

Research Topics & Ideas: Data Science and Analytics

About this document

This PDF contains 50 **data science and analytics research topics** for dissertations and theses, organized into 10 areas of data science and analytics, from analytics in health and medicine to recommender systems and personalization. Each topic is written as a researchable question and linked to a completed dissertation (2022–2026, 19 universities), so you can see how it was studied in practice. Use it to shortlist data science and analytics dissertation topics and thesis ideas, then narrow one down to your own population, setting and timeframe.

By [Derek Jansen](#) (MBA) · Reviewed by [Eunice Rautenbach](#) (DTech) · Grad Coach · October 2026

Online version: gradcoach.com/research-topics-ideas-data-science-analytics/

© 2026 Grad Coach. Free to share and adapt under [CC BY 4.0](#) with credit to Grad Coach and a link to gradcoach.com/research-topics-ideas-data-science-analytics/. Research questions and thesis titles quoted here remain their authors'; each dissertation keeps its own license.

Contents

Analytics in health and medicine · 5 topics	2
Analytics in industry and operations · 5 topics	2
Anomaly detection and monitoring · 5 topics	3
Causal inference and experimentation · 5 topics	3
Data quality, pipelines and discovery · 5 topics	3
Data visualization and communication · 5 topics	4
Explainable and trustworthy analytics · 5 topics	4
Learning from complex and shifting data · 5 topics	5
Predictive modeling and forecasting · 5 topics	5
Recommender systems and personalization · 5 topics	6
How to turn a topic into your research question	7

Analytics in health and medicine

1. Which methodological choices decide what a bedside monitor's data appears to say?

See it in a real dissertation: [Exploring Methodological Decisions for Deriving Clinical Insights from Inpatient Time Series Physiomarkers](#) · Reanna Panagides · University of Virginia · 2026

2. Can outcomes after pediatric cardiac surgery be predicted well enough to change a decision?

See it in a real dissertation: [Predictive Models for Pediatric Cardiac Surgery Outcomes](#) · Humphrey, Maeve Elizabeth · Harvard University · 2024

3. How far in advance can the risk of injury during childbirth be predicted?

See it in a real dissertation: [A predictive model for the risk of fetal and maternal injury during childbirth](#) · Alice May Collier · University of Oxford · 2025

4. Do social determinants of health explain who survives sepsis?

See it in a real dissertation: [Retrospective Study and Predictive Modeling of the Impact of Social Determinants of Health on Sepsis Outcomes](#) · Mwembezi A. Nyelele · Louisiana State University · 2025

5. Whose outcomes does a clinical risk score get wrong, and can that be corrected?

See it in a real dissertation: [Advancing Equitable and Patient-Centered Clinical Decision Support in Acute Myocardial Infarction: Insights from the CADILLAC Risk Score](#) · Emily García · Purdue University · 2026

Analytics in industry and operations

6. Which capabilities does a supply chain need before AI is worth integrating?

See it in a real dissertation: [A Capability Maturity Model for Artificial Intelligence Integration in Supply Chain Management](#) · Becklines Lordt · Dakota State University · 2026

7. What makes a farmer switch a brand they have bought for a decade?

See it in a real dissertation: [Determinants of Seed Brand and Retailer/Dealer Loyalty and Switching Behavior Among U.S. Corn and Soybean Farmers](#) · Luke Lonny Grieser · Purdue University · 2026

8. Which signals predict that an employee is about to leave, and how early?

See it in a real dissertation: [Analyzing Factors Influencing Employee Turnover in Tech Companies: A Predictive Modeling Approach](#) · Shinjon Ghosh · Illinois State University · 2025

9. Does a real-time information system change where trucks end up parking?

See it in a real dissertation: [Evaluating and improving the performance of Iowa truck parking information and management system](#) · Yilun Yang · Iowa State University · 2023

10. Can a refrigeration unit learn what normal looks like without anyone labeling it?

See it in a real dissertation: [Self-supervised Learning for Refrigeration System Monitoring in Iiot](#) · Xavier Alexander Lopez · Purdue University · 2026

Anomaly detection and monitoring

11. Where should the threshold sit on an anomaly detector that has to run continuously?

See it in a real dissertation: [Stability-aware Dynamic Thresholding for Online-like Unsupervised Network Anomaly Detection](#) · Dong Gun Kim · Purdue University · 2026

12. How do you find the odd event in a smart building that logs everything?

See it in a real dissertation: [Anomaly Detection in Smart Iot Environments Using Unsupervised Event Filtering and Graph-Based Forecasting](#) · Md Asif Tanvir · Missouri State University · 2026

13. What counts as unusual on a campus, and who decides the label?

See it in a real dissertation: [Anomaly Detection & Classification Using Machine Learning for Campus Surveillance](#) · Shakshi Richhariya · San Jose State University · 2024

