ULTIMATE RECIPE – COOKING AND GROCERY SHOPPING HELPER USING ML

A Capstone Proposal

Carlos de la Roca, Ksenia Koldaeva, Paulo Wu Liu, Miki Fukushima

Presented to the

Faculty of CST 499 at

California State University, Monterey Bay

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Science

in

Computer Science

---

by

Orion Analytics

Fall 2022

EXECUTIVE SUMMARY OF PROPOSAL

Ultimate Recipe - Cooking and Grocery Shopping Helper using ML

by

Orion Analytics

Bachelor of Science in Computer Science

California State University Monterey Bay, 2022

The purpose of this project is to provide a web-based application that can change your experience with cooking. Based on a survey done by Huffpost in 2011, approximately one third of Americans do not know how to cook which is ironic considering eating is an essential part of our everyday lives. Luckily, we live in a modern society where food is accessible at places like restaurants and grocery delis, but eating out can be both expensive and unhealthy. This application; Ultimate Recipe will enhance your cooking experience.

Ultimate Recipe will be powered by an already trained machine learning model and it will be able to suggest the type of dishes you can make based on the ingredients (or keywords of ingredients) you have access to. The application will contain an easy-to-use graphical user interface with a lot of graphics so people from any age should feel comfortable navigating through. This project has the potential to impact a large number of Americans who currently feel intimidated by cooking and most importantly, make cooking a fun activity.

TABLE OF CONTENTS

PART I

BACKGROUND AND APPROACH

INTRODUCTION

Ultimate Recipe - Cooking and Grocery Shopping Helper using ML

This project is for developing a web-based application called Ultimate Recipe which will enhance cooking experiences for the users. Regardless of age, gender, ethnicity, food is something we all enjoy. It is also a source of energy and a source of nutrition for our bodies. In today's world, luckily food is accessible at places like grocery stores and restaurants but eating out can bring unique challenges. It can be expensive and unhealthy. Ultimate Recipe is an application that is perfect for users that are new to cooking or users who simply do not have enough time in a day to cook.

The application is accessible through the web, and it provides a simplistic graphical user interface which is easy to use and navigate. The typical use case is that users enter a keyword related to food or ingredients and the application returns results consisting of dish recommendations. After the user enters the keywords, they are processed and aggregated through an already trained machine learning model. The benefit of utilizing an external machine learning model is that the results and output will always be dynamic. The output is always expected to change and update as more data gets trained. We believe that this project has the potential to impact a large number of Americans who currently feel intimidated by cooking and most importantly, make cooking a fun activity.

Problem / Issue in Technology

Issue: Current recipe applications are too general causing choice overload.

The market for recipe applications is vast with hundreds of sites available. Upon trying a few, we noticed that many sites have a user interface that feels overwhelming to navigate with several links to choose from. In addition, the search functionality from major recipe sites may return hundreds of results but may have limited use for the client. Search results do not take into consideration budget, recipe difficulty, or healthy options. Users can run a search and scroll for several pages without finding a good recipe to keep. The main issue with recipe sites is that recipe suggestions are generic by trying to appeal to the broader audience. Doing so leaves users with too many options that they might not even need.

Solution: Optimize user experience, add interesting functionality, and improve filtering search.

The solution will be on creating an intuitive and easy to use web application where a user is able to check different options when doing a search. An improved search will yield better results according to what the user needs. To further improve application searches, options will target a user base that is looking for recipes that are tasty, not expensive, and healthy. One possible demographic our project can cater to are underprivileged families and college students.

Other improvements can be made with the use of machine learning. A feature can be added to take in pictures of ingredients that once uploaded will automatically parse it into keywords for search purposes. The same implementation could also function as a dish recognition system.

Evidence of Need

Most recipe applications are not a one stop solution for users looking to cook at home. As mentioned earlier, most recipe sites yield search results that are general and do not address the specific recipe criteria for a user. Home cooks might need to find recipes from various apps, blogs, and general web searches to find all the recipes they need. With Ultimate Recipe, we plan to address those concerns by providing more accurate searches targeting a more specific demographic. Novel and useful features using machine learning will also provide improvements in user experience by making the web app more fun and exciting.

