Objectives

- To become familiar with software tools available for data management.
- To explore the software tools, frameworks for real world applications by designing the modules.
- To become familiar with the functionalities provided by the data analysis tools, framework.
- To make a comparison of data management tools based on their applicability and limitations.

In this fully practical subject, the major focus is on the usage of open source tools, libraries, software for pre-process, analyse, visualize the data.

Following are suggested tools/libraries.

- Python Libraries (StatsModels, Pandas, Scikit-learn, Keras, PyTorch, TensorFlow, Matplotlib, Seaborn, Altair, Plotly)
- Visualization using Tableau, PowerBI
- Hadoop, Pg, Hive, MomgoDB
- Apache Spark and Stor

In each of above framework/tool/library, minimum 3 applications are to be developed. The applications are developed in such a way as to cover maximum functionalities provided by particular framework/tool/library.

Outcomes

After completion of subject, students would be able to:

- design and develop applications which analyse the data and visualize the outcomes from it.
- compare the functionalities provided by different software and choose the appropriate as per requirement.