

**Program Name** : Diploma in Artificial Intelligence and Machine Learning  
**Program Code** : AN  
**Semester** : Third  
**Course Title** : Data Story Telling and Visualization  
**Course Code** : 22396

### 1. RATIONALE

This course will cover the basics of Data Story Telling and Data Visualization. This course also covers the constructing data stories, extracting meaningful data and representing data for effective visualization. This course gives better visualization to attract audiences. This course will enable the students to inculcate concept of Data Story Telling and methodology to create various visual effects.

### 2. COMPETENCY

The aim of this course is to help the student to attain the following *industry identified* competency through various teaching learning experiences:

- **Construct Data Story Telling and create various visual effects for better visualization**

### 3. COURSE OUTCOMES (COs)

The theory, practical experiences and relevant soft skills associated with this course are to be taught and implemented, so that the student demonstrates the following *industry oriented* COs associated with the above mentioned competency:

- Understand the context and visual effects of data story telling.
- Eliminate clutter and focus audience attention. (Focus audience attention after eliminating clutter)
- Prepare data storytelling according to designer's view.
- Construct/Build data storytelling.
- Create data visualization via single distribution.
- Create data visualization via many distributions.

### 4. TEACHING AND EXAMINATION SCHEME

Teaching Scheme				Credit (L+T+P)	Examination Scheme											
L	T	P	Theory						Practical							
			Paper Hrs.		ESE		PA		Total		ESE		PA		Total	
				Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	
3	-	2	5	3	70	28	30*	00	100	40	25@	10	25	10	50	20

(\*\*) marks should be awarded on the basis of internal end semester theory exam of 50 marks based on the specification table given in S. No. 9.

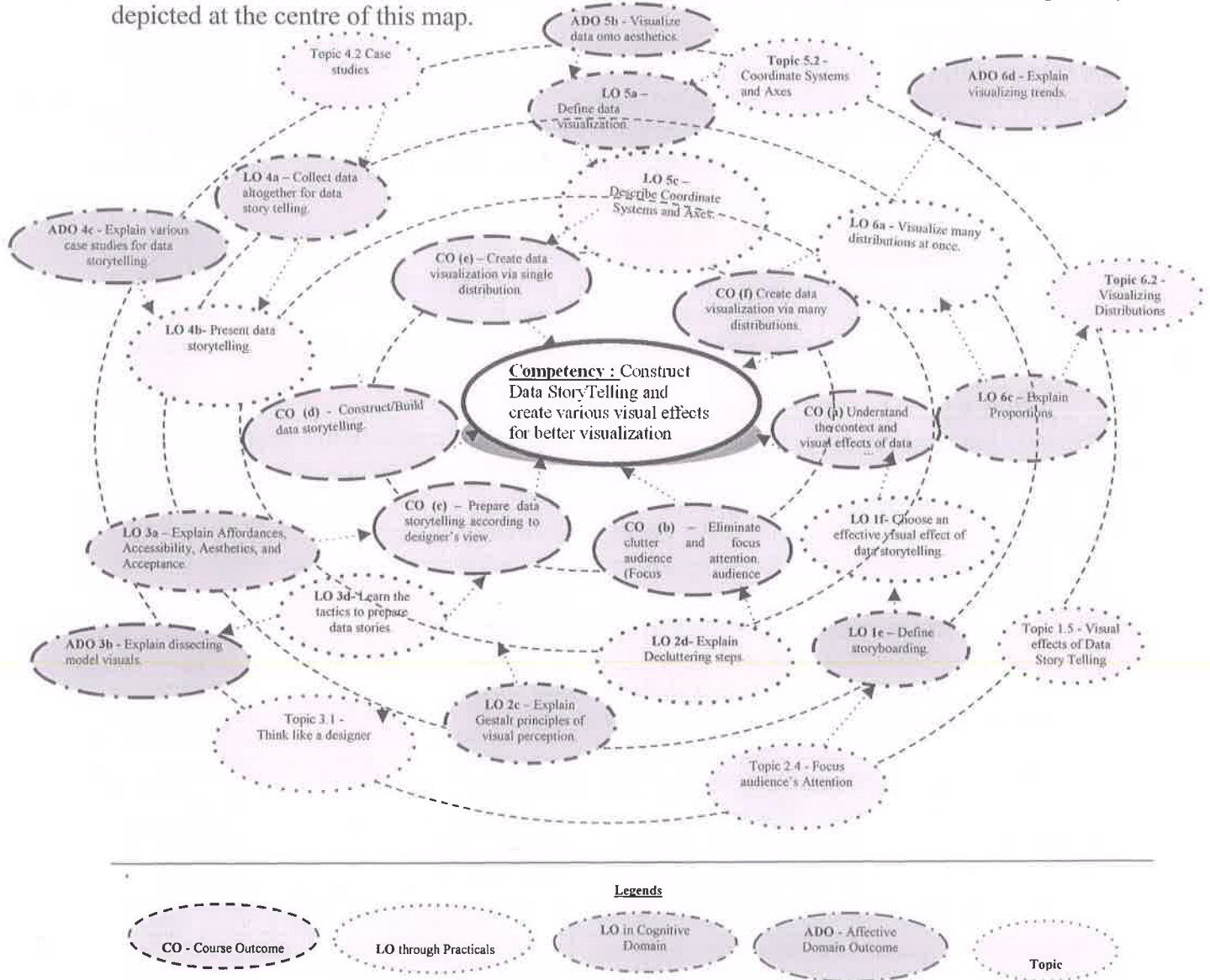
(~<sup>2</sup>): For the **practical only courses**, the PA has two components under practical marks i.e. the assessment of practicals (seen in section 6) has a weightage of 60% (i.e. 30 marks) and micro-project assessment (seen in section 12) has a weightage of 40% (i.e. 20 marks). This is designed to facilitate attainment of COs holistically, as there is no theory ESE.



