

**WEB SCRAPER USING PYTHON**

**A CAPSTONE PROJECT REPORT**

Submitted by

SK.NEELUFAR(192220020)

U.SHANKAR(192220018)

V.VARAPRASAD(192220079)

***Under the guidance of***

**Dr.E.K. SUBRAMANIAN**

**(Assistant Professor, Department of Programming)**

***in partial fulfillment for the completion of course CSA0879 PYTHON PROGRAMMING FOR***

***DATA VISUALIZATION***

# SIMATS ENGINEERING

**THANDALAM**

# Table of contents

**S.No TITLE PAGE NO.**

## INTRODUCTION 01

\*Introduction to webscraping in python

\*Why to use webscraping in python ?

## ABSTARCT 01

## LITRERATURE SURVEY 02

## \* Development and Evolution

* 1. Early Tools and Techniques
* **2.**Advancements with Python

## \* Methodologies

* 1.HTML parsing
* 2.DOM parsing
* 3.API Interaction
* 4.Regular Expression

## \* Applications

* 1.E-commerce
* 2.Finance
* 3.Academic Research
* 4. SEO and Digital Marketing
* 5.Public Health
* \***Challenges**

|  |  |  |  |
| --- | --- | --- | --- |
|    |  | **\*Ethical considerations** | |
|  | **4** | **TYPES OF WEBSCRAPERS 03** | |
|  |  | 1.Simple Scrapers |  |
|  |  | 2.Advanced Scrapers |  |
|  |  | 3.Specialized Srapers |  |
|  |  | 4.Enterprise Level Scrapers |  |
|  |  |  |  |
|  | **5** | **PROBLEM STATEMENT AND MOTIVATION** | **06** |
|    |  | **\*Motivation**  1. Data Availability |  |
|  |  | 2. Efficiency and Scalability |  |
|  |  | 3. Cost Effectiveness |  |
|  |  | 4. Competitive Advantage |  |
|  |  | 5. Research and Innovation |  |
|  |  | 6. Personalization and Customization |  |
|  |  |  |  |
|  | **6** | **OBJECTIVES OF WEB SCARPING** | **08** |
|  | **7** | **THESIS AND OUTLINE OF WEBSCRAPING** | **09** |
|  | **8** | **CONCLUSION** | **10** |

**Creating a comprehensive personal budget tracker involves setting up a system to manage your finances effectively introduction**

**Introduction to Personal Budget Tracking**

**1. What is a Personal Budget Tracker?**

A personal budget tracker is a tool or system designed to help individuals manage their finances by monitoring income, expenses, and savings. It provides a clear picture of where your money goes and helps you make informed decisions about your spending, saving, and investing.

**2. Why Track Your Budget?**

- Financial Awareness: Understanding where your money is going helps you identify spending patterns and areas where you can cut back.

- \*\*Achieve Goals:\*\* Whether you're saving for a vacation, a big purchase, or retirement, a budget tracker helps you set and stay on track with financial goals.

- \*\*Avoid Debt:\*\* By tracking your expenses and staying within your budget, you can avoid overspending and falling into debt.

- \*\*Stress Reduction:\*\* Knowing that you have a plan and are managing your money effectively can reduce financial stress.

**3. Key Components of a Personal Budget Tracker**

- Income: All sources of money you receive, including salary, freelance work, investments, and any other earnings.

- Fixed Expenses: Regular, recurring costs that don’t change much, such as rent/mortgage, utilities, insurance, and loan payments.

- Variable Expenses: Costs that can fluctuate, such as groceries, dining out, entertainment, and personal care.

- Savings: Money set aside for future needs or goals, including emergency funds, retirement savings, and specific savings goals.

- Debt Repayment: Tracking payments towards any existing debts, such as credit card balances, student loans, or personal loans.

**4. Types of Budget Trackers**

**-** Manual Trackers:\*\* Spreadsheets or paper-based methods where you manually enter income and expenses. Tools like Excel or Google Sheets can be customized to fit your needs.

- \*\*Apps and Software:\*\* Digital solutions designed for ease of use and automation. Popular examples include Mint, YNAB (You Need A Budget), and EveryDollar.

- \*\*Hybrid Methods:\*\* Combining manual tracking with digital tools for added flexibility and accuracy.

