



**Masanao Yajima,**

Director of MSSP Consulting

Associate Professor of the Practice,  
Department of Mathematics & Statistics,  
Boston University

# Statistics Practicum: Placing 'Practice' at the Center of Data Science Education

Joint work with



[Eric Kolaczyk](#) and



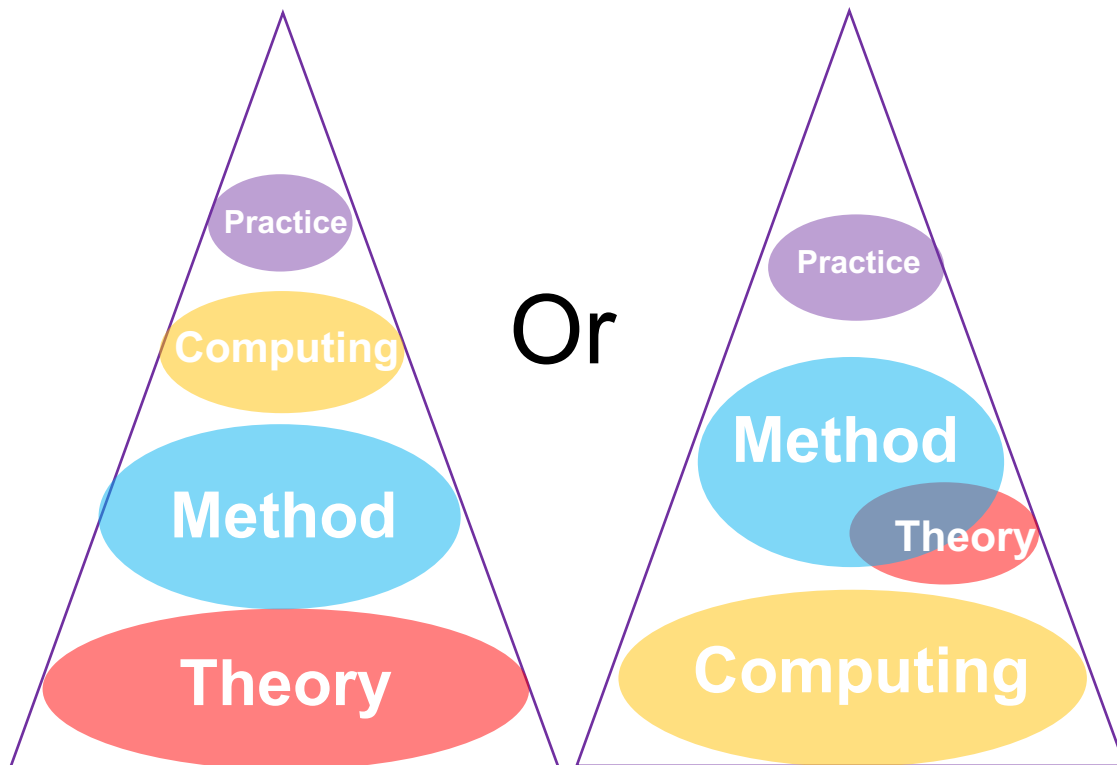
*Haviland Wright.*

## Background (USA)

- *Data Scientists have been in high demand.*
- *However, most companies want “experienced” Data Scientists.*
- *Massive Open Online Course (MOOC) did not prove to be the solution*
- *Lack of practical experience is one of the concern.*
- *Many Data Science Programs has flourished in the past 3-5 years...  
(500 Universities and 830 Programs)*
  - *Applied Statistics*
  - *Data Science*
  - *Business Analytics*
- *Etc.*

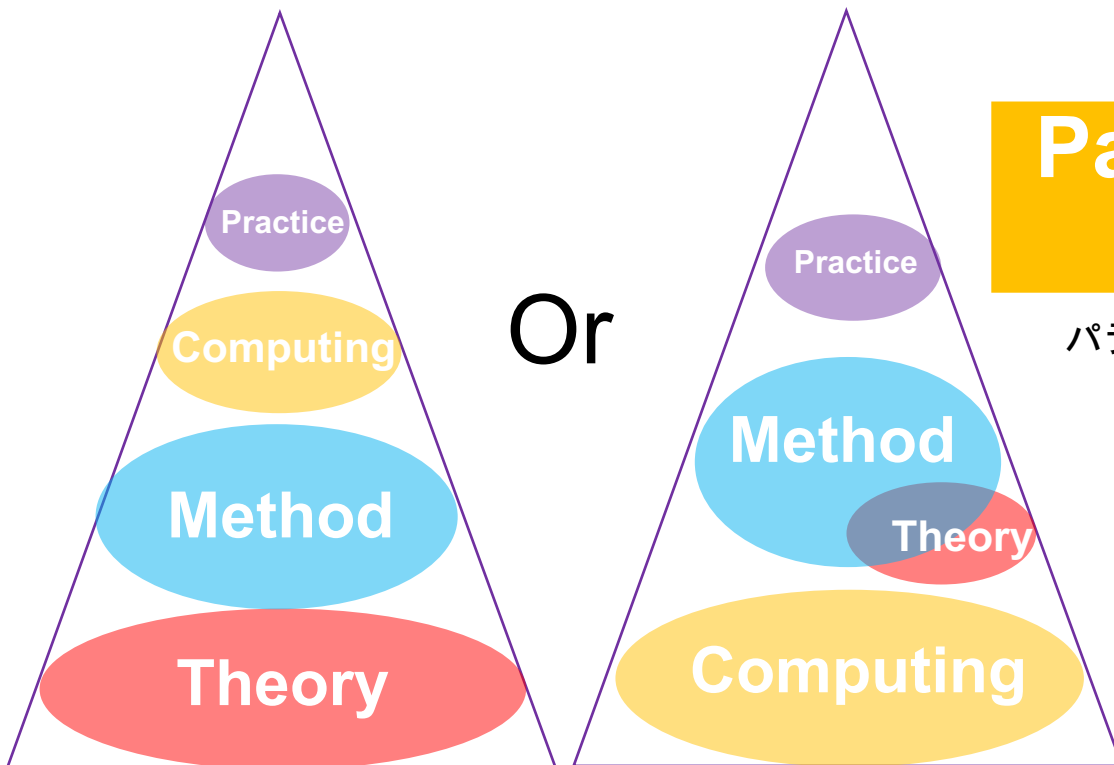
# Data Science (DS) Education Program Design

- Traditional DS program



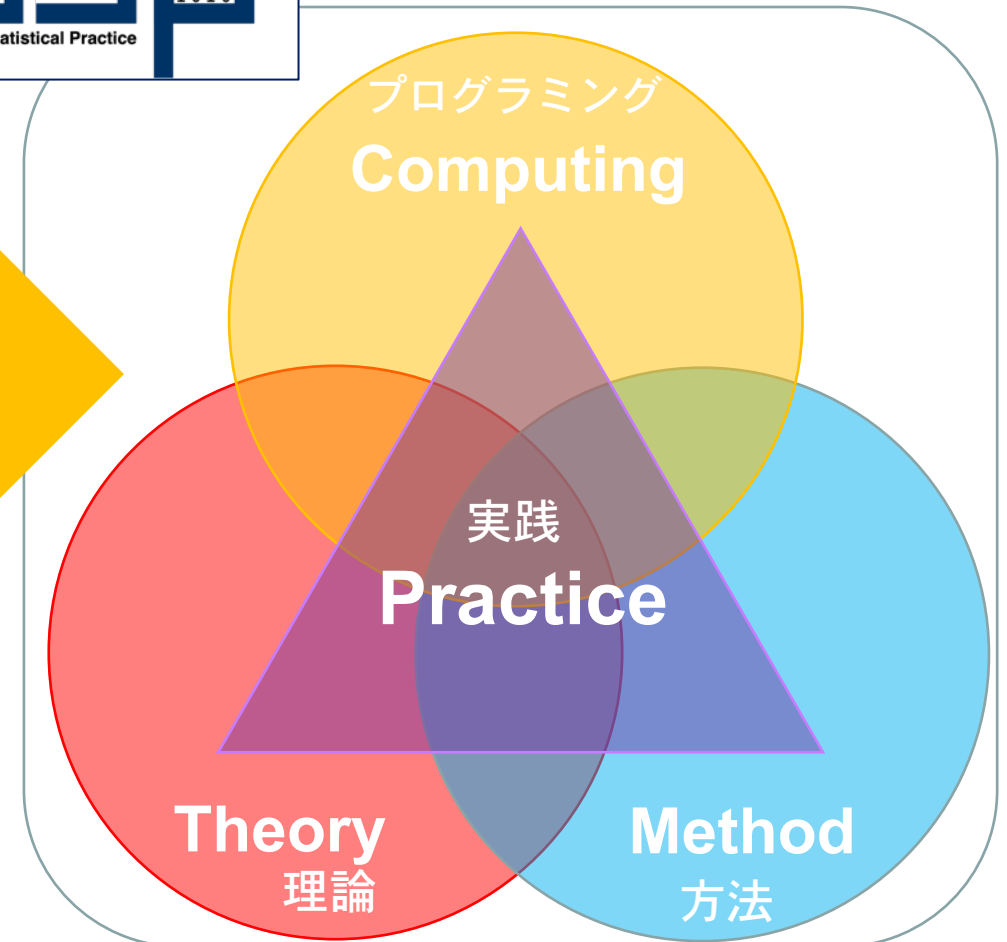
# Data Science (DS) Education Program Design

