

Data Science in the Wild – Group Project



Project Title: Gambling Harm Prevention Strategy

Client: Newcastle City Council

Company: Gambit Wise

Project Purpose: Gambit Wise is a data-driven company supporting Newcastle City Council in detecting and preventing problem gambling behaviours early. We help implement responsible gambling practices through real-time data analysis while ensuring user privacy.

Team Members of Group 1:

Monisha Dabbara

Andrieta Irene Georghiadou

Pranav Sunil Raja

Priydarshani Sanjay Patil

ZiQing Wang

Kartik Goswami

Date: 20th March 2025

Word Count: 2098 (excluding References and Appendix)

Table of Contents

Executive Summary.....	2
Introduction (Context)	2
Problem Definition	3
Organisational Goals	3
Key Stakeholders Involved.....	3
Previous Attempts & Their Challenges	3
Problem Analysis (Need).....	4
Customer Journey & Key Pain Points	4
Challenges in Data Collection	4
Impact on Newcastle City Council & Stakeholders	5
Key Findings from Our Research.....	5
Proposed Solution (Vision): A Data-Driven Gambling Harm Prevention Strategy.....	5
A United Approach: Centralised Gambling Harm Database	5
How the Database Works.....	6
How Our Solution Solves Challenges	7
Critical Assumptions	7
Outcomes (Expected results).....	8
Data Analysis on Synthesised Data	8
Minimal Viable Product Proposal.....	9
Limitations of Proposed Ideas	9
Conclusions and Future Scope	10
References	10
Appendix	11
Appendix A: Customer Journey Map (CJM).....	11
Appendix B: Value Proposition Canvas	12
Appendix C: Survey Questions (Example)	12
Appendix D: Business Model Canvas	16
Appendix E: Analysis of Synthesised Data	17
1. Dashboard.....	17
2. Code for the analysis of Synthesised Data	17
Appendix F: Minimum Viable Product (MVP) Implementation Plan.....	21

Executive Report on Enhancing Online Gambling Data Capture for Newcastle City Council

Executive Summary

Newcastle City Council's main challenge is improving the capture of online gambling data. This issue is difficult to address if real-time data is not available.

Gambling addiction is a growing public health concern, leading to both financial and mental distress (NHS, 2024). Addressing these issues is important to protect vulnerable users and overall reduce the burden on local help services (Newcastle City Council, 2025). We propose a data-driven gambling harm prevention strategy that includes integrating data from key stakeholders such as banks, healthcare systems, the gambling commission, and the people of the UK, more specifically Newcastle residents. Along with these should include campaigns to spread awareness of harms related to gambling. This centralised database will include anonymised gambling-related data, enabling targeted interventions and policy betterment, which will take one year to implement and requires huge funding.

This results in better gambling regulation, earlier identification of problem gamblers, and stronger cross-sector collaboration. The next steps focus on securing partnerships with GamCare and GambleAware to improve data accuracy and expand prevention efforts.

Introduction (Context)

The rise of online gambling has increased recently, making it difficult to track and regulate. Adding to this, the data collected presently is from self-reporting and operator disclosure, which is unreliable (Newcastle City Council, 2025; Gambling Commission, 2022; InNewcastle, 2024).



Figure 1: Gambling participation rates from 2018-2022, by type of gambling. Source: Gambling Commission (2022).

The lack of tracking systems for gambling activities makes it complicated for early detection for the council (Gambling Commission, 2024). Most of the local councils do not have direct access to detailed information on gambling activities like the national regulators, making it difficult for the Newcastle City Council to identify at-risk users. Further challenges include fragmented data systems, data privacy issues and lack of public awareness

of gambling harms. To address these challenges, the council must adopt a data-driven approach. (Newcastle City Council, 2024).

Problem Definition

Most available data is self-reported, fragmented, or restricted by privacy laws, obstructing interventions (GamCare, 2024). The distribution of licensed gambling depicts the players of different online gambling leads to financial hardship, but the council lacks evidence to confirm these developments (Information Commissioner's Office, 2024; Gambling Commission, 2024). Hence, to regulate gambling efficiently, the Council needs insights on:

- Who is gambling? (age, gender, income)
- Where are they located? (socioeconomic indicators, geographic areas)
- Risk Levels? (casual vs. problem gamblers)
- Behavioural Patterns? (mental health links, spending habits)

The present data sources are surveys, financial transactions, and incomplete operator reports, cause major drawbacks.

Organisational Goals

To address these issues, we seek to:

1. Enhance data collection to gain gambling behaviours, demographics, and risk levels.
2. Identify at-risk users for early interventions.
3. Fortify regulation despite jurisdictional challenges.
4. Collaborate with major stakeholders for better data-sharing.

Key Stakeholders Involved

1. Internal: Newcastle City Council
2. External:
 - Residents – Provides data of affected users
 - NHS & Healthcare Systems – Provides Gamblers with demographics
 - Gambling Commission – Holds gambling operator data.
 - Financial Institutions – provide anonymized transaction data to identify high-risk gambling behaviours

Previous Attempts & Their Challenges

Several previous attempts to address the issue include NHS and Gambling Commission guidelines (e.g., NICE), health surveys (e.g., Health Survey for England, Gambling Survey for Great Britain), expanded gambling clinics, and professional training initiatives (GamCare, GambleAware) (GambleAware, 2024). However, these attempts faced multiple challenges, such as self-reported data & biases, and the training remained insufficient for healthcare professionals (GamCare, 2024). The public awareness sector, through different campaigns, was not addressed satisfactorily due to weak enforcement and poor coordination (GambleAware, 2024), while

councils face fragmented data and limited jurisdiction, stopping effective regulation and resulting in a lack of localised data.

Problem Analysis (Need)

Significant ways to improve online gambling data collection include Newcastle residents, healthcare, finance, and policymakers, which are limited as data gaps and reactive responses persist (Newcastle City Council, 2024).



Figure 2: This flowchart illustrates the root cause of problems in obtaining online gambling data.
Source: Newcastle City Council, 2024

Customer Journey & Key Pain Points (See the customer journey map and related details in Appendix A.)

1. Exposure & Engagement: Gambling is targeted by aggressive gambling ads at young adults and financially vulnerable individuals (GambleAware, 2024).
2. Problem Development: Financial losses → debt → decline in mental health → relationship breakdowns (NHS Northern Gambling Service, 2024)
3. Crisis & Impact: With minimal NHS care, causes harm such as debt and mental health crises (NHS, 2024).
4. Barriers from Support: Support-related obstacles include underreporting, societal stigma, and slow NHS support (GamCare, 2024).

Challenges in Data Collection

Newcastle City Council faces major problems in obtaining gambling data (refer Figure 2):

- Fragmented & Incomplete Data (GamCare, 2024).

- Restricted access to data due to GDPR (Information Commissioner's Office, 2024)
- Under-representation of online gambling trends
- Lack of Public Awareness (GamCare, 2024).
- Minimal Healthcare Training (NHS, 2024)

Impact on Newcastle City Council & Stakeholders

Limited real-time data and legal restrictions impede proactive harm reduction, leading to regulatory issues (Newcastle City Council, 2024). Access to gambling transaction data is restricted under GDPR. Gambling addiction puts pressure on mental health services and leads to financial hardship, driving demand for social assistance. Financial institutions must act fast, but there are legal barriers to these actions before gambling is identified (UK Government, 2024).