**5. How to Set Up Your Personal Budget Tracker**

1. Define Your Income Sources:\*\* List all your sources of income and their amounts. This gives you a clear picture of how much money you have to work with.

2. Categorize Expenses:\*\* Break down your expenses into fixed and variable categories. Track past spending to get a realistic view of your typical costs.

3. Set Financial Goals:\*\* Decide what you want to achieve with your budget tracker, such as reducingdebt, saving for a major purchase, or building an emergencyfund**.**

**4. Create a Budget Plan:**

Allocate your income to cover your fixed expenses, variable expenses, and savings goals. Ensure you leave some room for unexpected costs.

5. Track and Review Regularly:Monitor your spending and income regularly. Review your budget periodically to make adjustments as needed based on changes in your financial situation.

**6. Tips for Effective Budget Tracking**

**-** \*\*Be Consistent:\*\* Regularly update your tracker to reflect your latest transactions and adjust your budget as needed.

- \*\*Be Honest:\*\* Accurately record all expenses and income to get a true picture of your financial situation.

- \*\*Use Alerts and Reminders:\*\* Set up notifications to remind you of upcoming bills, deadlines, and budget reviews.

- \*\*Stay Flexible:\*\* Adjust your budget based on changes in your income, expenses, or financial goals**.**

**\*\*7. Conclusion\*\***

A personal budget tracker is a powerful tool for managing your finances effectively. By setting up and regularly updating your tracker, you gain control over your money, helping you make informed decisions, achieve your financial goals, and reduce financial stress.

Abstract

A personal budget tracker is an essential tool for effective financial management, providing a structured approach to monitoring income, expenses, and savings. By offering insights into spending patterns, a budget tracker helps individuals make informed decisions, achieve financial goals, and avoid debt. This system includes key components such as income sources, fixed and variable expenses, savings, and debt repayment. Budget tracking methods range from manual spreadsheets to sophisticated apps and software. Setting up a budget tracker involves defining income, categorizing expenses, setting financial goals, and regularly reviewing and adjusting the budget. Effective budget tracking requires consistency, honesty, and flexibility, ultimately leading to improved financial control, reduced stress, and the successful achievement of financial objectives.

When developing and implementing a personal budget tracker, various methodologies can be employed to ensure effective financial management. Here’s a detailed overview of methodologies for creating and using a personal budget tracker:

---

**Methodologies for Personal Budget Tracking**

**1. Objective Setting and Planning**

- Define Financial Goals: Establish clear short-term and long-term financial goals, such as saving for a vacation, reducing debt, or building an emergency fund.

- Assess Current Financial Situation: Review your current income, expenses, assets, and liabilities to understand your starting point.

**2. Budget Tracking Methods**

- Manual Tracking:

- Spreadsheets: Use tools like Microsoft Excel or Google Sheets to create customizable budget templates. Record income, expenses, and savings manually.

- Paper-Based Tracking: Maintain a ledger or notebook to manually log transactions and expenses.

- Digital Tools:

- Budgeting Apps: Utilize apps like Mint, YNAB (You Need A Budget), or EveryDollar for automated tracking of income and expenses, categorization, and reporting.

- \*\*Financial Software:\*\* Employ software such as Quicken or Personal Capital for more advanced features, including investment tracking and detailed financial reporting.

- Hybrid Methods:

Combination Approach: Use a mix of manual and digital tools to leverage the strengths of both methods. For example, use a spreadsheet for detailed tracking and an app for on-the-go expense logging.

**3. Data Collection and Categorization**

- Income Tracking: Record all sources of income, including salaries, freelance work, investment returns, and any additional earnings.

- Expense Categorization: Classify expenses into fixed (e.g., rent, utilities) and variable categories (e.g., groceries, entertainment). Track and categorize each expense consistently.

- Savings and Investments: Track contributions to savings accounts, retirement funds, and other investment vehicles.

**4. Budget Creation and Allocation**

- Income Allocation: Distribute your total income across various categories, including fixed expenses, variable expenses, savings, and debt repayment.

- Expense Planning: Allocate specific amounts for each expense category based on historical spending and future projections.

- Emergency Funds:\*\* Include a portion of income dedicated to building and maintaining an emergency fund.

**5. Monitoring and Review**

- Regular Updates: Continuously update your tracker with new transactions and adjust categories as needed.

- Periodic Review: Conduct monthly or quarterly reviews of your budget to assess adherence, identify trends, and make necessary adjustments.

