Initial Project Proposal

Team No. 24

Members:
Jacob Davies - 2920296
Jianpeng Li - 2953297
Evan Powell - 2892588
Jarrett Zeliff- 2957780
Isaac Kuhlmann- 2886770

Project Name: SmartPharma

Project Synopsis:
Web interface to service better and more direct communication between Hospitals and pharmaceutical distributors/representatives

Project Description:
The underlying goal of this project is two-fold; to better connect pharmaceutical distributors to the Hospitals they service, and to provide both the hospitals and distributors with data from prediction models that may offer considerations as to which drugs to buy and when. Currently, pharmaceutical representatives make direct calls/visits to Hospitals in their district, which is time consuming, difficult to manage, and all around impractical. At this time, there is little on the market to service this need. As stated above, the system will also double as prediction software with the ability to predict when outbreaks will occur, what kind of outbreaks they will be, and what kind of medicine will most likely be in high demand when the outbreak occurs. This prediction will be based on national data for the time being, and will notify Hospitals if there are any significant predictions. The prediction will allow Hospitals to reach out to representatives well in advance of potential outbreaks, and order the recommended drugs as they see fit. The system will also serve as a trending service, keeping track of what time of the year specific drugs are prescribed and the volume of those drugs prescribed at that time. This data will be compiled into an easy to read graph for each drug in inventory, and can be viewed by both the pharmaceutical representative and the Hospital.

Preliminary Project Design:
Overview

As stated in the above description, this project will service both **interface** needs between pharmaceutical distributors / hospitals, and will offer **predictions** as to what the future viral climate may look like, to allow these parties to make decisions accordingly. Representatives of the distributors and hospitals will be able to directly interact over a centralized web portal, as well as view all prediction data.

In the industry currently, representatives visit their hospitals either in person, or over the phone. Representatives usually have no idea about current medication inventory without first inquiring, and hospitals are put in the position to make their best guess possible based on what and how much they **think** they will need. Obviously physical inventory manifests do exist, and medical professionals are very much qualified in making decisions regarding medication, but it is painfully obvious just how inefficient these interactions are.

Aside from its predictive functionality, SmartPharma is just a better medium for communication. It is a centralized location for interaction between hospitals / representatives, allows for real-time logging of inventory levels, uses those inventory levels to show prescription trends over the course of a year, and will provide predictions for possible future upticks in sickness.

The project should be thought of in terms of **three** components that will later be discussed in more detail.

- Front-End
- Back-End
- Prediction Models

Front-End

The Front-End will allow the user to perform the following actions:

- Update/view Inventory
- View Trend Data
- View Prediction Data
- Message either a hospital or representative

These actions can be performed by either representatives or hospitals. Figure 1 shows the preliminary design for the layout of the landing page.

