

Data Science Center of Excellence (DSCoE)

AI for all

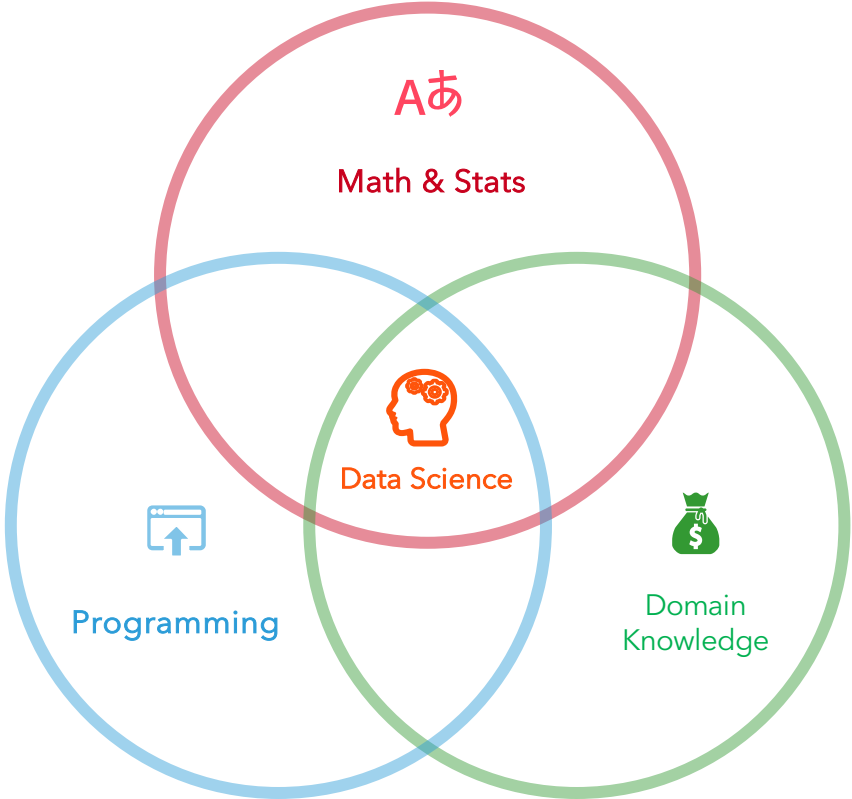
Agenda

- What is Data Science and why it matters
- How to democratize Data Science
- Data Science center of excellence
- Questions