Key Findings from Our Research

To address data gaps, a cross-sector strategy is necessary. This includes NHS for monitoring and early diagnosis, banks to spot dangerous gambling habits, and gambling commission that provide data from gambling websites and information from the public through surveys. (Responsible Gaming Council, 2024). To address the data gaps as depicted in Table 1, centralised database utilisation is very important. As sensitive data is involved, accessing data is difficult, but the ICO supports the use of anonymised data for public interest reasons (Information Commissioner's Office, 2024). While initiatives like GamProtect improve real-time risk data exchange (GamProtect, 2024).

Proposed Solution (Vision): A Data-Driven Gambling Harm Prevention Strategy

Centralised Gambling Harm Database — the Council must take a data-driven approach to gambling harm prevention.

A United Approach: Centralised Gambling Harm Database

What are the fragmented data issues (see Table 1)? To address this challenge, a singular, GDPR-compliant database must be in place (i.e., data must be centralised), that collates gambling-related data from relevant stakeholders - Gambling Commission, banks, healthcare services, and public surveys (Newcastle City Council, 2024). This is implemented by other countries like Australia and Sweden (Dowling, N. A., 2021).

	Public Surveys	Gambling Transaction & Spending Trends	Mental Health & Addiction Records	Self-Exclusion & Consumer Protections	Participants Info
NHS	✓		✓		
Gambling Commission	✓	✓		✓	
Banks		✓		✓	
Public	✓				✓

Table 1: The table shows existing data. Source: Gambling Commission, 2024

Provides a thorough understanding of gaming practices, financial hardship, and hazards to the public's health. The Value Proposition Canvas (see Appendix B for details related to Value Proposition Canvas) explains how this solution meets stakeholder needs and tackles important pain areas. Important sources of data are public surveys and resident reports; banking and financial data; the Gambling Commission; and NHS and healthcare data

How the Database Works

The system will act as a secure hub to detect risk patterns and trigger early interventions.

Key features include:

- Automated Risk Detection
- Real-time Data Integration (continuous updates from key sources)
- Predictive Insights for Early Intervention
- GDPR-Compliant Anonymisation

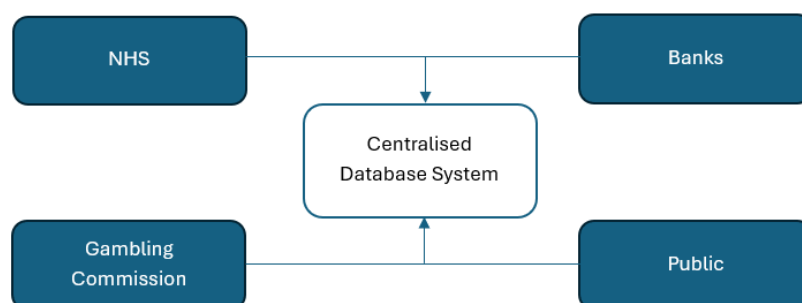


Figure 3: The figure gives a pictorial representation of the database and its contents

In addition to the Centralised Gambling Harm Database, we aim to include a multifaceted strategy incorporating regulatory reforms, public awareness, support systems, and technology.

Strengthening Regulations & Policy Advocacy

To ensure limited exposure to gambling, strong regulations are necessary, such as including self-exclusion programmes by banks. The council must majorly work with the Gambling Commission of the UK to enforce restrictions on gambling-related advertisements (Gambling Commission, 2022). Enhance user monitoring interfaces on gambling websites to check financial stability before it's too late (UK Government, 2024). Should implement a centralised self-exclusion system for better regulation, and most importantly, to council must secure funding to implement all the suggested programs (GamStop, 2024).

Public Awareness & Support Programs

To completely eradicate this problem, a nationwide approach is needed. To implement this Multiple public awareness campaigns should be conducted like those for smoking and alcohol addiction; this can be initiated by including harm warnings on ads and spending limit alerts on websites. The public must be educated on financial literacy and responsible gambling, which must be integrated into schools, universities, and workplaces (Responsible Gambling Council, 2024). Multiple support is to be provided

to individuals suffering from gambling issues by providing 24/7 helplines, online support groups, peer mentoring, Gamblers Anonymous, and financial counselling (GamCare, 2024).

Improving Surveys to Analyse Online Gambling Behaviour

The major drawback of not having enough data is due to issues in surveys, so introducing new survey systems and questions is necessary to explore more on online gambling concerns. This survey can include both direct and indirect questions (refer to Appendix C for an example of survey questionnaire) to understand the effects of gambling as well as to ensure the anonymity of the subject, as this encourages honest responses (National Council on Problem Gambling, 2024; GamCare, 2024).

Along with surveys, GP practices, and hospitals should conduct routine tracking and monitoring of gambling addictions to ensure better data collection. To inculcate this, we can employ multiple survey formats, such as online, app-based, or paper-based to reach wider audience. (NHS, 2024).

Leveraging Technology for Gambling Harm Prevention

Along with all the ideas mentioned above, it is better to always use technological aids, such as AI-driven risk detection, automated spending alerts, and mobile apps for gambling management (Newcastle City Council, 2024).

Integrating these solutions into the Centralised Gambling Harm Database will ensure a proactive approach to gambling harm prevention. A detailed breakdown of the sustainability and partnership framework supporting this is detailed in the Business Model Canvas (for details related to Business Model Canvas refer Appendix D).

How Our Solution Solves Challenges

Challenge	Solution
Fragmented data limits monitoring.	Centralized Database – Establish data-sharing platforms with the Gambling Commission, NHS, public information, and banks.
NHS lacks early gambling addiction tracking.	Healthcare Integration – Train professionals to screen for gambling harm and update NHS records.
GDPR restricts access to transaction data.	Banking Partnerships – Collaborate with banks for anonymized data.
Low public awareness of risks.	Education Campaigns – Target schools, workplaces, and online platforms to create gambling awareness.
Self-exclusion tools are ineffective.	Stronger Self-Exclusion – Advocate for a unified system across all platforms.
Gambling Commission lacks real-time insights.	AI Risk Detection – Use predictive analytics to flag high-risk gamblers early.

Table 2: This table includes major challenges and their respective solutions. Source: Newcastle City Council, 2024

Critical Assumptions

1. Regulatory bodies will enable GDPR-compliant data sharing.
2. Banks will provide anonymised transaction data.
3. AI tools will detect high-risk gamblers accurately.

- Awareness campaigns will increase self-exclusion and reduce financial & mental distress.

Outcomes (Expected results)

Data Analysis on Synthesised Data

Post implementing the solutions mentioned above, the council must investigate the data for some valuable insights. To implement this, we tried to synthesise data as to how it should be collected to provide the information needed by the council. The results obtained from the analysis must include demographic factors such as age, income, and employment status; behavioural data like betting frequency; gambling type; self-exclusion history, mental health status (NHS Digital, 2024; Aspers Newcastle, 2024); debt levels; and addiction history (Newcastle City Council, 2024).

Analysis of the synthesised data displayed in Figure 3 gives some valuable insights, such as the peak hours for gambling being in the evening, with Ouseburn being a primary hotspot. Sports betting is the most popular gambling activity. Furthermore, most gamblers reside in Sandyford, followed by Ouseburn, indicating major changes must be implemented at these places. For better visualisation, we made the dashboard as well. (refer to Appendix E: Dashboard and Code related to Analysis).

