Week-1 PSOSM 2023

1. Which is the correct online platform, as depicted in Figure 1?

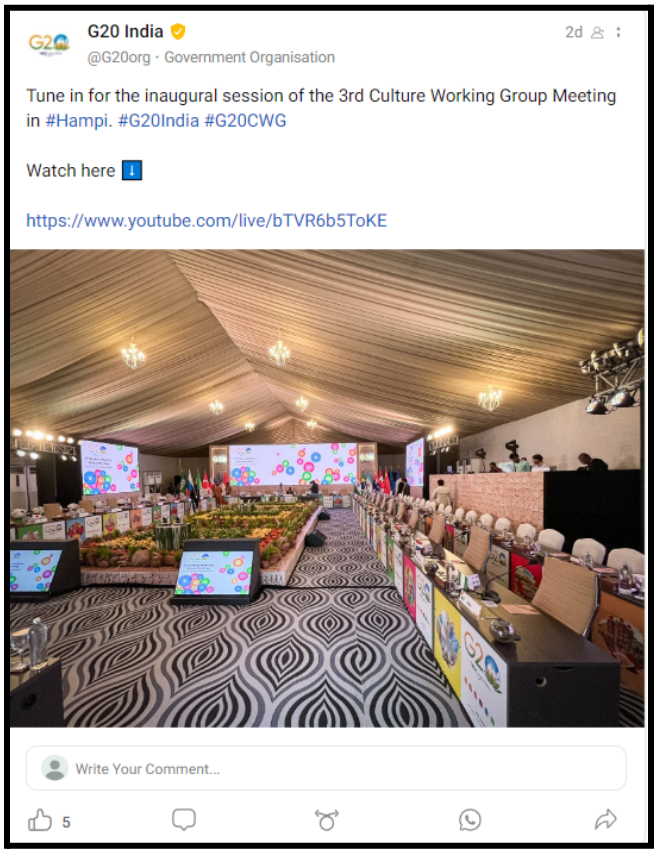


Figure 1

a) Koo
b) Blind
c) Reddit
d) LinkedIn

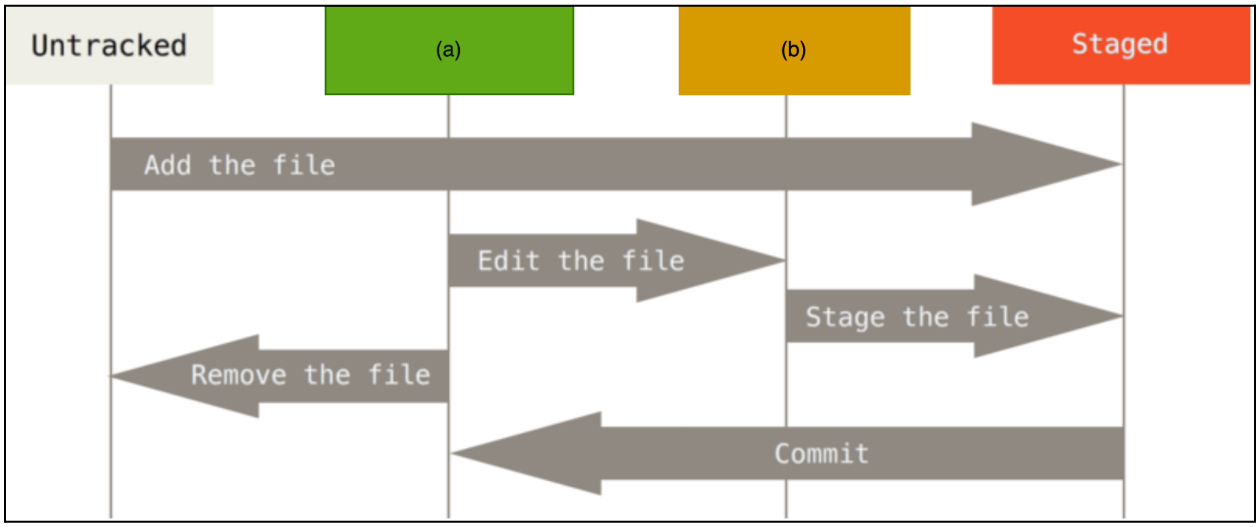
2. **Which of the following statement(s) is/are true about Anonymous social networks?**

a. Anonymous social networks help in providing enhanced privacy and anonymity.
b. Anonymous social networks provide more freedom of expression.
c. Anonymous social networks reduce the fear of judgment or social consequences.
d. None of the above

3. **Choose the correct option to create a virtual environment in Python**
    a. python3 -m venv <path_to_virtual_environment>
    b. python3 venv -m <path_to_virtual_environment>
    c. virtualpython3 -m venv
    d. virtualenv python3 -m venv myenv

4. **Why is it important to create a virtual environment in Python?**
    a. Creating a virtual environment isolates project dependencies and prevents conflicts between different projects.
    b. Creating a virtual environment enhances the performance of Python programs.
    c. Creating a virtual environment to ensure compatibility with different operating systems.
    d. Creating a virtual environment optimises memory usage in Python applications.

5. Fill in the correct option to stage files/directories.



a) Modified, Unmodified
b) Unmodified, modified
c) modified, unstaged
d) Modified, tracked

6. Which command lists all the remote addresses associated with a local repository?
a. git remote -v
b. git pull origin master
c. Git stash
d. Git checkout

7. Which of the following concept is closely related to the fact that infrequent, arms-length relationships are more beneficial for employment opportunities, promotions, and wages than strong ties?
a. Six degrees of separation
b. Small world phenomenon
c. Strength of weak ties
d. None of the above

8. Which of the following is/are shortcomings of online social media platforms?
a) Spread of misinformation
b) Cyberbullying and harassment
c) Addiction and time consumption
d) limited control over privacy

9. When a piece of content, such as a video, meme, or news article, gains attention on social media platforms, it can quickly reach a large audience within a short period. Which V of social media does it refer to?

a. Velocity
b. Veracity
c. Value
d. Volume

10. What is the primary purpose of the Gab platform?

A) To promote free speech
B) To facilitate online shopping experience
C) It provides educational resources
D) support for professional networking

Week-2 PSOSM 2023

1. What is the primary data format utilised by the Reddit API when returning information?
    a. JSON
    b. XML
    c. HTML
    d. CSV

2. JSON stands
    a) JavaScript Object Notation
    b) JavaScript Object Network
    c) JavaScript Oriented Notation
    d) JavaScript Oriented Namespace

3. Suppose you want to retrieve the top 5 hottest posts from the subreddit "r/technology" using the Reddit API. Which Python library would you most likely use to achieve this?
    a. PRAW
    b. Requests
    c. Urllib
    d. JSON

Explanation: PRAW (Python Reddit API Wrapper) is a Python library designed explicitly for interacting with the Reddit API. Using PRAW, you can easily retrieve the top 5 hottest posts from the subreddit "r/technology" and perform various other interactions with the Reddit platform.

4. You want to fetch the details of the 10 hot posts from the subreddit "r/worldnews" using the Reddit API in Python. Which is the correct code snippet from below options?

    a) import praw
    reddit = praw.Reddit(client_id='MY_CLIENT_ID', client_secret='MY_CLIENT_SECRET', user_agent='MY_USER_AGENT')
    posts = reddit.subreddit('worldnews').hot(limit=10)
    for post in posts:
    print(post.title)

    b) import requests
    url = "https://www.reddit.com/r/worldnews/new.json?limit=10"
    response = requests.get(url)
    data = response.json()
    for post in data['data']['children']:
    print(post['data']['title'])

c) import praw
reddit = praw.Reddit(client_id='MY_CLIENT_ID',
client_secret='MY_CLIENT_SECRET', user_agent='MY_USER_AGENT')
posts = reddit.subreddit('worldnews').new(limit=10)
for post in posts:
    print(post.title)
d) None of the above

5. Which of the following tweet content falls in category of linguistic features?
    a. Number of followers
    b. Presence of negative emotion words
    c. Presence of smiley
    d. Presence of pronouns

6. Why do we do network analysis on social media platforms?
    a) To analyze the performance of computer networks and improve data transmission speeds.
    b) To understand the social structures and relationships between individuals or entities in a network.
    c) To study the behavior of network protocols and optimize their efficiency.
    d) To identify and analyze patterns of diseases spreading within a population.

7. Which of the following is a recent example of misinformation spread on social media platforms?
    a) Tweet claiming that a particular brand of COVID-19 vaccine causes severe side effects, with no evidence to support the claim.
    b) Facebook post sharing information from a well-known health organization about the effectiveness of vaccination.
    c) Instagram story containing accurate information about a recent political development with credible sources cited.
    d) YouTube video providing a detailed explanation of climate change supported by scientific evidence.

8. Given a MongoDB collection named "employees" with the following document structure:

```
{
  "_id": ObjectId("615243d88f0ae43f255874c1"),
  "name": "Ram",
  "department": "CSE",
  "age": 35,
  "salary": 50000
```

}

What MongoDB query would you use to retrieve all employees with a salary greater than or equal to 60000?

a) db.employees.find({ "salary": { $gte: 60000 } })
b) db.employees.find({ "salary": { $gt: 60000 } })
c) db.employees.find({ "salary": { $lte: 60000 } })
d) db.employees.find({ "salary": { $lt: 60000 } })

9. You are developing a social media platform, and you have a database table, "posts" to store user posts. The "posts" table schema is given below:

```
CREATE TABLE posts (
    post_id INT PRIMARY KEY,
    user_id INT,
    post_content TEXT,
    post_date DATETIME
);
```

Now, you want to retrieve the 5 most recent posts made by a specific user with the user ID "12345". Which SQL query would you use?

a) SELECT * FROM posts WHERE user_id = 12345 ORDER BY post_date DESC LIMIT 5;

b) SELECT * FROM posts WHERE user_id = 12345 ORDER BY post_id DESC LIMIT 5;

c) SELECT * FROM posts WHERE user_id = 12345 ORDER BY post_date ASC LIMIT 5;

d) SELECT * FROM posts WHERE user_id = 12345 ORDER BY post_id ASC LIMIT 5;

10. What is the Facebook Graph API as discussed in lecture?

a) An API that provides access to a user's Facebook friends list.
b) An API that allows developers to create and manage Facebook pages.
c) An API that enables developers to integrate Facebook login into their applications for user authentication.
d) An API that provides access to a wide range of user data, interactions, and connections on Facebook.