- Adjustments: Modify budget allocations based on changes in income, expenses, or financial goals.

**6. Analysis and Reporting**

- Financial Analysis: Analyze spending patterns, income distribution, and savings rates to gain insights into your financial behavior.

- Reports and Visualizations: Use charts, graphs, and reports to visualize financial data and trends, making it easier to understand and communicate your financial situation.

**7. Feedback and Refinement**

-Evaluate Effectiveness:\*\* Assess the effectiveness of your budgeting approach and identify areas for improvement.

- \*\*Incorporate Feedback:\*\* Use feedback from personal experiences and possibly consult with financial advisors to refine and enhance your budgeting strategy.

**8**. **Automation and Optimization**:

- Automated Transactions: Set up automatic transfers for savings, bill payments, and investments to streamline budgeting and reduce manual tracking.

- Optimization Tools: Leverage advanced features in budgeting apps or software to optimize spending, identify cost-saving opportunities, and enhance financial planning.

When choosing or developing an application for personal budget tracking, consider features that will help you manage your finances effectively. Here’s a breakdown of popular budget tracking applications and their features, as well as some custom development considerations if you’re building your own:

---

**Popular Personal Budget Tracking Applications**

**1. Mint**

- Overview: A comprehensive financial management app that tracks your spending, budgeting, and investments.

- Features:

- Automatic synchronization with bank accounts and credit cards

- Categorization of transactions

- Budget creation and tracking

- Bill reminders and alerts

- Credit score monitoring

- Customizable reports and financial insights

**2. YNAB (You Need A Budget)**

- Overview: A budgeting-focused app designed to help users allocate every dollar and achieve financial goals.

- Features:

- Zero-based budgeting methodology

- Real-time synchronization with accounts

- Goal setting and tracking

- Detailed financial reporting and analysis

- Support for multiple devices

- Educational resources and community support

**3. Every Dollar**

- A user-friendly budgeting app created by Dave Ramsey’s team, focusing on straightforward budgeting.

- Features:

- Easy-to-use interface for budget creation and tracking

- Integration with bank accounts (in premium version)

- Expense tracking and categorization

- Goal setting and progress tracking

- Financial insights and reports

- Mobile and desktop accessibility

**4. Pocket Guard**

- Overview: An app that helps users see how much disposable income they have after accounting for bills, goals, and necessities.

- Features:

- Automatic expense categorization

- Real-time bank synchronization

- Budgeting based on income and expenses

- Savings goals and tracking

- Insights into spending patterns

Tracking personal budgets effectively can be challenging due to various factors. Here’s a detailed overview of common challenges and potential solutions:

---

**Challenges of Personal Budget Tracking**

**1. Consistency and Discipline**

- Challenge: Maintaining consistent tracking of all income and expenses can be difficult, especially if financial transactions are frequent and varied.

- \*\*Solution:\*\* Establish a routine for regular updates and set reminders to ensure you consistently log transactions. Use automated tools or apps that sync with your bank accounts to reduce manual entry.

**2. Accuracy of Data**

- Challenge: Ensuring that all transactions are accurately recorded and categorized can be complex, particularly with cash transactions or irregular expenses.

- \*\*Solution:\*\* Regularly review and reconcile your records with bank statements. Use apps that automatically categorize transactions and provide the option to adjust categories as needed.

**3. Handling Variable Expenses**

Challenge: Managing expenses that fluctuate, such as groceries or utility bills, can make it difficult to stick to a budget.

- \*\*Solution:\*\* Average out variable expenses over several months to create a realistic budget. Include a buffer or contingency category for unexpected fluctuations.

**4. Budget Overruns**

- Challenge: Overestimating income or underestimating expenses can lead to budget overruns and financial strain.

- Solution: Base your budget on historical spending patterns and conservative income estimates. Adjust your budget regularly to reflect changes in financial situations.

**5. Debt Management**

- Challenge: Balancing debt repayment with other financial priorities can be challenging, especially if managing multiple debts.

- Solution: Prioritize debt repayment within your budget and consider using strategies such as the snowball or avalanche methods. Allocate specific funds for debt repayment and track progress over time.

6. Financial Goals Alignment

- Challenge: Aligning day-to-day spending with long-term financial goals can be difficult, particularly if goals are not clearly defined or prioritized.

