

Business Data Analysis

MGM509

Lecture 10

Strategic Data Science: Engineering Solutions, Advanced Techniques, and Business Strategy

Prepared by
Eng. Ahmed Hussein, MBA, PfMP®,
PMP®, CMP APMG

Lecture
10

What we will cover

1. Other Data Science Tasks and Techniques
2. Analytical Engineering
3. Data Science and Business Strategy

Lecture
10

Other Data Science Tasks and Techniques



Lecture
10

Overview of Broader DS Tasks

- Not all data science tasks are classification, clustering or regression.
- Many problems focus on understanding structure, group behavior, or novelty.
- Other valuable methods:

Co-occurrence
grouping

Profiling

Link prediction

Casual
InferenceLatent
Information
Discovery

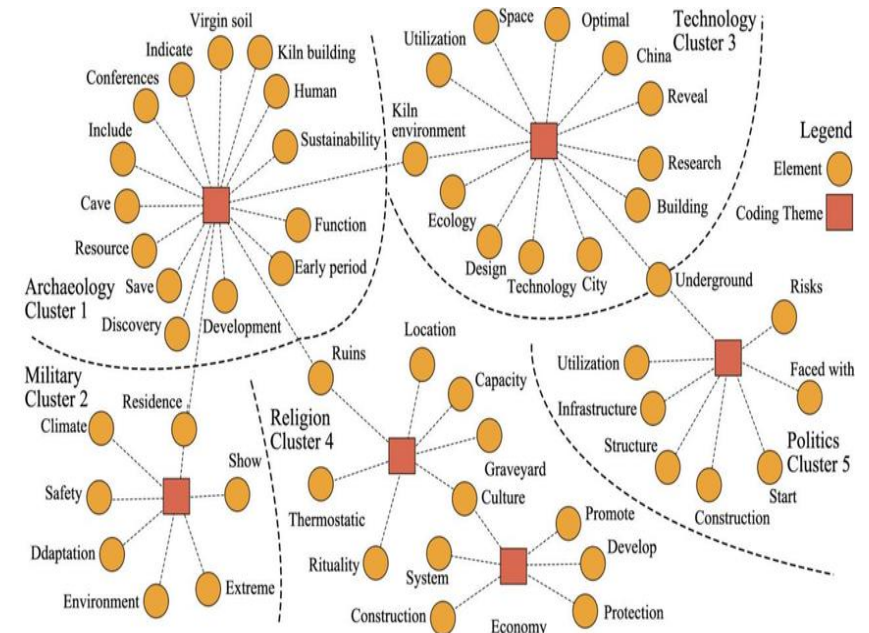
**Lecture
10****Co-Occurrence Grouping**

- Identifies items or events that tend to appear together in data, revealing meaningful associations or patterns.
- When two or more elements (e.g., words, products, events) frequently appear together in the same context (e.g., document, transaction, session).
- Classic examples:
 - Market basket analysis (e.g., “customers who bought X also bought Y”)
 - Text mining (e.g., identifying topic-related word pairs)
 - Recommendation systems
- Uses association rules (e.g., Apriori algorithm).
- Applied in retail, recommendations, and content bundling.

Lecture 10

Co-Occurrence Grouping

- Techniques Used:
 - Association Rule Mining: e.g., Apriori, FP-Growth
 - Pointwise Mutual Information (PMI): for text
 - Co-occurrence matrices and heatmaps
- Insights Provided:
 - Uncover latent relationships or groupings
 - Support feature engineering and similarity detection
 - Drive personalization and targeting strategies



Lecture
10

Profiling - Summarizing Behavior

- Profiling involves creating a summarized representation of an individual, group, or entity's behavior.
- Uses statistical and categorical features to describe users/customers (e.g., average purchase value, churn rate, visit frequency).
- Group entities with similar behavior (e.g., “frequent buyers,” “inactive users,” “deal seekers”).
- Personalization and Targeting: Profiles inform tailored recommendations, offers, or communications.

