

INCOME ACTIVITIES USING MACHINE LEARNING

AN INDUSTRY ORIENTED MINI REPORT

Submitted to

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY, HYDERABAD

In partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING (CSE)

Submitted By

MOHAMMED ASIF

ADEPU AMULYA

UDUTHA AJAY KUMAR

SYED ABULLAIS

21UK1A05F3

21UK1A05J8

21UK1A05E2

21UK1A05D3

Under the guidance of

B.Rajitha

Assistant Professor



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VAAGDEVI ENGINEERING COLLEGE

Affiliated to JNTUH, HYDERABAD
BOLLIKUNTA, WARANGAL (T.S) – 506005

DEPARTMENT OF

COMPUTER SCIENCE AND ENGINEERING (CSE)

VAAGDEVI ENGINEERING COLLEGE(WARANGAL)



CERTIFICATE OF COMPLETION

INDUSTRY ORIENTED MINI PROJECT

➤ This is to certify that the UG Project Phase-1 entitled “INCOME ACTIVITIES USING MACHINE LEARNING” is being submitted by MOHAMMED ASIF(21UK1A05F3), ADEPU AMULYA(21UK1A05J8), UDUTHA AJAY KUMAR(21UK1A05E2), SYED ABULLAIS(21UK1A05D3), in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science & Engineering to Jawaharlal Nehru Technological University Hyderabad during the academic year 2024- 2025.

Project Guide

B.Rajitha

(Assistant Professor)

HOD

Dr.R.Naveen kumar Rangaraju

(Professor)

External

ACKNOWLEDGEMENT

We wish to take this opportunity to express our sincere gratitude and deep sense of respect to our beloved **Dr.P.PRASAD RAO**, Principal, Vaagdevi Engineering College for making us available all the required assistance and for his support and inspiration to carry out this UG Project Phase-1 in the institute.

We extend our heartfelt thanks to **Dr.NAVEEN KUMAR RANGARAJU**, Head of the Department of CSE, Vaagdevi Engineering college for providing us necessary infrastructure and thereby giving us freedom to carry out the UG project Phase-1.

We express heartfelt thanks to Smart Bridge Educational service Private Limited, for their constant supervision as well as for providing necessary information regarding the UG project Phase-1 and for their support in completing the UG project Phase-1.

We express heartfelt thanks to the guide, **B.RAJITHA** ,Assistant professor, Department of CSE for his constant support and giving necessary guidance for completion of this UG project Phase-1.

Finally, we express our sincere thanks and gratitude to my family members, friends for their encouragement and outpouring their knowledge and experience throughout the thesis.

MOHAMMED ASIF	(21UK1A05F3)
ADEPU AMULYA	(21UK1A05J8)
UDUTHA AJAY KUMAR	(21UK1A05E2)
SYED ABULLAIS	(21UK1A05D3)

ABSTRACT

In recent years, machine learning (ML) techniques have revolutionized various aspects of income-generating activities across industries. This abstract explores the application of ML in enhancing income activities, focusing on key areas such as predictive analytics, automation, and personalized recommendation systems. Predictive analytics enables businesses to forecast trends and optimize strategies, thereby increasing profitability. Automation through ML algorithms streamlines processes, reducing operational costs and enhancing efficiency. Additionally, personalized recommendation systems leverage ML to tailor products and services to individual preferences, thereby boosting customer engagement and sales. This abstract concludes by emphasizing the transformative impact of ML on income activities and its potential for future innovation and growth.

TABLE OF CONTENTS:-

1. INTRODUCTION	5
1.1 OVERVIEW... ..	5
1.2 PURPOSE	5
2. LITERATURE SURVEY	8
2.1 EXISTING PROBLEM	8
2.2 PROPOSED SOLUTION	8-9
3. THEORITICAL ANALYSIS	10
3.1 BLOCK DIAGRAM	10
3.2 HARDWARE /SOFTWARE DESIGNING	10-11
4. EXPERIMENTAL INVESTIGATIONS	12-13
5. FLOWCHART... ..	14
6. RESULTS... ..	15-18
7. ADVANTAGES AND DISADVANTAGES	19
8. APPLICATIONS	20
9. CONCLUSION	20

10. FUTURE SCOPE...	21
11. BIBILOGRAPHY	22-23
12. APPENDIX (SOURCE CODE)&CODE SNIPPETS ...	24-30

1.INTRODUCTION

1.1. OVERVIEW

Machine Learning (ML) has emerged as a pivotal technology reshaping income-generating activities across various sectors. By leveraging vast amount of data, ML Algorithms can extract valuable insights and drive actionable decisions, ultimately enhancing profitability and efficiency. Here's an overview of how ML is transforming income activities. In conclusion, the integration of machine learning in income activities offers significant advantages by enabling data-driven decision-making, enhancing operational efficiency, and improving customer satisfaction. As businesses continue to adopt and innovate with ML technologies, the landscape of income generation is expected to evolve further, driving sustainable growth and competitive advantage in the global market. This overview outlines how machine learning technologies are transforming income activities across sectors, highlighting key benefits and applications.

1.2. PURPOSE

The purpose of incorporating machine learning (ML) into income activities is multifaceted and aims to achieve several critical objectives:

1.Enhanced Decision Making: ML enables businesses to make data-driven decisions by analysing large volumes of structured and unstructured data. This capability helps in optimizing pricing strategies, forecasting demand, and identifying opportunities for cost reduction, thereby increasing profitability.

2.Improved Efficiency and Productivity: Automation through ML algorithms streamlines processes, reduces manual intervention, and enhances operational efficiency. Tasks such as data entry, customer support, and inventory management can be automated, allowing employees to focus on more strategic initiatives.

3.Personalization and Customer Satisfaction: ML-powered recommendation systems analyse customer behaviour and preferences to deliver personalized experiences. By

offering tailored product recommendations and targeted marketing campaigns, businesses can enhance customer satisfaction, increase retention rates, and drive higher sales.

4.Risk Management and Fraud Detection: ML algorithms can detect anomalies and patterns indicative of fraudulent activities in real-time. This capability helps businesses mitigate financial risks, prevent fraud losses, and protect their assets.

5.Competitive Advantage: ML provides insights into market trends, competitor strategies, and consumer behavior, enabling businesses to stay ahead of the competition. By leveraging these insights, organizations can innovate faster, adapt to changing market conditions, and capitalize on emerging opportunities.

6.Scalability and Adaptability: ML models are scalable and adaptable to evolving business needs and market dynamics. They can handle large datasets and continuously learn from new data, ensuring that insights and recommendations remain relevant and accurate over time.

7.Cost Optimization: ML-driven predictive analytics helps businesses optimize resource allocation, inventory management, and supply chain operations. By reducing waste, improving efficiency, and minimizing operational costs, businesses can improve their profit margins and financial sustainability.

