

Business Data Analysis

MGM509

Lecture 10

Advanced Techniques, Analytical Engineering,
and Growing Trends in Data Science

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. Trends in Business Data Analytics

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

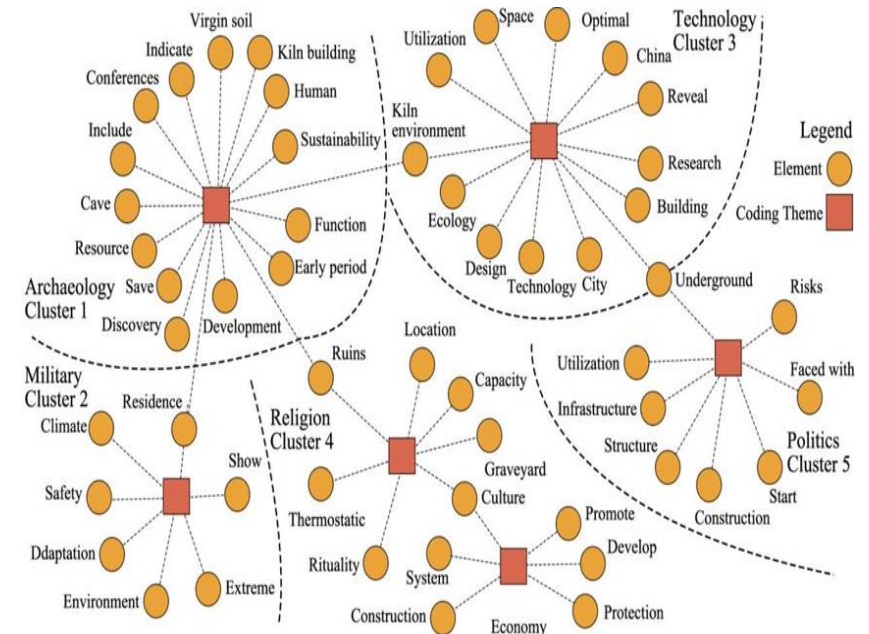
**Lecture
10****Co-Occurrence Grouping**

- 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

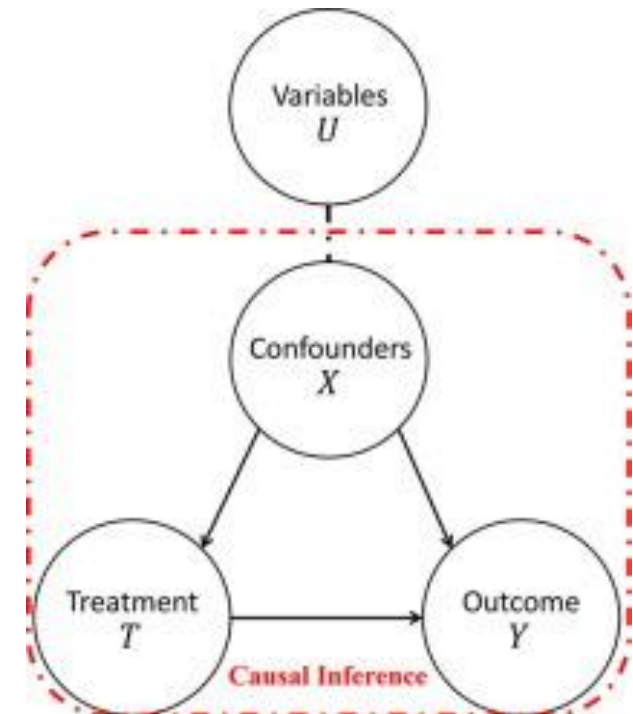
Embedding-Based
Node2Vec, DeepWalk,
Graph Neural Networks
(GNNs)

