Final Report

● Ungraded

1 Day, 17 Hours Late

Group

Sydney Simon
Aarti Bhamidipati
Nicholas Molina
...and 3 more
View or edit group

Total Points

- / 15 pts

Question 1

Report 15 pts

Question assigned to the following page: 1

WeighIn: A Crowdsourced Anonymous Polling Platform

NETS 213 Final Report

Nicholas Molina Sydney Simon Aarti Bhamidipati Brian Lu
Henry Casper Michael Wallison

December 11, 2024

1. Project Overview

Figure 1: WeighIn Logo

WeighIn is a crowdsourcing platform where users can post and respond to meaningful polls anonymously. The project was developed by a team comprising Henry Casper, Sydney Simon, Aarti Bhamidipati, Brian Lu, Nicholas Molina, and Michael Wallison. The primary focus of WeighIn is to address the challenge of indecisiveness by allowing users to ask questions and get the opinions of the crowd. The platform provides a structured way to gather collective insights while maintaining user privacy through anonymity. This approach enables better understanding of community perspectives and facilitates informed decision-making based on crowd wisdom.

There are some applications that are similar to WeighIn such as Yik Yak, SideChat, Reddit Polls, Survey Monkey, and Slido. WeighIn stands out by offering a unique combination of features and a strong focus on user anonymity. Yik Yak and SideChat are anonymous social platforms focusing on community engagement where users make posts. Reddit Polls has community-driven polling with discussion threads. Survey Monkey is for professional survey creation and analysis tools and Slido is an interactive polling app for presentations and events. WeighIn distinguishes itself through strong emphasis on user anonymity, its focus on meaningful questions to help users and real-time result aggregation and visualization of results.

WeighIn functions through a blend of crowd contributions and automated processes. Users can post polls with questions and answer choices anonymously, while other users vote on these polls, contributing to a shared pool of insights. The platform automatically aggregates the results in real-time, visualizes data trends, and ensures user anonymity by stripping identifying information from poll responses. This seamless interaction between crowd and automation enables the platform to deliver accurate and actionable results.

Question assigned to the following page: 1

The bulk of the team's effort went towards engineering the complex system of data collection with the creation of both the website and database. The team created a comprehensive presentation detailing the project's features, methodology, and outcomes. The presentation can be accessed at `https://youtu.be/f2Br3uVawEo`

Finally, the report delves deeper into two critical aspects of the project: Project Analysis and our Technical Challenges. These sections highlight the methods employed to collect and process data as well as the strategies used when creating the application.

2. The Crowd

The WeighIn project utilized a hybrid approach to crowd participation, combining both controlled classroom participants and organic social media users. This dual-pronged strategy enabled the team to gather diverse perspectives while maintaining quality control standards.

The crowd for the WeighIn project consisted of a diverse group, primarily drawn from classmates, friends, and users of the Sidechat platform. For the final project, we used a combination of both simulated and real crowds to gather data. The classroom crowd was simulated, where classmates participated on the website as part of an assignment given by the professor. This allowed the team to test the platform with an initial set of participants. In addition to this simulated crowd, the team also ran a real experiment using Sidechat to reach a broader audience.

To simulate the crowd, the class was tasked with engaging with the website, completing polls, and providing feedback. However, the real crowd component involved reaching out beyond the classroom by posting on Sidechat and encouraging friends to upvote the posts for increased visibility. This process allowed for a more organic and diverse set of participants who were genuinely interested in the polls. The participants who interacted with the platform through Sidechat were those who noticed the posts and were curious about the website and its content. This hybrid approach to crowd sourcing proved effective in gathering meaningful data while maintaining engagement quality. In total, the project had 74 unique participants who contributed to the dataset by responding to polls and interacting with the platform.

3. Incentives and User Engagement

The crowd participates in the WeighIn project for several motivations, including enjoyment, altruism, and the desire to help make decisions. Many participants engage with the polls because they find it fun to read through the questions posted by others and share their own opinions. In addition to enjoyment, some users are motivated by a sense of altruism, wishing to provide feedback that may help others make decisions, especially in cases where the polls present dilemmas or decisions that need to be made. The main incentive for creating a poll is the crowdsourced responses, which provide valuable insights that can help the poll creator make informed choices about what to do in their described situation. While the crowd's participation was driven primarily by intrinsic motivation, such as enjoyment and helping others, the project did not specifically analyze or compare different types of incentives. However, if the crowd was real rather than simulated, incentives like gamification, rewards, or recognition could be considered to further enhance participation. These could include offering points for answering polls or featuring popular users in the app's rankings to encourage ongoing engagement.