2 .LITERATURE SURVEY

2.1 EXISTING PROBLEM

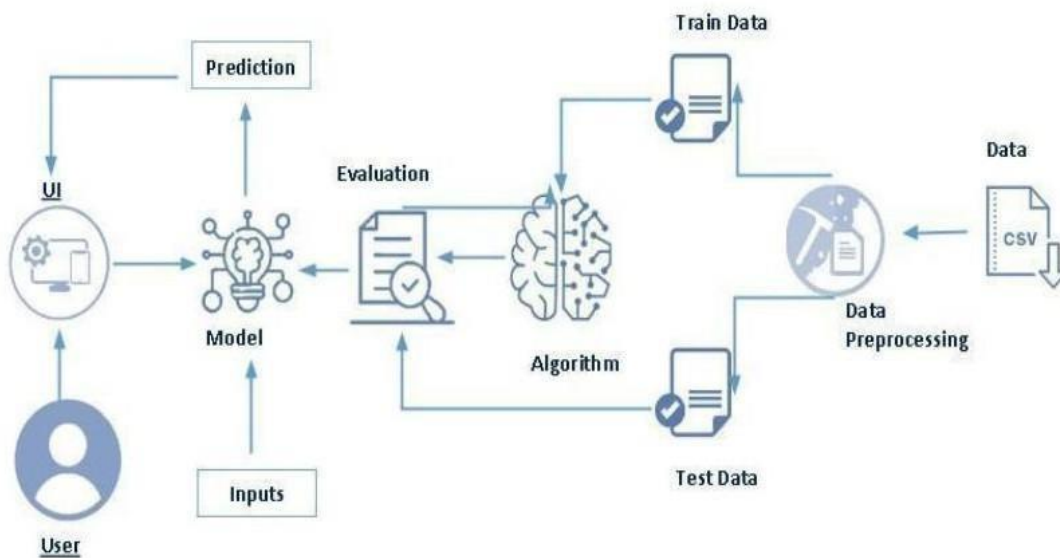
- ML models require high-quality, relevant data to make accurate predictions and decisions. However, data may be incomplete, noisy, or biased, impacting the reliability of ML outcomes.
- ML models, especially complex ones like deep neural networks, are often perceived as "black boxes," making it challenging to understand how decisions are made.
- ML algorithms can inadvertently learn biases from training data, leading to unfair outcomes across demographic groups.
- ML models often require access to sensitive personal or financial data, raising concerns about data privacy breaches and regulatory compliance.

2.2 PROPOSED SOLUTION

- Implement data preprocessing techniques such as cleaning, normalization, and feature engineering to improve data quality. Use techniques like data augmentation or synthetic data generation to supplement limited datasets. Employ bias detection and mitigation strategies to ensure fairness in model predictions.
Utilize interpretable ML techniques such as decision trees, rule-based models, or model-agnostic methods like SHAP (Shapley Additive explanations) to provide explanations for model predictions. Incorporate transparency frameworks that document model architecture, data used, and reasoning behind predictions to enhance trust and accountability.
 - Conduct thorough bias analysis during model development using fairness metrics. Implement techniques such as fairness-aware learning algorithms, data augmentation with diverse samples, and bias mitigation strategies like reweighing or adversarial debiasing. Regularly audit and monitor models for fairness to mitigate unintended biases. Employ privacy-preserving techniques such as differential privacy, federated learning, or homomorphic encryption to protect sensitive data during model training and inference. Implement robust data anonymization and access control measures to minimize the risk of data breaches. Ensure compliance with relevant data protection regulations and industry standards.
- By implementing these proposed solutions, organizations can effectively harness the power of ML in income activities while mitigating risks, ensuring fairness, and maintaining compliance with regulatory standards. Continuous monitoring, evaluation, and adaptation of ML strategies are essential to drive sustainable growth and innovation in today's competitive landscape.

3.THEORITICAL ANALYSIS

3.1. BLOCK DIAGRAM



3.2. SOFTWARE DESIGNING

The following is the Software required to complete this project:
Designing software for income-related activities involves several steps. Here's a broad outline of how you might approach this:

1. Requirement Analysis :

Identify the users: Who will be using the software? (e.g., freelancers, small business owners, large enterprises)

Define the scope: What specific income activities will the software manage?
(e.g.,invoicing, expense tracking, budgeting, financial reporting)

Gather requirements: Talk to potential users, stakeholders, and domain experts to understand their needs and pain points.

2. Planning and Design :

Architecture Design: Decide on the software architecture (e.g., client-server, microservices) and technology stack (e.g., programming languages, frameworks).

UI/UX Design: Create wireframes and mockups to design a user-friendly interface. Ensure ease of use and accessibility.

Database Design: Design the database schema to store and manage data efficiently (e.g., users, transactions, invoices).

3. Development :

Backend Development: Develop the server-side logic, APIs, and database interactions.

Frontend Development: Implement the user interface using appropriate web or mobile technologies.

Integration: Ensure seamless communication between the frontend and backend.

4. Testing :

Unit Testing: Test individual components to ensure they work as intended.

Integration Testing: Verify that different modules work together correctly.

User Acceptance Testing (UAT): Have actual users test the software to ensure it meets their needs.

5. Deployment :

Hosting: Choose a hosting platform (e.g., cloud services like AWS, Azure).

Deployment Pipeline: Set up continuous integration and deployment pipelines to streamline updates and maintenance.

6. Maintenance and Support :

Monitor Performance: Use monitoring tools to track performance and user activity.

Regular Updates: Continuously update the software to fix bugs, add features, and improve security.

User Support: Provide customer support channels to assist users with any issues.

Key Features to Consider

User Authentication: Secure login and user management.

Invoicing: Create, send, and track invoices.

Expense Tracking: Record and categorize expenses.

Financial Reporting: Generate reports to provide insights into income and expenses.

Integration with Payment Gateways: Enable users to receive payments online.

Data Security: Implement robust security measures to protect sensitive financial data.

Tools and Technologies:

Frontend: HTML, CSS, JavaScript, React, Angular, Vue.js

Backend: Node.js, Django, Flask, Ruby on Rails

Database: MySQL, PostgreSQL, MongoDB

APIs: REST, GraphQL

Hosting: AWS, Azure, Google Cloud

CI/CD: Jenkins, GitHubActions, CircleCI

3. EXPERIMENTAL INVESTIGATION

In this project, we have used Air Quality Dataset. This dataset is a csv file consisting of labelled data and having the following columns-

City: Location or urban area for which income activities is recorded.

Date: The specific date on which income activities were taken.

Initial steps involve creating a subject profile, including details such as place and date of birth, family information, contact details, and aliases.

Various intergovernmental bodies and organizations, including Financial Action Task Force (FATF) reports, Beneficial Ownership Guides, country reports, and Asset Recovery Guides, are used in the investigation.

Additionally, non-governmental organizations such as Global Financial Integrity and Global Witness provide invaluable insights for deeper analysis.

Previous information is then shared with various national and international jurisdiction links, facilitating collaboration with foreign counterparts.

Diverse data, including employment and academic history, criminal records, and real estate documents, is collected.