Week-3 PSOSM 2023

1. Consider the below response:
   "To this segment, consumer privacy is very important; they feel that they have been victims of privacy invasions, they are pessimistic about the future of privacy protection, and about a third of them favour creating a general federal regulatory agency on consumer privacy."

   To which of the three Westin Privacy categories is "they" in the above sentence most likely to fall?
   a. Fundamentalists
   b. Pragmatists
   c. Unconcerned
   d. None of the above

2. Which of the following features is used by the credibility model- TweetCred?
   a. Tweet links
   b. Tweet Author
   c. Tweet network
   d. Tweet meta data

3. Cronbach's alpha
   a. determines the statistical significance of a research study.
   b. assess the reliability and internal consistency of a questionnaire.
   c. analyze the effect size of an intervention.
   d. calculate the standard deviation of a sample population.

4. Suppose we have five tweets (TW1, TW2, TW3, TW4, and TW5) with their relevances based on user feedback or relevance labels:

   TW1: Highly relevant tweet (Relevance = 2)
   TW2: Moderately relevant tweet (Relevance = 1)
   TW3: Not relevant tweet (Relevance = 0)
   TW4: Relevant tweet (Relevance = 1)
   TW5: Highly relevant tweet (Relevance = 2)

   Now, let's assume the ranking algorithm generates a ranked list of tweets:

   Rank 1: TW5 (Highly relevant tweet)
   Rank 2: TW1 (Highly relevant tweet)
   Rank 3: TW3 (Not relevant tweet)
   Rank 4: TW2 (Moderately relevant tweet)
   Rank 5: TW4 (Relevant tweet)

What would be NDCG@5 for this ranked list of tweets?

a. 0.5
b. .513
c. .975
d. None of the above

Solution:
Step 1: Calculate DCG@5 (Discounted Cumulative Gain at rank 5).
Step 2: Calculate IDCG@5 (Ideal Discounted Cumulative Gain at rank 5).
Step 3: Divide DCG@5 by IDCG@5 to get NDCG@5.

Step 1: Calculate DCG@5 DCG@5 = rel1 + rel2/log2(2) + rel3/log2(3) + rel4/log2(4) + rel5/log2(5)

Given relevances: rel1 = 2 (for TW5) rel2 = 2 (for TW1) rel3 = 0 (for TW3) rel4 = 1 (for TW2) rel5 = 1 (for TW4)

DCG@5 = 2 + 2/log2(2) + 0/log2(3) + 1/log2(4) + 1/log2(5) DCG@5 = 2 + 2/1 + 0 + 1/2 + 1/2
DCG@5 = 4 + 0.5 + 0.5 DCG@5 = 5

Step 2: Calculate IDCG@5 (Ideal DCG@5) To calculate IDCG@5, we consider the ideal ranking based on relevance:

Ideal ranking: TW5 (Highly relevant tweet) > TW1 (Highly relevant tweet) > TW2 (Moderately relevant tweet) > TW4 (Relevant tweet) > TW3 (Not relevant tweet)

IDCG@5 = rel1 + rel2/log2(2) + rel3/log2(3) + rel4/log2(4) + rel5/log2(5)
For ideal ranking: rel1 = 2 (for TW5) rel2 = 2 (for TW1) rel3 = 1 (for TW2) rel4 = 1 (for TW4) rel5 = 0 (for TW3)
IDCG@5 = 2 + 2/log2(2) + 1/log2(3) + 1/log2(4) + 0/log2(5) IDCG@5 = 2 + 2/1 + 1/1.585 + 1/2 + 0 IDCG@5 = 2 + 2 + 0.63 + 0.5 IDCG@5 = 5.13
Step 3: Calculate NDCG@5 NDCG@5 = DCG@5 / IDCG@5
NDCG@5 = 5 / 5.13
NDCG@5 ≈ 0.975 (rounded to three decimal places)
So, the NDCG@5 for the ranked list of tweets is approximately 0.975.

Read through the report “Privacy in India: Attitudes and Awareness V 2.0” at https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/research/privacyindia/PI_2012_Complete_Report.pdf And answer the following questions [5-7].

5. How many participants have changed their default settings on their social network?
   a. 75%
   b. 60%
   c. 50%
   d. 25%

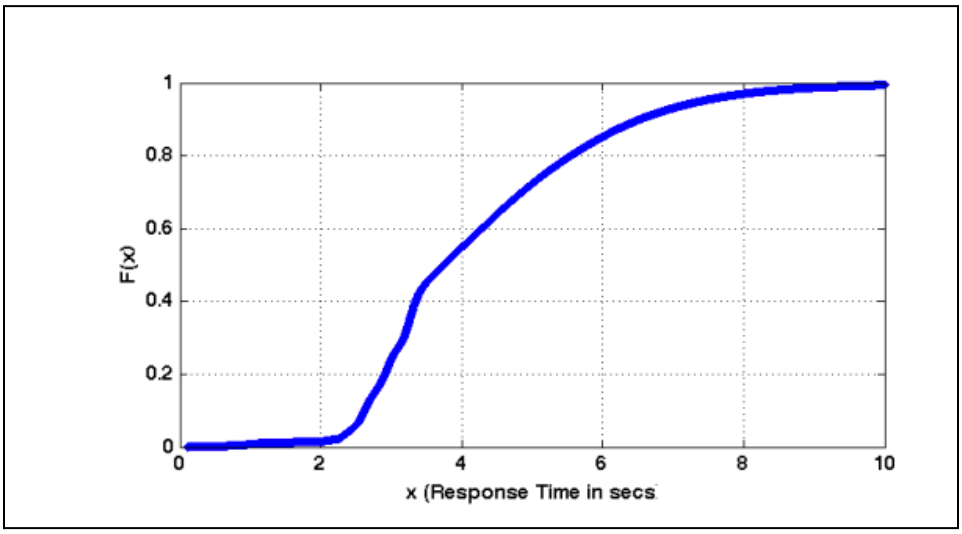
6. What is the primary purpose of a Focus Group Discussion (FGD) in the above report?

   a. To conduct one-on-one interviews with participants.
   b. To collect quantitative data through surveys and questionnaires.
   c. To observe participants' behaviours in a controlled environment.
   d. To facilitate group discussions and gather qualitative insights.

7. “When you hear the word privacy, what comes to mind”. Which is the most common respondent answer according to the above report?

   a. Information Privacy
   b. Bodily privacy
   c. Communication privacy
   d. Territorial privacy

8. Consider the below graph and choose the correct option:



a. For 82% of the users, the response time was less than 2 seconds, and for 99%, the response time was under 10 seconds.
b. For 99% of the users, the response time was less than 6 seconds, and for 100%, the response time was under 10 seconds.
c. For 82% of the users, the response time was less than 6 seconds, and for 99%, the response time was under 10 seconds.
d. None of the above

9. What is the Likert scale?

a. rating scale in surveys and questionnaires
b. gives the score for any real-time tweets
c. Interannotator agreement
d. None of the above

10. TweetCred is

a. a real-time web-based system to automatically evaluate the credibility of content on Twitter.
b. a real-time app-based system to automatically evaluate the score of tweets on Twitter.
c. a real-time app-based system to automatically assess the quality of content on Twitter.
d. None of the above

Week-4 PSOSM 2023

Question 1: What is the critical challenge that Latanya Sweeney's research on k-anonymity addresses?

a) make datasets larger for analysis
b) eliminate data from datasets
c) anonymise data without losing utility
d) ignore research utility for better privacy

Suppose you're working on a project that involves analysing sales data for a chain of stores. You are provided with a list of daily sales figures for a particular product over some time. The task is to calculate various statistics such as the total sales, average sales, and maximum sales for that product. Answer the following questions [2-4]:

Question 2: Which data structure can significantly enhance the efficiency and simplicity of your calculations?

a. Python numpy list
b. Python numpy array
c. Python variables
d. None of these

Question 3: Fill in the correct function calculate_total_sales(sales_list) which calculates the total sales from a list of daily sales figures using NumPy in the below code

```
import numpy as np

def calculate_total_sales(sales_list):
  <<<write your code here>>>>

    return total_sales

daily_sales = [1200, 1500, 1300, 1400, 1800, 1600]
total_sales = calculate_total_sales(daily_sales)
print("Total sales:", total_sales)
```

a. sales_array = np.array(sales_list)
   total_sales = np.sum(sales_array)
b. sales_array = np.list(sales_list)
   total_sales = np.sum(sales_array)
c. sales_array = np.array(sales_list)

total_sales = np.sum(np.concatenate(sales_array,sales_list))

d. None of the above

Question 4: Fill in the correct function calculate_maximum_sales(sales_list) which calculates the maximum sales from a list of daily sales figures using NumPy in the below code

```
import numpy as np

def calculate_maximum_sales(sales_list):
  <<<write your code here>>>>
   return max_sales

daily_sales = [1200, 1500, 1300, 1400, 1800, 1600]
maximum_sales = calculate_maximum_sales(daily_sales)
print("Maximum sales:", maximum_sales)
```

a. sales_array = np.array(sales_list)
   max_sales = np.max(sales_array)
b. sales_array = np.list(sales_list)
   max_sales = np.max(sales_array)
c. sales_array = np.array(sales_list)
   max_sales = np.max(np.concatenate(sales_array,sales_list))
d. None of the above