Question assigned to the following page: 1

4. Crowd Contributions

The crowd contributes to the project by submitting multiple-choice questions and possible responses in addition to voting on these questions. This process is somewhat automated since the platform aggregates responses and visualizes the results in real time. The automation of poll creation and responding could be extended through the use of large language models (LLMs), which could generate questions and answers. However, while responses can be automated to some degree, the subjective nature of the questions and answers makes full automation difficult.

WeighIn does not involve training machine learning components from crowd data, as the platform primarily focuses on gathering user-generated content for aggregation and analysis. Although the project could benefit from machine learning in the future, such as using algorithms to better understand crowd preferences or predict trends, this was not a part of the current system.

The platform includes a simple but functional user interface designed for crowd interaction. The main page displays all of the active polls on the app, while a profile page shows the polls a user has created and those they have responded to. Additionally, users can create new polls through a dedicated page, register a new account via a sign-up page, or log into their existing accounts. Each poll has its own dedicated page where users can view the poll and see the aggregated responses. This intuitive layout ensures that users can easily navigate the platform and participate in polls with minimal effort.

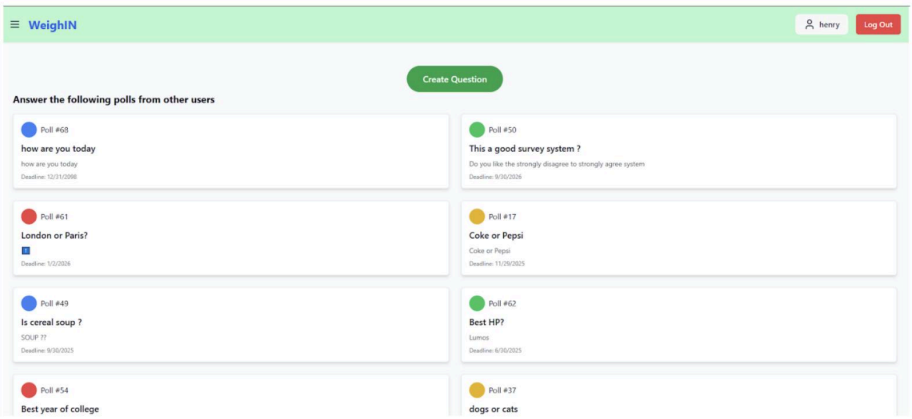

Figure 2: WeighIn Homepage Interface

Question assigned to the following page: 1

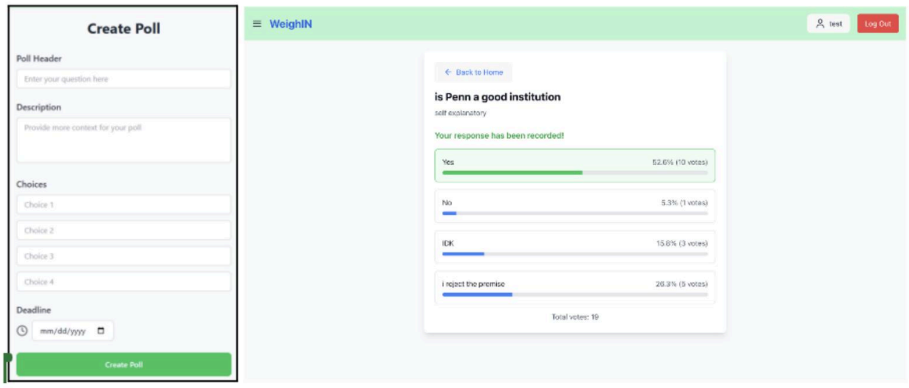

Figure 3: poll creation and results

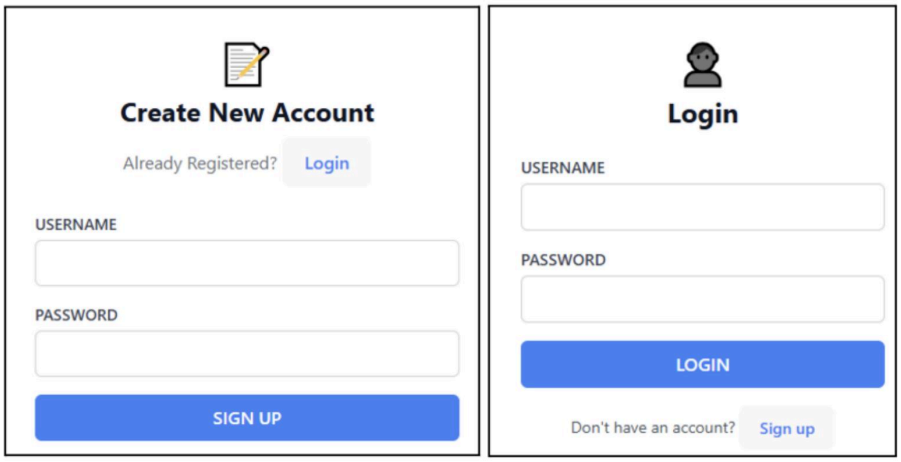

Figure 4: create account and login

5. Ethical Considerations

5.1. Platform Justification and Value Proposition

The team agrees that the application should exist, as long as it provides value by allowing individuals to share and gather opinions on meaningful questions. However, there are some potential ethical concerns related to the platform's operation, mainly the risks posed to participants and the treatment of data.

5.2. Concerns

One significant concern is that workers—users participating in polls—may be exposed to harmful text or content that slips through our quality control filters. While we have implemented measures to minimize this risk, there is still a possibility that harmful or distressing content could reach users. Such content could cause psychological distress, particularly if it involves sensitive topics

Question assigned to the following page: 1

or offensive material. Therefore, it is important to continually improve our content moderation systems to mitigate this potential harm.