Project Goals and Objectives

Goals

The goals of this project are to:

- Provide users with a platform to learn about new recipes.
- Create a platform for community members to share and learn about recipes.
- Reduce indecisiveness when choosing a recipe.
- Increase awareness of different cultural foods.
- Learn how to create a production-like microservices infrastructure (explore cloud deployment options such as AWS and Google Cloud).
- Learn how to work as a team to build a minimal viable product that could potentially benefit many people.
- Learn how to build a web application with an intuitive and quality user experience and functionality.

Objectives

- Ability for users to create an account and display their favorite recipes.
- Use machine learning models to determine recipes based on list ingredients.
    - Allow users to snap a picture of their ingredients and/or receipt.
    - Allow users to input a text list of ingredients.
- Build a recipe recommendation engine to cater recipes to certain user groups.
- Provide profiles and "spaces" for users to share recipes.
- Allow users to "Remix" recipes to personalize their own recipes.
- Choose a robust tech stack to build out these objectives

Environmental Scan and Literature Review

According to our research the current young adults generation heavily relies on technology when it comes to cooking. For example, an article on BYU School of Communications website states that "Millennials are more likely to use technology when they cook [...]. Familiarity with technology is driving the way young adults cook because the answers to their questions, or an exciting recipe, are just a click away" (Matthews, 2018). Technology can help young adults to become more interested in cooking, as well as to help them to save some in their day-to-day activity.

Additionally, in the post-pandemic environment where more employers offer flexible schedules and colleges offer virtual classes, more people cook at home. Thus recipe websites such as Allrecipes and Tasty gain greater popularity (Li & McAuley, 2020).

As a part of the environmental scan, we explored relevant web applications and their features to identify key areas where Ultimate Recipe can differentiate.

https://tasty.co/ – is a cooking website that categorizes recipes by occasion, ingredients, diet or meal. The layout of the website mostly uses square cards and blogs, like articles. The UI can feel crowded at times, possibly due to content aggregation.

https://www.allrecipes.com/ – "All Recipes" website feature blog post - like recipes and categorization by meals, occasions, ingredients and cooking tips. "All Recipe" has unique content created by "food bloggers, registered dieticians, and even food scientists" (*About allrecipes* 2022). Users can search websites by ingredients, including ingredient exclusions and by the name of a dish. "All Recipes" website allows favoriting and user accounts.

https://www.supercook.com/ – SuperCook web application focuses on the search by ingredients that users have at home. It feels less like a blogging platform, and can be considered a progressive web app. User can use powerful search by meal type and ingredients simultaneously with exclusions and filters on the results. User can customize what is in their pantry or fridge to get better recipe suggestions. Overall, SuperCook functionality aligns the most with the Ultimate Recipe application goal. Ultimate Recipe app will focus on a less cluttered UI design while maintaining similar powerful functionality.

https://www.bigoven.com/ – the "BigOven" website differentiates by additional functionality such as planner, grocery shop list, and "Use Up Leftovers" section.

The Ultimate Recipe will focus on simplicity and extended Machine Learning based functionality as key differentiators among similar products. To reduce friction among such user groups as kids learning to cook, college students, and busy adults, Ultimate Recipe will offer

excellent page load times. By using a single page web application layout Ultimate Recipe will offer an intuitive and cohesive user interface. The Machine Learning extended version of the app will allow users to upload images of the ingredients or receipts to set up the initial search, which will allow users to save even more time.

PART II

Internal and External Stakeholders

There are primarily two major stakeholders for the application we are building for our final Computer Science Capstone at CSU Monterey Bay. The two stakeholders are identified based on the type of involvement they have to the application. One is us; the Orion Analytics team who will be directly developing this application. Two is the external party who is defining the project requirement. In this case, we believe that the external party indicates the faculties of the Computer Science program at CSU Monterey Bay who are specifying the rubric for the CST499 course. Though generally speaking, the external stakeholders for this application would be the customers who will be using the application.

Through the process of designing and developing this application, the internal stakeholders; Orion Analytics will gain a significant amount of knowledge and skill set. As Capstone is the last project before graduation, these collaborative experiences will become the fundamental base for the software engineering career we are about to embark on. We also believe that there are real-life use cases for the application we will develop, allowing us to potentially productize the application to the public users that are the alternate external stakeholders.