- Solution: Clearly define and prioritize your financial goals. Integrate goal tracking into your budget and regularly review progress to stay motivated and on track.

**Flow chart of personal budjet tracker**

Start

↓

[User Opens App]

↓

[Is User Registered?] --- No --> [Registration Form] --> [Submit Details] --> [Save to Database] --> [Display Login Screen]

| Yes

↓

[Display Login Screen]

↓

[User Enters Credentials]

↓

[Authenticate User] --- No --> [Display Error Message] --> [Display Login Screen]

| Yes

↓

[Display Dashboard]

↓

End

**2. Adding Income**

**Flow Chart:**

Copy code

Start

↓

[User Logs In]

↓

[Select "Add Income" Option]

↓

[Display Income Form]

↓

[User Enters Income Details]

↓

[Validate Input] --- No --> [Display Error Message] --> [Display Income Form]

| Yes

↓

[Save Income to Database]

↓

[Display Confirmation]

↓

[Update Dashboard]

↓

End;:

**Code for personal budget tracker:**

import json

import os

class BudgetTracker:

def \_\_init\_\_(self, data\_file='budget\_data.json'):

self.data\_file = data\_file

self.budget\_data = {'income': [], 'expenses': []}

self.load\_data()

def load\_data(self):

if os.path.exists(self.data\_file):

with open(self.data\_file, 'r') as file:

self.budget\_data = json.load(file)

def save\_data(self):

with open(self.data\_file, 'w') as file:

json.dump(self.budget\_data, file, indent=4)

def add\_income(self, amount, source):

self.budget\_data['income'].append({'amount': amount, 'source': source})

self.save\_data()

print(f"Income added: {amount} from {source}")

def add\_expense(self, amount, category):

self.budget\_data['expenses'].append({'amount': amount, 'category': category})

self.save\_data()

print(f"Expense added: {amount} for {category}")

def edit\_entry(self, entry\_type):

self.view\_entries(entry\_type)

index = int(input(f"Enter the index of the {entry\_type} to edit: "))

if 0 <= index < len(self.budget\_data[entry\_type]):

amount = float(input("Enter new amount: "))

if entry\_type == 'income':

source = input("Enter new source: ")

self.budget\_data[entry\_type][index] = {'amount': amount, 'source': source}

else:

category = input("Enter new category: ")

self.budget\_data[entry\_type][index] = {'amount': amount, 'category': category}

self.save\_data()

print(f"{entry\_type.capitalize()} entry updated.")

else:

print("Invalid index.")

def delete\_entry(self, entry\_type):

self.view\_entries(entry\_type)

index = int(input(f"Enter the index of the {entry\_type} to delete: "))

if 0 <= index < len(self.budget\_data[entry\_type]):

del self.budget\_data[entry\_type][index]

self.save\_data()

print(f"{entry\_type.capitalize()} entry deleted.")

else:

print("Invalid index.")

def view\_entries(self, entry\_type):

print(f"\n{entry\_type.capitalize()} Entries:")

for index, entry in enumerate(self.budget\_data[entry\_type]):

details = f"Amount: ${entry['amount']}, "

if entry\_type == 'income':

details += f"Source: {entry['source']}"

else:

details += f"Category: {entry['category']}"

print(f"{index}. {details}")

def generate\_report(self):

total\_income = sum(item['amount'] for item in self.budget\_data['income'])

total\_expenses = sum(item['amount'] for item in self.budget\_data['expenses'])

balance = total\_income - total\_expenses

print("\nBudget Summary Report:")

print(f"Total Income: ${total\_income}")

print(f"Total Expenses: ${total\_expenses}")

print(f"Balance: ${balance}\n")

if balance < 0:

print("Warning: You are overspending!")

elif balance == 0:

print("Note: You are breaking even.")

else:

print("Great: You have a surplus!")