- Traditional DS program



Paradigm  
Shift

パラダイムシフト

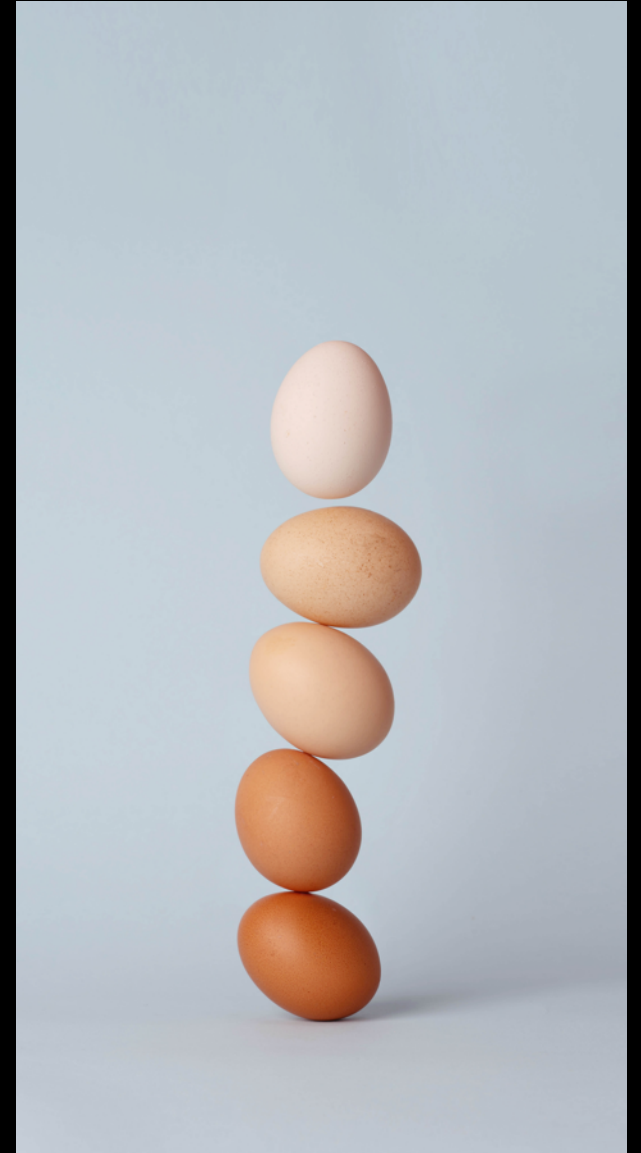






卵が先？それとも鶏？

**Chicken or Egg!?**



船越義珍



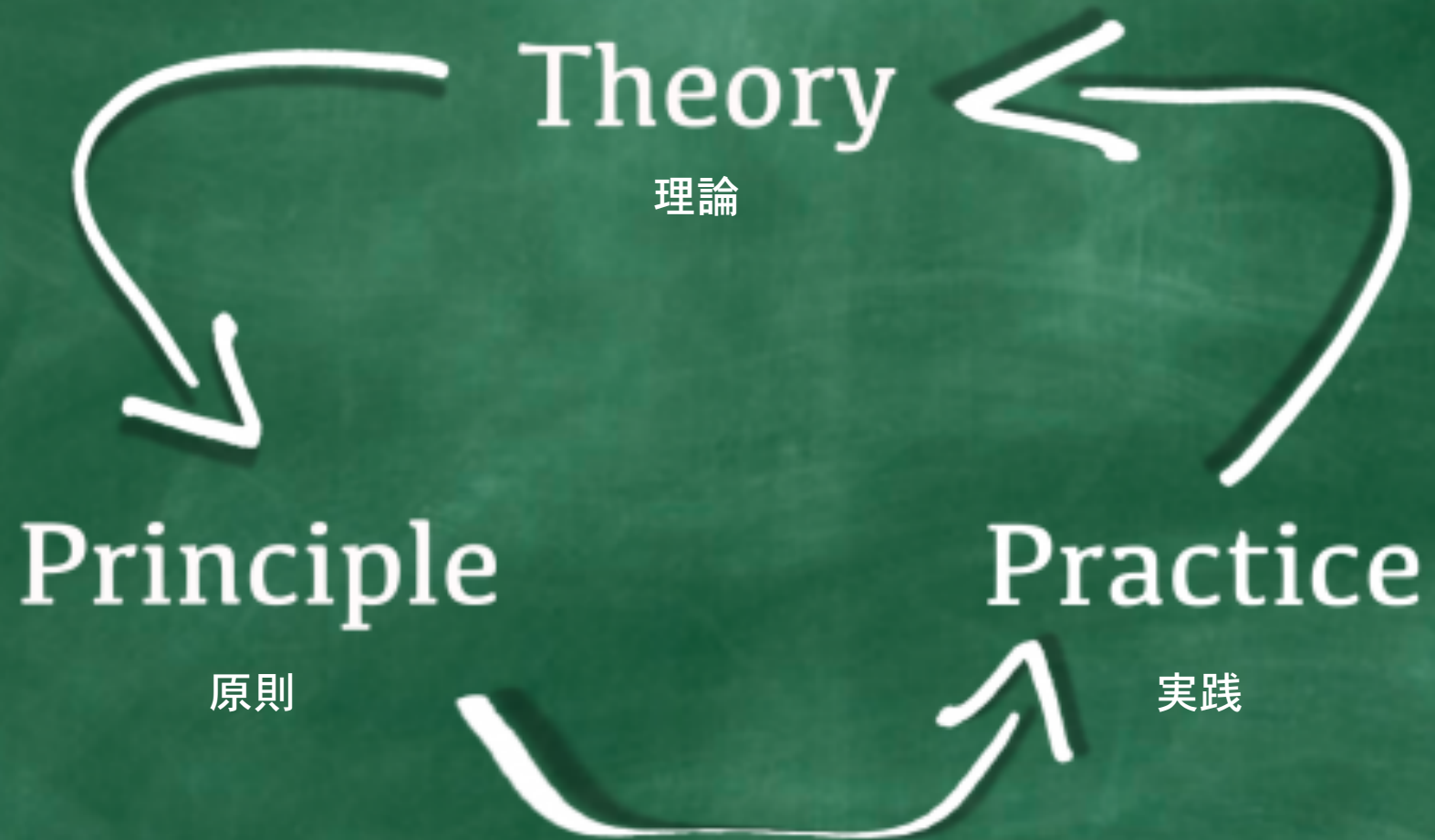
本部朝基



## Form or Practice?

型か実践か？

Our core belief:



How we make it work in the university setting

Since 2015



Statistical  
Consulting  
Service

External  
Partner  
Projects

Computing

Practicum

Theory

Method

今日の生徒



Today's  
Students

明日のデータサイエンティスト

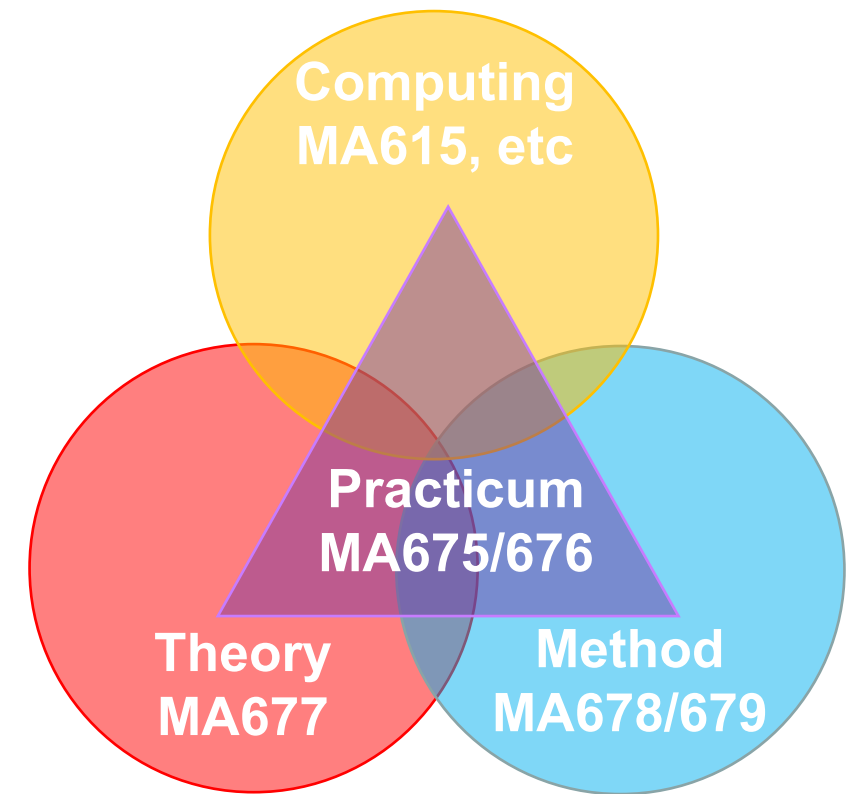


Tomorrow's  
Data Scientists

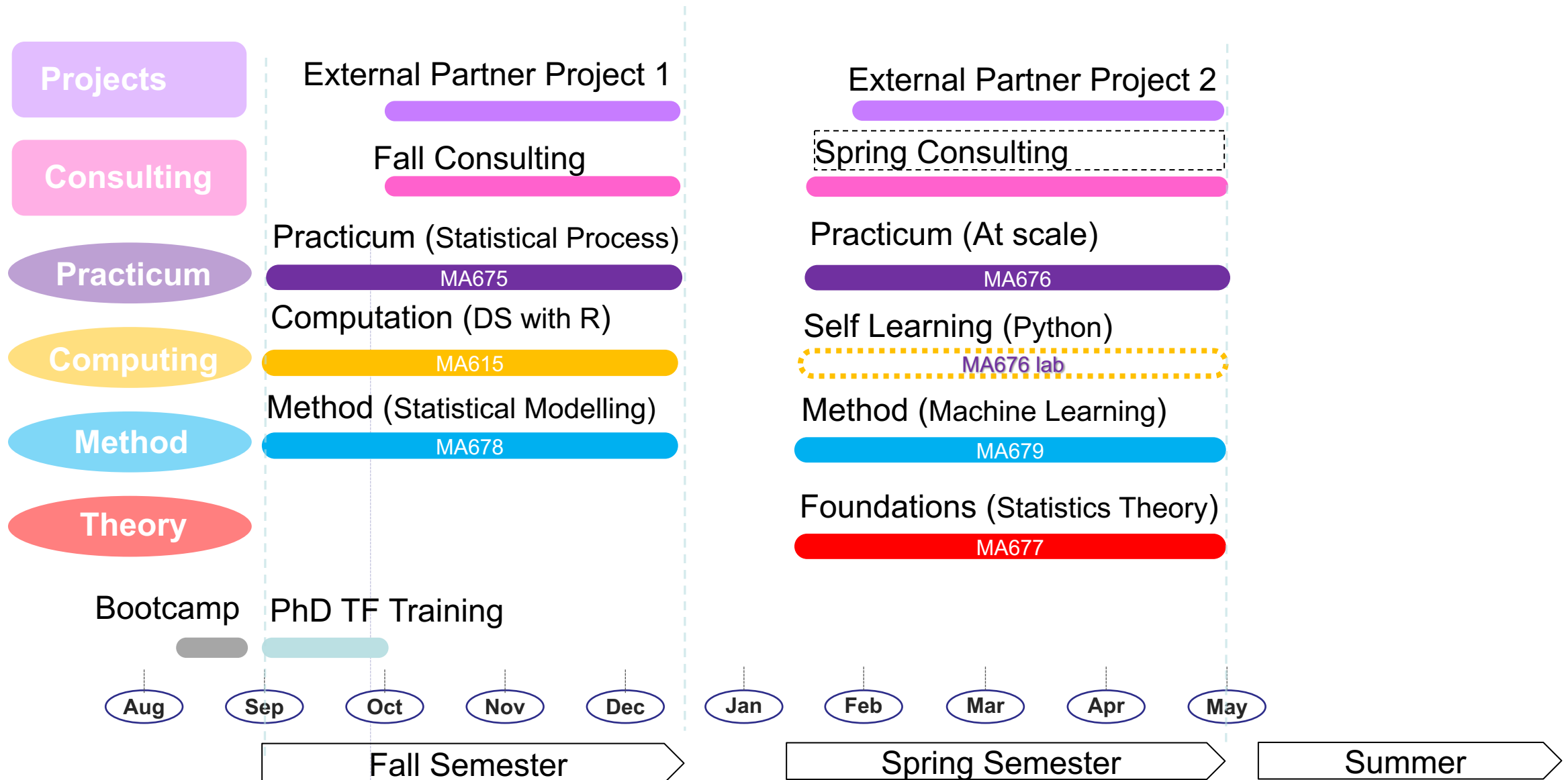
# MS in Statistical Practice (MSSP)

## A Practice Centric Statistics Focused Data Science Program

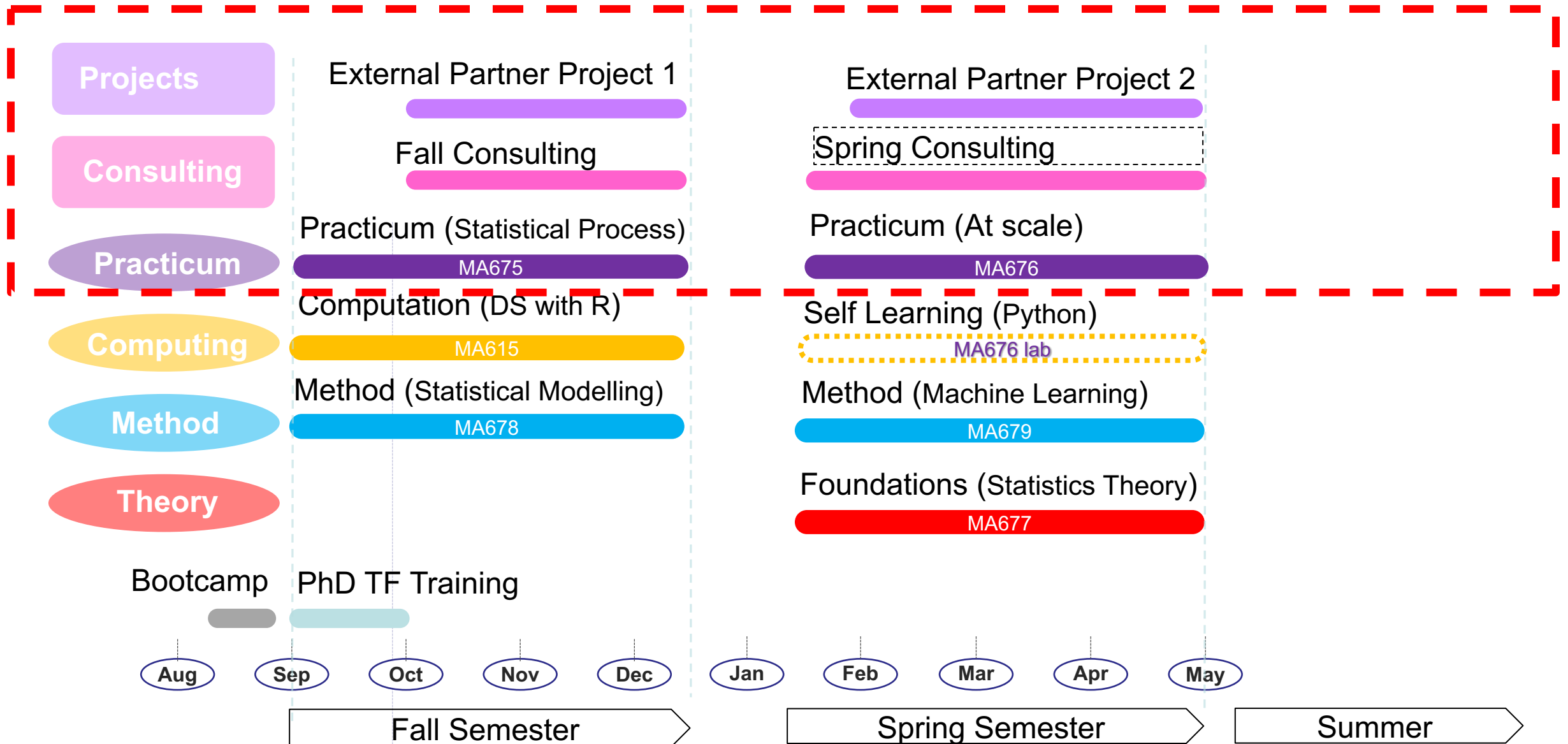
- Cohort based (~45-60 students), diverse background
- Duration: 9 month - 1.5 year
- 8 Courses (32 credit)
  - MA675/676 Statistical Practicum I & II
  - MA678/679 Statistical Modeling / Machine Learning
  - MA677 Foundation of Statistics
  - 3 Electives
    - MA615 Data Science in R, etc
- Additional offerings:
  - 2 Week Bootcamp
  - Career Seminars
  - Stats@Work Seminars
- Graduation requirement: Portfolio



