COSTCO WHOLESALE O

DATA ANALYTICS

Orientation for DAS
Orientation For Course
Beginning of course!

Matthew Morris
Git: DASTraining
msmorris@costco.com

Before Getting Started

Breaks

Housekeeping

Class Requirements

Support(Carts+)

ORIENTATION

DATA ANALYTIC SERVICES

Orientation Data Analytics Services

MARTHA PAL0MINO

- ➔ Operations Logistics Lead
- ➔ Classroom/ Course Logistics
- ➔ marthapalomino@costco.com

She's the one who sent your welcome emails and she's the Operations Manager for all Data analytic Services course logistics. She'll be cheering you on during the program, behind-the-scenes.

If you have any questions about program logistics (Certification, completion requirements, infosessions, office hours, rescheduling or about getting involved) - she's the one to go to. You can reach her via email.

Orientation Data Analytics Services

MATTHEW MORRIS

- ➔ Education Programs Lead
- ➔ Instructor Support
- ➔ msmorris@costco.com

Talk to this person about course content,
classes you would like to see. They may be in classes to assist the instructors from time to time.

INSTRUCTOR NAME - INSTRUCTIONAL LEAD

- Matthew Morris
- msmorris@costco.com
- DATA Analytics Enablement Analyst

Orientation Data Analytics Services

- Develop Data Analytic teams
- Special projects
- Stand up training
  - Training Materials
  - SQL
- Online Support/Content
- Ad-hoc(data pulls)
- Technical triage
  - Answer basic data usage questions
  - Source Systems
  - Data Flow/Applications
  - Diagnostics
- Tools

Assist user with fast turnaround Analytic Solutions and expedite long term solutions to appropriate departments

ORIENTATION

DATA ANALYTIC Roles and teams

Orientation Data Analytics Services

ROLES	DEFINITION OF ROLES	TOOLS	What They Can Do
Data Solutions Lead	Business modeling, Business lifecycles, Diagnostic reporting Predictive solutions	Webi Alteryx Tableau Lumira	Can publish reports to public folders Publish freehand SQL
Data Lead	Creates descriptive & diagnostic reporting Removes redundant reporting	Webi Alteryx Tableau Lumira	Can publish reports to public folders Publish reports with freehand SQL option
Data Analyst	Descriptive analytics Creates recurring and ad-hoc reporting	Webi Tableau Server Lumira	Can execute existing Alteryx reports Publish WEBI reports Create/Modify Dashboards on Tableau Server
Business User	Views reports	Webi Tableau Server	View reports

Orientation Data Analytics Services

BUSINESS USER

MANAGERS/EXECUTIVE TRAINING

RESPONSIBILITIES
- IDENTIFY PROBLEMS
- MAKE DECISIONS

TRAINING
- EXECUTIVE LEVEL OVERVIEW OF REPORTING/DECISION MAKING/STATS
- INTRO TO BUSINESS ANALYTICS ONLINE/STANDUP
- BUILDING AN ANALYTICS TEAM

Orientation Data Analytics Services

DATA ANALYST

RESPONSIBILITIES

- Creates/Maintains recurring and ad-hoc reporting using Webi, Lumira, Google sheets, Tableau server, Basic Statistics, Data visualization
- Applied Descriptive statistics
  - Similarity and differentiation comparisons
  - Summary statistics
  - Classification
- Helps identify redundant reporting and opportunities for automation
- Exploratory vs Operational

TRAINING

- WEBI SKILLS PORTAL
- WEBI FUNDAMENTALS- ADVANCED
- TABLEAU SERVER SKILLS PORTAL
- DATA ANALYTICS COHORT
- INTRO TO STATISTICS
- OPTIONAL(ANY INTRO CLASSES)

Orientation Data Analytics Services

DATA LEAD

RESPONSIBILITIES

- Publishes Data sets for Data analyst on Tableau Server
- Publishes Alteryx Workflows
- Identifies and removes redundant reporting
- Builds automation efforts
- Helps Business modeling efforts
    - Dimension, Measures and views
- Diagnostic Reporting
    - creating strong why statements
    - probabilities, likelihoods, and the distribution of outcomes
    - classification and regression
    - attribute importance, principle components analysis, sensitivity analysis, and conjoint analysis
- Assist with Special Projects

TRAINING

- Alteryx Fundamentals to Advanced
- Tableau Fundamentals to Advanced
- Github
- Statistics Fundamentals to Intermediate

Orientation Data Analytics Services

SOLUTION LEAD

RESPONSIBILITIES

- Business Modeling
  - Dimensions, Measures, Subject Areas
- Business Lifecycles
- Predictive Analytics
  - Classification, Regression, Clustering, Dimensionality Reduction
- Special Projects