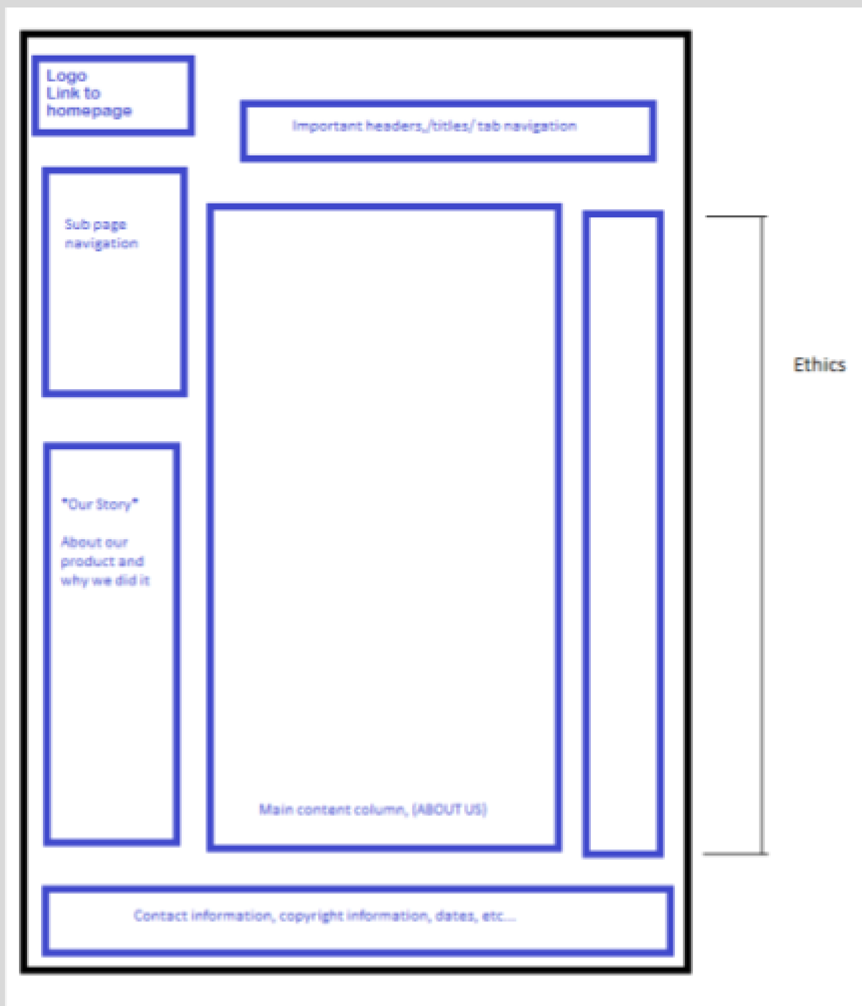

Figure 1

The inventory system should work relatively the same for both reps and hospitals with **slight** variations between the two. As for their commonality, both will be able to toggle the data shown via filters; view-full-inventory, view-low-inventory-only, view-most-prescribed, and view-least-prescribed. These filters can be applied on