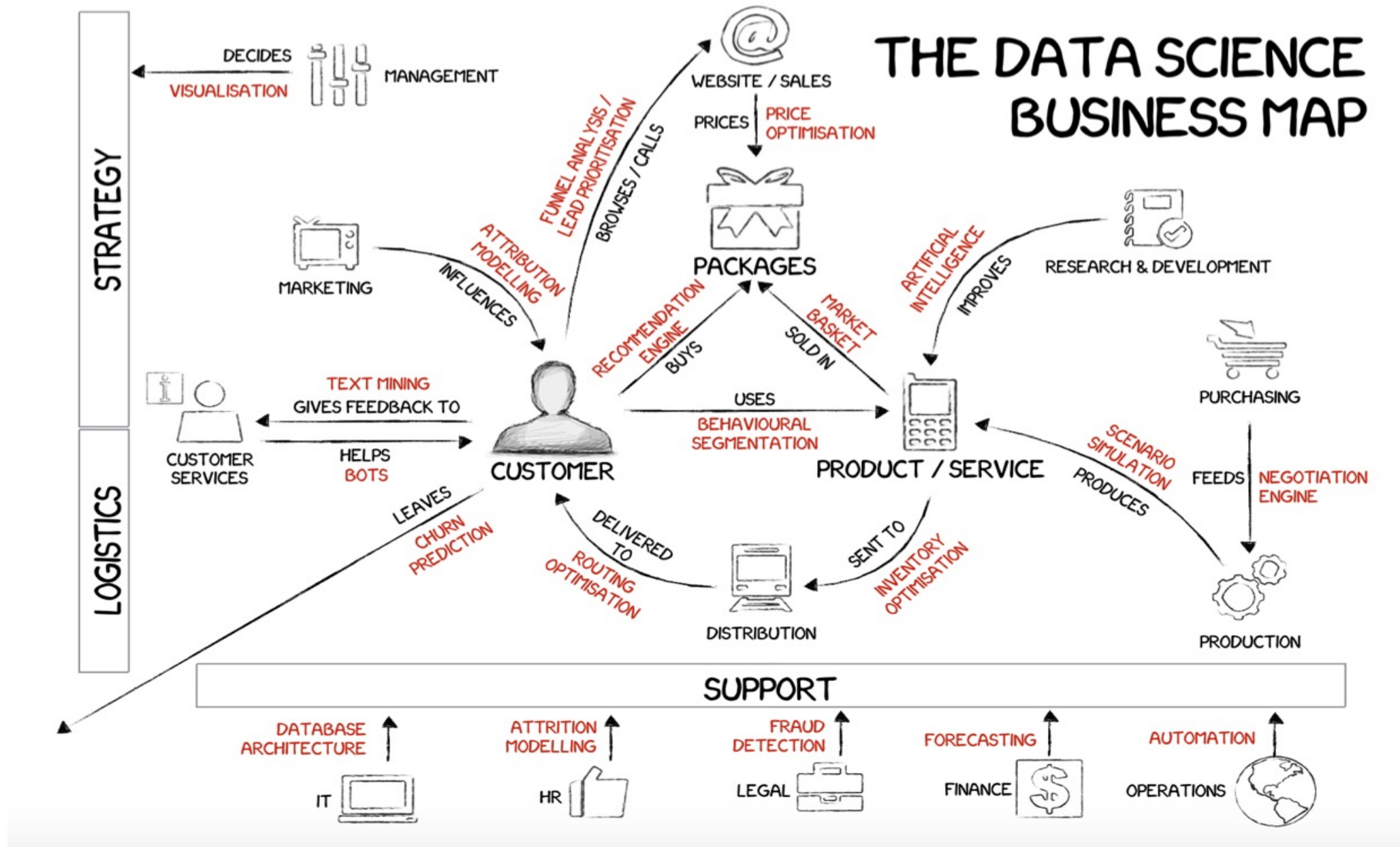
What is Data Science

“Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from noisy, structured and unstructured data, and apply knowledge and actionable insights from data across a broad range of application domains.”