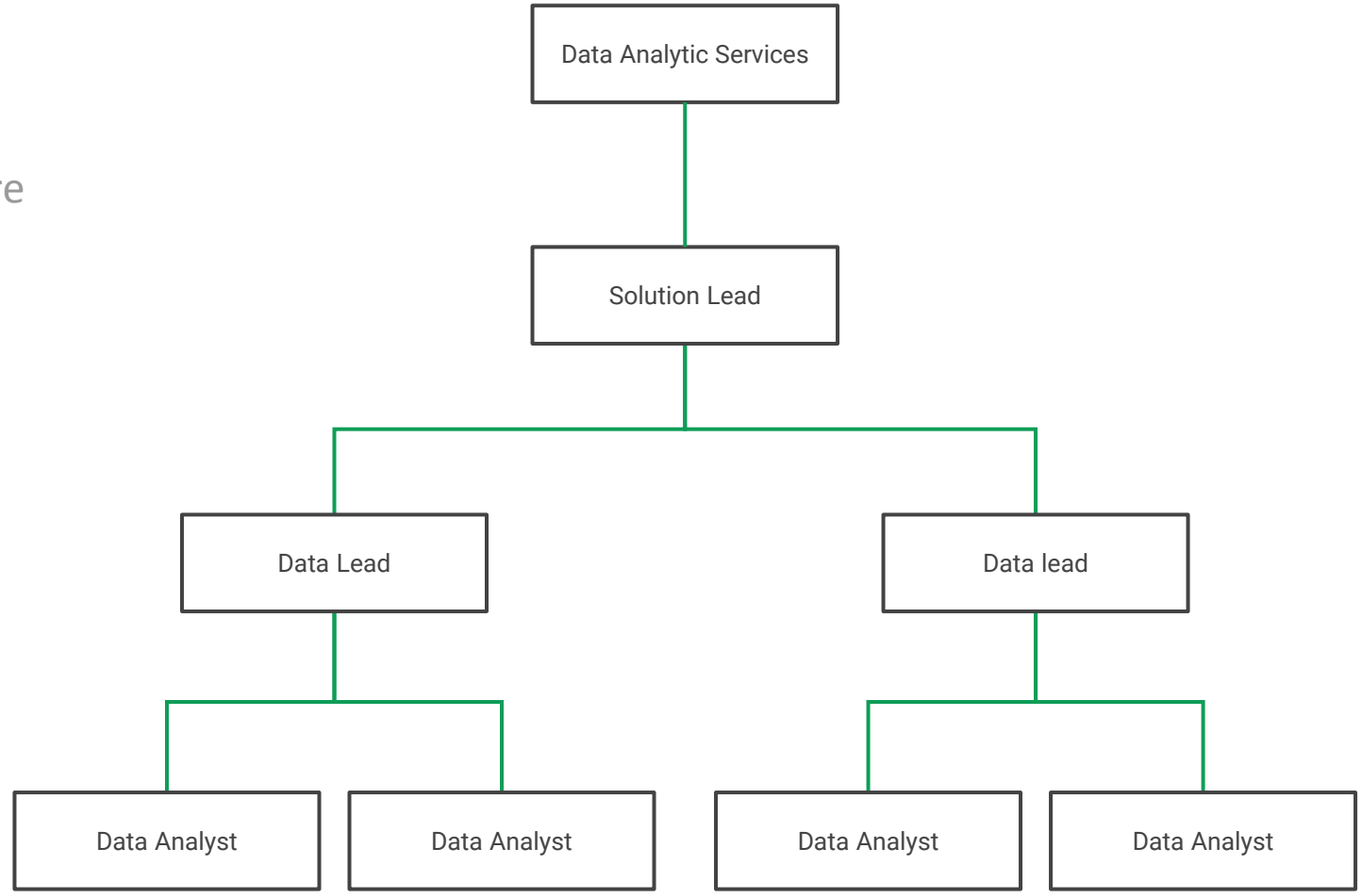
TRAINING

- Advanced Statistics
- R/Python
- Data Science

Orientation Data Analytics Services

Where Can I get help?