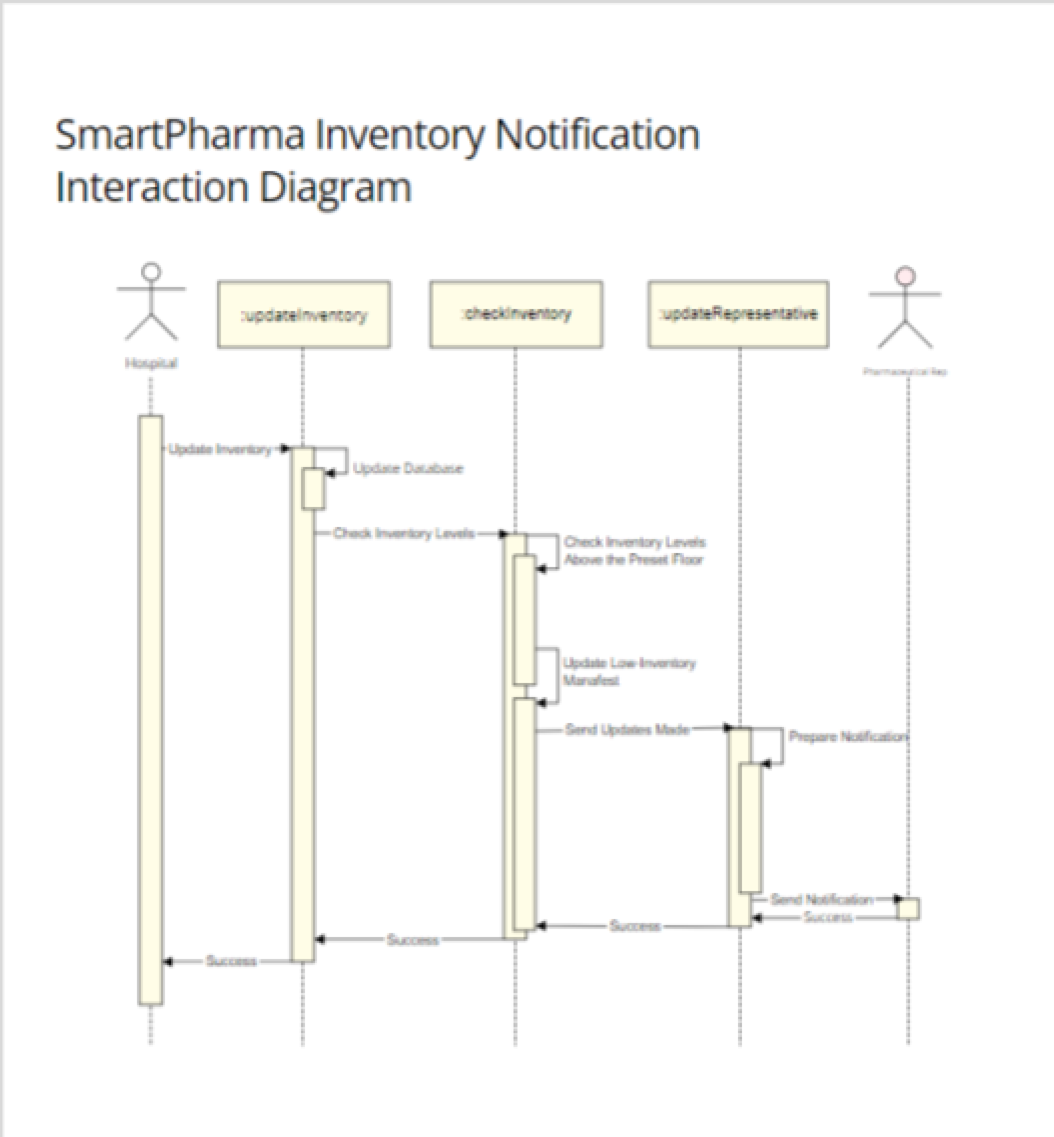

Figure 2

top of each other, for further control over shown data. Drugs that are below the **inventory floor**, which is preset by the hospital and can always be changed, will be shown in red to distinguish them from other drugs. As for the differences between the user views, hospitals will obviously only be able to see their own inventory, while representatives will be able to see every hospital that they represent. Figure 2[1]

[1] https://drive.google.com/file/d/1FIJ6KlngWU9lgfvGkIi_VzmySFtnn3Wv/view?usp=sharing

shows the interaction between the hospital and representative using the inventory system.

Trend Data will be available to both reps and hospitals. User views will be split like the inventory system where hospitals can view themselves, and reps can view all of

