



**Ahmedabad
University**

Course	CSD102 Data Science	Semester	Monsoon Semester 2024		
Faculty Name(s)	Anuja Gupta, Kuntalkumar Patel, Shefali Naik, Sonali Sahai, Virali Vora	Contact	gupta.anuja@ahduni.edu.in, kuntal.patel@ahduni.edu.in, shefali.naik@ahduni.edu.in, sonali.sahai@ahduni.edu.in, voravirali@gmail.com		
School	SEAS	Credits	3		
GER Category:	Mathematical and Physical Sciences	Teaching Pedagogy Enable:NO	P/NP Course: Can not be taken as P/NP		
Schedule	Section 1	08:00 am to 09:30 am	Mon	29-07-24 to 26-11-24	
		08:00 am to 09:30 am	Fri	29-07-24 to 26-11-24	
	Section 4	01:00 pm to 02:30 pm	Wed	29-07-24 to 26-11-24	
		01:00 pm to 02:30 pm	Fri	29-07-24 to 26-11-24	
	Section 2	04:00 pm to 05:30 pm	Tue	29-07-24 to 26-11-24	
		04:00 pm to 05:30 pm	Thu	29-07-24 to 26-11-24	
	Section 3	05:30 pm to 07:00 pm	Tue	29-07-24 to 26-11-24	
		05:30 pm to 07:00 pm	Thu	29-07-24 to 26-11-24	

Prerequisite	Not Applicable
Antirequisite	CSD100 Introduction to Data Science & CSD101 Fundamentals of Data Science or CSD101 Intermediate Level Data Science
Corequisite	Not Applicable
Course Description	<p>This course will introduce data science that will be useful in data analytics and visualization. Students will learn basics of statistics that they will apply for data collection, data cleaning, data modeling, data analysis and data visualization using the tools MS Excel, Tableau, Picktochart and QGIS. Students will be introduced to the Python programming for data science.</p> <p>Data science is an interdisciplinary area that involves recording, storing and analyzing data to gain insights and knowledge for decision making. This course is offered to first year students that emphasis on python programming and statistics. The course covers data science life cycle, data collection using sampling/surveys, organizing, processing and visualization of data using maps, charts and infographics. Descriptive statistics, probabilistic approaches, cases and practical applications of data science are discussed using spreadsheet modeling and python programming.</p>
Course Objectives	<p>The main objectives of the course are to make students</p> <ul style="list-style-type: none"> • Familiar with concepts of statistics useful in data science. • Learn tools for data analysis such as MS Excel, Tableau and Piktochart. • Acquaint with basics of GIS and GIS mapping. • Understand methods of data collection, data organization, data analysis and data presentation. • Study fundamentals of Python programming.
Learning Outcomes	<p>Upon the completion of this course, the students will be able to :</p> <ul style="list-style-type: none"> • Create data file, clean and organize the data, analyze data and visualize the data in Excel or Tableau. • Prepare GIS maps using the tool Quantum GIS. • Prepare infographics using the tool Piktochart. • Do statistical analysis and write small programs using Python.

Pedagogy	<p>Classroom Teaching: Students will be taught practical implementation of statistical methods used for data science.</p> <p>Flipped Classroom: Classroom teaching which will be flipped with practical demonstration.</p> <p>Activity Based Learning: Concept wise interesting activities will be given to individuals or groups which they can complete within few days.</p>
Expectation From Students	<ul style="list-style-type: none"> • Interactive during the sessions • Curious to learn new concepts • Readiness for hands on using various tools and for programming • Read prescribed books, reference books and reading material • Submit assignments and projects on time
Assessment/Evaluation	<ul style="list-style-type: none"> • Mid-Semester Examination: <ul style="list-style-type: none"> ◦ Theory Examination - 20% • End Semester Examination: <ul style="list-style-type: none"> ◦ Theory Examination - 30% • Other Components: <ul style="list-style-type: none"> ◦ Infographics Project - 20% ◦ Assignment - 30%
Attendance Policy	As per Ahmedabad University Policy.
Project / Assignment Details	<ul style="list-style-type: none"> • Assignments • Infographics Project

