



Dow Core Research & Development

ENABLING CITIZEN DATA SCIENTISTS WITH RSTUDIO ACADEMY

James Wade

RStudio Meetup | June 21st, 2022

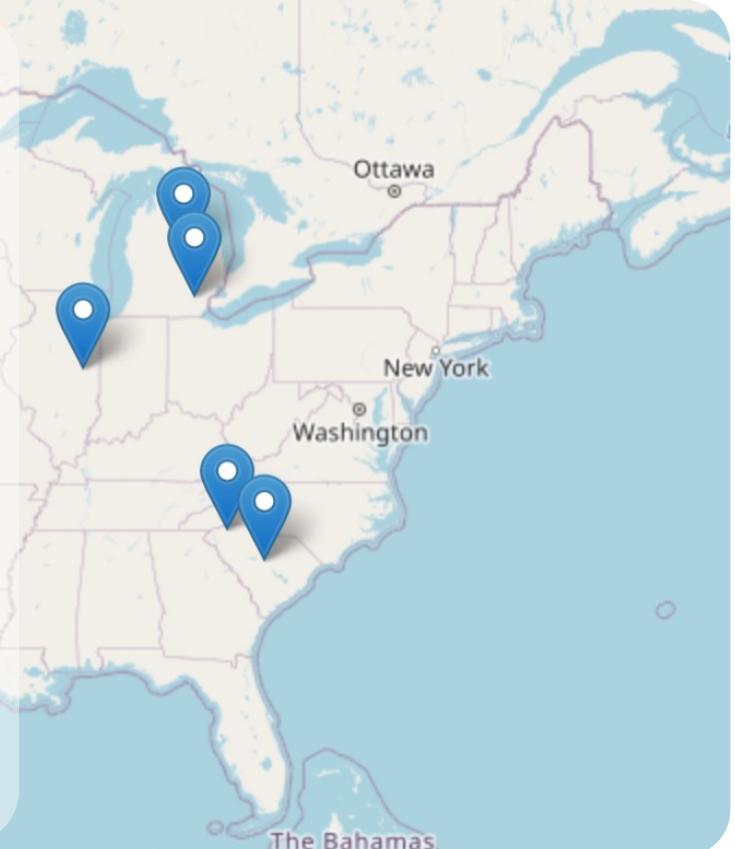
ABOUT ME

Education and Work Experience

- 👨‍🔬 2018, Chemist at Dow, Midland, MI
- 👨‍🔬 2017, Postdoc, University of Michigan
- 🎓 2017, PhD in Chemistry, University of Illinois
- 🎓 2012, BS in Chemistry, Furman University

Interests and Research Focus

- ⚖️ Structural Characterization of Materials
- 🌿 Discovery & Analysis of Sustainable Materials
- 🤖 High Throughput Research & Automation
- 📈 Data Science, Statistics, & AI / ML



ABOUT DOW

GLOBAL MEGATRENDS

CLIMATE CHANGE AND RESOURCE SCARCITY

↑50% GHG emissions, driven by energy demand in emerging economies

↑60% food needed to feed 9.7B people by 2050



MIDDLE CLASS GROWTH

Global middle class is expected to reach

5.5B
by 2030



DIGITAL TRANSFORMATION

Worldwide spending on digital transformation will reach

\$6.8 TRILLION
between 2020 and 2023



URBANIZATION

By 2045, the world's urban population will increase by

1.5X

to 6 billion people. Cities face systemic urban challenges in a post-COVID society.



CHANNELING DOW'S MATERIALS SCIENCE EXPERTISE AS WE COLLABORATE AND INNOVATE WITH CUSTOMERS AND PARTNERS TO **CREATE SOLUTIONS THAT POSITIVELY IMPACT THE WORLD**

PACKAGING

Providing options to keep foods fresher and deliver high-performing, durable, lighter-weight and recyclable packaging

INFRASTRUCTURE

Enhancing energy efficiency, circularity and durability in building and construction, appliances, adhesives, lubricants, and wire and cable

CONSUMER

Based on consumer specifications for innovative, convenient and more sustainable home and personal care options