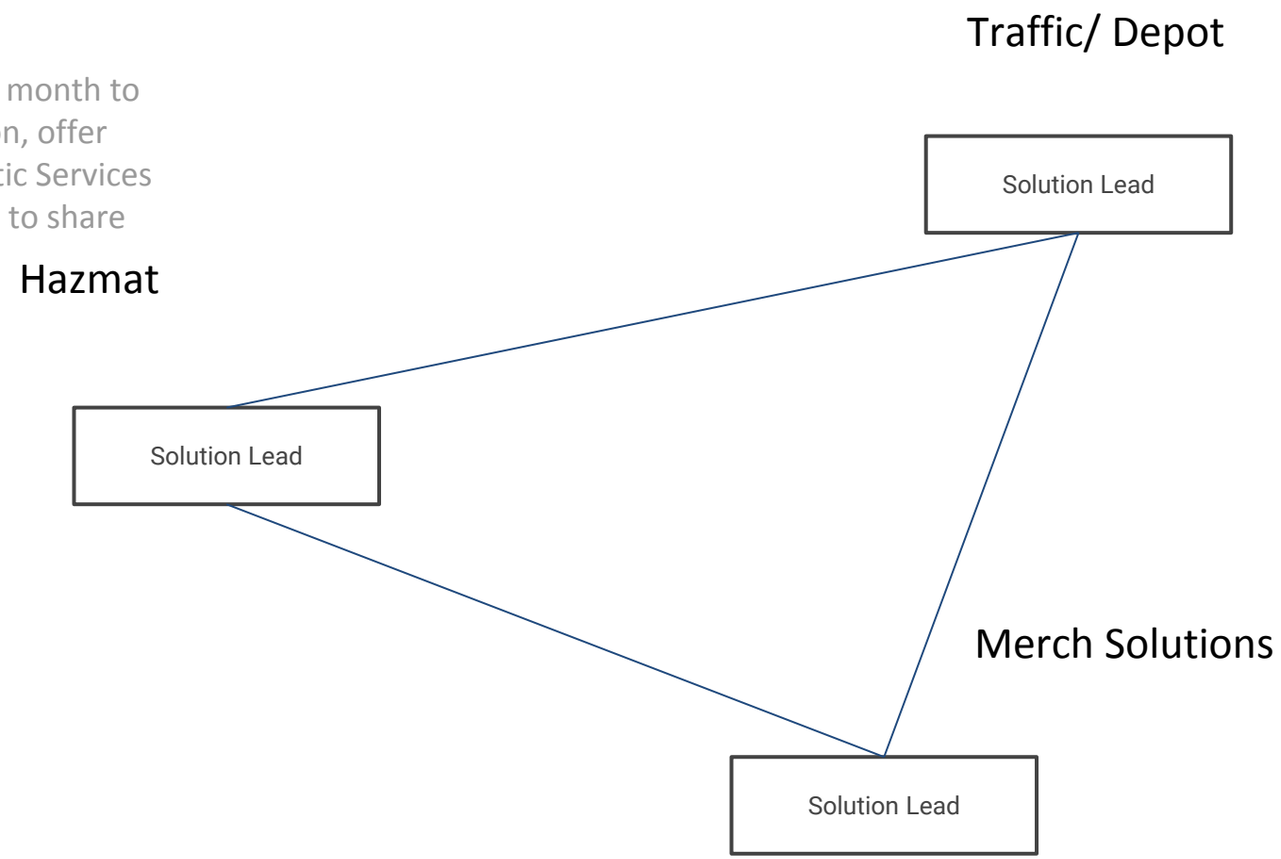
Before filling out a CARTs request or emailing anyone, Start with either talking to your Data Lead or your Solution lead if you do not know who talk to your manager or anyone in your area that works with lots of reporting. If you are a Solution or Data Lead network with your peers.



Orientation Data Analytics Services

SOLUTION LEADS SUPPORT NETWORK

For best results Solution leads should meet once a month to share tips, tricks, what projects they are working on, offer support and general knowledge share. Data Analytic Services is happy to have a representative attend meetings to share what is new, coming up or tips and tricks.



Orientation Data Analytics Services

How many Solution leads should I have in my Division or Department?
Different departments will have different needs.

Example A. You have a division with 90 people. 6 are technically savvy. It would not make sense to make them all Solution leads. Just one will work. The other 5 would become your data leads and then they would work with the rest of your staff to identify the needs for Data analyst.

In a large division like Accounting you may need additional Solution leads just for the sheer number of employees and volume of reports plus the division of the division.

Orientation Data Analytics Services

How do I get Alteryx or Tableau?

Directors approval and Data Analytics Service eval or your Solution leads eval.

Can I get Alteryx or Tableau without being a Data or Solution Lead?

No.

How do I become a Data or Solution lead?

Currently this is determined by your Director and management team. You should have extensive knowledge of SQL and support your departments reporting needs. You should also have the time to dedicate to the future state. Take courses and self teach in the areas needed to be a Solution lead.

ORIENTATION

DATA ANALYTIC COURSE

INTRODUCE THE PERSON NEXT TO YOU

- ➔ Name?
- ➔ What are they up to these days?
- ➔ Why are they taking this course?

- ➔ Something about them! Favorite show, Hobby, Activity you enjoy, Fun Fact…

Orientation for Course

- ➔ Be on time (i.e. early)
- ➔ Complete your presentations and submit your work
- ➔ Participate + ask questions
- ➔ Share with your peers
- ➔ Make friends :)

- ➔ You will **get out** what you **put in** to the class

ATTENDANCE

Let your instructor know if you will be absent

Make plans to catch up if you know you are going to miss class.

Certification eligibility requires not missing more than 6 hours of class time.