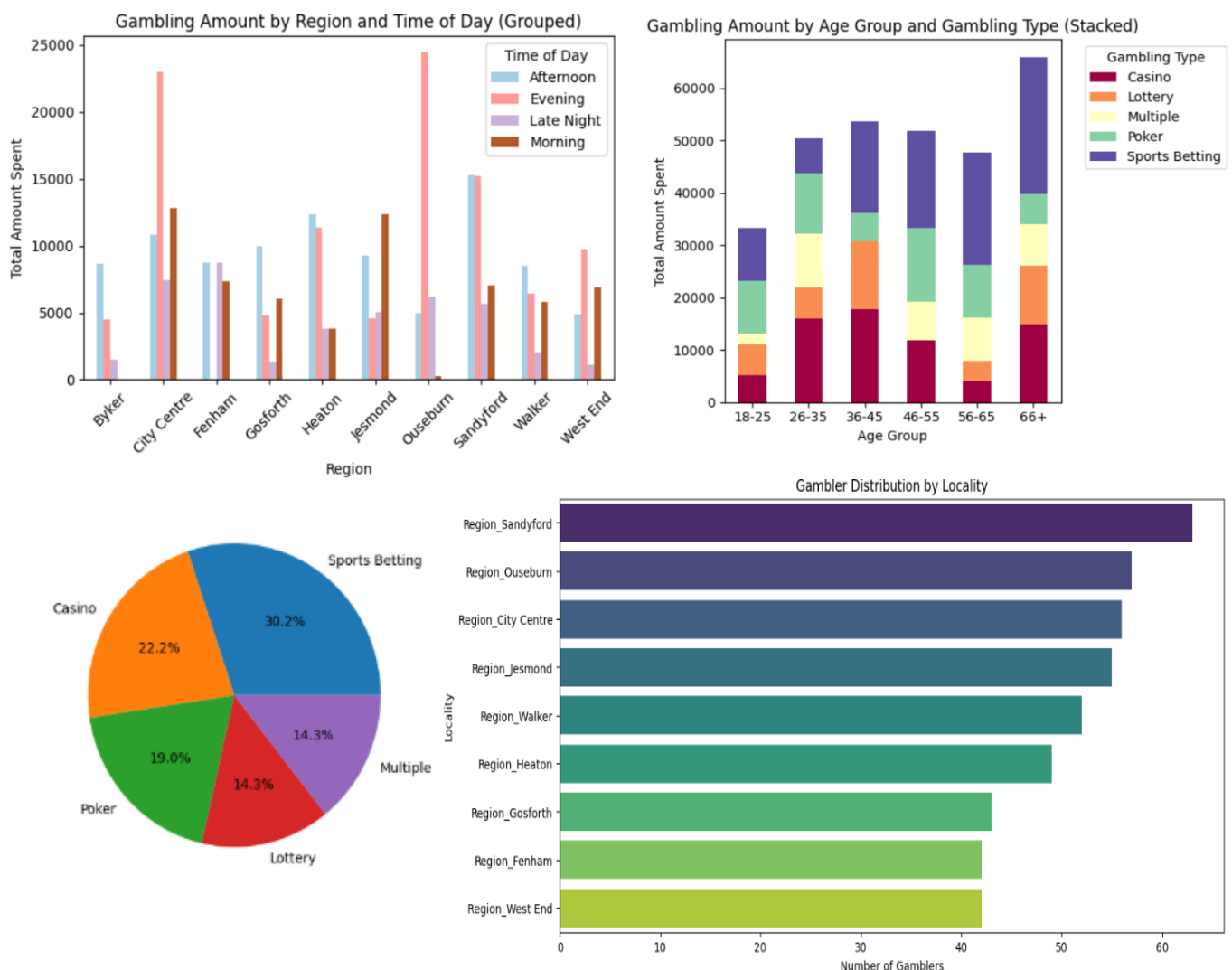


Figure 4: Outputs generated from synthesized data

Minimal Viable Product Proposal

We proposed a centralised database system to the council for leveraging data-sharing across all sectors, such as healthcare systems, financial institutions, gambling commission, and public information obtained from surveys for early intervention.

Key Outcomes of MVP

This proposed product not only facilitates the early identification of risky individuals but also helps in reducing the burden on targeted affected users (NHS, 2024; UK Parliament, 2024). As the database is anonymised, we get better insights into gambling trends, which enables data-driven policymaking. Along with this, organising public awareness campaigns improves peer engagement, eventually increasing participation in self-exclusion programmes and financial monitoring tools (GamCare, 2024). But this is not complete without strengthening regulatory enforcement which mandates affordability checks and stricter self-exclusion mechanisms across gambling platforms. (For details on implementation and policy initiatives, see Appendix F: MVP Implementation Plan).

Beneficiaries

This MVP benefits both the residents from better mental health and financial support and public health teams for better-regulating pupils through early intervention. Along with this, health professionals are now trained and equipped to screen for gambling-related harm, which improves data collection and improves patient care (NHS, 2024; Responsible Gambling Council, 2024). Banks will be using risk analytics to flag harmful gambling behaviour. (More on stakeholder contributions can be found in Appendix F: Ownership & Maintenance).

Measuring Success

On implementing all the suggested solutions, there will be a reduction in reported gambling harm cases. Self-exclusion tool adoption would increase by 30%, and public survey participation would rise by 20% (Newcastle City Council, 2025) (Detailed measurements are provided in Appendix F: Success Metrics)

Next Steps

Future steps include major partnerships with GamCare and Gam Protect to improve accuracy of the data and increase public outreach. (The full implementation plan is in Appendix F: Implementation Plan & Next Steps).

Limitations of Proposed Ideas

Though there are many benefits from the proposed solutions, the MVP still faces challenges, such as a major challenge in reaching all at-risk individuals. Data privacy regulations may limit access to critical information (ICO, 2024). A centralised database includes a lot of security risks, such as data breaches which could compromise sensitive data (Nottingham Trent University, 2013). Funding constraints hinder the implementation of all the suggested measures ((Bird & Bird LLP, 2024; NHS, 2024; UK Gambling Commission, 2024).

Conclusions and Future Scope

Key takeaways are that a data-driven strategy is necessary to address gambling-related harm, as this improves early intervention. Collaboration with NHS, the UK Gambling Commission, financial institutions, and behavioural experts enhances and protects at-risk individuals (UK Gambling Commission, 2024; UK Parliament, 2024).

For long-term impact, perpetual policy advocacy, public education, and responsible gambling initiatives are necessary (NHS, 2024). Financial monitoring and predictive analytics can help early harm detection. A centralised self-exclusion system and stronger legislation, including affordable checks, are important. For stronger impact schools should include gambling awareness, and targeted social media ads must be restricted. (National College, 2024; The Guardian, 2017).

