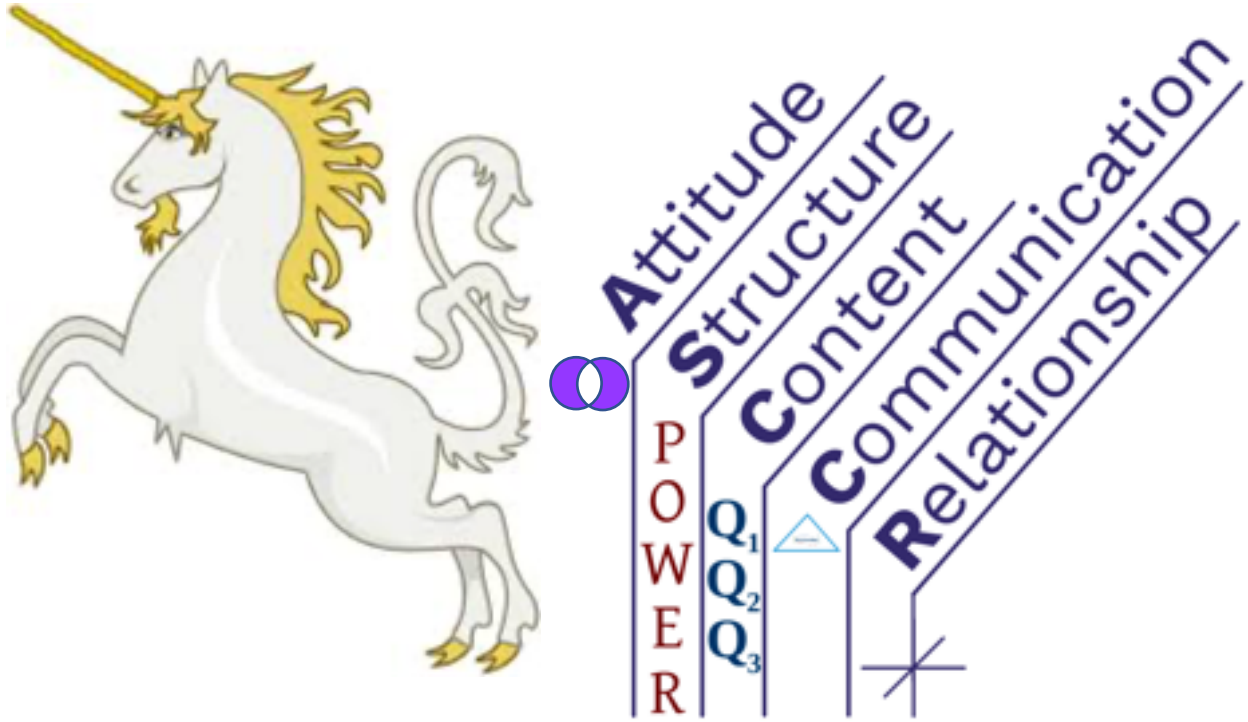
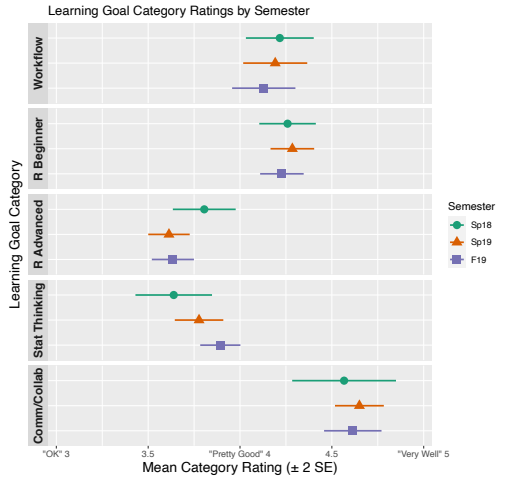
Using Team-Based Learning to Teach Data Science



