

Course Name	Foundations of Data Analytics
Course Number	BO CDA 105/ ACS 105 / MCS 105
Course Credit	3-0-3-9

Data sources and collection: Review of existing structured and unstructured data sources, Data collection techniques using sensors, surveys, and different instruments with examples of data collection and storing for different domains such as IoT, Audio and Video, Web and Social Networks etc.

Data Pre-processing: Importance of data correction, Basic features for data analysis, Descriptive data summarization, data cleaning, normalization, data integration and transformation, data reduction.

Data representation: Importance of data representations, Extracting salient features from data, Examples include MFCC from audio signals, histogram representation for text, feature representations for images, encoded representations, Spatial data representation: cartography, GIS paper maps to ArcGIS ArcMap symbolizing, Time-series data representations and curve fitting.

Data visualization: Basic charting, examples with real world weather data, extract and manipulate the data to display the maximum information, various types of graphs like pie chart, bar graphs, 3-D plots using R/Python.

Practical component: Lab to be conducted on a 3-hour slot weekly. It will be conducted with the theory course so the topics for problems given in the lab are already initiated in the theory class. The topics taught in the theory course should be synchronized with the laboratory classes by using R/Python.

Text Books:

1. Tufte, Edward R. The visual display of quantitative information. Vol. 2. Cheshire, CT: Graphics press, 2001.
2. Yau, Nathan. Visualize this: the Flowing Data guide to design, visualization, and statistics. John Wiley & Sons, 2011.