MOHAMED SATHAK ENGINEERING COLLEGE
KILAKARAI-623 806, RAMANATHAPURAM DIST.
Approved by AICTE, COA, New Delhi, DGS Mumbai, Affiliated to Anna University,
An ISO 9001:2015 Institution, Sponsored by Mohamed Sathak Trust, Chennai - 06.
(Recognized under section 2(f) & 12B of UGC, NewDelhi)

DEPARTMENT OF INFORMATION TECHNOLOGY

Innovative Teaching Methodology –Animated Videos

Course Code & Name : C203/CS3352-Foundation of Data Science

Name of the Course Instructor(s) : Mrs.K.Samundeeswari

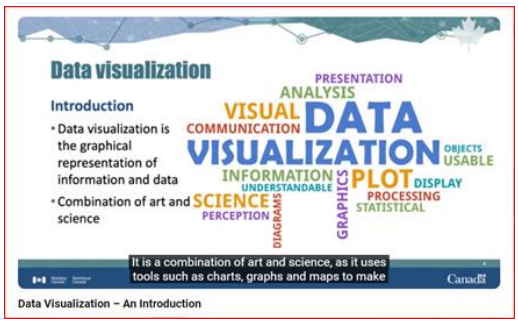
Class : II IT A

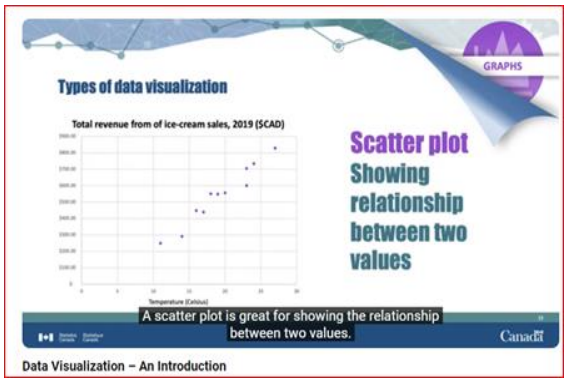
CO's: Apply visualization Libraries in Python to interpret and explore data

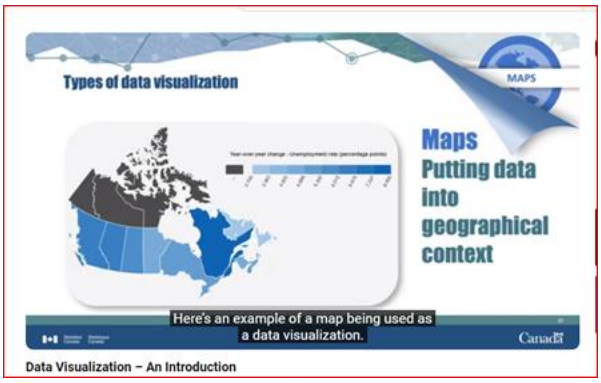
Topic Name: Date Visualization

URL: https://www.youtube.com/@StatisticsCanada

Sample Images:







Signature of Faculty **HOD**

DEPARTMENT OF INFORMATION TECHNOLOGY

Academic Year 2023 – 2024 (Odd Semester)

Degree, Semester & Branch : III Semester B.Tech IT
Course Code & Title : CS3352 & Foundations of Data Science
Name of the Faculty : Mrs.K.Samundeeswari

Innovative Practice Description

- **Unit / Topic: Unit V / Data Visualization /** Histograms and Legends
- **Course Outcome:** CO5
- **Topic Learning Outcome:**
- **Activity Chosen:** Flipped Classroom
- **Justification:**

  Flipping the classroom is an inverting the classroom approach to teaching. In this approach, the traditional in-class teaching is "flipped" to better meet the needs of individual learners. Students gain control of the learning process through studying course material outside of class, using readings, pre-recorded video lectures. It helps the faculty/lecturer to redefine in-class activities and include homework problems and keep the students engaged in the content.

