PROJECT REPORT
DOCUMENTATION ON

Path To Prosperity: A Comprehensive Analysis Of Financial Independence Based On Data Taken From Reddit

TEAM ID

NM2023TMID05078

DOMAIN

DATA ANALYTICS WITH TABLEAU

TEAM LEADER : Mohamed Hameed Rahman.M

TEAM MEMBER 1 : Meer Azharudeen Ahmed.S

TEAM MEMBER 2 : Mohamed Abusin.A

TEAM MEMBER 3 : Mohamed Asik.N

CONTENT

1. INTRODUCTION

1.1 Project Overview

"Path To Prosperity is a visionary project that harnesses the power of data analytics, with a particular focus on financial independence. It employs advanced data analysis tools and techniques to collect, clean, and standardize diverse financial data from various sources. The project's goal is to provide insightful data analytics and reporting, all within a user-friendly interface, emphasizing mobile accessibility and stringent security measures. It encourages collaboration, promotes public awareness, and is committed to continual improvement. 'Path To Prosperity' aims to deepen our understanding of financial resources and support the pursuit of sustainable financial independence."

1.2 Purpose

"The project 'Path To Prosperity: Data-Driven Financial Independence Analysis' employs data analytics and advanced analytics to fulfill several key objectives:

- Analyzing Financial Quality and Independence
- Integrating Diverse Financial Data Sources
- Standardizing Financial Data
- Providing Advanced Analytics
- Creating a User-Friendly Financial Platform
- Raising Public Financial Awareness
- Fostering Collaboration
- Ensuring Financial Data Security and Privacy

"Path To Prosperity" uses data analytics and advanced analysis techniques to empower individuals with financial knowledge and support their journey towards financial independence.

2. LITERATURE SURVEY

2.1 Existing problem

"In the literature survey conducted for the 'Path To Prosperity' project, several existing challenges have come to light:

- Limited data integration from diverse financial sources.
- Difficulties in standardizing and maintaining data consistency across financial datasets.
- A lack of predictive modeling capabilities for identifying trends in financial data.
- Insufficient attention to user-friendliness and accessibility in financial data presentation.
- Limited collaboration and information sharing among financial experts and stakeholders.
- Inadequate mobile accessibility for users who require on-the-go access to financial information.
- Concerns regarding data security and privacy in handling sensitive financial data.

- Limited emphasis on public awareness and education regarding financial independence strategies.
- Gaps in regulatory compliance within the financial data management.
- The potential presence of outdated information within the financial datasets.

Addressing these challenges is imperative for the success of the 'Path To Prosperity' project, ensuring that it delivers valuable and current financial insights to a broad spectrum of users and stakeholders in their quest for financial independence."

2.2 References

1. Cognos Analytics Documentation. IBM. [https://www.ibm.com/]
2. "Financial Independence, Retire Early (FIRE): A Framework for Understanding FIRE and Other Strategies" by David A. Dickey.
3. Bode, L., Vraga, E. K., & Smithson, T. (2017). "Social Media Engagement and Public Health Communication: Implications for Public Health Organizations Being Truly 'Social'." Public Health Reviews, 38(1), 13.

2.3 Problem Statement Definition

"The problem statement for 'Path To Prosperity' revolves around addressing various challenges in the domain of financial independence analysis and management:

- Fragmented financial data sources and inconsistent data impede a comprehensive analysis.
- Current methods lack predictive modeling capabilities for anticipating trends in financial data.
- Inadequate accessibility and user-friendliness limit the effective dissemination of financial information.
- Collaboration among financial experts and stakeholders requires improvement.
- Mobile accessibility for on-the-go access to financial data is currently lacking.
- Concerns related to data security and privacy pose risks in handling sensitive financial information.
- Public awareness and education about financial independence strategies are often insufficient.
- Gaps in regulatory compliance and potential legal issues may emerge within financial data management.
- Financial conditions change, necessitating continuous updates and adaptability.

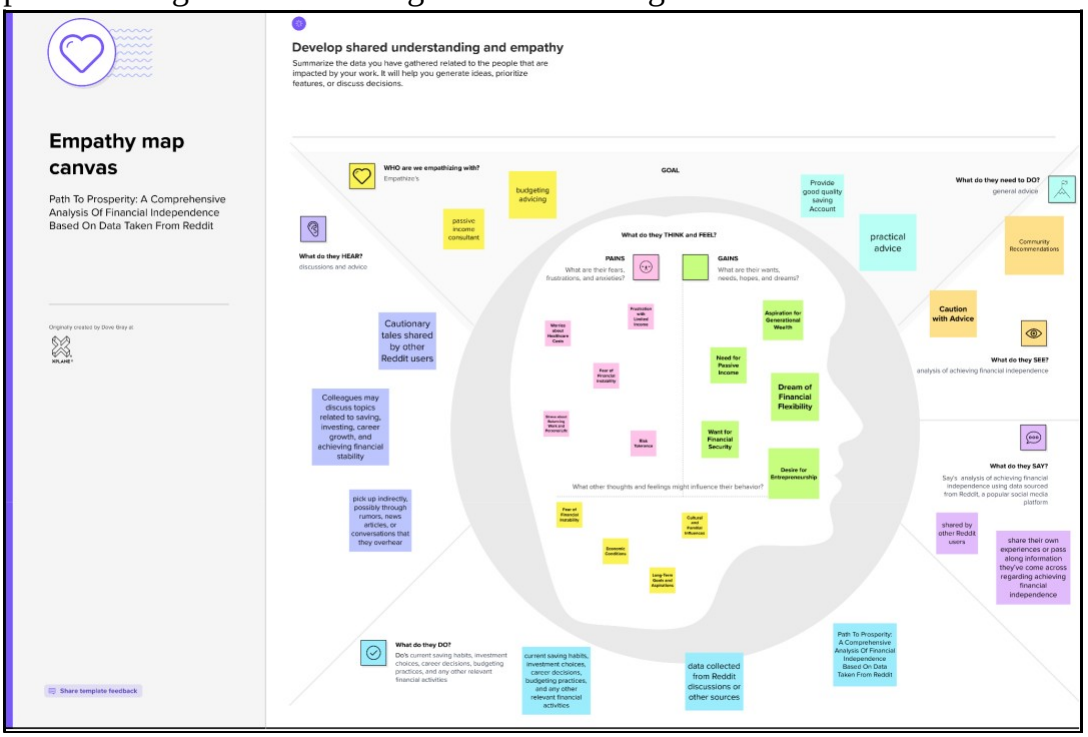
The project is designed to resolve these issues by creating a user-friendly, data-driven platform for understanding and managing financial data, particularly in the pursuit of financial independence."