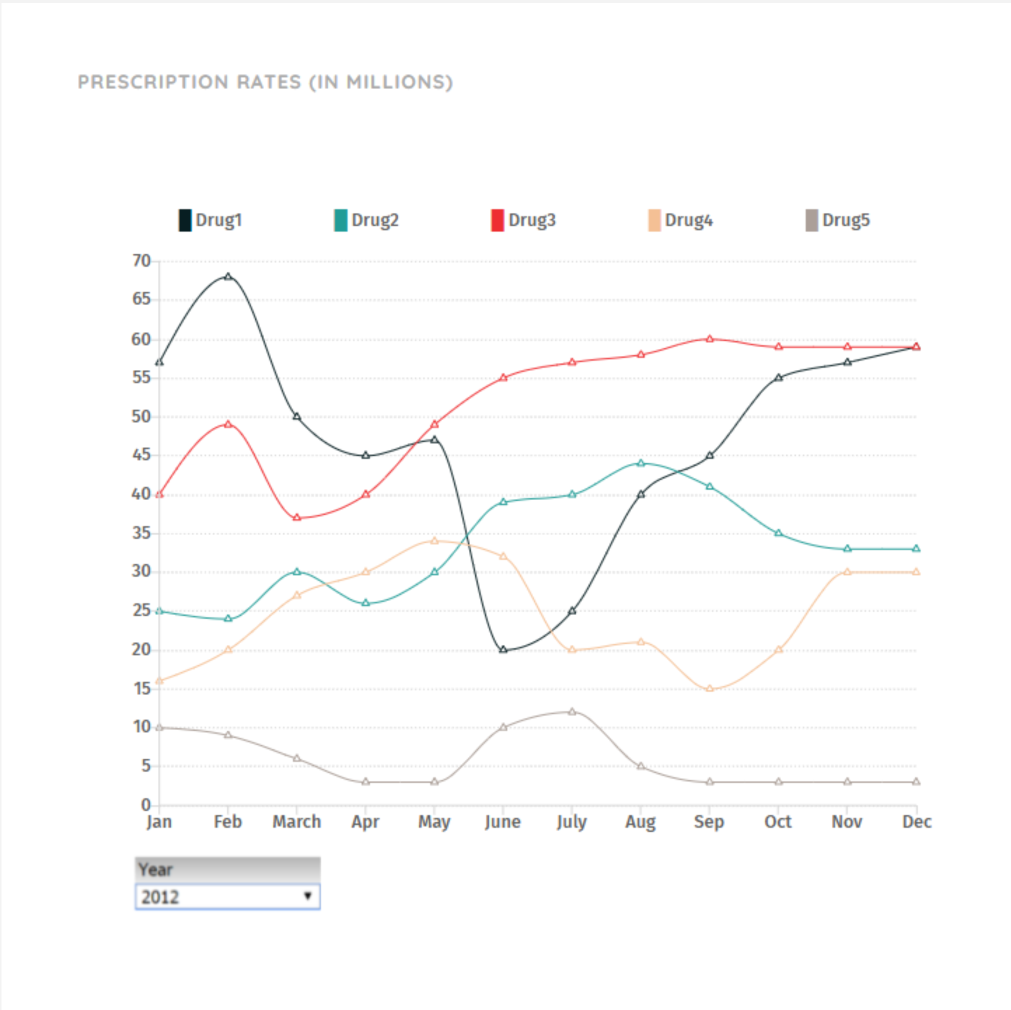

Figure 3

their hospitals. The trends that will be logged are prescription rates per month for 12 months and prescription rates per year. There are certainly many more trends that could be logged in the future, however we would want to wait for user feedback so we don't go logging useless trends. The graph that will be used to view the trends can be filtered to show as many or few drugs as the user wishes. This is possible because the data is logged for each individual drug. The user would also be able to filter with all of the filters outlined in the inventory system. Figure 3[2] illustrates how logged data may be displayed in order to visualize trends. The chart is showing the rates of prescriptions for different drugs over the Course of a year. Below the chart is a modifier where the user would be able to select which year to view. Other modifiers would include toggles to choose which drugs are displayed and a toggle to switch between viewing a single year, to viewing multiple years (zoom in - zoom out).

Prediction data can be viewed by both types of users, with the same rep / hospital dynamic. The prediction models will make predictions based on *infection rate per time of year* datasets. The only interaction the users will be able to perform with this data is to view it. At this time, there is no plan to implement any interaction with the prediction models.

The messaging system will just be a basic notification system. Keeping this system fairly topical and not going into message encryption etc, the system will allow for either a rep or hospital to initiate a DM, and will notify the recipient when the message has been sent. The recipient will then be able to go to their message box, view the message, and send a reply. This is a fairly simple system and there is nothing necessarily special about it. It is worth noting that reps can message multiple hospitals while hospitals can only message their rep (as above, so below). It is not apparent at this time that hospitals will benefit from the ability to message other hospitals, but maybe a general message board may be beneficial.

Back-End

The back-end will employ Django, a free and open source web framework. This framework will allow us to worry more about the specifics of developing the project rather than working on boiler-plate web app code. Django is python based, which is one of the main reasons for choosing it. There are many reasons that Django is a good framework, but it being python based makes it a very attractive choice, as it will work well with the prediction, which is also written in python.

---

[2] Chart is a concept, and does not necessarily reflect the final design.
https://drive.google.com/file/d/1wRqz4tg1SXipl0p-hY-8Gsq6rvgQmbrf/view?usp=sharing

Considering both the prediction model and web framework are both python based, we made the obvious decision to write all of the back-end services in python as well.

The database we have decided to use is postgres as it is free and is very user friendly. It is also very commonly used, so there is plenty of documentation and user-feedback which will be helpful for us to get things up and running.

