



Introduction to DSLA Methods

WELCOME TO BOOTCAMP IN R

ROHIT PADEBETTU



Our Team



Cynthia Correa



Rohit Padebettu



Linesh Dave



Suresh Nambiar



Prachi Bedi Nambiar



Jyoti Pulimamidi

Agenda for the Day

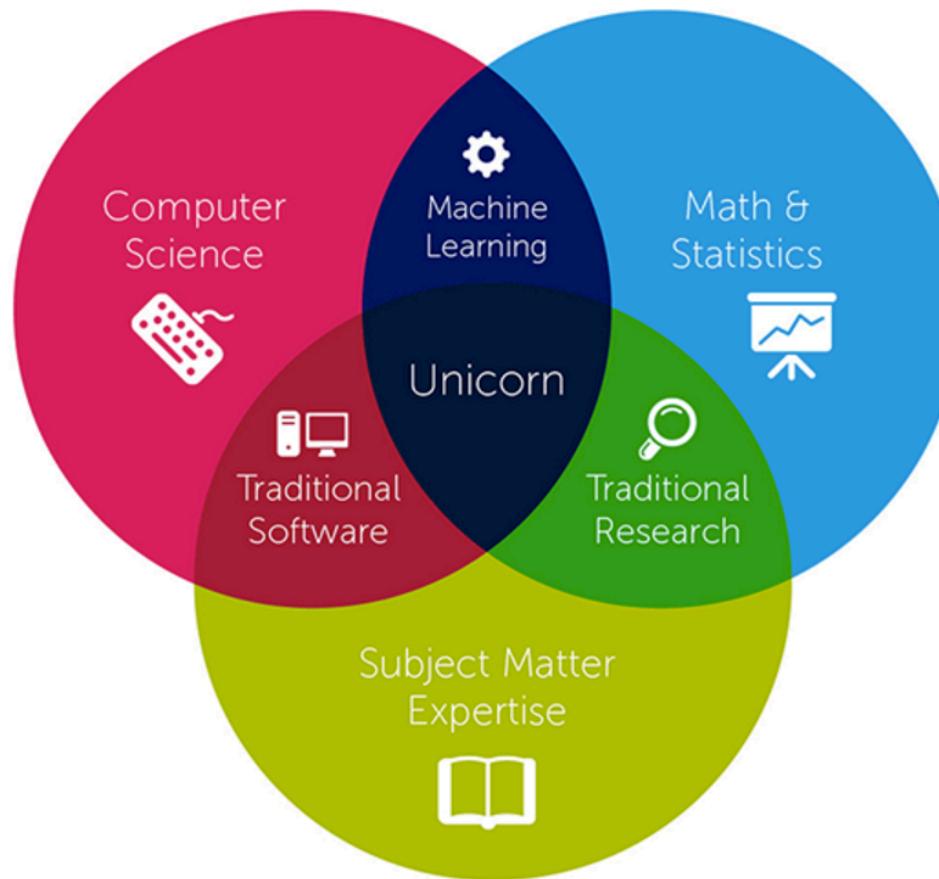
Morning Session

- Introduction to DSLA
- Overview of Data Science
- Student Introductions
- DSLA Methods
- Short Break
- Course Structure Discussion
- Assignments

Afternoon Session

- Introduction to Machine Learning
- Introduction to Linear Regression
- Short Break
- Instructor Case Demo
- DSLA Website & Tools Workshop

Who is a Data Scientist?



What do Data Scientists do?

- **Obtain the data**
- **Clean the data**
- **Explore the data**
- **Seek to understand the data**
- **Build models to explain the data**
- **Predict using the models**
- **Report the results**

Student Introductions

1. Name
2. What attracted you to data science?
3. What is your professional background?
4. Which of the target skills will challenge you most?
5. Which one comes most naturally to you?
6. Hobby
7. Name

Walk to the front of the
class

Current state of Data Science

Many good online courses

Many university offered classroom courses

Many more private classroom courses

Lots of freely available material online too !!