The examples of the external stakeholders' benefits are 1) successful demonstration of our knowledge which Orion Analytics team gained through the CSU Monterey Bay education. 2) proof that the academic curriculums for the CSU Monterey Bay CS Online Program adequately prepared Orion Analytics team for building an application with highly complex requirements. As CSU Monterey Bay being an academic institution, we believe that one of the most important goals is to educate and train their students to prepare for their careers. The successful delivery of

this project will certainly be the achievements for the external stakeholders as well as for the internal ones.

Lastly, this project will make significant differences for both the internal and external stakeholders. As described previously, experiences we gain through the Capstone project is incalculable. For the external stakeholders, a successful demonstration of the students' skills increases the value of education which CSU Monterey Bay offers. There are also potential benefits for the alternate stakeholders if we productize the application. An application like Ultimate Recipe can truly change the cooking experience for many users around the world. This application has the great potential to promote healthier lifestyles by changing the way people think about cooking. Overall, this project can affect the various types of stakeholders in positive ways.

Ethical Considerations

User privacy and anonymity are major ethical concerns in web applications and will be considered for this capstone project. During production of the project and to address these concerns, an authentication system will be implemented for users to safely login and store their data. In addition, a privacy policy may be included on the footer section of the website to provide users with clear information as to how their data will be used. Another ethical consideration concerns biases from the data being used (whether methods or sampling) in machine learning as well as the API used for recipe search. To address bias issues, many models and APIs will be reviewed and analyzed. The final selection will be chosen as the one that best addresses the needs of the user base for this project.

Since this project is aimed at people looking to cook healthy meals at home it should benefit most users, especially underprivileged groups as cooking at home is more economical than eating at restaurants or buying ready made meals. For that reason, this project should not negatively impact underprivileged families. Access to the web app, however, could be a hurdle for this group as it requires an internet connection and a computing device. Nevertheless, internet access can be found for free at a local library or school. Disadvantaged users can use computers and printers at these facilities without paying out of pocket.

The project will take into consideration people with disabilities. For those visually impaired, the web app will have a pleasant and evenly spaced text area where text size can be increased or decreased. Proper contrast in color and texture will be provided as well. For blind users, images will be properly labeled for text to speech with a text layout that is both organized and sectioned.

Legal Considerations

Similar to the ethical considerations, users' privacy is the key concern when it comes to the legal considerations. Virtually every website and web application will have a legal and privacy notice. The reason for that is the rapidly changing laws, and growing public concern about how their data is used by online businesses. According to Brookings.edu "More and more data about each of us is being generated faster and faster from more and more devices, and we can't keep up" (Kerry, 2022). Brookings.edu state that users need to understand how their data is used, and that ultimately new legislation is needed "to ensure that individuals can trust that data about them will be used, stored, and shared in ways that are consistent with their interests and the circumstances in which it was collected" (Kerry, 2022). It will be our responsibility to investigate

applicable laws and guidelines and show correct legal notices according to the user's geolocation. For example, some users could be located in the EU, where more strict GRPC laws can be applicable.

Our application will not collect such personally identifiable information as social security numbers, health information, or account numbers. However, according to the guide published by the U.S Department of Labor, information that can be used "to identify specific individuals in conjunction with other data elements, i.e., indirect identification" (*Guidance on the protection of personal identifiable information*) is also considered personal information. That means that if a user can be identified with a collection of data pieces that we would store on the website, we should treat this data especially carefully. Primary data points that need attention in the user authentication, passwords and email protection. As mentioned above, we should be able to safely login to our application. We will be using 3rd party authentication providers such as Google OAuth, or will implement a safe encrypted storage of emails and passwords.