Compensation-wise, the platform does not offer financial rewards for participation. Instead, users are incentivized by intrinsic motivations such as enjoyment and the desire to help others make decisions as well as interest in other people's lives. Participants engage because they find it fun to contribute to others' decision-making processes. While this approach fosters voluntary engagement, it may also raise questions about fair compensation, especially if users are dedicating significant time to the platform.

5.3. Machine Learning

Since WeighIn does not involve machine learning components, we do not need to have concerns related to the ethical use of data in machine learning. However, should the project evolve to include machine learning, we will take steps to ensure that the data used in development will only come with consent or other sources that are not related to the users.

5.4. Conclusion of Evaluation

Finally, our evaluation of the platform's effectiveness, including sentiment analysis, appears to be scientifically sound. After checking the questions, there were not any strong negative emotions for the questions so majority does seem either positive or no real correlation between negative sentiment. As the project evolves, we will continue to assess its impact and address any emerging ethical concerns to ensure that the platform remains fair, safe, and beneficial for all users.

6. Skills

The crowd workers participating in WeighIn do not require specialized skills to engage with the platform. The primary requirement for participation is simply the ability to form and express an opinion on the polls. To contribute, users need basic reading skills to understand the poll questions and mild technology skills to navigate the platform. These skills are accessible to a broad range of individuals, ensuring that the platform can engage people of varying backgrounds, ages, and levels of expertise.

Given the simplicity of the task—answering multiple-choice questions or creating polls—there is no formal limit on the types of skills users can possess. Participants can be of any age or intelligence level, which leads to a wide variation in the skills and backgrounds of individual workers. This diversity means that the crowd's capabilities are not uniform, but instead reflect a broad spectrum of skills and experiences.

We did not analyze the skills of the crowd in detail for this project. Since participation is based on user opinion and engagement rather than specific competencies, there was no need to assess individual skills or performance. However, had we conducted a deeper analysis, we might have explored factors such as the quality of responses or the time taken to answer questions.

Question assigned to the following page: 1

7. Quality Control

7.1. Overview

Quality control is a big component of the WeighIn platform, ensuring data integrity and user safety. Our quality control framework addresses two primary challenges: authentication of legitimate users (no bots) and content moderation.