--- Wikipedia

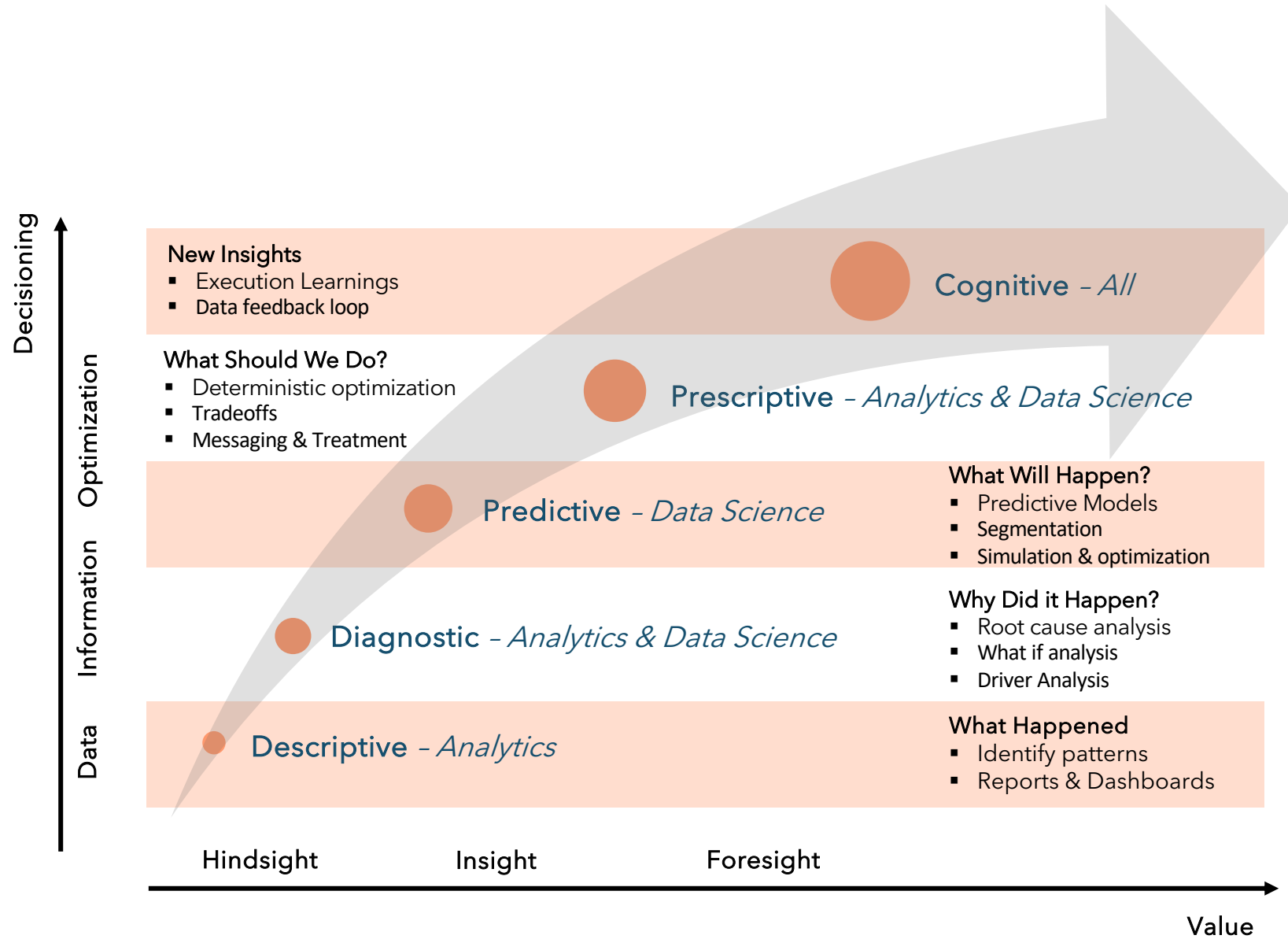


Why Data Science matters to business



Source - Medium.com (phelixjuma)
https://miro.medium.com/max/1400/1*lo_e9vvgMrEwCr3ApDTulQ.png