Orientation Data Analytics Course

DATA ANALYST CERTIFICATION

COMPLETE AND SUBMIT ALL PRESENTATIONS BY DEADLINES

SCORE 75% OR HIGHER ON CERTIFICATION TEST

Orientation Data Analytics Course

THREE PRESENTATIONS

Gathering requirements and understanding of the problem(5 min)

Status Update (Obtain, prep, initial analysis and recommendations)(5 min)

Recommendation and Next steps (10 min)

Orientation Data Analytics Course

Excel – Preparing, relating, normalizing and adding value to data

OracleSQL Developer – SQL Editor
(TRAINING ONLY DO NOT USE AT DESK CAN CAUSE SYSTEM ISSUES)

Webi –Drag and Drop tool used over prebuilt views

Alteryx – Self-Service Data Analytics. Repeatable workflow for self-service BI, data preparation, data blending, and advanced analytics.

Tableau – Connecting to data, telling stories with data, Tableau workflow

FEEDBACK FORMS

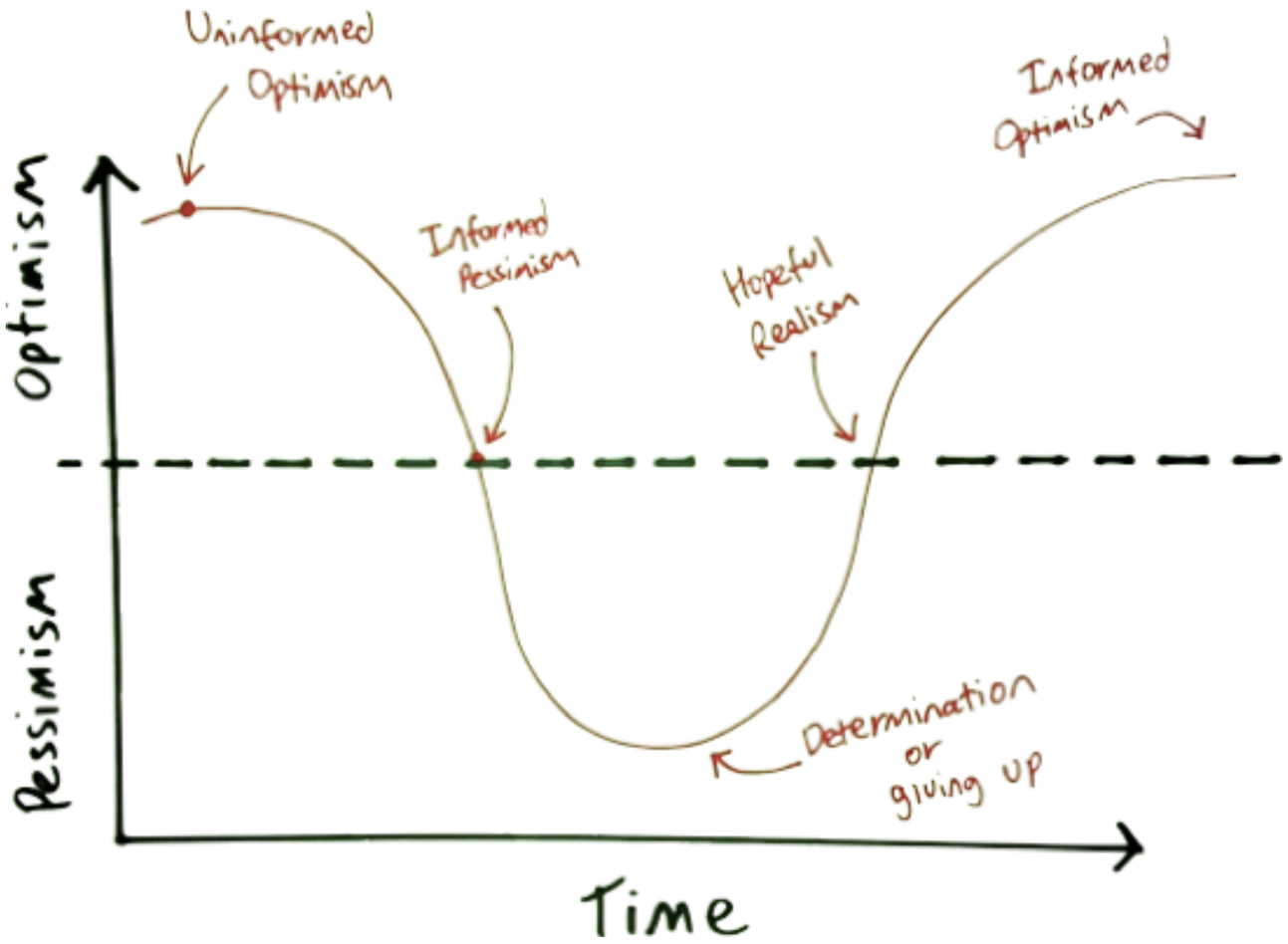
Let us know how it's going

Are you learning?
Is there something missing from the course?
Or is it awesome?

Feel free to share with our team reviews the feedback to look for ways to improve or to keep things that are working.