# Practice Centric: Interwoven with Practical Experience

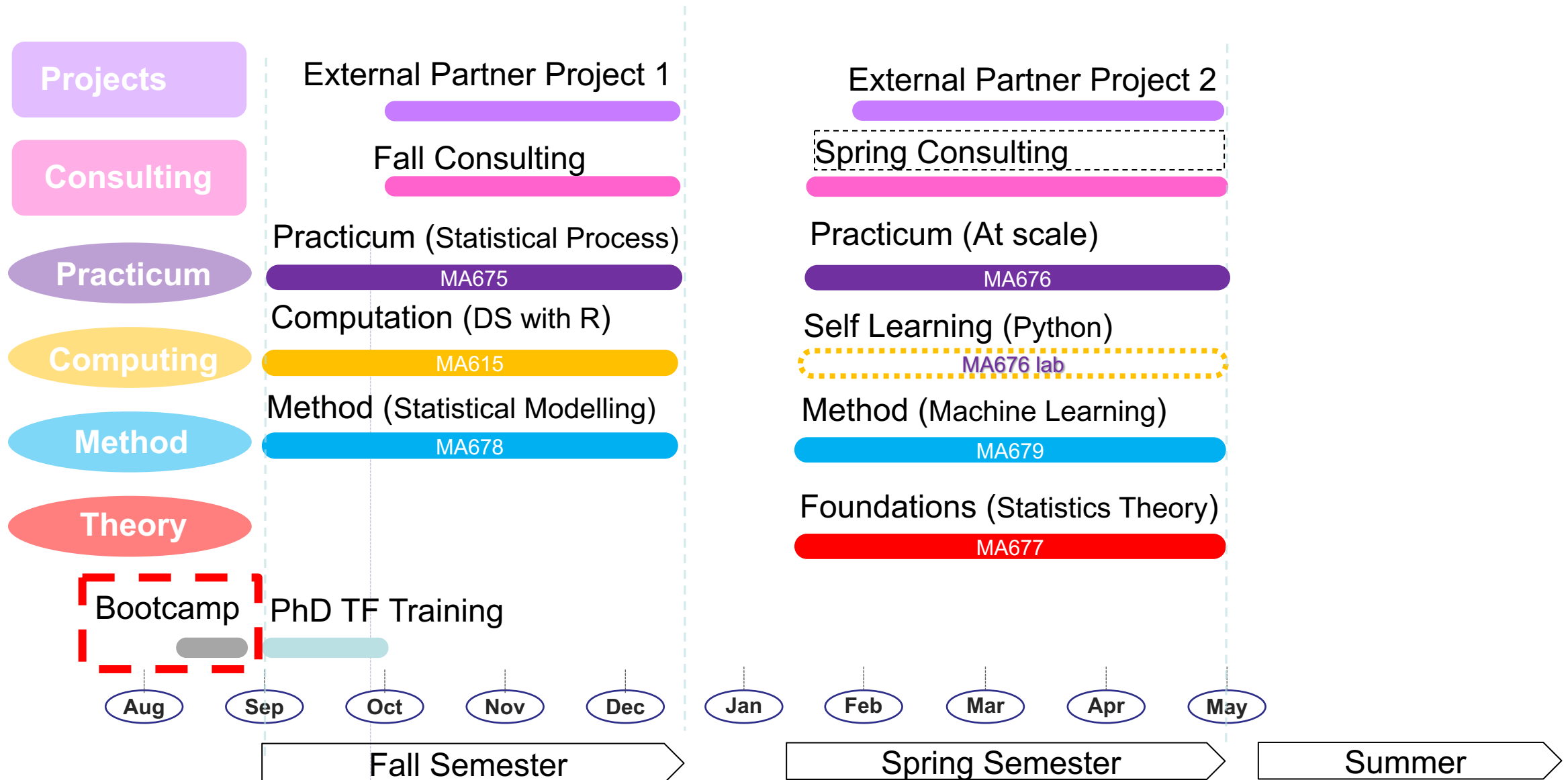


# Practice Centric: Interwoven with Practical Experience





# Practice Centric: Interwoven with Practical Experience



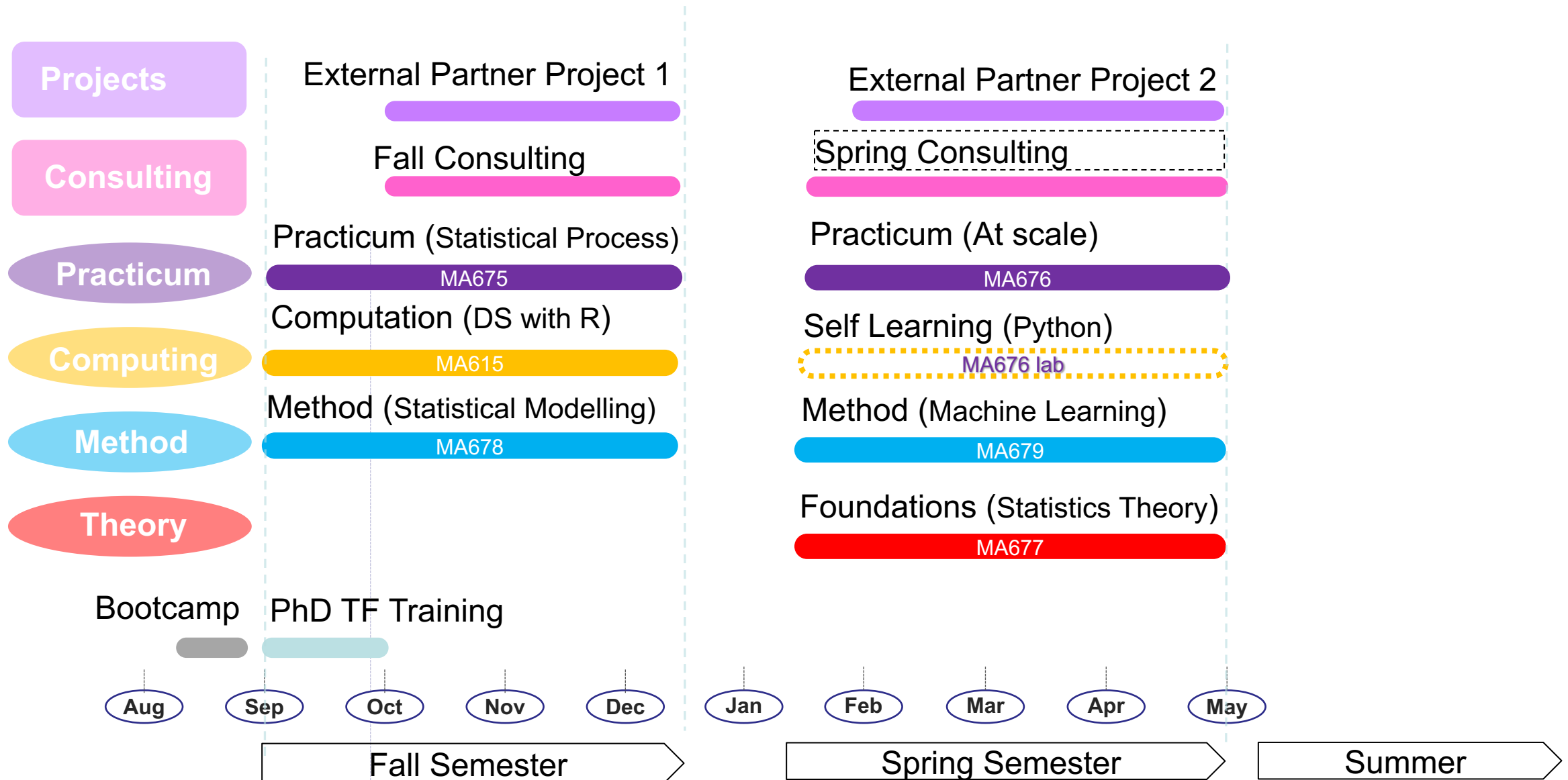


# Bootcamp

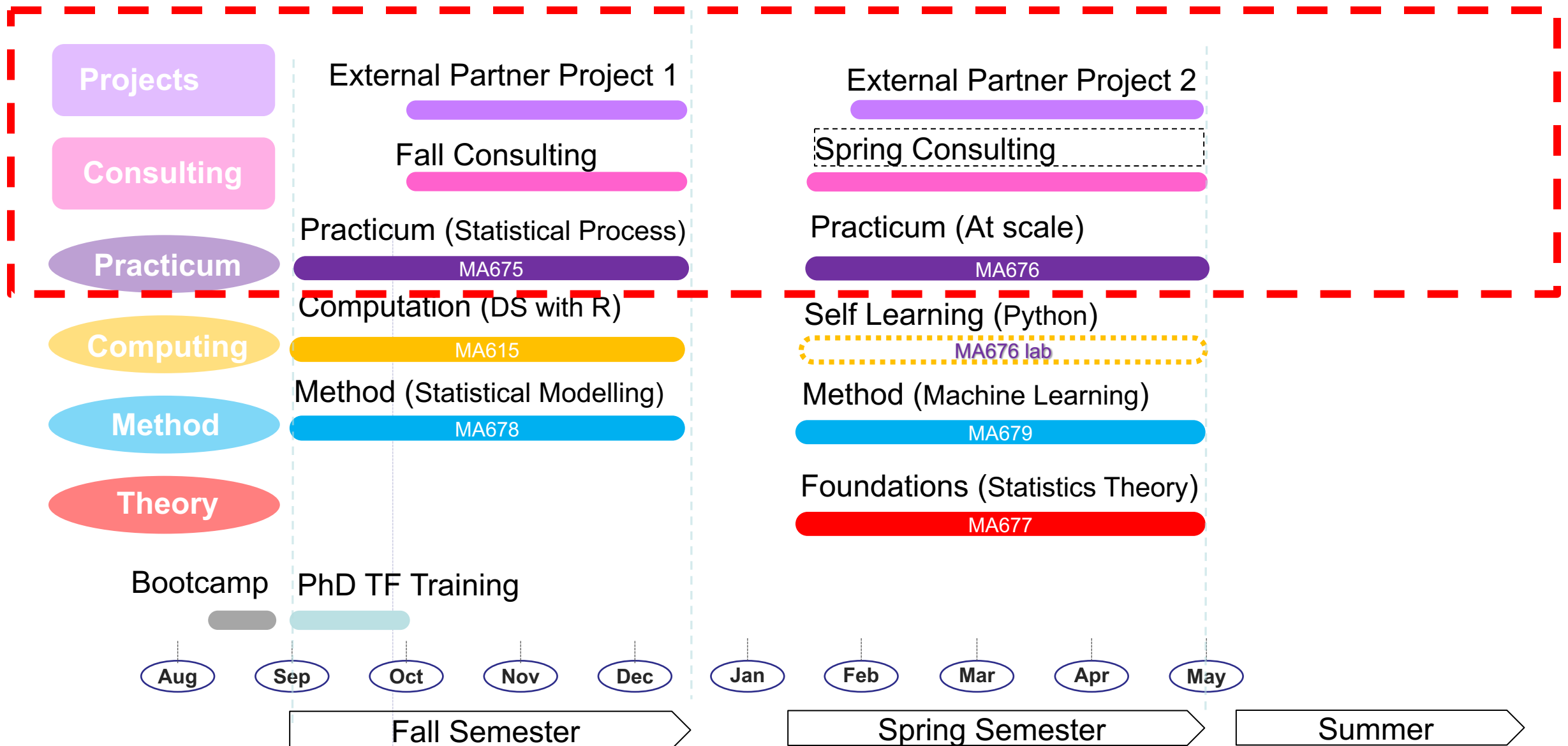
- Two Weeks of intense program
  - Basic R
  - Probability and Statistics
  - Linear Algebra
- Multiple activities
  - Paired introduction
  - Team building exercises
  - Social events
- Purpose
  - Pre-test and post-test to assess student performances.
  - Team exercises to identify communication and group work performance.
  - Set up the students so that they can hit the road running.



