Final Report
Technology and Operations 399: Independent Study
Student: Philip Ruffini
Professor: Sanjeev Kumar

Executive Summary

This final report details the learning outcomes and results of Philip Ruffini's independent study that was conducted through Winter Semester of 2018 under Sanjeev Kumar. Philip spent the semester learning how to develop analytical tools for Amazon Sellers to analyze raw data. The main tool that Philip developed is a dynamic heat map that allows Amazon Sellers to visually display their sales for any product over a specific range of dates. This tool will allow Amazon Sellers to reconfigure their marketing strategies to focus on key geographic areas where they are either trying to increase sales or attempting to grow a new market. The rest of this final report will outline the steps Philip took to complete the project.

Goal of Action Based Learning Project

The main goal of this action based learning project is for the student, Philip Ruffini, to develop data analysis tools for not only his personal Amazon Business, but also the businesses of other Amazon Sellers. These tools will allow Philip to gain key market insights from the raw data that Amazon supplies. This data will primarily focus on customer locations, product seasonality, and product profitability. Thus this project will give Philip a significant advantage for both his business and his future career as an entrepreneur as he will gain first-hand experience building and using his own data analysis tools to draw insights. The initial completed proposal can be found in Appendix A.

Initial Market Research

Before beginning to develop tools for Amazon Sellers the current market was analyzed to see where needs were missing. After researching and talking to other Amazon Sellers, key takeaways from research and interviews were that despite the vast offering of Amazon Seller Software applications that many people still want better solutions. People are looking for better solutions pertaining mostly to accounting, customer relations, and product research. In terms of account, a lot of sellers complained about issues including bugs and poor UI design for their current software solutions. With Customer Relations sellers want to know more about their customer. They do not know anything about the people they are selling too and want to know more. They also want help getting more reviews. Lastly, people still wanted as many resources as possible for product research. This pertains to sales tracking, momentum, and arbitrage opportunities. Key assumptions that were validated were that sellers want a one stop shop platform that offers all of the capabilities and sellers want to know more about their customers.

Due to Amazon's API, Amazon sellers have access to sales data, but cannot easily analyze or aggregate it to read or observe trends. The proposed solution to this issue is a heat map of customer locations. Their currently no direct competitors as this current solution does not exist. However, competitors would consist of Amazon, HelloProfit, and Seller Labs. Amazon is a competitor because they could easily implement this solution and day or shut it down without hesitation. However, Amazon has chosen to allow their seller community to develop their own tools. HelloProfit is a competitor because a heat map is listed on their features as "coming soon" therefore they will eventually bring a similar product to market. Seller Labs is an indirect competitor because they are a well-known firm that is focused on developing tools for amazon

sellers. They currently offer accounting, inventory tracking, and email feedback software. A business model canvas can be found in Appendix B.

Purpose of Heat Map

The heat map is going to allow an Amazon Sellers to see where their customers are located as a whole and for each individual product. The benefit of the application is that it gives Amazon Sellers more information on their customers. One such practical use for this knowledge is that it will allow Amazon Sellers to better target Facebook and Social Media ads by targeting specific geographic areas. Therefore, allowing Amazon Sellers to make concise conclusions and draw analytical insight from the raw data

Progress of Project

This project can be broken down into four main steps: research, collecting data from Amazon's API, cleaning the data, and displaying the data. The first part of the process of research is documented above. Multiple Amazon Sellers were interviewed to identify what tools were missing and the one that continued to appear was visualizing customer location.

The next step in the project was to collect the data that was needed to create the heat map. This involved working with Amazon's API that uses Java Programming. The first month of the project was learning how to pull data from Amazon's API, which was quite difficult. The data that was needed from Amazon's API was order number, date of order, product name, product's ASIN (unique identifier), and the customer's address. Amazon's API will only output about ~50 orders per minute, so the program had to be modified to continually make requests to Amazon's API. It takes about 30 minutes per 1000 orders. The data from the API is then outputted to a txt file that stores all of the information.