Imagine you have a NumPy array representing your monthly expenses for the past year. Each array element corresponds to the total expenses (in dollars) for a specific month. The task is to analyze and extract specific information.

```
import numpy as np

expenses_array = np.array([1200, 1500, 1300, 1400, 1800, 1600, 1400, 1250, 1350, 1500, 1700, 1900])
```

Answer the following questions from [5-7]:

Question 5: How would you extract expenses for the second half of the year (July to December)?

a. second_half_expenses = expenses_array[6:]
b. second_half_expenses = expenses_array[5:]
c. second_half_expenses = expenses_array[:12]
d. second_half_expenses = expenses_array[7:]

Question 6: Identify the months where expenses exceeded $1600.

a. high_expense_months = np.where(expenses_array < 1600)[0]
b. high_expense_months = np.where(expenses_array > 1600)[0]
c. high_expense_months = np.where(expenses_array > 1600))
d. high_expense_months = np.where(expenses_array > 1600)[0] + 1

Question 7: Calculate the average expenses for the first quarter of year.

a. first_quarter_average = np.mean(expenses_array[:4])
b. first_quarter_average = np.mean(expenses_array[:2])
c. first_quarter_average = np.mean(expenses_array[:3])
d. first_quarter_average = np.mean(expenses_array[:6])

Question 8: Choose the correct option for “result”

```
 import pandas as pd
data = {
   'Product': ['A', 'B', 'A', 'B', 'C', 'A', 'C', 'B', 'C', 'A'],
   'Category': ['Electronics', 'Clothing', 'Electronics', 'Clothing', 'Home', 'Electronics', 'Home',
'Clothing', 'Home', 'Electronics'],
   'Price': [500, 40, 600, 35, 100, 550, 80, 30, 90, 480]
}

sales_df = pd.DataFrame(data)
grouped = sales_df.groupby('Category')
result = grouped.agg({
   'Price': ['sum', 'mean']
}).reset_index()
# Rename the columns
result.columns = ['Category', 'Total Revenue', 'Average Price']
print(result)
```

a.

	Category	Total Revenue	Average Price
0	clothing	105	35.00
1	electronics	1630	543.33
2	home	270	90.00

b.

	Category	Total Revenue	Average Price
0	clothing	1630	35.00
1	electronics	105	543.33
2	home	270	90.00

c.

	Category	Total Revenue	Average Price
0	clothing	105	35.00
1	electronics	270	90.00
2	home	1630	543.33

d. None of the above

Question 9: What is the primary goal of k-anonymity in data privacy?

A) Ensure that every attribute in the dataset is unique.
B) Prevent unauthorised access to the dataset.
C) Minimize the amount of data collected.
D) Making it difficult to link specific individuals to their records.

Question 10: What will be the output of the below code:

```
import numpy as np
import matplotlib.pyplot as plt
time = np.linspace(0, 2 * np.pi, 1000)
amplitude = 1.0
frequency = 1.0
sine_wave = amplitude * np.sin(frequency * time)
plt.figure(figsize=(8, 6))
plt.plot(time, sine_wave, label='Sine Wave', color='blue')
plt.title('Sine Wave')
plt.xlabel('Time')
plt.ylabel('Amplitude')
```

```
plt.legend()
plt.grid(True)
plt.show()
```

a. A plot consisting of sine wave
b. A plot consisting of cosine wave
c. Logical error in code
d. None of the above

Solutions

1. Refer to lecture Week 4.1
2. In this example, by converting the sales data into a NumPy array, you can use the np.sum() function to easily calculate the total sales with just one line of code. NumPy's built-in functions are optimized for numerical computations and can provide better performance compared to traditional looping through a Python list. Converting the sales list to a NumPy array also opens up the possibility to perform various other mathematical operations (such as calculating the average, finding the maximum sales, etc.) efficiently and with concise code.

Question 3 to 8 and Q10 -please run the code

Q8: correct answer (code)

```
import pandas as pd
data = {
    'Product': ['A', 'B', 'A', 'B', 'C', 'A', 'C', 'B', 'C', 'A'],
    'Category': ['Electronics', 'Clothing', 'Electronics', 'Clothing', 'Home', 'Electronics', 'Home', 'Clothing', 'Home', 'Electronics'],
    'Price': [500, 40, 600, 35, 100, 550, 80, 30, 90, 480]
}

sales_df = pd.DataFrame(data)
grouped = sales_df.groupby('Category')
result = grouped.agg({
    'Price': ['sum', 'mean']
}).reset_index()
# Rename the columns
result.columns = ['Category', 'Total Revenue', 'Average Price']
print(result)
```

```
      Category  Total Revenue  Average Price
0     Clothing            105           35.0
1  Electronics           2130          532.5
2         Home            270           90.0
```

More cell actions

9. D) Making it difficult to link specific individuals to their records.

K-anonymity is a technique used in data privacy to protect individual identities within a dataset. The primary goal of k-anonymity is to make it difficult to identify specific individuals by ensuring that each record is indistinguishable from at least k-1 other records with respect to certain attributes.

Options A, B, and C are not accurate descriptions of the primary goal of k-anonymity. While ensuring uniqueness, preventing unauthorized access, and minimizing data collection are important considerations for data privacy, they are not the central focus of k-anonymity.

Week-5 PSOSM 2023

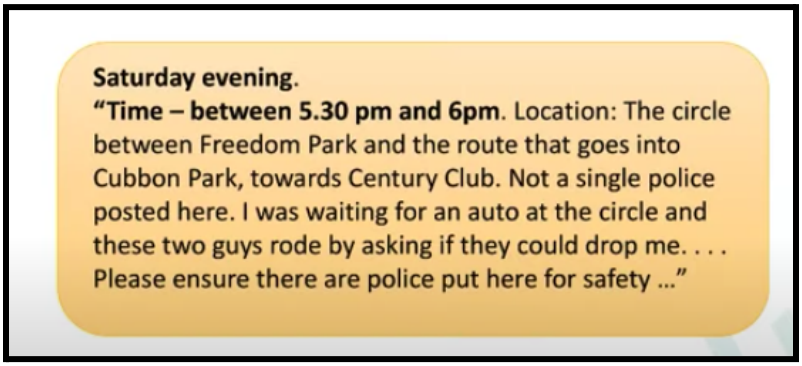
Question 1: How do Online Social Networks improve public Communication between police and residents?

a. provide actionable information and collective action
b. provide mutual accountability
c. understand fear and victimisation effects
d. help to contact concerned authorities

Question 2: Which of the following are examples of actionable information?

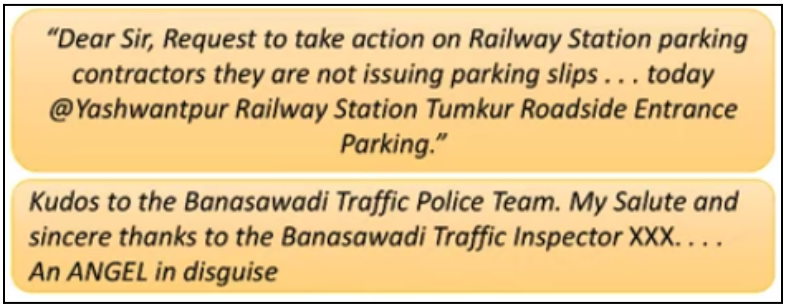
a. OSN post regarding traffic issues on the road
b. OSN post regarding open potholes on the road
c. OSN post sharing about Azaadi ka Amrit Mahotsav celebrations in the locality
d. None of the above

Question 3: What type of actionable information is highlighted in (bold) the below example?



a. Temporal data
b. Spatial data
c. Linguistic data
d. None of the above

Question 4: Given below are two types of communication as discussed in the lecture.



a. Formal and Formal
b. Informal and Formal
c. Informal and Informal
d. Formal and Informal

Question 5: How would you identify emotional and psychological states based on the written text?

a. LIWC
b. K-means clustering
c. N-gram analysis
d. All of the above

Question 6: Consider a sentence *"ISRO launched Chandrayaan and conducted experiments on thermophysical properties and environment"*. What is the thematic content and subject matter being discussed?

a. The linguistic diversity of the sentence.
b. The emotional tone is conveyed in the sentence.
c. The grammatical structure of the sentence.
d. The activities and research conducted by ISRO.

Question 7: What does an “n” represent in the n-gram analysis of a sentence?

a. The number of emotions in the analysis.
b. The number of words in each n-gram.
c. The number of characters in each n-gram.
d. The number of different languages in the sentence.

Question 8: A sentiment analysis study analyses a text passage for valence and arousal values. On a scale of -1 to 1, the valence and arousal score is 0.75 and 0.85. What do these values suggest about the emotional characteristics of the text?

a. Negative sentiment, low emotional intensity
b. Positive sentiment, High emotional intensity
c. Negative sentiment, high emotional intensity
d. Positive sentiment, low emotional intensity

Question 9: A social media analyst is studying the engagement characteristics (likes, comments, and shares) of three different posts on a platform. The engagement counts for each post over a week are as follows:

Post 1: 120 likes, 30 comments, 15 shares
Post 2: 80 likes, 45 comments, 20 shares

Post 3: 100 likes, 25 comments, 10 shares

Calculate the mean engagement (average likes, comments, and shares) for each post and then find the standard deviation of engagement for the entire set of posts. Based on the standard deviation, which post shows the highest variability in engagement characteristics?

A) Post 1
B) Post 2
C) Post 3
D) All posts show similar variability.

