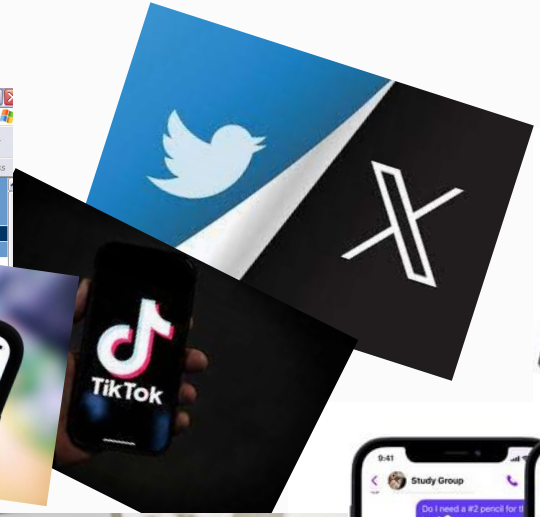
Intro

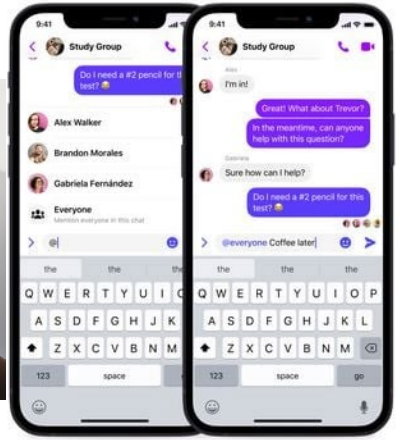
- 4MB of Ram
- Upgrades Include a MASSIVE 100MB Hard Drive

Fast Forward

- The **bloom** of new data



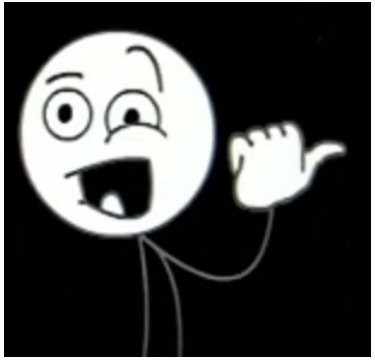


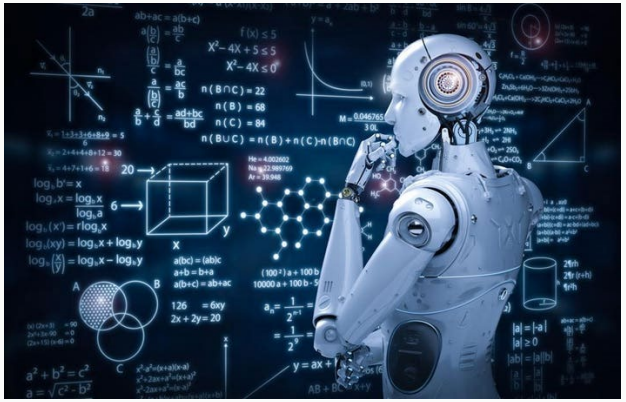


The Rise of Big Data

- In 2014, Big Data went mainstream
  - Specifically, by 2014, a Forbes study found that 70% of enterprise organizations had either deployed a solution for Big Data or were planning on deploying one in 2014

Demystifying AI & ML





What does this mean for Business?

- Unless your business is the technology itself, you care the most about….

- …not the analytics or the technology itself, right?

What does this mean for Business?