Course Material	<p>Text Book(s)</p> <ul style="list-style-type: none"> Statistics for Management, Business Statistics, J. Joseph Francis, Second Edition, Cengage learning, Year: 2024, <p>Reference Book</p> <ul style="list-style-type: none"> Statistics, David Freedman, Robert Pisani and Roger Purves, Fourth Edition, W. W. Norton & Company Ltd., MS Excel 2013 Bible, John Walkenbach, First Edition, Wiley and Sons Inc., Step by Step Microsoft Excel 2013, Curtis D. Frye, Microsoft Press, Year: 2013, Learning QGIS, Anita Graser, Third Edition, Packt Publishing, Year: 2016, <p>Coursepacks</p> <ul style="list-style-type: none"> Statistics for Management, Richard Levin, David Rubin, Masood Siddiqui and Sanjay Rastogi, Eighth, Pearson Education, Description, Course Pack for Computer Science related sessions, The course pack contains notes on introduction to data science, gis mapping using the tool QGIS and fundamentals of python programming., Statistics for Management, Richard Levin, David Rubin, Masood Siddiqui and Sanjay Rastogi, Eighth, Pearson Education, Description, Course Pack for Computer Science related sessions, The course pack contains notes on introduction to data science, gis mapping using the tool QGIS and fundamentals of python programming.,
Additional Information	Students are expected to bring laptops during the sessions which require programming.

Session Plan

NO.	TOPIC TITLE	TOPIC & SUBTOPIC DETAILS	READINGS,CASES,ETC.	ACTIVITIES	IMPORTANT DATES
1	Introduction to key terminology in data science	Data Science terminology Data Science in different domains Concepts of Big Data and Data Mining Importance of Datasets.	Web Ref : https://machinelearning-blog.com , https://www.edureka.co/blog/what-is-data-science	Discussion of Course Objectives and Expectations, Discussions on basic concepts of Data Science and its applications	
2	Introduction to key terminology in data science	Relationship between Artificial Intelligence, Machine and Deep Learning, Data Analysis Process.	Teaching Notes from the Book: Machine Learning using Python (2019), M Pradhan, UD Kumar Ch. 1 Introduction to Machine Learning Web Ref : https://machinelearning-blog.com	Discussion on importance of data, data life cycle and its applications through Machine and Deep Learning.	
3	Introduction to geographical information systems	Concepts of GIS and Maps, Examples of GIS Applications, Vector and Raster data, Overview of Projection	Teaching Notes from the Book: Practical GIS (2017), Gábor Farkas, ISBN : 9781787123328	Demonstration of vector and raster data from various sources and understanding the formats	
4	Introduction to geographical information systems	Understanding Attribute Table, Geo-referencing	Teaching Notes from the Book: Practical GIS (2017), Gábor Farkas, ISBN : 9781787123328	1. Demonstration of geo-referencing on the map of Gujarat 2. Students will perform an exercise for the given map of India	
5	Introduction to geographical information systems	Creating vector point, vector line and polygon data on map	Teaching Notes from the Book: Practical GIS (2017), Gábor Farkas, ISBN : 9781787123328	Demonstration of plotting points, lines and polygons on the map to represent Cities, Highways and Lakes respectively.	

6	Introduciton to geographical infromation systems	Examples of Creating vector point, vector line and polygon data on map	Teaching Notes from the Book: Practical GIS (2017), Gábor Farkas, ISBN : 9781787123328	Students will perform an exercise on the given map of India	
7	Types of data, scales of measurement, and methods for collection	Categorical/Qualitative data, Numerical / Quantitative data, Nominal, Ordinal, Interval and Ratio scales of data			
8	Types of data, scales of measurement, and methods for collection	Collecting raw data, arranging data using arrays, frequency tables, Grouping data			
9	Data cleaning using MS Excel	Whats\' and \'Whys\' of Data Cleaning? Methods of Data Cleaning\n \n Data Cleaning using Excel - functions and tactics	https://support.microsoft.com/en-us/office/top-ten-ways-to-clean-your-data-2844b620-677c-47a7-ac3e-c2e157d1db19	Demonstration of the functions and exercise for students on a given dataset	
10	Simple descriptive statistics	Measures of central tendency: Arithmetic mean, Weighted mean, Geometric mean, Median, Mode	• Teaching Notes\n• Levin, Chapter-3 Measures of Central Tendency and Dispersion in Frequency Distributions	• Classroom teaching\n• Classroom discussion\n• Problem solving\n• Practical Demonstration using Excel	
11	Simple descriptive statistics	Measures of Dispersion: Range, Standard deviation, Variance, Mean Absolute deviation, Mean Absolute deviation from the median, Quartiles, Deciles, Percentiles, Interquartile range, Quartile deviation, Coefficient of range, Coefficient of quartile deviation, Coefficient of variation, Coefficient of dispersion, Box-plots	• Teaching Notes\n• Levin, Chapter-3 Measures of Central Tendency and Dispersion in Frequency Distributions	• Classroom teaching\n• Classroom discussion\n• Problem solving\n• Practical Demonstration using Excel	