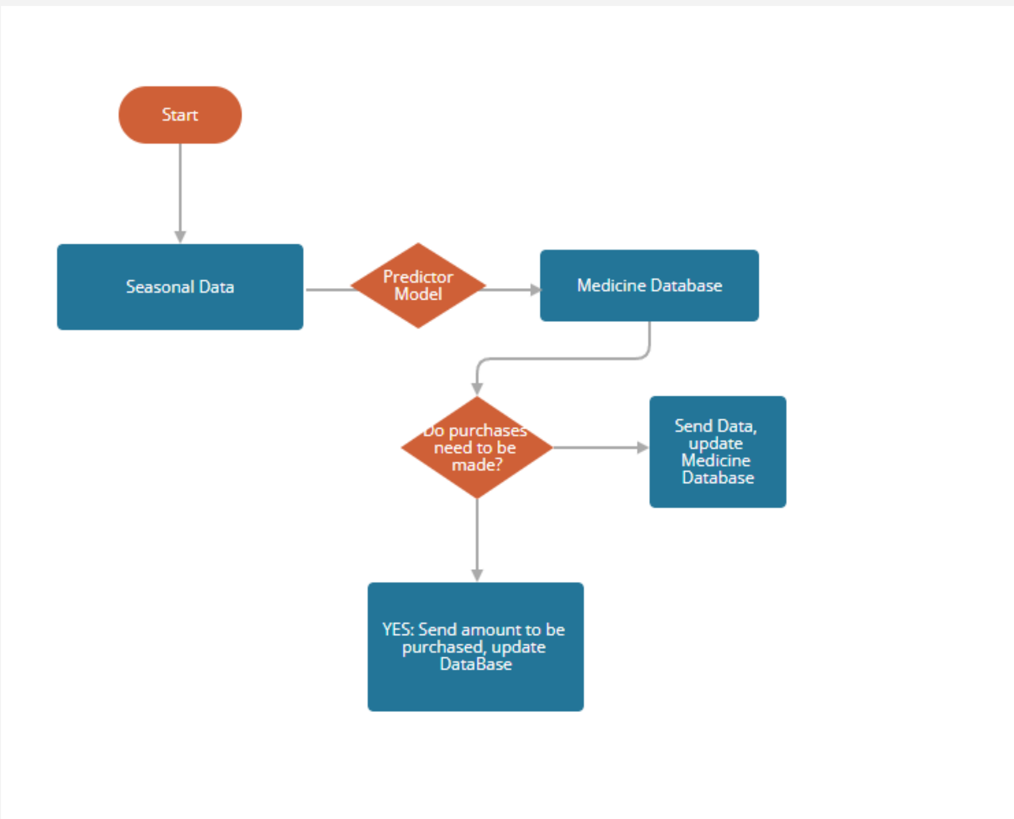

Figure 4

Prediction Model

The predictor model will serve to <describe>. It will take, as input, a set of features created to define the yearly patterns of common illnesses. Using these features, the predictor will come to conclusions on what diseases may or may not be prevalent in the coming future. These results will then be used to make informed decisions regarding the distribution of medicinal supplies.

<u>Features</u>

The set of features used as input to the predictor model will be formed from disease and outbreak data accrued in preparation for this project. The very nature of this sort of data in the medical world lends itself to a large variety of entries pertaining to what may cause a disease to be prevalent at a given time. This can grow to include climate, population age, average net worth, and many more. This amount of variety all being used for the feature set would lead to an ineffective model with contradicting features.

To combat this, the feature set will undergo a process of annealing to determine which entries will prove to be most impactful. It is this annealed set that will be used in the final implementation of the predictor model. This undertaking will boost the quality of the input to the model, as well as sharpen the decisiveness of the output.

<u>Model</u>

To obtain an accurate and effective prediction model, an ensemble of different machine learning models will be formulated. The exact combination will be continually tested and perfected in order to determine a viable solution. That being said, this ensemble will likely include the neural network learning model. This will be thoroughly tested as well in order to conclude its practicality.

This model will be fed features as described above. A separate feature set will be created for each disease tracked by this system, but each set will contain the same feature labels. This allows for ease of scalability. When new diseases are tracked, the features that describe them will follow the template defined by the feature labels.

The output of the prediction model will correspond to a level of predicted severeness of a given disease over a period of time. A prediction for a higher level of severeness could be used to make the decision to order medicinal supplies that treat said disease.

design constraints

We have chosen to program the back-end of this project in python, as it will be consistent with DJango (our web framework), and with our prediction models. We have put this constraint on ourselves to ensure there is little to no conflict between back-end components.

We have also decided to limit the data used in the prediction model to be National Data. That is, there will be region-specific predictions at this time. We have chosen to constrain ourselves here simply because of the sheer scale of data we would have to accumulate to make these predictions with accuracy. This project is not in a position to handle this at this time, but we will not discount revisiting this functionality in the future.

Project Milestones:

First Semester

- Identify data sets to use for training of prediction models
- Identify a registry of medications on hand at Hospitals
- Design document for database
- Design document for website hosting framework (front-end)
- Design document for website hosting framework (back-end)
- Design document for prediction model
- Gather data sets for use in training prediction models

Second Semester

- Filter data sets for use in training prediction models

- Create a database, and fill it with different medications used at Hospitals
- Build prediction model and train with collected data sets
- Set up back-end (basic API)
- Set up front-end (cosmetics)
- Connect front-end and back-end

Work Distribution:

Jake:

- Web framework / back-end in general
- Inventory system
- Database

Jianpeng:

- Front End
- Back End
- Database

Evan:

- Predictor Model Design
- Data Gathering and Engineering
- Database

Jarrett:

- Front-end web page design
- Predictor design
- Database

Isaac:

- Predictor Model Datasets and Design
- Data Gathering and Engineering
- Back-end Django

Project Budget:
We will be developing this project using free, open-source software, and will look to host the site after development. We have no plans for a budget at this time.

