# Data Privacy Practicals

**Question 1 :** Write a program to perform encryption and decryption using Caesar cipher (substitutional cipher)

**Solution :**

**Time Complexity Analysis:**
1. Encryption/Decryption: O(n), where n is the length of the input text. Each character is processed exactly once.

**Space Complexity Analysis:**
1. Space Complexity: O(n), as the result string requires storage proportional to the length of the input.

**Output :**

```
Caesar Cipher: Encryption and Decryption
Choose an option (encrypt/decrypt): encrypt
Enter the text: JiyaSharma
Enter the shift value: 4
Encrypted text: NmceWlevqe
```

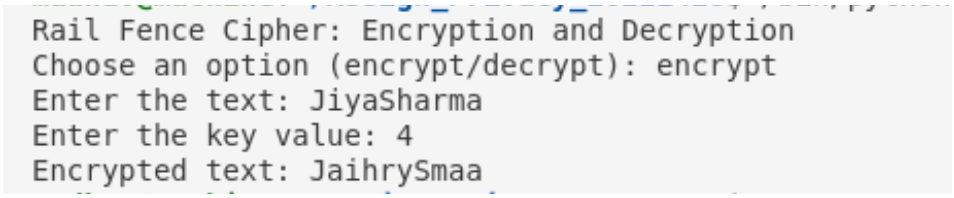
**Question 2 :** Write a program to perform encryption and decryption using Rail Fence Cipher (transpositional cipher)

**Solution :**

**Time Complexity Analysis:**
1. Encryption/Decryption: O(n), where n is the length of the input text. The algorithm iterates over each character in the text.

**Space Complexity Analysis:**
1. Space Complexity: O(n), due to the storage required for the 2D rail matrix.

**Output :**



**Question 3 :** Write a Python program that defines a function and takes a password string as input and returns its SHA-256 hashed representation as a hexadecimal string.

**Solution :**

**Time Complexity Analysis:**
1. Hashing the Password: O(n), where n is the length of the password. The hash function processes each byte of the input password.

**Space Complexity Analysis:**
1. Space Complexity: O(1), as the hash object itself uses a constant amount of memory irrespective of the password length.

**Output :**

```
Password Hashing with SHA-256
Enter the password: Queensharma
SHA-256 Hashed Password: e334031d113014e61448af33d63adb2eefb952490a92890449e81c32984aa775
```

**Question 4 :** Write a Python program that reads a file containing a list of usernames and passwords, one pair per line (separated by a comma). It checks each password to see if it has been leaked in a data breach. You can use the "Have I Been Pwned" API (https://haveibeenpwned.com/API/v3) to check if a password has been leaked.

**Solution :**

**Time Complexity Analysis:**
Overall Time Complexity: O(n + m).

**Space Complexity Analysis:**
1. Space Complexity: O(1)

**Dummy Test File :**

password_file.txt

```
user1,password123
user2,helloworld
user3,mightyraju
user4,goodnight
user5,pikachu
```

**Output :**

```
The password for user 'user1' has been pwned 294857 times!
The password for user 'user2' has been pwned 30566 times!
The password for user 'user3' has been pwned 17 times!
The password for user 'user4' has been pwned 9521 times!
The password for user 'user5' has been pwned 89312 times!
```

**Question 5 :** Write a Python program that generates a password using a random combination of words from a dictionary file..

**Solution :**

**Time Complexity Analysis:**
Overall Time Complexity: O(n).

**Space Complexity Analysis:**
O(n), where n is the number of lines in the dictionary file, stored in memory for sampling.

**Dummy Test File :**

dictionary.txt

```
huckleberry
jackfruit
kumquat
lychee
mulberry
nutmeg
olive
persimmon
pomegranate
rhubarb
salak
tamarind
walnut
yam
zebrafruit
```

**Output :**

```
Generated Password: salak-lychee-pomegranate-kumquat
```

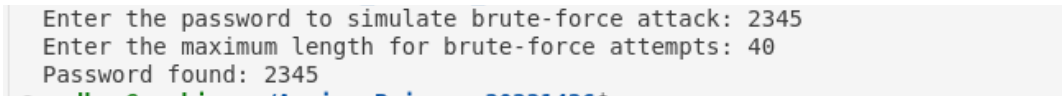
**Question 6 :** Write a Python program that simulates a brute-force attack on a password by trying out all possible character combinations.