References

- Gambling Commission (2022) *New rules boosting safety and consumer choice*. Available at: <https://www.gamblingcommission.gov.uk/news/article/new-rules-boosting-safety-and-consumer-choice> (Accessed: 02 March 2025).
- Nottingham Trent University (2013) *Gambling addiction research*. Available at: <https://irep.ntu.ac.uk/id/eprint/22666> (Accessed: 05 March 2025).
- Dowling, N. A. (2021) 'The impact of gambling addiction on mental health', *Journal of Behavioural Addictions*, 10(2), pp. 234-250. Available at: <https://doi.org/10.1007/s11469-021-00686-w> (Accessed: 06 March 2025).
- Responsible Gambling Council (2024) *Understanding responsible gambling*. Available at: <https://www.responsiblegambling.org/> (Accessed: 08 March 2025).
- Newcastle City Council (2024) *Gambling regulations and licensing*. Available at: <https://www.newcastle.gov.uk/services/licences-and-permits/gambling-act-2005/gambling-act-2005> (Accessed: 12 March 2025).
- Newcastle City Council (2025) *Gambling statement of policy 2025-2028*. Available at: <https://moderngov.newcastle-staffs.gov.uk/documents/s43679/Gambling+Statement+of+Policy+2025-2028+-+for+Council+approval.pdf> (Accessed: 12 March 2025).
- Gambling Commission (2022) *Taking a more in-depth look at online gambling*. Available at: <https://www.gamblingcommission.gov.uk/statistics-and-research/publication/taking-a-more-in-depth-look-at-online-gambling> (Accessed: 13 March 2025).
- Gambling Commission (2022) *Guidance on using statistics from the Gambling Survey for Great Britain*. Available at: <https://www.gamblingcommission.gov.uk/about-us/page/guidance-on-using-statistics-from-the-gambling-survey-for-great-britain> (Accessed: 14 March 2025).
- National College (2024) *Online gambling and young people: Understanding the risks*. Available at: <https://nationalcollege.com/news/online-gambling-and-young-people-understanding-the-risks> (Accessed: 14 March 2025).
- GamStop (2024) *Self-exclusion from online gambling*. Available at: <https://www.gamstop.co.uk> (Accessed: 14 March 2025).
- Betting and Gaming Council (2024) *Safer gambling initiatives: Research, education, and treatment*. Available at: <https://bettingandgamingcouncil.com/safer-gambling-initiatives/research-education-treatment> (Accessed: 14 March 2025).

- Aspers Newcastle (2024) *Safer gambling help & advice*. Available at: <https://www.aspersnewcastle.co.uk/safer-gambling-help-advice.html> (Accessed: 15 March 2025).
- InNewcastle (2024) *The rise of online casino gambling in Newcastle*. Available at: <https://innnewcastle.co.uk/the-rise-of-online-casino-gambling-in-newcastle-a-locals-guide-to-the-best-platforms/> (Accessed: 15 March 2025).
- Bird & Bird LLP (2024) *Gambling reform insights*. Available at: <https://www.twobirds.com/en/trending-topics/gambling-reform> (Accessed: 15 March 2025).
- NHS (2024) *Gambling addiction support and mental health*. Available at: <https://www.nhs.uk/live-well/addiction-support/gambling-addiction/> (Accessed: 01 March 2025).
- UK Government (2019) *Local authorities bulletin on gambling regulation*. Available at: <https://www.gamblingcommission.gov.uk/authorities/guide/page/la-bulletin-december-2019-case-studies> (Accessed: 15 March 2025).
- Information Commissioner's Office (2024) *Data protection and gambling regulation*. Available at: <https://ico.org.uk/about-the-ico/media-centre/news-and-blogs/2023/07/ico-backs-new-data-sharing-schemes-to-protect-gamblers-from-harm/> (Accessed: 12 March 2025).
- The Guardian (2017) *Gambling addiction and professional training for mental health services*. Available at: <https://www.theguardian.com/healthcare-network/2017/oct/09/gambling-addiction-professional-training-mental-health> (Accessed: 06 March 2025).
- UK Parliament (2024) *Gambling regulation and policy recommendations*. Available at: <https://committees.parliament.uk/writtenevidence/2331/html/> (Accessed: 08 March 2025).
- National Council on Problem Gambling (2024) *Training and surveys on gambling addiction*. Available at: <https://www.ncpgambling.org/training/ngage-survey/ngage-2/> (Accessed: 07 March 2025).
- UK Gambling Commission (2024) *An update on the Single Customer View industry challenge*. Available at: <https://www.gamblingcommission.gov.uk/news/article/an-update-on-the-single-customer-view-industry-challenge> (Accessed: 12 March 2025).
- UK Gambling Commission (2024) *Gambling regulation and GDPR compliance*. Available at: <https://www.gamblingcommission.gov.uk/licensees-and-businesses/guide/gambling-regulation-and-the-general-data-protection-regulation-gdpr> (Accessed: 12 March 2025).

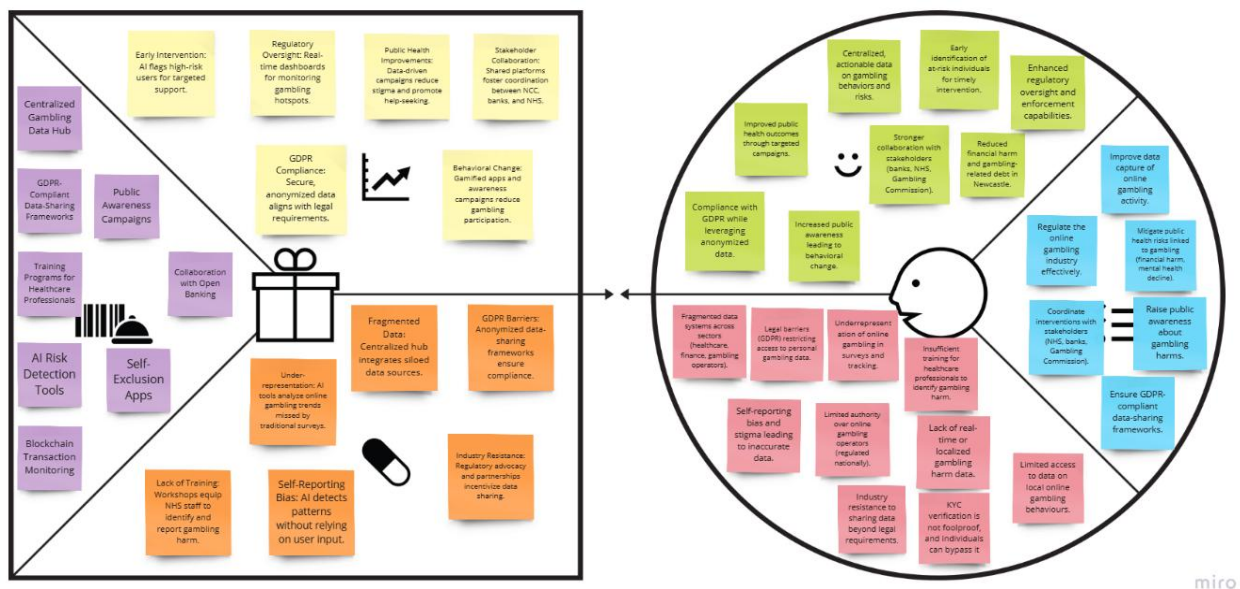
Appendix

Appendix A: Customer Journey Map (CJM)

Customer Journey Map - Newcastle City Council						
Stages	Data Integration & Sharing	Data Collection	Policy Making	Implementation	Monitoring & Evaluation	Policy Adjustment & Feedback
Job(s)	Establish data-sharing agreements, advocate for regulatory changes	Gather insights on gambling impact, improve self-reporting accuracy	Develop regulations, interventions, and data-driven policies	Enforce regulations, implement harm-reduction measures	Track effectiveness of policies using real-time data dashboards	Revise policies based on new insights, advocate for national policy changes
Pains	Data fragmentation, jurisdictional limitations, GDPR restrictions, industry reluctance to share data	Self-reporting bias, lack of reliable online gambling data, resistance from gambling operators	Resource limitations, slow policy response to emerging trends	Data accuracy challenges, lack of training for healthcare professionals	Underreporting of online gambling, limited behavioral insights beyond transactions	Delayed policy adjustments, gaps in addressing online gambling harms
Gains	Improved data accessibility, collaboration between agencies, potential for anonymized data sharing	Comprehensive data for decision-making, more accurate gambling harm metrics	Evidence-based policies, targeted interventions using predictive analytics	Reduced gambling harm, increased access to support programs	Real-time dashboards for data-driven decisions, improved public health outcomes	More adaptive and effective regulation, better national policy alignment
Feeling	Positive					
	Negative					
Effort	High					
	Low					

The image illustrates the Customer Journey Map (CJM) which shows the approach we followed to address the harms related to gambling and more specifically online gambling. Using this structured approach, we went through six stages more specifically Data Integration & Sharing, Data Collection, Policy Making, Implementation, Monitoring & Evaluation, and Policy Adjustment & Feedback. For each of the stages, we identified the specific jobs needed, the pains and gains, and the feeling and effort needed to get the required gambling data.

