# DS 7347
# High-Performance Computing (HPC) and Data Science
# Session 21

Robert Kalescky
Adjunct Professor of Data Science
HPC Research Scientist

July 5, 2022

Research and Data Sciences Services
Office of Information Technology
Center for Research Computing
Southern Methodist University

Group Discussion

Session Question

Common Data Structures

Readings and Assignments

# Group Discussion

- Explore various file formats for your data and compare performance

# Session Question

Describe as many places as you can to store data on an HPC system.

# Common Data Structures

- Disks
- Distributed file systems
- Memory
- Caches

- Physical data location
- Data access patterns

- Lists
- Dictionaries
- "Array of Structures"
- "Structure of Arrays"
- Tables

# Readings and Assignments

### Readings
None

### Project

- Complete non-optimized and basic workflow, reduce data or analysis complexity if needed.
- Commit the code to your project repo
- Due Tuesday, July 12.