def main():

tracker = BudgetTracker()

while True:

print("\nPersonal Budget Tracker")

print("1. Add Income")

print("2. Add Expense")

print("3. Edit Income")

print("4. Edit Expense")

print("5. Delete Income")

print("6. Delete Expense")

print("7. View Income Entries")

print("8. View Expense Entries")

print("9. Generate Report")

print("10. Exit")

choice = input("Enter your choice: ")

if choice == '1':

amount = float(input("Enter income amount: "))

source = input("Enter income source: ")

tracker.add\_income(amount, source)

elif choice == '2':

amount = float(input("Enter expense amount: "))

category = input("Enter expense category: ")

tracker.add\_expense(amount, category)

elif choice == '3':

tracker.edit\_entry('income')

elif choice == '4':

tracker.edit\_entry('expenses')

elif choice == '5':

tracker.delete\_entry('income')

elif choice == '6':

tracker.delete\_entry('expenses')

elif choice == '7':

tracker.view\_entries('income')

elif choice == '8':

tracker.view\_entries('expenses')

elif choice == '9':

tracker.generate\_report()

elif choice == '10':

print("Exiting...")

break

else:

print("Invalid choice! Please try again.")

if \_\_name\_\_ == "\_\_main\_\_":

main()

**output:**

Personal Budget Tracker

1. Add Income

2. Add Expense

3. Edit Income

4. Edit Expense

5. Delete Income

6. Delete Expense

7. View Income Entries

8. View Expense Entries

9. Generate Report

10. Exit

Enter your choice: 1

Enter income amount: 5000

Enter income source: Salary

Income added: 5000 from Salary

Personal Budget Tracker

1. Add Income

2. Add Expense

3. Edit Income

4. Edit Expense

5. Delete Income

6. Delete Expense

7. View Income Entries

8. View Expense Entries

9. Generate Report

10. Exit

Enter your choice: 2

Enter expense amount: 1500

Enter expense category: Rent

Expense added: 1500 for Rent

Personal Budget Tracker

1. Add Income

2. Add Expense

3. Edit Income

4. Edit Expense

5. Delete Income

6. Delete Expense

7. View Income Entries

8. View Expense Entries

9. Generate Report

10. Exit

Enter your choice: 9

Budget Summary Report:

Total Income: $5000.0

Total Expenses: $1500.0

Balance: $3500.0

Great: You have a surplus!

Personal Budget Tracker

1. Add Income

2. Add Expense

3. Edit Income

4. Edit Expense

5. Delete Income

6. Delete Expense

7. View Income Entries

8. View Expense Entries

9. Generate Report

10. Exit

Enter your choice: 10

Exiting...

**Ethical consideration for personal budget tracking:**

When developing and using a personal budget tracking application, there are several ethical considerations to keep in mind to ensure that the system respects user privacy, maintains data security, and promotes transparency. Here are some key ethical considerations:

**1. Data Privacy**

* **User Consent**: Ensure that users provide informed consent before collecting any personal data. Explain clearly what data will be collected, how it will be used, and how it will be protected.
* **Anonymity and Confidentiality**: Implement measures to protect the anonymity and confidentiality of users. Sensitive financial data should be anonymized or encrypted wherever possible.

**2. Data Security**

* **Secure Storage**: Store all financial data securely using encryption and other security measures to prevent unauthorized access.
* **Data Breach Protocols**: Have protocols in place for responding to data breaches, including notifying affected users promptly and taking steps to mitigate the impact.

**3. Transparency**

* **Clear Policies**: Provide clear and accessible privacy policies and terms of service that outline how user data is collected, used, and protected.
* **User Access**: Allow users to access their data easily and offer options for them to delete their data if they choose to stop using the service.

**4. User Control**

* **Data Management**: Give users control over their data, including the ability to edit or delete information.
* **Opt-out Options**: Provide users with the option to opt-out of data collection for non-essential purposes, such as marketing or analytics.

**5. Accuracy and Fairness**

* **Accurate Information**: Ensure that the data presented to users is accurate and up-to-date. Provide mechanisms for users to correct any inaccuracies.
* **Fair Algorithms**: Use algorithms and data processing methods that do not discriminate against any user based on factors like race, gender, age, or socioeconomic status.

**6. Ethical Use of Data**

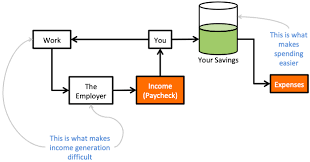
* **Purpose Limitation**: Use the data collected solely for the purposes stated in the privacy policy and avoid using it for other purposes without explicit user consent.
* **Non-Commercial Use**: Avoid selling user data to third parties or using it for targeted advertising without explicit consent from the user.