Lecture 10

Profiling - Summarizing Behavior

- Use Cases:
 - Customer relationship management (CRM)
 - Fraud detection
 - Risk profiling in finance
 - Patient behavior tracking in healthcare
- Techniques Involved:
 - Aggregation (sums, averages, frequencies)
 - Clustering (e.g., K-Means for customer segments)
 - Feature extraction (e.g., recency, frequency, monetary value – RFM)

Customer Profile



John Doe

Age: 45
Location: New York
Gender: Male

Recency
7 days since last purchase

Frequency
12 purchases in last year

Monetary Value
\$1200 total spent

Segment
Frequent Buyer

**Lecture
10****Link Prediction**

- Predicts whether a connection (link) between two entities in a network should exist or will exist in the future.
- Common in social networks, e-commerce, and citation analysis.
- Example: Suggesting new friends on Facebook.
- Common Techniques:

Heuristic-Based
Common Neighbors,
Jaccard Coefficient,
Preferential
Attachment

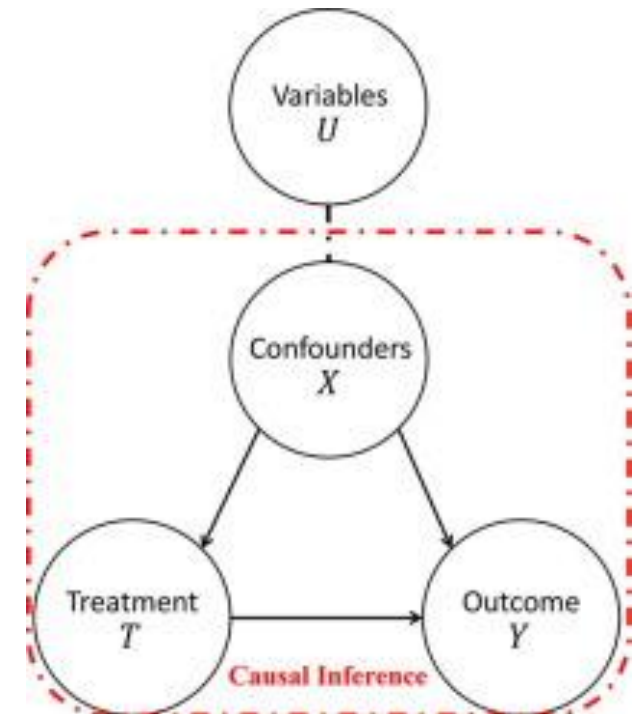
Embedding-Based
Node2Vec, DeepWalk,
Graph Neural Networks
(GNNs)

Matrix Factorization
Decompose adjacency
matrix to predict
unseen links

Lecture
10

Causal Inference

- Determines whether and how one variable causes a change in another, not just whether they are correlated.
- Moves beyond correlation to explain and predict interventions
- Helps make strategic business decisions (e.g., “Did the ad campaign cause increased sales?”)
- Essential in policy-making, medicine, economics, and A/B testing.
- Techniques:
 - Randomized Controlled Trials (RCTs):
 - Observational Methods: Propensity Score Matching. Difference-in-Differences. Instrumental Variables, Causal Graphs (e.g., DAGs)



**Lecture
10****Latent Information Discovery**

- Process of uncovering hidden patterns, structures, or relationships in data that are not immediately observable.
- Reveals underlying drivers of behavior (e.g., user intent, hidden topics)
- Supports personalization, clustering, and dimensionality reduction
- Useful in NLP, marketing, and content analytics.