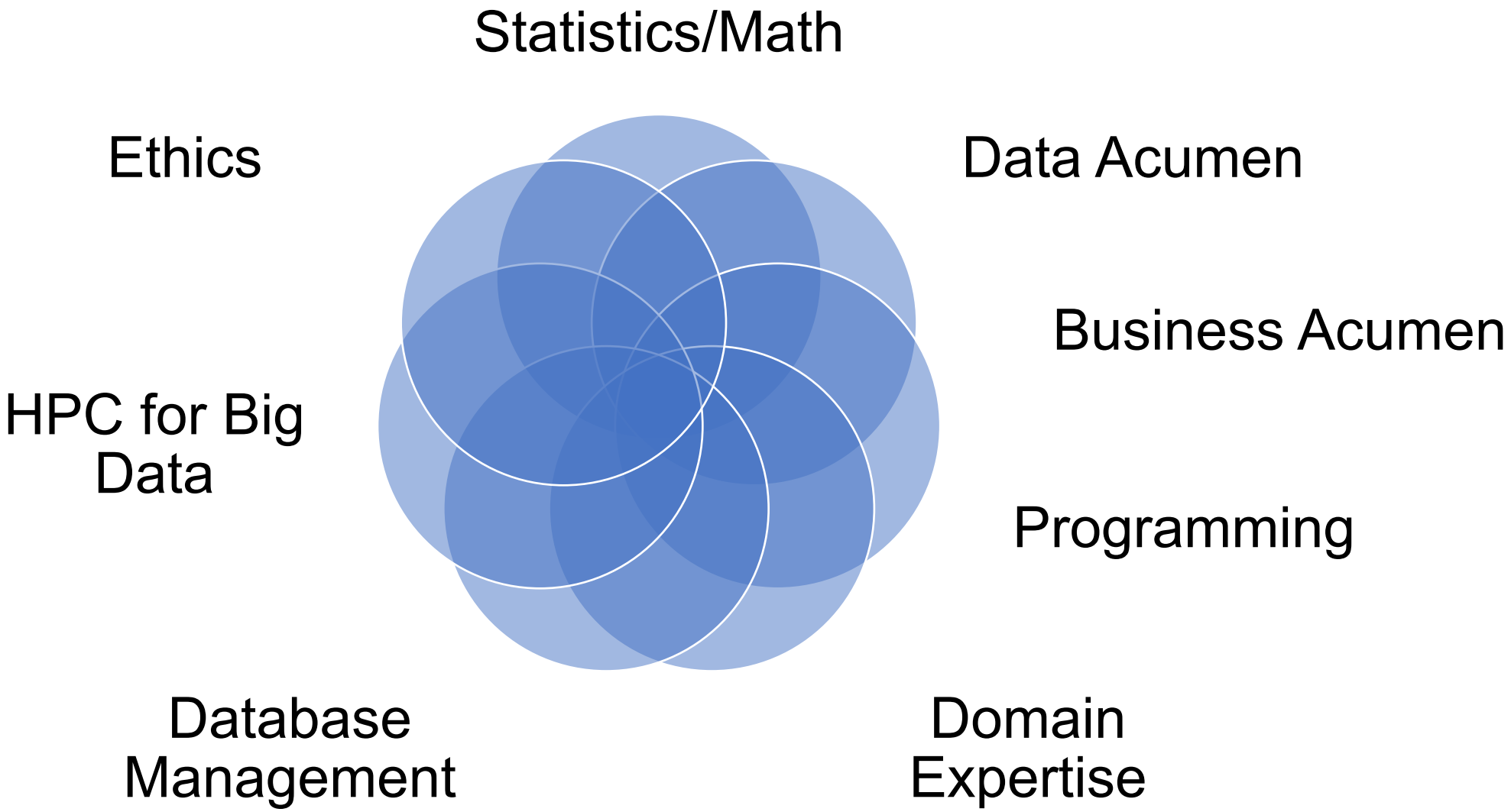
TBLC Team-Based Learning™ Collaborative



Eric Vance
Dept. Applied Mathematics
LISA Laboratory for Interdisciplinary Statistical Analysis
University of Colorado **Boulder**