- **Time Allotted for the Activity:** 45 Minutes
- **Details of the Implementation:**
  - **Plan:** Basics of Data visualization - Histograms and Legends is given as self-learning through self-exploration and laboratory exercises. Separate student's teams are formed and made them to explore as a group.
  - **Identify and Share:** Related materials are identified and I posted relevant references/materials/notes of Data visualization - Histograms and Legends to the students and given as self-learning. Also book content is also shared with the students to explore more on Data visualization. Students are requested to make a PPT presentation of the Data visualization - Histograms and Legends.
  - **Evaluate:**
    - I have prepared few questions related to the content shared and ask those questions in next class session and make the students to write/present in class.
    - Make the students to find out the answer by their own by learning.

- **CO – PO / PSO mapping:**

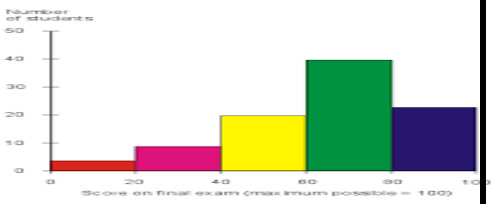
CO5	Apply visualization Libraries in Python to interpret and explore data

Innovative practice	**PO1**	**PO2**	**PO 3**	**PO4**	**PO5**	**PO9**	**PO10**	**PO11**	**PO12**	**PSO1**	**PSO2**
	2	**2**	**3**	**3**	**3**	**1**	**2**	**1**	**2**	**3**	**2**
Justificatio nfor correlation	Apply basic Knowledge and fundamenta ls in Data visualizatio non	Identify the need of Data visualization n in Data Science process	Able to design and develop thevarious charts and graphs using Data visualizati onn	Functional individually in identifying the representation of data with Data visualization methods	Able to Create, select, and apply appropriate techniques, resources, and modern tools for Data Visualization	Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary setting of concepts	Communicat e / share the ideas with other students in visualization the Data	Demonstr ate knowledg e and understan ding Data visualizati on Process in Data Science Projects	Ability to reproduce the contents gathered through self-learning	To identify and implement appropriate techniques, resources, modern tools for providing solution to new idea and innovation for Data Visualization	To manage complex IT projects in Data Science and an understanding of risk management

(1 – Low 2 – Moderate 3 – High)

- **Images / Screenshot of the practice:**

- **Reflective Critique:**



  - ***Feedback of practice from students and other stakeholders:***
    - Student felt good, since they can study at their own pace/time.
    - They felt that through such learning, they can explore more.
  - ***Benefit of the practice:***
    - More one-to-one time between teacher and student.
    - More collaboration time for students.
    - Students learn at their own pace.
    - Practical things – like missing class due to illness – become less problematic.
    - It encourages students to come to class prepared.
  - ***Challenges faced in implementation:***
    - The depth of the subject can be dictated by the student themselves or the group the student is working with.
    - The time and effort required from a teacher's perspective initially when creating the flipped class material is higher than for a traditional class.

References:

1. https://www.w3schools.com/python/pandas/default.asp

Signature of Faculty **HOD**

DEPARTMENT OF INFORMATION TECHNOLOGY

Academic Year 2023 – 2024 (Odd Semester)

Degree, Semester & Branch **:** III Semester B.Tech IT
Course Code & Title **:** CS3352 & Foundations of Data Science
Name of the Faculty **:** Mrs.K.Samundeeswari

Innovative Practice Description

- **Unit / Topic: Unit III / Describing Relationships /** Standard error of estimate
- **Course Outcome:** CO3
- **Activity Chosen:** Problem based Learning
- **Justification:**

  Problem-based learning is an educational approach in which the students learn by doing/ solving some problems and understand the core concepts clearly. Standard error of estimate is an evaluation in the finding the correlation of two variables. To better understand the error calculation process, the students are made to solve a problem. The prime objective of learning by doing is to enhance the problem solving and creative thinking of the students.
- **Time Allotted for the Activity:** 45 Minutes
- **Details of the Implementation:**