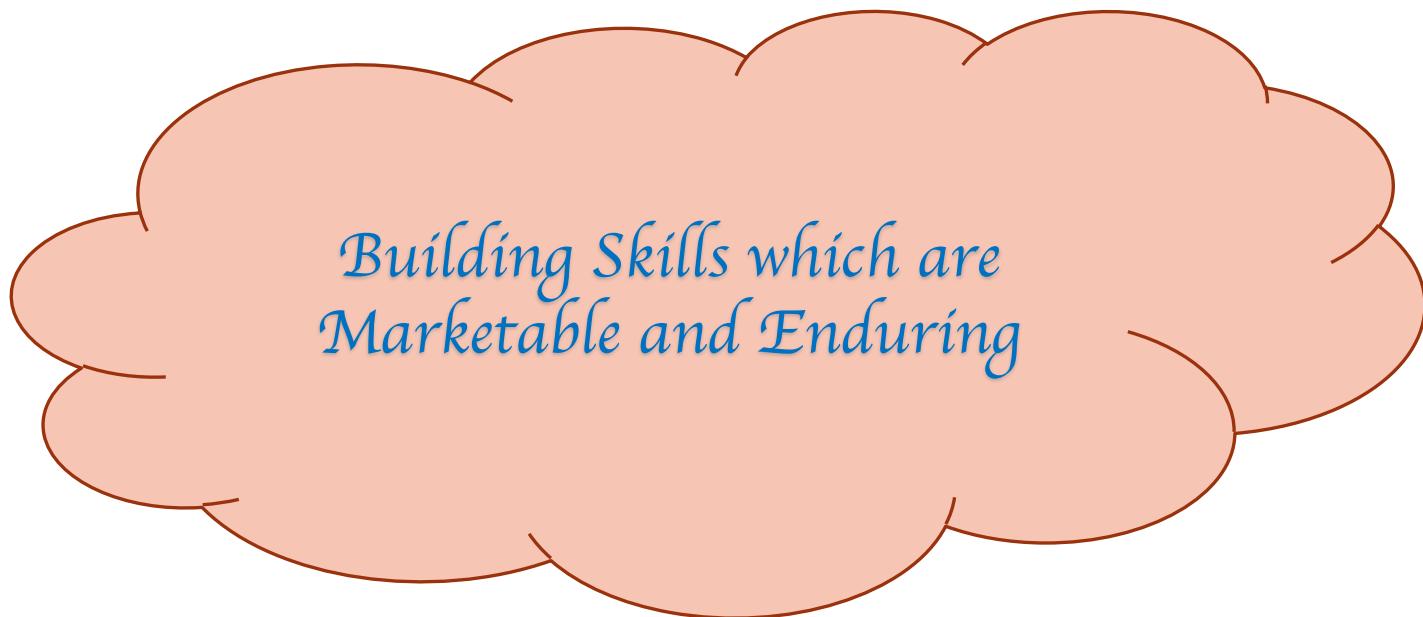
Most of these focus on

- Technical tools
- Heavy in Statistics and Math
- Almost always work on clean, tidy data

All this is necessary, but not sufficient!

DSLA Courses

Our Goal



*Building Skills which are
Marketable and Enduring*

Student Introductions

1. Name
2. What attracted you to data science?
3. What is your professional background?
4. Which of the target skills will challenge you most?
5. Which one comes most naturally to you?
6. Hobby
7. Name

Walk to the front of the
class

Target Skill # 1

Conceptualization



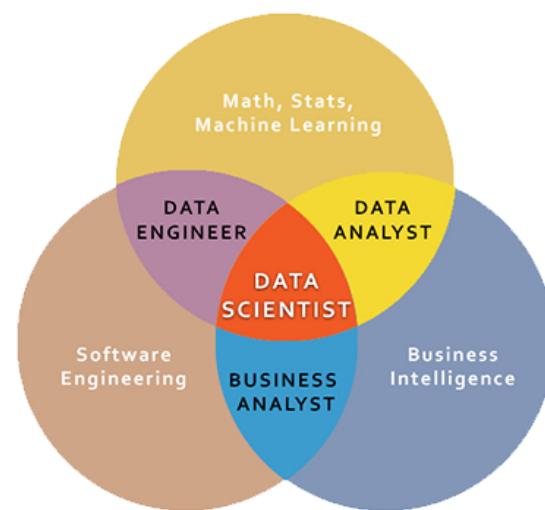
Target Skill # 2

*Creative Problem Solving
&
Divergent Thinking*



Target Skill # 3

Effective Collaboration



IF YOU WANT TO GO FAST,
GO ALONE.
IF YOU WANT TO GO FAR,
GO TOGETHER.

- African Proverb

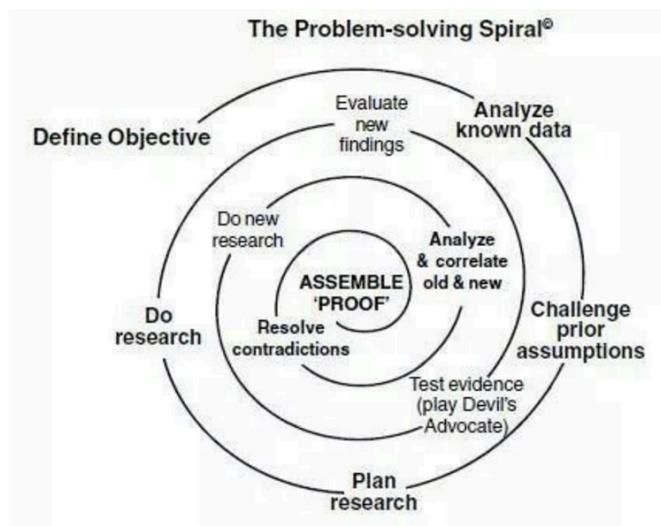
Student Introductions

1. Name
2. What attracted you to data science?
3. What is your professional background?
4. Which of the target skills will challenge you most?
5. Which one comes most naturally to you?
6. Hobby
7. Name