Question 10: What kind of post and sentiment is depicted in the below example?
“Absolutely disgusted by the state of our local park! It's been months since we reported the broken playground equipment, and nothing has been done. Our kids deserve better!”

a. Citizen-initiated and positive
b. Public authority-initiated and negative
c. Citizen-initiated and negative
d. Public authority-initiated and positive

Solutions:

Question 9:

Mean Engagement for each post:

Post 1: (120 + 30 + 15) / 3 = 55 engagements
Post 2: (80 + 45 + 20) / 3 = 48.33 engagements (rounded off)
Post 3: (100 + 25 + 10) / 3 = 45 engagements

Calculate the Standard Deviation of Engagement for all posts:
Calculate the mean engagement across all posts: (55 + 48.33 + 45) / 3 ≈ 49.44 engagements
Calculate the squared differences from the mean for each post:
Post 1: (55 - 49.78)^2 ≈ 27.07
Post 2: (48.33 - 49.78)^2 ≈ 2.11
Post 3: (45 - 49.78)^2 ≈ 22.98
Calculate the average of the squared differences: (27.07 + 2.11 + 22.98) / 3 ≈ 17.05

Calculate the square root of the average (which is the standard deviation): √17.05 ≈ 4.13

Based on the standard deviation values, Post 1 shows the highest variability in engagement characteristics among the three posts.

	Likes	comments	shares	Mean Engagement
Post 1[	120	30	15]	55
Post 2[	80	45	20]	48.33
Post 3 [	100	25	10]	45

Across all rows → Average of matrix= 49.78 → overall engagement
Across all rows → Standard deviation= 4.13

X

(https://swayam.gov.in) (https://swayam.gov.in/nc_details/NPTEL)

garvit.aggarwal.ug20@nsut.ac.in

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Privacy And Security In Online Social Media (course)

If already registered, click to check your payment status

Course outline

- **How does an NPTEL online course work? ()**
- **Prerequisite Assignment ()**
- **Welcome to PSOSM class ()**
- **Introduction to Social Media API ()**
- **Misinformation and Privacy ()**
- **Privacy and Pictures on**

Week 6: Assignment 6

The due date for submitting this assignment has passed.

Due on 2023-09-06, 23:59 IST.

Assignment submitted on 2023-09-06, 19:19 IST

1) Consider the following scenario on a social network ***1 point***

A is friends with B and C.
B is friends with A and D.
C is friends with A and D.
D is friends with B and C.

Calculate the betweenness centrality for A, B, C, and D. Identify the individual having the highest betweenness centrality.

- A
- B
- C
- All of them have the same betweenness centrality

Yes, the answer is correct.
Score: 1
Accepted Answers:
All of them have the same betweenness centrality

2) What does indegree centrality measure in the social media and network analysis context? ***1 point***

- The level of activity a user has on the platform.
- The number of outgoing connections a user has with other users.

Online Social Media ()

Policing and Social Media ()

E-crime and social media ()

- Week 6.1: eCrime on Online Social Media (unit?unit=47&lesson=48)
- Week 6.2 eCrime on Online Social Media (unit?unit=47&lesson=49)
- Tutorial 4 Social Network Analysis (unit?unit=47&lesson=50)
- **Quiz: Week 6: Assignment 6 (assessment?name=130)**
- Week 6 Feedback Form : Privacy and Security in Online Social Media (unit?unit=47&lesson=51)

Social media and ecrime ()

Identity resolution and social media ()

Research papers: Location

- (selected) A user's popularity based on their followers or incoming connections.
- The diversity of content shared by a user.

Yes, the answer is correct.
Score: 1
Accepted Answers:
A user's popularity based on their followers or incoming connections.

3) What does betweenness centrality depict about the node? ***1 point***

- One who is quickly approachable
- Who is close to everyone
- One who is most influential
- (selected) All of the above

No, the answer is incorrect.
Score: 0
Accepted Answers:
One who is quickly approachable

4) What is link farming? ***1 point***

- Agricultural practice that involves cultivating crops in a vertical arrangement.
- (selected) Rapidly growing links on a website through unethical or manipulative means.
- A strategy for developing social connections by attending networking events.
- All of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Rapidly growing links on a website through unethical or manipulative means.

5) Spamming the index of search engines is called ***1 point***

- (selected) Spamdexing
- (selected) Spamexing
- Spamming
- None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Spamdexing
Spamexing

6) Calculate the Klout score based on the following engagement metrics for a user on online social media. ***0 points***

Number of retweets: 50
Number of likes: 150
Number of comments: 100
Number of followers: 500
Number of mentions: 30

based Privacy ()

Research Papers Part - II ()

Week 11: Summary ()

Lecture materials/Notes ()

Text Transcripts ()

Download videos ()

Books ()

Problem Solving Session - July 2023 ()

- 230
- 330
- 430
- 530

Yes, the answer is correct.
Score: 0
Accepted Answers:
330

7) Which of the following is considered as “social etiquette” on online social media such as Twitter? ***1 point***

- Following back
- Comment on the post of an influential person
- Giving likes to the post of an influential person
- Saying “Thank you” to any appreciation by an influential person

Yes, the answer is correct.
Score: 1
Accepted Answers:
Following back

8) A psychological principle where individuals feel compelled to return favours and interactions they receive is called ***1 point***

- Spamming
- Reciprocity
- Link farming
- All of the above

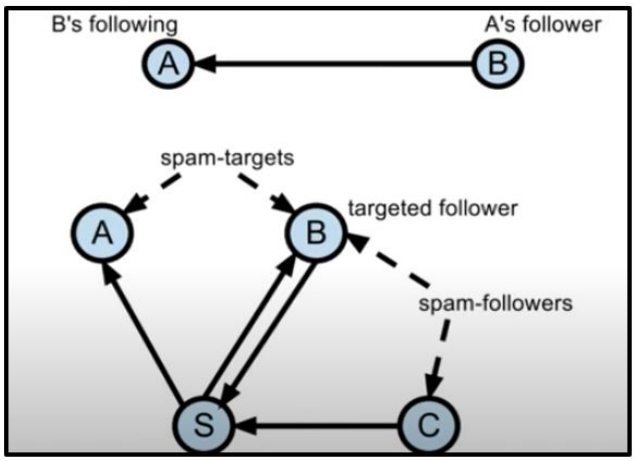
Yes, the answer is correct.
Score: 1
Accepted Answers:
Reciprocity

9) Which of the following is/are graph centrality measures? ***1 point***

- Eigen centrality
- Node degree
- Betweenness centrality
- Closeness centrality

Yes, the answer is correct.
Score: 1
Accepted Answers:
Eigen centrality
Node degree
Betweenness centrality
Closeness centrality

10) Which of the following are spam targets? ***1 point***



- [x] A
- [x] B
- [ ] C
- [ ] S

Yes, the answer is correct.
Score: 1

Accepted Answers:

A

B

X

(https://swayam.gov.in) (https://swayam.gov.in/nc_details/NPTEL)

garvit.aggarwal.ug20@nsut.ac.in

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Privacy And Security In Online Social Media (course)

If already registered, click to check your payment status

Course outline

- How does an NPTEL online course work? ()
- Prerequisite Assignment ()
- Welcome to PSOSM class ()
- Introduction to Social Media API ()
- Misinformation and Privacy ()
- Privacy and Pictures on

Week 7: Assignment 7

The due date for submitting this assignment has passed.

Due on 2023-09-13, 23:59 IST.

Assignment submitted on 2023-09-12, 00:35 IST

1) Consider the following statement: ***"Hello everyone!!!, Are you enjoying psosm course --- and learning new skills."*** Choose the correct code to convert it to unpunctuated string *1 point*

```
punctuations = '''!()-[]{};:'"\,<>/?@#$%^&*_~'''
no_punct = ""
for char in my_str:
   if char not in punctuations:
      no_punct = no_punct + char
```

```
punctuations = '''!()[]{};:'"\,<>./?@#$%^&*_~'''
no_punct = ""
for char in my_str:
   if char not in punctuations:
      no_punct = char
```

```
punctuations = '''!()-[]{};:'"\,<>./?@#$%^&*_~'''
no_punct = ""
for char in my_str:
   if char not in punctuations:
      no_punct = no_punct + char
```

Online Social Media ()

Policing and Social Media ()

E-crime and social media ()

Social media and ecrime ()

- Week-7.1: Link Farming in Online Social Media (unit?unit=52&lesson=53)
- Week-7.2: Nudges (unit?unit=52&lesson=54)
- Week-7.3: Semantic attacks: Spear phishing (unit?unit=52&lesson=55)
- Tutorial 5: Analyzing text using Python NLTK (unit?unit=52&lesson=56)
- **Quiz: Week 7: Assignment 7 (assessment?name=131)**
- Week 7 Feedback Form : Privacy and Security in Online Social Media (unit?unit=52&lesson=57)

Identity resolution

None of these

Yes, the answer is correct.
Score: 1

Accepted Answers:

```
punctuations = '''!()-[]{};:'"\,<>./?@#$%^&*_~'''
no_punct = ""
for char in my_str:
   if char not in punctuations:
      no_punct = no_punct + char
```

2) Consider the following code: ***1 point***

```
from collections import Counter
x = [2, 1, 3, 4, 5, 2]
token = ['social', 'media', 'nptel', 'nltk', 'python', 'tutorial']
pairs= Counter(dict(zip(token, x))).most_common(2)
```

Which of the following is the correct option?

- This code is used to retrieve an ascending sorted list having two tuples.
- This code is used to retrieve a descending sorted list having two tuples.
- This code retrieves two tuples in a sorted or unsorted manner.
- This code returns a logical error.

Yes, the answer is correct.
Score: 1

Accepted Answers:
This code is used to retrieve a descending sorted list having two tuples.