Appendix B: Value Proposition Canvas



We used the Miro tool to design the above plot to help us define our value proposition canvas. On the right side, the customer pains and gains are highlighted. On the left side, the solutions and the implementations are mentioned. These aim to improve data collection strategies, and regulatory compliance and to enhance public awareness.

Appendix C: Survey Questions (Example)

Gambling and You: A Research Study on Gambling Behaviour

Gambling can take many forms, from playing the lottery or scratchcards to betting on sports, using online casinos, or even engaging with gambling-like features in video games. This survey aims to understand gambling habits, motivations, and potential harms in a confidential and non-judgmental way.

All responses are anonymous and will help improve understanding and support for those who may need it. If any question does not apply to you, feel free to skip it.

General Information

What is your age?

- ☐ Under 18
- ☐ 18-24
- ☐ 25-34
- ☐ 35-44
- ☐ 45-54
- ☐ 55-64
- ☐ 65+

What is your gender?

- ☐ Male

- Female
- Non-binary
- Prefer not to say
- Other (please specify)

In which region do you live?

- England
- Scotland
- Wales
- Northern Ireland
- Other (please specify)

What is your highest level of education?

- No formal education
- GCSEs or equivalent
- A-Levels or equivalent
- Bachelor's degree
- Postgraduate degree
- Other (please specify)

What is your current employment status?

- Employed full-time
- Employed part-time
- Self-employed
- Unemployed
- Student
- Retired
- Other (please specify)

Do you have any financial dependents? (*e.g., children, partner, relatives financially dependent on you*)

- Yes
- No

What is your marital status?

- Single
- Married
- In a relationship
- Divorced
- Widowed
- Other (please specify)

Do you rent or own your home?

- Rent
- Own
- Live with family/friends
- Other (please specify)

Do you have any long-term health conditions, including mental health issues, that affect your daily life?

- Yes
- No
- Prefer not to say

How would you describe your financial situation?

- Comfortable
- Managing okay
- Struggling

- In financial difficulty
- 1. In the last 12 months, have you spent money on any of the following? (Tick all that apply)**
 - National Lottery draws
 - Scratchcards
 - Bingo (online or in-person)
 - Sports betting (with a bookmaker, in-shop or online)
 - Online slot machines, casino games, or poker
 - Horse or greyhound racing bets
 - Private betting (e.g., with friends, poker games)
 - Other (please specify)
 - I have not spent money on gambling in the past 12 months
- 2. In the last 4 weeks, have you spent money on any gambling activities? (Tick all that apply)**
 - National Lottery draws
 - Scratchcards
 - Bingo (online or in-person)
 - Sports betting (with a bookmaker, in-shop or online)
 - Online slot machines, casino games, or poker
 - Horse or greyhound racing bets
 - Private betting (e.g., with friends, poker games)
 - Other (please specify)
 - I have not spent money on gambling in the past 12 months
- 3. How do you usually place your bets?**
 - Online
 - In a physical betting shop
 - Both online and in-shop
- 4. How often do you participate in any form of gambling?**
 - Every day
 - A few times a week
 - Once a week
 - A few times a month
 - Once a month or less
- 5. In the past 7 days, how many days have you gambled?**
- 6. When you gamble, how do you usually decide how much to spend? (Tick all that apply)**
 - I set a budget in advance and stick to it
 - I sometimes exceed my planned budget
 - I adjust based on wins/losses during play
 - I don't think about a budget when gambling
- 7. Have you ever continued gambling longer than you originally intended?**
 - Yes, occasionally
 - Yes, often
 - No
- 8. Have you seen gambling advertisements in the past month on any of the following?**
 - Social media (Facebook, Instagram, Twitter, TikTok)
 - TV or online streaming (YouTube, Netflix, etc.)
 - Websites or apps unrelated to gambling
 - Billboards or posters
 - None of the above

9. Have you ever felt encouraged to gamble because of an ad, bonus offer, or promotion?

- ☐ Yes, multiple times
- ☐ Yes, but only once or twice
- ☐ No

10. Do you believe gambling is generally fair and trustworthy?

- ☐ Yes
- ☐ No
- ☐ Not sure

11. Have you heard of self-exclusion programs that allow people to block themselves from gambling?

- ☐ Yes
- ☐ No

12. Have you ever used deposit limits or gambling block features to control your spending?

- ☐ Yes
- ☐ No

13. Have you ever gambled to escape stress, anxiety, or problems in your life?

- ☐ Yes, frequently
- ☐ Yes, sometimes
- ☐ No

14. Have you ever tried to cut back on gambling but found it difficult?

- ☐ Yes
- ☐ No

15. Have you experienced arguments or relationship issues because of gambling?

- ☐ Yes
- ☐ No

Youth Gambling

Have you spent your own money on any of the following in the past 12 months? (Tick all that apply)

- ☐ Fruit machines (arcade, claw grabbers, etc.)
- ☐ Betting with friends/family (e.g., FIFA, poker, card games)
- ☐ Online loot boxes or in-game purchases with chance-based rewards
- ☐ Other (please specify)
- ☐ I have not spent money on gambling

Have you ever tried to win back money you lost while gambling?

- ☐ Yes
- ☐ No

Have you ever hidden or lied about your gambling from a parent or teacher?

- ☐ Yes
- ☐ No

Impact on Others

Have you ever felt affected by someone else's gambling (e.g., a family member, friend, or partner)?

- ☐ Yes, negatively
- ☐ Yes, but not negatively
- ☐ No

If yes, how has someone else's gambling affected you? (Tick all that apply)