January 25, 2022
***JSDSE* Webinar**

Statisticians and data scientists need to know many things



Statisticians and data scientists need to know many things

Statistics/Math

Ethics

Data Acumen

Business Acumen

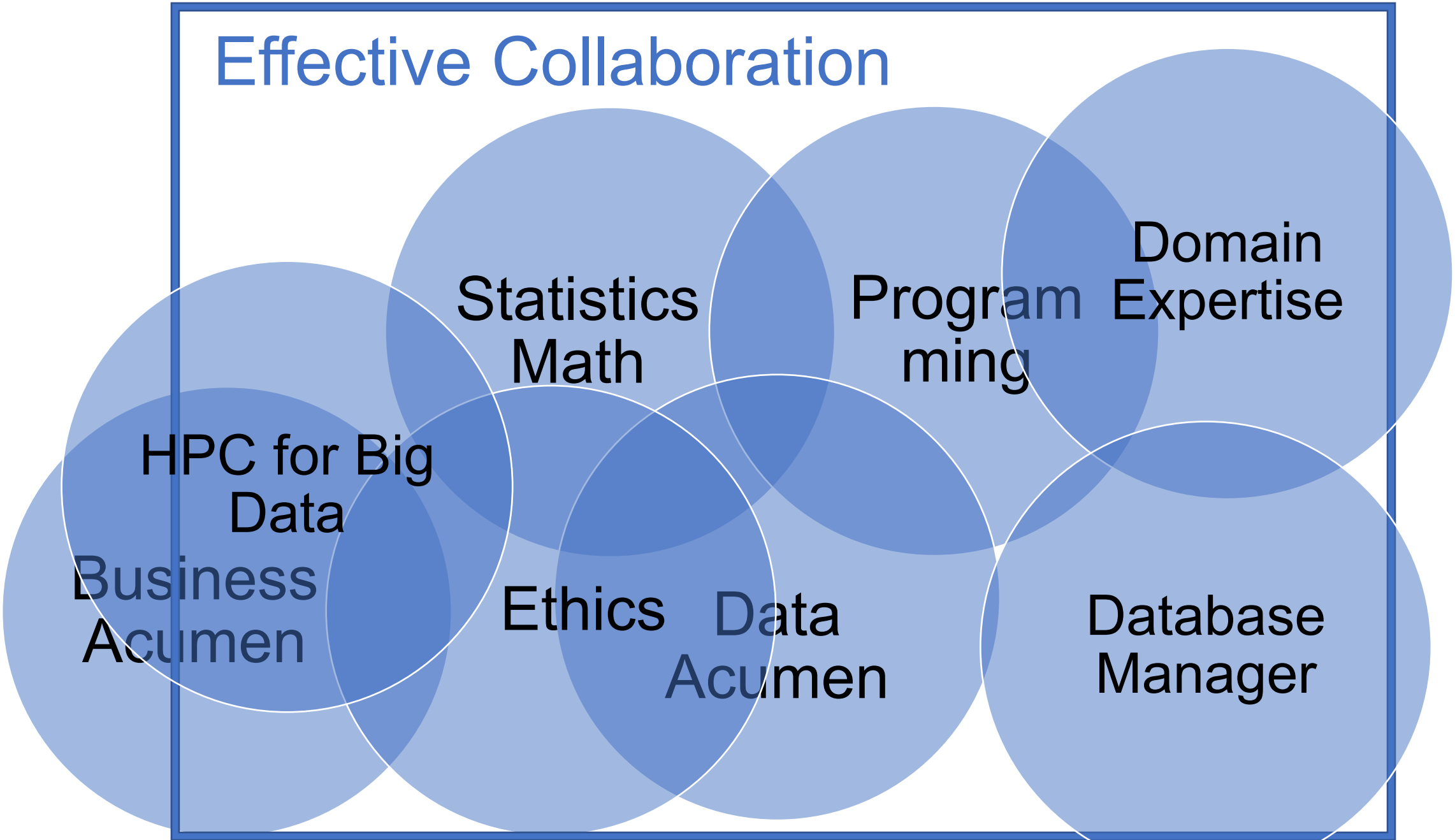