# Practice Centric: Interwoven with Practical Experience

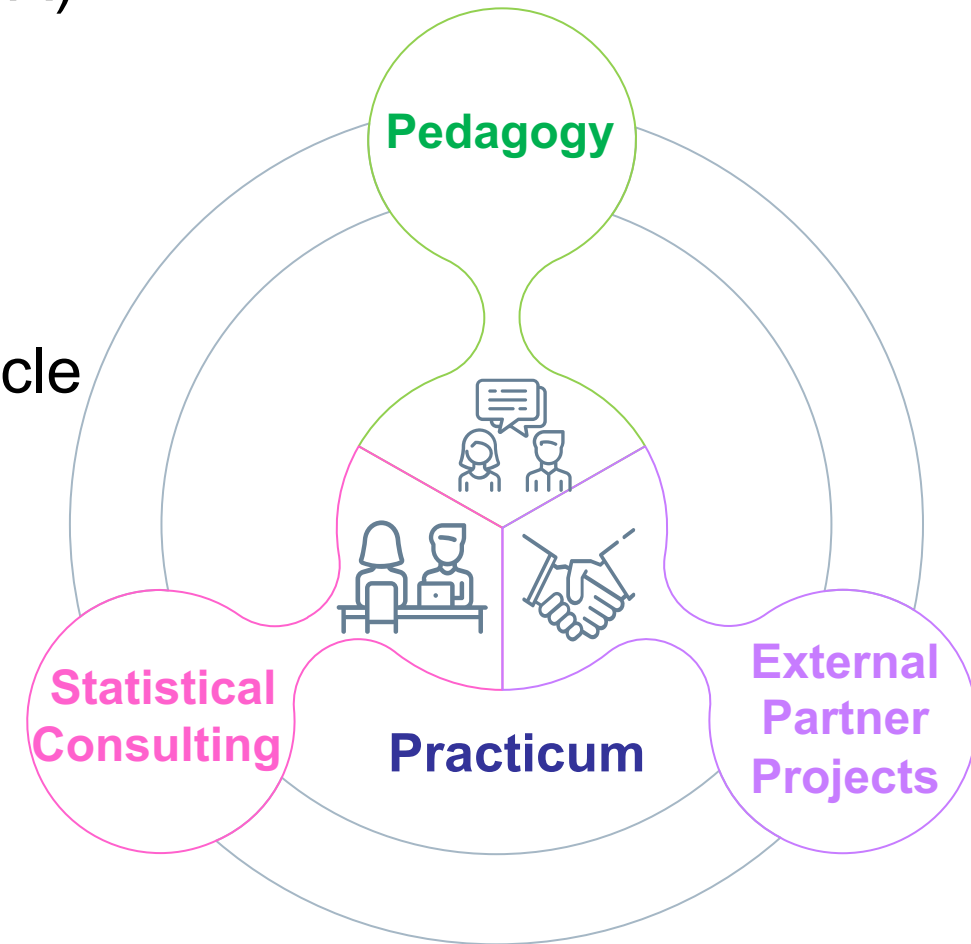


# Practice Centric: Interwoven with Practical Experience



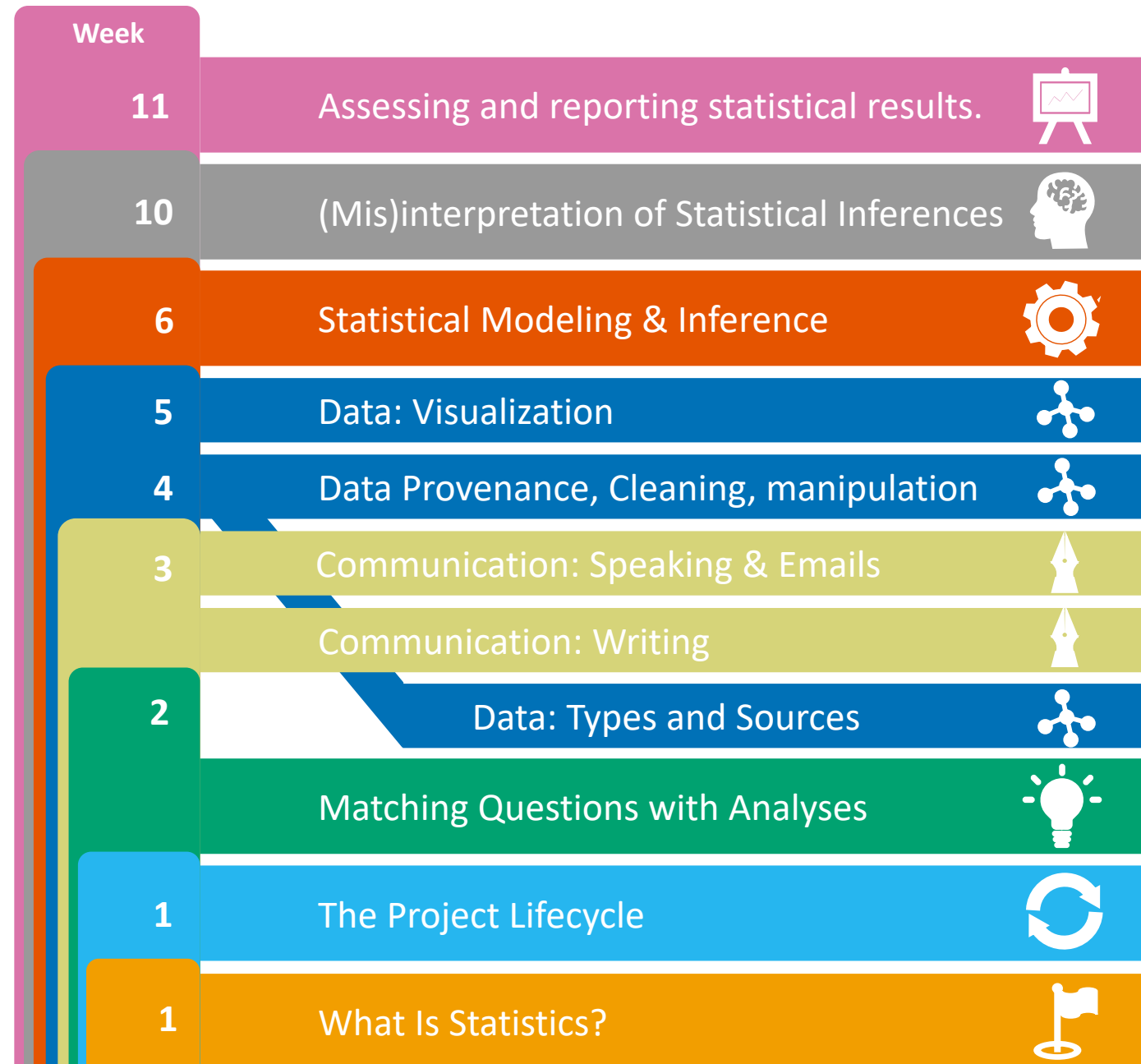
# Statistical Practicum -- Where it all comes together.

- Team taught (3 Faculties, 3 PhD TFs and 1 MS TA)
- Run as a "regular class"
  - In person class (1h15m \* 2)
  - Discussion section (technical topics)
  - Lab (50m) + On call consulting
- MA675: Statistics Focused Data Science Life Cycle
- MA676: Statistician in Broader Data Science
- Three Pillars of the practicum
  - Pedagogy (Weekly readings, discussion, quiz)
  - Consulting Projects
  - External Partner Project



# Pedagogy

- Weekly reading
- Reading quiz
- Discussion
- Homework assignments
  - Target Resume
  - Writing
  - Networking



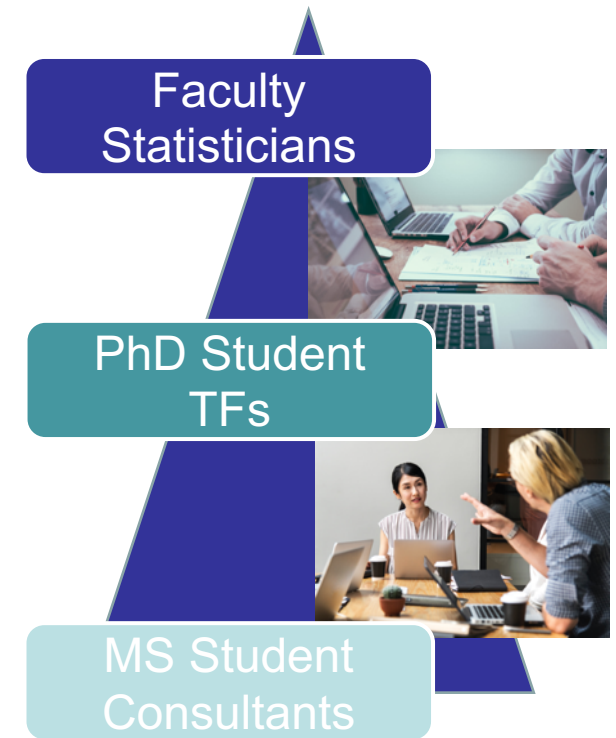
# Pedagogy

- Weekly reading
- Reading quiz
- Discussion
- Homework assignments
  - Target Resume
  - Writing
  - Networking