Walk to the front of the
class

Target Skill # 4

Iterative Problem Solving



Target Skill # 4

Iterative Problem Solving



Target Skill # 5

Persuasive Communication



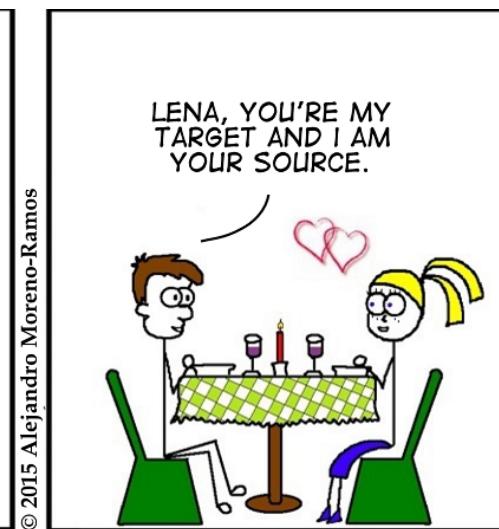
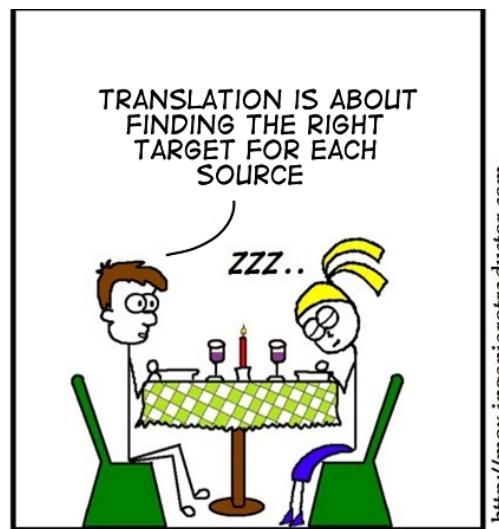
Student Introductions

1. Name
2. What attracted you to data science?
3. What is your professional background?
4. Which of the target skills will challenge you most?
5. Which one comes most naturally to you?
6. Hobby
7. Name

Walk to the front of the
class

Target Skill # 6

Translation Analytics to Business



Target Skill # 7

Selling your Ideas



It's not about you.

Decision makers aren't interested in your pain.

THEY'RE INTERESTED IN THEIR PAIN.

They want to know how your idea will ease their pain.
Solve their problem. Provide worry-free sleep.

Or maybe they want to know how the idea will make
their lives fun and joyful. Make them prosperous.
Make life easier.

The last thing they want to hear are your problems.
The overtime you put into the idea. Your hassles along
the way. Your sleepless nights and supreme sacrifices.

Don't whine or complain during your pitch.
Keep it positive.

STAY OUT OF YOUR PROBLEMS.
STAY IN THEIR SOLUTIONS.

Student Introductions

1. Name
2. What attracted you to data science?
3. What is your professional background?
4. Which of the target skills will challenge you most?
5. Which one comes most naturally to you?
6. Hobby
7. Name