14. What representation lets a model recognize something it has never seen?

See it in a real dissertation: [Learning Representations for Novelty and Anomaly Detection](#) · Ranya Almohsen · West Virginia University · 2023

15. Can augmented data teach a model about events that almost never happen?

See it in a real dissertation: [Prediction of Anomalous Events with Data Augmentation and Hybrid Deep Learning Approach](#) · Ahmed Shoyeb Raihan · West Virginia University · 2024

Causal inference and experimentation

16. What has to change before causal inference works on data nobody designed?

See it in a real dissertation: [New solutions for real-world causal inference](#) · X Lin · University of Oxford · 2025

17. Can the design of an experiment carry the inference, without a model of the outcome?

See it in a real dissertation: [Design-Based Causal Inference: Applications to Social Sciences and Industry](#) · Ham, Dae Woong · Harvard University · 2024

18. Which causal questions in political economy can an instrument actually answer?

See it in a real dissertation: [Essays on Causal Inference and Its Applications in Empirical Political Economy](#) · Zeyang Yu · University of Chicago · 2023

19. How many different quantities could a study be estimating without ever saying which?

See it in a real dissertation: [Enumeration of Estimands for Causal Effect Identification](#) · Hyunchai Jeong · Purdue University · 2026

20. Where does a treatment effect change, and how would you show that it did?

See it in a real dissertation: [Statistical Methods for the Study of Effect Modification and Spatial Causal Inference: Theory and Applications](#) · Cohn, Eric Ryan · Harvard University · 2024

Data quality, pipelines and discovery

21. Where did this table come from, and can the answer be reconstructed after the fact?

See it in a real dissertation: [Reconstructing the Lineage of Artifacts in Data Lakes](#) · Mohammed Suhail Rehman · University of Chicago · 2023

22. What would a system have to do to find the dataset worth joining to yours?

See it in a real dissertation: [Navigating the Open World of Data: Proactive Systems for Data Discovery and Correlation Exploration](#) · Yue Gong · University of Chicago · 2025

23. Can a pipeline diagnose its own data quality problems before the model sees them?

See it in a real dissertation: [AutoCleanML: An Automated Data-Centric Framework for Diagnosing and Repairing Data Quality Issues in Big-Data Machine Learning Pipelines](#) · Rahul Kamma · Purdue University · 2026

24. Which additions to a training set improve a model, and which only enlarge it?

See it in a real dissertation: [Improving machine learning models through enrichment of training data](#) · Yuanzhe Jin · University of Oxford · 2025

25. How much does a co-authorship network reveal about how a research field works?

See it in a real dissertation: [Social Network Analysis of a Biomedical Research Co-Authorship Network](#) · Amrina Ferdous · Boise State University · 2023

Data visualization and communication

26. Which color and shape combinations let a reader tell categories apart reliably?

See it in a real dissertation: [Automating Effective Categorical Palette Design: Empirically Grounded Tool for Color, Shape, and Redundant Encodings](#) · Chin Tseng · University of North Carolina at Chapel Hill · 2026

27. What does a person need to see to keep control of a system that decides faster than they do?

See it in a real dissertation: [Visual Interfaces for Enhancing Transparency and Control in Human-AI Collaboration](#) · Ian Thomas · University of North Carolina at Chapel Hill · 2026

28. What is the right way to show a robot's decisions over space and time?

See it in a real dissertation: [Visualizing Spatial-temporal Data in Robotics: An Overview](#) · Yan Zhu · University of North Carolina at Chapel Hill · 2026

29. How much context can a chart carry before it stops helping the decision?

See it in a real dissertation: [Balancing Context and Clarity Through Visualizations for Better Decision-Making](#) · Bhavana Doppalapudi · University of South Florida · 2024

30. Can visual analytics surface a material property relationship a model missed?

See it in a real dissertation: [Visual Analytics and Modeling of Materials Property Data](#) · Diwas Bhattarai · Louisiana State University · 2022

Explainable and trustworthy analytics

31. Would an operator trust an autoscaler more if it explained its decisions?

See it in a real dissertation: [An Explainable Ai-driven Approach to Cloud Autoscaling Using Machine Learning and Variable Selection](#) · Meghana Gorripati · Purdue University · 2026

32. What makes an explanation of a recommendation good enough to act on?

See it in a real dissertation: [Human-in-the-Loop Framework for eXplainable Recommender Systems Based on Explanation Information Quality](#) · Yeonbin Son · University of Virginia · 2026

33. Can you get more out of the data you have without trusting the model less?

See it in a real dissertation: [Enhanced Data Utilization for Efficient and Trustworthy Deep Learning](#) · Zhuokai Zhao · University of Chicago · 2024