- Set goals that will help you both understand potential value and make sure you realize that value

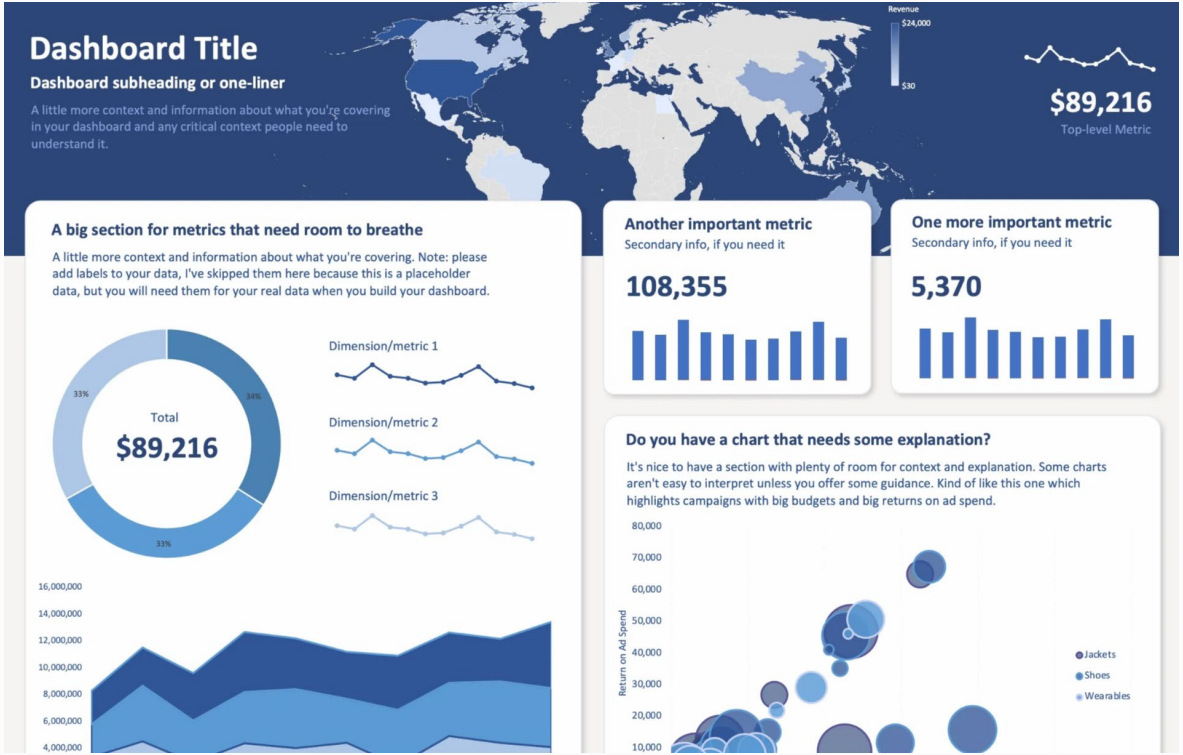
- Examples of how to think negative: **(Think Negative)** $^{-1}$

- We will not achieve or realize any value if:
  - We don't have a way to track newly acquired customers
  - We do not associate customer ID's to transactions
  - We don't ever get executive support
  - We don't store historical data long enough to see if our model improved over time
  - etc.

- Next, just invert your list:
  - Make sure you have a way to track newly acquired customers
  - Make sure you associate the customer ID to all transactions post launch
  - Figure out how to get executive support
  - Make sure to store historical data long enough

Let’s Play a Game!

- Guess the Tech!
- Type of Analytics:
  - Descriptive
  - Diagnostic
  - Predictive
  - Prescriptive
- Technology:
  - Name That Tool!