HPC for Big Data

Programming

Database Management

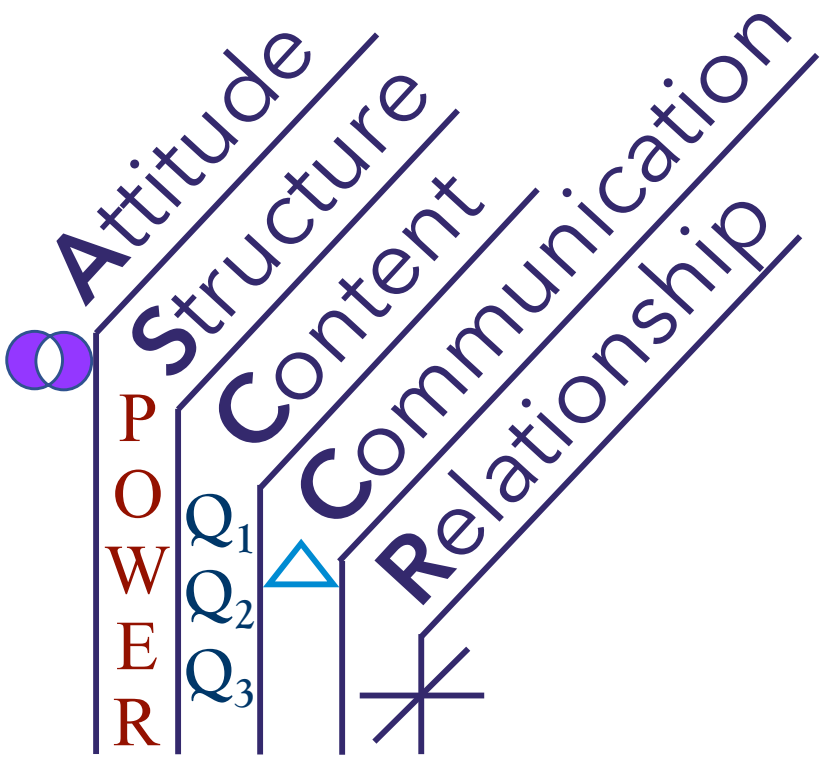
Domain Expertise

An alternative to trying to be a Unicorn is learning



How to teach effective collaboration in Data Science curriculum?