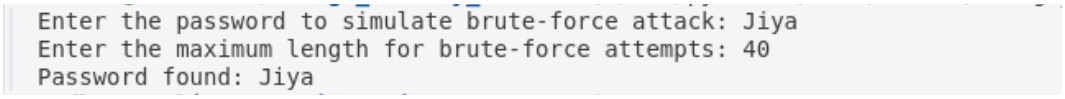
**Solution :**

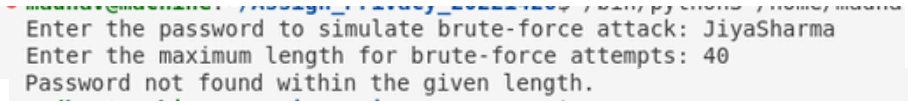
**Time Complexity Analysis:**
1. Brute-force Attack: O(|C|^L), where |C| is the number of characters in the character set (62 here) and L is the length of the password.

**Space Complexity Analysis:**
1. Space Complexity: O(1), as the function generates combinations on the fly without storing them.

**Output :**



**Question 7 :** Demonstrate the usage/sending of a digitally signed document.

**Solution :**

**Time Complexity Analysis:**
1. Key Generation: O(n^3), where n is the key size. Key generation in RSA involves generating large prime numbers.
2. Signing the Document: O(n^2), where n is the key size. Signing involves modular exponentiation with the private key.
3. Verifying the Signature: O(n^2), where n is the key size. Verification also involves modular exponentiation with the public key.

**Space Complexity Analysis:**
1. Space Complexity: O(n),

**Output :**

```
Document signed successfully.
Sending document and signature...
The document is verified successfully. The signature is valid.
```

**Question 8 :** Students needs to conduct a data privacy audit of an organization to identify potential vulnerabilities and risks in their data privacy practices.

**Output :**

```
Starting Data Privacy Audit...

Category: DATA_COLLECTION
 - Are users informed about data collection? (Yes/No): yes
 - Is data collection limited to what's necessary? (Yes/No): yes
 - Are consent mechanisms in place? (Yes/No): yes


Category: DATA_STORAGE
 - Is data encrypted at rest? (Yes/No): no
 - Is sensitive data stored securely? (Yes/No): yes
 - Are backup policies in place? (Yes/No): yes


Category: DATA_ACCESS
 - Is access to data restricted based on roles? (Yes/No): yes
 - Are access logs maintained and monitored? (Yes/No): no
 - Are strong authentication mechanisms used? (Yes/No): yes


Category: COMPLIANCE
 - Is the organization GDPR compliant? (Yes/No): yes
 - Are data retention policies clearly defined? (Yes/No): no
 - Is there a process for handling data subject requests? (Yes/No): no



Analyzing Responses...


Summary of Findings:
 - DATA_COLLECTION: 0 issues identified.
 - DATA_STORAGE: 1 issues identified.
   * Is data encrypted at rest?
 - DATA_ACCESS: 1 issues identified.
   * Are access logs maintained and monitored?
 - COMPLIANCE: 2 issues identified.
   * Are data retention policies clearly defined?
   * Is there a process for handling data subject requests?

Audit report generated: data_privacy_audit_report_2024-11-26_16-24-29.json
```

**Question 9 :** Students needs to explore the requirements of the Data Protection Regulations and develop a plan for ensuring compliance with the regulation

**Output :**

```
Available Regulations:
 - GDPR
 - HIPAA

Enter the regulation to comply with (e.g., GDPR, HIPAA): GDPR

Categories for GDPR:
 - data_processing
 - data_security
 - compliance_monitoring

Enter the categories to include in the compliance plan (comma-separated): data_processing, data_security

Developing Compliance Plan for GDPR...


Compliance Plan:
 - Data_processing:
   * Ensure lawful, fair, and transparent processing of personal data.
   * Obtain explicit consent from data subjects.
   * Provide data subjects with access to their data and the right to correct or delete it.
 - Data_security:
   * Implement appropriate technical and organizational measures.
   * Ensure encryption and pseudonymization of data.
   * Maintain data integrity and confidentiality.

Compliance plan report generated: compliance_plan_GDPR_2024-11-26_16-27-33.json
```

**Question 10 :** Students needs to explore ethical considerations in data privacy, such as the balance between privacy and security, the impact of data collection and analysis on marginalized communities, and the role of data ethics in technology development.