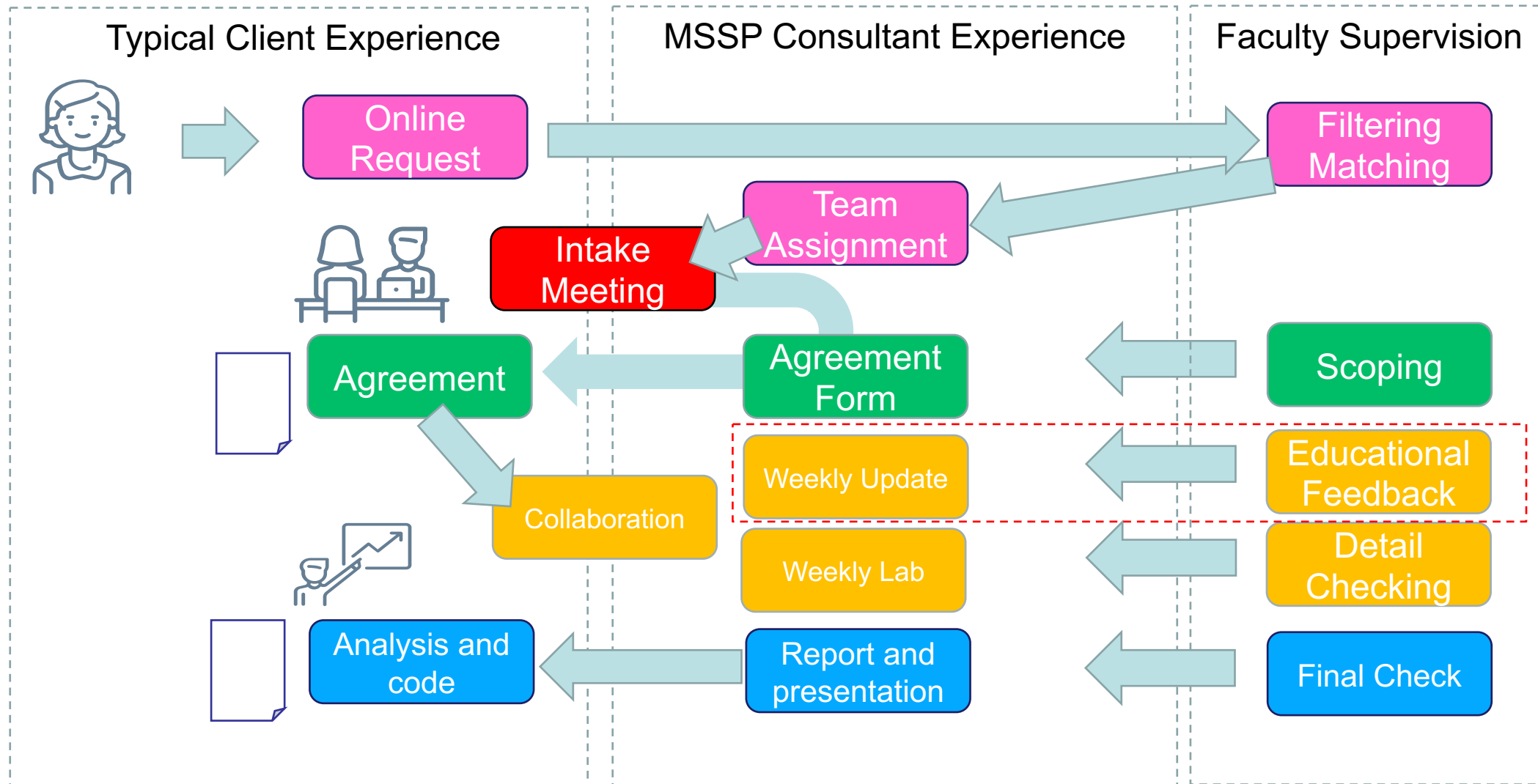
Week		
11	分析の発表方法	
10	分析の理解	
6	統計モデル、推論	
5	データ： 視覚化	
4	データ： 出所の考慮、クリーニング、操作	
3	コミュニケーション：会話、E-mailの書き方	
	コミュニケーション：文書の書き方	
2	データ：様々なデータと特徴	
	質問と分析を一致させる	
1	プロジェクトライフサイクル	
1	統計って何？	

# Consulting $\approx$ Residency Program

- Consulting service provided to BU research community.
- Every Thursday class is a Consulting day.
- MS students + PhD TF execute the work.
- Roughly 15 students per TF meet during a lab
- Students are expected to be “on call” for meetings.
- Three levels of service
  1. **“One and Done”**: at most one hour of consultation
  2. **Limited Duration (LD)**: 10~20 hours of student time
  3. **Extended Duration (ExD)**: length depends on the project.



# MSSP Consulting: How it works





# Impact (~100 projects / per year)

## Genetic Counseling Program

Analysis of genetic test choice for BMC patients

## Biology

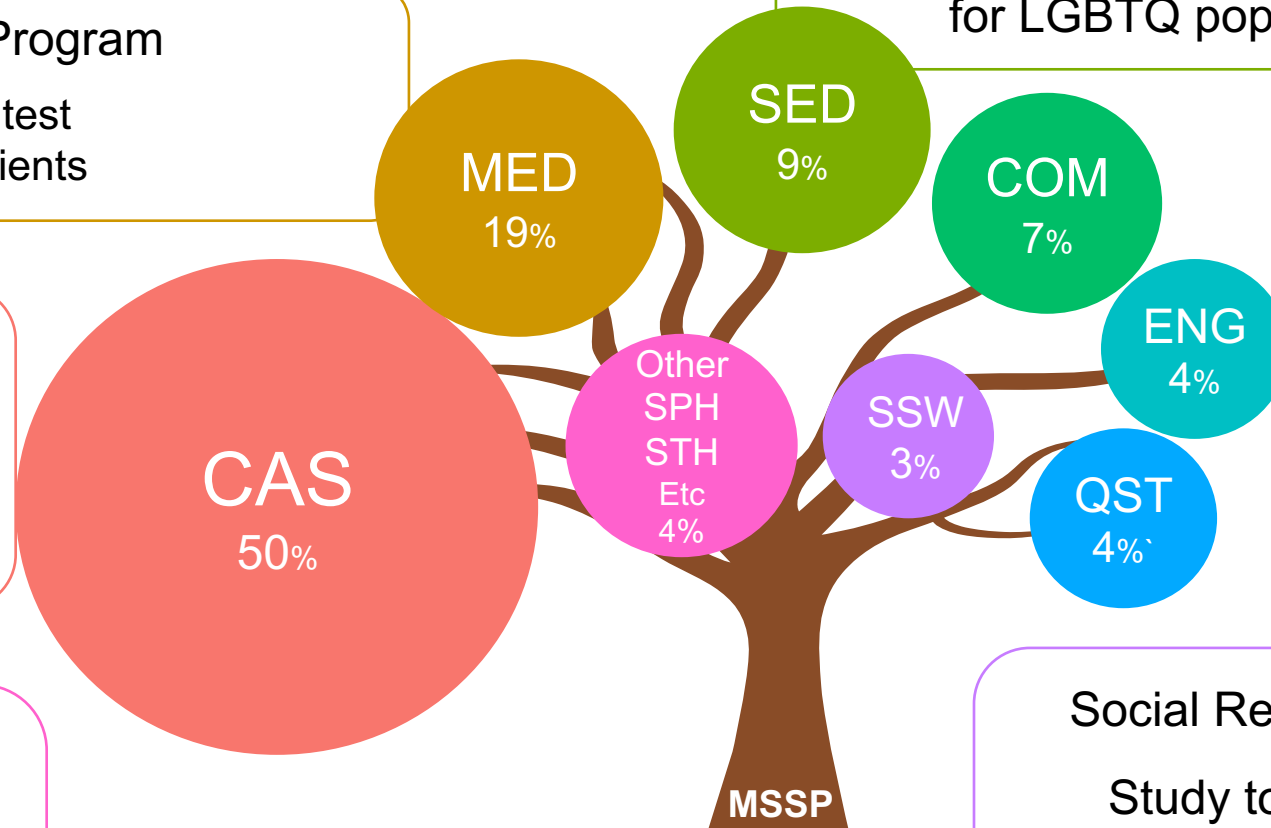
Analyses of biogeochemistry experiments and CO2 emission models

## Occupational Therapy

teaching young individuals how to resolve environmental barriers to participation

## Environmental Health

Assessment of particulate matters around Logan Airport



## Counseling Psychology

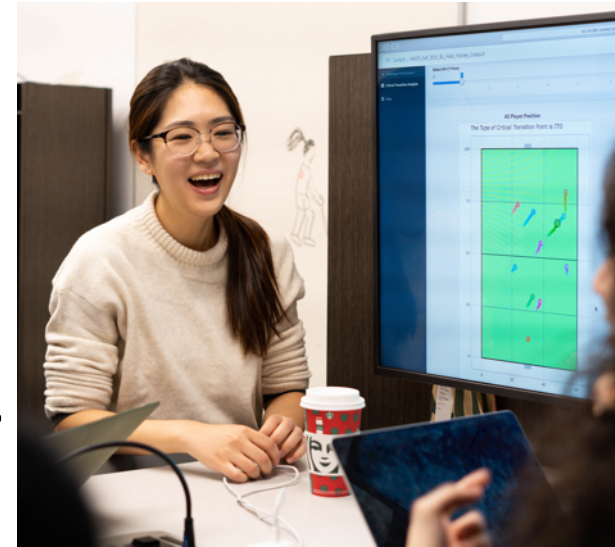
Exploration of large-scale study on violence exposure for LGBTQ population

## Social Research

Study to improve the mental and sexual health of Asian American Women

# External Partner Project

- R&D for the partners; guided internship experience.
- One semester long , student team supervised by a Faculty.
- Real experience;
  - business context,
  - direct (constructive) feedback, and
  - real (messy) data
- What we deliver
  - Code/software,
  - written report,
  - detailed explanation of the idea.