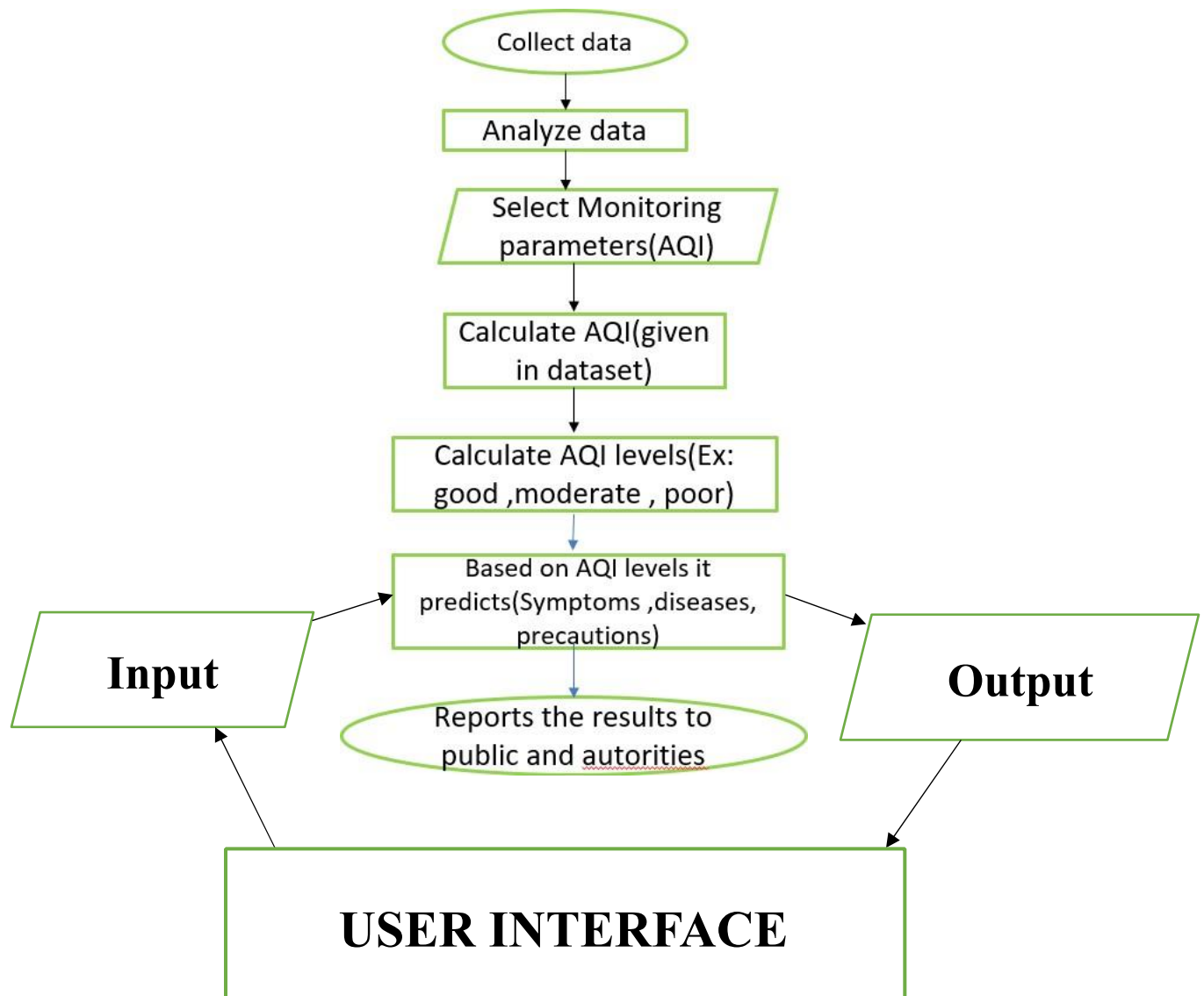
All financial information regarding assets, income, liabilities, and expenses is collected from different sources like law firms, real estate agents, government agencies, banks, and other financial agencies.

Evidence collection is crucial for initiating legal proceedings. It involves gathering relevant information and financial data related to identified targets to formulate a comprehensive investigation strategy.

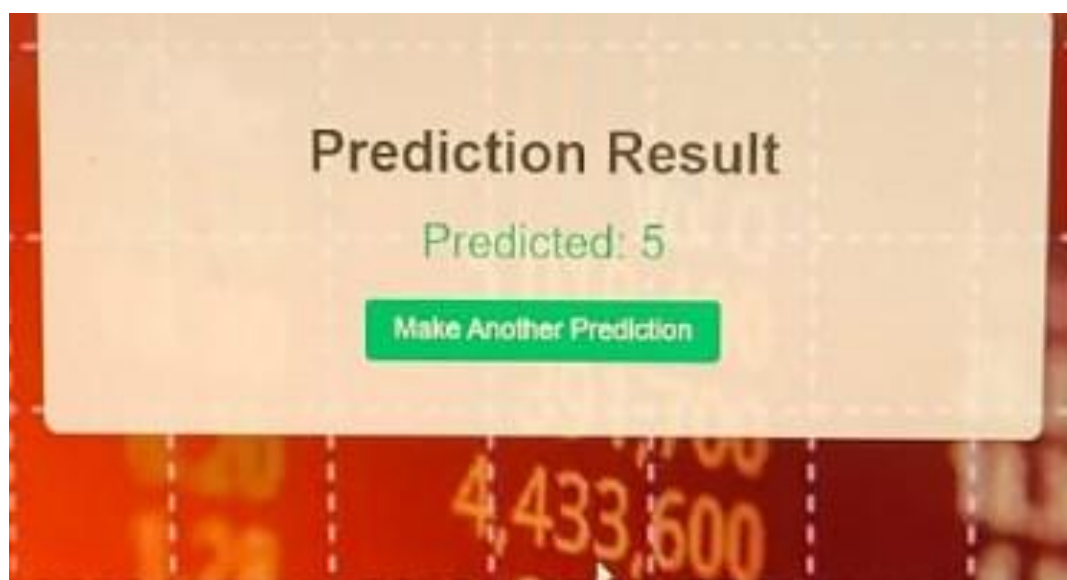
Domestic and foreign accounts are meticulously searched through document production, physical surveillance, and search warrants at businesses and residences.

Asset identification, case assessment, and subject profiling require the integration of social media and the Internet. Various open sources of information, such as civil litigation, patent applications, and publicly reported financial investigations, are used.

5. FLOWCHART



6.RESULT



Enter Details

Country Name:

Region Name:

Income Category:

Income Year:

Income Value:

Income Data Source:

Predict

14 July 2024

7.ADVANTAGES AND DISADVANTAGES

ADVANTAGES:

1. Enhanced Public Health Advantages:
2. Provides financial stability and regular income.
3. Offers the potential for career advancement and skill development.
4. Can be used to qualify for various financial products and services..

DISADVANTAGES:

Software designing of income activities .Designing software for income-related activities involves several stepsHere's a broad outline of how you might approach this:

1 .Requirement Analysis

Identify the users: Who will be using the software? (e.g., freelancers, small business owners, large enterprises)

Define the scope: What specific income activities will the software manage? (e.g., invoicing, expense tracking, budgeting, financial reporting)

Gather requirements: Talk to potential users, stakeholders, and domain experts to understand their needs and pain points.

2 .Planning and Design: **Architecture Design:** Decide on the software architecture (e.g., client-server, micro services) and technology stack (e.g., programming languages, frameworks). **UI/UX Design:** Create wireframes and mock ups to design a user-friendly interface. Ensure ease of use and accessibility.

Database Design: Design the database schema to store and manage data efficiently (e.g., users, transactions, invoices)

3. Development

Backend Development: Develop the server-side logic, APIs, and database interactions.

Frontend Development: Implement the user interface using appropriate web or mobile technologies.

Integration: Ensure seamless communication between the frontend and backend.

Unit Testing: Test individual components to ensure they work as intended.

Integration Testing: Verify that different modules work together correctly.

User Acceptance Testing (UAT): Have actual users test the software to ensure it meets their needs.

5. Deployment

Hosting: Choose a hosting platform (e.g., cloud services like AWS, Azure).

Deployment Pipeline: Set up continuous integration and deployment pipelines to streamline updates and maintenance.

6. Maintenance and Support

Monitor Performance: Use monitoring tools to track performance and user activity.

Regular Updates: Continuously update the software to fix bugs, add features, and improve security.

User Support: Provide customer support channels to assist users with any issues.

Key Features to Consider:

User Authentication: Secure login and user management.

Invoicing: Create, send, and track invoices.

Expense Tracking: Record and categorize expenses.

Financial Reporting: Generate reports to provide insights into income and expenses.

Integration with Payment Gateways: Enable users to receive payments online.

Data Security: Implement robust security measures to protect sensitive financial data.

Tools and Technologies

Frontend: HTML, CSS, JavaScript, React, Angular, Vue.js

Backend: Node.js, Django, Flask, Ruby on Rails

Database: MySQL, PostgreSQL, MongoDB

APIs: REST, GraphQL

Hosting: AWS, Azure, Google Cloud

CI/CD: Jenkins, GitHub Actions, CircleCI

Would you like more detailed information on any of these steps or features?