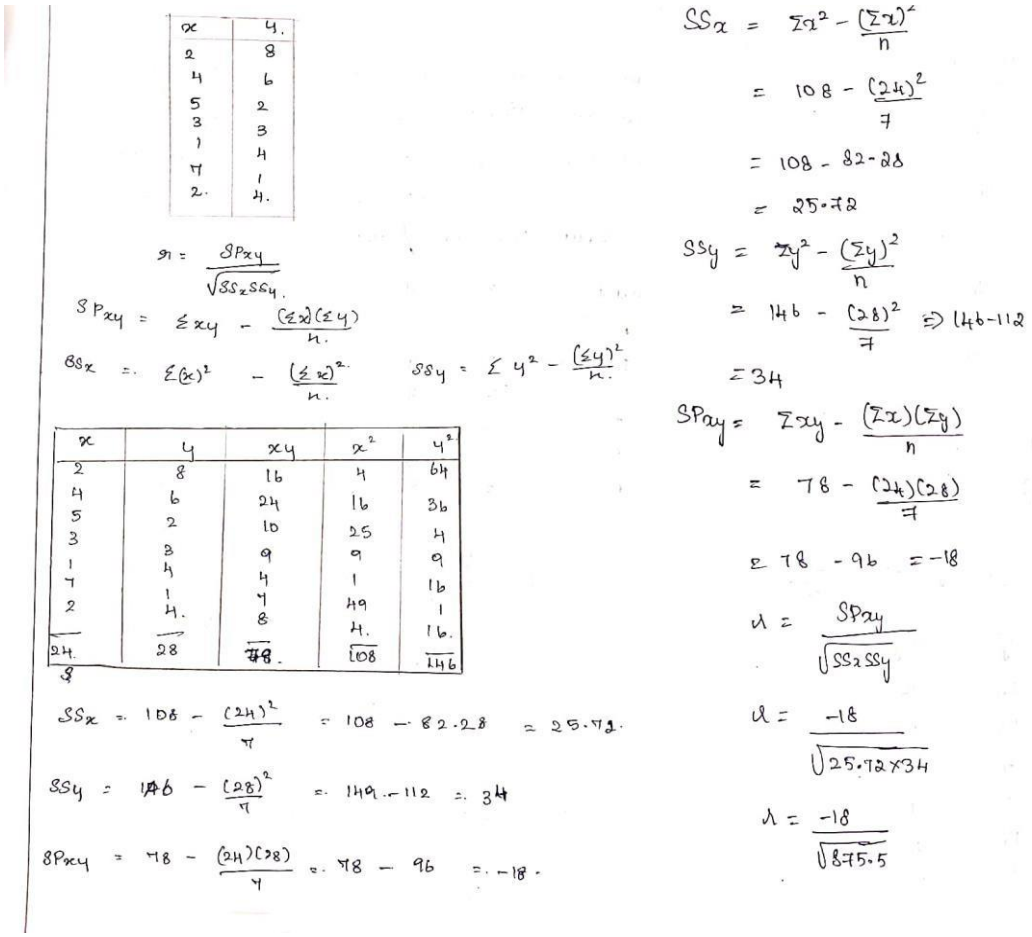
  In the previous class, the formulas for calculating the correction values and standard error estimate value are discussed to the students. A problem statement was given to the students. In the class, the formula and steps were revised once again and the students were supposed to solve the given problem with their teammates. Students were asked to interact and share their ideas in calculating the error estimate methods. Finally, at the end of the session, answers were explained and discussed.

(1 – Low 2 – Moderate 3 – High)

- **PO / PSO mapped:**

Innovative practice	**PO1**	**PO2**	**PO9**	**PO10**	**PO12**
	2	**2**	**2**	**2**	**1**
Justification for correlation	Apply the knowledge of mathematical formulation for Standard error of estimate	Identify the appropriate methods to calculate the Standard error of estimate	Individually identify the problem and finding solutions	Deliver efficient document and solve the problem step by step	Recognize the need for Standard error of estimate in correlation problems in data analysis