3) Which of the following are not stopwords in the following list? ***1 point***

stopwords = ["i", "me", "my", "myself", "we", "our", "ours", "ourselves", "you", "your", "yours", "yourself", "yourselves", "he", "him", "his", "himself", "she", "her", "hers", "herself", "it", "its", "itself", "they", "them", "their", "theirs", "themselves", "what", "which", "who", "whom", "this", "that", "these", "those", "am", "is", "are", "was", "were", "be", "been", "being", "have", "has", "had", "having", "do", "does", "did", "doing", “apple”, "a", "an", "the", "and", "but", "if", "or", "because", "as", "until", "while", "of", "at", "by", "for", "with", "about", "against", "between", "into", "through", "during", "before", "after", "above", "below", "to", "from", "up", "down", "in", "out", "on", "off", "over", "under", "again", "further", "then", "once", "here", "there", "when", "where", "why", "how", "all", "any", "both", "each", "few", "more", "most", "other", "some", "such", "no", "nor", "not", "only", "own", "same", "so", "than", "too", "very", "s", "t", "can", "will", "just", "don", "should", "now"]

- All of the above are stopwords.
- Further, again, apple
- While, apple
- apple

Yes, the answer is correct.
Score: 1

Accepted Answers:
apple

and social media ()

Research papers: Location based Privacy ()

Research Papers Part - II ()

Week 11: Summary ()

Lecture materials/Notes ()

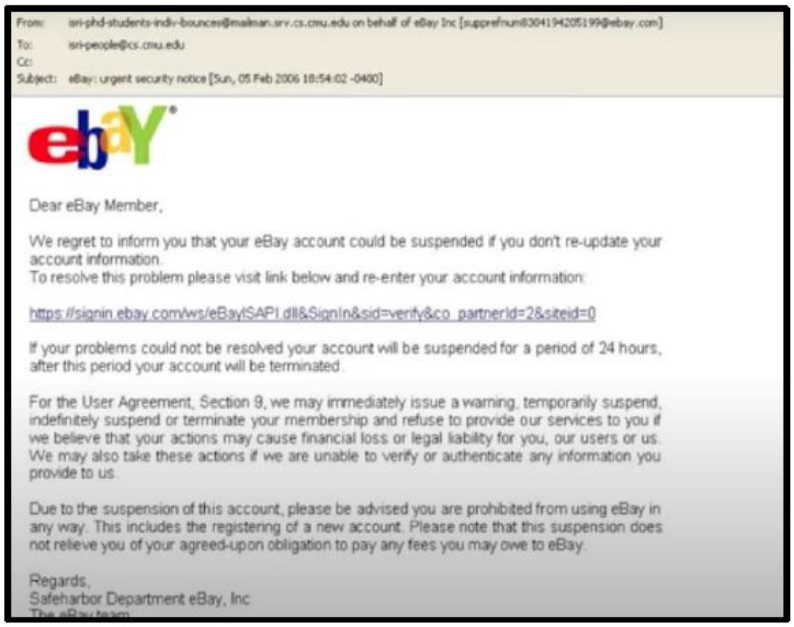
Text Transcripts ()

Download videos ()

Books ()

Problem Solving Session - July 2023 ()

4) Which elements in the below sample mail make it potentially phishing email, as discussed in the lecture? ***1 point***


From: isri-phd-students-indiv-bounces@mailman.srv.cs.cmu.edu on behalf of eBay Inc [supprefnum830419420519 9@ebay.com]
To: isri-people@cs.cmu.edu
Cc:
Subject: eBay: urgent security notice [Sun, 05 Feb 2006 18:54:02 -0400]

Dear eBay Member,

We regret to inform you that your eBay account could be suspended if you don't re-update your account information.
To resolve this problem please visit link below and re-enter your account information:

https://signin.ebay.com/ws/eBayISAPI.dll&SignIn&sid=verify&co_partnerId=2&siteid=0

If your problems could not be resolved your account will be suspended for a period of 24 hours, after this period your account will be terminated.

For the User Agreement, Section 9, we may immediately issue a warning, temporarily suspend, indefinitely suspend or terminate your membership and refuse to provide our services to you if we believe that your actions may cause financial loss or legal liability for you, our users or us. We may also take these actions if we are unable to verify or authenticate any information you provide to us.

Due to the suspension of this account, please be advised you are prohibited from using eBay in any way. This includes the registering of a new account. Please note that this suspension does not relieve you of your agreed-upon obligation to pay any fees you may owe to eBay.

Regards,
Safeharbor Department eBay, Inc

- Urgency in subject
- Regrettable information in the body
- Potential phishing link along with body
- All of the above

Yes, the answer is correct.
Score: 1

Accepted Answers:
All of the above

5) Which of the following is correct statement as discussed in the lecture? ***1 point***

- Younger participants were more vulnerable to phishing attacks.
- Older participants were more vulnerable to phishing attacks.
- Younger and older participants were equally vulnerable to phishing attacks.
- No one is vulnerable to phishing attacks.

Yes, the answer is correct.
Score: 1

Accepted Answers:
Younger participants were more vulnerable to phishing attacks.

6) In a scientific experiment, which group refers to manipulating the independent variable? ***1 point***

- Observational group
- Comparison group
- Control group
- Experimental group

Yes, the answer is correct.

Score: 1
Accepted Answers:
Experimental group

7) Which of the following is/are the advantages of having privacy policies? ***1 point***

- To help users make informed decisions with the information that is presented to them
- To help individuals avoid regrettable online disclosures
- Both a and b are correct
- None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Both a and b are correct

8) Which of the following nudges is depicted in the figure below? ***1 point***

- Timer nudge
- Picture nudge
- Sentiment nudge
- All of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Picture nudge

9) As discussed in the lecture, which of the following is/are not a social media nudge? ***1 point***

- Timer Nudge
- Picture Nudge
- Sentiment Nudge
- Voice Nudge

Yes, the answer is correct.
Score: 1
Accepted Answers:
Voice Nudge

10) Which of the following has the property of having high indegree and high outdegree on online social media? ***1 point***

- Top Link farmers

- Random social network users
- Top Spammers
- All of the above

Yes, the answer is correct.
Score: 1

Accepted Answers:
Top Link farmers

X

(https://swayam.gov.in) (https://swayam.gov.in/nc_details/NPTEL)

garvit.aggarwal.ug20@nsut.ac.in

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Privacy And Security In Online Social Media (course)

If already registered, click to check your payment status

Course outline

- **How does an NPTEL online course work? ()**
- **Prerequisite Assignment ()**
- **Welcome to PSOSM class ()**
- **Introduction to Social Media API ()**
- **Misinformation and Privacy ()**
- **Privacy and Pictures on**

Week 8 : Assignment 8

The due date for submitting this assignment has passed.

Due on 2023-09-20, 23:59 IST.

Assignment submitted on 2023-09-19, 01:04 IST

1) What are the different approaches to profile linking? ***1 point***

- [x] Study and analyse common attributes (place of work, profile picture, name, etc.) across different social media platforms
- [x] Check important attributes such as the same emails, locations, and contact numbers (if available) used across multiple platforms
- [ ] Only looking at the user posting behaviour on social media platforms will help in profile linking.
- [ ] It is not possible to find the same user profiles on different social media platforms.

Yes, the answer is correct.
Score: 1
Accepted Answers:
Study and analyse common attributes (place of work, profile picture, name, etc.) across different social media platforms
Check important attributes such as the same emails, locations, and contact numbers (if available) used across multiple platforms

2) Which of the following statements is/are true regarding username change on social media? ***1 point***

S1: Changes to any profile attribute other than username do not lead to unreachability to the user profile.
S2: The potential benign reasons for username change include space gain, suiting a trending event, gaining/losing anonymity, adjusting to real-life events, and avoiding boredom.

Online Social Media ()

Policing and Social Media ()

E-crime and social media ()

Social media and ecrime ()

Identity resolution and social media ()

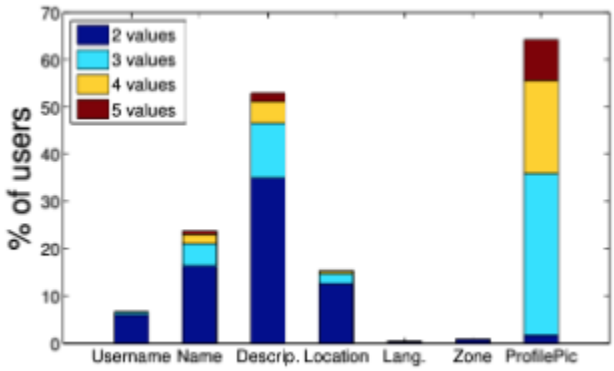
- Week 8.1: Profile Linking on Online Social Media (unit?unit=58&lesson=59)
- Week 8.2: Anonymous Networks (unit?unit=58&lesson=60)
- Tutorial 6: Gephi Network Visualization (unit?unit=58&lesson=61)
- **Quiz: Week 8 : Assignment 8 (assessment?name=133)**
- Week 8 Feedback Form : Privacy and Security in Online Social Media (unit?unit=58&lesson=62)

S3: The malicious intentions for username change may include obscured username promotion and username squatting.
S4: Search by the user's old username results in a failed attempt to reach the user's profile, potentially making others falsely believe that the user account has been deactivated.

- Only S1 and S4 are correct
- Only S1, S2, and S3 are correct
- S1, S2, and S4 are correct
- All S1, S2, S3, S4 are correct.

No, the answer is incorrect.
Score: 0
Accepted Answers:
All S1, S2, S3, S4 are correct.

3) Below is the graph of username change. Which of the following is correct inference ***0 points*** from the graph (as discussed in the lecture)?



- Around 35 per cent, of the users change their profile picture at least 3 times.
- Around 60-70 percent of the users change their usernames atleast once.
- Around 35 percent of the users change their profile picture atmost 3 times.
- Around 50-55 per cent of users have changed their descriptions thrice.

Yes, the answer is correct.
Score: 0
Accepted Answers:
Around 35 per cent, of the users change their profile picture at least 3 times.

4) Why do whisper users have a low clustering coefficient over the network? ***1 point***

- Whisper users are likely to interact with complete strangers who are highly unlikely to interact with each other.
- Users interact with a large sample of other users
- New users make 20% of the contribution in the whisper content.
- All of the above

No, the answer is incorrect.
Score: 0
Accepted Answers:
Whisper users are likely to interact with complete strangers who are highly unlikely to interact with each other.