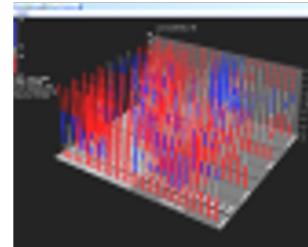
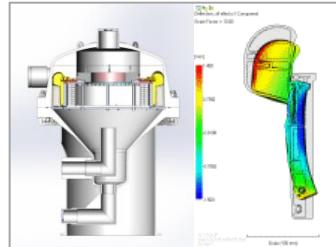
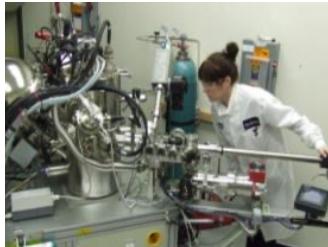
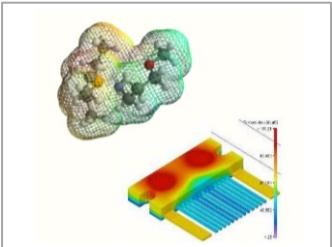
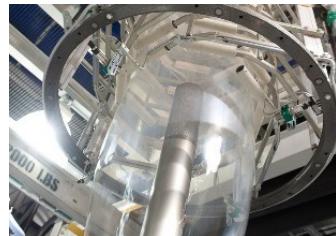
MOBILITY

Delivering options for lighter weight, longer range, comfort, safety and lower carbon footprint for vehicles that are connected, autonomous, shared and electric

ABOUT Dow R&D

High quality capabilities

high throughput • analytical • application testing • modeling and computation • data science