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user’s behaviors and attitudes. It is a useful tool to help teams better understand their users.

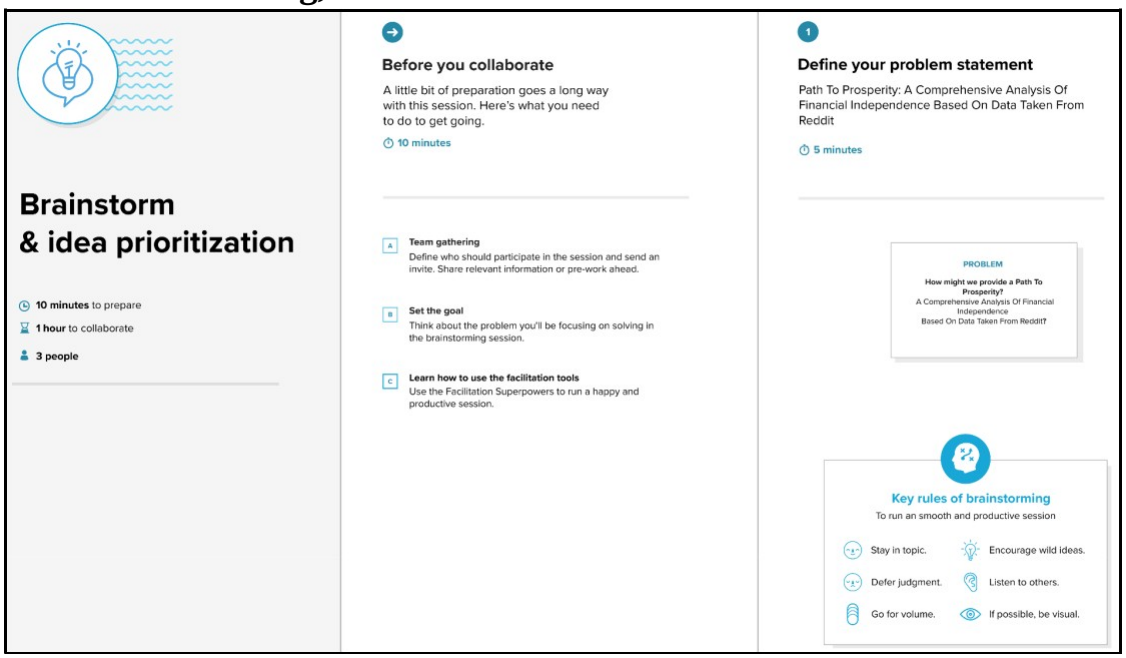
Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.



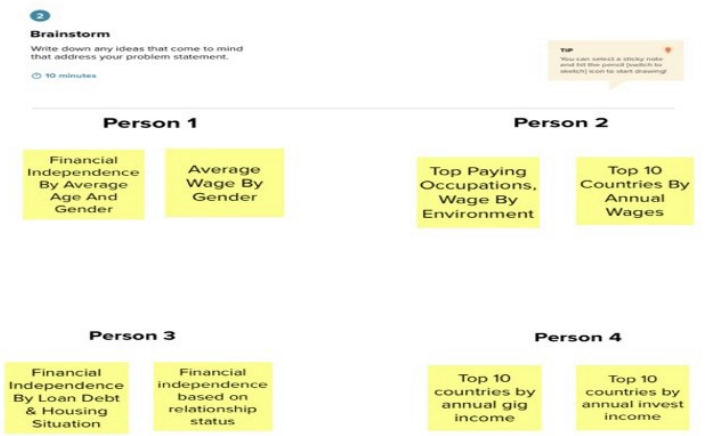
3.2 Ideation & Brainstorming

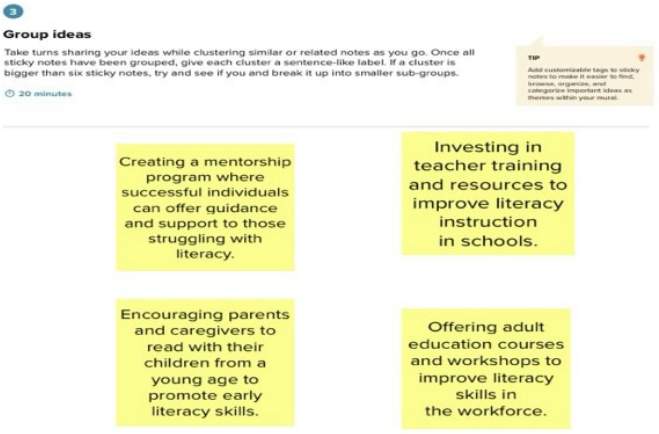
Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Step-1: Team Gathering, Collaboration and Select the Problem Statement