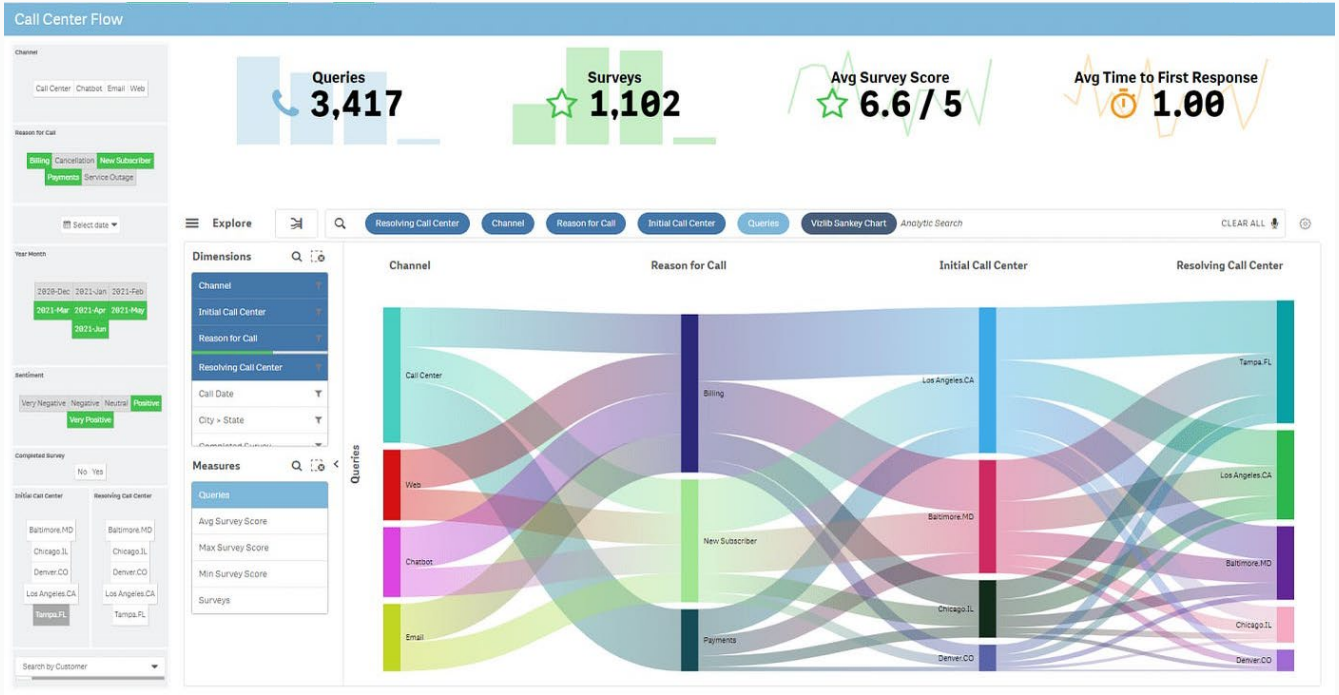


Let’s Play a Game!

- Guess the Tech!

- Type of Analytics :
  - Descriptive
  - Diagnostic
  - Predictive
  - Prescriptive

- Technology:
  - Name That Tool!