Walk to the front of the
class

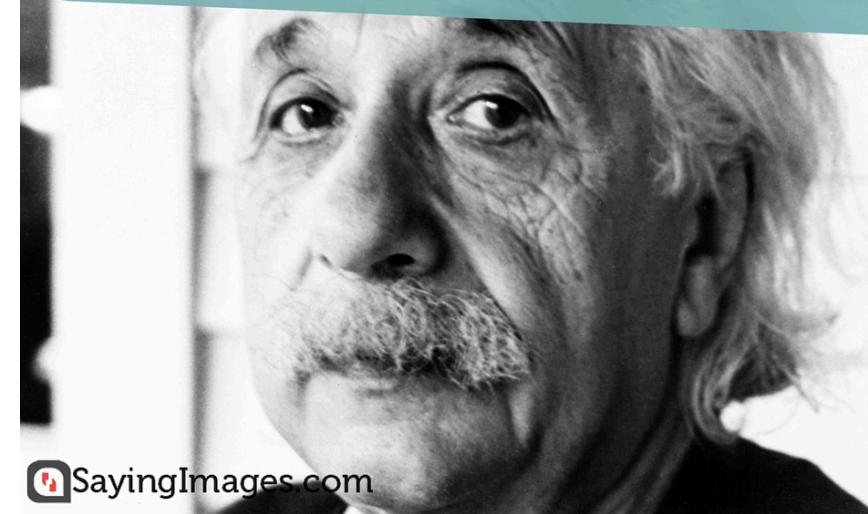
Training Approach

Team Oriented

Learn by Doing

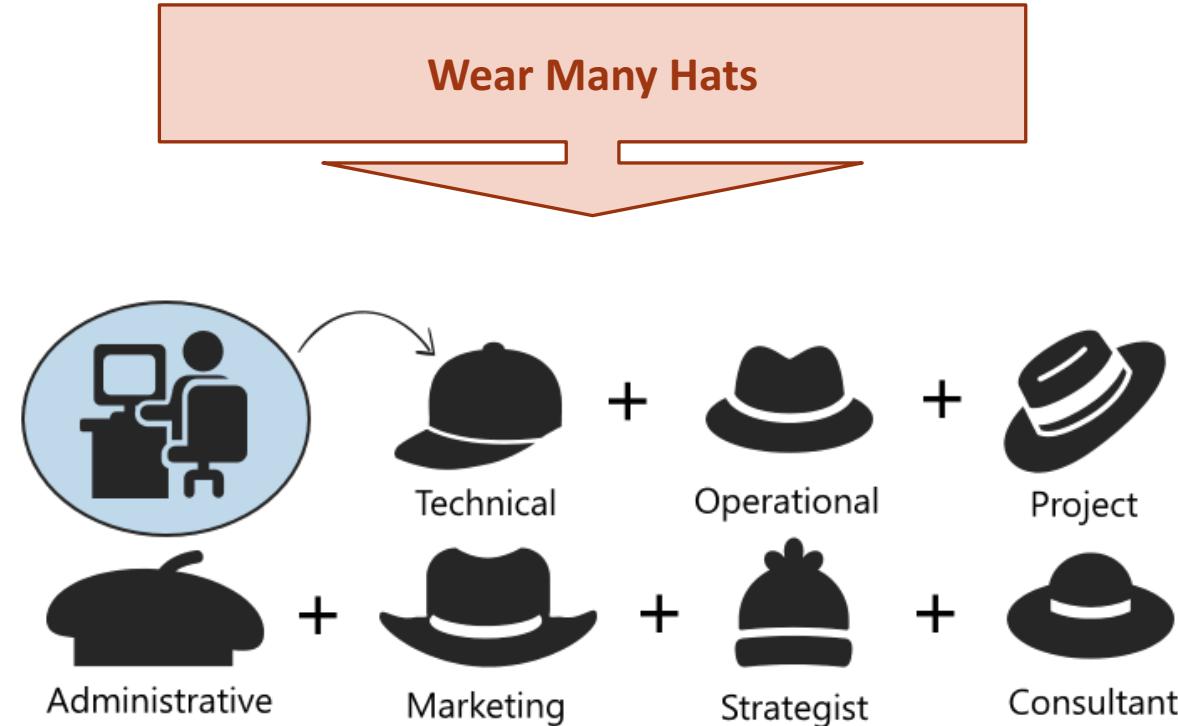
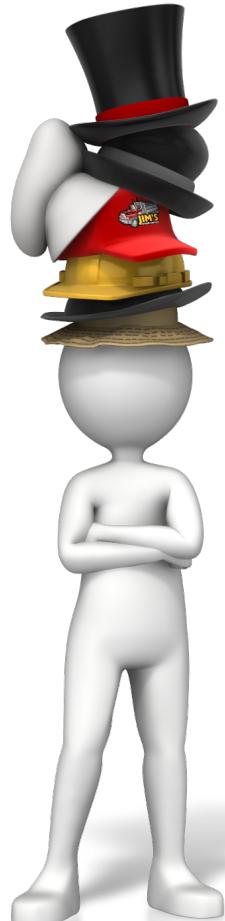
Social Learning

