Indian Institute of Information Technology, Nagpur

Academic Year 2023-2024 (ODD Semester)

Course Plan

Course Code: CSL 109
Course Name: Introduction to Data & Analytics
Credits: 4
Pre-requisites: DSA
Overlap: None
Course Coordinator: Dr. Anil Kumar Kushwah (akushwah@iiitn.ac.in)
Time Table Slot: Tuesday (09:00 - 10:00 am), Wednesday (11:00 am - 12:00 pm), and Thursday (12:00 - 01:00 pm)

Course Objectives:

1. Identify and describe the methods and techniques commonly used in data science.
2. Demonstrate proficiency with the methods and techniques for obtaining, organizing, exploring, and analyzing data.
3. Recognize how data analysis, inferential statistics, modeling, machine learning, and statistical computing can be utilized in an integrated capacity.
4. Create and modify customizable tools for data analysis and visualization per the evaluation of characteristics of the data and the nature of the analysis.
5. Demonstrate the ability to clean and prepare data for analysis and assemble data from a variety of sources.

Text Books:

1. An introduction to Data Science by Jeffrey Stanton
2. The Elements of Data Analytic Style by Jeff Leek
3. Exploratory Data Analysis with R, by Roger Peng
4. OpenIntro Statistics, by Diez, Barr, and Centinkaya-Rundel
5. R Programming for Data Science, by Roger Peng

Reference Books:

1. UC Irvine Machine Learning Repository https://archive.ics.uci.edu/ml/index.php
2. Variety of consumer datasets https://www.kaggle.com/datasets
3. World Bank https://data.worldbank.org/data-catalog/
4. US Government Data https://www.data.gov/

Lecture Plan

Module No.	Lecture No.	Topic Name	Planned Date
I	1.	Introduction to Data	22 Aug 23
I	2.	Introduction to Data Science	23 Aug 23
I	3.	Evolution of Data Science	24 Aug 23
I	4.	Data Science Roles	29 Aug 23
I	5.	Stages in a Data Science Project	30 Aug 23
I	6.	Applications of Data Science in various fields	31 Aug 23
I	7.	Data Security Issues	05 Sep 23
II	8.	Data Collection Strategies	12 Sep 23
II	9.	Data Pre-Processing Overview	13 Sep 23
II	10.	Data Cleaning	14 Sep 23
II	11.	Data Integration and Transformation	19 Sep 23
II	12.	Data Reduction, Data Discretization	20 Sep 23
		Sessional I Examination (23rd to 26th Sep. 2023) (Syllabus up to II Module)	

III	13.	Descriptive Statistics	27 Sep 23
III	14.	Mean, Standard Deviation	28 Sep 23
III	15.	Skewness, and Kurtosis	03 Oct 23
III	16.	Box Plots	04 Oct 23
III	17.	Pivot Table	05 Oct 23
III	18.	Heat Map	10 Oct 23
III	19.	Correlation Statistics	11 Oct 23
III	20.	ANOVA	12 Oct 23
III	21.	One Way ANOVA	17 Oct 23
III	22.	Two Way ANOVA	18 Oct 23
IV	23.	Data Definitions and Analysis Techniques	19 Oct 23
IV	24.	Data Elements and Variables	25 Oct 23
IV	25.	Data categorization	26 Oct 23
IV	26.	Levels of Measurement	31 Oct 23
IV	27.	Data management	01 Nov 23
		Sessional II Examination (2nd to 4th Nov. 23) (Syllabus up to III and IV Module)	
IV	28.	Indexing	07 Nov 23
IV	29.	Introduction to statistical learning	08 Nov 23
IV	30.	Introduction to statistical learning	09 Nov 23
V	31.	Measures of central tendency	16 Nov 23
V	32.	Measures of the location of dispersions	21 Nov 23
V	33.	Practice and analysis	22 Nov 23
V	34.	Statistical hypothesis generation	23 Nov 23
V	35.	Testing: Chi-Square test, t-Test	28 Nov 23
V	36.	Analysis of variance	29 Nov 23
V	37.	Correlation analysis	30 Nov 23
V	38.	Maximum likelihood test	05 Dec 23
V	39.	Practice, and analysis	06 Dec 23
	40.	Revision	07 Dec 23

Theory and Lab Evaluation Policy

Theory Evaluation Policy

Evaluation Component	Marks	Date	Remarks
Sessional I and II	30	As per academic calendar	Module 1 to 4
End Semester Exam	60	As per academic calendar	Entire Course
Teacher's Assessment	10	Throughout the Semester	Attendance and MCQ QUIZ

Lab Evaluation Policy

Evaluation Component	Marks	Tentative Date	Remarks
Sessional I Assessment	10	03/10/2023 to 05/10/2023	Module 1 and 2
Sessional II Assessment	10	07/11/2023 to 09/11/2023	Module 3 and 4
Mini Project	20	05/12/2023 to 07/12/2023	Entire Course
Attendance	10	Entire Semester	75% Attendance is Required