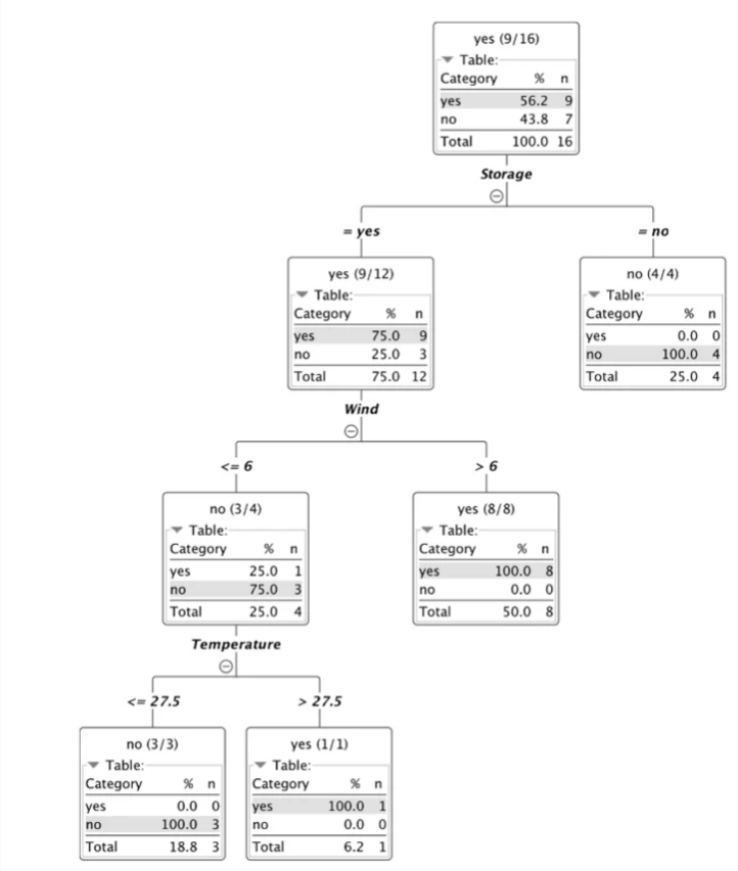


Let’s Play a Game!

- Guess the Tech!

- Type of Analytics :
  - Descriptive
  - Diagnostic
  - Predictive
  - Prescriptive

- Technology:
  - Name That Tool!

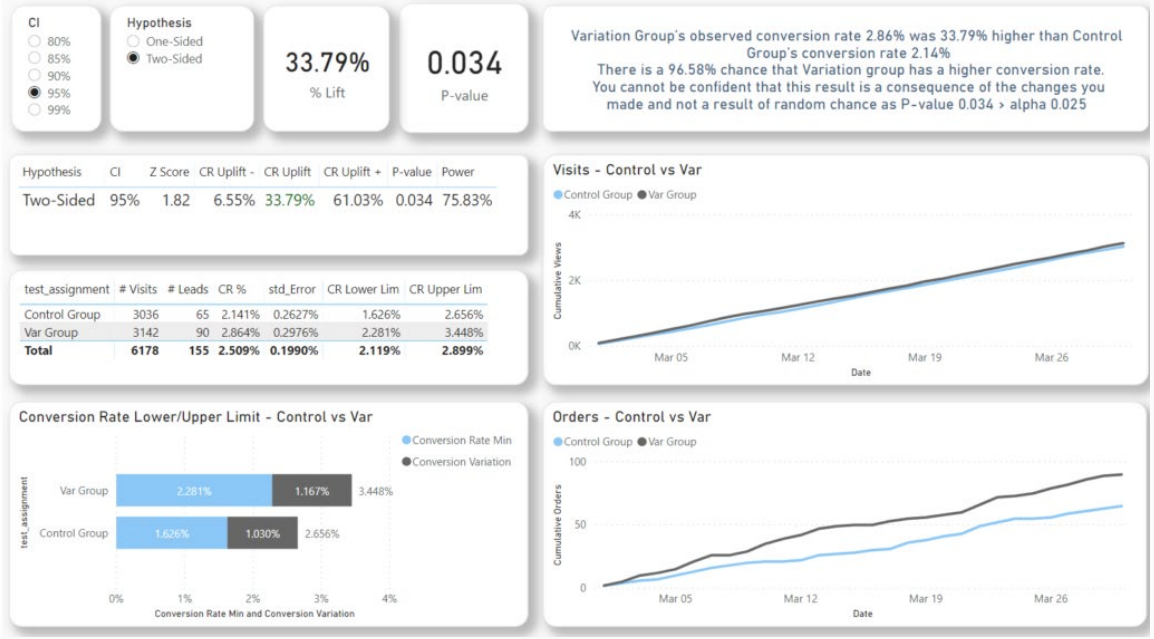


Let’s Play a Game!

- Guess the Tech!

- Type of Analytics :
  - Descriptive
  - Diagnostic
  - Predictive
  - Prescriptive

- Technology:
  - Name That Tool!



So what is AI?