7.2. Quality Control Architecture

7.2.1 Authentication System

The platform implements a two-layer authentication approach:

- Username/password verification
- Randomized Quality Control Question after login

7.2.2 Content Moderation System

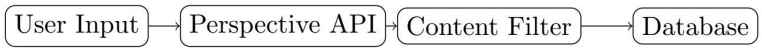

Figure 5: Content Moderation Pipeline

7.3. Implementation Details

To prevent botting and ensure that responses are from real users, we introduced a quality control question during the login process. This randomized question helps verify that participants are paying attention and not automated bots. Additionally, to prevent toxic content and uphold the platform's goal of fostering a social and fun community, we utilized Perspective API to filter out profanity and inappropriate language from both the poll questions and responses. This ensures that our database remains clean and free from harmful material.

While these quality control measures are effective in filtering out inappropriate language and ensuring the legitimacy of users, they are not perfect. For example, the system is not able to detect and filter new slang or controversial topics that might still carry offensive or inappropriate meanings. Nevertheless, our quality control process was crucial in maintaining a safe and reliable platform for all users. As we continue to develop the platform, future improvements could involve expanding our ability to detect and handle emerging language trends or more complex forms of toxic content.

Since the questions and answers are filtered before being submitted to the database, we are confident that the filtering process ensured that the majority of content was appropriate and that responses came from legitimate users.

8. Data Aggregation and Analysis

8.1. Aggregation Methodology

The WeighIn platform implements a real-time data aggregation system that processes and displays crowd responses dynamically. This section explains our approach to data collection, processing,

Question assigned to the following page: 1

and visualization.

The results from the crowd on the WeighIn platform are aggregated by calculating the number of votes for each choice in a poll and displaying the corresponding percentages once users have voted. This allows participants to see the distribution of responses in real-time, giving them insight into how others feel about the question posed in the poll. The aggregation process is straightforward, with the platform automatically updating the percentages as users continue to submit their votes.

8.2. Content Category Analysis

We aggregated the questions into categories and found the average number of responses for each group of questions under 6 categories:

1. **Philosophical and Conceptual Debates**
2. **Academic and Career Guidance**
3. **Food and Dining Choices**
4. **Institutional, Societal and Political Issues**
5. **Entertainment and Cultural Preferences**
6. **General Preferences and Lifestyle**

