

ASSIGNMENT – III

TITLE :

Create a data security model for encrypting and anonymizing sensitive information in a Big Data system.

NAME : Shinde Shubham Dnyandev.

ROLL NO : 23107121.

BATCH : B

```
(base) admin1@dslcomp03:~$ nano Data_security.py
```

```
GNU nano 4.8 Data_security.py
import pandas as pd
from cryptography.fernet import Fernet
import hashlib

# Load dataset
data = pd.read_csv("Bank Customer Churn Prediction.csv")
print(data.head())

# Sensitive fields
print("Sensitive Fields: customer_id, balance, estimated_salary")

# Generate AES key
key = Fernet.generate_key()
cipher = Fernet(key)
print("Encryption Key Generated Successfully")

# Encrypt customer_id
def encrypt_value(value):
    return cipher.encrypt(str(value).encode()).decode()

data["customer_id_encrypted"] = data["customer_id"].apply(encrypt_value)
print("Customer ID Encrypted")

# Hash balance
def hash_value(value):
    return hashlib.sha256(str(value).encode()).hexdigest()

data["balance_hashed"] = data["balance"].apply(hash_value)
print("Balance Hashed Using SHA-256")

# Mask estimated salary
def mask_salary(value):
    value = str(int(value))
    return "XXXXXX" + value[-2:]

data["estimated_salary_masked"] = data["estimated_salary"].apply(mask_salary)
print("Estimated Salary Masked")

# Drop original sensitive columns
secured_data = data.drop(
    columns=["customer_id", "balance", "estimated_salary"]
)

print(secured_data.head())

# Save secured dataset
secured_data.to_csv("secured_bank_churn_data.csv", index=False)
print("Secured dataset saved as secured_bank_churn_data.csv")
□

^G Get Help      ^O Write Out     ^W Where Is      ^K Cut Text      ^J Justify       ^C Cur Pos       M-U Undo
^X Exit          ^R Read File     ^\ Replace       ^U Paste Text    ^T To Spell      ^ Go To Line     M-E Redo
```

```
(base) admin@dslcomp03:~$ python Data_security.py
customer_id credit_score country gender age tenure balance products_number credit_card active_member estimated_salary churn
0 15634602 619 France Female 42 2 0.00 1 1 1 101348.88 1
1 15647311 608 Spain Female 41 1 83807.86 1 0 1 112542.58 0
2 15619304 502 France Female 42 8 159660.80 3 1 0 113931.57 1
3 15701354 699 France Female 39 1 0.00 2 0 0 93826.63 0
4 15737888 850 Spain Female 43 2 125510.82 1 1 1 79084.10 0
Sensitive Fields: customer_id, balance, estimated_salary
Encryption Key Generated Successfully
Customer ID Encrypted
Balance Hashed Using SHA-256
Estimated Salary Masked
credit_score country gender age tenure ... active_member churn customer_id_encrypted balance_hashed estimated_salary_masked
0 619 France Female 42 2 ... 1 1 gAAAAABpgYEQrN64mHTpnSQ0AkoOP8n2SSVuc7zUtrTN2J... 8aed642bf5118b9d3c859bd4be35ecac75b6e873cce34e... XXXXX48
1 608 Spain Female 41 1 ... 1 0 gAAAAABpgYEQVGE7Qe6ru44DhHQBf3asIngCPTzsIBr5-... c5954bfe1dd28b3196d60db9c91e868954914a2a37810b... XXXXX42
2 502 France Female 42 8 ... 0 1 gAAAAABpgYEQ77SewvPvVodiffVqxh6TYG1595RUvg1wxo... e5af7c8c23a99127b4c0670a6aeced0f7f916e042a3ca0... XXXXX31
3 699 France Female 39 1 ... 0 0 gAAAAABpgYEQacA91CJp7GwnypQwlslyAmdDTynnShM9ou... 8aed642bf5118b9d3c859bd4be35ecac75b6e873cce34e... XXXXX26
4 850 Spain Female 43 2 ... 1 0 gAAAAABpgYEQ04dnd3_QFPDFM2XBAS6AbrQ8o46NxFyIVk... e5413faf86091ac4f45bfadf842e42fdded8b5c0ca0347f... XXXXX04

[5 rows x 12 columns]
Secured dataset saved as secured_bank_churn_data.csv
```