TIPS FOR SUCCESS

Orientation Data Analytics Course



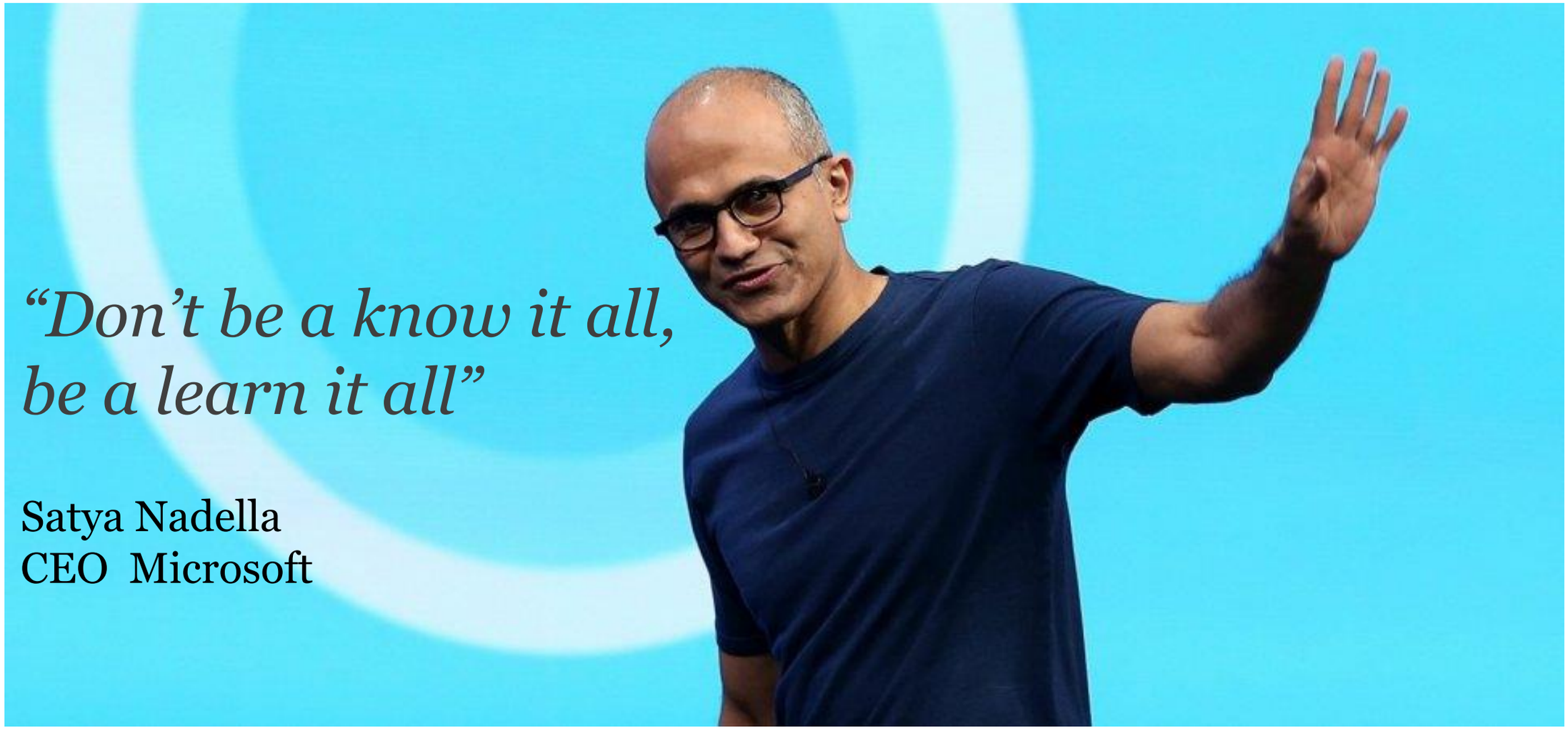
TIPS FOR SUCCESS

WORK WITH EVERYONE IN THE ROOM

TEACH OTHERS BACK IN YOUR DIVISION

BOOK OFFICE HOURS

Orientation Data Analytics Course



Orientation Data Analytics Course

"I dedicated 1 to 3 hours a day, determined to master the material"

DATA ANALYTIC SERVICES - TRAINING

DATA ANALYST AND THE DATA ANALYTIC WORKFLOW

Data Analytic Summary
Iseries Options Source
Validating Data

Matthew Morris
Git: Morrisdata
Msmorris@costco.com

WHAT IS A DATA ANALYST?

Data

Facts and statistics collected together for reference or analysis

Things known or assumed as facts, making the basis of reasoning or calculation.

Analysis

Detailed examination of the elements or structure of something, typically as a basis for discussion or interpretation.

Data Analyst

Professional whose focus of analysis and problem solving relates to data, types of data and relationships among data elements within business or IT systems.

So What does all of this mean to you?

