



# Community-Based Research in a Regression Analysis Course

**Debra L. Hydorn**  
**University of Mary Washington**  
**MathFest 2024**

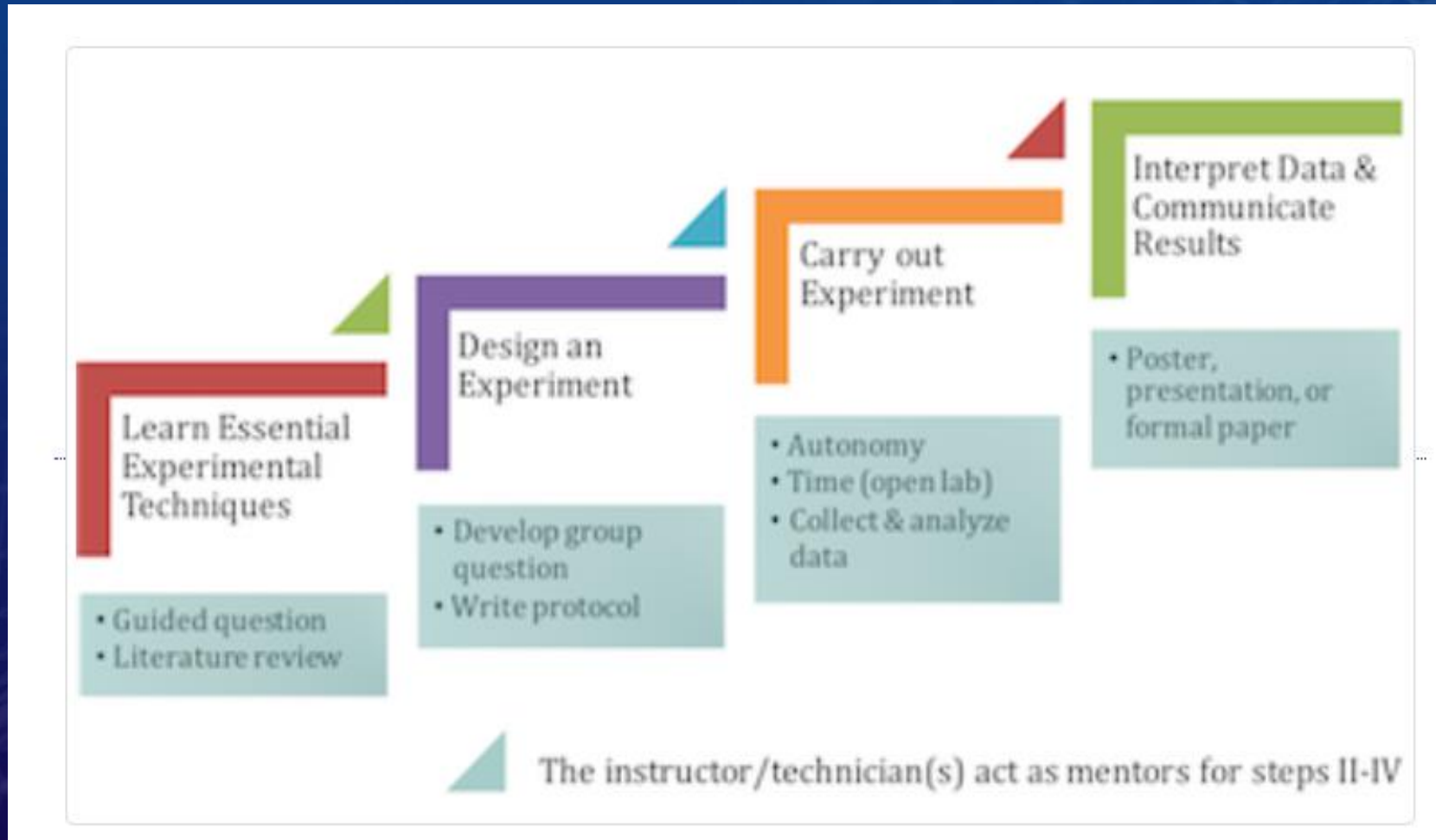
# Community-Based Undergraduate Research