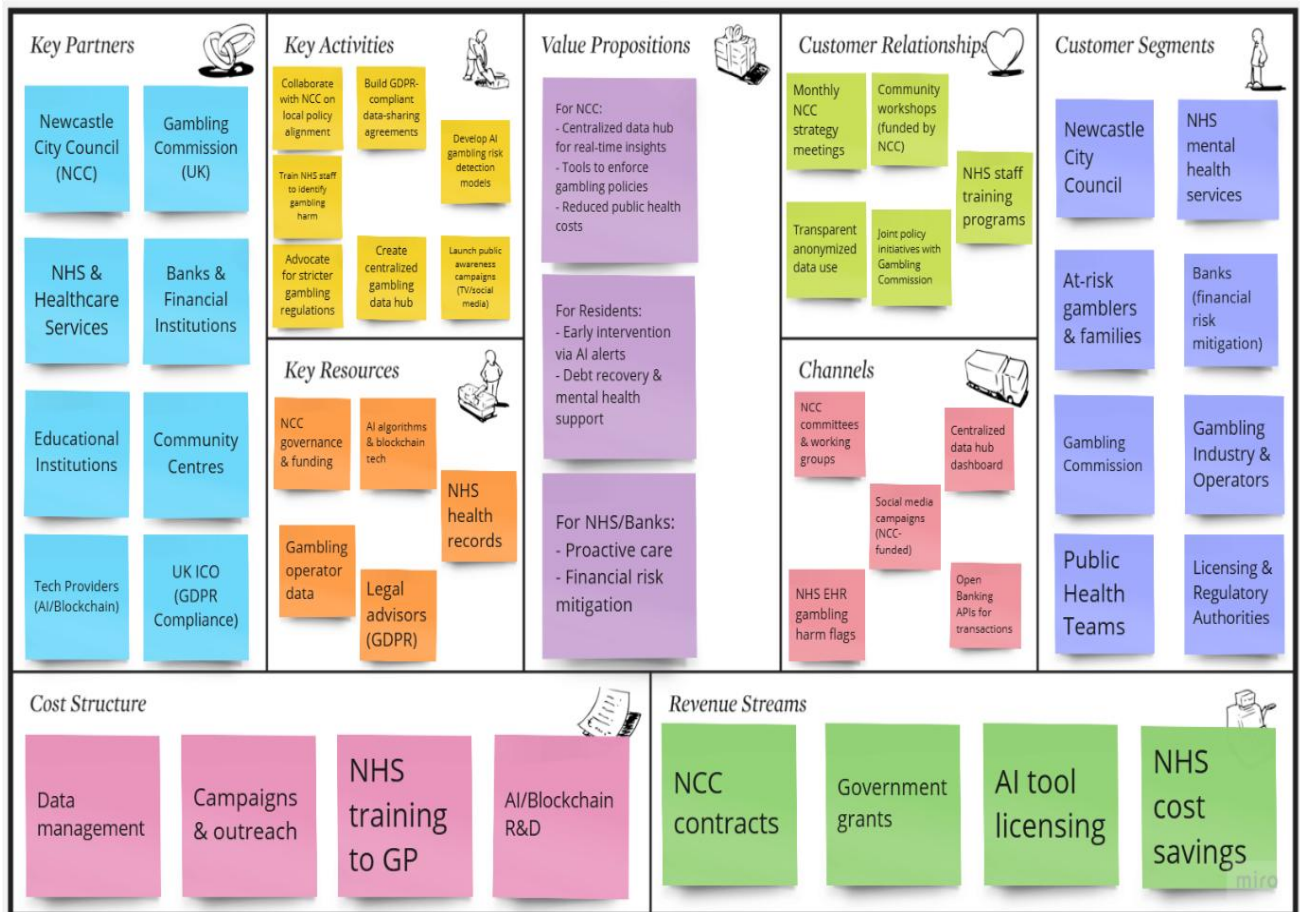
- ☐ Financial stress (e.g., family money spent on gambling)
- ☐ Emotional stress (e.g., arguments, anxiety)
- ☐ Relationship strain (e.g., trust issues, broken commitments)

- Other (please specify)

Thank you for taking the time to complete this survey. Your responses will help researchers better understand gambling habits, the role of advertising, the impacts on individuals and families, and ways to improve responsible gambling support.

If you are concerned about your gambling or someone else's, resources such as GamCare and BeGambleAware provide confidential advice and support.

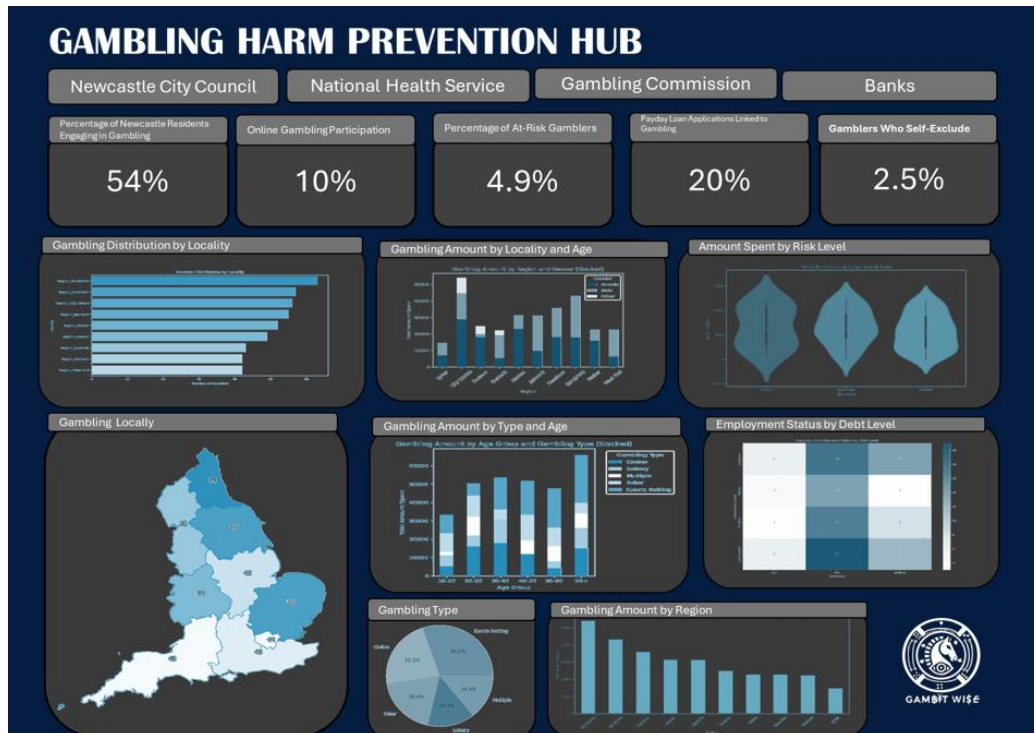
Appendix D: Business Model Canvas



Using the Miro tool, we designed a Business Model Canvas as shown below to create a framework for our solution through a multi-stakeholder approach. More specifically, the model highlights our key partners (e.g., Newcastle City Council, Gambling Commission, NHS, financial institutions), the key activities that are needed for the implementation of our proposed solution (developing AI-driven risk detection tools, advocating for stricter regulations, and creating a centralized gambling data hub), the Customer Relationships and Channels (transparent data use, NHS training programs, and public awareness campaigns), the Cost Structure and Revenue Streams (government funding, NHS cost savings, and AI tool licensing) and our value proposition (real-time insights, early intervention for at-risk gamblers, and financial risk mitigation).

Appendix E: Analysis of Synthesised Data

1. Dashboard



This dashboard represents part of the implementation of our solution and more specifically a data-driven approach to prevent gambling-related harm. Data for the database we suggested are being collected from Newcastle City Council, the National Health Service, the Gambling Commission, and Banks. Then through analysis, we get the visualisations shown in the dashboard showing statistics related to gambling behaviours and harms. The visualisation includes regional gambling distribution, spending trends by age and risk level, employment status correlations, and self-exclusion rates. Through this, useful insights can be extracted that can be important to make further changes in regulations, request funding, and enhance awareness campaigns. Generally, the dashboard can result in better decision-making strategies regarding gambling-related activities.