Data Analyst

Daily Activity Summary

File Edit View Insert Format Data Tools Add-ons Help Last edit was yesterday at 5:12 AM

5/19/2014

	A	B	C	D	E	F	G	H
1	5/19/2014	10013	9	2067	90	feet	4.89	miles
2	5/20/2014	16291	18	2415	180	feet	7.52	miles
3	5/21/2014	9659						
4	5/23/2014	10269						
5	5/24/2014	16400						
6	5/25/2014	13323						
7	5/26/2014	11858						
8	5/27/2014	13235						
9	5/28/2014	9686						
10	5/29/2014	14052						
11	5/30/2014	11270						
12	5/31/2014	17764						
13	6/2/2014	12253						
14	6/3/2014	13622						
15	6/11/2014	9089						

	A	B	C	D	E	F	G
1	Date	Weight	BMI	Fat	Resting Heart Rat	Normal Heart Rat	Exertive Heart Rate
2	2016-06-01	170.56	24.47	14.388	0	0	0
3	2016-06-02	169.86	24.37	14.555	0	0	0
4	2016-06-03	169.17	24.27	14.722	0	0	0
5	2016-06-04	168.47	24.17	14.889	0	0	0
6	2016-06-05	167.78	24.07	15.056	0	0	0
7	2016-06-06	167.08	23.97	15.223	0	0	0
8	2016-06-07	166.39	23.87	15.39	0	0	0
9	2016-06-08	165.69	23.78	15.557	0	0	0
10	2016-06-09	165	23.68	15.724	0	0	0
11	2016-06-10	164.31	23.58	15.894	0	0	0
12	2016-06-11	164.31	23.58	15.894	0	0	0
13	2016-06-12	164.31	23.58	15.894	0	0	0
14	2016-06-13	164.31	23.58	15.894	0	0	0
15	2016-06-14	164.31	23.58	15.894	0	0	0
16	2016-06-15	164.31	23.58	15.894	0	0	0
17	2016-06-16	164.31	23.58	15.894	0	0	0
18	2016-06-17	164.31	23.58	15.894	0	0	0
19	2016-06-18	164.31	23.58	15.894	0	0	0
20	2016-06-19	164.31	23.58	15.894	0	0	0
21	2016-06-20	164.31	23.58	15.894	0	0	0
22	2016-06-21	164.31	23.58	15.894	0	0	0
23	2016-06-22	164.31	23.58	15.894	0	0	0
24	2016-06-23	164.31	23.58	15.894	0	0	0
25	2016-06-24	164.31	23.58	15.894	0	0	0
26	2016-06-25	164.31	23.58	15.894	0	0	0
27	2016-06-26	164.31	23.58	15.894	0	0	0
28	2016-06-27	164.31	23.58	15.894	0	0	0
29	2016-06-28	164.31	23.58	15.894	106	174	185
30	2016-06-29	164.3	23.57	15.894	111	174	182
31	2016-06-30	162.32	23.29	15.894	0	0	0