- Two areas of interest:
- Undergraduate Research
  - Summer Science Institute
  - Independent Studies
- Community Engagement
  - Service Learning
  - CE Introduction to Statistics
- Goal: Community-based undergraduate research projects within one of my classes

# Fall 2023: CUR MIRIC Participation

- Mentoring the Integration of Research Into the Classroom
  - Biology Division of the Council on Undergraduate Research
    - Reading groups
    - Discussion groups
    - Advice
- Main resource:
  - “Developing Course-Based Undergraduate Research Experiences”, Jacqueline McLaughlin
  - Five characteristics of a CURE

# Four-step CURE Pedagogical Framework

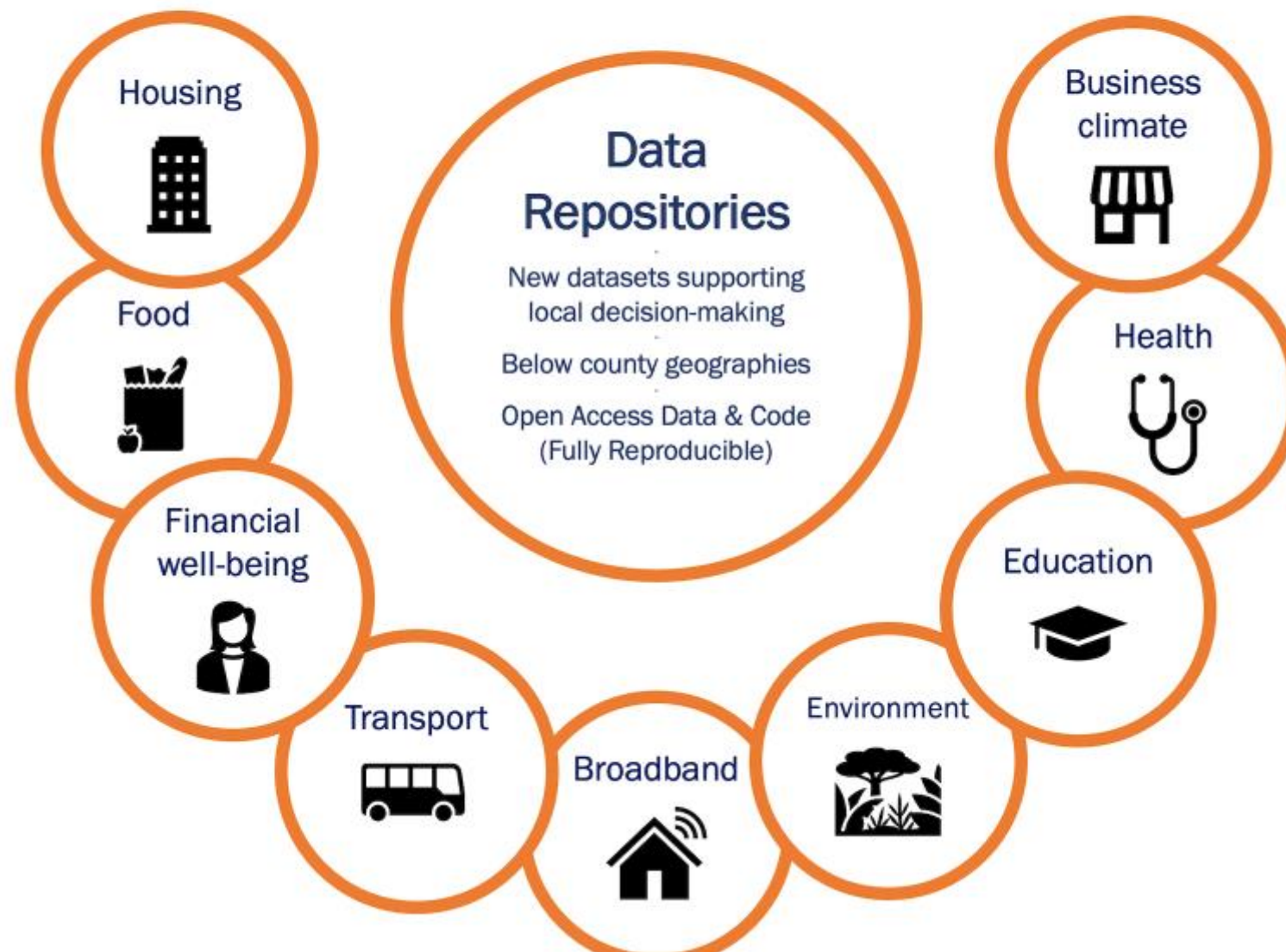


# Which course for Spring 2024?

- My original choice: MATH 300 Linear Algebra
  - Lots of applications that students could explore
  - Sensitivity Analysis
  - Modifications
  
- STAT 320 Applied Regression Analysis
  - Experience fitting models to real data
  - Data: University of Virginia's BioComplexity Institute
    - Social Impact Data Commons

# Data Repositories

Freely Available Open Source Data and Code



# Two Data Sources

- National Capital Region
  - Data by counties, tracts, zip code
  - Topics include Broadband Access, Education, Employment, Health
- Virginia Department of Health Data Commons
  - Counties in Virginia
  - Data used to create a Health Opportunity Index
  - Topics include Economic Opportunity, Built Environment, Consumer Opportunity, Social Impact

# Course organization

- Part I: Learn Essential Regression Analysis Methods
  - Review simple linear regression and using R
  - Introduce multiple regression and model selection methods
  - Homework assignments and in-class practice using BioComplexity data
- Part II: Design a Regression Project
  - Met with BioComplexity representatives – they described some areas of interest
  - Explore the data to identify dependent and independent variables
- Part III: Additional regression topics
  - Multicollinearity, Checking Assumptions, Remedial Measures, Assessing Goodness of Fit
- Part IV: Complete project and prepare report