*Education is not the learning
of facts, but the training of
the mind to think.*

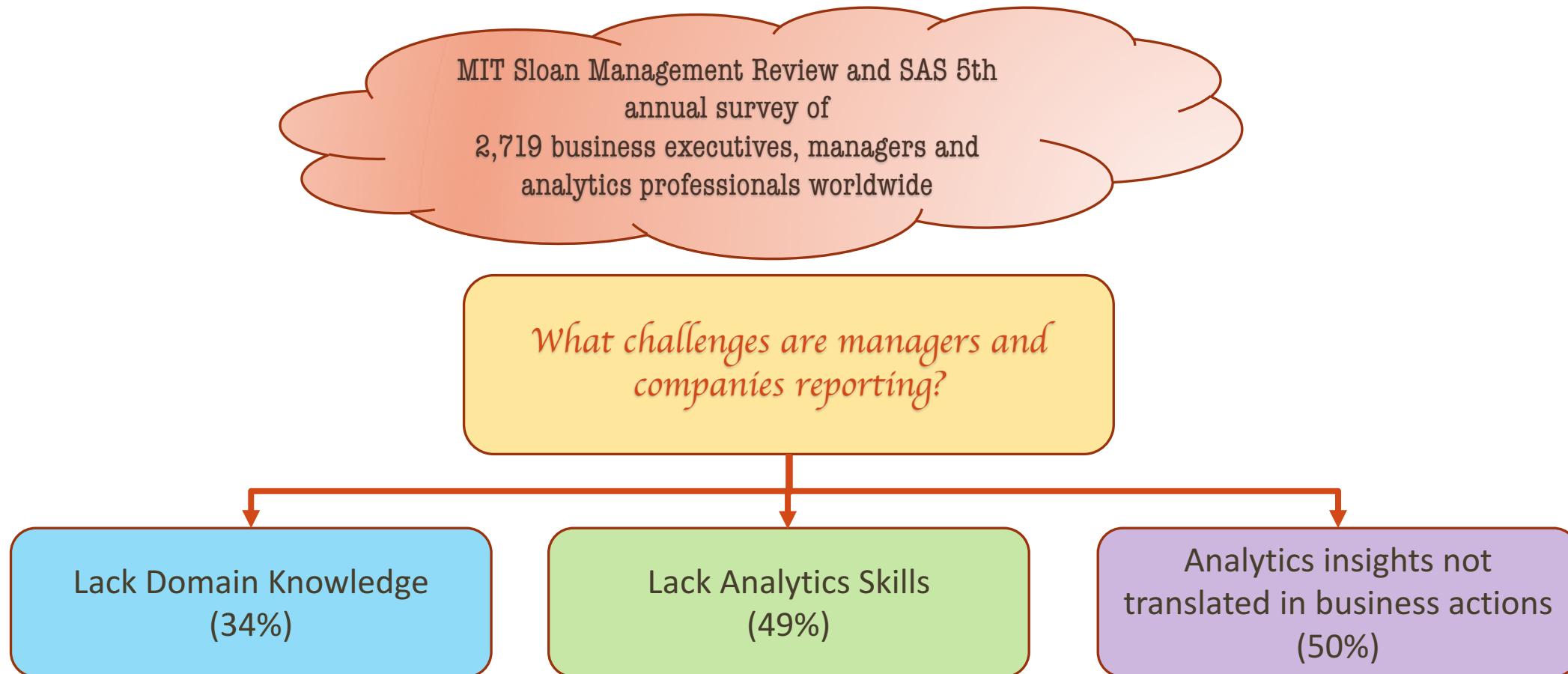


 SayingImages.com

Your Role



Recent Survey



We help you *Develop*

Conceptualization

*Creative Problem Solving
&
Divergent Thinking*

Effective Collaboration

Iterative Problem Solving

Persuasive Communication

*Translation
Analytics to Business*

Selling your Ideas



We help you *Prepare*

*Build a Portfolio of
Projects to showcase*

*Hone Technical Skills
in R, Python, Big
Data and others*



*Develop Effective
Interview Skills*

*Practice Soft skills such as
Report Writing,
Presentation, Project
Management, Problem
Solving*



10 minute break

You'll need an access card to
re-enter the glass doors

Problem Solving

Problem Solving Skills

Is there a magic formula for Problem Solving ?

