Abstracts

Bayesian Time-Dynamic Mixed-Effects Modeling of Longitudinal Data

Makena Grigsby, University of California, Riverside

Longitudinal studies often collect repeated measurements on subjects over time, making it important to account for within-subject dependence and changing covariate effects. In many applications, measurements are not collected on a common schedule; subjects may have different numbers of observations, irregular follow-up times, or covariates that vary over time. This work presents a Bayesian time-dynamic mixed-effects model framework for longitudinal outcomes with balanced or irregular observation schedules. Time-varying regression coefficients are modeled using spline basis expansions, allowing covariate effects to change smoothly over the observed time domain while subject-specific random effects account for dependence among repeated measurements. Model parameters, including coefficient functions, variance components, and random effects, are estimated using Markov chain Monte Carlo sampling, with posterior summaries used for inference and uncertainty quantification. The framework is implemented in JAGS for Gaussian, Bernoulli, and Poisson outcomes, allowing the approach to be applied to continuous, binary, and count longitudinal data. Simulation studies examine performance under balanced longitudinal data, irregular measurement times, and time-varying covariates. Overall, this work provides a flexible Bayesian framework for modeling longitudinal data when covariate effects are not constant over time and observation schedules may vary across subjects.

When the Data Goes Quiet: Using HDBSCAN to Reveal Gendered Street Crime Patterns Hidden in Official Crime Data

Farah Ahmed, HORUS VISTA LABS and OSU

Official crime taxonomies were not built to surface how gendered violence shapes daily life in public space. When sexual assault, stalking, and harassment are folded into broad categories dominated by property crime and theft, the patterns that determine how women move through cities disappear into statistical noise.
This talk presents an unsupervised learning approach applied to over 387,000 reported crimes in Chicago from 2023 to 2025, filtered to non domestic incidents involving gendered violence. Rather than relying on predefined categories or administrative boundaries, we used HDBSCAN, a density based clustering algorithm, to let spatial and temporal patterns emerge on their own. The model incorporated geographic coordinates alongside a composite risk score weighting factors such as location isolation and time of day.
The results showed that just over a third of gendered street crimes form dense, geographically stable clusters, meaning the same neighborhoods surface as high risk year over year. The remaining crimes appear as scattered incidents that traditional hotspot mapping would miss entirely, requiring different intervention strategies than concentrated crime does.
This methodology is not specific to Chicago or to any single reporting system. Every city that tracks crime using categories built around property offenses has the same blind spot: the crimes that most affect how women navigate public space are the crimes most likely to vanish in the data.

From Gut to Data: A Replicable Forecasting Framework for Rural Micro-Retailers in Data-Poor Environments

Farah Ahmed, HORUS VISTA LABS and OSU

Small independent retailers in rural and underserved communities often operate on intuition, not by choice, but because formal data infrastructure does not exist. There are no point-of-sale systems, no inventory software, and no historical transaction records, yet these stores are frequently the only food access point for their communities.
This talk presents a replicable three-layer data science framework developed with a SNAP-authorized independent grocery in rural South Dakota that had no existing data infrastructure. Starting from manually collected sales records, we built a demand forecasting pipeline that progressed through SARIMA, SARIMAX, XGBoost, and ultimately linear regression, chosen not for sophistication but because the store owner could validate its coefficients against decades of operational experience. Stakeholder adoption, rather than model complexity, defined success.
SHAP values applied during the XGBoost phase revealed which features actually drove demand and informed the decision to prioritize an interpretable final model. The framework's roadmap extends to inventory intelligence and a plain English LLM interface that would let owners query their data conversationally. The methodology was designed for portability: minimum viable data collection, interpretability over accuracy maximization, and owner validation as a feature rather than a limitation. This same framework could be applied to any small business or micro retail across the globe.

Collective Intelligence Lead to Open Innovations

Nisha Daniel Raju Daniel, Bayer AG

Open innovation hackathons have evolved from isolated corporate experiments into an essential strategic framework within the pharmaceutical sector to drastically reduce research costs, accelerate clinical pipelines, and democratize cross-disciplinary domain expertise. This abstract demonstrates how global leaders like Bayer, Novartis, Pfizer, Johnson & Johnson, and Sanofi utilize high-energy "vibe coding" environments to bridge traditional science-tech divides, transforming passive internal workforces into active solution builders who crowdsource critical advancements in data science, trial optimization, rare disease diagnostics, and automated regulatory processing. While these dynamic sprints offer immense advantages by enabling rapid prototyping, modular tool creation, and immediate risk validation without heavy capital expenditure, they simultaneously present unique organizational vulnerabilities, including complex intellectual property licensing, internal resistance to change management, and long-term technical debt. Ultimately, this study redefines how the life sciences industry evaluates innovation success, shifting the metric from finalized software products to the velocity of insights generated and long-term behavioral transformation.