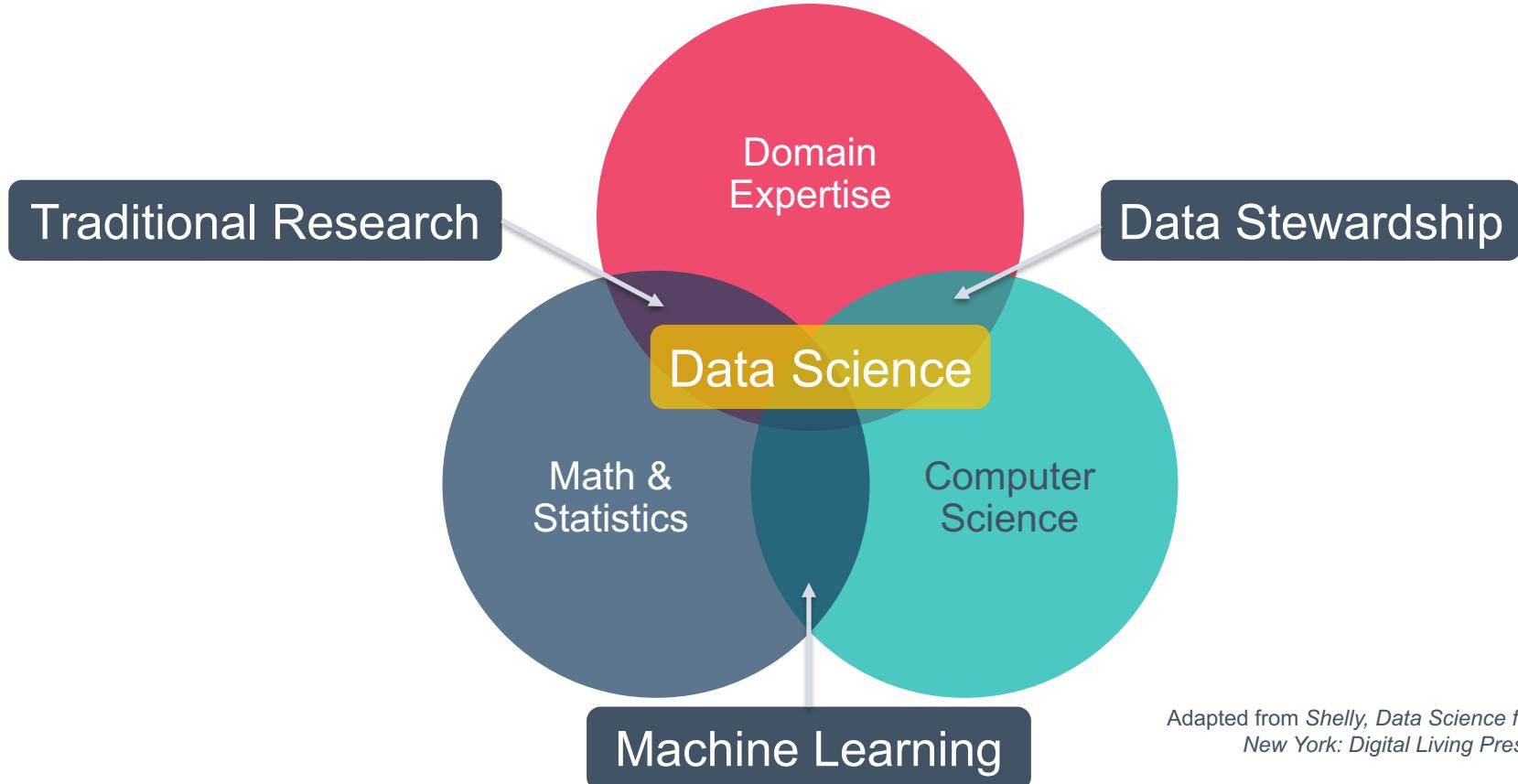


CITIZEN DATA SCIENCE AT Dow

As researcher in materials science research and development, we have **messy, unstructured, and scattered** data.

Creating value from our data requires researchers to **deeply embed in data analysis workflows** to impart their experience.

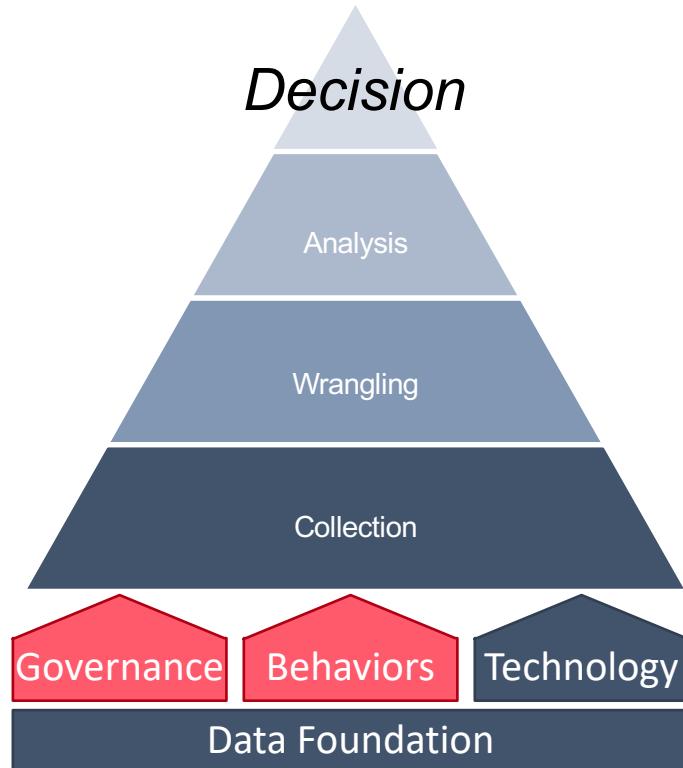
DATA SCIENCE AS AN INTERDISCIPLINARY ENDEAVOR



Adapted from Shelly, *Data Science for the C-Suite*.
New York: Digital Living Press, 2015. Print.



CITIZEN DATA SCIENCE FOR RESEARCH AND DEVELOPMENT



The **Citizen Data Scientist** enriches a data-centric cultural foundation by adhering to...

a **governance framework** for data best practices, and practicing **behaviors** that are data-centered.

The **Digital Infrastructure** forms...

a **technology base** focused on the needs of the “data pyramid.”

A CITIZEN DATA SCIENTIST'S GOAL

Construct a **tidy data** project that supports **collaborative**, **transparent**, and **reproducible** workflows.