Research papers: Location based Privacy ()

Research Papers Part - II ()

Week 11: Summary ()

Lecture materials/Notes ()

Text Transcripts ()

Download videos ()

Books ()

Problem Solving Session - July 2023 ()

5) What is the primary purpose of Gephi software? ***1 point***

- ○ Video Editing
- ○ 3D Modeling
- ◉ Network Visualization and Analysis
- ○ Music Composition

Yes, the answer is correct.
Score: 1
Accepted Answers:
Network Visualization and Analysis

6) Why studying username change behaviour is/are useful for social media researchers? ***1 point***

- ☑ This is one of the unique attributes of the user
- ☑ It is usually homogeneous
- ☑ The number of characters and length of the username is restricted
- ☑ It is the publicly available attribute

Yes, the answer is correct.
Score: 1
Accepted Answers:
This is one of the unique attributes of the user
It is usually homogeneous
The number of characters and length of the username is restricted
It is the publicly available attribute

7) In the network graph of Twitter, where nodes are users, the in-degree of the node would be ***1 point***

- ◉ Number of followers on Twitter
- ○ Number of people you follow
- ○ Both a and b
- ○ None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Number of followers on Twitter

8) Which of the following is/are anonymous social networks? ***0 points***

- ☑ Secret
- ☑ Wickr
- ☑ Yak yak
- ☐ Blind

Partially Correct.
Score: 0
Accepted Answers:
Secret
Wickr

Yak yak
Blind

9) A scale-free network can be one having ***1 point***

- (●) Power law degree distribution
- ( ) Zipf's law
- ( ) Pareto principle
- ( ) All of the above

No, the answer is incorrect.
Score: 0
Accepted Answers:
All of the above

10) Consider the below situation ***1 point***

Imagine you're a student in a university, and you decide to survey to determine the number of friends each of your classmates has on social media. You ask everyone to report the number of friends they have on one popular social media platform (say Facebook).

When you collect the data and calculate the average number of friends reported by your classmates, 300 friends. However, when you individually examine the number of friends reported by each of your classmates, you notice that many of them have fewer than 300 friends, perhaps ranging from 50 to 200 friends.

Why did you observe such a phenomenon?

- ( ) Most of the individuals have reported it wrong in the report.
- (●) Few individuals have exceptionally high friend counts (outliers) that skew the average upward.
- ( ) There is an error in individual calculations. It is impossible to have 300 friends on average if many individuals have fewer than 300 friends.
- ( ) None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Few individuals have exceptionally high friend counts (outliers) that skew the average upward.

X

If already registered, click to check your payment status

Course outline

How does an NPTEL online course work? ()

Prerequisite Assignment ()

Welcome to PSOSM class ()

Introduction to Social Media API ()

Misinformation and Privacy ()

Privacy and Pictures on

Week 9 : Assignment 9

The due date for submitting this assignment has passed.

Due on 2023-09-27, 23:59 IST.

Assignment submitted on 2023-09-27, 23:30 IST

1) Which of the following is/are location-based services? ***1 point***

- [x] Foursquare
- [x] Yelp
- [x] Gowalla
- [x] Uber

Yes, the answer is correct.
Score: 1
Accepted Answers:
Foursquare
Yelp
Gowalla
Uber

2) What is the difference between Pearson and Spearman rank correlation? ***1 point***

Hint: https://en.wikipedia.org/wiki/Pearson_correlation_coefficient (https://en.wikipedia.org/wiki/Pearson_correlation_coefficient)

- (•) Pearson correlation assesses linear relationships, while Spearman correlation evaluates monotonic relationships.
- ( ) Spearman correlation is always between -1 and 1, while Pearson correlation can have values outside this range.
- ( ) Spearman correlation is calculated using a formula involving the sum of squared differences, whereas Pearson correlation is based on the product of z-scores.

Online Social Media ()

Policing and Social Media ()

E-crime and social media ()

Social media and ecrime ()

Identity resolution and social media ()

Research papers: Location based Privacy ()

- Week 9.1: Privacy in Location Based Social Networks Part 1 (unit?unit=63&lesson=64)
- Week 9.2: Privacy in Location Based Social Networks Part 2 (unit?unit=63&lesson=65)
- Tutorial 7: Visualization - Highcharts (unit?unit=63&lesson=66)
- **Quiz: Week 9 : Assignment 9 (assessment?name=139)**

All of the above

Yes, the answer is correct.
Score: 1

Accepted Answers:
Pearson correlation assesses linear relationships, while Spearman correlation evaluates monotonic relationships.

3) You have two datasets and want to measure the strength and direction of the relationship between them. One set of data consists of the ranks or ordinal values of the observations, while the other set consists of continuous numerical data. Which correlation coefficient is most appropriate for this scenario? ***1 point***

- Spearman's rank correlation coefficient
- Pearson's correlation coefficient
- Either Spearman or Pearson correlation can be used interchangeably.
- None of the above because correlation coefficients are not applicable in this scenario.

Yes, the answer is correct.
Score: 1

Accepted Answers:
Spearman's rank correlation coefficient

4) Which of the following are open text fields, whose validity is not enforced by the system and may carry noise/invalid locations? ***1 point***

- [x] Venue location
- [ ] tips
- [x] User home city
- [ ] dones

Yes, the answer is correct.
Score: 1

Accepted Answers:
Venue location
User home city

Week 9 Feedback Form : Privacy and Security in Online Social Media (unit?unit=63&lesson=67)

Research Papers Part - II ()

Week 11: Summary ()

Lecture materials/Notes ()

Text Transcripts ()

Download videos ()

Books ()

Problem Solving Session - July 2023 ()

5) Which of the following is correct inference from Figure 2? ***1 point***

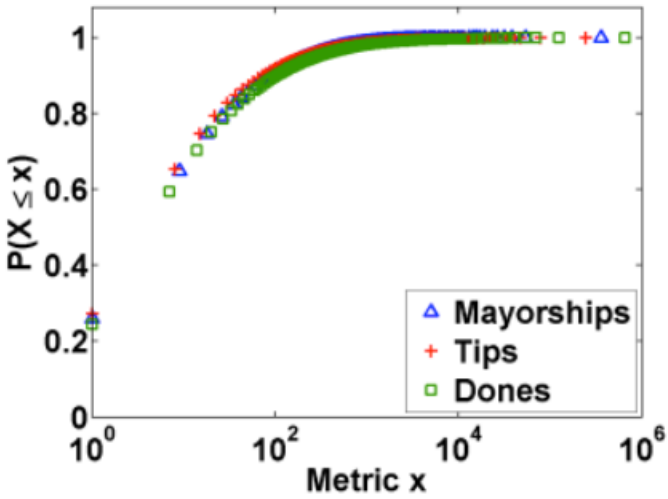


Figure 2. Cumulative Distribution of the Number of Mayorships, Tips and Dones per City.

- The distributions are very skewed, with a few cities having as many as 100 mayorships, tips or dones.
- The distributions are evenly balanced, with cities having a nearly equal number of mayorships, tips, or dones.
- The distributions are normally distributed, with no cities having an exceptionally high number of mayorships, tips, or dones.
- None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
The distributions are very skewed, with a few cities having as many as 100 mayorships, tips or dones.

6) What is the key difference between a scatter plot and a bar chart? ***1 point***

- Scatter plots display categorical data, while bar charts are used for numerical data.
- Scatter plots show the relationship between two numerical variables, while bar charts display the frequency or distribution of categorical data.
- Scatter plots are only used for displaying data with a single variable, while bar charts can visualize data with multiple variables.
- All of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Scatter plots show the relationship between two numerical variables, while bar charts display the frequency or distribution of categorical data.

7) Which of the following is publicly available data on Foursquare? ***1 point***

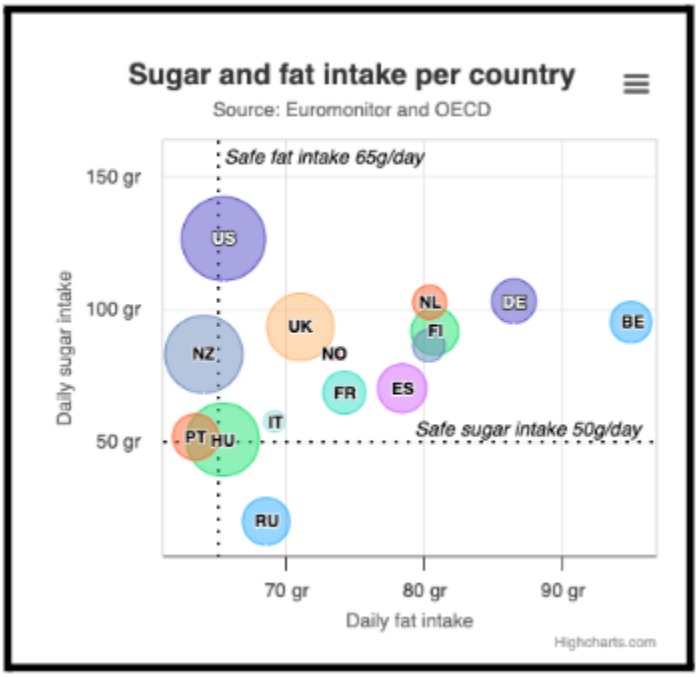
- Mayorships
- Tips

- [x] Dones
- [ ] Checkins

Yes, the answer is correct.
Score: 1

Accepted Answers:
Mayorships
Tips
Dones

8) Which plot is shown in the below figure? ***1 point***



- [ ] Scatter plot
- [ ] Bar chart
- [x] Bubble plot
- [ ] Histogram

Yes, the answer is correct.
Score: 1

Accepted Answers:
Bubble plot

9) What is the primary characteristic of Location-Based Social Networks (LBSNs)? ***1 point***

- [ ] They allow users to share text-based posts with their friends and followers.
- [ ] They provide a platform for online gaming and virtual reality experiences.
- [x] They enable users to connect with others and share their physical locations and activities.
- [ ] They focus exclusively on professional networking and job searching.