Ethical Issues:
Having a predictor model that deals with potential life or death raises a handful of potential problems (i.e. intentionally not prepping for a specific time of the year based on our model predicting low infection rates). It should be abundantly clear that these predictions are **recommendations only**, and decisions about medication should not be made based solely on our prediction model. We cannot claim enough accuracy with the model to justify the **encouragement** of one drug over the other. We will address this by putting a disclaimer on our site, as well as putting it by the data itself.

If by some means in the future if we are getting personal data on a population (of a city, for example), namely any underlying health conditions, medical records, or GPS location (for contact tracing), we would exercise the utmost caution in guarding this information as well as refusing to sell or share this data to anyone. We will comply with all HIPAA codes and ensure that there are no discrepancies. As we do not intend to store any of this information at this time (it will be strictly pharmaceutical medication, which is public record), we will ensure that there are checks in place if this does become an issue in the future.

We will remain agnostic towards any particular drug production company, leaving any and all purchasing decisions to the individual user (hospitals and pharmacies) as to the specific brands of any purchased drugs. SmartPharma does not claim to know any exact number of drugs to purchase and any recommendations made to that point should be exactly that, recommendations. As the accuracy of the prediction model improves, we might be able to make more and more robust recommendations, but for the time being, the model should be used as recommendation only. All of the final decisions made should be solely the users.

Intellectual Property Issues:
For the back-end: all data used to train the model will be free for use, publicly-available information (mostly provided by the World Health Organization as of now). If in the future we are given any data to train our model with/use with our model we will claim no ownership as well as give proper citation on our site to any contributions. Any such data provided to us will remain secure, we will not sell or share any given data to third parties. If it is needed in the future, we may employ protection services to keep all of our data safe.

We will honor the ACM code of ethics to the best of our ability, prioritizing user privacy and security. The code will be at the forefront of all decisions we make, from the open-source software we choose to use to the code we write ourselves. Any external libraries or datasets that are not built in-house will be open source and free for our use. Any and all data that we receive shall be used as it was received; it will not be altered in any way in order to falsely represent the accuracy of our model. If necessary in the future, our model will be kept as a trade secret, and remain protected under copyright law as such.

Change Log:

We have kept with the primary idea from our initial project description, however we have new ideas we would like to implement if time permits:

1.) Outbreak predictors to predict when an outbreak will occur, potentially per state/city
2.) Hospital bed optimizer, after seeing the issue with hospitals running out of space for patients we want to try to optimize the amount of people who can be cared for at a time.

Gantt Chart:

				Design Semester																																		
				Rough Drafts					1st Revised Drafts					2nd Revised Drafts					3rd Revised Drafts					Final Design Docs					Video Rough Draft					Final Video				
TASK TITLE	TASK OWNER	DUE DATE	PCT OF TASK COMPLETE	M	T	W	R	F	M	T	W	R	F	M	T	W	R	F	M	T	W	R	F	M	T	W	R	F	M	T	W	R	F	M	T	W	R	F
Database Design Document																																						
Row/Column Diagram	Jake	11/12/21	0%																																			
Data Flow Diagram	Jake	11/12/21	0%																																			
Front End Design Document																																						
Sketch/model of landing page	Jarrett	11/12/21	0%																																			
	Brian	11/12/21	0%																																			
decide on front end framework	Brian	11/12/21	0%																																			
Cache files	Brian	11/12/21	0%																																			
Maybe build a prototype?	Brian	11/12/21	0%																																			
Back End Design Document																																						
Framework	Jake	11/12/21	0%																																			
Object Interaction Diagram	Jake	11/12/21	0%																																			
Data Flow Diagram	Jake	11/12/21	0%																																			
Prediction Model Design Document																																						
Research/Test Effective Models	Isaac	11/12/21	0%																																			
	Evan	11/12/21	0%																																			
Establish flow of data / data mining	Evan	11/12/21	0%																																			
	Jarrett	11/12/21	0%																																			
	Isaac	11/12/21	0%																																			
Establish Training and Test Set Sources	Isaac, Evan	11/12/21	0%																																			
End To End Model Design	Isaac, Evan	11/12/21	0%																																			
Video Presentation																																						
Final Video	all	11/13/21	0%																																			