Name of the Programme: MSc Integrated
Course Code: IMC- 402
Title of the Course: Data Modeling and Visualization
Number of Credits: 6(4L-0T-2P)
Effective from AY: 2021-22

Prerequisites for the course:	A basic understanding of data management concepts and knowledge of relationship database tables	
Objective:	1. Learn to understand practical techniques to analyze and model data as part of the overall data management lifecycle 2. To expose students to visual representation methods and techniques that increases the understanding of complex data. 3. Learn to design good design practices for visualization, tools for visualization of data from a variety of fields and visualization software like Processing, GapMinder and Tableau.	
Content Theory:	**Data modeling fundamentals:** The purpose and role of data modelling- basic data modeling concepts and terminology, data modeling building blocks- hierarchies for the entities, data model Constraints for your attributes: specify cross-entity dependencies through strong and weak entities -summary of real-world entity and attributes complexities. real-world complexities to relationships why relationship cardinality and complexities matter - build real-world complexities into data model relationships-define the maximum cardinality of a relationship -define the minimum cardinality of relationship -use crow's foot notation for minimum and maximum cardinality -summary of cardinality and complex relationships. move across the different levels of data model: Harmonize different levels of data model - brief look a relational database normalization -forward-engineering your conceptual data model - more data model forward engineering - reverse engineer a physical model back into conceptual model - summary - how to work with different levels of data model	19 hours
	Software for data modeling: The importance of data modeling software -build a data model with a drawing program - build model with data modeling software tool	9 hours
	Visualization: Right graph for right data, Components of a Data Visualization-Different Types of Graphs, Deadly Sins of Graph Design, How to Avoid Being Mislead with Graphs Session. The Value of Visualization Sessions - Effective Use of Form and Space. Fundamentals of Graphs - Integrity in Visualization-Visual Perception and Quantitative Communication Reading - Effective Use of Form and Space	13 hours
	Detailed Design of Tables and Graphs Readings: Summary at a Glance: Table Design Summary at a Glance: Graph Design Session. Additional Constructs and Multivariate Analysis- Escaping 2 Dimensions: Animated Scatter-Plots-Introduction to Information Design.	7 hours

Content Practical:	• **Data Modelling part - lab hrs – 24 hrs** Suggested Data Modelling and visualization lab assignments These assignments focus on different aspects of data modeling, allowing students to understand and practice conceptual, logical,	

	physical, dimensional, and NoSQL data modelling techniques. They provide hands-on experience in translating real-world scenarios into structured data models. **Assignment 1 - Conceptual Data Modeling:** • Task: Choose a real-world scenario (e.g., online marketplace, banking system) and create a conceptual data model. • Requirements: Identify the main entities, attributes, and relationships in the scenario. Use an appropriate notation (e.g., Entity-Relationship Diagram) to represent the conceptual model. • Deliverables: Conceptual data model diagram, along with a description of the entities, attributes, and relationships. **Assignment -2 -Logical Data Modeling:** • Task: Take the conceptual data model created in the previous assignment and transform it into a logical data model. • Requirements: Specify the tables, columns, primary keys, foreign keys, and relationships based on the conceptual model. Normalize the logical data model to eliminate redundancy. • Deliverables: Logical data model diagram, including table structures, primary and foreign keys, and a brief explanation of the normalization process. **Assignment -3 -Physical Data Modeling:** • Task: Convert the logical data model into a physical data model suitable for implementation in a specific database management system. • Requirements: Choose a database management system (e.g., MySQL, PostgreSQL) and map the logical data model elements to the corresponding database objects (e.g., tables, columns, data types, constraints). • Deliverables: Physical data model diagram, including the database objects, data types, and constraints. **Assignment -4 -Dimensional Modeling for Data Warehousing:** • Task: Design a dimensional model for a data warehousing scenario. • Requirements: Identify the fact tables, dimension tables, and their attributes.Establish relationships and define hierarchies between dimensions. Consider the design principles of star schema or snowflake schema. • Deliverables: Dimensional model diagram (e.g., star schema or snowflake schema), including fact tables, dimension tables, and their attributes. **or** **Assignment – 5 -NoSQL Data Modeling:** • Task: Choose a NoSQL database (e.g., MongoDB, Cassandra) and design a data model for a specific use case. • Requirements: Identify the entities, attributes, and relationships in the use case. Determine the document structure, collections, and indexing strategies based on the NoSQL database's features and query requirements. • Deliverables: Data model representation (e.g., JSON-like documents, key-value pairs) and a brief explanation of the design choices made.	5 hours 5 hours 5 hours 5 hours 4 hours