Lecture 10

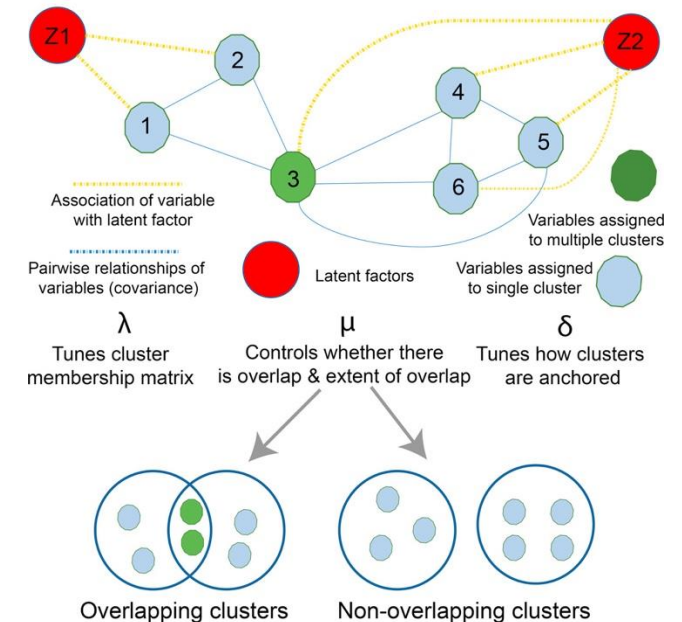
Latent Information Discovery Techniques

Topic Modeling (e.g., LDA): Discover abstract topics in a collection of documents

Matrix Factorization (e.g., SVD, NMF): Used in recommender systems to uncover user preferences and product features

Embedding Models: Capture latent semantic meanings in text (e.g., Word2Vec, BERT)

Autoencoders / Deep Learning: Reduce dimensions and extract features from images, text, or tabular data



Lecture
10

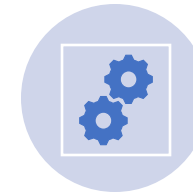
Choosing the Right Task

**Business Goal:**

What are you trying to improve or predict?



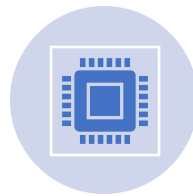
Data Type: Is the data labeled?
Structured?
Textual?



Desired Output:
Category, number, group, pattern?

**Interpretability**

Needs: Will decisions be audited or explained?



Deployment Context: Real-time vs. batch processing?

Lecture
10

Analytical Engineering



**Lecture
10****From Ad Hoc to Engineering**

Many data science projects are ad hoc, not reusable or scalable.



Analytical engineering aims to make data science solutions systematic and repeatable.



Think in terms of pipelines, standards, and operations.



Move from one-off insights to embedded decision systems.

**Lecture
10**

What is Analytical Engineering

- **Definition:** Analytical Engineering is the design, implementation, and integration of data-driven solutions into operational systems that support real-world decision-making.
- **Key Characteristics:**

**System-Oriented Thinking**

Focuses on building end-to-end analytical pipelines that connect data inputs to decision outputs.

**Combines Logic + Prediction**

Merges business rules with predictive models to automate or enhance decisions.

**Operational Integration**

Ensures that models and logic are embedded into business workflows, not left on the shelf.

**Built for Scale and Sustainability**

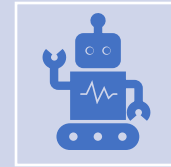
Prioritizes reliability, maintainability, and performance in real-world conditions (e.g., APIs, dashboards, alerts).

**Lecture
10**

Elements of Analytical Engineering



Input data pipelines:
Clean, structured, timely
data.



Decision logic
integration: Rule-based +
predictive.



Feedback loops: Capture
outcomes and retrain.



Evaluation metrics: Tie
model performance to
business KPIs.

Lecture
10

Principles of Decision-Centric View of Modeling



**Start with the
Decision**

- Clearly define the decision the model will inform (e.g., approve loan, target customer with offer).

**Align Metrics with
Impact**

- Evaluate models based on how they improve decision outcomes, not only accuracy.

**Integrate with
Business
Processes**

- Ensure model outputs are actionable and embedded into workflows.

**Balance
Complexity vs.
Usability**

- A slightly less accurate but easier-to-use model may be more valuable.