Advantages of income activities : Income-generating activities provide several advantages, both for individuals and organizations. Here are some key benefits:

For Individuals:

Financial Stability: Regular income from these activities can provide financial security and stability, helping individuals meet their day-to-day expenses and save for the future.

Skill Development: Engaging in income-generating activities often requires developing new skills or honing existing ones, which can enhance personal and professional growth.

Entrepreneurial Opportunities: For those who start their own businesses or freelance, income activities can provide a path to entrepreneurship and self-employment, offering greater independence and control over one's work life.

Diversified Income Streams: By engaging in multiple income activities, individuals can diversify their income sources, reducing reliance on a single job or source of income and mitigating financial risks.

Retirement Planning: Regular income allows individuals to contribute to retirement savings, ensuring a more comfortable and secure retirement.

For Organizations:

Revenue Generation: Income-generating activities are essential for businesses to generate revenue, ensuring their operations can be sustained and expanded.

Business Growth: Successful income activities can fuel growth by providing the necessary capital for investment in new projects, technology, and human resources.

Market Presence: Engaging in diverse income activities can help a business establish a strong market presence and brand recognition, attracting more customers and opportunities.

Innovation and Development: Income-generating activities often require innovation and continuous improvement, driving organizations to develop new products, services, and solutions.

Employee Benefits: Increased revenue from income activities allows businesses to offer better salaries, benefits, and career development opportunities, helping to attract and retain top talent.

Social Impact: Successful businesses can contribute to the economy by creating jobs, supporting local communities, and engaging in corporate social responsibility initiatives.

General Benefits:

Economic Contribution: Both individuals and organizations contribute to the economy through their income activities, fostering economic growth and development.

Improved Quality of Life: Increased income can lead to a better quality of life, enabling access to better education, healthcare, and living conditions.

Financial Independence: Regular income allows individuals and businesses to be financially independent, reducing reliance on loans and external financial support.

Resilience: Diverse income activities can provide a buffer against economic downturns, as reliance on multiple income sources can mitigate the impact of financial crises.

Income-generating activities are crucial for personal development, business growth, and overall economic health, offering numerous advantages across different aspects of life and society.

Disadvantages of income activities

while income-generating activities offer many advantages, they also come with certain disadvantages that can impact individuals and organizations. Here are some key drawbacks:

For Individuals:

Time and Stress Management: Managing multiple income activities can be time consuming and stressful, leading to burnout and negatively impacting mental and physical health.

Financial Risk: Not all income-generating activities are guaranteed to be successful. Investments in new ventures can result in financial losses if the activities fail to generate expected returns.

Work-Life Balance: Engaging in multiple income activities can interfere with personal time, affecting relationships and overall quality of life.

Unstable Income: Freelancing or self-employment can lead to irregular income streams, making financial planning and stability more challenging.

Skill Mismatch: Individuals may have to engage in activities that do not align with their skills or interests, leading to job dissatisfaction and reduced productivity.

8.APPLICATIONS

1. **Public-Private Partnerships:** Collaborating with private enterprises to develop infrastructure or services that generate income and benefit the community.
2. **Economic Development Programs:** Implementing initiatives to promote entrepreneurship, attract investment, and create job opportunities.
3. **Tourism Promotion:** Developing and promoting tourism attractions to generate income for local economies.
4. **Renewable Energy Projects:** Investing in renewable energy sources such as solar or wind power to create sustainable income streams.
5. **Cultural and Creative Industries:** Supporting arts, culture, and creative industries to generate revenue and preserve cultural heritage.

9 .CONCLUSION

In conclusion, the integration of machine learning into income-generating activities can drive significant growth and efficiency improvements, making it a valuable tool for businesses looking to enhance their profitability and competitiveness. Machine learning (ML) has the potential to significantly enhance income-generating activities by optimizing processes, reducing costs, and uncovering new opportunities. In theory, the application of ML to income activities can lead to the enhanced Decision-Making, Personalization, Operational Efficiency, Predictive Analytics, Market Segmentation, Fraud Detection, Innovation.

10. FUTURE SCOPE

The integration of machine learning into income-generating activities is likely to continue growing, driven by advancements in technology and increased data availability. Future applications may include more sophisticated predictive models, real-time financial analytics, and even greater automation of business processes. As ML technologies become more accessible, a broader range of users will benefit from their capabilities, further democratizing opportunities for income generation. By embracing machine learning, individuals and organizations can enhance their income-generating efforts, drive economic growth, and improve overall quality of life. The potential is vast, and the journey towards fully realizing these benefits is just beginning.

11. BIBLIOGRAPHY

Creating a bibliography for income activities using machine learning involves citing various sources that discuss the application of machine learning in generating income, whether through financial trading, business optimization, marketing strategies, or other innovative uses.

Here are some key references that you can consider for your bibliography:

Books

1. "Machine Learning for Asset Managers" by Marcos López de Prado This book explores the application of machine learning techniques in financial markets and asset management.
2. "Machine Learning: A Probabilistic Perspective" by Kevin P. Murphy Provides a comprehensive introduction to the field of machine learning, including its application in different income-generating activities.
3. "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Geron A practical guide that discusses various machine learning techniques and their applications, including business and marketing strategies.

Articles and Papers

1. "Financial Applications of Machine Learning" by Ilya Kuntsevich and Malcom C. Baker
This paper discusses how machine learning algorithms are applied in finance for risk management, trading, and investment strategies.
2. "Machine Learning in Marketing" by Werner Reinartz and V. Kumar A comprehensive overview of how machine learning techniques are transforming marketing strategies and improving customer targeting.
3. "Machine Learning for Predictive Maintenance" by R. Liu and M. Yang This paper highlights the application of machine learning in industrial settings for predictive maintenance, which can optimize operations and reduce costs.

Websites and Online Resources

1. Towards Data Science - A popular Medium publication with numerous articles on practical applications of machine learning in business and finance.
Example article: "How Machine Learning is Revolutionizing the Stock Market" [link] (<https://towardsdatascience.com/machine-learning-stock-market-forecasting-59b41dd11a8>)
2. Kaggle - A platform for data science competitions, which often feature projects related to financial modelling, marketing strategies, and other income-generating activities.
Example project: "Predict Future Sales" [link] (<https://www.kaggle.com/c/competitive-data-sciencepredict-future-sales>)
3. Data Camp - Offers courses and tutorials on machine learning applications in various fields, including finance and marketing.
- Example course: "Machine Learning for Time Series Data in Python"
[link] (<https://www.datacamp.com/courses/machine-learning-for-time-series-data-in-python>)

Journals

1. Journal of Financial Economics

Regularly publishes papers on the application of machine learning in finance, such as algorithmic trading and risk assessment.

2. Marketing Science

Features research on how machine learning can be utilized for customer segmentation, targeting, and marketing optimization.