## Output :

Exploring Ethical Considerations in Data Privacy...

Topic: Privacy vs Security

- How should organizations balance individual privacy with the need for security?
Your response: they should collect only the minimum necessary data, implement strong data protection measures, ensure transparency in data usage, and comply with privacy regulations. Security practices should be proportionate to the risks, while respecting individuals' rights and ensuring informed consent.

- What measures can be implemented to protect privacy without compromising security?
Your response: Use encryption, minimize data collection, enforce strict access controls, apply privacy-by-design, anonymize data, and ensure compliance with privacy regulations.

- What are examples of when privacy and security have conflicted?
Your response: Mass surveillance vs. individual privacy, workplace monitoring vs. employee privacy, contact tracing apps vs. user privacy, and data retention laws vs. confidentiality.

---

Topic: Impact on Marginalized Communities

- How can data collection practices disproportionately affect marginalized communities?
Your response: Biased data collection can reinforce stereotypes, increase surveillance, lead to discriminatory profiling, and amplify inequities, disproportionately impacting marginalized communities' rights.

- What steps can be taken to ensure inclusivity and fairness in data analysis?
Your response: Diversify data sources, address biases, involve diverse stakeholders, ensure transparency, validate models for fairness, and conduct regular impact assessments.

- Are there cases where biased data has caused harm? If so, how could it have been prevented?
Your response: Facial recognition biases led to wrongful arrests; preventable through diverse training data, bias audits, and improved algorithm transparency and accountability measures.

---

Topic: Role of Data Ethics in Technology Development

- What ethical principles should guide the development of data-driven technologies?
Your response: Ethical principles include transparency, fairness, accountability, privacy protection, minimizing bias, ensuring inclusivity, informed consent, social benefit, and avoiding harm.

- How can organizations ensure accountability in the use of AI and machine learning?
Your response: Organizations can ensure accountability by implementing clear governance frameworks, conducting regular audits, ensuring model explainability, establishing ethical guidelines, and involving diverse oversight committees.

- What are the consequences of neglecting data ethics in technology development?
Your response: Neglecting data ethics can lead to privacy violations, biased outcomes, discrimination, erosion of trust, reputational damage, legal penalties, and societal harm.

---

Summary of Ethical Considerations:

Topic: Privacy vs Security
- How should organizations balance individual privacy with the need for security?
* they should collect only the minimum necessary data, implement strong data protection measures, ensure transparency in data usage, and comply with privacy regulations. Security practices should be proportionate to the risks, while respecting individuals' rights and ensuring informed consent.
- What measures can be implemented to protect privacy without compromising security?
* Use encryption, minimize data collection, enforce strict access controls, apply privacy-by-design, anonymize data, and ensure compliance with privacy regulations.
- What are examples of when privacy and security have conflicted?
* Mass surveillance vs. individual privacy, workplace monitoring vs. employee privacy, contact tracing apps vs. user privacy, and data retention laws vs. confidentiality.

Topic: Impact on Marginalized Communities
- How can data collection practices disproportionately affect marginalized communities?
* Biased data collection can reinforce stereotypes, increase surveillance, lead to discriminatory profiling, and amplify inequities, disproportionately impacting marginalized communities' rights.
- What steps can be taken to ensure inclusivity and fairness in data analysis?
* Diversify data sources, address biases, involve diverse stakeholders, ensure transparency, validate models for fairness, and conduct regular impact assessments.
- Are there cases where biased data has caused harm? If so, how could it have been prevented?
* Facial recognition biases led to wrongful arrests; preventable through diverse training data, bias audits, and improved algorithm transparency and accountability measures.

Topic: Role of Data Ethics in Technology Development
- What ethical principles should guide the development of data-driven technologies?
* Ethical principles include transparency, fairness, accountability, privacy protection, minimizing bias, ensuring inclusivity, informed consent, social benefit, and avoiding harm.
- How can organizations ensure accountability in the use of AI and machine learning?
* Organizations can ensure accountability by implementing clear governance frameworks, conducting regular audits, ensuring model explainability, establishing ethical guidelines, and involving diverse oversight committees.
- What are the consequences of neglecting data ethics in technology development?
* Neglecting data ethics can lead to privacy violations, biased outcomes, discrimination, erosion of trust, reputational damage, legal penalties, and societal harm.

---