Strategic Marketing for Personal Loans using Python and SQL.

In the world of modern banking, where every transaction and interaction is recorded, there exists a treasure trove of data, waiting to be unlocked. Behind the numbers and accounts, behind the screens and financial statements, lies a wealth of information that holds the key to enhancing marketing strategies and driving revenue growth. This is where you step in.

You are the data virtuoso, armed with cutting-edge tools and techniques in data analysis. Your mission? To embark on a transformative journey through the bank's data, armed with a trove of customer information. This dataset is like a vault of secrets, containing a multitude of details about the bank's customers, their financial profiles, behaviors, and more. Your task is to transform this raw data into actionable insights that will pave the way for a brighter future for the bank.

The dataset is a goldmine, and with your expertise, you will unlock its potential. By diving deep into the data, you will uncover the fascinating stories of the bank's customers - from their financial habits to their potential for personal loans. You will identify the customer segments with the highest likelihood of conversion, the hidden gems among the liability customers waiting to be discovered.

But your role goes beyond just numbers and statistics. You will bring all the data together to craft a comprehensive picture of customer segments and their potential for conversion. Your analysis will reveal trends and patterns that can drive marketing strategies and campaigns. You will help the bank understand its customers on a whole new level, enabling personalized messaging and targeted marketing efforts.

As you collaborate with the project team, you'll create data visualizations that breathe life into these insights. Your charts and graphs will make abstract data tangible, transforming it into a compelling narrative. Your work will be a source of inspiration for not only the project team but also for the entire banking community. Your dedication to extracting meaningful insights from data will help the bank thrive, retain customers, and boost its revenue.

In the realm of data-driven marketing, you are the unsung hero, the one who unveils the hidden opportunities within the dataset. Your dedication to data and your ability to transform it into illuminating insights contribute to the ongoing saga of modern banking, making every customer interaction, financial decision, and marketing campaign that much more meaningful.

Module 1

Task 1: Unlocking Banking Potential.

In the relentless quest to empower banks with data-driven strategies, you embark on a journey to unleash the potential of personal loan conversion. Armed with the Bank_Personal_Loan_Modelling dataset, your mission is clear: transform raw data into actionable insights. By analyzing customer demographics and behavior, you aim to identify key segments with a high likelihood of converting liability customers to personal loan customers.

The objective is simple yet profound: to optimize marketing strategies, reduce costs, and fuel customer retention and revenue growth.

```python
#--- Import Pandas ---
import pandas as pd
#--- Read in dataset ----
df = pd.read_csv('./Bank_Personal_Loan_Modelling.csv')

# ---WRITE YOUR CODE FOR TASK 1 ---
#--- Inspect data ---
```

Task 2: The Quest for Accuracy.

In the realm of data analysis, our mission is clear: to identify and eliminate duplicates within the dataset. By uncovering and rectifying these redundancies, we aim to enhance data accuracy and reliability. The task at hand is pivotal in ensuring that insights drawn from this data are trustworthy, making it a cornerstone of informed decision-making.

```python
# --- WRITE YOUR CODE FOR TASK 2 ---
duplicates = df.duplicated().sum()

#--- Inspect data ---
```

Task 3: Chasing Data Perfection.

Our journey revolves around data perfection, where we confront the issue of null values within the dataset. By identifying and addressing these gaps, we pave the way for more robust analysis. This mission is vital to ensure the dataset's completeness, reliability, and integrity, forming the bedrock for data-driven decision-making.

```python
# --- WRITE YOUR CODE FOR TASK 3 ---
null_values = df.isnull().sum()
#--- Inspect data ---
```

Task 4: Streamlining Data for Precision.

In the pursuit of precision, we embark on the task of data refinement. By removing non-essential columns like 'ID' and 'ZIP Code,' we streamline the dataset for more focused analysis. This action enhances data clarity and simplifies our quest for insights, allowing us to make efficient, data-driven decisions.

```python
#--- WRITE YOUR CODE FOR TASK 4 ---
df.drop(columns = ['ID', 'ZIP Code'] , axis = 1 , inplace = True)

#--- Inspect data ---
```

Task 5: Removing Negative Experience.

Our mission is to ensure data integrity by eliminating entries with negative 'Experience.' This task purifies the dataset, making it more reliable for analysis. By removing inconsistencies, we pave the way for precise insights and data-driven decisions, ensuring our journey is built on a solid foundation of accurate information.

```python
#--- WRITE YOUR CODE FOR TASK 5 ---

condition_mask = df['Experience'] >= 0

filtered_df = df[condition_mask]

df = filtered_df

#--- Inspect data ---
```

Module 2

Task 1: Transforming Education Levels.