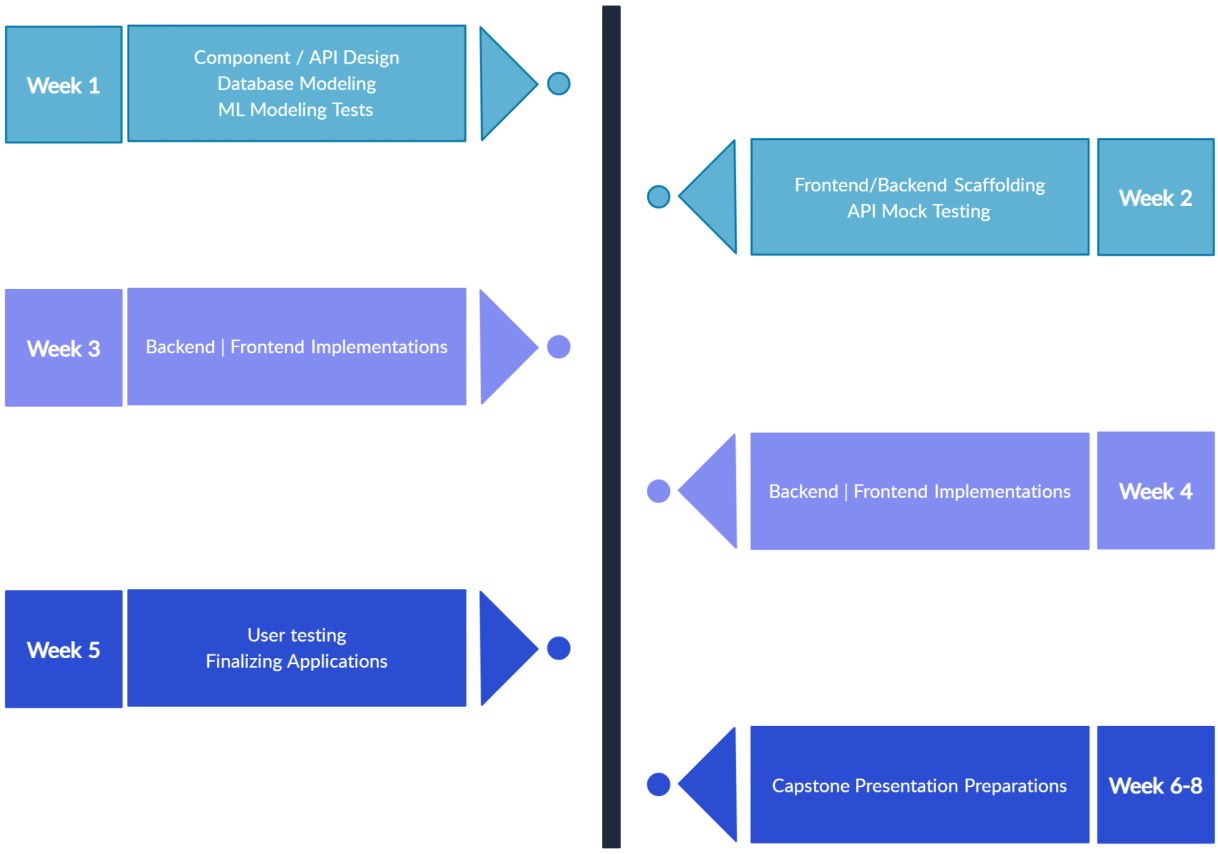
Secondary data points that could potentially identify the user or be sensitive data are user's search queries and photographs or groceries, and ingredients that users might upload to a website. For the search queries, terms and conditions should be clearly defined on the website, especially if the search queries are stored, cached or reused between the users. For the image upload, ideally, those images should be only used at the runtime to retrieve the list of keywords to be passed to a search engine. No user images should be stored in the Ultimate Recipe's database, which should make dealing with this type of data easier.

Overall, the application should be complied with the privacy laws in the United States, and other countries, where users might use the app. At the same time, effort to protect user's data, any PII and any additional sensitive information should be applied across the application.

PART III

Project Scope

Timeline



We divided the scope of our project into parts that fit neatly into five weeks. The first two weeks will consist of designing components, modeling our data, testing out our ML models, and building out the scaffolding of our applications. We will then work on implementing our designs and models in the third and fourth weeks. Our plan is to host a test version of our front and back ends by the end of the fourth week. Afterwards, we will dedicate the fifth week to finalize our

application and run user tests to get feedback on our application. Finally, the remainder of our time will be spent preparing for the capstone project.

Milestones

The major steps of the process toward completing the proposed project are outlined in Figure 1. Since Orion Analytics is not working with an external stakeholder for the CST 499 project, we analyze and determine the expected user stories as a part of our first step. We then assess the feasibility to support those user stories as we need to rely on the already trained machine learning models available on Hugging Face website. We expect several iterations of design and development as well as testing before the completion of our prototype application.