2. Code for the analysis of Synthesised Data

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
# Load the dataset
gambling_data = pd.read_csv('/content/gambling_data.csv')
# Data Cleaning and Preparation (Handle Missing Values and Data Types)
gambling_data.dropna(inplace=True) # Remove rows with missing values
# Ensure 'Amount_Spent' is numeric
gambling_data['Amount_Spent'] = pd.to_numeric(gambling_data['Amount_Spent'], errors='coerce')
gambling_data.dropna(subset=['Amount_Spent'], inplace=True)
# Convert 'Age' to integer
gambling_data['Age'] = gambling_data['Age'].astype(int)
# Define age groups
def categorize_age(age):
    if age <= 25:
        return '18-25'
    elif age <= 35:
        return '26-35'
    elif age <= 45:
        return '36-45'
    elif age <= 55:
        return '46-55'
```

```

elif age <= 65:
    return '56-65'
else:
    return '66+'
gambling_data['Age_Group'] = gambling_data['Age'].apply(categorize_age)

# 1. Employment Status and Debt Level (Heatmap)
employment_debt = gambling_data.groupby(['Employment_Status', 'Debt_Level'])['User_ID'].count().unstack().fillna(0)

plt.figure(figsize=(12, 8))
sns.heatmap(employment_debt, annot=True, fmt=".0f", cmap="YlGnBu")
plt.title('Heatmap of Employment Status vs. Debt Level')
plt.xlabel('Debt Level')
plt.ylabel('Employment Status')
plt.tight_layout()
plt.show()

# 2. Region and Time of Day (Grouped Bar Chart)
region_time = gambling_data.groupby(['Region', 'Time_of_Day'])['Amount_Spent'].sum().unstack()

plt.figure(figsize=(16, 8))
region_time.plot(kind='bar', colormap='Paired')
plt.title('Gambling Amount by Region and Time of Day (Grouped)')
plt.xlabel('Region')
plt.ylabel('Total Amount Spent')
plt.xticks(rotation=45)
plt.legend(title='Time of Day')
plt.tight_layout()
plt.show()

# 3. Correlation Heatmap of Numerical Features
numerical_data = gambling_data[['Age', 'Betting_Frequency', 'Amount_Spent']]
correlation_matrix = numerical_data.corr()

plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm", linewidths=.5)
plt.title('Correlation Heatmap of Numerical Features')
plt.tight_layout()
plt.show()

# -- Advanced Visualizations and Statistical Analysis --

from sklearn.preprocessing import LabelEncoder
from scipy.stats import chi2_contingency
import numpy as np

# 4. Parallel Coordinates Plot (Multiple Factors)
# Select relevant categorical columns
categorical_cols = ['Region', 'Gender', 'Employment_Status', 'Family_Status', 'Risk_Level']
data_subset = gambling_data[categorical_cols + ['Amount_Spent']]

# 5. Encode categorical variables to numeric
for col in categorical_cols:
    le = LabelEncoder()
    data_subset[col] = le.fit_transform(data_subset[col])

# Normalize Amount_Spent
data_subset['Amount_Spent'] = (data_subset['Amount_Spent'] - data_subset['Amount_Spent'].min()) /
(data_subset['Amount_Spent'].max() - data_subset['Amount_Spent'].min())

plt.figure(figsize=(16, 8))
pd.plotting.parallel_coordinates(data_subset, class_column='Risk_Level', colormap='viridis')
plt.title('Parallel Coordinates Plot of Gambling Factors')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()

# 6. Residual Analysis for Visualization
residuals = (contingency_table - expected) / np.sqrt(expected)
plt.figure(figsize=(12, 8))
sns.heatmap(residuals, annot=True, cmap='coolwarm', center=0)
plt.title('Residual Heatmap: Region vs. Risk Level')

```

```

plt.tight_layout()
plt.show()

# 7. Factor Analysis inspired Visualization (requires more involved statistical modeling - conceptual example)
# This is a placeholder for a more complex analysis. Real Factor Analysis would require
# more data preparation, scaling, and consideration of factor loadings. This example creates a
# simplified scatter plot based on a conceptual "risk factor" score.

# 8. Violin Plot of Amount Spent by Risk Level
plt.figure(figsize=(10, 6))
sns.violinplot(x='Risk_Level', y='Amount_Spent', data=gambling_data, palette='muted')
plt.title('Violin Plot of Amount Spent by Risk Level')
plt.xlabel('Risk Level')
plt.ylabel('Amount Spent')
plt.tight_layout()
plt.show()

# Data Cleaning and Preprocessing
# Handle missing values
data.fillna(data.mode().iloc[0], inplace=True)

# 9. Perform one-hot encoding on categorical features
categorical_features = ['Gender', 'Employment_Status', 'Income_Level', 'Family_Status', 'Region', 'Gambling_Device',
'Self_Exclusion_History', 'Debt_Level', 'Mental_Health_Issues', 'Gambling_Type', 'Time_of_Day', 'Alcohol_Addiction',
'Drug_Addiction', 'Smoker']
data = pd.get_dummies(data, columns=categorical_features, drop_first=True)

# Define features and target
features = [col for col in data.columns if col not in ['User_ID', 'Risk_Level']]
target = 'Risk_Level'

# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(data[features], data[target], test_size=0.3, random_state=42)

# Train a Random Forest Classifier model
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# Make predictions
predictions = model.predict(X_test)
# Evaluate the model
print(classification_report(y_test, predictions))

# Descriptive Analysis

# Most frequent locality
most_frequent_locality = data[[col for col in data.columns if 'Region_' in col]].sum().idxmax()
print(f"The most frequent locality is: {most_frequent_locality}")
# Average age of gamblers
average_age = data['Age'].mean()
print(f"The average age of gamblers is: {average_age:.2f}")

# Gender distribution
gender_distribution = data[[col for col in data.columns if 'Gender_' in col]].sum() / len(data) * 100
print("Gender distribution:")
print(gender_distribution)

# Addict behavior based on Alcohol and Drug Addiction
alcohol_addiction_percentage = data[[col for col in data.columns if 'Alcohol_Addiction_' in col]].sum() / len(data) * 100
print("Alcohol addiction percentage:")
print(alcohol_addiction_percentage)
drug_addiction_percentage = data[[col for col in data.columns if 'Drug_Addiction_' in col]].sum() / len(data) * 100
print("Drug addiction percentage:")
print(drug_addiction_percentage)

# Risk Level distribution
risk_level_distribution = data['Risk_Level'].value_counts(normalize=True) * 100
print("Risk Level distribution:")
print(risk_level_distribution)
# Important Features
feature_importances = pd.Series(model.feature_importances_, index=X_train.columns)
feature_importances = feature_importances.sort_values(ascending=False)

```

```

# Print the most important features
print("Most important features:")
print(feature_importances.head(10))

# Visualizations
# Locality Distribution
locality_cols = [col for col in data.columns if 'Region_' in col]
locality_counts = data[locality_cols].sum().sort_values(ascending=False)
plt.figure(figsize=(12, 6))
sns.barplot(x=locality_counts.values, y=locality_counts.index, palette='viridis')
plt.title('Gambler Distribution by Locality')
plt.xlabel('Number of Gamblers')
plt.ylabel('Locality')
plt.show()

# Risk Level Distribution
plt.figure(figsize=(8, 6))
data['Risk_Level'].value_counts().plot(kind='bar')
plt.title('Risk Level Distribution')
plt.xlabel('Risk Level')
plt.ylabel('Number of Gamblers')
plt.xticks(rotation=45)
plt.show()
# Addiction Distribution
alcohol_cols = [col for col in data.columns if 'Alcohol_Addiction_' in col]
drug_cols = [col for col in data.columns if 'Drug_Addiction_' in col]