# Current and Competed MSSP Partner Projects



**CITY of BOSTON**



TechWeb



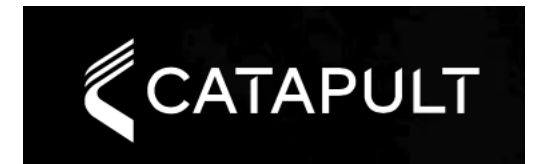
BOSTON PUBLIC SCHOOLS

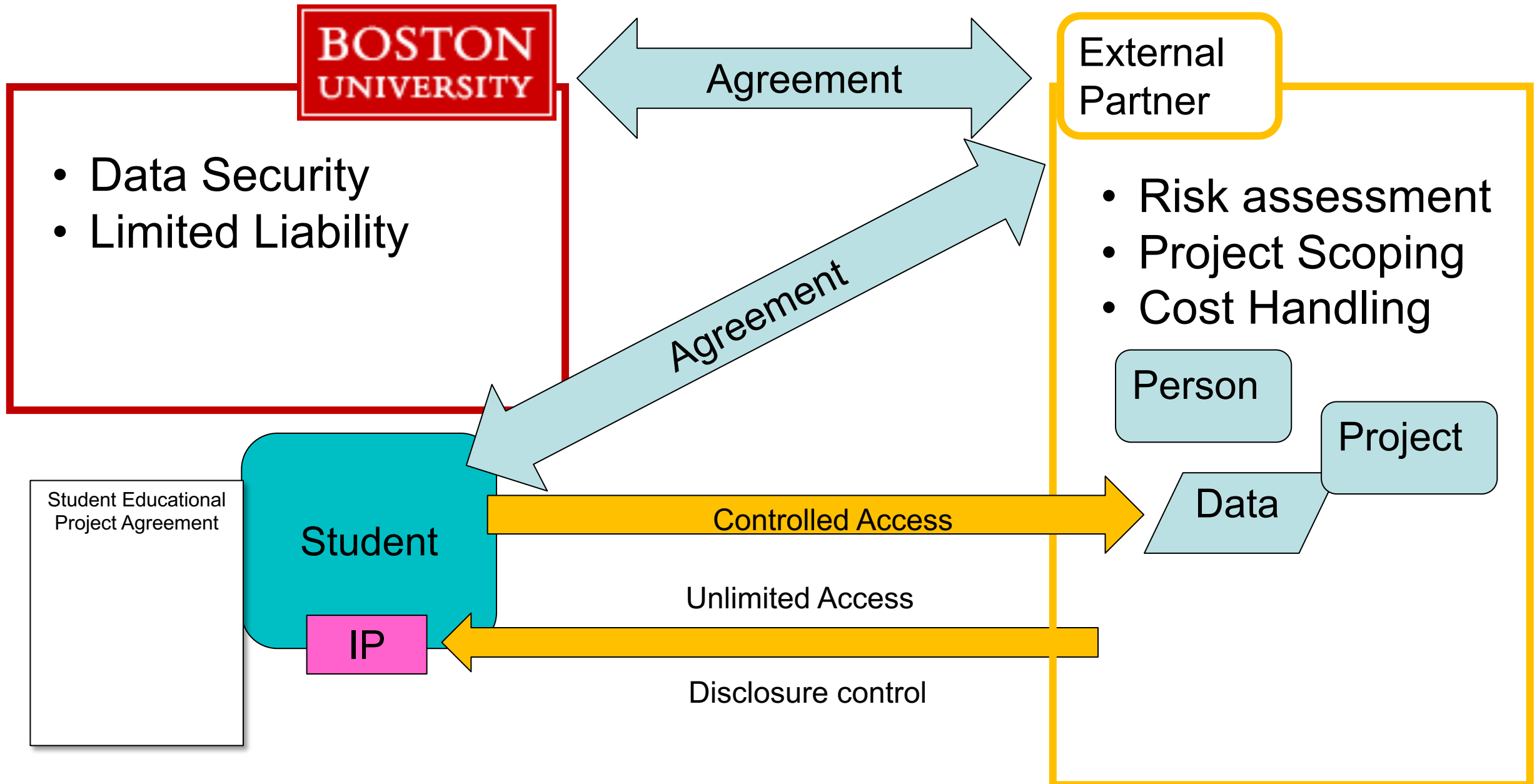


BOSTON UNIVERSITY ATHLETICS



**MANGO  
SCIENCES**





Since 2015



Statistical  
Consulting  
Service

External  
Partner  
Projects

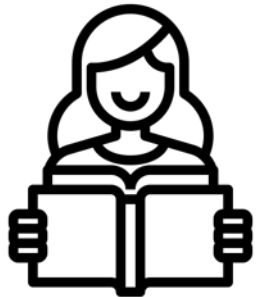
Computing

Practicum

Theory

Method

今日の生徒



Today's  
Students

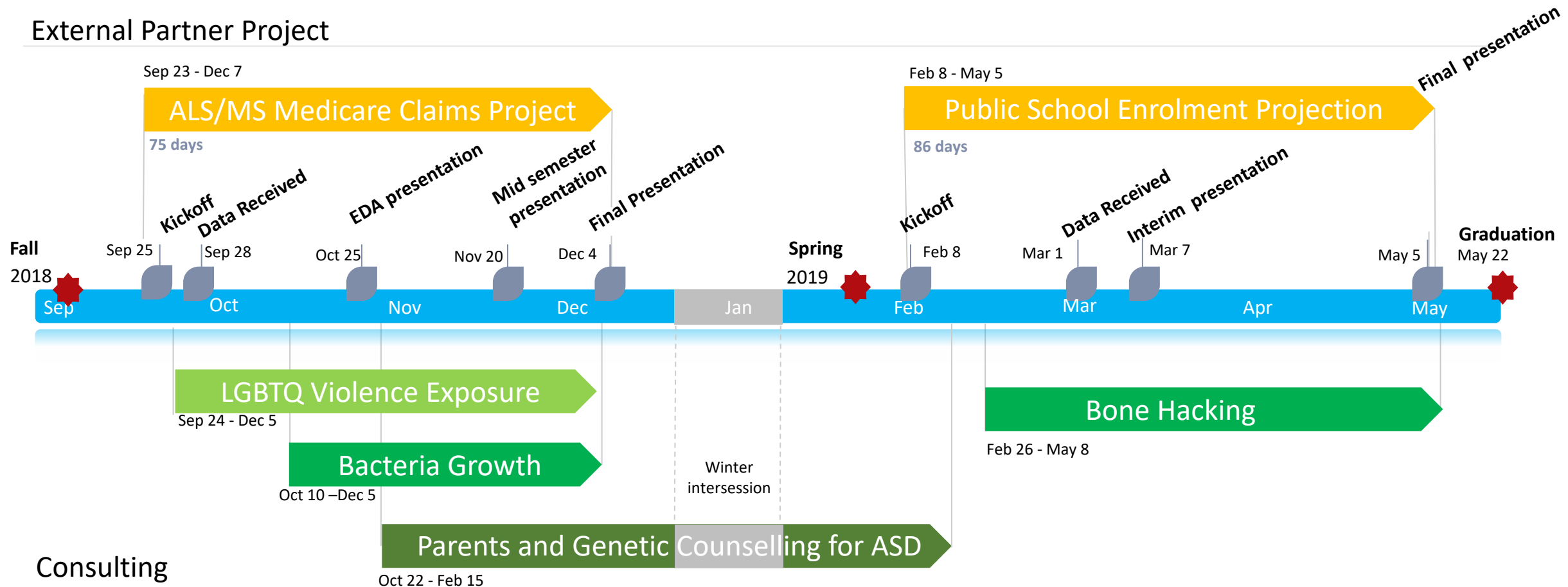
明日のデータサイエンティスト



Tomorrow's  
Data Scientists

# Example Student Experience

## External Partner Project



# Secret Sauce?

- Reality: real client, real project, real interest
- Culture of Trust: Bootcamp, Cohort, Work together
- Communication (faculty, TFs, students, clients)
- Just In Time (JIT) teaching & training, Staying Agile.
- Manage expectations and strive to provide value.
- Instructors are put under pressure in both teaching and management.
- Human resource:

	Quantity	Class Role	Consulting
Professors	3	+	+
Research Faculty	1+		+
PhD TF	3	+	+
TA	(1)	+	



# Other things we did not go over

- MA677 (Foundations of Statistical Practice)
  - Look back theory. Why does it work? Not how it should be.
- Practicum Labs and self guided programming trainings
  - Kept open for just in time (JIT) learning
- Deal with demand fluctuations for projects.