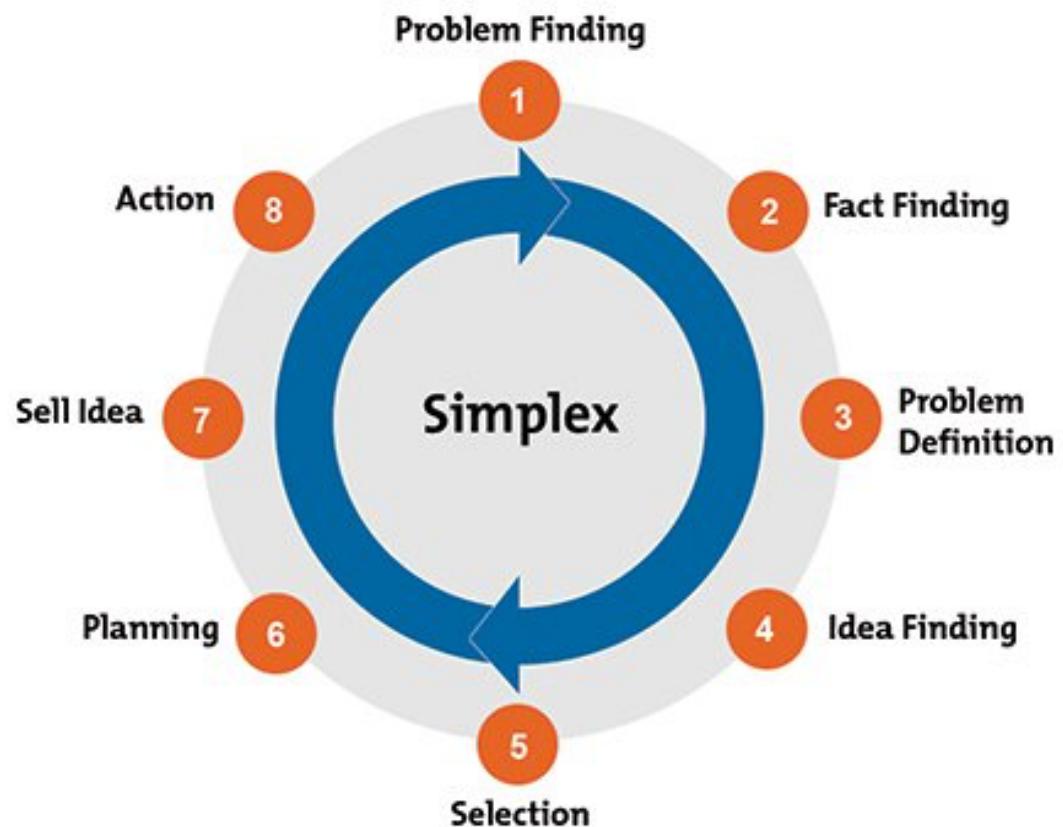
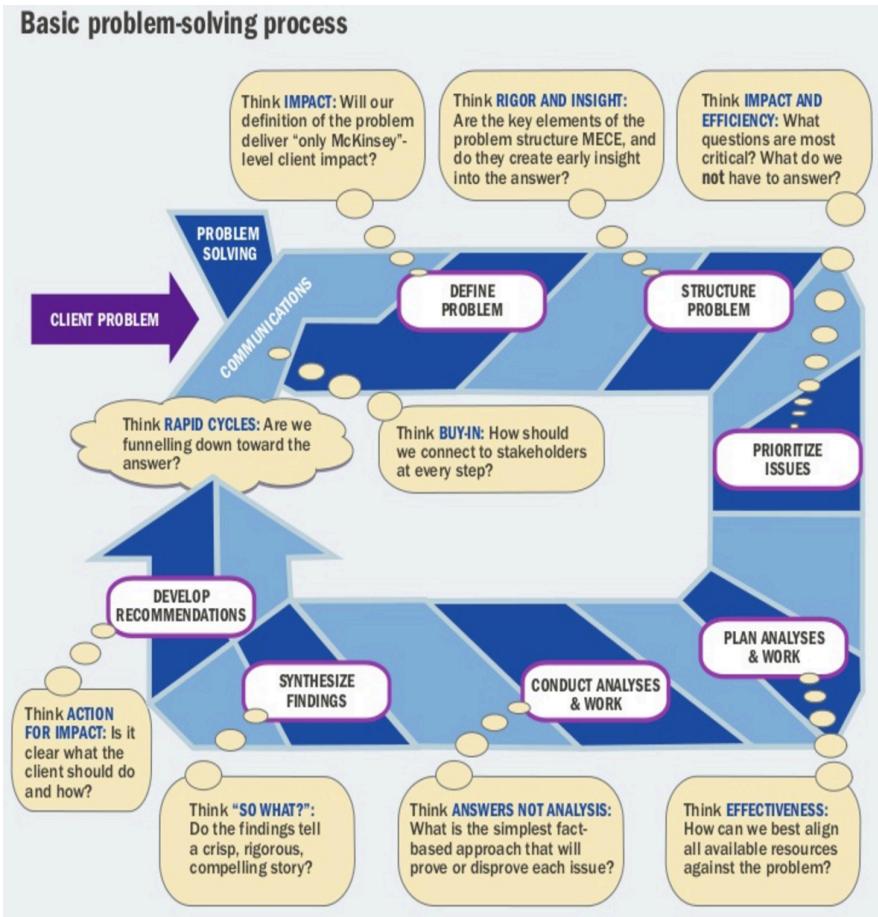


or...