Data Organization

1. Store & organize data at the level of a project
2. Store raw data in a managed system
3. Organize data to reflect observational hierarchy
4. Use tidy data tables
5. Document the structure of the data

Data Access

1. Make data accessible to both humans and machines
2. Remove artificial barriers to data access
3. Provide easy means to add and edit data
4. Visualize your data as soon & as much as possible

Data Analysis

1. Separate data storage and analyses
2. Create transparent analysis routines
3. Divide analysis into minimal components
4. Implement version control for data, analysis, & platform

Value Preservation

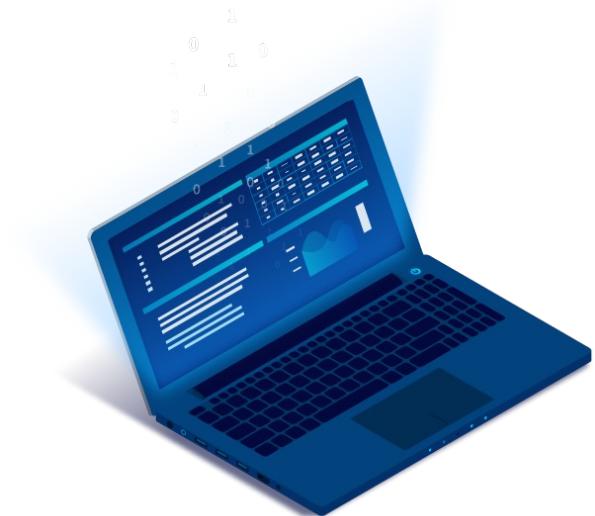
1. Package projects for long term storage
2. Transend platforms & languages with shareable interfaces



A WORKBENCH FOR CITIZEN DATA SCIENCE

As a Citizen Data Scientist in R&D, I need a data science workbench that offers

- an **approachable & collaborative** workspace that welcomes new researchers
- capabilities & convenience of a modern development environment with **extremely low barriers to exploring ideas** and starting projects
- safe & secure **data connectivity with minimal overhead**
- **simple, secure, & rapid** deployment requiring no new permissions or reviews



Our solution: RStudio Team

What to build?

Tidy data project

Where to build it?

RStudio Team

How to build it?



how to code in R |

RSTUDIO ACADEMY



Lessons

Learn the fundamentals for the project with interactive lessons.



Project milestones

Each week, recreate a progressive piece of the project and expand on it.



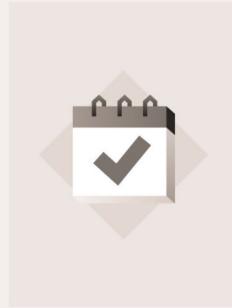
Group sessions

Share your work and ideas with your group at weekly sessions.



Mentor meetings

Meet regularly with a mentor to get personalized guidance.



Daily practice

Build long-lasting habits with spaced-repetition drills.

rstudio.com/academy



VS



WHAT YOU NEED FOR ACADEMY

- 👨‍🎓 Academy Learners – 5 to 7 per cohort
- 👩‍🏫 Cohort Mentors – from RStudio & Your Group
- 🔧 A Project – The closer to your work, the better

WHO IS A **GOOD** CANDIDATE?

1. **Intrinsic motivation:** you need to solve a data problem
2. You have a project to apply your knowledge
3. You have time and support to participate

WHO IS A **BAD** CANDIDATE?

- You are an experienced R user
 - "I know R, but I want to learn the 'right' way to code"
 - You have built and deployed a custom shiny app
 - You participate in #tidytuesday on a regular basis
 - You are good at tidy evaluation?
 - You have written or contributed to an R package?
- You want to learn a new skill
 - "It would be fun to learn how to code"
 - You don't have any data to use

WHAT MAKES A GOOD COHORT?