**Legends:** L-Lecture; T – Tutorial/Teacher Guided Theory Practice; P - Practical; C – Credit, ESE - End Semester Examination; PA - Progressive Assessment, '#': No Theory Examination

**5. COURSE MAP (with sample COs, Learning Outcomes i.e. LOs and topics)**

This course map illustrates an overview of the flow and linkages of the topics at various levels of outcomes (details in subsequent sections) to be attained by the student by the end of the course, in all domains of learning in terms of the industry/employer identified competency depicted at the centre of this map.



**Figure 1 - Course Map**

**6. SUGGESTED PRACTICALS/ EXERCISES**

The practicals/exercises/tutorials in this section are psychomotor domain LOs (i.e. sub-components of the COs) are to be developed and assessed in the student to lead to the attainment of the competency.

S. No.	Practical Exercises (Learning Outcomes to be achieved through practicals)	Unit No.	Approx. Hrs Required
1	Identify a project you are working on where you need to		



S. No.	Practical Exercises (Learning Outcomes to be achieved through practicals)	Unit No.	Approx. Hrs. Required
	communicate in a data-driven way. Reflect upon and write the answers of following questions. 1) WHO IS YOUR AUDIENCE? a. List the primary groups or individuals to whom you'll be communicating. b. If you had to narrow that to a single person, who would that be? c. What does your audience care about? d. What action does your audience need to take?		
2	Identify a project you are working on where you need to communicate in a data-driven way. Reflect upon and write the answers of following questions. 1) WHAT IS AT STAKE? a. What are the benefits if your audience acts in the way that you want them to? b. What are the risks if they do not? 2) FORM YOUR BIG IDEA It should: a. Articulate your point of view b. Convey what's at stake c. Be a complete (and single!) sentence.	I	2*
3	Buses Bunching (Part-I) - Make a data story on bus bunching, (Bus Bunching means when a bus gets delayed and later causes multiple buses to arrive at a single stop at the same time.) a. Find audience, communication mechanism, and desired tone. b. Select effective visual.	I	2*
4	Buses Bunching (Part-II) a. Identify and eliminate clutter. b. Select suitable pre attentive attributes. c. Explore affordances, accessibility, and aesthetics as per designers view.	II, III	2*
5	Create a video (up to 5 minutes) telling your story on practical no. 3 and 4.	I, II, III, IV	2*
6	Make a data story on Percentage of Indian Population by Age Group.	I, II, III, IV	2*
7	Create a data story w.r.t. following observations: a. What are the most popular pick-up locations across the city for Citi Bike rental? b. How does the average trip duration vary across different age groups, and over time? c. Which age group rents the most bikes? d. How does bike rental vary across the two user groups (one-time users vs. long-term subscribers) on different days of the week?	I, II, III, IV	2