In our relentless pursuit of clarity, we're transforming the 'Education' data into more comprehensible categories. By mapping the values to 'Undergraduate,' 'Graduate,' and 'Professional,' we simplify the dataset for easier interpretation. This task makes the data more accessible and ensures that our journey towards insights and informed decisions is smoother and more insightful.

```python
def edu(x):
    if x==1:
        return "Undergraduate"
    elif x==2:
        return "Graduate"
    else:
        return "Professional"


#--- WRITE YOUR CODE FOR TASK 1 ---
#df = ...
df['EDU'] = df['Education'].apply(edu)
#--- Inspect data ---
```

Task 2: Categorizing Account Holders.

Our mission delves into categorizing account holders based on their financial choices. By defining categories like "Holds Securities & Deposits," "Does not hold Securities & Deposits account," "Holds only Securities account," and "Holds only Deposits account," we provide a clear insight into customers' financial preferences. This task streamlines data interpretation, offering valuable insights for strategic decision-making.

```python
def security(y):
    if(y['Securities Account'] == 1) & (y['CD Account'] == 1):
        return "Holds Securities & Deposits"
    if(y['Securities Account'] == 0) & (y['CD Account'] == 0):
        return "Does not hold Securities & Deposits account"
    if(y['Securities Account'] == 1) & (y['CD Account'] == 0):
        return "Holds only Securities account"
    if(y['Securities Account'] == 0) & (y['CD Account'] == 1):
        return "Holds only Deposits account"


#--- WRITE YOUR CODE FOR TASK 2 ---
df['Account_holder_category'] = df.apply(security, axis=1)

#--- Inspect data ---
```

Task 3: Archiving the Transformed Data.

Our mission is to safeguard the progress we've made in data refinement. By saving the transformed dataset as 'liability.csv,' we ensure that the valuable insights and clean data are readily available for future analysis and decision-making. This task preserves the fruits of our labor, ensuring that our journey towards data-driven decisions is built on a solid foundation of archived information.

```python
# ...WRITE YOUR CODE FOR TASK 3 ...
#export the cleaned data

df.to_csv('liability.csv', index=False)

#--- Inspect data ---
```

Task 4: Data Download, Import, and Database Connection.

```python
# -- Load the sql extention ----
%load_ext sql

# --- Load your mysql db using credentials from the "DB" area ---

from urllib.parse import quote
encoded_password = quote('Cab#22se')
%sql mysql+pymysql://<b36eea86>:<Cab#22se>@localhost/<b36eea86>
```

Module 3

Task 1: Quantifying Progress.

In our quest for data-driven decisions, we embark on an assessment of our transformed dataset. By counting the number of records in 'liability,' we quantify the extent of our data transformation efforts. This task measures our success in streamlining and refining the data,

providing an essential metric to gauge our progress. It is a vital step in ensuring the reliability and effectiveness of our data-driven decision-making journey.

```sql
%%sql SELECT COUNT(*) from liability
```

Task 2: Seeking Average Income.

In our data-driven journey, we are now focused on understanding the financial landscape. By executing the SQL query, we seek to calculate the average income of the customers from the 'liability' dataset. This task is pivotal as it provides a key metric for assessing the economic profiles of our customer base. The average income serves as a compass, guiding us in tailoring marketing strategies, personalized messaging, and financial product offerings. Ultimately, this data-driven decision aids in ensuring that our strategies resonate with the financial realities of our customers, driving success in personal loan conversion and revenue growth.

```sql
%%sql
SELECT AVG(Income) AS AverageIncome
FROM liability
```

Task 3: Identifying High-Potential Customers.

In our relentless pursuit of data-driven excellence, we turn our focus to identifying high-potential customers. By executing this SQL query, we retrieve the top 10 customers with the highest income from the 'liability' dataset. This task is paramount because it allows us to pinpoint individuals with significant financial capacity. These customers represent a prime target for personalized marketing and personal loan conversion efforts. The insights gained from this analysis are instrumental in optimizing campaigns to attract and retain these high-value customers, ultimately driving revenue growth and banking success.

```sql
%%sql
SELECT * FROM liability
ORDER BY Income DESC
LIMIT 10;
```

Task 4: Uncovering Educational Financial Trends.

In our quest for data-driven insights, we now explore the intersection of education and income. By executing this SQL query, we calculate the average income for customers within different education levels using the 'liability' dataset. This task is of paramount importance as it reveals financial trends and disparities among various education categories. It equips us with crucial knowledge to fine-tune marketing strategies and personalized messaging that resonates with the unique financial needs of each group. This data-driven approach paves the way for enhanced personal loan conversion strategies, empowering the bank to cater effectively to diverse customer segments and drive revenue growth.

```
%%sql
SELECT Education, AVG(Income) FROM liability
GROUP BY Education
```

Task 5: Top Income Earners by Education.