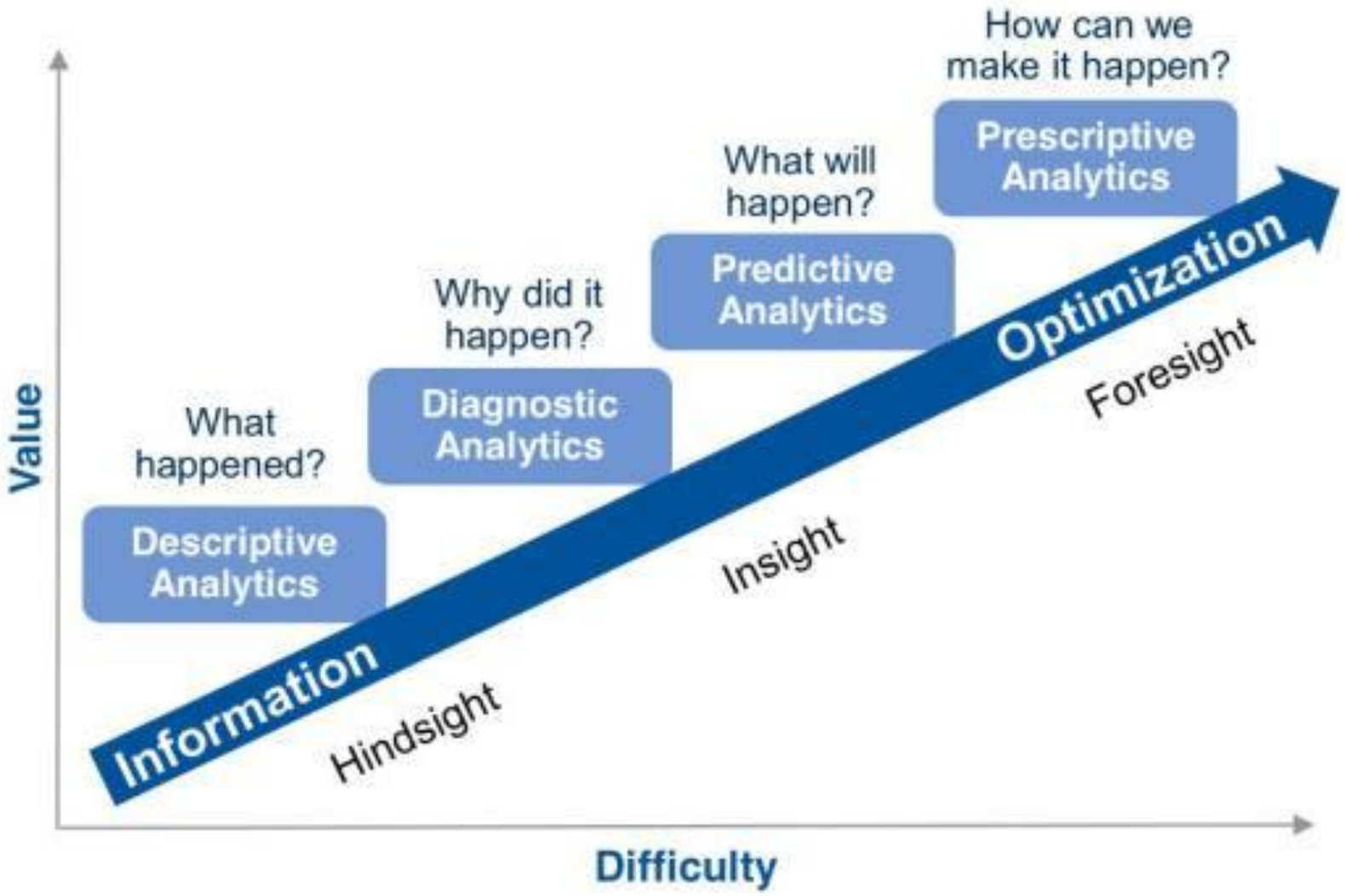
Data Analyst

There are a lot of questions.

How
Who
Where
What
When
Why

We develop those questions and find answers.

Data Analyst



Orientation Data Analytics Course

IDENTIFY THE PROBLEM

Identify the problem

Here are some examples of questions you can use during the inquiry phase to enhance your understanding of the situation:

"What are these reports trying to solve?"

"What concerns you the most about _____?"

"What actions do these reports provoke?"

"What seems to be your main obstacle to _____?"

"What makes your boss/customers angry enough to contact you?"

"What is contributing to lost deals or opportunities?"

"What would make your report more appealing?"

What series of spreadsheets or reports do you look at to make decisions?

What are some of the more complex decisions you are trying to make?

"What will you have to do to get the job done?"

"Is there something I should have asked that you need me to know?"

"How do you want _____ to turn out?"

"What do you want?"

"What is your desired outcome?"

"What benefits would you like to get out of X?"

"What do you propose?"

"What reports link with your goals this period/quarter/year?

"If you do this, how will it affect _____?"

"What else do you need to consider?"

Identify the Problem

REVIEW CURRENT REPORTS

- Why are fields copied and pasted
- Why are some numbers highlighted?
- Are their tick marks, Why?
- What would they see on the report and add to a slide deck or make a phone call about?
- Are there reports that seem to answer a similar question but from different angles?
- What are the dimensions, measures, calculations and filters

Identify the Problem

Here is a thin framework that is meant to be developed changed and built out to support your overall problem statement:

- We believe **X**. We're examining **Y** to determine if it a) supports the claim, b) refutes the claim or c) is inconclusive.

OR

- We are hoping to accomplish **X**. We're examining **Y** as measured by **Z** to determine if there appears to be a relationship/opportunity.

UNDERSTAND THE DATA

Understand the Data

Steps for creating an analytic summary (Understanding your Data)

1. Bring in field descriptions(VLOOKUP, HLOOKUP, INDEX MATCH)
2. First assumptions about data types
3. Basic statistics (COUNT, COUNTA, COUNTB, MIN, MEAN, MEDIAN, MODE, MAX, SUM)
4. Pivot Summary
5. Review data and ask How, Who, Where, When, Why
6. Use filters to explore answers
7. Look for blanks, outliers and data that needs to be prepped

Understand the Data

VLOOKUP HLOOKUP, INDEX, and MATCH

Often considered advanced topics, and knowing how to use them effectively can make you stand out in the world of Excel users.

Understand the Data

VLOOKUP (lookup_value, table_array, col_index_num, [range_lookup])

Value you're searching for

Table searching for value in

Data to retrieve from searched Table

TRUE = Closest
FALSE = Exact

Understand the Data

=MIN(number1, [number2]…)

- Finds the minimum value of a range of numbers.

=MAX(number1,[number2]…)

- Finds the maximum value of a range of numbers.

=SUM()

- Adds a range of numbers.

=AVERAGE()

- Adds all numbers in a range and divides by the number of values.