- “AI is just making a computer act like a human”
- John McCarthy was one of the founding fathers of AI, and with Alan Turing developed the idea in 1956.
- Types:
  - Artificial narrow Intelligence
    - Netflix watch next recommendations
    - Natural Language Processing
    - Chatbots, face recognition, etc.
  - Artificial General Intelligence
    - IBM Watson, OpenAI CPT-3
  - Artificial Super Intelligence
    - Stuff for the movies (Iron Man’s “Jarvis”, or “I, Robot”)

This is where most of us play

So what is ML?

- Machine Learning uses Artificial Intelligence and is a subset of AI
- Example: email spam detection, medical or automotive diagnosis

- Types:
  - Machine Learning
    - Supervised & Unsupervised
  - Deep Learning
    - Subset of Machine Learning, where the model learns and improves in accuracy over time by examining computer algorithms that have been trained and tested
    - Involves artificial neural networks, which try to imitate human thought

When should AI be used?

- When you want to automate tasks that could be done by a human

- Healthcare:
  - Help lab technicians & doctors identify diseases
- Education:
  - Help teachers automate grading of papers and exams
- Robotics & Automation:
  - Robotic arm route detection, path/route adjustments in real time
- Banking & Finance:
  - Detecting anomalies in transactions & reducing fraud

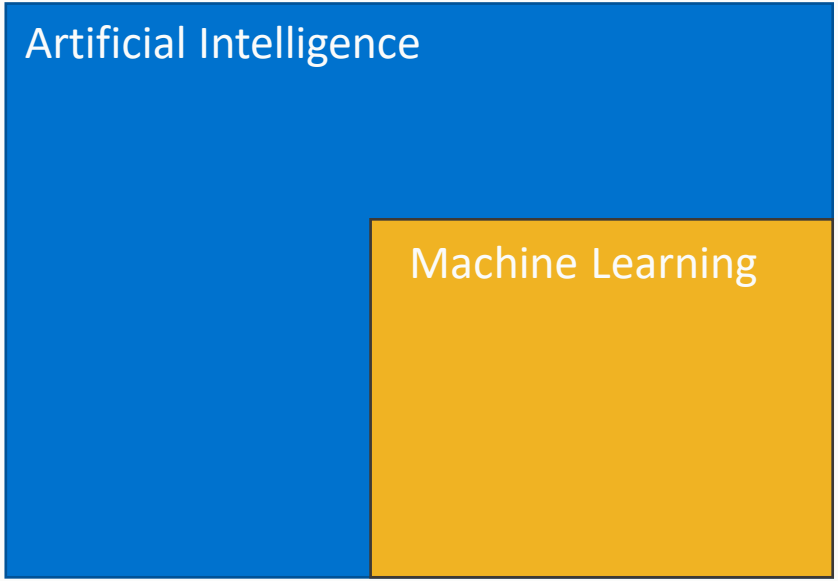
When should ML be used?

- When you have a hard thing to solve that involves a lot of data and a lot of variables