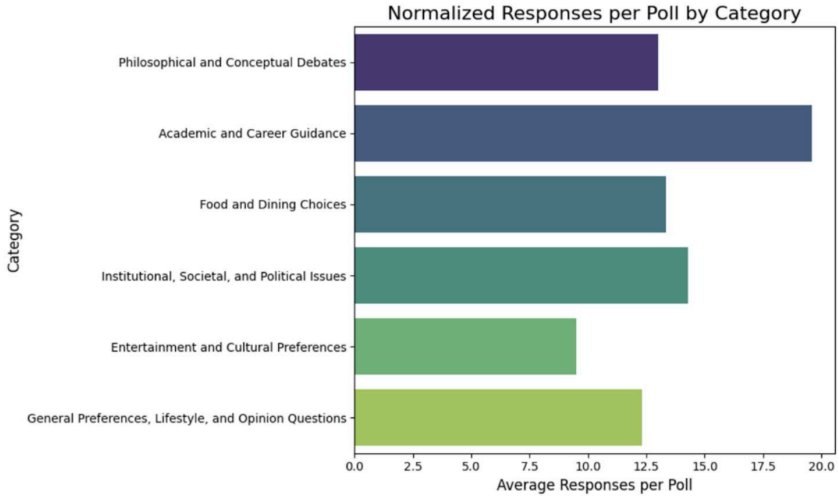

Figure 6: Average Response Distribution Across Categories

Question assigned to the following page: 1

8.3. Visualization System

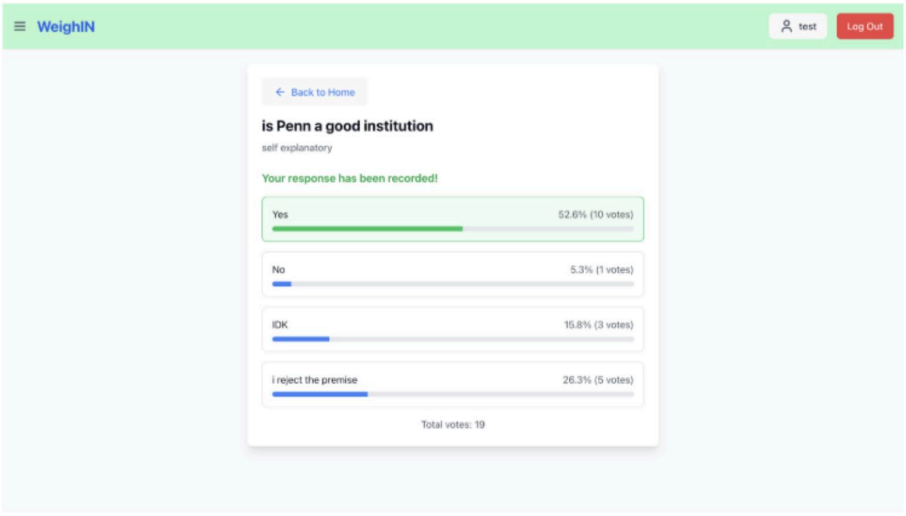

Figure 7: Poll Results User Interface

Once a user answers one of the polls, they will see the above interface. Each of the 4 choices will have the number of votes and the percentage of people who selected it. The choice with the highest percentage of responses will be highlighted in green.

9. Scalability Analysis

9.1. Global Scale Potential

WeighIn's mission is to address the challenge of indecisiveness by allowing users to ask questions and get the opinions of the crowd. This can be extended to anyone, anywhere in the world, where they can ask questions and answer questions posed by others for a more vast and diverse user-base. Our platform has the potential to facilitate the aggregation of diverse opinions from a global crowd, making it an open and scalable system for collecting and sharing insights. With this universal accessibility, the potential for large-scale participation is significant.

9.2. Benefits of Scale

Our project would certainly benefit from contributions by thousands, or even millions, of people. A larger crowd would mean a more diverse set of questions, with greater variety and complexity. Additionally, having more responses for each question would enrich the aggregated results, drawing from a wider pool of intelligence and life experience. This increased engagement would ultimately enhance the quality of insights generated on the platform.

Question assigned to the following page: 1

9.3. Technical Scaling Challenges

9.3.1 Infrastructure Requirements

Scaling the platform to accommodate such a large crowd would introduce several challenges. One of the key challenges is managing the vast amount of data that would come from a growing number of polls and users. The platform would need efficient systems for storing and processing this data, including responses, user profiles, and poll results. Scaling infrastructure would be necessary to handle the increased load while maintaining performance and reliability.

9.4. Infrastructure Cost Analysis

9.4.1 AWS Service Requirements

Service	**Purpose**	**Scaling Considerations**
EC2	Compute	Auto-scaling groups for dynamic load handling
RDS	Database	Scalable storage and read replicas
Cognito	Authentication	User pool management and federation
CloudWatch	Monitoring	Performance tracking and alerting

Table 1: AWS Infrastructure Components

9.4.2 Cost Projections

We conducted a cost analysis to assess the feasibility of scaling up the platform. We explored the use of AWS services such as EC2 (compute), RDS (database), Cognito (authentication), and CloudWatch (monitoring) to support the infrastructure needs of a larger user base. Using the AWS Pricing Calculator, we estimated the monthly costs for infrastructure needed to support a user base of up to 250,000 users—the approximate size of Reddit's r/polls subreddit, which shares similar functionality to our platform. Based on rough calculations, the monthly infrastructure spend for this user base would likely be in the range of a few thousand dollars, covering the compute resources, storage, authentication, and monitoring required for smooth operation.

For a target scale of 250,000 users:

- **Monthly Infrastructure Costs:**
  - Compute: $1,000-$1,500
  - Storage: $500-$800
  - Authentication: $200-$300
  - Monitoring: $100-$200
  - **Total Estimated Range:** $2,000-$3,000/month

While the cost of scaling up is manageable, the challenge lies in ensuring that the system remains efficient and cost-effective as user numbers grow. Further analysis and planning would be needed to refine the infrastructure and optimize costs, ensuring that WeighIn can scale sustainably as the user base expands.

Question assigned to the following page: 1

10. Project Analysis and Results

The code to conduct the analysis is here: `https://colab.research.google.com/drive/1IAKaiNvtGM5Oornehlk5?usp=sharing` and the polls.db file can be found in the GitHub.