Yes, the answer is correct.
Score: 1

Accepted Answers:
They enable users to connect with others and share their physical locations and activities.

10) As discussed in lecture, Choose the correct inference (s) for figures below (Figure 4 and Figure 5). Here mayorships (blue), Tips (red), and Dones (green). ***1 point***

Reference: https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/Publications_files/TP_lbsn_2012.pdf (https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/Publications_files/TP_lbsn_2012.pdf)

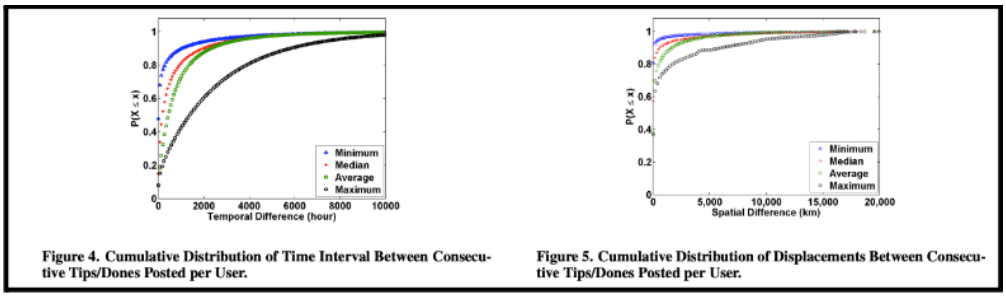


Figure 4. Cumulative Distribution of Time Interval Between Consecutive Tips/Dones Posted per User.

Figure 5. Cumulative Distribution of Displacements Between Consecutive Tips/Dones Posted per User.

- [x] In Figure 4, The distribution of minimum inter-activity times is very skewed towards short periods of time, with almost 50% of the users posting consecutive tips/dones 1 hour apart.
- [x] In Figure 4, On average, median and maximum, users do tend to experience very long periods of time between consecutive tips and dones. For instance, around 50% of the users have an average interactivity time of at least 450 hours, whereas around 80% of the users have a maximum inter-activity time above 167 hours (roughly a week).
- [x] In Figure 5, Around 36% of the users have average and maximum displacements of 0 kilometer, indicating very short distances (within a few metres)
- [x] In Figure 5, About 10% of the users have a maximum displacement of at least 6,000 kilometers.

Yes, the answer is correct.
Score: 1

Accepted Answers:
In Figure 4, The distribution of minimum inter-activity times is very skewed towards short periods of time, with almost 50% of the users posting consecutive tips/dones 1 hour apart.
In Figure 4, On average, median and maximum, users do tend to experience very long periods of time between consecutive tips and dones. For instance, around 50% of the users have an average interactivity time of at least 450 hours, whereas around 80% of the users have a maximum inter-activity time above 167 hours (roughly a week).
In Figure 5, Around 36% of the users have average and maximum displacements of 0 kilometer, indicating very short distances (within a few metres)
In Figure 5, About 10% of the users have a maximum displacement of at least 6,000 kilometers.

X

(https://swayam.gov.in) (https://swayam.gov.in/nc_details/NPTEL)

garvit.aggarwal.ug20@nsut.ac.in

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Privacy And Security In Online Social Media (course)

If already registered, click to check your payment status

Course outline

- **How does an NPTEL online course work? ()**
- **Prerequisite Assignment ()**
- **Welcome to PSOSM class ()**
- **Introduction to Social Media API ()**
- **Misinformation and Privacy ()**
- **Privacy and Pictures on Online Social Media ()**

Week 10 : Assignment 10

The due date for submitting this assignment has passed.

Due on 2023-10-04, 23:59 IST.

Assignment submitted on 2023-10-01, 20:25 IST

Please go through the below paper: “Beware of What You Share Inferring Home Location in Social Networks” at

https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/Publications_files/TP_GM_MV_AG_JA_PK_VA_Pinsoda_2012.pdf (https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/Publications_files/TP_GM_MV_AG_JA_PK_VA_Pinsoda_2012.pd

Answer the following questions from 1-3.

1) Which statement(s) is/are true? ***1 point***

- [x] The authors perform a large-scale inference study in three popular social networks: Foursquare, Google+ and Twitter.
- [x] The paper is looking at different social networks not only Foursquare.
- [x] The dataset used in the paper comprises many attributes including venues where the location must be defined filling the open text fields, (limited in 30 and 127 characters, respectively), and setting a pin in the map.
- [ ] None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
The authors perform a large-scale inference study in three popular social networks: Foursquare, Google+ and Twitter.
The paper is looking at different social networks not only Foursquare.
The dataset used in the paper comprises many attributes including venues where the location must be defined filling the open text fields, (limited in 30 and 127 characters, respectively), and setting a pin in the map.

Policing and Social Media ()

E-crime and social media ()

Social media and ecrime ()

Identity resolution and social media ()

Research papers: Location based Privacy ()

Research Papers Part - II ()

- Week 10.1: Beware of What You Share Inferring Home Location in Social Networks (unit?unit=68&lesson=69)
- Week 10.2: On the dynamics of username change behavior on Twitter (unit?unit=68&lesson=70)
- Week 10.3: Boston Marathon Analyzing Fake Content on Twitter (unit?unit=68&lesson=71)
- **Quiz: Week 10 : Assignment 10**

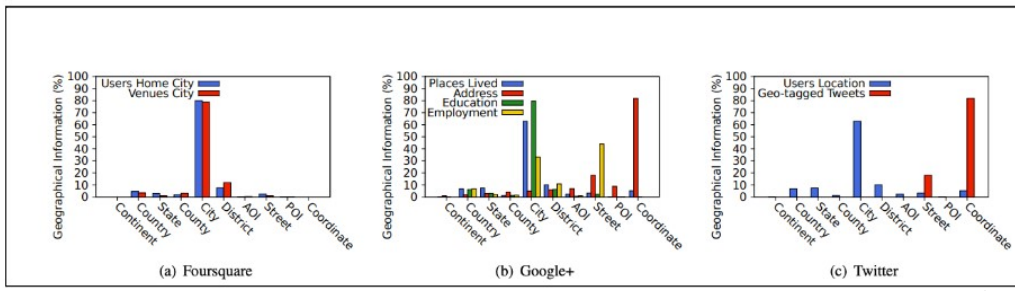
2) What is a valid AGI? ***1 point***

- ○ Valid Artificial General intelligence
- ◉ Valid Ambigious Geographical Information
- ○ Valid Artificial Geographical Interface
- ○ None of the above

Yes, the answer is correct.
Score: 1

Accepted Answers:
Valid Ambigious Geographical Information

3) Given the following figure depicts the quality of valid UGI, what is the correct inference(s) from the figure ? ***1 point***



- ☑ In figure (a), the vast majority (80%) of Foursquare users and venues have location information at the city level.
- ☐ In figure (b), majority of users provide home location information at country level
- ☐ In figure (c), 60% of the users have provided their location information at the district level
- ☑ In figure (b), 80% of the users have provided location information at city level.

Yes, the answer is correct.
Score: 1

Accepted Answers:
In figure (a), the vast majority (80%) of Foursquare users and venues have location information at the city level.
In figure (b), 80% of the users have provided location information at city level.

Please go through the following paper, “On the dynamics of username change behavior on Twitter”

https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/Publications_files/04-Jain.pdf (https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/Publications_files/04-Jain.pdf)

Answer the following questions from 4-6

4) From the overall conclusion of the paper, which of the statements is true ***0 points***

- ○ The set of people who change their handle many times is slightly larger than those who change their handle a very small number of times.
- ◉ The set of people who change their handle many times is slightly smaller than the set of people who change their handle very less number of times.
- ○ The set of people who change their handle many times is roughly equal to the set of people who do not change their handle.

(assessment?name=140)

Week 10 Feedback Form : Privacy and Security in Online Social Media (unit?unit=68&lesson=72)

Week 11: Summary ()

Lecture materials/Notes ()

Text Transcripts ()

Download videos ()

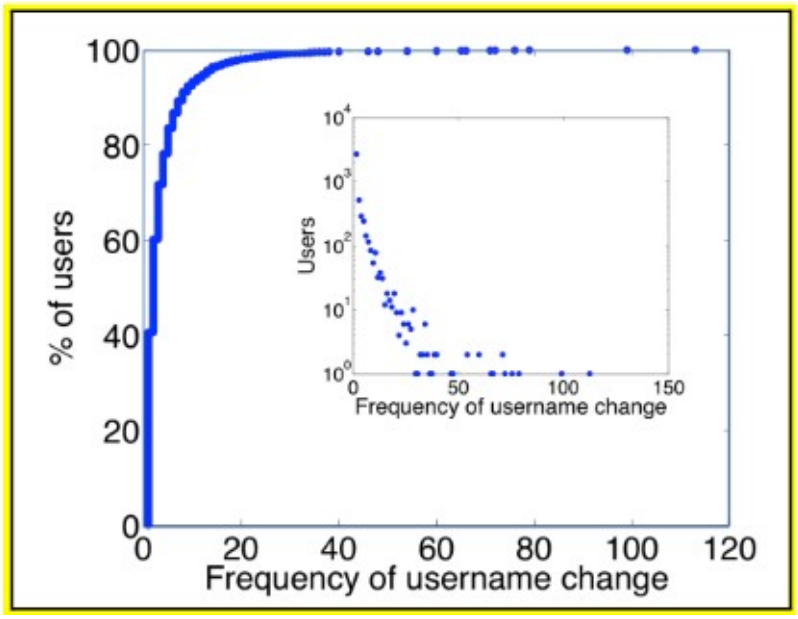
Books ()

Problem Solving Session - July 2023 ()

None of the above

Yes, the answer is correct.
Score: 0
Accepted Answers:
The set of people who change their handle many times is slightly smaller than the set of people who change their handle very less number of times.

