ASMS 2019
ANNUAL
CONFERENCE
WORKSHOP

SOCALBIOINFORMATICS.COM

WORKSHOP OVERVIEW

PROGRAMMING SKILLS ARE IMPORTANT FOR ALL SCIENTISTS

- The data we routinely encounter is large and complex
- Software packages with canned analysis routines don't always cut it
- Having even basic programming skills can go a long way, and empower you to do new things and ask new questions
- Our world is becoming increasingly computational and we need to be a part of it

FROM LAST YEAR'S PARTICIPANT POLL

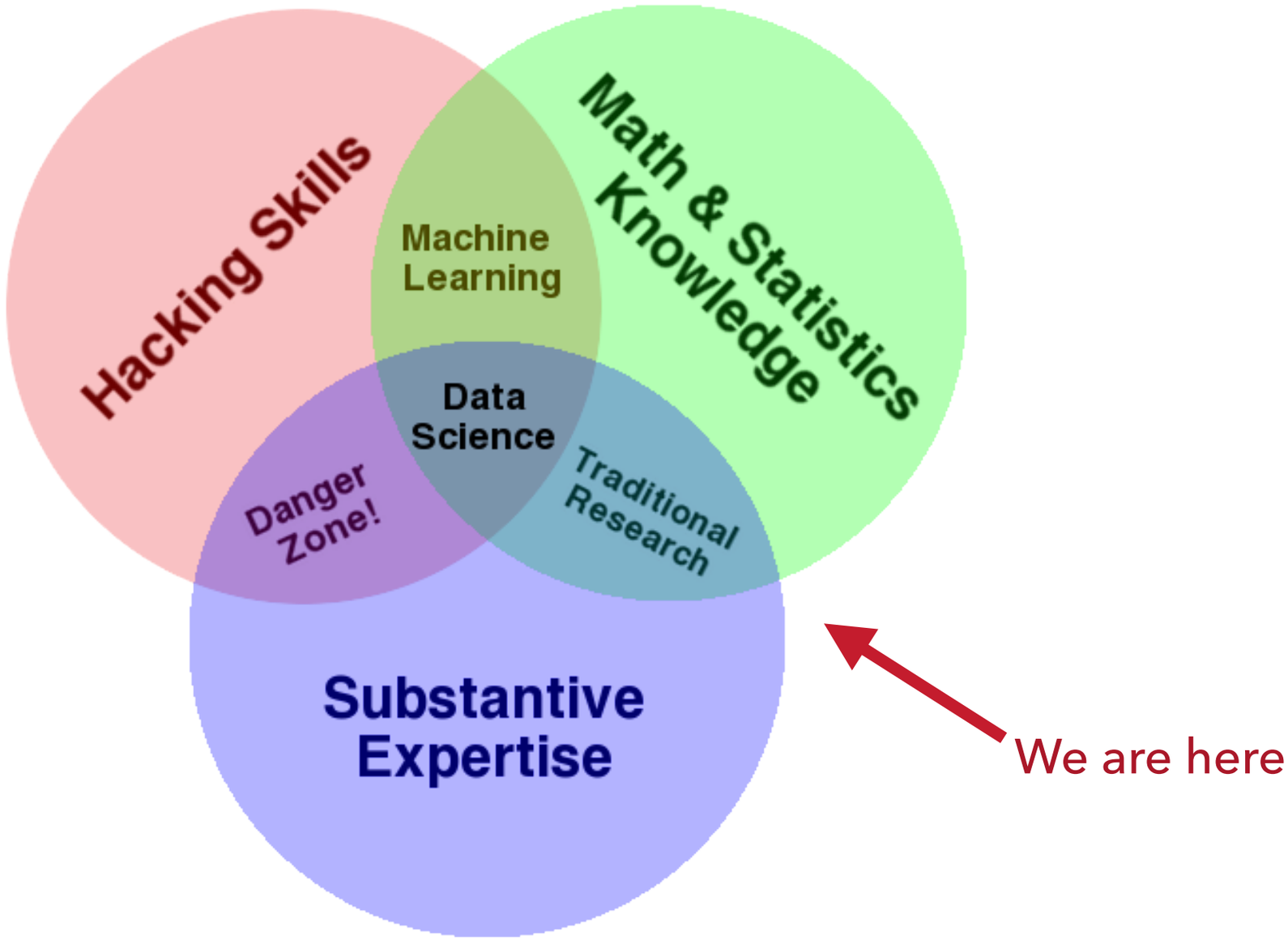
- >60% of responders (68 total) identified as beginners or having no experience with programming tools
- 42% identified Excel as their main analysis tool

Suggests that helping new users learn a tool like R is worthwhile

GOAL OF THE WORKSHOP

Help new and aspiring R users get through some of the initial hurdles often encountered when learning R and how to program

THE DATA SCIENCE VENN DIAGRAM



WHY NOT PYTHON?

- Sure, why not!
- Python is a great programming language
  - One of the most popular programming languages
  - Used extensively in the sciences (including data science)
  - Probably a better choice to learn to program with
  - Is a more general programming language
- You can’t go wrong with learning Python

SO THEN, WHY R?

- R was built specifically for doing data analysis: data-centric
- R has a great ecosystem for all aspects of the scientific data analysis process
  - data review, cleaning, munging, manipulation
  - data analysis, statistics, modeling
  - visualization, reporting
  - reproducibility
- R has an amazing, active, welcoming community

AGENDA

- Workshop overview
- Brief introduction to R & RStudio
- Data analysis with tidyverse packages
- Overview of mass spec R packages for data analysis
- Discussion and questions