12.APPENDIX

Model Building :

- 1)Dataset
- 2)Google colab and VS code Application Building
 - 1.Html file(Index file, predict file)
 - 1.CSS file
 - 2.Models in pickle format

SOURCE CODE :

INDEX.HTML

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <title>Welcome to Income Activities</title>
  <style>
    body {
      background-image:
url('https://t4.ftcdn.net/jpg/02/62/17/49/360_F_262174953_wob2JN075TWTPT
QbHNFr8AAJNM8K3Zgp.jpg');
      background-size: cover;
      font-family: Arial, sans-serif;
      color: #ded0d0;
    }
    .container {
      text-align: center;
      padding: 50px;
```

```

    }
    h1 {
        font-size: 3em;
        margin-bottom: 20px;
    }
    .btn {
        display: inline-block;
        padding: 10px 20px;
        font-size: 1.2em;
        color: white;
        background-color: #4CAF50;
        border: none;
        border-radius: 5px;
        text-decoration: none;
    }
</style>
</head>
<body>
    <div class="container">
        <h1>Welcome to Income Activities</h1>
        <a class="btn" href="/prediction">Make a Prediction</a>
    </div>
</body>
</html>

```

PREDICT.HTML

```

<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <title>Income Activities Form</title>

```



```

<style>
  body {
    background-image:
url('https://media.istockphoto.com/id/1932356331/vector/stressed-investor-
businessman-looking-at-falling-graph-forecast-or-predict-the-next-
economic.jpg?s=612x612&w=0&k=20&c=sD2iR-H82yrBqSbS-
aPN1M8GhnpsbQuUwgU135-jF-U=');
    background-size: cover;
    font-family: Arial, sans-serif;
    color: #333;
  }
  .container {
    max-width: 500px;
    margin: auto;
    padding: 20px;
    background-color: rgba(255, 255, 255, 0.8);
    border-radius: 10px;
    box-shadow: 0 0 10px rgba(0, 0, 0, 0.1);
  }
  h1 {
    text-align: center;
    margin-bottom: 20px;
  }
  form {
    display: flex;
    flex-direction: column;
  }
  label {
    margin-bottom: 5px;
  }
  input, select {
    margin-bottom: 15px;
    padding: 10px;
  }

```

```

    font-size: 1em;
    border: 1px solid #ccc;
    border-radius: 5px;
}
.btn {
    padding: 10px;
    font-size: 1.2em;
    color: white;
    background-color: #4CAF50;
    border: none;
    border-radius: 5px;
    cursor: pointer;
}
.btn:hover {
    background-color: #45a049;
}
</style>
</head>
<body>
    <div class="container">
        <h1>Enter Details</h1>
        <form action="/submit" method="POST">
            <label for="ADM0_NAME">Country Name:</label>
            <input type="text" id="ADM0_NAME" name="ADM0_NAME"
required>

            <label for="ADM1_NAME">Region Name:</label>
            <input type="text" id="ADM1_NAME" name="ADM1_NAME"
required>

            <label for="Income_Category">Income Category:</label>
            <input type="number" id="Income_Category"
name="Income_Category" required>

```

```

        <label for="Income_Year">Income Year:</label>
        <input type="number" id="Income_Year" name="Income_Year"
required>

        <label for="IncomeValue">Income Value:</label>
        <input type="number" id="IncomeValue" name="IncomeValue"
required>

        <label for="Income_DataSource">Income Data Source:</label>
        <input type="text" id="Income_DataSource"
name="Income_DataSource" required>

        <button type="submit" class="btn">Predict</button>
    </form>
</div>
</body>
</html>

```

APP.PY

```

from flask import Flask, request, render_template
import pickle
import numpy as np
import pandas as pd

app = Flask(__name__)
model = pickle.load(open('model.pkl', 'rb'))

@app.route('/')
def index():
    return render_template("index.html")

```

```

@app.route('/prediction')
def prediction():
    return render_template("prediction.html")

@app.route('/submit', methods=["POST"])
def submit():
    # Reading the inputs given by the user
    form_values = request.form.to_dict()
    input_feature = [
        form_values['ADM0_NAME'],
        form_values['ADM1_NAME'],
        int(form_values['Income_Category']),
        int(form_values['Income_Year']),
        int(form_values['IncomeValue']),
        form_values['Income_DataSource']
    ]

    # Creating the DataFrame
    data = pd.DataFrame([input_feature], columns=[
        'ADM0_NAME', 'ADM1_NAME',
        'Income_Category', 'Income_Year', 'IncomeValue',
        'Income_DataSource'
    ])

    # Making predictions using the loaded model
    prediction = model.predict(data)
    prediction = round(int(prediction[0]))

    return render_template("result.html", result=prediction)

if __name__ == "__main__":
    app.run(debug=True, host='0.0.0.0')

```

CODE SNIPPETS

MODEL BUILDING

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelEncoder
from sklearn.linear_model import LinearRegression
from sklearn.neighbors import KNeighborsRegressor
from sklearn.ensemble import RandomForestRegressor, AdaBoostRegressor, GradientBoostingRegressor
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error, r2_score
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, make_scorer
from sklearn.model_selection import GridSearchCV
import pickle
import warnings
warnings.filterwarnings('ignore')
```

```
[ ] df = pd.read_csv('/content/income.csv', encoding='latin-1') # Try reading the file with 'latin-1' encoding.

[ ] df.head()
```

df.head()

	ADM0_NAME	ADM1_NAME	ADM2_NAME	Income_Category	Income_SubCategory	Income_Month	Income_Year	IncomeValue	Indicator_Type	Income_DataSource
0	Armenia	Aragatsotn	NaN	Agro - Food&Cash Crop	0.0	8	2009	33.592162	Baseline	WFP VAM Impact of the Global Financial Crisis ...
1	Armenia	Aragatsotn	NaN	Asset Sales/Savings/Credit	0.0	8	2009	0.301408	Baseline	WFP VAM Impact of the Global Financial Crisis ...
2	Armenia	Aragatsotn	NaN	Commerce - Business/Trading	0.0	8	2009	3.810976	Baseline	WFP VAM Impact of the Global Financial Crisis ...
3	Armenia	Aragatsotn	NaN	Labor - Artisan/Skilled	0.0	8	2009	9.058316	Baseline	WFP VAM Impact of the Global Financial Crisis ...
4	Armenia	Aragatsotn	NaN	Labor - Daily/Unskilled - Mixed	0.0	8	2009	10.475054	Baseline	WFP VAM Impact of the Global Financial Crisis ...

df.tail()

	ADM0_NAME	ADM1_NAME	ADM2_NAME	Income_Category	Income_SubCategory	Income_Month	Income_Year	IncomeValue	Indicator_Type	Income_DataSource
4365	Yemen	NaN	NaN	Labor - Salary/Regular	0.0	12	2011	33.510186	Baseline	WFP VAM The State of Food Security and Nutriti...
4366	Yemen	NaN	NaN	Non-Wage - Aid/Welfare	0.0	12	2011	2.537681	Baseline	WFP VAM The State of Food Security and Nutriti...
4367	Yemen	NaN	NaN	Non-Wage - Pension/Benefit	0.0	12	2011	2.047048	Baseline	WFP VAM The State of Food Security and Nutriti...
4368	Yemen	NaN	NaN	Non-Wage - Remittance/Fam. Aid	0.0	12	2011	8.875287	Baseline	WFP VAM The State of Food Security and Nutriti...
4369	Yemen	NaN	NaN	Other	0.0	12	2011	2.077772	Baseline	WFP VAM The State of Food Security and Nutriti...

```

qrlbe: Դրեզ
tucow6 Դրեզոուլո 0
tucw1c9rou_1lbe 0
tucow6A9jne 0
tucow6_169u 0
tucow6_10u7u 0
tucow6_20p9c9e8ouL 1880
tucow6_29e8ouL 0
VDM5_1NME 3041
VDM7_1NME 398
VDM9_1NME 0