**Consider Costs
and Payoffs**

- Account for the economic or strategic consequences of correct vs. incorrect predictions.

Lecture 10

Modeling for Business Constraints

- Designing predictive models that **align with operational, financial, legal, and strategic limitations** of the business.
- Ensures that models are **realistic, implementable, and cost-effective**

- Optimize for maximum ROI, balancing model accuracy with cost of action.



- Model predictions must fit within available resources (e.g., only 500 customer service calls possible per week).

- Ensure that models align with the organization's long-term vision, market positioning, and growth priorities,

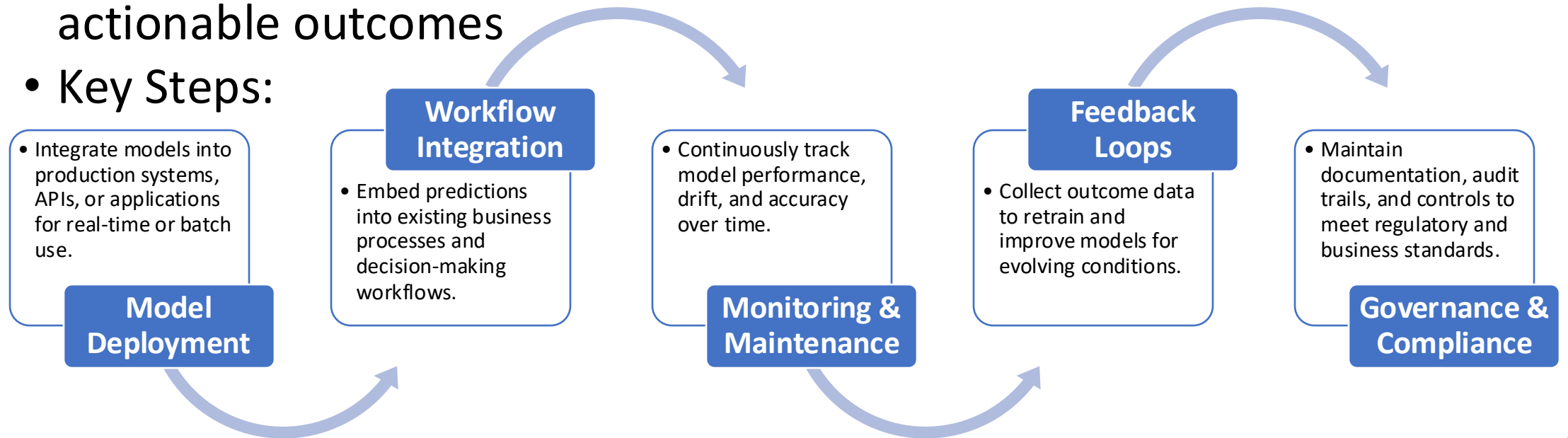
- Comply with privacy laws, fairness, and non-discrimination guidelines.

Lecture 10

Operationalizing Data Science

- The process of deploying, integrating, and managing data science models in real-world business environments.
- Focuses on turning analytical insights into repeatable, scalable, and actionable outcomes

- Key Steps:



Lecture
10

Managing Uncertainty in Practice

- All predictions and decisions are made with imperfect information.
- Key Approaches:

Quantify Uncertainty

- Use **confidence intervals**, probability distributions, or scenario ranges to represent prediction reliability.

Sensitivity Analysis

- Test how changes in input variables affect outcomes to identify high-impact factors.

Scenario Planning

- Develop “best-case,” “worst-case,” and “most likely” projections for decision-making.

Threshold Adjustment

- Adapt decision thresholds based on business risk tolerance (e.g., fraud detection sensitivity).

Continuous Monitoring

- Reassess uncertainty as new data arrives or market conditions shift.