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

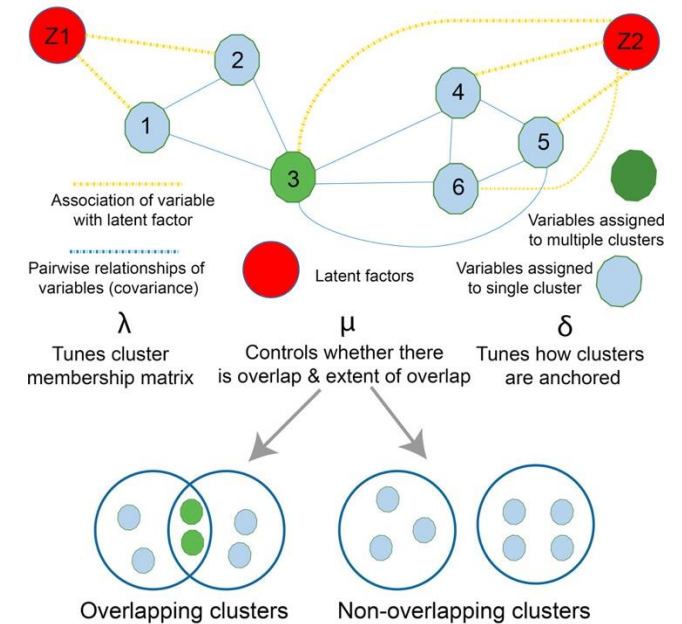
Latent Information Discovery Techniques

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

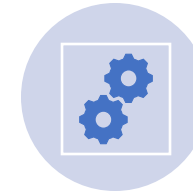
Choosing the Right Task

**Business Goal:**

What are you trying to improve or predict?



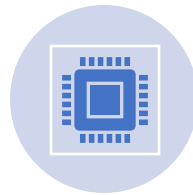
Data Type: Is the data labeled?
Structured?
Textual?



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**

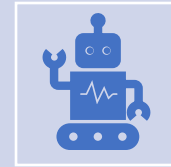
Prioritizes reliability, maintainability, and performance in real-world conditions (e.g., APIs, dashboards, alerts).

**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.

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



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

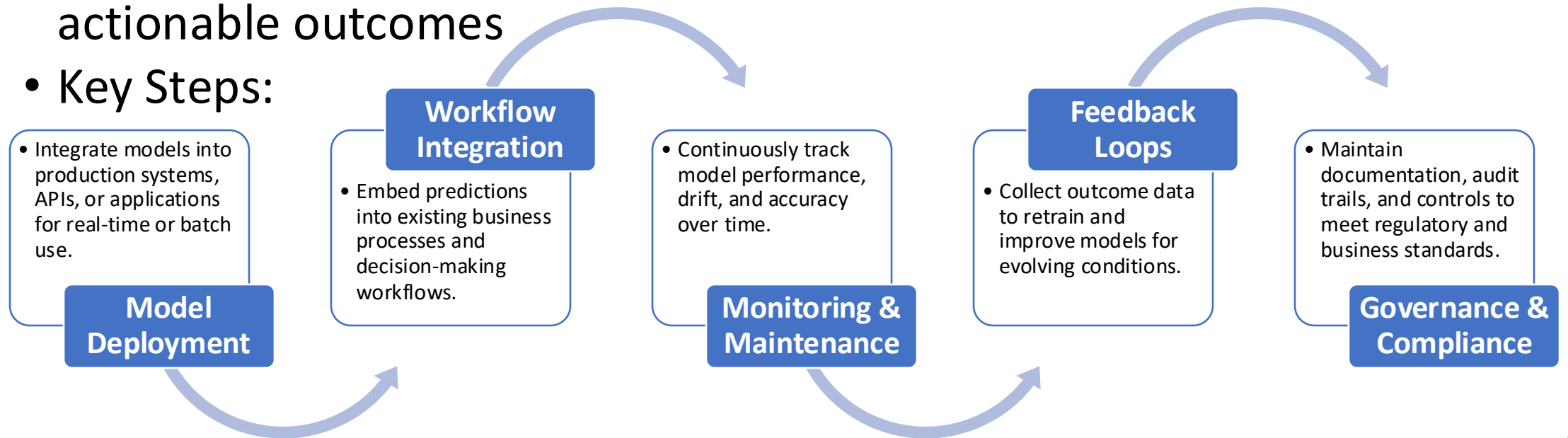
- Comply with privacy laws, fairness, and non-discrimination guidelines.