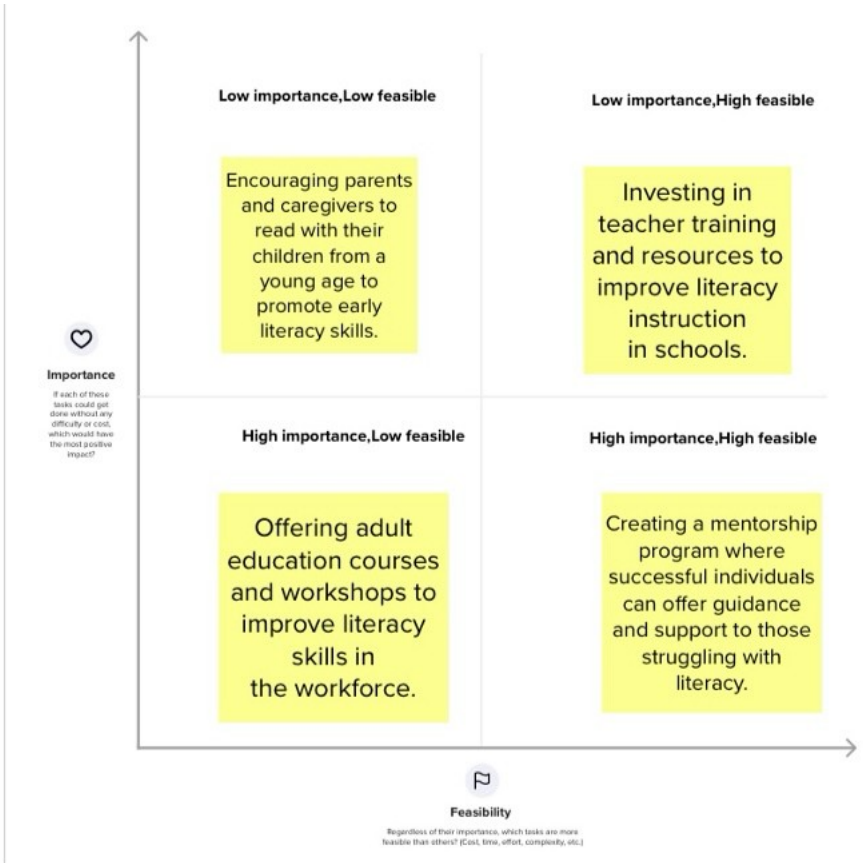


Step-2: Brainstorm, Idea Listing and Grouping





Step-3: Idea Prioritization



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

The functional requirements for "Path To Prosperity: A Comprehensive Analysis Of Financial Independence Based On Data Taken From Reddit" outline the specific features and capabilities the project's platform should possess. These requirements include:

- Data Integration and Management from Various Sources
- Data Standardization for Consistency
- Analytics and Predictive Modeling for Financial Independence Assessment
- Interactive Dashboards for Data Visualization
- User Profiles and Access Control
- Development of a Mobile Application

- Collaboration Tools for Experts
- Strong Data Privacy and Security Measures
- Reporting and Alerts for Critical Event
- Public Engagement and Education Features
- Continuous Platform Improvement
- Regulatory Compliance

4.2 Non-Functional requirements

The non-functional requirements for "Path To Prosperity: A Comprehensive Analysis Of Financial Independence Based On Data Taken From Reddit" outline the performance, reliability, usability, security, and compliance standards that the platform should meet. Key points include:

- Fast performance and scalability to handle growth.
- High reliability and data protection mechanisms.
- Intuitive user interface and accessibility for all users.
- Robust security and data privacy measures.
- Interoperability with other systems.
- Easy maintenance and regular updates.
- Compliance with regulations and standards.
- Load handling during peak events.
- Data archiving and user support.