In our data-driven quest, we delve into a deep exploration of financial excellence and education. Through this SQL query, we create a ranked dataset that identifies the top two income earners within each education level category from the 'liability' dataset. This task is of immense significance as it uncovers the brightest financial stars within distinct education groups. The insights extracted here are invaluable for crafting tailored marketing strategies and personalized messaging. By recognizing and nurturing these high-earning individuals, we empower the bank to foster customer loyalty, stimulate personal loan conversion, and drive remarkable revenue growth.

```
%%sql
WITH RankedData AS (
    SELECT
        *,
        RANK() OVER (PARTITION BY Education ORDER BY Income DESC) AS
IncomeRank
    FROM liability
)

SELECT * FROM RankedData WHERE IncomeRank <= 2;
```

Task 6: Profiling Customer Demographics.

In our data-driven journey, we're shifting the spotlight to customer age demographics. Through this SQL query, we categorize customers into distinct age groups and count their representation within each group using the 'liability' dataset. This task is pivotal as it offers a granular view of the customer base, helping to craft personalized marketing strategies and pinpoint specific segments for personal loan conversion efforts. By understanding the unique characteristics of each age group, the bank can tailor its approach, ensuring that every customer, whether aged 18-30, 31-45, 46-60, or 61+, receives targeted and effective messaging. This data-driven approach promises to enhance customer engagement and drive revenue growth.

```
%%sql
SELECT
    CASE
        WHEN Age BETWEEN 18 AND 30 THEN '18-30'
        WHEN Age BETWEEN 31 AND 45 THEN '31-45'
        WHEN Age BETWEEN 46 AND 60 THEN '46-60'
        ELSE '61+'
    END AS AgeGroup,
    COUNT(*) AS RecordCount
FROM
    liability
```

```
GROUP BY
    AgeGroup;
```

Task 7: Analyzing Age vs. Credit Card Spending.

In our relentless pursuit of data-driven insights, we are now exploring the relationship between age and credit card spending. Through this SQL query, we calculate the average age of customers whose credit card spending (CCAvg) exceeds the overall average CCAvg within the 'liability' dataset. This task is essential as it helps us identify and understand customers who are above-average spenders. The insights garnered here provide a deeper understanding of the demographics of these high-CCAvg individuals. By tailoring marketing efforts and personalized messaging to this specific group, the bank can further enhance personal loan conversion strategies and revenue growth.

```
%%sql
SELECT
    AVG(Age) AS AverageAge
FROM
    liability
WHERE
    CCAvg > (SELECT AVG(CCAvg) FROM liability);
```

Task 8: Unveiling High-Income Elite.

In our quest for data-driven excellence, we aim to unveil the high-income elite among our customers. By executing this SQL query, we extract customer records with incomes exceeding 1.5 times the average income within the 'liability' dataset. This task is crucial as it identifies the financially well-endowed individuals within our customer base. These high-income customers present a prime opportunity for personal loan conversion, as their financial capacity opens doors for higher-value financial products. By tailoring marketing strategies and personalized messaging to this elite group, the bank stands to maximize its revenue growth and foster lasting customer relationships.

```
%%sql
SELECT *
FROM liability
WHERE Income > 1.5 * (SELECT AVG(Income) FROM liability);
```

Task 9: Family Dynamics Analysis.

In our data-driven journey, we're now delving into family dynamics and age. Through this SQL query, we determine the youngest family member within each family category from the 'liability' dataset. This task holds significance as it allows us to understand the distribution of age within different family structures. The insights gleaned here provide a nuanced view of customer demographics, which can be instrumental in tailoring marketing strategies and personalized messaging. By recognizing the dynamics of each family type and the age of the youngest member, the bank can create targeted campaigns that resonate with the unique needs and

aspirations of these distinct groups. This data-driven approach promises to enhance customer engagement and drive revenue growth.

```sql
%%sql
SELECT
    Family,
    MIN(Age) AS YoungestAge
FROM
    liability
GROUP BY
    Family;
```

Task 10: Mortgage Holders.

In our quest for data-driven insights, we're directing our attention to customers who hold mortgages. Through this SQL query, we retrieve customer records from the 'liability' dataset where the mortgage amount is greater than zero. This task is pivotal as it identifies customers who are indebted through mortgages, offering a specific demographic segment for in-depth analysis. Understanding the characteristics of these mortgage holders is crucial for tailoring marketing strategies, personalized messaging, and personal loan conversion efforts. By focusing on this customer group, the bank can provide solutions and financial products that address their unique needs, ultimately driving revenue growth and customer satisfaction.

```sql
%%sql
SELECT *
FROM liability
WHERE Mortgage > 0;
```

Task 11: Understanding Customer Distribution

In our data-driven journey, we're shifting the spotlight to customer educational backgrounds. Through this SQL query, we categorize customers by their education levels and count their representation within each group using the 'liability' dataset. This task is pivotal as it provides an in-depth view of the educational diversity among the customer base. These insights enable the bank to create targeted marketing campaigns and personalized messaging that resonate with the unique financial needs and aspirations of each education category. Understanding the characteristics of each group empowers the bank to foster customer engagement and drive revenue growth through informed decision-making and tailored strategies.

```sql
%%sql
SELECT
    Education,
    COUNT(*) AS CustomerCount
FROM
    liability
GROUP BY
    Education;
```