WeighIn was successful in achieving its core objectives, as the platform is up and running and we received substantial engagement from users. The crowd was active, providing a variety of questions and responses that demonstrated both creativity and involvement. This feedback indicates that the platform resonated with users and met their needs for a space to engage in meaningful and anonymous polling. Despite challenges in scaling, the system worked as expected, and the project provided valuable insights into how people interact with crowdsourced polls.

One of the main challenges we faced was determining how to gather data from external users, outside of our classroom. Finding ways to incentivize the broader public to participate in the polls was difficult, as we lacked a well-established user base initially. However, by posting on platforms like Sidechat and encouraging friends to share the app, we managed to gather a real crowd.

In terms of changes, the final product remained largely aligned with our original proposal, but there were some adjustments made during development. For example, we initially planned to implement quality control checks during the poll process itself. However, we realized this could be intrusive and potentially annoying for users, so we moved the quality control check to the login process instead. Additionally, we had planned to use deadlines for polls, but we opted not to implement this feature in the final version due to concerns over the availability of enough questions to keep the platform active.

There were some limitations to the product that hindered its full potential. Time constraints prevented us from marketing the platform as effectively as we would have liked, limiting our ability to attract a larger crowd outside of the classroom. Additionally, the platform was hosted on free resources, which posed scalability challenges. With more users, the system could face performance bottlenecks and require more robust infrastructure to handle the increased load. These limitations would need to be addressed as the project moves forward, particularly if we aim to scale the platform for broader use.

In terms of analysis, the questions posted were split into 6 categories based on question type and the number of questions that fit into each category were counted up. Then they were presented in a bar chart.

Question assigned to the following page: 1

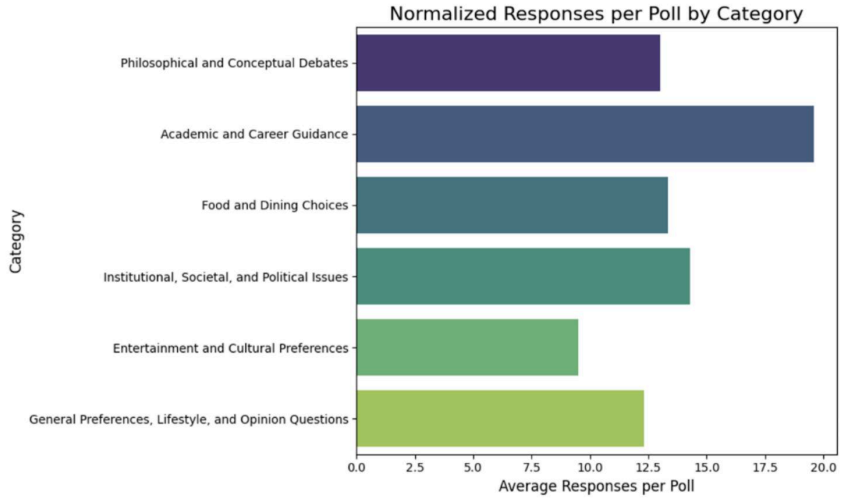

Figure 8: bar chart with categories

Additionally, libraries like textblob and nltk were used to perform a sentiment analysis on the text from the questions and the text from the answer choices (both were analyzed separately in two bar graphs). As we can see, the majority of the descriptions are neutral (around 0) as many did not elaborate on the question, or have one word descriptions. For the sentiment on the questions, majority are also neutral, however we can also see some questions that have positive sentiment, which supports our goal of fostering a community around helping promote emotional question framing (making sure users are expressing their emotions in an adequate way for others to understand).