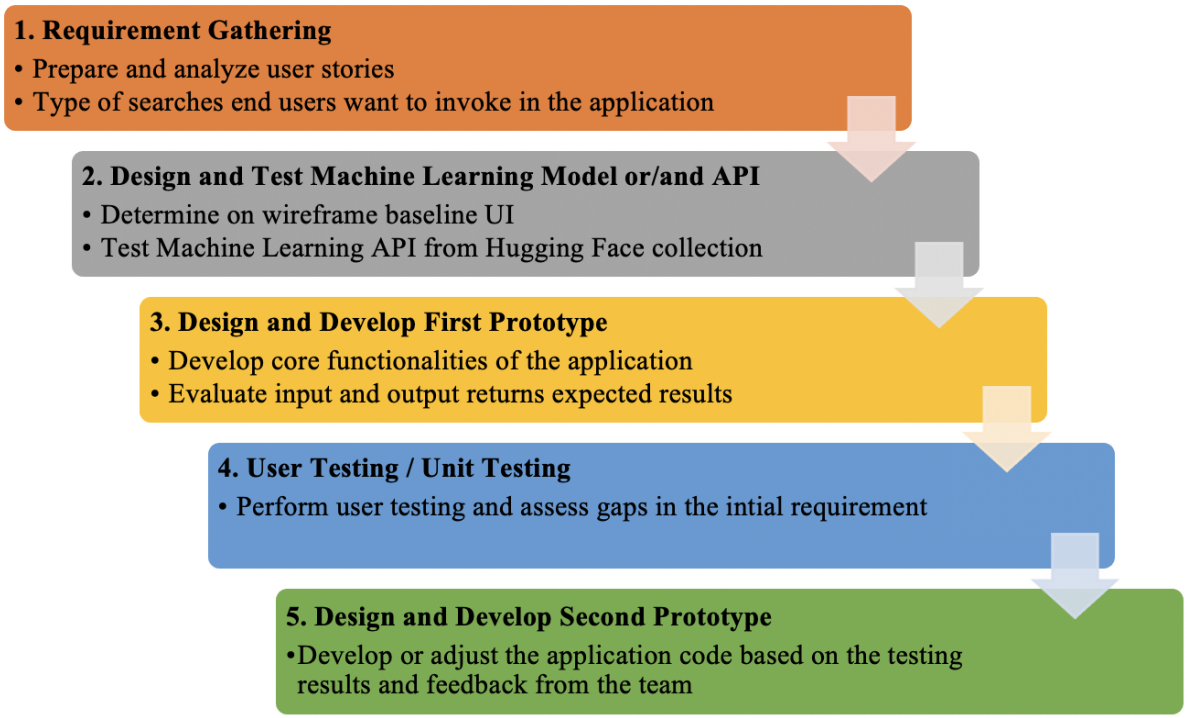


Figure 1. Project Milestones

Resources

We will be utilizing free versions of software to host our applications and codebases. We will utilize Github to host our frontend and backend code repositories. We will also use Heroku to host both our frontend and backends.

Risks and Dependencies

The machine learning implementation to recognize food ingredients from pictures is one risk that may affect the completion of this project. To mitigate the risk, more research will be needed to develop the model. If the process proves to be very time intensive a suitable API will be used instead. In terms of dependencies, the steps shown in the timeline will outline the steps required before moving to the next stage of the project. For example, tasks for week 1 and part of week 2 regarding the architectural design will first need to be first designed and put in place in order to start the backend and frontend.

Final Deliverables

The main goal of this project is to deliver a fully end-2-end functional version of the application. even though some non-critical components may be descoped. Users should be able to follow a minimal user story scenario to successfully login to the Ultimate Recipe application, perform recipe search and get a comprehensive list of results.

The second concern to describe final deliveries is to enable continuous work on this project, future team members onboarding, and if needed a hand off to a different supporting team. The project should have a code well organized and documentation supporting its components.

As such, we set the following final deliveries.

1. User Facing Deliveries
    a. A website with a registered domain name that can be accessible by the users over the Internet.
2. Development Team Deliveries
    a. System design and database entity relationships diagrams for the service based architecture.
    b. GitHub repositories or a mono repository containing applications components code – main backend server, ML model backend server, frontend server. Each repository must have a README.md document with the dependencies and setup instructions.
    c. User group testing sign off documentation. This should be a document describing the results of the manual testing of the user scenarios. The document should call out any blocking issues – such as inability for the user to access the website, and improvements and nice-to-haves.
    d. Final presentation video with the slides accompanying the project with additional background documentation, research and discoveries.

Usability Testing and Evaluation

Usability testing will be conducted during week 5 to a group of 5 participants who are interested in cooking at home. The focus group will be chosen among friends and family members from this project's team. On week 5, participants will be sent a link to test the website and at the end of testing will complete a Google form survey to evaluate the application.