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

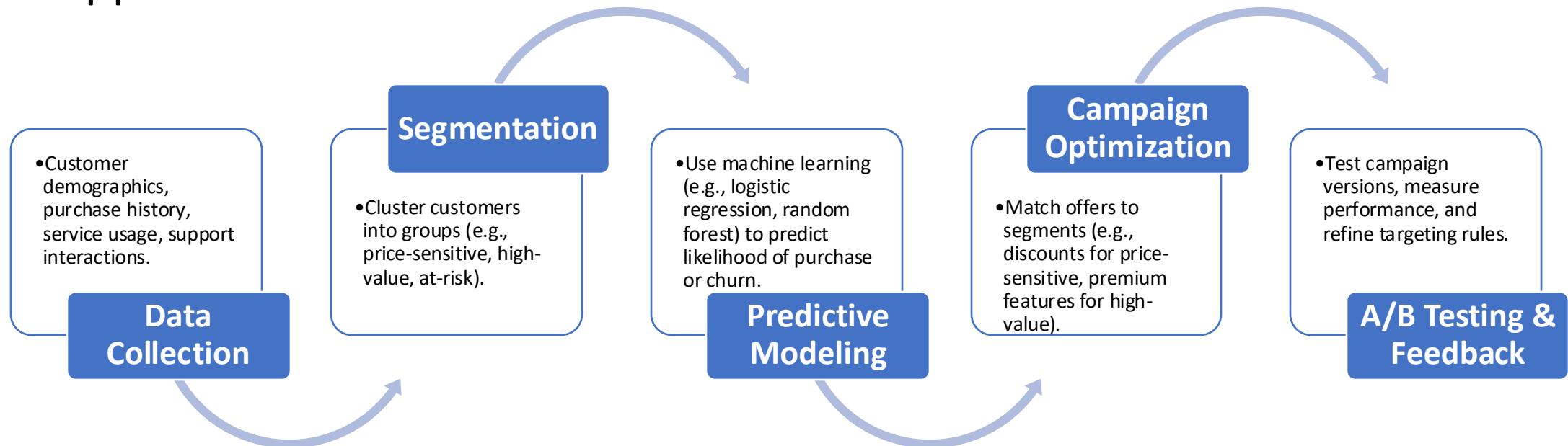
- Data science models degrade over time due to changes in data, behavior, or the market (data drift).
- Feedback loops enable models to adapt and improve continuously.
- Key Components:

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

Trends in Business Data Analytics



Lecture
10

Why Business Data Analysis Is Transforming Organizations?

- Data is now a core strategic asset
- Decisions increasingly depend on insights, not intuition
- Explosion of enterprise data from apps, IoT, and digital systems
- Analysts play a critical role in converting raw data → business outcomes

Lecture
10

1. AI-Driven Analytics & Automation

- Rapid adoption of machine learning for forecasting and pattern detection
- Automated insight generation through AI copilots and agents
- Natural language querying (“ask the data”) democratizes data access
- Analysts evolve into AI-assisted decision partners

Lecture
10

2. Self-Service Analytics for Business Teams

- Growing use of tools like **Power BI, Tableau, and Qlik**
- Non-technical users explore data independently
- Analysts shift from report builders → data enablers and governance stewards
- Increased focus on data literacy across departments

Lecture
10

3. Real-Time & Streaming Data

- Increased demand for real-time dashboards and alerts
- Use cases include customer behavior monitoring, fraud detection, operations tracking
- Organizations move from monthly reports → instant insights
- Technologies: **event streaming platforms, APIs, IoT data hubs**

Lecture
10

4. Data Governance, Quality, & Ethics

- Stronger focus on data ownership, privacy, and lineage
- Rise of enterprise data catalogs and governance frameworks
- Analysts contribute to ethical data use and model transparency
- High-quality data becomes a competitive differentiator

Lecture
10

5. Predictive & Prescriptive Analytics

- Shift from “What happened?” → “What will happen?” → “What should we do?”
- Adoption of predictive models to forecast risks, demand, churn, and performance
- Prescriptive analytics recommends decisions based on optimization models
- Expanding need for analysts who understand statistics + business context

Lecture
10

6. Cloud Analytics & Data Platforms

- Migration from on-premise systems to cloud platforms (**Azure, AWS, Google Cloud**)
- Use of cloud data warehouses/lakes (**Snowflake, BigQuery, Databricks**)
- Scalable computing enables deeper analysis on massive datasets
- Analysts work more closely with data engineers and architects