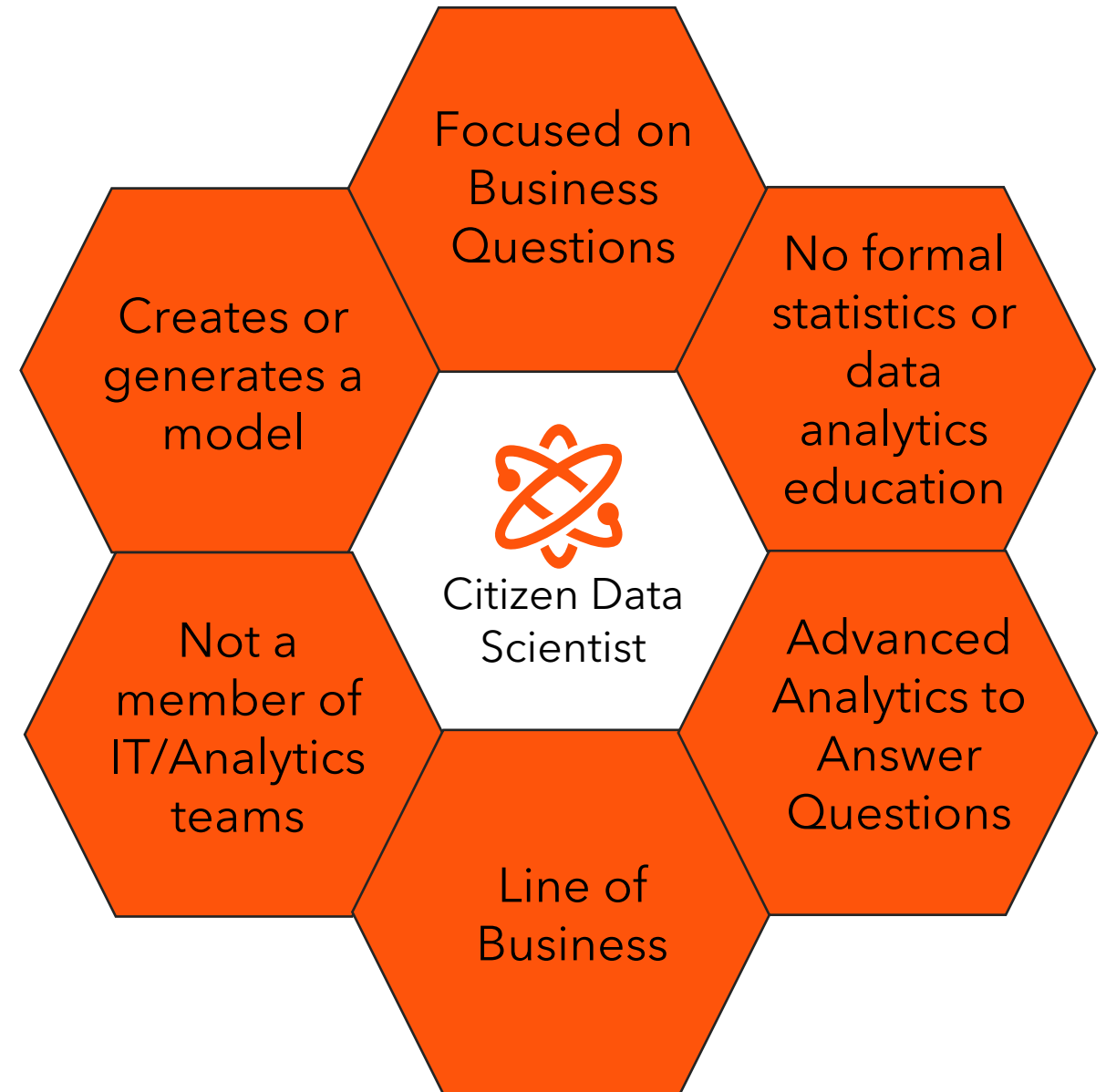
Role of Data Science



Citizen Data Scientist

A person who creates or generates machine learning models that use advanced diagnostic, predictive, and prescriptive capabilities, but whose primary job function is outside the field of statistics and analytics

-- Gartner



How to democratize Data Science

Data science is transforming business and generating new value across enterprise. Data science empowering management and officers to make better decision



AI for business

AI and ML to Empowering management and officers to make better decision



Trusted Data

Access to quality data



Knowledge/Training

Training



Scalable compute

Platform and Tools to discover insights



Trusted AI

AI Governance

Challenges in democratizing Data Science

Data science is transforming business and generating new value across enterprise. Data science still often requires dedicated technology and sophisticated teams to execute successfully.



AI for business

Majority of the business stakeholders want to use AI but lack ML skills



Trusted Data

No single source of truth of the customer and business metrics



Knowledge/Training

Knowledge assets are built in silos with little reproducibility and are gradually lost



Scalable compute

No Standardized infrastructure, tools, capabilities, best practices



Trusted AI

No central governance to ensure models are not biased, accessible, valuable, and safe.

Data Science Center of Excellence (DSCoE) - Pillars



Governance

- Centralized Model catalog
- Standardized AI best practices
- Model Bias and fairness best practices
- Audit review
- Model discovery
- Training
- Privacy Complaint

Model Performance

- Model drift
- Model stability
- Model dashboard
- Explainability
- Model canvas

Scalable and Resilient

- Design to withstand failures
- High Availability
- Security and monitoring
- Authentication and authorization

Self Serve and MLOps

- Provide self serve capabilities
- Tools for model deployment and scoring

No Vendor Lock-in


- Provide multiple tools
- Reference Architectures
- Hybrid Cloud environment

Trusted Data

- Reliable source of data
- Data analytics and DQ capabilities


Training and best practices

- Collaboration and training sessions
- Best practices




Trusted AI

Centralized model policies, model risk management, discovery, catalog, model fairness and reliable data



AI/ML Platform

Standardized infrastructure, tools, capabilities, best practices, and training



Integrated Innovation & Training

Drives innovation and success through research, collaboration, and training

DSCoE - Trusted AI - Benefits



Trusted AI

Centralized model policies, model risk management, discovery, catalog, model fairness and reliable data

Objective: Centralized model policies, model risk management, discovery, catalog, model fairness and ethics

Benefits:

- Collaborated standardized standards for AI and ML fairness and bias.
- Standardized model canvases to explain AI effectiveness
- Define processes and elements to increase synergies
- Repeatable process and evaluation
- Model risk management (MRM)
- Best practices and adopt industry standards
- Created trusted data features

How it works:

- Collaborate and set standards to make AI fair and ethical
- Share best practices across all Data Science teams
- Create standard AI communication canvas
- Cross functional visible

AI/ML Tools - DSCoE - Benefits



Objective: Standardized infrastructure, tools, capabilities, best practices, and training

Benefits:

- Share methodologies and tools.
- Create sandbox environments to help kick start different data science initiatives
- Collaborate on create stand deployment and AI frameworks
- Data quality and model governance standards
- Best practices and adopt industry standards

How it works:

- Collaborate and create standard tool set to help kick start data science initiatives
- Create scalable, resilient environment to share
- Follow Privacy standards

Integrated Innovation & Training - DSCoE - Benefits



Objective: Drives innovation and success through research, collaboration, and training

Benefits:

- Collaboration and training sessions
- Shared innovation initiatives
- Helping hand when needed
- Continuous upskilling our data scientists

How it works:

Share research and development

Share best practices across all Data Science teams

Data Science democratization

Conclusion

Data Science center of excellence has distinct data science, data management, analytics, and governance teams that contribute unique skills and resources to help an organization become data-driven. These teams can enable fast discovery and more profound insights when designed around federated techniques, such as tiger teams, agile methodologies, and MLOps tools and techniques.

Thank you