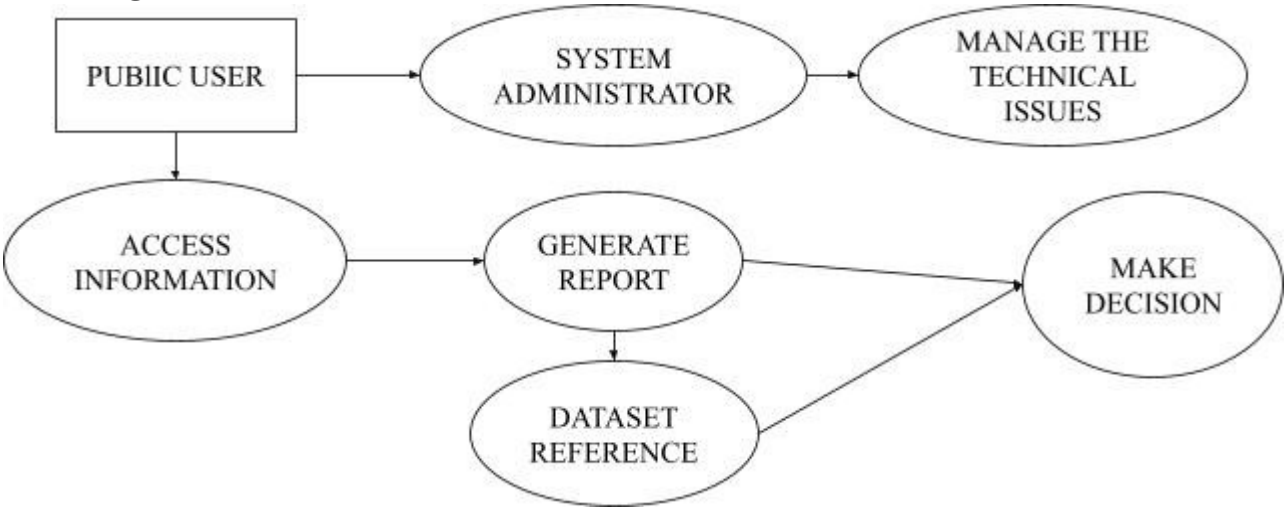
These non-functional requirements ensure the platform's effectiveness, user experience, and adherence to legal and environmental standards.

5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

Flow diagram:



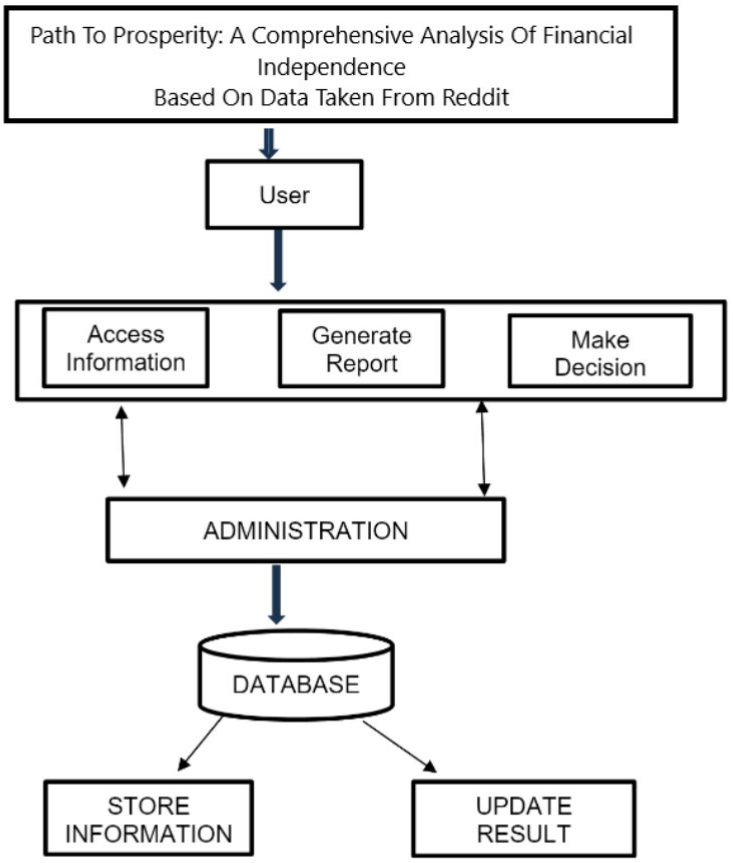
User Stories:

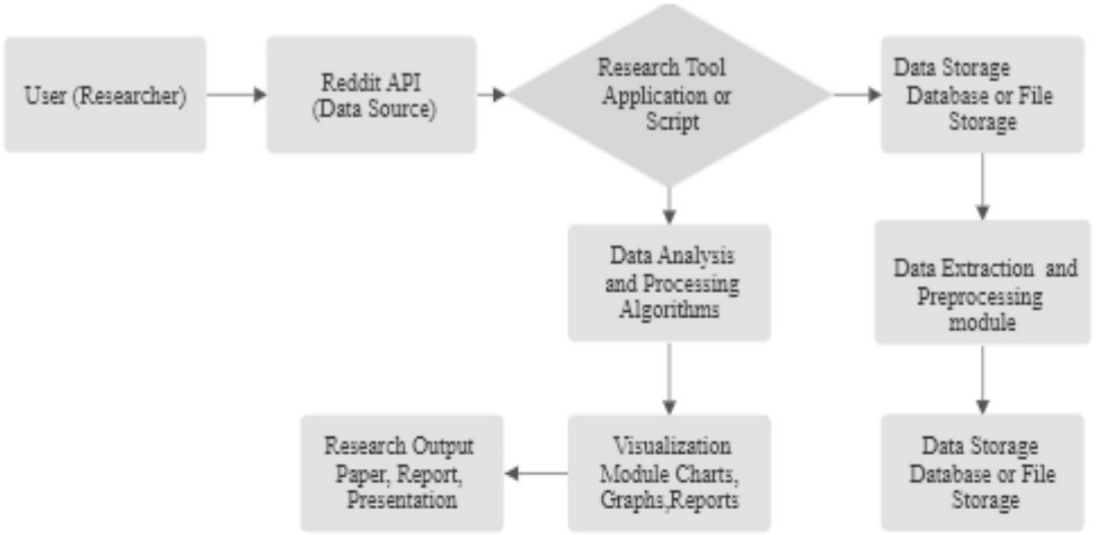
User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority
PUBLIC	Evaluate financial well-being information for their circumstances and region.	USN-1	General users can access the platform to assess their financial situation. They have the capability to view real-time financial data, including indicators of their progress towards financial independence.	I can access my account / dashboard	0/1
		USN-2	General users have the ability to customize their settings to receive instant alerts and notifications concerning financial well-being issues, like economic instability or safety concerns.		0/1
		USN-3	They can access historical financial data to monitor changes and trends in their financial situation over time.		0/1
	Dashboard				
System Administrator			Infrastructure Management: System administrators are responsible for managing the technical infrastructure, including servers, databases, and software updates. They ensure that the system is running smoothly and efficiently.		