- **Images / Screenshot of the practice:**



x	y
2	8
4	6
5	2
3	3
1	4
7	1
2	4

$r = \frac{SP_{xy}}{\sqrt{SS_x SS_y}}$

$SP_{xy} = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}$

$SS_x = \Sigma(x)^2 - \frac{(\Sigma x)^2}{n}$ $\quad SS_y = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$

x	y	xy	x^2	y^2
2	8	16	4	64
4	6	24	16	36
5	2	10	25	4
3	3	9	9	9
1	4	4	1	16
7	1	7	49	1
2	4	8	4	16
24	28	78	108	146

$SS_x = 108 - \frac{(24)^2}{7} = 108 - 82.28 = 25.72$

$SS_y = 146 - \frac{(28)^2}{7} = 146 - 112 = 34$

$SP_{xy} = 78 - \frac{(24)(28)}{7} = 78 - 96 = -18$

$SS_x = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$

$= 108 - \frac{(24)^2}{7}$

$= 108 - 82.28$

$= 25.72$

$SS_y = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$

$= 146 - \frac{(28)^2}{7} \Rightarrow 146 - 112$

$= 34$

$SP_{xy} = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}$

$= 78 - \frac{(24)(28)}{7}$

$= 78 - 96 = -18$

$r = \frac{SP_{xy}}{\sqrt{SS_x SS_y}}$

$r = \frac{-18}{\sqrt{25.72 \times 34}}$

$r = \frac{-18}{\sqrt{875.5}}$

❖ **Reflective Critique:**

- ❖ ***Feedback of practice from students and other stakeholders:***
  Students enjoyed in solving the problem as a team and effectively competed with other student members and there was a healthy competition within the students.
- ❖ ***Benefit of the practice:***
  - ➢ Students learned a few things individually that leads to self-learning. They also actively involved in this session rather than it being a one-way communication.
  - ➢ This activity encouraged the students to share their knowledge with others.
  - ➢ Students were able to connect ideas through discussion.
- ❖ ***Challenges faced in implementation:***
  Slow learners were not comfortable in solving because of their individual (slow) learning background. However, they understood the concepts through the interaction within the team.

Signature of Faculty **HOD**

DEPARTMENT OF INFORMATION TECHNOLOGY

Academic Year 2023 – 2024 (Odd Semester)

Degree, Semester & Branch : III Semester & IT
Course Code & Title : CS3352 & Foundations of Data Science
Name of the Faculty member : Mrs.K.Samundeeswari

Innovative Practice Description

- **Unit / Topic: Unit I /** Basic Statistical descriptions of Data
- **Course Outcome:** CO1
- **Activity Chosen:** Reflection
- **Justification:**

  Reflection is a process where teachers can identify their own teaching process and analyzing how their teaching (content delivery) might be improved for the better understanding of the students. The prime objective of Reflection is to make the students write their doubts in a paper on topics which they are not clear, so as the teacher can discuss and clarify the concepts that are not understood by the students.

- **Time Allotted for the Activity:** 10 Minutes
- **Details of the Implementation:**

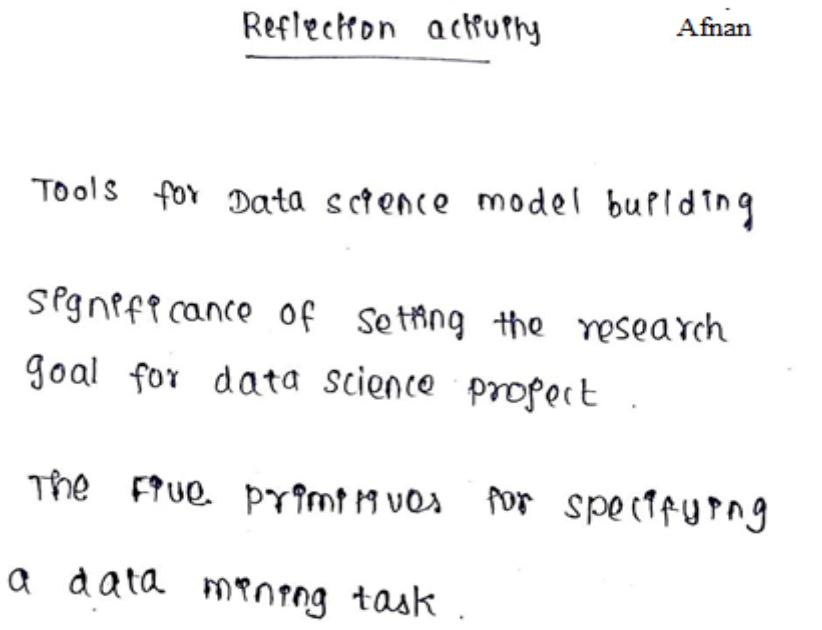
  After I completed the Unit-1, in order to address the doubts and to know the level of understanding of the students, they are insisted to honestly describe the concepts that they didn't understand well. I asked the students to write their doubts in Data Science concepts.