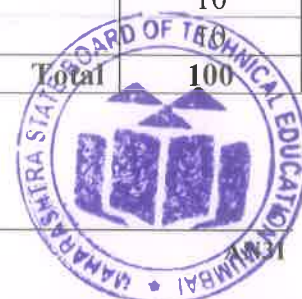
S. No.	Practical Exercises (Learning Outcomes to be achieved through practicals)	Unit No.	Approx. Hrs. Required
	e. Do factors like weather and user age impact the average bike trip duration?		
8	Create a video (up to 5 minutes) telling your story. You might want to record a video of yourself speaking, or narrate while showing visual props or sketches, or screencast a PowerPoint presentation, etc. You get to choose how you present the story. You should produce a single video file (formatted as a .mov file)	I, II, III, IV	2*
9	Implement python program that loads any dataset and plot the grouped bars.	V	2*
10	Implement python program that perform data cleaning on any dataset.	V	2*
11	Implement Python program to perform data pre-processing, analysis and visualization.	V	2*
12	Apply Data analysis using Python for available data set. (Apply Supervised machine learning algorithm.)	V	2
13	Implement Python program to apply Image Processing functions like display, plot and colour by using sklearn or other libraries	V	2
14	Connecting to Database and extracting data from database table in Python.	V	2*
15	Develop worksheet, add filters and create chart using dataset in Python.	VI	2
16	Create and organize above (Practical No. 6) data visualizations into the bar chart.	VI	2*
17	Create and organize above (Practical No. 7) data visualizations into the bar chart.	VI	2
<b>Total</b>			<b>34</b>

\*: compulsory practicals to be performed.

### Note

- i. Given in above tables is suggestive list of practical exercises. Teachers can design other similar exercises.
- ii. Assessment of the 'Process' and 'Product' related skills in the laboratory/workshop/field work should be done as per suggested sample below:

S. No.	Performance Indicators	Weightage in %
1	Gather information for data story and select proper visualization method.	20
2	Use standard statements to prepare data story and represent it pictorially.	40
3	Present data story and execute visualization programs.	20
4	Able to answer oral questions.	10
5	Submission of report in time.	
<b>Total</b>		<b>100</b>



Additionally, the following affective domain LOs (social skills/attitudes), are also important constituents of the competency which can be best developed through the above mentioned laboratory/field based experiences:

- Present data story.
- Demonstrate working as a leader / a team member.
- Maintain tools and equipment.
- Follow ethical practices.

The development of the attitude related LOs of Krathwohl's 'Affective Domain Taxonomy', the achievement level may reach:

- 'Valuing Level' in 1<sup>st</sup> year
- 'Organizing Level' in 2<sup>nd</sup> year and
- 'Characterizing Level' in 3<sup>rd</sup> year.

### 7. MAJOR EQUIPMENT/ INSTRUMENTS REQUIRED

The major equipment with broad specification mentioned here will usher in uniformity in conduct of experiments, as well as aid to procure equipment by authorities concerned.

S. No.	Equipment Name with Broad Specifications	Expt. S. No.
1	Hardware: Personal computer, (i3-i5 preferable), RAM minimum 2 GB onwards.	For all Experiments
2	Operating system: Windows 7 onwards	
3	Software: Python	

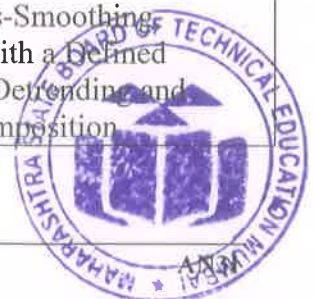
### 8. UNDERPINNING THEORY COMPONENTS

The following topics/subtopics should be taught and assessed in order to develop LOs in cognitive domain for achieving the COs to attain the identified competency.