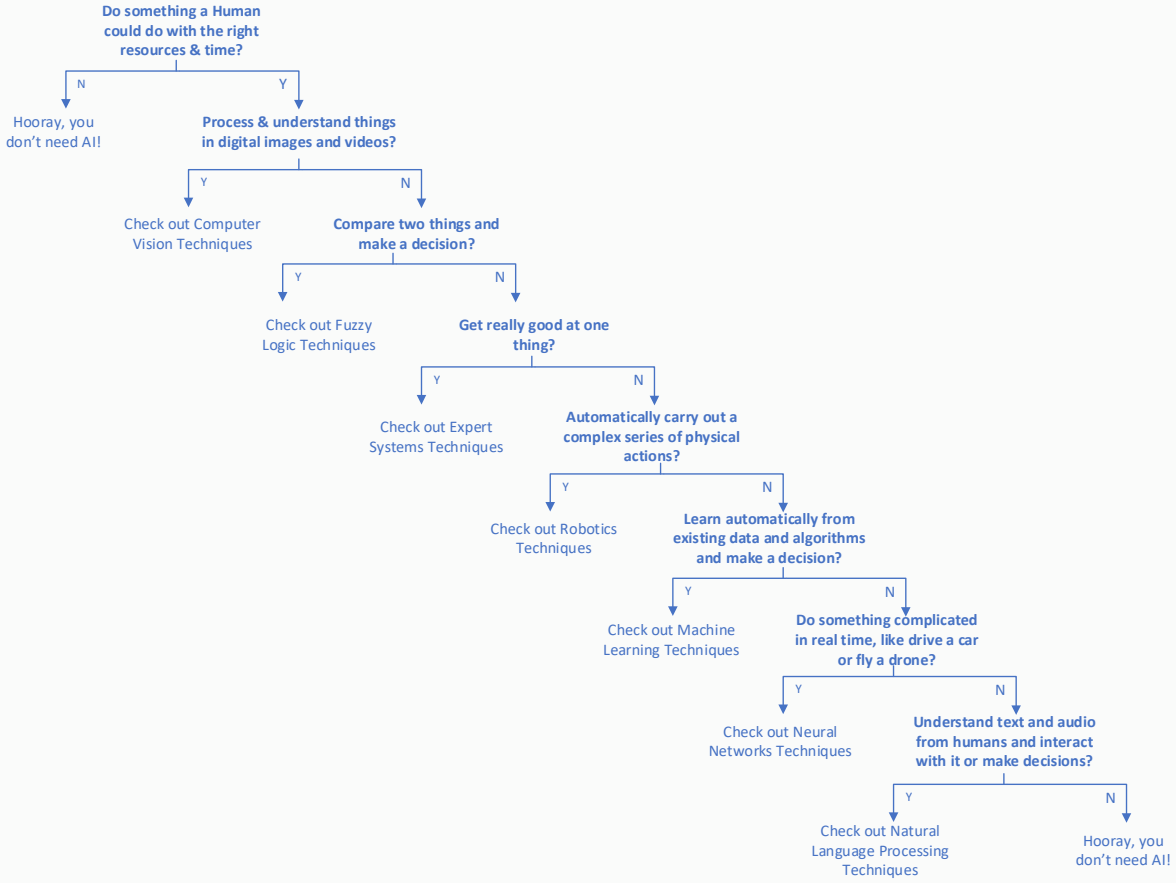
- A fancy form of statistical analysis
- Machine Learning is a capability. It gives predictions or decisions based on data.
- These results are commonly visualized using data visualization tools

When do you need shiny new tech?

- AI vs ML?
- AI = ML?
- AI <> ML?
- AI > ML?
- AI < ML?



How to decide where to start



Value, Value, Value

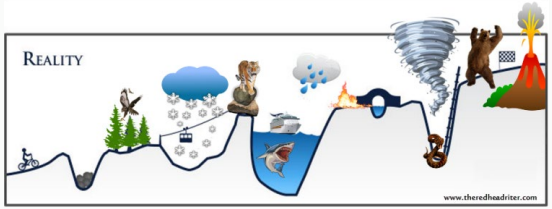
- “Pick your battles and *then* pick your tech”

- “If the model had just been 1% more accurate this project would have been a success”
  - Nobody

- **Most of the time, you don’t need a huge dataset or the latest technology to realize the value you need**
- …and even if you do, I challenge you to find a way to gain immediate value with what you have now, as a steppingstone, on your way to something greater.

Getting Started

- “Begin small but with the end in mind”
    1. Define the objective & value
    2. Gather Requirements
    3. Data Collection & Preparation
    4. **Explore – Do a POC, etc.**
    5. Form a Hypothesis
    6. Begin Building
    7. Evaluate Steps along the way
    8. Communicate your findings
    9. Iterate and Refine
    10. **Deploy**
    11. Repeat Steps 9 & 10



Thank You!