34. Which interpretability method survives contact with an actual clinician?

See it in a real dissertation: [Interpretable AI for Medical Data: A Machine Learning Approach](#) · Irem Öztürk · San Francisco State University · 2025

35. Does an offline evaluation predict how a recommender performs in front of people?

See it in a real dissertation: [Best Practices for Offline Evaluation for Top- N Recommendation](#) · Ngozi Ihemelandu · Boise State University · 2024

Learning from complex and shifting data

36. Why does a learned representation fail on data that looks the same to a person?

See it in a real dissertation: [Towards Generalizable Representation Learning Across Domains](#) · Hossein Kashiani · Clemson University · 2026

37. How much can a pretrained model be compressed before adaptation stops working?

See it in a real dissertation: [Efficient Reuse of Pretrained Transformers: Closed-Form Gradient-Free Compression and Its Composition with Downstream Adaptation](#) · Boxiang Zhang · Purdue University · 2026

38. Which assumptions have to be built in before a model learns structure rather than correlation?

See it in a real dissertation: [Causal inductive biases for learning Structured World Models](#) · I Lei · University of Oxford · 2025

39. What do you see in a complex system when you look at it one agent at a time?

See it in a real dissertation: [From micro to macro: using a granular lens to analyze global complex systems](#) · Valentina Semenova · University of Oxford · 2024

40. What does predictive analytics look like across devices that cannot pool their data?

See it in a real dissertation: [Advancing Predictive Analytics in Connected Systems: Federated Learning, Auxiliary Information Integration, and Adaptive Architectures](#) · Jingyi Gao · University of Virginia · 2026

Predictive modeling and forecasting

41. How should uncertainty be handled when a model is chosen by its own predictions?

See it in a real dissertation: [Prediction-powered machine learning for model selection and uncertainty](#) · Vikrant Mukul Shirvaikar · University of Oxford · 2025

42. What breaks in statistical learning when biomedical data arrive in three shapes at once?

See it in a real dissertation: [Advances in statistical learning for complex data in biomedical research](#) · Hyunjae Cho · University of Virginia · 2026

43. Can generated data stand in for measurements you were never going to get?

See it in a real dissertation: [Generative AI and Simulation-Driven Predictive Modeling Framework Under Data Scarcity](#) · Seung Ho Woo · Purdue University · 2026

44. When is a Bayesian model worth its cost on healthcare data?

See it in a real dissertation: [Predictive Modeling for Healthcare Data Using Nonlinear Bayesian Methods](#) · Prince Kofi Asare · Illinois State University · 2025

45. Does a battery model built in a lab survive field testing?

See it in a real dissertation: [Acceleration Battery Predictive Model Utilizing Field-based Testing](#) · Peterson Bohannon · Fort Hays State University · 2026

Recommender systems and personalization**46. Can synthetic data make a recommender private without making it useless?**

See it in a real dissertation: [Synthetic Data Generation for Accurate, Fair, and Private Recommender Systems](#) · Derek Lilienthal · San Jose State University · 2024

47. What does a language model add to a recommender that matrix factorization cannot?

See it in a real dissertation: [Toward Generative Recommender Systems with Large Language Models: Foundation, Augmentation, and Alignment](#) · Yaochen Zhu · University of Virginia · 2026

48. How do you learn from clicks without also learning the bias in what got shown?

See it in a real dissertation: [New accurate, explainable, and unbiased machine learning models for recommendation with implicit feedback](#) · Khalil Damak · University of Louisville · 2022

49. Which biases in a recommender come from the users and which from the system?

See it in a real dissertation: [New debiasing strategies in collaborative filtering recommender systems: modeling user conformity, multiple biases, and causality](#) · Mariem Boujelbene · University of Louisville · 2022

50. Should a hashtag be recommended from the content, or from who else used it?

See it in a real dissertation: [Graph Deep Learning Based Hashtag Recommender for Reels on Social Media](#) · Sriya Balineni · San Jose State University · 2023

How to turn a topic into your research question

A topic is a starting point, not something you can research yet. These four steps take you from a topic on this list to a focused question you can actually answer.

- 1. Shortlist three, not one.** Pick three topics from this list that genuinely interest you and fit the access you have to people and data.

Recommended reading: [How to find and choose a dissertation topic](#)

- 2. Read the dissertation behind each one.** Its literature review and its "recommendations for future research" section show you what is already known and what is still open.

Recommended reading: [How to find a research gap](#)

- 3. Narrow it to your setting.** Name one population, one setting and one thing you will measure or explore. That turns a broad topic into a question you can finish.

Recommended reading: [How to write a research question](#)

- 4. Line up your aims, objectives and questions.** Your committee will check that they agree.

Recommended reading: [Research aims, objectives and questions](#)

Work through these once and you will have a focused question rather than a broad topic, which is what the rest of your proposal is built on.