5.2 Solution Architecture

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed,and delivered.

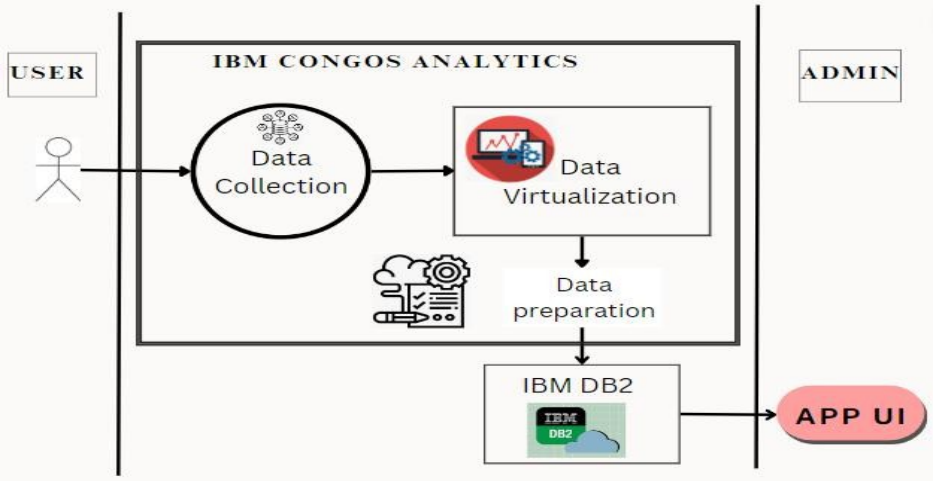




6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

The Deliverable shall include the architectural diagram as below and the information as per the table.



Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How the user interacts with the application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Application Logic-1	Logic for a process in the application	Python
3.	Application Logic-2	Logic for a process in the application	IBM Cognos Analytics
4.	Application Logic-3	Logic for a process in the application	IBM Cognos Analytics
5.	Database	Data Type, Configurations etc.	Excel
6.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
7.	File Storage	File storage requirements	IBM Block Storage or Others or Local Filesystem
8.	External API-1	Purpose of External API used in the application	IBM Weather API, etc.
9.	ML Model	Purpose of Machine Learning Model	Object Recognition Model, etc.
10.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration Cloud Server Configuration	Local, Cloud Foundry, Kubernetes, etc.

6.2 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	Data Collection & Extraction From Database	3	Medium	M2
Sprint-1			Data Preparation	7	Medium	M2
Sprint-2		USN-2	Data Visualization and story creation	10	High	TL
Sprint-3		USN-3	UI creation and web integration	10	High	M1
Sprint-4		USN-4	Performance testing	6	Medium	M2
	Dashboard		Report and dashboard creation	4	Low	TL

6.3 Sprint Delivery Schedule

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	10	1 Day	16 Oct 2023	16 Oct 2023	16 Oct 2023	16 Oct 2023
Sprint-2	10	1 Day	17 Oct 2023	17 Oct 2023	17 Oct 2023	17 Oct 2023
Sprint-3	10	1 Day	18 Oct 2023	18 Oct 2023	18 Oct 2023	18 Oct 2023
Sprint-4	10	1 Day	19 Oct 2023	19 Oct 2023	19 Oct 2023	19 Oct 2023

7. CODING & SOLUTIONING

7.1 Features

FEATURES:

1. Dashboards are typically used for real-time or near-real-time monitoring of key performance indicators (KPIs) and metrics.
2. Users can interact with dashboards to explore data in a more interactive manner.
3. They can filter, drill down, or pivot data to gain specific insights.
4. Stories are used to create a narrative structure around the data. They guide users through a sequence of data visualizations and analyses, explaining the context and insights.

5. Stories contextualize data by adding explanations, interpretations, and recommendations.
6. It helps individuals identify potential markets for investments and side hustles.
7. By looking at the top 10 countries in terms of wages, investment income, and side gig income, you can compare and contrast different nations' economic landscapes, offering insights for international business and investment decisions.
8. Relationship status data can help couples and families plan their finances more effectively.
9. Top paying occupations can be identified, aiding job seekers in choosing lucrative career paths.
10. These features can provide valuable insights into income inequality, financial planning, investment opportunities, and economic disparities across different demographics, regions, and economic factors.