	• **Visualization part - lab hrs -24 hrs** These assignments focus on different aspects of data visualization, allowing students to practice creating various types of visualizations and effectively communicating insights. They provide hands-on experience in data exploration, interactive dashboard design, geospatial analysis, network visualization, and storytelling with data. **Assignment -1 Exploratory Data Visualization:** • Task: Choose a dataset of your choice and create a set of visualizations to explore and understand the data. • Requirements: Use a visualization library (e.g., Matplotlib, Seaborn, Plotly) to create a variety of charts and plots, such as scatter plots, line charts, bar charts, and heatmaps. Highlight important patterns, trends, and relationships in the data. • Deliverables: Jupyter notebook or script with code, along with a report explaining the insights gained from the visualizations. **Assignment -2 Interactive Dashboard Design:** • Task: Design an interactive dashboard for a specific business scenario or data analysis task. • Requirements: Use a dashboarding tool like Tableau, Power BI, or Plotly Dash to create a visually appealing and interactive dashboard. Include multiple visualizations, filters, and interactive elements to allow users to explore and analyze the data. • Deliverables: Interactive dashboard, documentation explaining the design choices, and a user guide. **Assignment -3 -Geospatial Data Visualization:** • Task: Visualize geospatial data on maps to uncover insights and patterns. • Requirements: Utilize libraries like Folium, Plotly, or D3.js to create maps and plot geospatial data. Represent data using markers, choropleth maps, heatmaps, or other appropriate visualizations. Explore relationships between the data and geographic locations. • Deliverables: Interactive map visualizations, code snippets, and a report summarizing the findings. **Assignement -4 Network Visualization:** • Task: Visualize relationships and networks within a dataset. • Requirements: Use network visualization libraries like NetworkX, Gephi, or Cytoscape.js to create visual representations of nodes and edges. Explore connectivity, centrality, and other network metrics to analyze and understand the underlying structure. • Deliverables: Network visualization diagrams, code snippets, and a report explaining the insights gained from the visualizations. **Assignment – 5 Storytelling with Data:** • Task: Create a data-driven story using visualizations to convey a narrative or message. • Requirements: Choose a topic or dataset and design a series of visualizations that support your story. Use appropriate charts, images, and annotations to guide the audience through the narrative. Ensure a logical flow and effectively communicate the intended message.	5 hours 5 hours 5 hours 5 hours 4 hours

	• Deliverables: A presentation or report with a coherent narrative, visualizations, and accompanying explanations.	
Pedagogy:	lab assignments/ theory assignments /mini case study/capstone project	
References/ Readings	1. Hoberman, Steve. Data modeling made simple: a practical guide for business and IT professionals. Technics Publications, 2015. 2. Edward Tufte, The Visual Display of Quantitative Information 3. Tufte, Edward R., Nora Hillman Goeler, and Richard Benson. Envisioning information. Vol. 2. Cheshire, CT: Graphics press, 1990. 4. Fry, Ben. Visualizing data: Exploring and explaining data with the processing environment. " O'Reilly Media, Inc.", 2008. **Data Modeling:** 1. "Database Systems: The Complete Book" by Hector Garcia-Molina, Jeffrey D. Ullman, and Jennifer Widom 2. "The Data Model Resource Book: A Library of Universal Data Models for All Enterprises" by Len Silverston 3. "Data Modeling Essentials" by Graeme Simsion and Graham Witt **Data Visualization:** 1. "The Visual Display of Quantitative Information" by Edward R. Tufte 2. "Data Visualization: A Practical Introduction" by Kieran Healy 3. "Storytelling with Data: A Data Visualization Guide for Business Professionals" by Cole Nussbaumer Knaflic 4. "Information Visualization: Perception for Design" by Colin Ware	
Course Outcomes	1. Understand data modeling principles and create effective data models. 2. Design databases based on data models and optimize database structures. 3. Use data visualization tools and software to create informative visualizations. 4. Communicate insights and findings through visually appealing data visualizations.	