1. Teach ASCCR Framework for Collaboration in a Capstone Course



2019 “ASCCR Frame” paper in *JSE*:
bit.ly/asccrframe

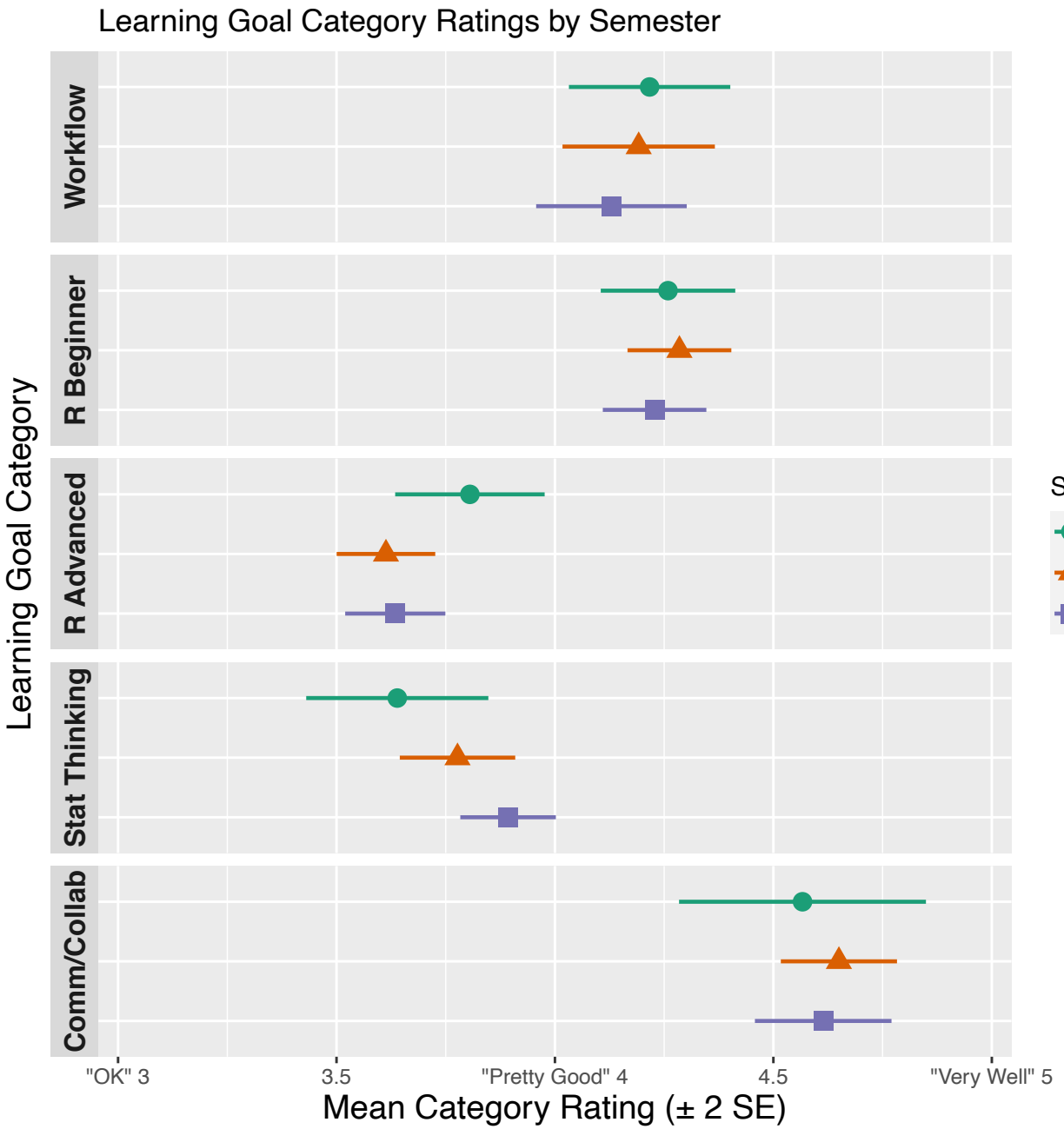
2. Teach Collaboration throughout the curriculum using TBL

“Intro to Data Science”

TBLC

Team-Based Learning™ Collaborative

- Work through 29 of 30 chapters in *R4DS* in 7 modules
- In-class team application exercises to learn Statistics concepts
- Weekly team projects
- Team-Based Learning (TBL) combines flipping the classroom, problem-based learning, small-group learning (Collaboration)



Students reported learning Workflow, Collaboration, and R: Beginner topics: Pretty Good—Very Well; Statistics and R: Advanced: OK—Pretty Good

Collaborating w/ teammates and Communicating findings/recs were top- ranking topics. Students engaged in this type of Collaboration every day.

Transforming data, Data visualization, and Importing data were 3 top-ranking R: Beginner topics students engaged in nearly every week.

Students got much less practice with R: Advanced topics like `purrr`, `forcats`, `lubridate`