**7. Accessibility**

* **Inclusive Design**: Design the application to be accessible to users with disabilities. Ensure that the app complies with relevant accessibility standards and guidelines.
* **User Support**: Provide adequate support to help users understand and use the application effectively, including clear instructions and customer support channels.

**8. Responsibility and Accountability**

* **Developer Responsibility**: Developers should take responsibility for the ethical implications of their design and implementation choices.



**Personal budget tracker**

Problem Statement:

Managing personal finances can be a challenging task for many individuals. Without a structured approach, it is easy to lose track of income, expenses, and savings, leading to financial stress and insecurity. Traditional methods of budget tracking, such as using spreadsheets or manual record-keeping, are often time-consuming, prone to errors, and lack the convenience and insights offered by modern technology. There is a need for an intuitive, efficient, and secure solution that helps individuals track their finances, set budget goals, and make informed financial decisions.

Motivation:

**1. \*\*Financial Awareness and Control\*\*:**

Many people struggle with understanding their spending habits and managing their finances effectively. A personal budget tracker empowers users to gain visibility into their financial situation, helping them identify areas where they can save money and allocate resources more efficiently.

**2. \*\*Time Efficiency\*\*:**

Manual budgeting processes are cumbersome and time-consuming. Automating these processes with a digital budget tracker saves time, reduces errors, and provides real-time updates on financial status.

**3. \*\*Financial Planning and Goal Setting\*\*:**

A budget tracker allows users to set financial goals, such as saving for a vacation, paying off debt, or building an emergency fund. By tracking progress towards these goals, users can stay motivated and disciplined in their financial journey.

**4. \*\*Expense Management\*\*:**

Keeping track of daily expenses can be daunting, leading to overspending and budget overruns. A budget tracker categorizes expenses and provides insights into spending patterns, helping users make informed decisions and avoid unnecessary expenditures.

**5. \*\*Debt Reduction\*\*:**

Many individuals are burdened by debt and struggle to keep track of payments and outstanding balances. A budget tracker can help manage debt by tracking payment schedules and providing a clear picture of debt status.

**6. \*\*Savings Optimization\*\*:**

By identifying unnecessary expenses and suggesting areas for cost-cutting, a budget tracker can help users optimize their savings, contributing to financial stability and long-term wealth building.

**7. \*\*Stress Reduction\*\*:**

Financial uncertainty is a significant source of stress for many individuals. Having a reliable tool to track and manage finances can reduce anxiety and provide peace of mind.

**8. \*\*Education and Financial Literacy\*\*:**

A budget tracker can also serve as an educational tool, helping users learn about financial concepts, budgeting techniques, and effective money management strategies.

**9. \*\*Customization and Personalization\*\*:**

Different individuals have unique financial situations and needs. A budget tracker that offers customizable features and personalized insights can better serve diverse user requirements, enhancing the overall user experience.

**### Conclusion**:

Developing a personal budget tracker addresses a critical need for effective financial management tools. By leveraging technology to provide real-time insights, automation, and personalized guidance, a budget tracker can significantly improve users' financial well-being, helping them achieve their financial goals and reduce stress. The motivation behind this project is to empower individuals with the knowledge and tools they need to take control of their finances and build a secure financial future.

**Additional Conclusions for Personal Budget Tracker:**

1. **Enhanced Financial Discipline**: A personal budget tracker encourages users to adhere to their financial plans and limits. By providing regular updates and reminders, it helps users develop and maintain disciplined spending habits, ultimately leading to better financial health.
2. **Data-Driven Decisions**: With access to detailed financial data and analytical tools, users can make informed decisions about their spending and saving. This data-driven approach helps users identify trends, forecast future financial needs, and adjust their budgets accordingly.
3. **Seamless Integration**: Modern budget trackers can integrate seamlessly with bank accounts, credit cards, and other financial tools. This integration ensures that users have a comprehensive view of their finances, eliminating the need to manually input data and reducing the risk of errors.
4. **Support for Diverse Financial Goals**: Whether users are looking to save for short-term goals, such as a vacation, or long-term objectives, like retirement, a budget tracker can accommodate a wide range of financial aspirations. By providing tailored advice and tracking progress, it helps users stay on track to meet their goals.
5. **Accessibility and Convenience**: Available on various platforms, including mobile devices, budget trackers offer unparalleled convenience. Users can access their financial information anytime, anywhere, making it easier to stay on top of their finances even on the go.