# Challenges and Future

# MSSP in COVID19

- Learn From Anywhere (LfA)
  - Weekly Testing
  - Socially Distanced Seating
  - Attendance Monitor App
- 52 Students in at least 4 time zones
  - Total of 12 hour time zone differences
- Heavy use of technology
  - Zoom, MS Teams, Flipgrid, etc
- Fall 2020
  - 5 Partner projects, 33 Consulting projects
- Spring 2021
  - 4 Partner Projects



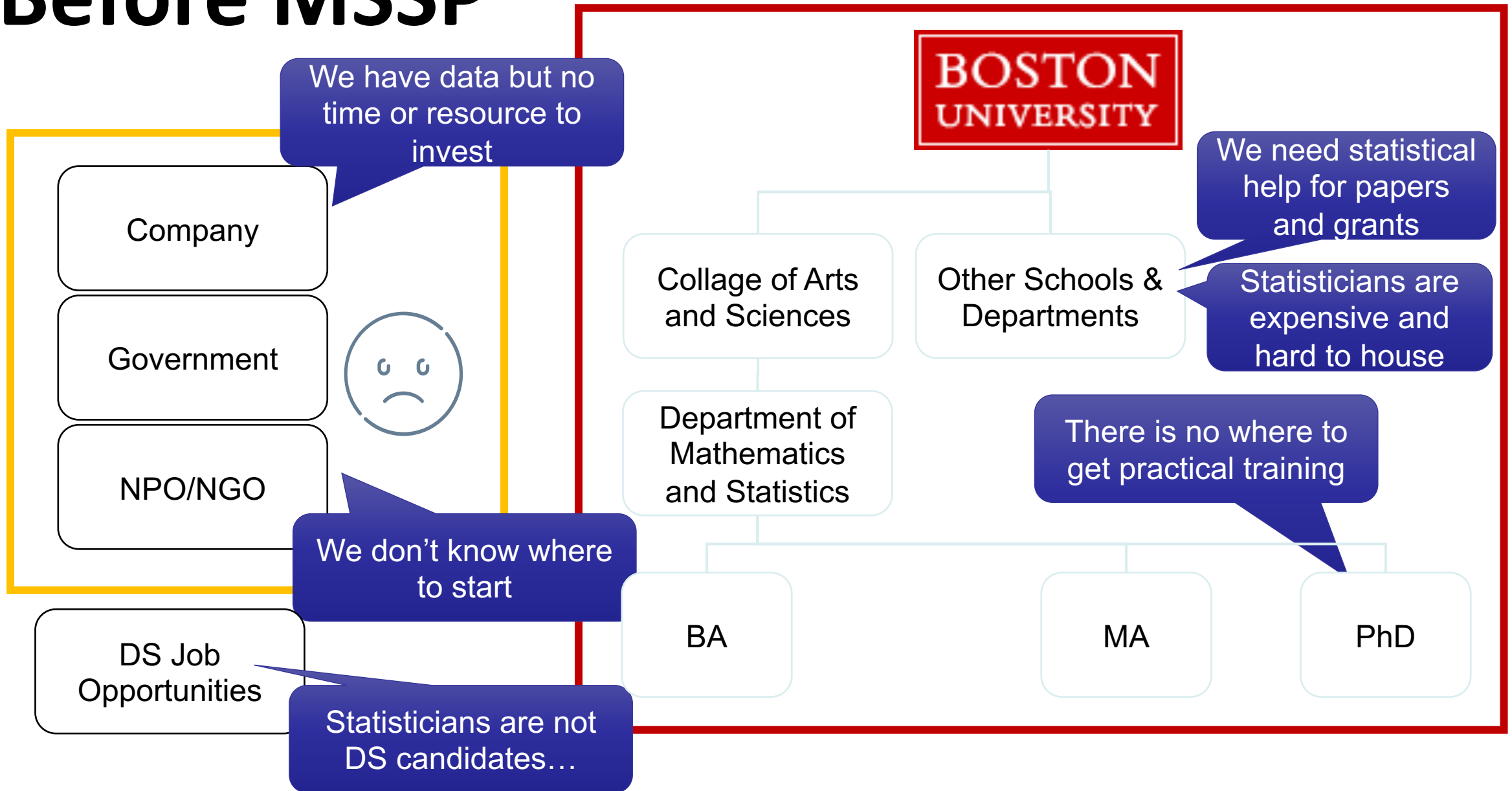


# Challenges

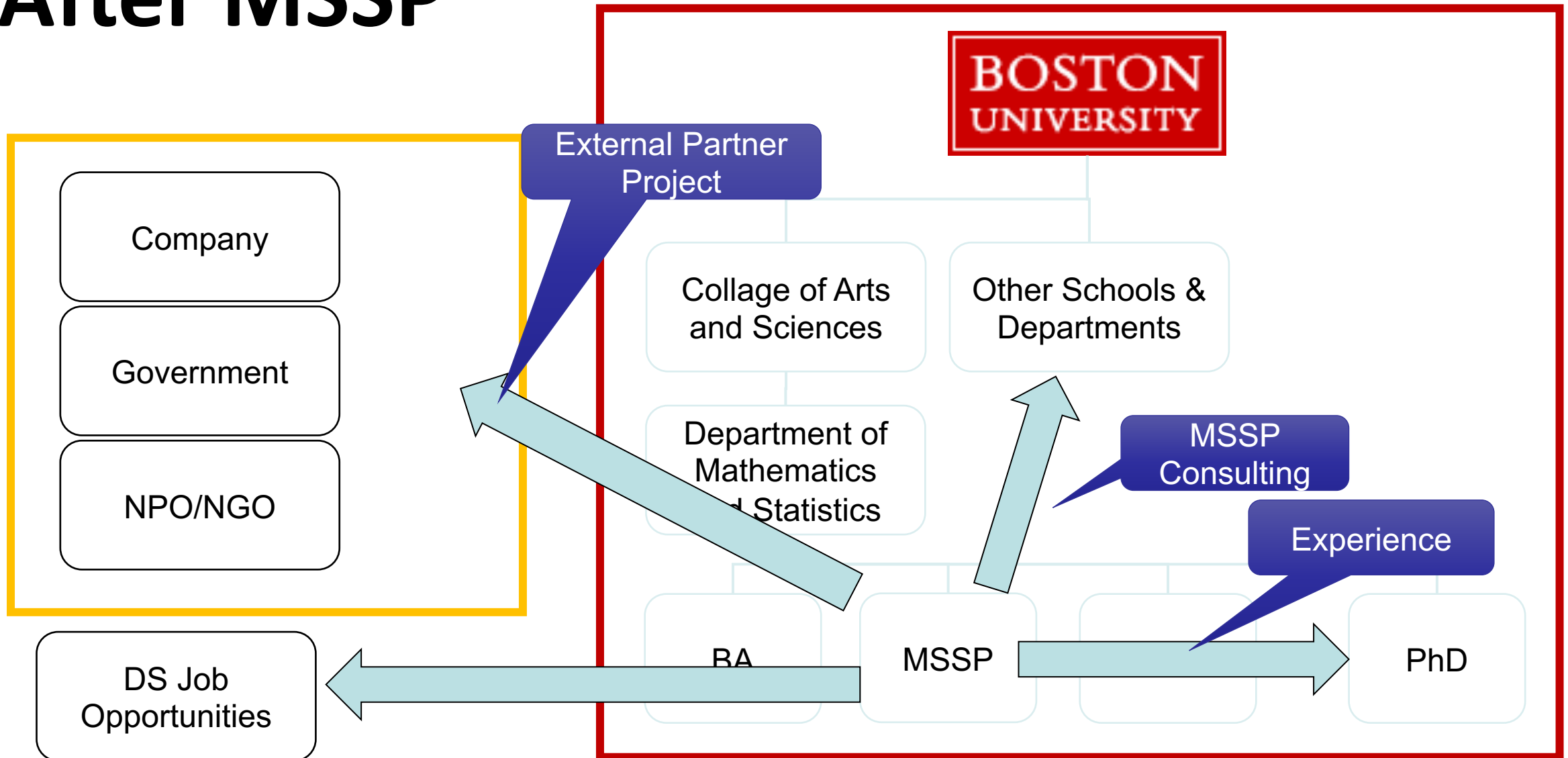
- Student Assessment
  - Peer-Self-Co-Instructor (PSCo) Assessment
- Diversity
  - “Student's success” for diverse students
- Recruiting and Working with Working Professionals
  - University and industry schedule miss match
- Staying agile but also staying principled
  - Fast-changing landscape of Data Science
- Scaling to a larger cohort with inclusivity



# Before MSSP



# After MSSP



# Special Thanks

- Collaborators



Eric Kolaczyk



Haviland Wright

- Members of MSSP



Luis Carvalho



Sean Grogg





Please feel free to reach out if you have any question.

Masanao Yajima  
yajima@bu.edu

More Information at  
<https://www.bu.edu/mssp/>

HDSR paper  
<https://hdsr.mitpress.mit.edu/pub/twyc748y/release/2>



Home >

**COVID-19:** View the **latest updates** for researchers or visit our [recovery toolkit](#).

# Corporate-Student Engagement in the Classroom, Guidelines for



PRINT



# Agreement Template

**Note:** This form agreement is an example only. It contains terms and conditions to protect BU and its students. It assumes that the corporate participant in the student educational project seeks rights to intellectual property that may be created in the course of the project and that the project requires the corporate participant to share information it considers confidential. Not all corporate-sponsored student educational projects require the sponsor to provide confidential information and not all corporate sponsors will seek intellectual property rights. The form may need to be changed to reflect that. This form also assumes that the student's participation is optional, which is reflected in the Student Agreement and Acknowledgment.

## Student Educational Project Agreement

This Student Educational Program Agreement (including all attachments, exhibits, and amendments, the “Agreement”), dated \_\_\_\_\_, 20\_\_ (the “Effective Date”), is made by and between Trustees of Boston University (“BU”) and [Sponsor Name], a [Sponsor state of organization and type, e.g. corporation, LLC, LP] with its principal office at [Sponsor address] (the “Sponsor”) (BU and Sponsor, collectively, the “Parties,” each a “Party”).

### I. Purpose

- a. This Agreement outlines the terms and conditions that will govern the Student Educational Project described in the Project Description attached hereto as Exhibit A (the “Project”). The Project Description describes the roles and expectations of the Parties for the Project. The overall purpose of this Project is to provide students with an educational experience requiring application of statistics, machine learning, predictive analytics, and other aspects related to data science practices to a real-world problem.
- b. The parties will also enter into a Mutual Confidential Disclosure Agreement, in the form

THANK YOU





# Acknowledgement

- The pictures in the slides are from <https://unsplash.com/>
- Icons are from [Slidesgo](#)