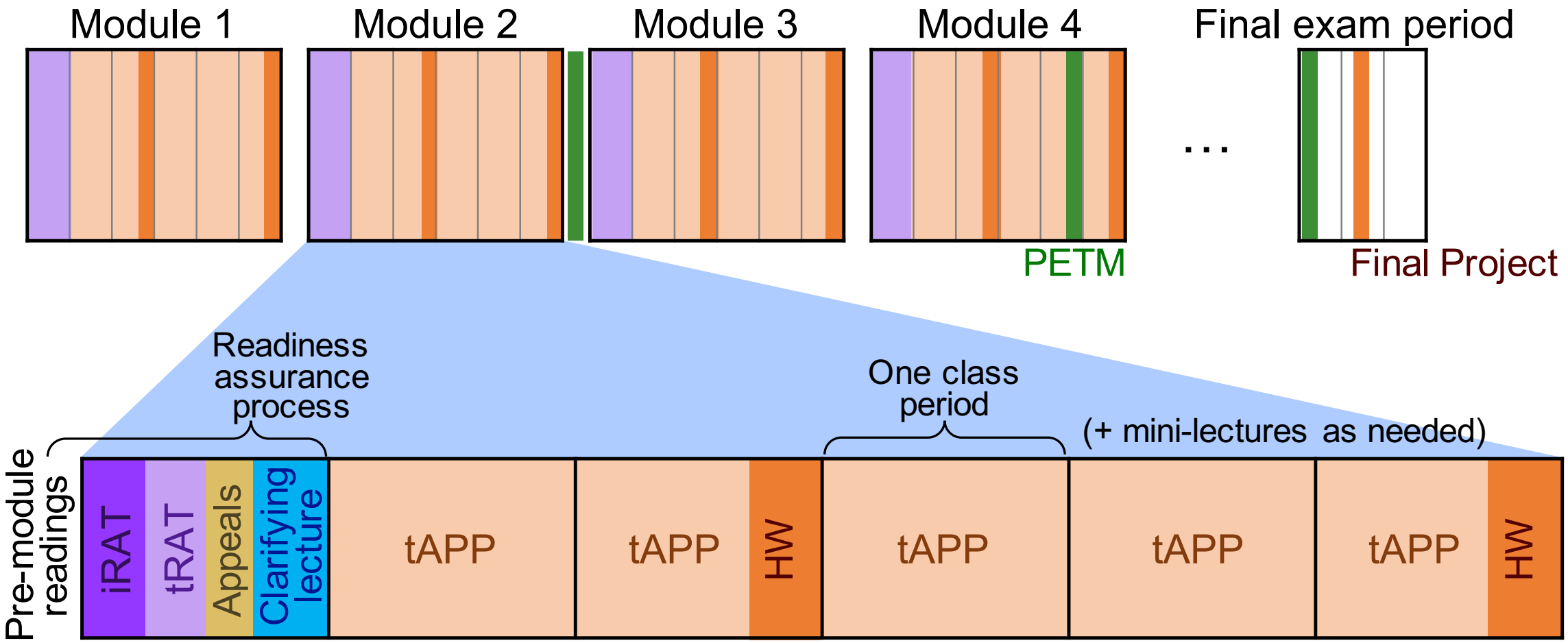
Some students had difficulty with Statistics topics like Simpson's paradox, hypothesis testing via simulating p-values, and confounding variables

Team-Based Learning in a nutshell

1. Create permanent teams of 4–5 students

2. Readiness Assurance Process:
   Readings, iRAT, tRAT, clarifying lecture, …
3. Team application exercises:
   Students work together to do data science
4. Peer Evaluation and Team Maintenance (PETM):
   PETM scores are 5–20% of the students' final grade

The Rhythm of TBL



7 Tips for Overcoming Barriers to Using TBL in Data Science

1. I don't know anything about TBL
2. Students don't come to class prepared
3. Some students are just not learning the programming skills
4. Students display poor time management
5. Loafers; students not contributing to their teams
6. It's difficult to teach programming as well as statistics concepts
7. Students forget everything they learned after your class

1. Read articles, take online and in-person workshops: teambasedlearning.org
2. Get a good book. Use iRATs, tRATs, and clarifying lectures
3. Let students share favorite Googled resources beyond the book
4. Reiterate student advice (from past semesters) throughout the course
5. Peer evaluation and team maintenance
6. I'm not so sure. I do only OK on this.
7. Coordinate and collaborate with other instructors. Reinforce concepts in future courses.

Resources to help teach Collaboration in Data Science

www.teambasedlearning.org

"The ASCCR Framework for Collaboration"
by Vance and Smith (2019) **2020 ASA Jackie Dietz Award**
bit.ly/asccrframe

New paper in *JSDSE* on Using TBL to Teach Data Science
https://www.tandfonline.com/doi/full/10.1080/26939169.2021.1971587

R for Data Science: *r4ds.had.co.nz*

These slides and more Collaboration resources: *osf.io/xmtce*

Eric.Vance@Colorado.EDU