```

```

q4.72u9().շոմ()
[ ] ևսրեզյոն արդյուն արդյուն

```

```

m6ouL n9Re: 347.2+ KB
qrlbe: Էյոթեզ(5), Էյեզ(5), օրթեզ(0)
0 tucow6 Դրեզոուլո 4330 ուս-ույյ օրթեզ
8 tucw1c9rou_1lbe 4330 ուս-ույյ օրթեզ
1 tucow6A9jne 4330 ուս-ույյ Էյոթեզ
2 tucow6_169u 4330 ուս-ույյ Էյեզ
3 tucow6_10u7u 4330 ուս-ույյ Էյեզ
4 tucow6_20p9c9e8ouL 3984 ուս-ույյ Էյոթեզ
5 tucow6_29e8ouL 4330 ուս-ույյ օրթեզ
6 VDM5_1NME 4330 ուս-ույյ օրթեզ
7 VDM7_1NME 4033 ուս-ույյ օրթեզ
8 VDM9_1NME 4330 ուս-ույյ օրթեզ
---
# Եյոթու ուս-ույյ օրթեզ
0949 Եյոթու (Եյոթ 10 Եյոթու):
0906Եյոթ: 4330 Եյոթու, 0 10 4300
<Եյոթ, Եյոթու, Եյոթ, Եյոթու, Եյոթու, Եյոթու>

```

```

q4.72u9()
#ևսրեզյոն արդյուն արդյուն

```

```

[ ] #removing missing values
df.dropna(inplace=True)

```

```

[ ] df.isna().sum()

```

```

ADM0_NAME 0
ADM1_NAME 0
ADM2_NAME 0
Income_Category 0
Income_SubCategory 0
Income_Month 0
Income_Year 0
IncomeValue 0
Indicator_Type 0
Income_DataSource 0
dtype: int64

```

```

df.isnull().sum()

```

```

ADM0_NAME 0
ADM1_NAME 0
ADM2_NAME 0
Income_Category 0
Income_SubCategory 0
Income_Month 0
Income_Year 0
IncomeValue 0
Indicator_Type 0
Income_DataSource 0
dtype: int64

```

```

import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Function to remove outliers
def remove_outliers(df, column):
    Q1 = df[column].quantile(0.25)
    Q3 = df[column].quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5 * IQR
    upper_bound = Q3 + 1.5 * IQR
    return df[(df[column] >= lower_bound) & (df[column] <= upper_bound)]

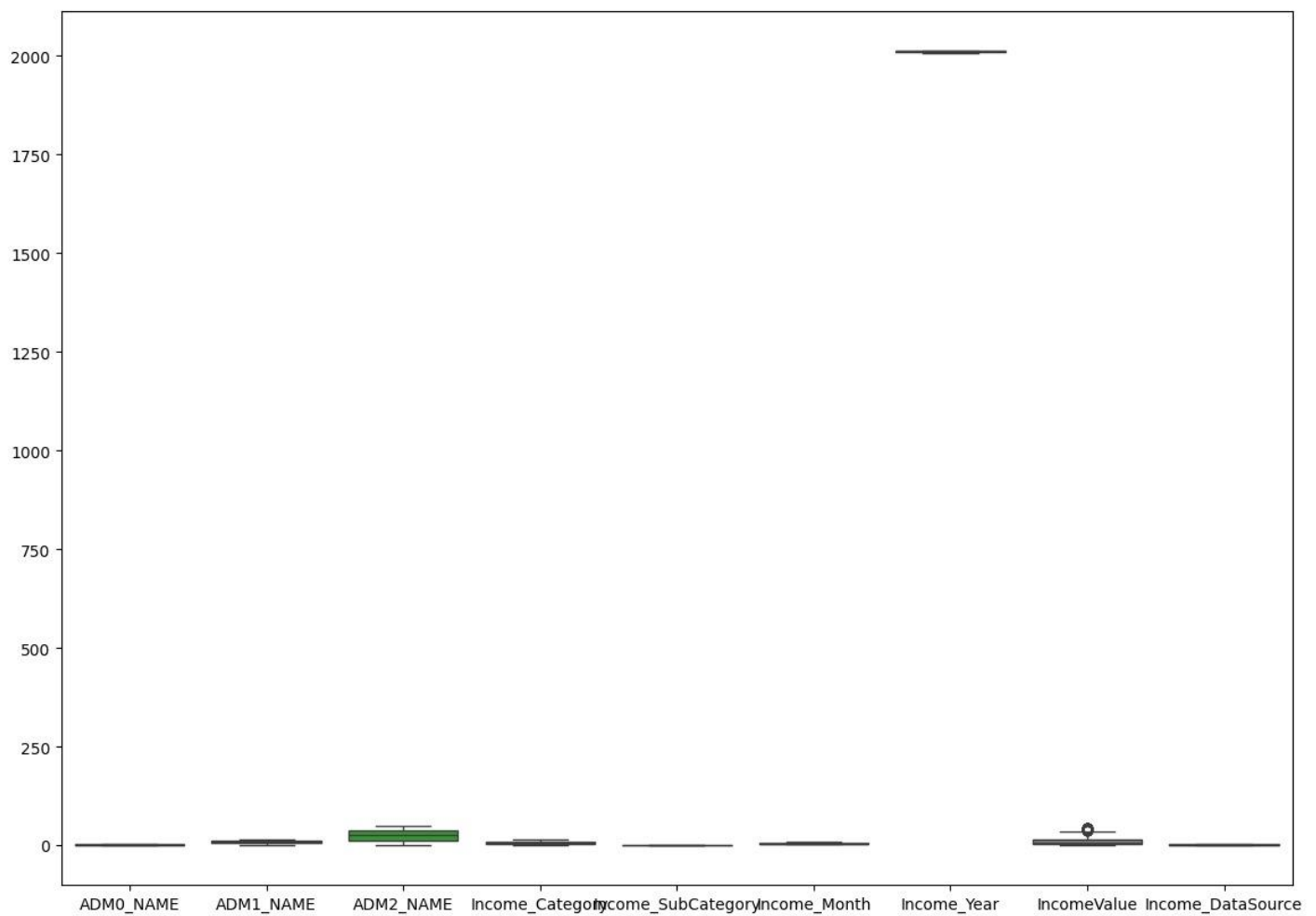
# Apply the remove_outliers function to all numeric columns
numeric_columns = df.select_dtypes(include=['float64', 'int64']).columns
for column in numeric_columns:
    df = remove_outliers(df, column)

# Increase the size of the plot
plt.figure(figsize=(14, 10))

# Plot the boxplot without outliers
sns.boxplot(data=df)

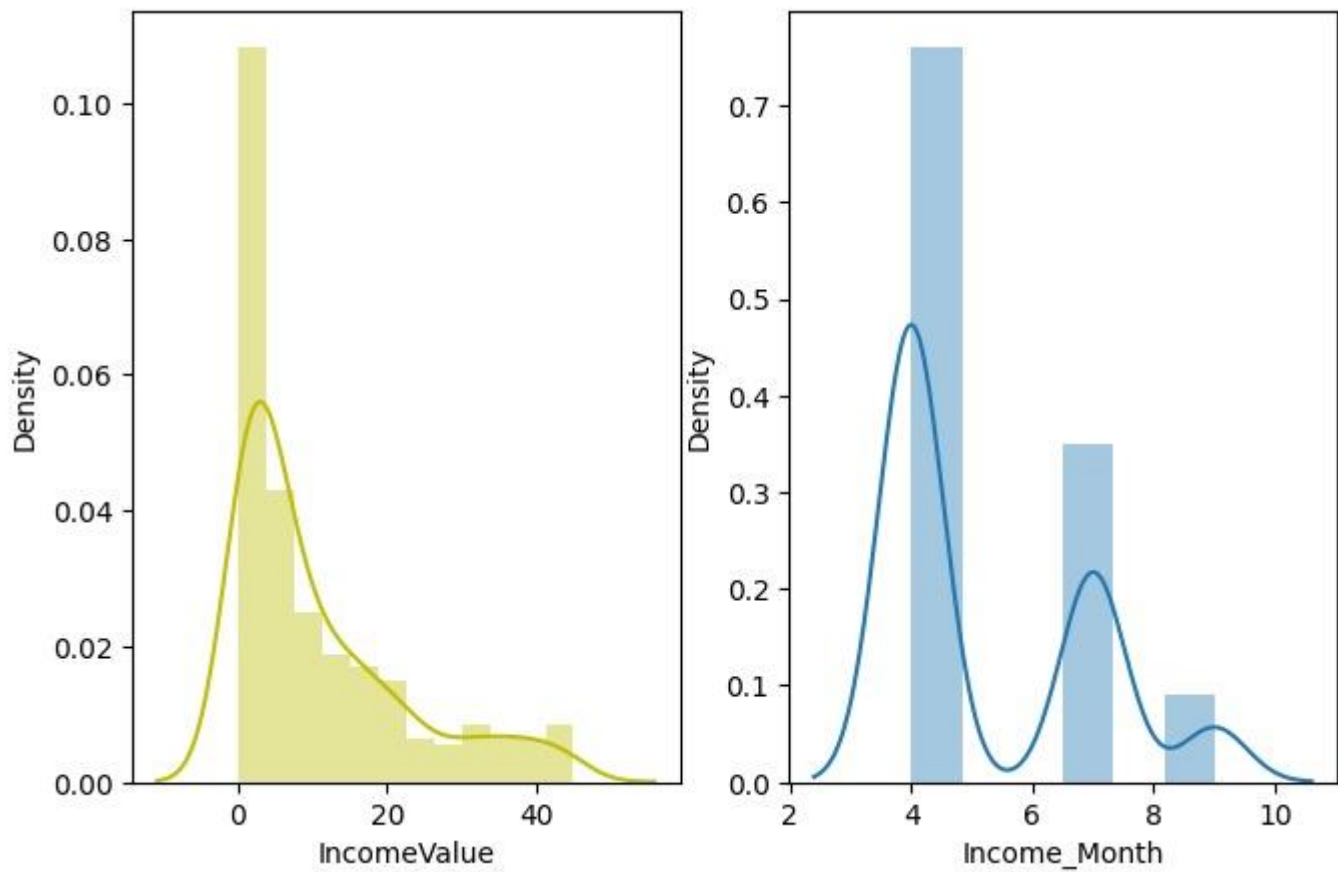
# Show the plot
plt.show()