Lecture
10

Managing Uncertainty in Practice

- Example:

In loan approval modeling, a predicted default probability of $8\% \pm 3\%$ may require different action than a precise 8%—uncertainty affects risk strategy

Lecture 10

Feedback Loops & Continuous Learning

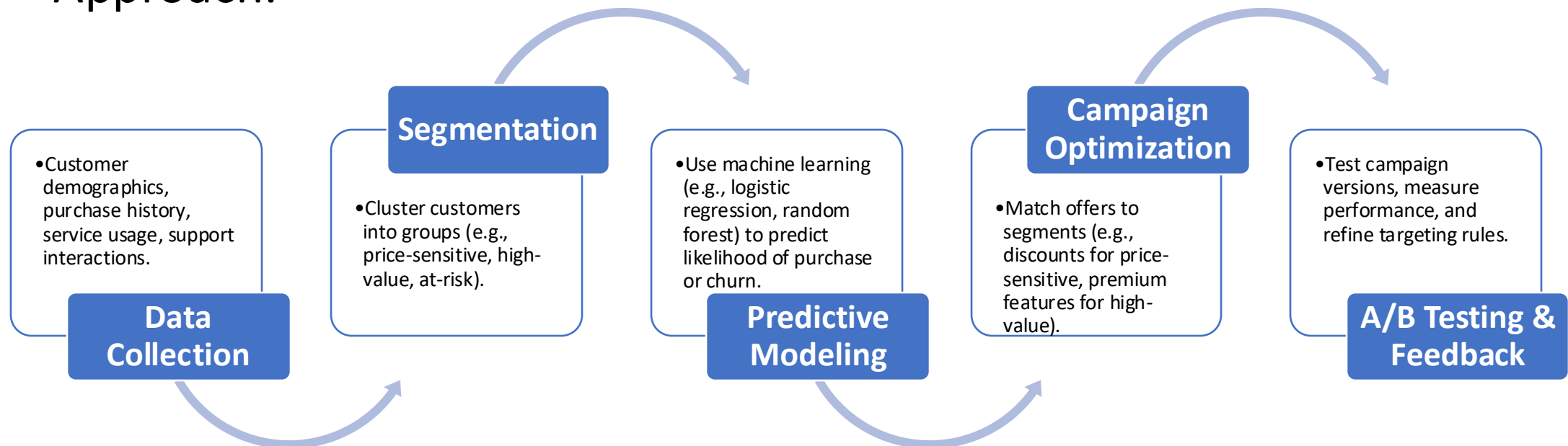
- Data science models degrade over time due to changes in data, behavior, or the market (data drift).
- Feedback loops enable models to adapt and improve continuously.
- Key Components:

Performance Monitoring	Error Analysis	Data Refresh	Model Retraining	Automated Pipelines
• Track KPIs (e.g., accuracy, precision, ROI) in production.	• Identify misclassifications or poor predictions to guide retraining.	• Incorporate new, relevant data into model updates.	• Schedule regular or trigger-based retraining cycles.	• Use MLOps tools to automate feedback collection and model updates.

Lecture 10

Case Example – Targeted Marketing System

- A telecom company wants to increase customer retention and boost upsell conversions.
- Approach:



Lecture
10

Data Science and Business Strategy



Lecture
10

Why Strategy Matters?

- Data science doesn't operate in a vacuum, Models are only as useful as the business impact they generate.
- Predictive analytics should directly support goals like revenue growth, cost reduction, or customer retention
- Focus on value creation, not just predictive accuracy: A less accurate model that drives higher profit or better decisions may be more valuable than a perfectly accurate one with no business impact.
- Strategy provides direction and decision-making context, and ensures that analytics efforts are prioritized where they deliver the most ROI.

**Lecture
10**

Strategic Alignment Model

- Ensures that data science initiatives directly contribute to the organization's strategic goals.
- Aligns technical execution with business priorities for maximum value.



Lecture 10

Embedding Data Science into Business

- Integrate analytics into daily operations.
- Embedding ensures **faster decision-making, greater adoption, and continuous improvement**
- Key Practices:

Cross-Functional Collaboration

- Involve domain experts, analysts, engineers, and decision-makers from the start.

