only one of the ve traditional positions (PG, SG, SF, PF, C) and instead spend a percentage of their playing time at multiple positions, making positional data compositional. Further, given the desirability for versatile players, an argument can be made that traditional positions themselves are archaic. Using data from the 2016-17, 2017-18, and 2018-19 seasons, I explore how Bayesian hierarchical models can be used to estimate team defensive strength in three ways. First, only considering players classi ed by their majority traditional position. Second, by using compositional traditional positional data. Third, using compositional data from modern positions (archetypes) de ned by fuzzy k-means clustering. I nd that the fuzzy k-means approach leads to a modest improvement in both the root mean squared error and median 95% posterior predictive interval width for the test data, and, more importantly, identi es 11 modern archetypes that, when combined, are correlated with team win total and adjusted team defensive rating. The modern archetype compositions can be used by stakeholders to better understand team defensive strength.

February 9

Speaker: Larry D'Agostino
A liation: Stellantis Financial and R User Group
Title: Knowledge and Insight From Career in Data Science
Abstract: A career in data science and analytics can lead to many places. This talk is one example of many of those experience. The goal of the talk is to discuss four main topics and advice on getting into a career in data science. The hope is to impart some wisdom on the future generation in this eld of study. The four parts of the talk are 1) Starting Your Career, 2) Marketing Yourself, 3) Assembling Your Toolbox, and 4) Developing Data Skills.

February 23

Speaker: Alessandro Rinaldo
A liation: Department of Statistics and Data Sciences, University of Texas at Austin
Title: Sequential Change-point Detection for Network Data
Abstract: We study the change point detection settings in which we are presented with a stream of independent, labeled networks on a xed node set, whose distributions are piece-wise constant over time. Our goal is to determine, as soon as we acquire a new observation, whether the data collected so far have provided su cient evidence to infer that the underlying distribution has changed at the present time or in the near past. For sequences of Bernoulli networks, we formulate polynomial time CUSUM-based procedures and derive high-probability bounds on the corresponding detection delays with an explicit dependence on the network size, the entrywise and rank sparsity, and the magnitude of the change. We complement our analysis with minimax lower bounds, which we show are realized by NP-hard procedure. We also consider change point detection for sequences of multilayer hein8xegly product networks with xed and static latent positions but time-varying connectivity matrices. To handle this more complex and subtle scenario, we develop a sequential change point algorithm based on tensor methods and analyze its properties.

March 1

Speaker: Difeng Cai
A liation: Department of Mathematics, Southern Methodist University
Title: Data-Driven Kernel Matrix Computations: Geometric Analysis and Scalable Algorithms
Abstract: Dense kernel matrices arise in a broad he of disciplines, such as potential theory, molecular biology, statistical machine learning, etc. To reduce the computational cost, low-rank or hierarchical low-rank techniques are often used to construct an economical approximation to the original matrix. In this talk, we consider general m-by-n kernel matrices associated with possibly high dimensional data. We perform analysis to provide a straightforward geometric interpretation that answers a central question: what kind of subset is preferable for skeleton low-rank approximations. Based on the theoretical ndings, we present scalable and robust algorithms for approximating

general kernel matrices that arise in astrophysics, kernel ridge regression, Gaussian processes, etc. The e ciency and robustness will be demonstrated through extensive experiments for various datasets, kernels and dimensions.

March 22

Speaker: Michael Gallaugher
A liation: Department of Statistical Science, Baylor University
Title: Cluster Weighted Models with Heavy Tailed Matrix Variate Distributions
Abstract: Cluster weighted models (CWMs) are powerful clustering devices used in many regression-type analyses. Unfortunately, real data often present atypical observations that make the commonly adopted normality assumption of the mixture components inadequate. Thus, to robustify the CWM approach in a matrix-variate framework, we introduce ten CWMs based on the matrix-variate t and contaminated normal distributions. Furthermore, once one of our models is estimated and the observations are assigned to the groups, di erent procedures can be used for the detection of the atypical points in the data. Parameter estimation via an ECM algorithm will be discussed, and the method will be applied to simulated data, as well as a real data analysis on greenhouse emissions data.

April 5

Speaker: Yulun Liu
A liation: Peter O'Donnell Jr. School of Public Health, UT Southwestern
Title: Network meta-analysis made simple: A composite likelihood approach
Abstract: Network meta-analysis, also known as mixed treatments comparison meta-analysis or multiple treatments meta-analysis, expands upon conventional pairwise meta-analysis by simultaneously synthesizing multiple interventions in a single integrated analysis. Despite the growing popularity of network meta-analysis within comparative e ectiveness research, it comes with potential challenges. For example, within-study correlations among treatment comparisons are rarely reported in the published literature. Yet, these correlations are pivotal for valid statistical inferences. As demonstrated in earlier studies, ignoring these correlations can lead to in ated mean squared errors in estimates and inaccurate standard errors. In this talk, I will introduce a composite likelihood-based approach that guarantees accurate statistical inferences even without knowledge of these within-study correlations. The approach is computationally robust and e cient, with substantially reduced computational time compared to the state-of-the-science methods implemented in R packages. The proposed method has been evaluated through extensive simulations and applied to two important applications, including a network meta-analysis comparing interventions for primary open-angle glaucoma and another comparing treatments for chronic prostatitis and chronic pelvic pain syndrome.

April 26

Speaker: Pedro Maia
A liation: Department of Mathematics, University of Texas at Arlington
Title: Mathematical models and methods in computational neurology
Abstract: The emerging eld of computational neurology provides an important window of opportunity for modeling of complex biophys(eldex)-as