Lecture
107. Data Storytelling & Visual
Communication

- Insight is useless if not understood

- Growing emphasis on:

Narrative-driven
dashboards

Executive-ready insight
summaries



Data visualization best
practices

- Analysts become storytellers who influence business decisions

Lecture
10

8. Integration of BI with Enterprise Systems

- Connected analytics across **ERP, CRM, HRIS, and project systems**
- Unified data layers for end-to-end visibility
- Business analysts leverage cross-system insights to identify:

Revenue
leakages

Operational
inefficiencies

Customer
behavior
patterns

- Rise of “enterprise data fabric” architectures

Lecture
10

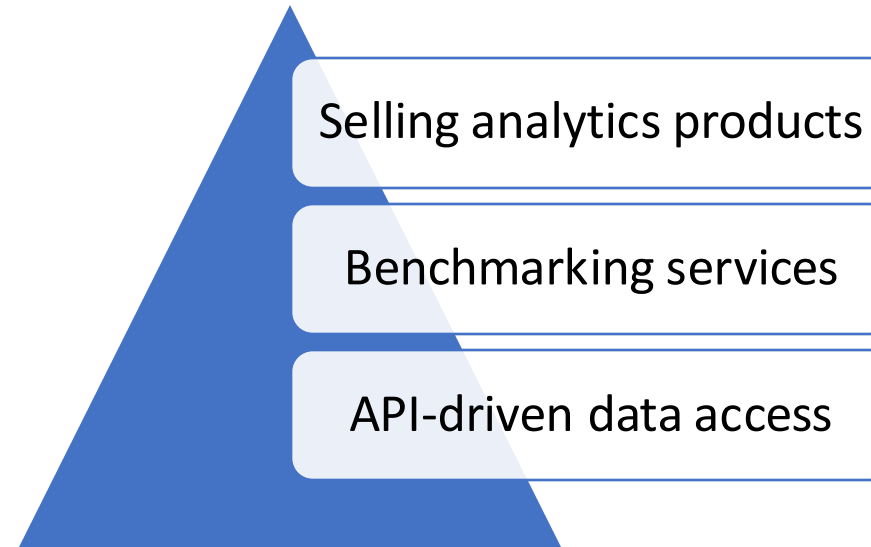
9. Low-Code / No-Code Data Applications

- Tools like **PowerApps, Alteryx, Knime, AppSheet, and Orange Data Mining** enable analysts to build solutions
- Faster prototyping of dashboards and workflow automations
- Reduces dependency on engineering teams
- Analysts act as citizen developers augmenting business processes

Lecture
10

10. Data Monetization & Value Realization

- Organizations turning data into revenue streams
- Models include:



- Analysts quantify business value of data through KPIs and ROI models
- Increased focus on data-as-a-product (DaaP)

Lecture
10

11. Scenario Planning & Simulation Modeling

- Use of “What-if” models for forecasting financial and operational outcomes
- Decision modelling to evaluate alternative strategies
- Popular in risk management, supply chain, and resource planning
- Analysts adopt tools like Power BI parameters, Monte Carlo simulations, and simulation engines

Lecture
10

12. Rise of the Analytics Center of Excellence (CoE)

- Organizations centralize analytics talent, processes, and standards
- CoEs deliver:
 - Standard dashboards
 - Data governance frameworks
 - Technology roadmaps
- Analysts benefit from structured training, templates, and reusable data assets

Lecture
10

13. Human + AI Collaboration Becomes the Norm

- Analysts use AI as a partner for:
 - Data preparation
 - Forecasting
 - Insight generation
 - Visualization
- Human expertise ensures context, ethics, and business relevance
- Future analysts are **AI-augmented decision makers**

Lecture
10

The Future Business Data Analyst

- As Business Data Analysis continues to evolve into a strategic, insight-driven discipline essential for competitive advantage, The Key competencies shaping the next-generation analyst are:
 - AI and machine learning literacy
 - Advanced analytics + business acumen
 - Mastery of visualization and insight storytelling
 - Ethical and responsible data use
 - Ability to translate data into measurable business value

**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.

Data Science is growing and integrating into every business area.

As Human and AI collaboration is growing, Data Analysts need to develop in technicalities, visualization and storytelling, and understand AI governance and ethics.

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