Workflow Integration

- Embed model outputs into existing systems (e.g., CRM, ERP, dashboards) so they directly support business actions.

Business Ownership

- Ensure stakeholders understand and trust the models enough to act on their recommendations.

Change Management

- Provide training and support to encourage adoption and address resistance.

Performance Measurement

- Continuously track how embedded models impact business KPIs.

Lecture
10

Framing Strategic Questions

- The quality of insights depends on the quality of the questions asked.
- Well-framed strategic questions ensure analytics efforts target high-value business problems.
- Key Guidelines
 1. Align with Business Goals
 - Start from the organization's strategic objectives.
 2. Be Specific & Action-Oriented
 - Ask questions that lead to clear decisions (e.g., "Which customers are at risk of churn next quarter?").

Lecture
10

Framing Strategic Questions

3. Incorporate Measurable Outcomes

- Define how success will be evaluated (e.g., “Increase upsell conversion rate by 10%”).

4. Prioritize by Impact & Feasibility

- Focus on areas where analytics can deliver significant value quickly.

5. Challenge Assumptions

- Ensure questions are not biased by existing beliefs or outdated data.

Lecture
10

Framing Strategic Questions

Examples

Retail:

- “How can we reduce inventory costs without affecting availability?”

Telecom:

- “Which customers should we proactively engage to prevent churn?”

Finance:

- “What is the likelihood of default for new loan applicants?”

Lecture
10

Defining Data-Driven Objectives

- Clear, measurable objectives ensure analytics projects stay focused and deliver tangible business value.

- Principles:

Align with
Strategy

Be SMART

Link to
Decisions &
Actions

Quantify
Expected
Impact

Ensure Data
Feasibility.

- Examples:

- Increase campaign ROI by 15%.
- Reduce fraud by 20% using anomaly detection.
- Link model outputs to business impact.

**Lecture
10****Operationalizing Analytics**

- Operationalizing Analytics: Moving analytics from the lab (data science team) into day-to-day business operations.
- Ensuring insights are delivered at the right time, to the right people, in the right format.
- Steps:

Integrate with
Business
Processes



Automate
Where
Possible



Ensure User
Adoption



Monitor and
Refine



Build
Feedback
Loops

Lecture
10

Operationalizing Analytics

• Examples:

Retail:

- Dynamic pricing recommendations integrated into POS systems.

Finance:

- Automated fraud alerts for transaction approval systems.

Telecom:

- Churn risk scores embedded into customer service tools.

**Lecture
10****Creating Competitive Advantage**

- How: Leveraging analytics to differentiate from competitors in a sustainable way.
- Turning data into a strategic asset that others can't easily replicate.
- Key Strengths:

**Unique Data
Assets****Faster
Decision-
Making****Personalized
Experiences****Operational
Excellence****Innovation
Enablement**

Lecture
10

Examples

Amazon

- Predictive product recommendations that drive repeat purchases.