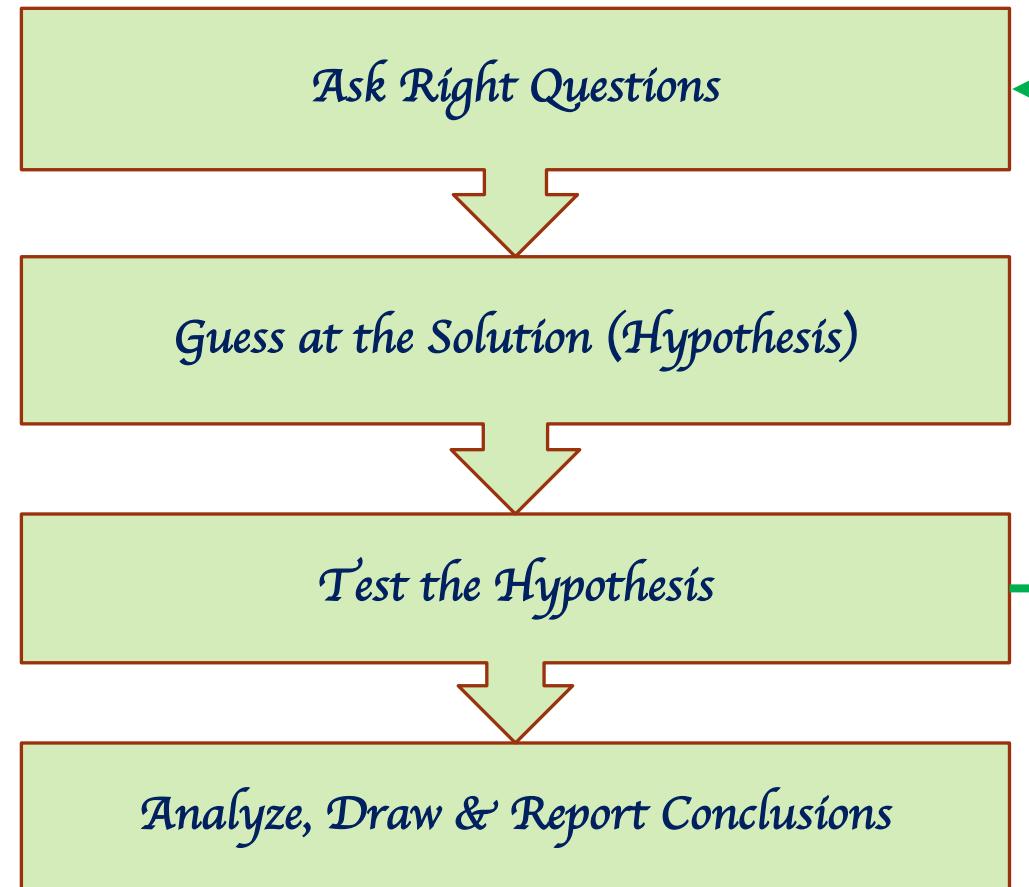


Problem Solving Process

Basic problem-solving process



Basic Problem Solving Framework



Course Structure

Week 1	
Morning	Introduction to DSLA course
Afternoon	Instructor Lead Case: MTcars (Excel)
Week 2	
Morning	Instructor Lead Case: Simple Classification (Rstudio local)
Afternoon	Intro to Group Case: Classification (Classification Case)
Week 3	
Morning	Presentation of Group Case: Classification
Afternoon	Intro to Instructor Lead Case: Mnist digit Classification (Rstudio AWS/Server) (Image Processing)
Week 4	
Morning	Data Science Theory - Concepts
Afternoon	Intro to Group Case: Sentiment Analysis (NLP, Using APIs, Exploratory Analysis)
Week 5	
Morning	Presentation of Group Case: Sentiment Analysis
Afternoon	Instructor Lead Case: Big Data Spark- Predictive Analytics (Machine Learning)
Week 6	
Morning	Instructor Lead Case: Big Data Spark- Predictive Analytics (shiny)
Afternoon	Intro to Group Case: Big Data Demand Analysis and Predictive Analytics (Hadoop, Hive, Spark)
Week 7	
Morning	Presentation of Group Case: Demand Analysis and Predictive Analytics (GUI)
Afternoon	Intro to Elective Group Case: Dynamic Pricing/ Default Predictions/ Recommendation Systems
Week 8	
Morning	Interview Questions/Concepts
Afternoon	Presentation of Elective Group Case: Dynamic Pricing/ Default Predictions/ Recommendation Systems
Week 9	
Morning	Individual Case
Afternoon	Individual Case
Week 10	
Morning	Individual Case Presentation
Afternoon	Individual Case Presentation