- Similar **work group**
- Similar **time zone**
- Similar **skill level**

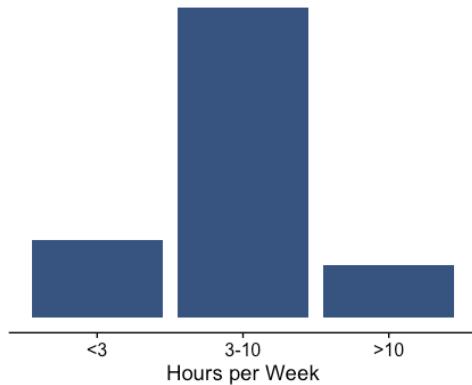
FEEDBACK FOR RSTUDIO ACADEMY

"The group sessions were a critical part of the accountability for me to ensure that I stayed on track and had inspiration from others that I could try in my own code. The format of the academy learning pages led to an outstanding user experience, with the bite-size learning messages and integrated code chunks to write and run."

"The RStudio Academy is an excellent way for researchers to get acquainted with the concept of using basic data science and programming to both analyze and communicate their data.... I would highly recommend to most researchers in Dow who are interested in any data science or data analytics roles."

Time Commitment of Academy

Most researchers spend between 3 and 10 hours per week



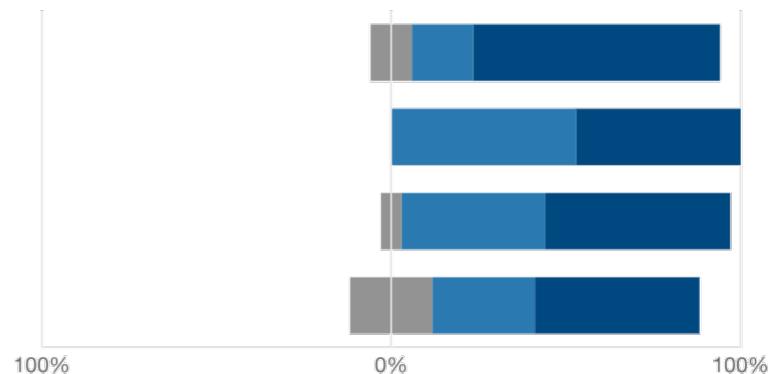
■ Strongly Disagree ■ Disagree ■ Neutral ■ Agree ■ Strongly Agree

RStudio Academy was a good use of my time.

I recommend other researchers participate in RStudio Academy.

Dow R&D should invest more in RStudio Academy or similar programs.

The RStudio mentor is a critical piece of RStudio Academy.



Academy only touches on reproducible reporting and deployment but sharing your analysis results with others is where much of the **value** in data analysis resides.



More Advanced Lessons



Transfer to Workbench



Community Building



RStudio Academy Success Story

Simplifying a Walk-up Lab with Shiny

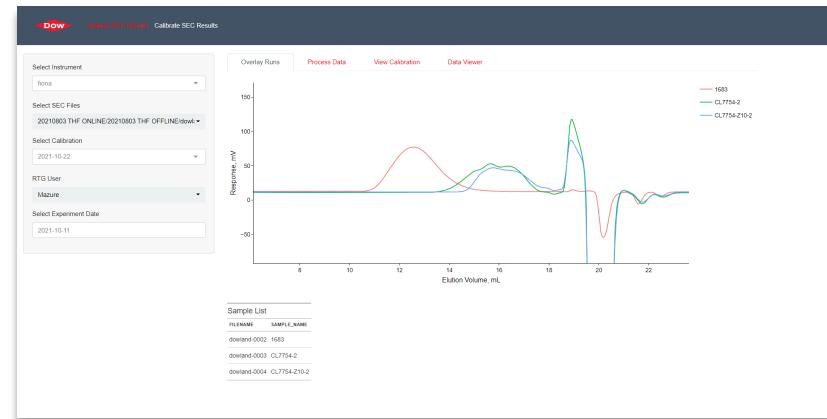
Data Analysis Tool for Polymer Analysis

Purpose

Empower walk-up users to **analyze their own data** without sacrificing integrity of data analysis

Benefit

Improved experience for walk-up users and researchers supporting walk up instrumentation