Understand the Data

=COUNT(value1,[value2]…)

- Counts the number of numeric values in a range

=COUNTA(value1, [value2]…)

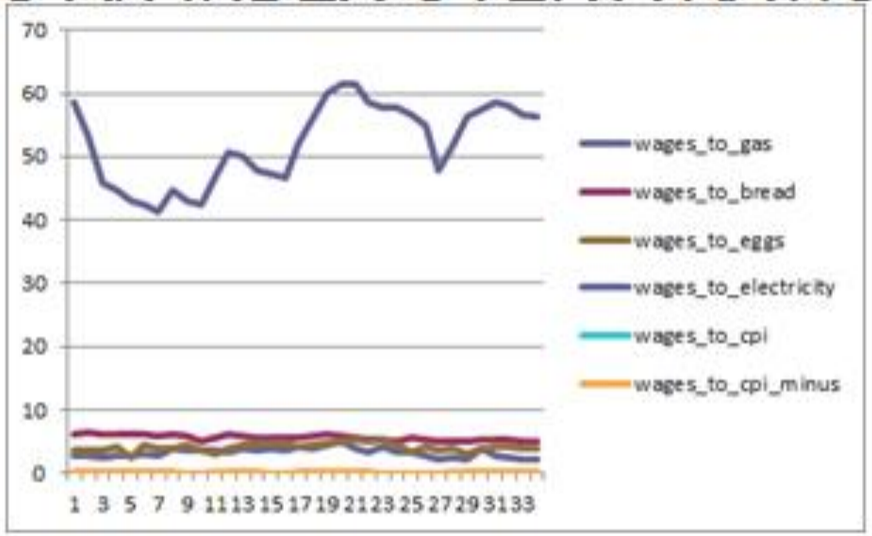
- Counts the number of alpha values in a range

=COUNTBLANK(range)

- Counts the number of blanks in a range

Understand the Data

CREATING AN INDEX OVER A RATIO



Understand the Data

IELD	DATA TYPI	DATA FOR	COUNT	COUNTA	COUNTBL	SUM	MIN	MA
OMPANY	Dimensio	int	109430	109430	0			
VHS5	Dimensio	int	109430	109430	0			
AME	Dimensio	char	0	109430	0			
EGN	Dimensio	char	0	109430	0			
DDRESS	Dimensio	char	0	109430	0			
ITY	Dimensio	char	0	109430	0			
TA	Dimensio	char	0	109430	0			
TRY	Dimensio	char	0	109430	0			
IP	Dimensio	int	109430	109430	0			
REA	Dimensio	int	109430	109430	0			
BDATE	Dimensio	date	109430	109430	0			
EPT	Dimensio	int	109430	109430	0			
M NO	Dimensio	int	109430	109430	0			
DA	Measure	double	105536	109430	0	2876917	0	46619
	Measure	double	105371	109430	0	21629235	0	97254.34

Analytic Summaries are the best place to begin any journey into a new data set.

Understand the Data

Review of MIN, MAX, SUM, AVERAGE, COUNT, COUNTIF COUNTA, COUNTIFS, COUNTBLANKS to summarize dat sets.

A fantastic quick deliverable!

EXERCISE: BUSINESS SCENARIO

Exercise: Business Scenario

UNDERSTAND THE DATA

- Read any documentation with the data
- Perform Exploratory Analysis
- Verify the Quality of the data

You have been given the INWITMP file and have been asked to create an analytic summary:

You will need to import the description

Make general assumptions about data types

Run general stats

Look for outliers

Review your data, ask questions and use filters to answer those

Questions.

Run some counts and visualize the data

Resources

A thorough guide to the steps of data cleansing:
www.siop.org/tip/backissues/Jan05/PDF/423_089to096.pdf

Numerous Data sets
www.kaggle.com

Q & A

“Data! Data! Data!
I can’t make bricks without clay. “

-Sir Arthur Conan Doyle, Author

Conclusion

- Identifying a problem will dictate the value of the answer
- Just cleaning data and creating percentages can add immense value to a dataset. Prepping data is a major part of data analytics.

PROJECT

Project

Prompt

You have been hired by Costco to review all of the reports in your department and identify what problems they are trying to solve to help make better decisions faster.

Project

THE TOOLS

Identify the Problem

- What are the objectives?
- What would measure success?
- How do you know you have the right data set?
- How will this affect change or action?

Obtain the Data

- Identify the right tables
- Export Sample data for summary
- Determine the best tools to work with the data

Understand the Data

- Read Any Documentation provided with the data
- Perform Exploratory Analysis
- Validate the quality of the data

Project

Business needs as per your interpretation of the scenario;

Data selected from the original file;

Cleaning methods used to remove erroneous data;

<u>Format:</u> Google Slides or PDF (Keynote/PPT need to be exported);

FEEDBACK FORM

CLASS : FUNDAMENTALS OF DATA ANALYTICS

QUESTION: What are the steps of the Data Analytics workflow?