  The students wrote their doubts and topics which they are not clear. The papers are collected by me and I allotted 10 minutes time in the next hours and discussed all the doubts in next class session and make them understand. I also gave a few simple tasks in classroom to gain more understanding in the topics. The students felt they understood the concepts better.

- **PO / PSO mapped:**

Innovative practice	PO9 1	PO10 2	PO12 2
Justification for correlation	**Student individuality in expressing their doubts**	**Written communicatio n/ Asking questions are improved**	**Student will implement the questioning skill in future in written format**

- **Images / Screenshot of the practice:**



Reflection activity Afnan

Tools for Data science model building

Significance of setting the research goal for data science project.

The five primitives for specifying a data mining task.

Reflection Activity

1) List the application of data science?
2) What is project charter?
3) List the stages of data Science process.

- **Reflective Critique:**
  - ❖ ***Feedback of practice from students and other stakeholders:***
    - Students felt the activity conducted gave them a chance to represent their common doubts among students and could get additional clarification from me.

  - ❖ ***Benefit of the practice:***
    - Through this activity the students can recollect the topics on Data Science process, models, and Data Mining, Data Warehousing. The students are made to understand the concepts in which they are not clear and ensured that they understood the topics thoroughly.

  - ❖ ***Challenges faced in implementation:***
    - Students were not able to frame or express their doubts in English sentence properly and some questions were not clear to understand.
    - Slow learners were not interested to participate in this innovative teaching and submitted the empty paper.
    - Some of the students did not even ready to write their doubts
    - Students started to discuss and shout with their neighbors during the activity.

Signature of Faculty **HOD**

DEPARTMENT OF INFORMATION TECHNOLOGY

Academic Year 2023 – 2024 (Odd Semester)

Degree, Semester & Branch : III Semester B.Tech IT

Course Code & Name : CS3352 & Foundations of Data Science

Name of the Faculty : Mrs.K.Sanundeeswari

Innovative Practice Description

- **Unit / Topic: Unit IV / Data Manipulation with Pandas**
- **Course Outcome:** CO4
- **Activity Chosen: Think-Pair-Share**
- **Justification:**
  - Active learning strategy, in which students work on a problem posed by instructor, first individually (Think), then in pairs (Pair) or groups, and finally together with the entire class (Share).
  - **T (Think):** Teacher asks a specific question about the topic. Students "think" about what they know or have learned, and come up with their own individual answer to the question. [Takes 1-3 Minutes].
  - **P (Pair):** Teacher asks another question, related to the previous one that is suitable to deepen the students' understanding of the topic. Each student is paired with another student. They share their thinking with each other and proceed with the task. [Takes 5-10 Minutes].
  - **S (Share):** Students share their thinking (or solution) with the entire class. Teacher moderates the discussion and highlights important points. [Takes 10-20 minutes].
- **Time Allotted for the Activity:** 45 Minutes
- **Details of the Implementation:**
  - Think-Pair-Share innovative practice conducted for II year IT A Section students, after explanation of the concept of Data Manipulation with Pandas.
  - First, I asked the students to think about the real time implementation of the Pandas Program. The students searched few webpages and Develop the program and ways of performing Panel data.

- Then I make them as a pair and share their thinking with their partners.
- Finally I asked any two or three pairs to explain the concept to whole class for further discussion. Abdul Ahad , Afnan , Nivetha and Bala nithis,II year IT A shared their ideas and knowledge of Panel Data and how to write the Given Pandas Program.