Netflix

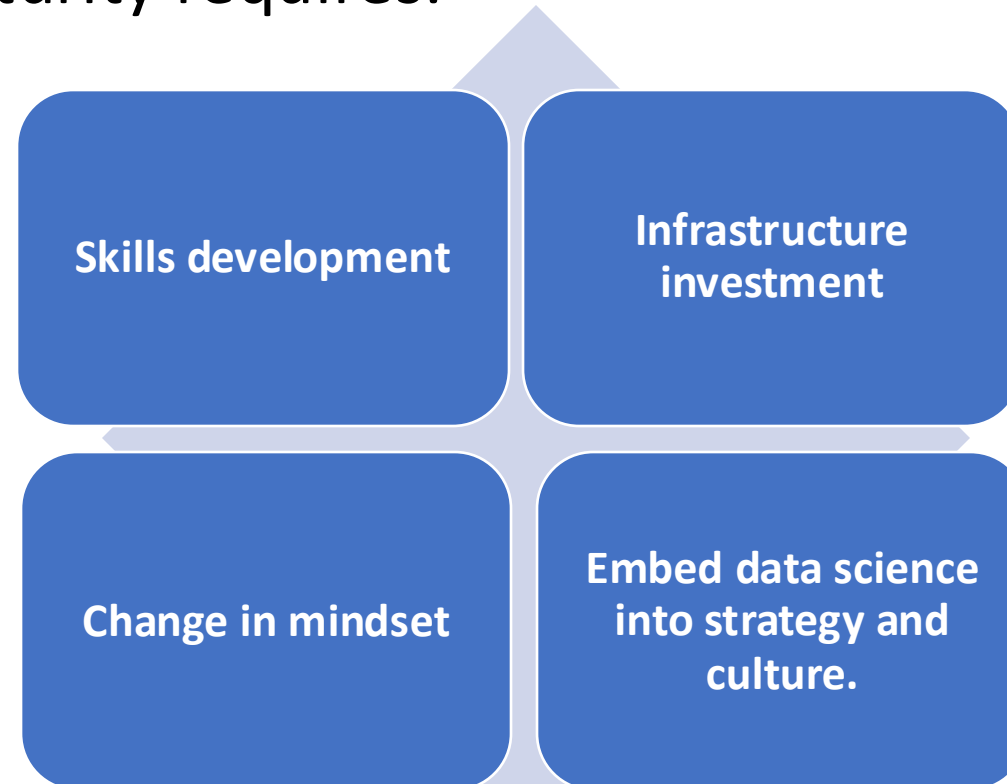
- Personalized viewing suggestions to keep users engaged.

UPS

- Route optimization saving millions in fuel costs.

**Lecture
10****Building Organizational Capability**

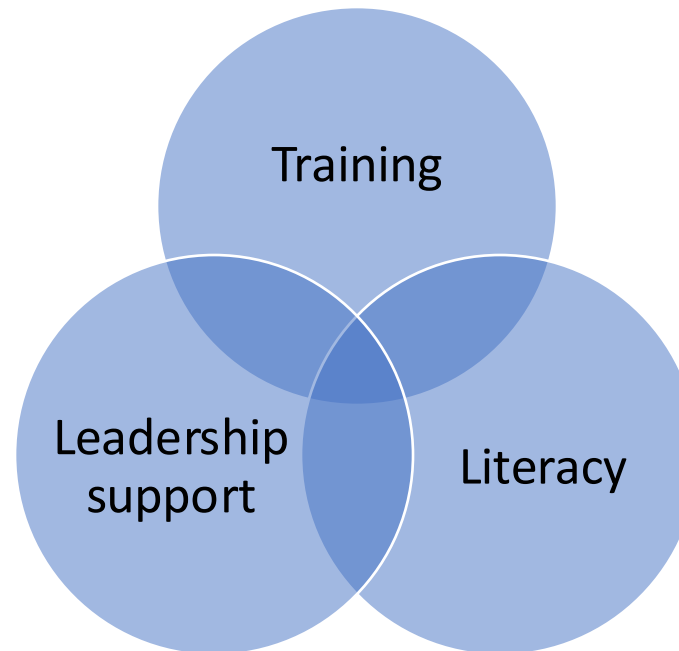
- Data science maturity requires:



Lecture
10

Building an Analytics Culture

- Promote data-driven decision-making at all levels.
- Celebrate data-informed wins across the org.
- Invest in:



**Lecture
10****Key Points**

Data science covers diverse tasks beyond basic modeling, such as profiling, link prediction, and causal inference.

Analytical engineering integrates scalable, reliable models into operations.

Decision-centric models deliver actionable outputs, not just predictions.

Align analytics with business strategy to drive competitive advantage.

Use feedback loops and uncertainty management for sustainable value.

For Discussion Panel and Next Lecture:

- Chapters 11, 12, and 13 in Text Book "Business Data Analysis"

**Lecture
10**



**For Discussion Panel
and Next Lecture**



Thank You