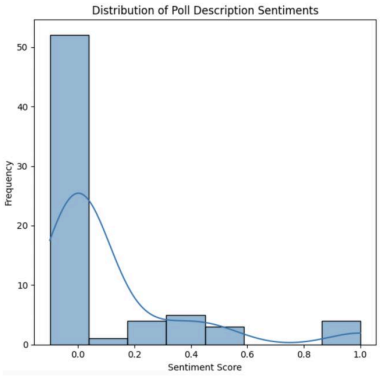

Figure 9: sentiment analysis for descriptions

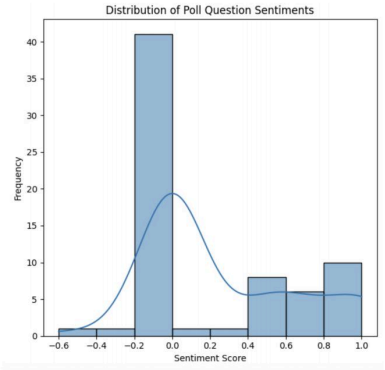

Figure 10: sentiment analysis for categories

We also did an analysis to see how poll creators align with their audience across different topic categories. Food and Dining Choices show nearly perfect agreement, suggesting universal preferences or clear consensus in food-related decisions. Institutional and Political Issues follow with strong agreement (∼60%), indicating shared viewpoints within the community on these typically controversial topics. General Preferences and Academic/Career Guidance show moderate agreement rates (35-40%), reflecting natural variation in personal choices and perspectives. En-

Question assigned to the following page: 1

tertainment and Cultural Preferences have the lowest visible agreement rate (∼25%), suggesting significant diversity in taste and preferences. This data helps understand where community perspectives align and differ, which could be valuable for improving future polls and understanding group dynamics.

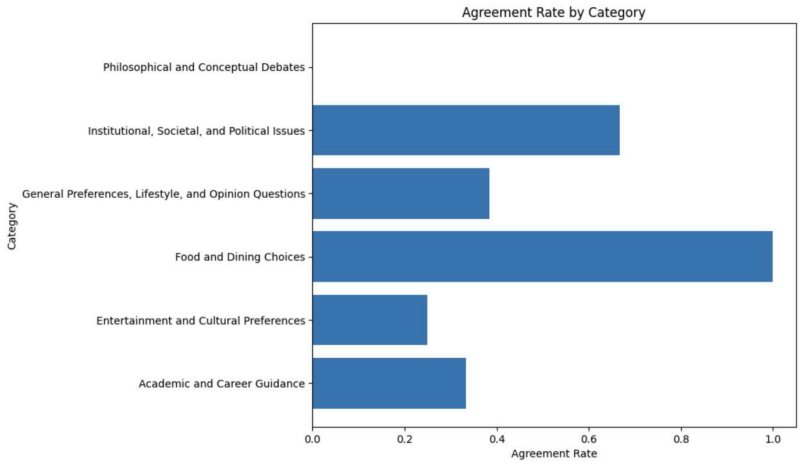

Figure 11: User agreement rate by category

Overall, the project was successful in its execution, but challenges related to scaling, user engagement, and resource limitations highlight areas for future improvement and development.

This comprehensive analysis demonstrates both the success of the WeighIn platform in meeting its core objectives and the opportunities for future enhancement and expansion.

11. Technical Implementation and Challenges

11.1. System Architecture Overview

The WeighIn project required a substantial technical component, with significant software engineering work involved in developing both the frontend and backend systems. We built the entire application, including data collection components, from the ground up. Although we did not need to learn entirely new programming languages, we did work with libraries and APIs that were unfamiliar to some team members at the start. These included SQLite, Flask, and Vite. Additionally, we integrated Google's Perspective API for quality control to ensure that content shared on the platform remained appropriate and non-toxic.

11.1.1 Technology Stack

- **Frontend Technologies:**
  - React.js framework
  - Vite build tool
  - TypeScript for type safety
  - Tailwind CSS for styling

Question assigned to the following page: 1

- **Backend Technologies:**
  - Flask web framework
  - SQLite database
  - Google Perspective API
  - RESTful API architecture