12	Simple descriptive statistics	Measure of Divergence from Normality: Skewness and Kurtosis	• Teaching Notes\n• Levin, Chapter-3 Measures of Central Tendency and Dispersion in Frequency Distributions	• Classroom teaching\n• Classroom discussion\n• Problem solving\n• Practical Demonstration using Excel	
13	Data visualisation using MS Excel and other software	Types of graphical displays available to visualize the data - Charts and graphs\n\nChoosing the type of Graphical display and Modifying the default graphical displays	https://www.juiceanalytics.com/writing/reading-visualizations-for-beginners , https://queue.acm.org/detail.cfm?id=1805128 \n\n https://datajournalism.com/read/handbook/one/understanding-data/using-data-visualization-to-find-insights-in-data	Hands-on application to real-time datasets	
14	Data visualisation using MS Excel and other software	Creating a Data Dashboard for story telling	https://www.toptal.com/designers/data-visualization/data-visualization-best-practices	Hands-on application to real-time datasets	
15	Introduction to computer programming	• Introduction to Problem solving and algorithmic thinking \n• Problem solving using Pseudocode, Flowchart and Algorithms	• Teaching Notes	Examples of problem solving,\n Writing Pseudocodes, Drawing Flowcharts	
16	Introduction to computer programming	• Understanding Python Programming Environment\n• Basic syntax and key terminologies\n• Data types	• Teaching Notes\n• docs.python.org/3	Classroom discussions,\n Activity based on simple python programs	
17	Introduction to computer programming	• Numbers and Strings\n• Operators	Python Data Science Handbook by Jake VanderPlas, Ch-2	Classroom discussions, Activities based on numbers and string processing	
18	Introduction to computer programming	• Branching statements\n• Looping statements	docs.python.org	Classroom teaching and demonstrations	

19	Introduction to computer programming	· Data Structures\n· Python Collections	· docs.python.org\n· Handouts	Classroom discussions, hands-on using collections	
20	Introduction to computer programming	Python libraries\n· Pandas Series\n· Pandas Data Frames objects	· Jake VanderPlas, Ch-3 Data Manipulation with Pandas	Demonstrations - creating series and data frames using Pandas	
21	Introduction to computer programming	Working with datasets\n· Importing data\n· Handling missing data\n· Combine, Concatenate, Append, Merge	· Jake VanderPlas, Ch-3 Data Manipulation with Pandas	Classroom teaching, Applying various operations on data sets using Pandas	
22	Introduction to computer programming	Data Visualization using Matplotlib\n· Line plots, Scatter plots, Histograms\n· Customizing plots	· Jake VanderPlas, Ch-4 Visualization with Matplotlib	Classroom teaching, Generating plots/charts using Matplotlib	
23	Infographics project	Developing a story with a strong narrative supported by various graphics. Projected to be selected based on the Foundation Studio.			
24	Infographics project	Developing a story with a strong narrative supported by various graphics. Projected to be selected based on the Foundation Studio.			
25	Infographics project	Developing a story with a strong narrative supported by various graphics. Projected to be selected based on the Foundation Studio.			
26	Infographics project	Developing a story with a strong narrative supported by various graphics. Projected to be selected based on the Foundation Studio.			
27	Examination and Evaluation				
28	Examination and Evaluation				

29	Review and Reflection	Review and Reflection			
30	Review and Reflection	Review and Reflection			