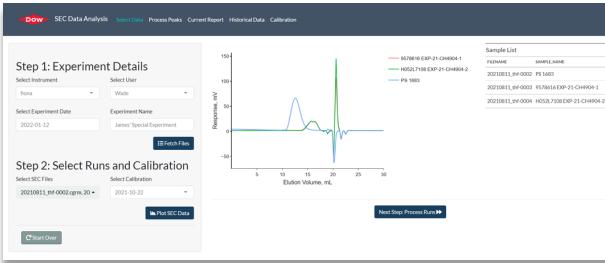


Polymer Characterization Data Analysis Tool

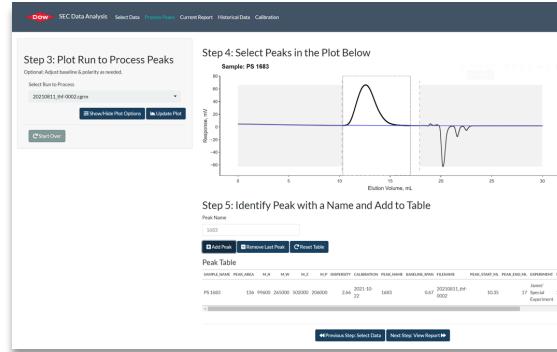
The graphic representations are presented here for illustrative purposes only and should not be construed as product specifications.

Data Analysis Tool for Polymer Analysis

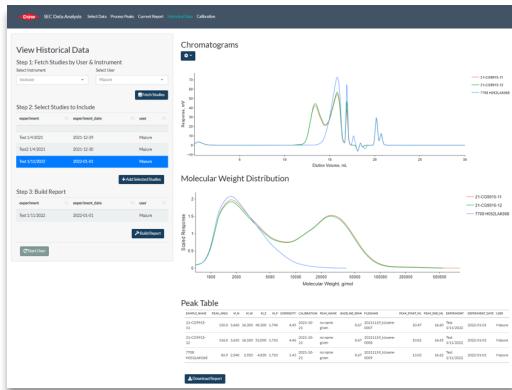
1 Select & Plot Data



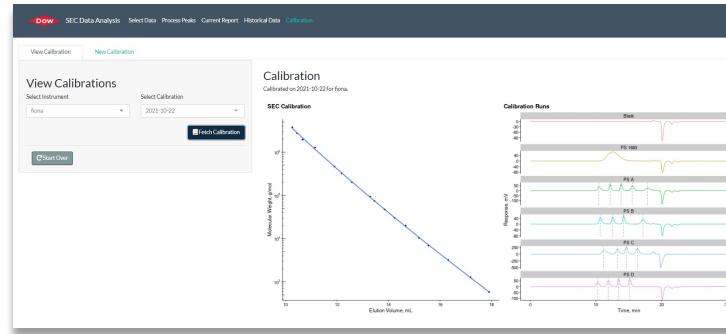
2 Analyze Data



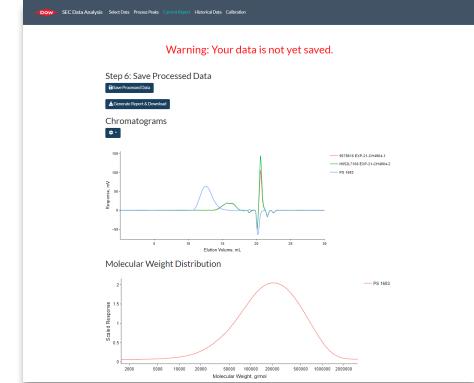
4 View Historical Data



5 Inspect Calibrations



3 Generate Report



The graphic representations are presented here for illustrative purposes only and should not be construed as product specifications.

ACKNOWLEDGMENTS



Tony Sokolov

linkedin.com/in/anatoliysokolov/

Thank you to our mentors & partners!

Dow

Michael Hausinger
Tim Licquia
Lynette Naler
Camila Saez
Academy Participants!

RStudio

Kristin Bott
Daniel Chen
Desiree De Leon
Hadrien Dykiel
Deirdre Francks
Brian Law
Katie Masielo
Eric Pite
Steven Smallberg





Seek

Together™



James Wade

-  jhwade@dow.edu
-  [@jameshwade](https://twitter.com/@jameshwade)
-  [JamesHWade](https://github.com/JamesHWade)
-  [in/james-h-wade](https://www.linkedin.com/in/james-h-wade)
-  jameshwade.com

Disclaimer: this presentation is provided in good faith for informational purposes only. Dow assumes no obligation or liability