4 or 5 Instructor Led Cases

3 Common Group Projects

1 Elective Group Project

1 Individual Project

Course Assignments

Programming Assignments

Reading Assignments

Presentation Assignments

Technical Skills Assignments

Writing Assignments

Technical Assignment

Install & Familiarize yourself with R and RStudio IDE

Programming Assignment

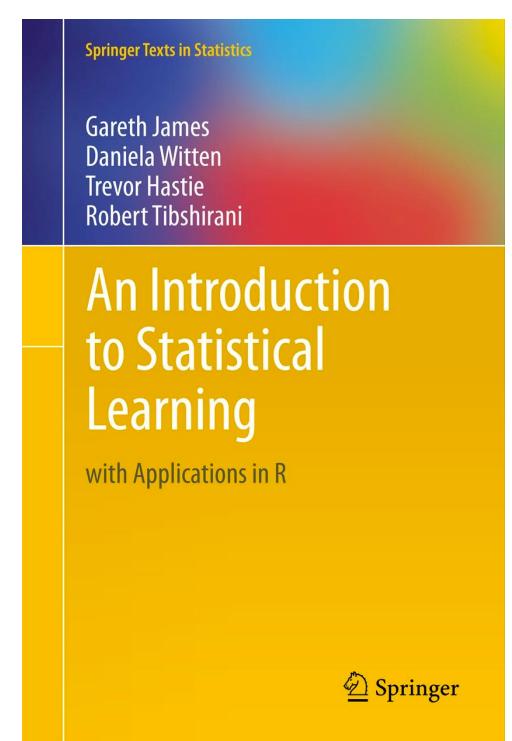
Install & Complete: Git / GitHub Tutorial

Install & Complete: Swirl - R Programming

Reading Assignment

*Read Chapter 1: Introduction
Read Chapter 2: Statistical Learning*

*Read Chapter 3: Linear Regression
Read Chapter 4: Classification*



Writing Assignment

*Create & Submit by Wednesday
Aspirational Resume*

What do you want your Resume to contain in 3 months ?

- Two pages maximum
- List 3 projects or jobs
- Explain the problem
- Describe mathematical or statistical techniques used
- Specific Metric you Improved

Presentation Assignment

By Tuesday Submit

1. 2 Industries hiring data scientists where you would like to work
2. 2 Types of data science problems they solve
3. Identify the Algorithms typically used to solve those problems

By Saturday Submit

3 examples of data science problems in your personal/professional life.

For each give

1. Relevant Question
2. Hypothesis
3. Data Source
4. Algorithm

Course Curriculum

Using R programming for machine learning | DSLA New Jersey 6/24/17, 11:10 PM

Who can benefit from this course:

- Anyone interested in learning more about R programming, machine learning, and data science.
- This course is suited for people who have never programmed as well as people already familiar with programming.
- Anyone interested in the rapidly expanding world of data science.

Course structure R course

		Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10
Morning	Introduction to DSLA course										
Afternoon	Instructor Lead Case: Mtcars (Excel)										
Morning	Instructor Lead Case: Simple Classification (Rstudio local)										
Afternoon	Intro to Group Case: Classification (Classification Case)										
Morning	Presentation of Group Case: Classification										
Afternoon	Intro to Instructor Lead Case: Mnist digit Classification (Rstudio AWS/Server) (Image Processing)										
Morning	Data Science Theory - Concepts										
Afternoon	Intro to Group Case: Sentiment Analysis (NLP, Using APIs, Exploratory Analysis)										
Morning	Presentation of Group Case: Sentiment Analysis										
Afternoon	Instructor Lead Case: Big Data Spark- Predictive Analytics (Machine Learning)										
Morning	Instructor Lead Case: Big Data Spark- Predictive Analytics (shiny)										
Afternoon	Intro to Group Case: Data Mining and Predictive Analytics (Hadoop, Hive, Spark)										
Morning	Presentation of Group Case: Demand Analysis and Predictive Analytics (GUI)										
Afternoon	Intro to Elective Group Case: Dynamic Pricing/ Default Predictions/ Recommendation Systems										
Morning	Interview Questions/Concepts										
Afternoon	Presentation of Elective Group Case: Dynamic Pricing/ Default Predictions/ Recommendation Systems										
Morning	Individual Case										
Afternoon	Individual Case										
Morning	Individual Case Presentation										
Afternoon	Individual Case Presentation										

COURSE CURRICULUM

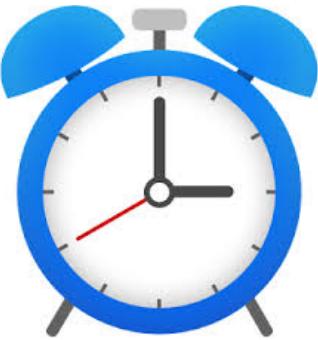
Pre-Week1 Preparation

Install R Studio	00:00:00
Programming Assignment: Complete Git/Github Tutorial	00:00:00
Reading Assignment: Introduction to Statistical Learning – Chapter 1	00:00:00
Optional Reading Assignment: Introduction to Statistical Learning – Chapter 2	00:00:00

Week 1

Philosophy of course and our vision	00:00:00
Student Introductions	00:00:00

<https://dsla1.com/course/r-programming-machine-learning/> Page 2 of 8



60 minute Lunch break

You'll need an access card to
re-enter the glass doors