```

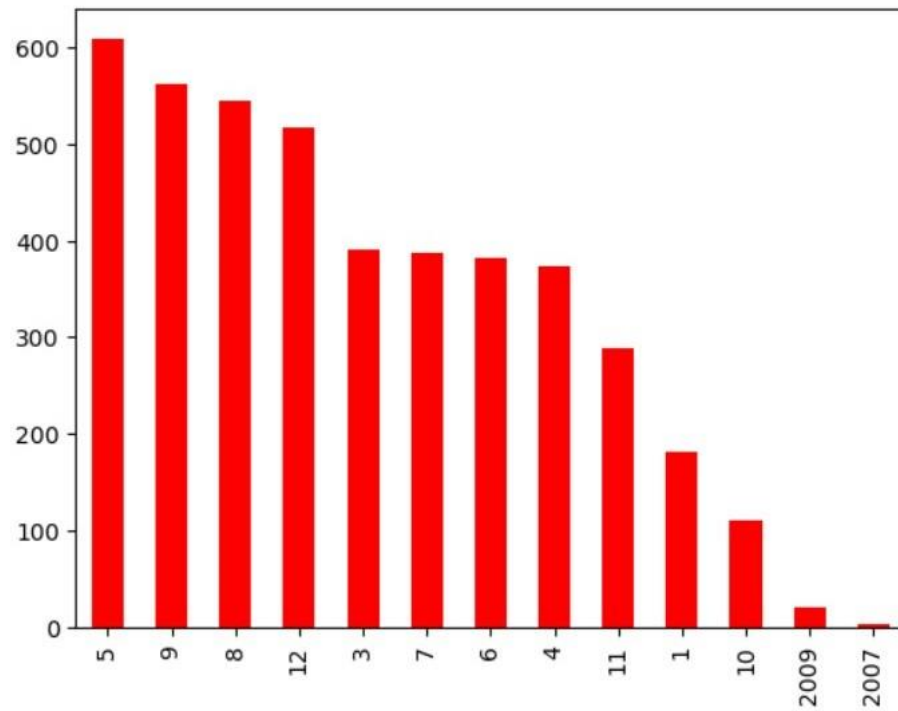


DATA VISUALIZATION

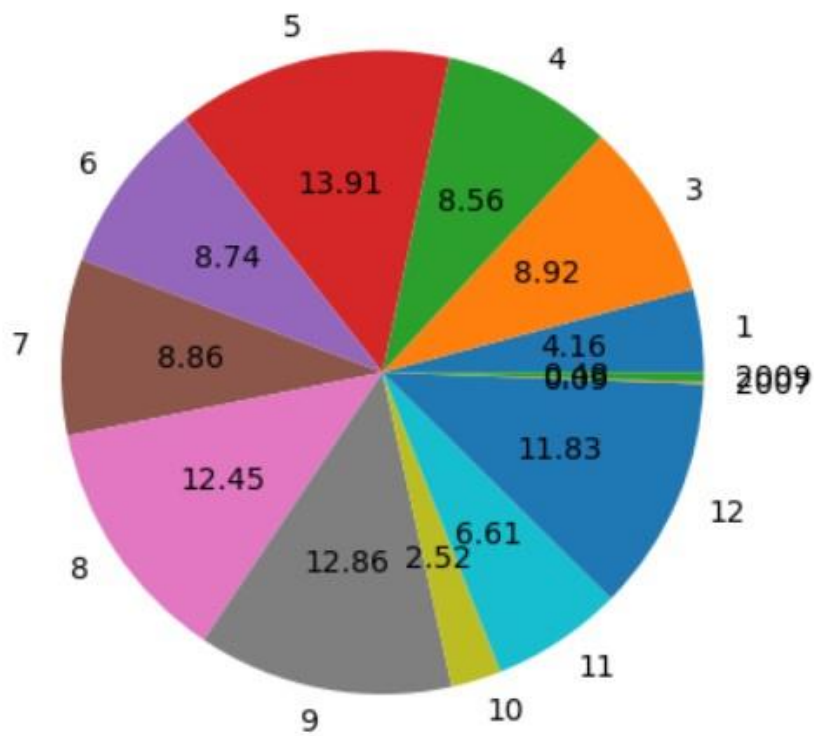
```
plt.figure(figsize=(12,5))
plt.subplot(131)
sns.distplot(df['IncomeValue'],color="y")
plt.subplot(132)
sns.distplot(df['Income_Month'])
plt.show()
```




```
df['Income_Month'].value_counts().plot(kind='bar')
```