Depending on time constraints another smaller subset of participants can be asked to help in providing feedback during the development stage during a screen sharing session on Zoom or Google Meet.

Approach and Methodology

Due to the rapid development timeline expected for the project, the adoption of agile methodologies outlined in Figure 2 will likely benefit Orion Analytics. In each milestone task, below subtasks are expected to be fulfilled.

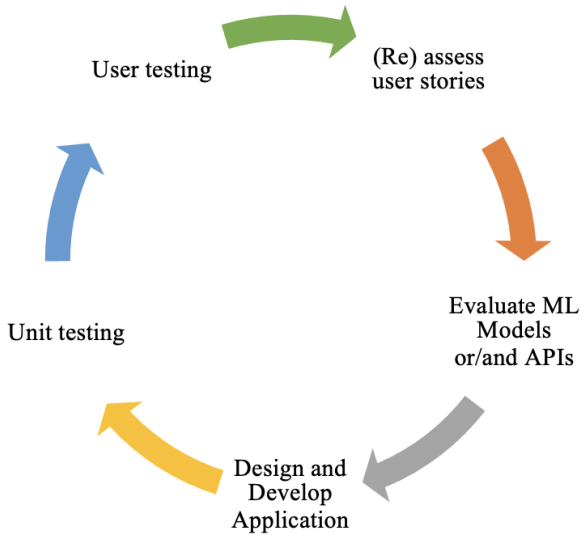


Figure 2. Agile Approach to Ultimate Recipe Development.

1. Requirement Gathering
    a. Analyze and solidify user stories around questions such as:
        i. What type of input does the user expect to apply (image, text)
        ii. How is the user invoking / executing the search?
        iii. What type of output does the user expect to receive?

2. Design and Test Machine Learning Model andAPIs
    a. Determine most suited framework for development
        i. Programming language
        ii. Website hosting server
        iii. GitHub repositories
    b. On Hugging Face website, search for machine learning models most adequately support user stories defined in the first step
        i. Evaluate the type of input and output expected for each machine learning model
        ii. Evaluate any dependencies associated with the machine learning model
    c. Design application components and services APIs
        i. Divise major application components such as main backend service, backend service serving ML model, frontend service, and any peripheral services such as authentication
        ii. Create database design and describe each service surface APIs
3. Design and Develop First Prototype
    a. Develop core functionalities of the application such as:
        i. Main landing page
        ii. Navigation on the application
        iii. Options for users to search (i.e. text box, dropdown list)
        iv. Search button
        v. Execution of machine learning models or API

4. User Testing / Unit Testing
    a. Unit Testing different functionalities on server side
        i. Execution of machine learning models
        ii. API
        iii. Verify input and output produce expected results
    b. User Testing
        i. Evaluate the application as if Orion Analytics is the end user to assess any bugs and enhancements
        ii. Evaluate areas that need adjustments in code or user story
5. Design and Develop Second Prototype
    a. Make adjustments to the code based on the discoveries from previous steps
    b. Make adjustments to the code to fix any bugs identified during the testing

References

Allrecipes. (2022, September 29). *About allrecipes*. Retrieved October 4, 2022, from https://www.allrecipes.com/about-us-6648102#toc-the-allrecipes-allstars

Guidance on the protection of personal identifiable information. United States Department of Labor. (n.d.). Retrieved October 11, 2022, from https://www.dol.gov/general/ppii#:~:text=Personal%20Identifiable%20Information%20(PII)%20is,either%20direct%20or%20indirect%20means.

Kerry, C. F. (2022, March 9). Protecting privacy in an AI-Driven World. *Brookings*. Retrieved October 11, 2022, from https://www.brookings.edu/research/protecting-privacy-in-an-ai-driven-world/

Li, S., & McAuley, J. (2020, July 30). Recipes for Success: Data Science in the home kitchen · issue 2.3, Summer 2020. *Harvard Data Science Review*. Retrieved October 4, 2022, from https://hdsr.mitpress.mit.edu/pub/wjrl1qsq/release/3

Matthews, S. (2018, July 5). BYU students find more taste with technology. *The Daily Universe.* Retrieved October 4, 2022, from https://universe.byu.edu/2018/07/03/byu-students-find-more-taste-with-technology1/