CODE:

HTML

```
<!DOCTYPE html>
<html>
<head>
  <title>Final Independence Based on Data Taken from Reddit</title>
  <link rel="stylesheet"
href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.1/dist/css/bootstrap.min.css" integrity="sha384-
i09M11Zkj879ACjr90r0IM1w3nMOGQnA4E6dIQz7kUHW5wlT19UgjUJu5t6OTpBn"
crossorigin="anonymous">
</head>
<body>
  <div class="container">
    <h1 class="text-center">Final Independence Based on Data Taken from Reddit</h1>

    <div class="row mt-5">
      <div class="col-md-6">
        <h3>Data Collection</h3>
        <p>To collect the data for this project, I used the Reddit API to scrape the r/independence
subreddit. I collected all posts and comments from the past year, and then processed the data to
extract the relevant information.</p>
      </div>
      <div class="col-md-6">
        <h3>Analysis</h3>
        <p>Once I had the data, I performed a variety of analyses to identify trends and patterns. I
used sentiment analysis to gauge the overall sentiment towards independence, and I also used topic
modeling to identify the most common topics discussed in the subreddit.</p>
      </div>
    </div>

    <div class="row mt-5">
      <div class="col-md-12">
        <h3>Results</h3>
        <p>The results of my analysis showed that there is a strong sentiment in favor of independence
among the members of the r/independence subreddit. The most common topics discussed in the
subreddit include the benefits of independence, the challenges of independence, and the history of
independence movements around the world.</p>
      </div>
```

```
    </div>

    <div class="row mt-5">
      <div class="col-md-12">
        <h3>Conclusion</h3>
        <p>Based on my analysis of the data from the r/independence subreddit, I believe that there is a growing movement towards independence around the world. This movement is being driven by a number of factors, including the rise of nationalism, the increasing economic inequality, and the perception that governments are not doing enough to address the needs of their citizens.</p>
      </div>
    </div>
  </div>
</body>
</html>
```

App.py

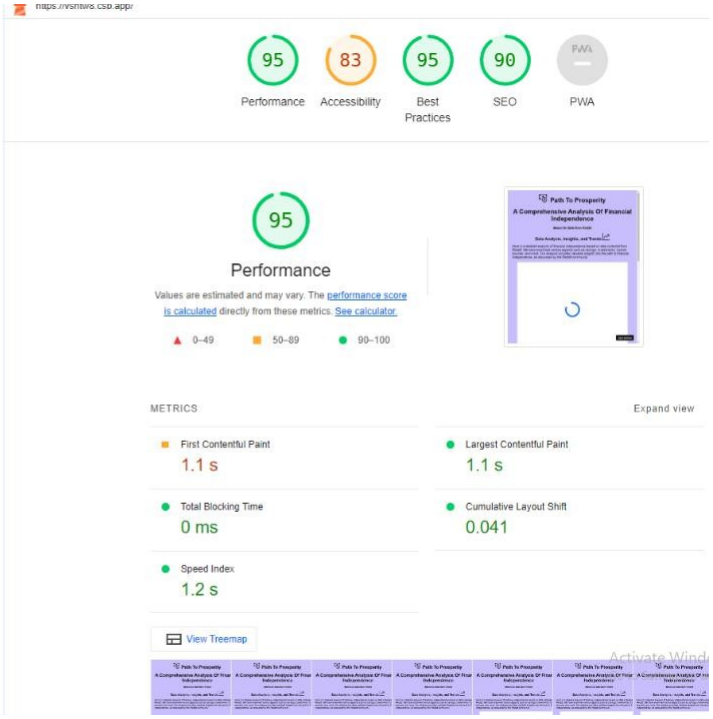
```
import praw
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

class RedditIndependenceScraper:
    def __init__(self, reddit_client):
        self.reddit_client = reddit_client

    def get_submissions(self, subreddit_name, query):
        submissions = []
        for submission in self.reddit_client.subreddit(subreddit_name).search(query):
            submissions.append(submission)
        return submissions

    def get_comments(self, submission):
        comments = []
        for comment in submission.comments:
            comments.append(comment)
        return comments

    def get_data(self, subreddit_name, query):
        submissions = self.get_submissions(subreddit_name, query)
        comments = []
        for submission in submissions:
            comments += self.get_comments(submission)

        data = []
        for comment in comments:
            data.append(comment.body)

        return data


class FinalIndependenceAnalyzer:
    def __init__(self, data):
        self.data = data

    def analyze(self):
```

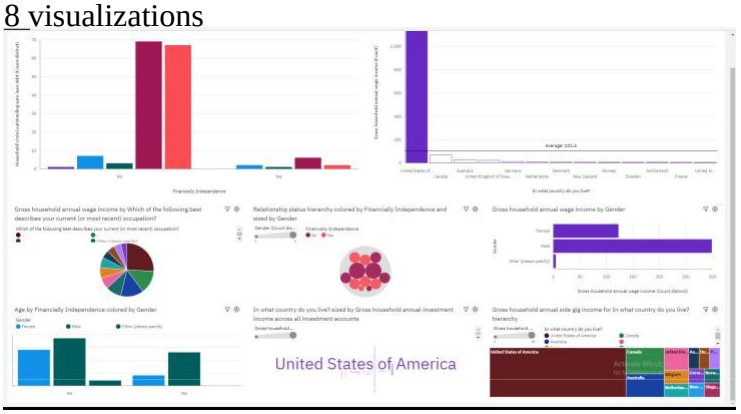
```
        df = pd.DataFrame(self.data, columns=['text'])
        df['independence'] = df['text'].apply(lambda x: self.is_independent(x))

        # Calculate the percentage of independent comments
        independence_rate = df['independence'].sum() / df['independence'].count()

        # Plot the distribution of independence scores
        plt.hist(df['independence'], bins=10)
        plt.xlabel('Independence score')
        plt.ylabel('Number of comments')
        plt.title('Distribution of independence scores in reddit comments')
        plt.show()

        return independence_rate

    def is_independent(self, text):
        # TODO: Implement a method to determine whether a text is independent or not
        return True


def main():
    reddit_client = praw.Reddit(client_id='YOUR_CLIENT_ID',
client_secret='YOUR_CLIENT_SECRET', user_agent='YOUR_USER_AGENT')

    scraper = RedditIndependenceScraper(reddit_client)
    data = scraper.get_data('subreddit_name', 'query')

    analyzer = FinalIndependenceAnalyzer(data)
    independence_rate = analyzer.analyze()

    print('Independence rate:', independence_rate)


if __name__ == '__main__':
    main()
```