Questions

Group 1:Ahad and Team

1.Image you have a series of data that represents the amount of precipitation each day for a year in a given city. Load the daily rainfall statistics for the city of Chennai in 2021. Which is given in a csv file Chennairainfall 2021.csv using Pandas generate a histogram for rainy days and find out the days that have high rainfall.

Group 2 :Nivetha and Team

2. Consider that an E commerce organization like Amazon have different region sales as Northsales, Southsales, Westsales.csv files. They want combine North and west region sales and south and east sales to find the aggregate sales os this collaborating region help them to do so using python code.

Group 3 :Afnan and Team

3.Create a data frame with key and data pairs as A-10,B-20,A-40,C=5,B=10,C=10.Find the sum of each key and display the results a each key group.

Group 4:Bala nithis and team

4.Name some methods you know to replace NaN value of a dataframe in pandas.

5.how are iloc() and loc() different.

- **CO – PO / PSO mapping:**

CO4	Use the Python Libraries for Data Wrangling

Innovative practice	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO9**	**PO10**	**PO11**	**PO12**	**PSO1**	**PSO2**
	2	**2**	**2**	**2**	**3**	**1**	**2**	**2**	**2**	**3**	**2**
Justification for correlation	Apply basic Knowledge and fundamentals in Data wrangling	Identify theneed of Data Manipulation with Pandas	Able to design and develop the Program for Pandas	Functional individually in identifyingthe representation of Panel Data	Able to Create, select, and apply appropriate techniques, resources, and modern tools for Data Wrangling	Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary setting of concepts	Communicate/ share their ideas with other students in Pandas	Demonstrate knowledge and understanding Data visualization Process in Data wrangling	Ability to reproduce the contents gathered through self-learning	To identify and implement appropriate techniques, resources, modern tools for providing solution to new idea and innovation for Data wrangling	To manage complex IT projects in Data Science and an understanding of risk management

(1 – Low 2 – Moderate 3 – High)

- **Images / Screenshot of the practice:**

```
import pandas as pd
ecom=pd.read_csv('../input/ecommerce-purchases-csv/Ecommerce Purchases.csv')
ecom.info( )
```

```
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Address           10000 non-null  object
 1   Lot               10000 non-null  object
 2   AM or PM          10000 non-null  object
 3   Browser Info      10000 non-null  object
 4   Company           10000 non-null  object
 5   Credit Card       10000 non-null  int64
 6   CC Exp Date       10000 non-null  object
 7   CC Security Code  10000 non-null  int64
 8   CC Provider       10000 non-null  object
 9   Email             10000 non-null  object
 10  Job               10000 non-null  object
 11  IP Address        10000 non-null  object
 12  Language          10000 non-null  object
 13  Purchase Price    10000 non-null  float64
```

```
First five lines of rain dataset:
                  precip    area
country
Afghanistan          327.0   652.2
Albania             1485.0    27.4
Algeria               89.0  2381.7
American Samoa         NaN     0.2
Andorra                NaN     0.5
Angola              1010.0  1246.7
Antigua and Barbuda 1030.0     0.4
Argentina            591.0  2736.7
Armenia              562.0    28.5
Aruba                  NaN     0.2
```

```
Program
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
rain = pd.read_csv("Chennai rainfall 2021.csv ")
 rain['country'].max()
rain.hist()
```

- **Reflective Critique:**
  - ❖ ***Feedback of practice from students and other stakeholders:***
    - Working as a pair, they can share their knowledge with them and come up with better understanding about the topic.
    - Different types of reasoning/justification based questions are discussed in the class room, so students come to know how to answer/draw diagrams whenever new real time system is asked in the question.
  - ❖ ***Benefit of the practice:***
    - Students are actively engaged and students learn from each other.
    - Makes class interactive and builds a friendly, yet academic atmosphere.
    - Includes all the students in the teaching-learning process.
  - ❖ ***Challenges faced in implementation:***
    - Students felt very difficult in this activity and could not complete in time whenever new real time system asked in the question.
    - Students are hesitating to speak in front of all.

Signature of Faculty **HOD**