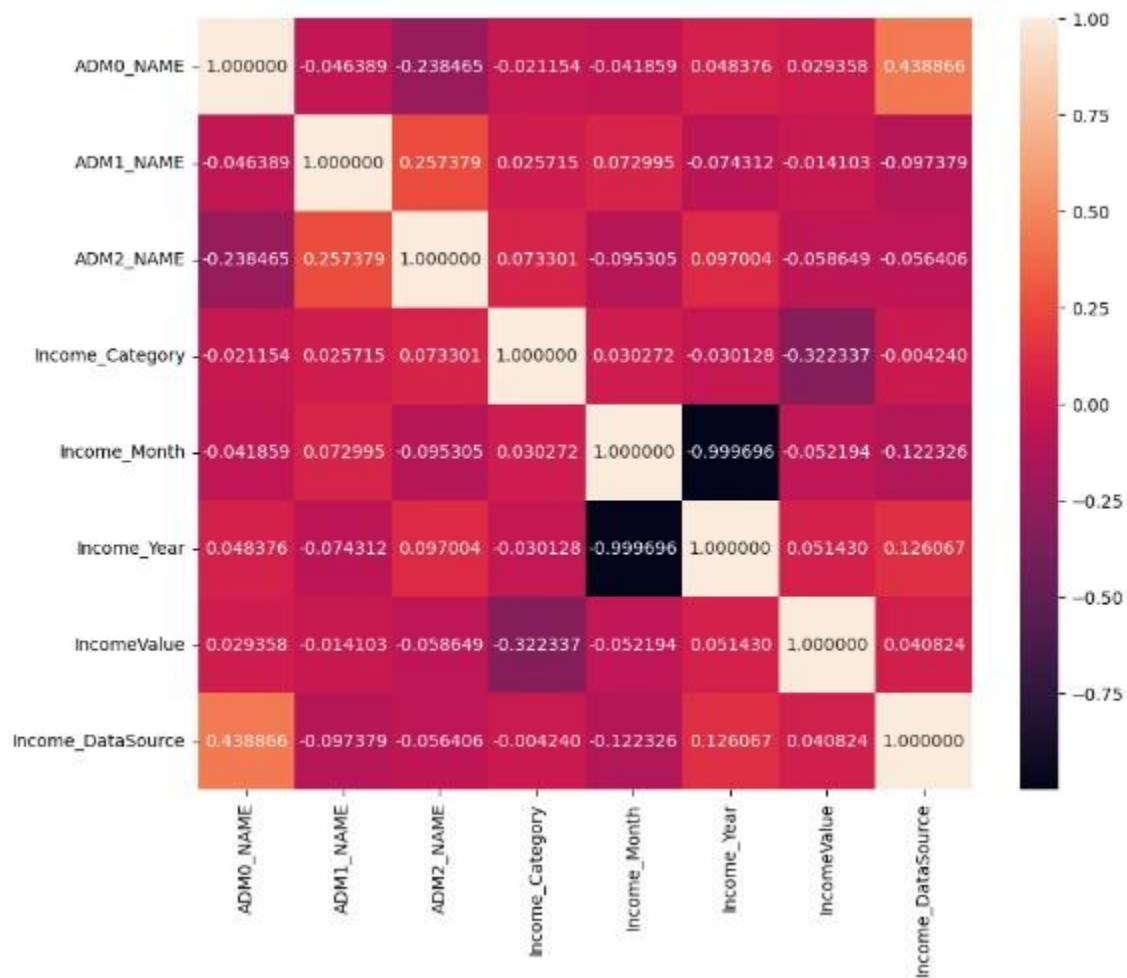
```
df.groupby('Income_Month').size().plot(kind='pie', autopct='%.2f')
```



```
# Convert relevant columns to numeric, handling non-numeric values
for col in df.columns:
    try:
        df[col] = pd.to_numeric(df[col], errors='coerce') # Convert to numeric, replace non-convertible with NaN
    except:
        pass # Skip columns that cannot be converted

plt.figure(figsize = (25, 18))
sns.heatmap(df.corr(), annot = True, cmap = 'viridis', linewidth = 0.5, fmt=".1f")
```

```
[ ] df.describe()
```



	ADM0_NAME	ADM1_NAME	ADM2_NAME	Income_Category	Income_SubCategory	Income_Month	Income_Year	IncomeValue	Indicator_Type	Income_DataSource
count	344.000000	344.000000	344.000000	344.000000	344.0	344.000000	344.000000	344.000000	0.0	344.000000
mean	0.656977	8.340116	25.715116	7.366279	0.0	5.250000	2011.040698	10.839412	NaN	1.215116
std	0.899503	3.715544	14.961182	4.176796	0.0	1.718108	1.462286	11.663107	NaN	0.566513
min	0.000000	0.000000	0.000000	0.000000	0.0	4.000000	2007.000000	0.116654	NaN	0.000000
25%	0.000000	6.000000	13.000000	4.000000	0.0	4.000000	2010.000000	1.966588	NaN	1.000000
50%	0.000000	8.000000	26.000000	7.000000	0.0	4.000000	2012.000000	6.086913	NaN	1.000000
75%	2.000000	11.000000	38.250000	10.000000	0.0	7.000000	2012.000000	15.692458	NaN	2.000000
max	2.000000	15.000000	50.000000	14.000000	0.0	9.000000	2012.000000	44.900651	NaN	2.000000

Double-click (or enter) to edit

```
[ ] y=df['Income_Month']
x=df.drop(columns = ['Income_Month','Income_SubCategory','Income_Month','Indicator_Type','ADM2_NAME'])

[ ] from sklearn.preprocessing import StandardScaler
sc = StandardScaler()

[ ] names = x.columns
```

```
[ ] x=sc.fit_transform(x)

[ ] x

array([[ -0.73144138, -2.24792527, -1.52642189,  0.65698479, -0.63682067,
         -0.38027304],
       [ -0.73144138, -2.24792527, -0.80712171,  0.65698479,  2.86834896,
         -0.38027304],
       [ -0.73144138, -2.24792527, -0.32758826,  0.65698479, -0.61929487,
         -0.38027304],
       ...,
       [ 1.49524743,  1.79504943, -0.08782153, -0.71272895,  1.33549659,
         1.38748271],
       [ 1.49524743,  1.79504943,  0.87124537, -0.71272895,  0.93268929,
         1.38748271],
       [ 1.49524743,  1.79504943,  1.59054555, -0.71272895, -0.53392931,
         1.38748271]])

[ ] from sklearn.model_selection import train_test_split
    x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2)
```

```
print(x_train.shape)
print(x_test.shape)
print(y_train.shape)
print(y_test.shape)

(275, 6)
(69, 6)
(275,)
(69,)
```

```
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error, r2_score
import numpy as np
import pandas as pd

# Convert numerical data back to a NumPy array for the model

# Split the numerical data into training and testing sets

rf_model = RandomForestRegressor(random_state=42) # You can adjust parameters here
rf_model.fit(x_train, y_train)

# Generate predictions on the training set
train_pred_rf = rf_model.predict(x_train)

# Calculate RMSE on the training set
train_rmse_rf = np.sqrt(mean_squared_error(y_train, train_pred_rf))
print("Train RMSE for Random Forest Regressor:", train_rmse_rf)

# Calculate R^2 score on the training set
r2_train_rf = r2_score(y_train, train_pred_rf)
print("Train R^2 score for Random Forest Regressor:", r2_train_rf)

# Optionally, you can also calculate RMSE and R^2 on the test set
test_pred_rf = rf_model.predict(x_test)
test_rmse_rf = np.sqrt(mean_squared_error(y_test, test_pred_rf))
print("Test RMSE for Random Forest Regressor:", test_rmse_rf)

r2_test_rf = r2_score(y_test, test_pred_rf)
print("Test R^2 score for Random Forest Regressor:", r2_test_rf)
```

```
[ ] r2_test_rf = r2_score(y_test, test_pred_rf)
print("Test R^2 score for Random Forest Regressor:", r2_test_rf)
```

```
Train RMSE for Random Forest Regressor: 0.0
Train R^2 score for Random Forest Regressor: 1.0
Test RMSE for Random Forest Regressor: 0.0
Test R^2 score for Random Forest Regressor: 1.0
```

```
[ ] rf_model.predict([[0,213,18,2009.0,20.118277,27]])
```

```
array([5.8])
```

```
[ ] rf_model.predict([[31,54,18,2011.0,7.022699,31]])
```

```
array([5.44])
```

```
[ ] Start coding or generate with AI.
```

```
[ ] import pickle
pickle.dump(rf_model,open('model.pkl','wb'))
```