Unit	Major Learning Outcomes (in cognitive domain)	Topics and Sub-topics
<b>Unit – I Context and Visual effects of Data Story Telling</b>	1a. Understand the importance of context. 1b. Differentiate between exploratory and explanatory analysis. 1c. Identify audience. 1d. Explain 3-minutes story. 1e. Define storyboarding. 1f. Choose an effective visual effect of data storytelling.	1.1 Concept / Importance of Context 1.2 Exploratory vs. explanatory analysis 1.3 Who - Your audience, You, What – Action, Mechanism, Tone, How, Example 1.4 The 3 minutes story, Big Idea, Storyboarding. 1.5 <b>Visual effects of Data Story Telling</b> - Choosing an effective visual - Simple text, Tables, Graphs, Points, Bars - Vertical bar chart, Horizontal bar chart
<b>Unit– II Focus audience's Attention by eliminating clutter</b>	2a. Define clutter. 2b. Define cognitive load. 2c. Explain Gestalt principles of visual perception. 2d. Explain Decluttering steps. 2e. Describe Preattentive attributes in text, graphs, page.	2.1 <b>Clutter is our enemy</b> - Cognitive load, Clutter 2.2 Gestalt principles of visual perception, Lack of visual order 2.3 Decluttering: step-by-step 2.4 <b>Focus audience's Attention</b> - Preattentive attributes in text look 2.5 Preattentive attributes in text



Unit	Major Learning Outcomes (in cognitive domain)	Topics and Sub-topics
		2.6 Preattentive attributes in graphs 2.7 Size, Color, Position on page
<b>Unit– III Storytelling from designers view</b>	3a. Explain Affordances, Accessibility, Aesthetics, and Acceptance. 3b. Explain dissecting model visuals. 3c. Understand the magic of story. 3d. Learn the tactics to prepare data stories.	3.1 <b>Think like a Designer</b> -Affordances 3.2 Accessibility 3.3 Aesthetics 3.4 Acceptance 3.5 Dissecting model visuals - line graph, 100% stacked bars 3.6 <b>Lessons in storytelling</b> - The magic of story 3.7 Constructing the story 3.8 The narrative structure 3.9 The power of repetition 3.10 Tactics to help ensure that your story is clear
<b>Unit-IV Pulling it all together for data storytelling</b>	4a. Collect data altogether for data story telling. 4b. Present data storytelling. 4c. Explain various case studies for data storytelling.	4.1 Pulling it all together for data storytelling 4.2 Case studies -Color considerations with a dark background, Leveraging animation in the visuals you present 4.3 Final Thoughts - Where to go from here, Building storytelling with data competency in your team or organization
<b>Unit –V Visualizati on Part - I</b>	5a. Define data visualization. 5b. Visualize data onto aesthetics. 5c. Describe Coordinate Systems and Axes. 5d. Explain Directory of Visualizations.	5.1 From Data to Visualization – Introduction, Visualizing Data: Mapping Data onto Aesthetics. 5.2 Coordinate Systems and Axes 5.3 Directory of Visualizations - Visualizing Amounts - Bar Plots, Grouped and Stacked Bars, Visualizing Distributions - Visualizing a Single Distribution, Visualizing Multiple Distributions at the Same Time, Empirical Cumulative Distribution Functions, Highly Skewed Distributions, Quantile-Quantile Plots
<b>Unit –VI Visualizati on Part - II</b>	6a. Visualize many distributions at once. 6b. Visualize distributions along the axes. 6c. Explain Proportions. 6d. Explain visualizing trends.	6.1 Visualizing Many Distributions at Once 6.2 Visualizing Distributions Along the Vertical Axis, Visualizing Distributions Along the Horizontal Axis 6.3 Proportions 6.4 Visualizing Trends-Smoothing, Showing Trends with a Defined Functional Form, Detrending and Time-Series Decomposition



*Note: To attain the COs and competency, above listed Learning Outcomes (LOs) need to be undertaken to achieve the 'Application Level' of Bloom's 'Cognitive Domain Taxonomy'.*