5) From the following figure, what is the correct inference? ***1 point***



20% of users rarely change user names in short intervals and 80% change user names frequently after longer intervals.

80% of users rarely change user names in short intervals, and 20% change user names frequently after longer intervals.

35% of users choose a user name unrelated to an old one, and 65% of users reuse an old user name.

20% of users frequently change user names in short intervals, and 80% change user name rarely after longer intervals.

Yes, the answer is correct.
Score: 1
Accepted Answers:
20% of users frequently change user names in short intervals, and 80% change user name rarely after longer intervals.

6) What is the correct observation in the paper? ***1 point***

This paper aims at finding how and why users change their usernames within a social network like Twitter.

Most users created new usernames unrelated to the old username when they changed usernames within a social network.

The authors believe that unrelated usernames over time could be credited to the absence of cognitive load to remember a past dumped username.

All of the above

Yes, the answer is correct.

Score: 1
Accepted Answers:
All of the above

Please go through the following paper, “Boston Marathon Analyzing Fake Content on Twitter”

https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/Publications_files/ecrs2013_ag_hl_pk.pdf
(https://cdn.iiit.ac.in/cdn/precog.iiit.ac.in/Publications_files/ecrs2013_ag_hl_pk.pdf)

Answer the following questions from 7-10

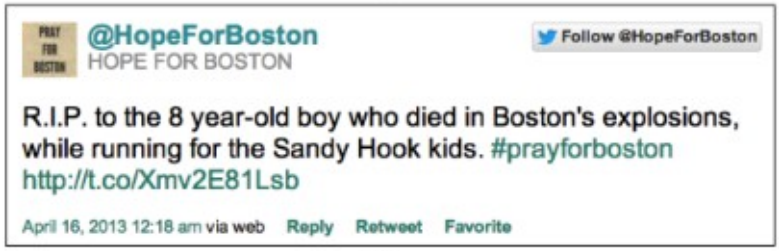
7) Which of the following statements is true about Boston blasts? ***1 point***

- Twin blasts occurred during the Boston Marathon on April 15th, 2013 at 18:50 GMT
- Twin blasts occurred during the Boston Marathon on April 15th, 2009 at 18:50 GMT
- Four people were killed and 264 were injured in the incident
- All of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Twin blasts occurred during the Boston Marathon on April 15th, 2013 at 18:50 GMT

8) What can be depicted from the tweets shared below? ***1 point***

(a)



(b)

- Both tweets contain Fake content.
- Figure (a) depicts a tweet from a fake charity profile.
- Figure (b) depicts a rumour about a child being killed in the blasts.
- All of the above

Yes, the answer is correct.

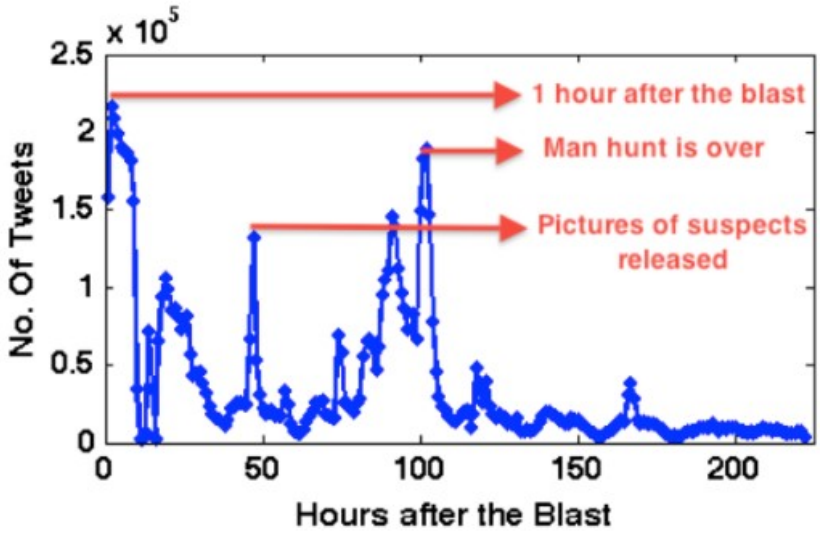
Score: 1
Accepted Answers:
All of the above

9) What are the major contributions of the figure? ***1 point***

- [x] The authors characterised the spread of fake content on Twitter using temporal, source and user attributes.
- [x] The authors used linear regression to predict how viral a rumour would be in the future based on its current user characteristics.
- [x] The authors analysed the activity and interaction graphs for the suspended user profiles created during the Boston blasts.
- [ ] None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
The authors characterised the spread of fake content on Twitter using temporal, source and user attributes.
The authors used linear regression to predict how viral a rumour would be in the future based on its current user characteristics.
The authors analysed the activity and interaction graphs for the suspended user profiles created during the Boston blasts.

10) What does the following figure depict? ***1 point***



- [x] The figure shows the temporal distribution of tweets after the Boston blast
- [ ] The figure shows the spatial distribution of tweets after the Boston blast
- [ ] Both of the above
- [ ] None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
The figure shows the temporal distribution of tweets after the Boston blast

X

(https://swayam.gov.in) (https://swayam.gov.in/nc_details/NPTEL)

garvit.aggarwal.ug20@nsut.ac.in

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Privacy And Security In Online Social Media (course)

If already registered, click to check your payment status

Course outline

- **How does an NPTEL online course work? ()**
- **Prerequisite Assignment ()**
- **Welcome to PSOSM class ()**
- **Introduction to Social Media API ()**
- **Misinformation and Privacy ()**
- **Privacy and Pictures on**

Week 11 : Assignment 11

The due date for submitting this assignment has passed.

Due on 2023-10-11, 23:59 IST.

Assignment submitted on 2023-10-11, 04:36 IST

1) Given below the statement. The statement is “An API is a set of rules and protocols that allows different software applications to communicate with each other, while a programming language is a formal language used to write instructions for a computer to perform specific tasks.” ***1 point***

- (●) True
- ( ) False
- ( ) Can’t say

Yes, the answer is correct.
Score: 1

Accepted Answers:
True

2) Which of the following features is supposed to be a part of tweet content on Twitter? ***1 point***

- [x] Presence of pronouns
- [ ] No. of followers
- [ ] No. of retweets
- [x] Mention of self words (I;my;mine)

Yes, the answer is correct.
Score: 1

Accepted Answers:
Presence of pronouns
Mention of self words (I;my;mine)

Online Social Media ()

Policing and Social Media ()

E-crime and social media ()

Social media and ecrime ()

Identity resolution and social media ()

Research papers: Location based Privacy ()

Research Papers Part - II ()

Week 11: Summary ()

- Week 11: Summary (unit?unit=73&lesson=74)
- **Quiz: Week 11 : Assignment 11 (assessment?name=141)**
- Week 11 Feedback Form : Privacy and Security in Online Social Media (unit?unit=73&lesson=75)

Lecture materials/Not

3) Who among the following are most likely to engage in link farming? ***1 point***

- [x] Highly active users
- [x] Legitimate users
- [x] Popular users
- [x] Bloggers and experts

Yes, the answer is correct.
Score: 1

Accepted Answers:
Highly active users
Legitimate users
Popular users
Bloggers and experts

4) Which of the following is username creation behaviour? ***1 point***

- [x] Static behavior patterns
- [ ] Occasional reuse patterns
- [x] Temporal behavior patterns
- [ ] Frequent reuse patterns

Yes, the answer is correct.
Score: 1

Accepted Answers:
Static behavior patterns
Temporal behavior patterns

5) Which of the following is a username reuse behaviour? ***1 point***

- [ ] Static behavior patterns
- [x] Occasional reuse patterns
- [ ] Temporal behavior patterns
- [x] Frequent reuse patterns

Yes, the answer is correct.
Score: 1

Accepted Answers:
Occasional reuse patterns
Frequent reuse patterns

6) Why is it called a "Graph" API? ***1 point***

- ( ) it allows users to draw and create graphs and charts.
- (•) it represents social connections and relationships as a graph data structure.
- ( ) it primarily deals with geographical mapping and location-based services.
- ( ) it is used exclusively for graph theory and mathematical computations.

Yes, the answer is correct.
Score: 1

Accepted Answers:
it represents social connections and relationships as a graph data structure.

es ()

Text Transcripts ()

Download videos ()

Books ()

Problem Solving Session - July 2023 ()

7) Latanya Sweeney's research highlighted how two seemingly unrelated pieces of information can be linked to reidentify an individual. Which term best describes this concept? ***1 point***

- Data Anonymization
- Data Aggregation
- Reidentification
- All of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Reidentification

8) What is the primary purpose of an API key on Twitter? ***1 point***

- To limit the number of characters in a tweet.
- To verify a user's identity on the platform.
- To authenticate and access Twitter's APIs for development purposes.
- To encrypt and secure direct messages on Twitter.

Yes, the answer is correct.
Score: 1
Accepted Answers:
To authenticate and access Twitter's APIs for development purposes.

9) What is the primary purpose of link farming in the context of SEO (Search Engine Optimization)? ***1 point***

- To create high-quality, relevant backlinks to improve a website's search engine ranking.
- To increase organic traffic by optimising on-page content and meta tags.
- To manipulate search engine algorithms by generating a large number of low-quality backlinks.
- To enhance the user experience by improving website navigation and design.

Yes, the answer is correct.
Score: 1
Accepted Answers:
To manipulate search engine algorithms by generating a large number of low-quality backlinks.

10) Which V of social media represents “Viral trends that spread rapidly across social media platforms.” ***1 point***

- Value
- Velocity
- Volume
- Veracity

Yes, the answer is correct.
Score: 1
Accepted Answers:
Velocity