11.2. Deployment Architecture

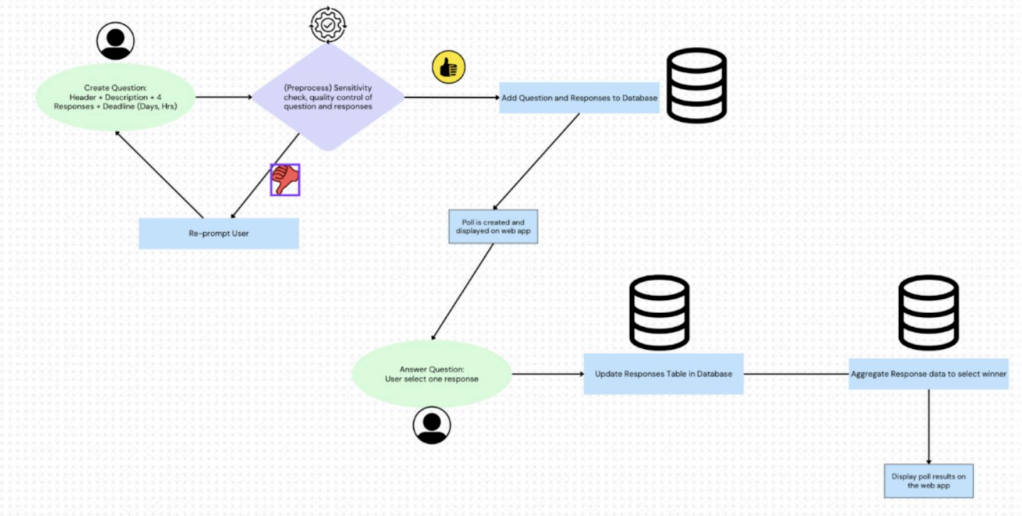

Figure 12: Production Deployment Architecture

11.3. Major Technical Challenges

The largest technical challenge we faced was hosting the app so that it would be accessible to users on the internet, rather than just running locally. The frontend, built with React and Vite, and the backend, developed with Flask, both needed to be hosted in a way that would support their interaction and allow for a seamless user experience. We had to ensure that the code was deployment-ready, which involved making adjustments to our development environment and finding suitable hosting services for both components.

1. **Environment Configuration**
   - Development to production transition
   - Environment variable management
   - Cross-origin resource sharing (CORS)
2. **Service Integration**
   - Frontend-backend communication
   - API endpoint configuration
   - Database connection management

Question assigned to the following page: 1

3. **Platform Selection**
    - Cost considerations
    - Service limitations
    - Scalability requirements

11.4. Technical Solutions

To overcome this challenge, we found free hosting solutions that met our needs. For the frontend, we used Vercel, a platform that offers easy deployment of websites like React applications built with Vite and Typescript. This service allowed us to quickly deploy and manage our frontend. For the backend, we turned to PythonAnywhere, a free platform that supports Flask applications. Using these services required us to familiarize ourselves with their deployment processes and tools, but they ultimately provided an effective, cost-free solution to hosting both the frontend and backend.

- **GitHub Repo:** `https://github.com/sydsimon/WeighIn`
- **Frontend Deployment (Vercel) at** `https://weigh-in.vercel.app/`**:**
    - Automatic build optimization
    - CDN distribution
    - SSL certification
    - Continuous deployment
- **Backend Deployment (PythonAnywhere) at** `https://ngmolina1.pythonanywhere.com/`**:**
    - WSGI server configuration
    - Database hosting
    - API endpoint exposure
    - Resource management

These hosting solutions allowed us to deploy and run the app without incurring significant costs, though we did face limitations in terms of scalability. Going forward, if we were to scale the project, we would need to explore more robust hosting options that could handle increased traffic and resource demands such as AWS as mentioned in the previous section.

11.5. Technical Learnings

Here are the key technical skills and knowledge acquired through the completion of the WeighIn project:

- **Framework Proficiency:**
    - Flask application development
    - React component architecture
    - Vite build system
- **API Integration:**

Question assigned to the following page: 1

  - REST API development
  - External service integration
  - Authentication implementation

- **Deployment Expertise:**
  - Environment configuration
  - Production deployment

Our technical implementation demonstrates the successful deployment of a full-stack crowdsourcing application while highlighting important learnings and areas for future development.

12. Conclusion

WeighIn has proven to be a robust platform that effectively addresses indecisiveness by enabling users to create and respond to meaningful, anonymous polls. Through a combination of innovative features, real-time result aggregation, and a strong emphasis on user privacy, the project has successfully demonstrated the value of crowdsourcing in decision-making. By engaging a diverse crowd and leveraging both simulated and real user interactions, the platform provided insightful data and highlighted the potential for scalable growth. As we look ahead, we see opportunities to expand its features, enhance content moderation, and make the platform even more engaging and accessible.