# 1. Which locality has the most online gamblers?
online_devices = ['Mobile', 'Desktop', 'Tablet']
online_gamblers = df[df['Gambling_Device'].isin(online_devices)]
top_locality = online_gamblers['Region'].value_counts().idxmax()
print(f"Locality with most online gamblers: {top_locality}")

# 2. Demographics of top locality
top_region_data = online_gamblers[online_gamblers['Region'] == top_locality]
print(f"\nAverage age in {top_locality}: {top_region_data['Age'].mean():.1f} years")
gender_dist = top_region_data['Gender'].value_counts(normalize=True)
print("\nGender distribution:")
print(gender_dist)

# 3. General behaviors
behavior_analysis = top_region_data[['Betting_Frequency', 'Amount_Spent', 'Gambling_Type', 'Time_of_Day']]
print("\nCommon gambling types:")
print(behavior_analysis['Gambling_Type'].value_counts())
print("\nAverage betting frequency:", behavior_analysis['Betting_Frequency'].mean())
print("Average amount spent:", behavior_analysis['Amount_Spent'].mean())

# 4. Addiction indicators
addiction_features = top_region_data[['Self_Exclusion_History', 'Debt_Level', 'Mental_Health_Issues',
'Risk_Level', 'Alcohol_Addiction', 'Drug_Addiction']]
print("\nAddiction indicators:")
print("Self-exclusion history:", addiction_features['Self_Exclusion_History'].mean())
print("Debt levels distribution:")
print(addiction_features['Debt_Level'].value_counts())
print("\nMental health issues distribution:")
print(addiction_features['Mental_Health_Issues'].value_counts())

# Visualization
plt.figure(figsize=(15, 10))
plt.subplot(2, 2, 1)
sns.countplot(data=online_gamblers, x='Region', order=online_gamblers['Region'].value_counts().index)
plt.title('Online Gamblers by Region')
plt.xticks(rotation=45)
plt.subplot(2, 2, 2)
sns.histplot(top_region_data['Age'], kde=True)
plt.title(f'Age Distribution in {top_locality}')
plt.subplot(2, 2, 3)
top_region_data['Gambling_Type'].value_counts().plot.pie(autopct='%1.1f%%')
plt.ylabel("")

```

```

plt.title('Gambling Types')
plt.subplot(2, 2, 4)
sns.countplot(data=addiction_features, x='Risk_Level')
plt.title('Risk Level Distribution')
plt.tight_layout()
plt.show()

# Clustering for behavior patterns (Example)
cluster_data = top_region_data[['Age', 'Betting_Frequency', 'Amount_Spent']]
scaler = StandardScaler()
scaled_data = scaler.fit_transform(cluster_data)
kmeans = KMeans(n_clusters=3, random_state=42)
cluster_labels = kmeans.fit_predict(scaled_data)
top_region_data['Cluster'] = cluster_labels

# Analyze clusters
cluster_analysis = top_region_data.groupby('Cluster').agg({
    'Age': 'mean',
    'Betting_Frequency': 'mean',
    'Amount_Spent': 'mean',
    'Gender': lambda x: x.mode()[0]
}).reset_index()
print("\nCluster Analysis:")
print(cluster_analysis)

```

Appendix F: Minimum Viable Product (MVP) Implementation Plan

Centralized Gambling Harm Prevention Database Proposed by Gambit Wise for Newcastle City Council

1. Objective

To establish a centralized, GDPR-compliant database integrating anonymized data from the NHS, banks, residents, and the UK Gambling Commission. This system will enable proactive regulation, early intervention, and data-driven policymaking to mitigate gambling-related harm. The initiative aligns with Newcastle City Council's broader public health strategy and fills gaps in current gambling harm prevention measures.

2. Key Components of the Centralized Database

Component	Data Sources	Purpose
NHS Health Insights	Mental health records, addiction services	Identify correlations between gambling and mental health decline.
Banking Transactions	Anonymized gambling spending trends	Detect risky financial behaviours (e.g., frequent large bets, debt patterns).
Resident Surveys	Expanded public health surveys	Capture self-reported behaviours and underrepresented demographics.
Gambling Commission	Participation rates, self-exclusion data	Monitor industry trends and compliance.
AI Risk Detection	Machine learning models	Flag high-risk individuals for early intervention.

3. Implementation Steps

Phase 1: Stakeholder Partnerships (Weeks 1-4)

- Secure data-sharing agreements with NHS, banks (e.g., Barclays), and the UK Gambling Commission.
- Develop GDPR-compliant anonymization protocols, audited by the UK Information Commissioner's Office (ICO).
- Establish a transparency policy to address public concerns about data privacy.

Phase 2: Data Integration (Weeks 5-8)

- Build APIs to aggregate anonymized data from NHS electronic health records (EHRs), banking systems, and Gambling Commission reports.
- Standardize data formats (e.g., age brackets, risk categories) for seamless integration.

Phase 3: AI & Analytics Deployment (Weeks 9-12)

- Train AI models to detect high-risk patterns (e.g., frequent large bets, mental health correlations).
- Implement fairness checks to minimize algorithmic bias.

- Pilot real-time alerts with two NHS clinics and one bank for early intervention.

Phase 4: Public Engagement (Weeks 13-16)

- Launch resident surveys via SMS, email, and community centres.
- Refine AI models based on survey insights and stakeholder feedback.

4. Expected Benefits

Stakeholder	Benefit
Newcastle City Council	Real-time oversight of gambling harm trends and evidence-based policy decisions.
Residents	Early access to support services (e.g., debt counselling, mental health care).
NHS & Banks	Tools to proactively mitigate financial and health risks.

5. Challenges & Mitigation

Challenge	Mitigation Strategy
Data fragmentation	Use APIs and ETL pipelines to harmonize data formats.
GDPR compliance risks	Partner with the UK ICO for anonymization audits and implement robust data governance.
Stakeholder resistance	Offer incentives (e.g., reduced regulatory scrutiny for compliant operators).
Public concerns about privacy	Implement an opt-in system for residents and ensure transparency in data usage.
Algorithm bias in AI models	Integrate fairness audits and continuously update models based on diverse data inputs.

6. Success Metrics

Success Metric	Target Outcome	Measurement Approach
Reduction in Gambling Harm Cases	Fewer problem gambling reports	NHS and mental health service data analysis
User Engagement in Self-Exclusion	Increased adoption of self-exclusion tools	Gambling operator reports and surveys
AI Risk Detection Accuracy	Higher precision in identifying at-risk gamblers	Comparison of AI predictions with real-world cases
Policy Impact & Compliance	Stricter affordability checks and advertising regulations	Legal audits, industry reports
Public Awareness & Education	Increased responsible gambling awareness	Survey responses, outreach event participation

7. Supporting Documents

1. Data-Sharing Agreement Template: GDPR-compliant framework for anonymized data exchange.
2. Sample AI Code: Python script for transaction analysis (attached separately).
3. Resident Survey Template: 10 questions on gambling behaviours and awareness.
4. Data Governance Framework: Ensuring compliance and ethical use of data.

8. Visual Representation

Centralized Database Workflow:

1. NHS → Mental health data
2. Banks → Transaction trends
3. Residents → Survey responses
4. Gambling Commission → Regulatory reports
5. AI Layer → Risk detection & alerts

This structured approach ensures that Newcastle City Council can take a proactive role in mitigating gambling-related harm through data-driven insights and early interventions.