# Project Process

- Exploratory Data Analysis
- Scatterplot Matrix and Correlations
- Fit the full model and investigate multicollinearity
- Use at least two model fitting methods
- Choose a best model
  - Check assumptions
  - Check for outliers and influential points
- Remedial measures
- Modify model as needed

# The Projects

- Nutrition and Food Security Project: How Economic and Educational Factors Affect the Average Meal Cost
- Understanding the Socioeconomic Determinants of Uninsured Populations (2 students)
- Economics – Exploration of Factors Impacting Median Household Income (3 students)
- Education – Factors Associated with the Number of 2-Year Schools with Computer Science Programs (2 students)
- Drugs – Factors Associated with Average Number of Trips to the ER for an overdose
- Education – Factors Associated with Average Years of Schooling
- Mental Health

# Lessons Learned

- More contact with our “client”
  - Specific research questions
- Provide project guidelines earlier in the course
- Provide specified format for the project report
- Be flexible, courageous even – make it work!
- Many High Impact Educational Practices tend to involve



“learning at the edge of  
chaos”

(Bertschinger and Natschläger 2004, Kleiman 2011)

# Future Plans

- Try again – regression analysis if possible
- Include Biocomplexity data in my classes this fall and Spring
  - Introduction to Statistics – data analysis labs
  - Statistical Methods – projects (regression, ANOVA,
  - Multivariate Statistics – CURE maybe

# Resources

- <https://urfm.psu.edu/mentors/developing-course-based-undergraduate-research-experiences>
- <https://uva-bi-sdad.github.io/sdc.intro/data.html>
- [https://uva-bi-sdad.github.io/capital\\_region/?selected\\_variable=avg\\_down\\_using\\_devices](https://uva-bi-sdad.github.io/capital_region/?selected_variable=avg_down_using_devices)
- [https://uva-bi-sdad.github.io/vdh\\_rural\\_health\\_site/?selected\\_variable=incarceration\\_rate\\_per\\_100000](https://uva-bi-sdad.github.io/vdh_rural_health_site/?selected_variable=incarceration_rate_per_100000)