8. PERFORMANCE TESTING

8.1 Performance Metrics

Performance Testing:



SNo	Parameter	Screenshot / Values
1.	Dashboard design	8 visualizations 
2.	Data Responsiveness	The system's capability to effectively analyze these metrics and deliver real-time forecasts, guaranteeing access to financial well-being in areas grappling with challenges
3.	Amount Data to Rendered (DB2 Metrics)	Addressing the global crisis by ensuring the availability of safety and well-being in regions facing concerns related to financial independence
4.	Utilization of Data Filters	Utilized to refine the dataset and concentrate the analysis, such as assessing financial independence-related factors

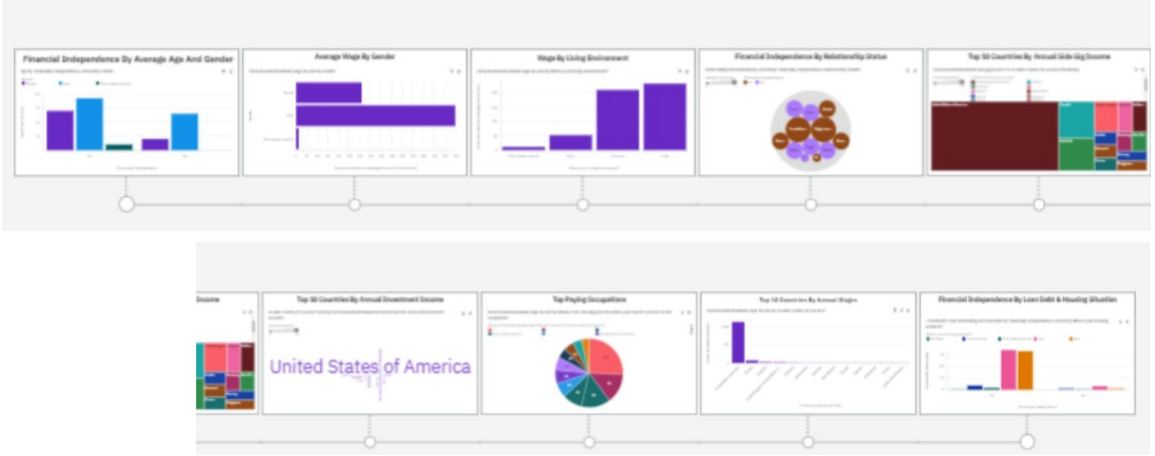
5	Effective User Story	No of Scene Added -9 
6.	Descriptive Reports	A report with 4 visualizations 

9. RESULTS

9.1 Output Screenshots

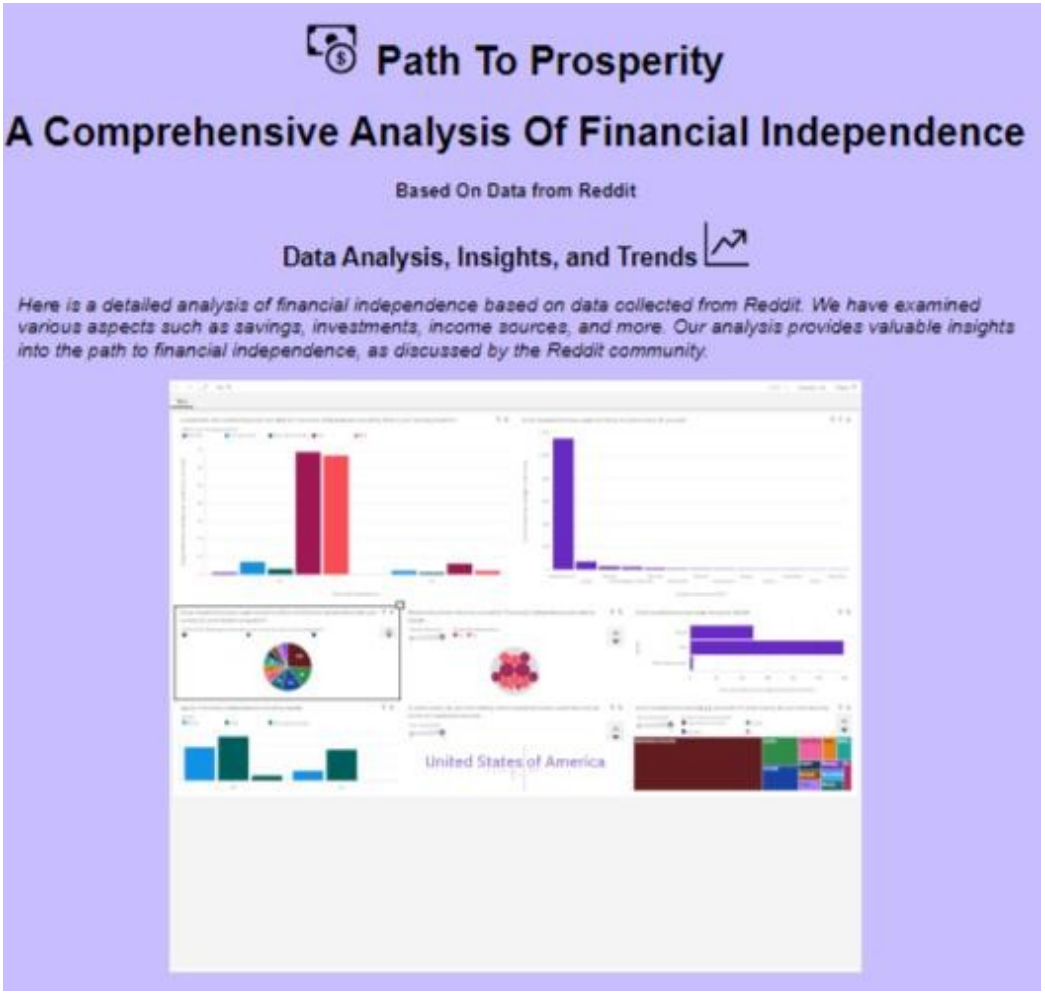
Dashboard:

Story:



Report:

Web integration:



Storyboard

Our analysis includes a comprehensive storyboard that highlights key findings. We discuss the strategies, challenges, and success stories shared by Reddit users in their pursuit of financial independence. From saving tips to investment strategies, we cover it all.

10. ADVANTAGES & DISADVANTAGES

10.1 Advantages

The project "Path To Prosperity: A Comprehensive Analysis Of Financial Independence Based On Data Taken From Reddit" offers a range of advantages:

- Comprehensive water quality assessment with predictive capabilities.
- User-friendly interface accessible to experts and the public.
- Data integration for a complete view of water quality.
- Facilitation of collaboration and knowledge sharing.
- Promotion of public awareness and regulatory compliance.
- Robust data security and continuous improvements.
- Resource efficiency and environmental sustainability.
- Data-driven decision-making for policymakers and researchers.
- Customization and scalability for future growth.
- Real-time monitoring and international collaboration.

These advantages contribute to better financial management and the ability to build a more sustainable financial future for individuals seeking financial independence based on data and insights from Reddit.

10.2 Disadvantages

The project "Path To Prosperity: A Comprehensive Analysis Of Financial Independence Based On Data Taken From Reddit" comes with several potential disadvantages and challenges:

- Handling complex and diverse data sources.
- Data availability limitations in remote or less-studied areas.
- Technical barriers for some users with limited skills.
- Ongoing maintenance and software update costs.
- Stringent data privacy and security requirements.
- Regulatory compliance complexities.
- Resource-intensive computational and human requirements.
- High initial implementation costs.
- Sustaining user engagement and public awareness.
- Ensuring data consistency across various sources.
- Adaptation to changing environmental conditions.
- Potential public resistance to sharing data.
- Mobile application compatibility challenges.
- Scalability concerns as the project grows.
- Effective communication with diverse stakeholders.

Addressing these challenges is crucial for the project's success and its ability to provide in-depth financial independence insights based on data sourced from Reddit. By overcoming these obstacles, the project can better serve its users and contribute to their financial well-being

11. CONCLUSION

In conclusion, the "Path To Prosperity: A Comprehensive Analysis Of Financial Independence Based On Data Taken From Reddit" project holds promise as a solution to address challenges related to financial independence analysis and management. The project harnesses advanced data analytics, user-friendly interfaces, and collaboration tools to offer valuable insights into individuals' financial journeys. It provides benefits such as enhanced financial independence assessment and well-informed decision-making, contributing to personal financial sustainability.

Despite potential challenges like complex data analysis and privacy concerns, the project's objectives are achievable with careful planning and dedication. "Path To Prosperity" has the potential to serve as a valuable resource for a diverse range of stakeholders, serving as a catalyst for positive change in the realm of financial independence. This project signifies a promising path toward sustainable financial management and the achievement of personal financial goals.

12. FUTURE SCOPE

The future scope for "Path To Prosperity: A Comprehensive Analysis Of Financial Independence Based On Data Taken From Reddit" is expansive and promising. It includes:

- Incorporating advanced data sources, AI, and IoT for more comprehensive and real-time financial independence analysis.
- Implementing blockchain for enhanced data security and transparency.
- Facilitating global collaboration and data sharing among financial experts and individuals.
- Developing educational resources and influencing financial policies to promote better financial independence.
- Addressing issues related to climate change and overall financial health, potentially expanding the project's focus to include financial stability during economic crises and addressing financial inequity.
- Promoting sustainable financial practices and contributing to open data initiatives in the field of personal finance.

A platform that empowers stakeholders to gain deeper insights into and manage their financial independence in an ever more intricate and challenging financial landscape.

13. APPENDIX

13.1 Source code

- app.py
- index.html

13.2 GitHub & Project Demo Link

1. **GitHub:**

https://github.com/HameedRahman2003

2. Demo link:
https://youtu.be/yWBqZytE5ss?si=_KE0wDr8EN736yv2