### 9. SUGGESTED SPECIFICATION TABLE FOR QUESTION PAPER (INTERNAL) DESIGN

Unit No.	Unit Title	Teaching Hours	Distribution of Theory Marks			
			R Level	U Level	A Level	Total Marks
I	Context and Visual effects of Data Story Telling	06	02	02	04	08
II	Focus audience's Attention by eliminating clutter	08	02	04	06	12
III	Storytelling from designers view	08	02	04	06	12
IV	Pulling it all together for data storytelling	06	02	02	06	10
V	Visualization Part - I	10	02	04	08	14
VI	Visualization Part - II	10	02	04	08	14
<b>Total</b>		<b>48</b>	<b>12</b>	<b>20</b>	<b>38</b>	<b>70</b>

**Legends:** R=Remember, U=Understand, A=Apply and above (Bloom's Revised taxonomy)

**Note:** This specification table provides general guidelines to assist student for their learning and to teachers to teach and assess students with respect to attainment of LOs. The actual distribution of marks at different taxonomy levels (of R, U and A) in the question paper may vary from above table.

This specification table also provides a general guideline for teachers to frame internal end semester practical theory exam paper which students have to undertake.

### 10. SUGGESTED STUDENT ACTIVITIES

Other than the classroom and laboratory learning, following are the suggested student-related **co-curricular** activities which can be undertaken to accelerate the attainment of the various outcomes in this course:

- Prepare journals based on practical performed in laboratory.
- Library/E-Book survey regarding Python programming used in Computer industries.
- Prepare power point presentation for showing different types of Python Programming Applications.

### 11. SUGGESTED SPECIAL INSTRUCTIONAL STRATEGIES (if any)

These are sample strategies, which the teacher can use to accelerate the attainment of the various outcomes in this course:

- Massive open online courses (**MOOCs**) may be used to teach various topics/sub topics.
- '**L**' in item No. 4 does not mean only the traditional lecture method, but different types of teaching methods and media that are to be employed to develop the outcomes.
- About **15-20% of the topics/sub-topics** which is relatively simpler or descriptive in nature is to be given to the students for **self-directed learning** and assess the development of the LOs/COs through classroom presentations (see implementation guideline for details).
- With respect to item No.10, teachers need to ensure to create opportunities and provisions for **co-curricular activities**.





- e. Guide student(s) in undertaking micro-projects.
- f. No. of practical's selection to be performed should cover all units.
- g.

## 12. SUGGESTED MICRO-PROJECTS

**Only one micro-project** is planned to be undertaken by a student assigned to him/her in the beginning of the semester. S/he ought to submit it by the end of the semester to develop the industry oriented COs. Each micro-project should encompass two or more COs which are in fact, an integration of practicals, cognitive domain and affective domain LOs. The micro-project could be industry application based, internet-based, workshop-based, laboratory-based or field-based. Each student will have to maintain dated work diary consisting of individual contribution in the project work and give a seminar presentation of it before submission. The total duration of the micro-project should not be less than **16 (sixteen) student engagement hours** during the course.

A suggestive list is given here. Similar micro-projects could be added by the concerned faculty:

- a. Car Racing in last 5 years
- b. Corona Virus Breakdown
- c. Regional Language Awareness
- d. Percentage of candidates in various countries on Twitter
- e. Indian Industry Revolution
- f. Cricket Scorecard of different matches for particular player

## 13. SUGGESTED LEARNING RESOURCES

S. No.	Title of Book	Author	Publication
1	Storytelling with data – a data visualization guide for business professionals	Cole Nussbaumer Knaflic	Wiley India Pvt. Ltd ISBN: 978-1-119-00225-3
2	Fundamentals of Data Visualization	Claus O. Wilke	O'Reilly Media, Inc. ISBN: 9781492031086
3	Fundamentals of PYTHON	Kenneth A Lambert, B.L. Juneja	CENGAGE Learning, ISBN:978-81-315- 2903-4

## 14. SOFTWARE/LEARNING WEBSITES

- a. <https://nptel.ac.in/courses/106106182>
- b. <https://careerfoundry.com/en/tutorials/data-analytics-for-beginners/storytelling-with-data/>
- c. <https://blog.hubspot.com/marketing/great-data-visualization-examples>