The next step of the project was organize the data to generate the heat map. The data that came out of Amazon's API was not organized by product, but by the date of when the sale occurred. Additionally, the heat map plots the sales data by latitude and longitude and not by the actual address. Therefore, the first step was to take the customer addresses and geocode them to latitude and longitude. Instead of doing this each time a heat map is created, the python script was written to run every address through Google Maps geocoder one time and cache all the data to add the latitude and longitude to the end of existing entries that came out of Amazon's Api. A second python script is then used to collect the latitude and longitude coordinates based on a product's ASIN, start date, and end date. These are the coordinates that will be plotted on the heat map. Those coordinates are then put into their own txt file.

The final step is to generate the heat map using those coordinates. This is done by using Google Maps' API that allows users to heat map at set of coordinates. The first step was to create a PHP script that runs a loop through the txt file full of latitude and longitude coordinates. The PHP script collects the coordinates, which are then uploaded to Google Maps' API to create the heat map. This process was then made dynamic. With all of the data cached on the backend the heat map was then equipped with three inputs. A drop down menu with all of the product ASINs and two calendars to select the start and end date of the data that would be visualized. These three inputs are then used to create a new txt file that will populate the heat map. Therefore, allowing users to check their sales over all of their products for various sets of dates.

Learning Outcomes

The main purpose of this project was for the student, Philip Ruffini, to develop analytical tools to analyze and gain key market insights from raw data. The heat map has been provided to three Amazon Sellers and run on their accounts to see how it works for others as well. Key insights that were drawn, for example, was finding that most of Philip's sales for 8 pack sign holders came from three major cities: Philadelphia, New York City, and San Francisco. It was interesting to see that most of those cities made up the sales. Philip will now be able to optimize and retarget his Facebook Ads for areas that are already driving a large amount of sales.

In addition to developing this analytical tool there were two key learning outcomes in terms of product management skills and software development skills. The first large hurdle that had to be overcome was pulling data from Amazon's API. The language required to do this in was Java. Philip had never worked with Java before, so this was quite difficult. After spending approximately two weeks attempting to pull data from Amazon's API, Philip ended up contracting out the work to a software development firm in Sri Lanka. Philip managed two developers in Sri Lanka to create the Java script file that would pull from Amazon's API. Philip ran into multiple issues working with them and ended up having to edit the code himself as well. From this Philip learned what it would take to be a product manager and some best practices to follow. The first learning insight from that experience being too clearly articulate what you need in the program as concise and clear as possible.

Finally, Philip ended up working with three languages he had never worked with before: PHP, Java, and JavaScript. In addition, to these three languages Philip also learned more about programming with Python, such as using dictionaries to organize data. While working with these languages Philip learned how to connect client and server side languages. PHP and Python had to

be built into the HTML code of the map to allow the server side functions to run while HTML displayed the map.

What Could Have Been Done better

Looking back on the project certain steps code have been executed more effectively. The first mistake was believing that Amazon's API would be easy to work with and work with Python. Not enough research was conducted on how the Amazon's API worked, which really slowed the early stage development of the project. The next mistake that was made was assuming that the map would generate by just displaying the HTML code on a browser. Philip was not aware that he would have to run a virtual server that linked the backend frameworks with the frontend. Philip then began to use MAMP to display the heat map, which allows both front and backend languages to run simultaneously. The final mistake of the project was utilizing Google Maps' to generate the heat map. Philip should have looked into cheaper alternatives such as R Shiny to create the heat map. Google Maps' only works with latitude and longitude coordinates and it becomes expensive to geocode addresses into latitude and longitude using Google's API. This could be a serious issue for Amazon Seller's that have hundreds of thousands of orders to convert.

Functionality of Project

This section will break down the technical pieces of the project. All files discussed below and be found in Appendix C, which is the attached folder.

Collecting Data from Amazon's API

The first step when running the program is to pull from Amazon's API. This is done by running the GetZipCodeAmazon.java file by running a javac command to compile the file and a java command to run it. The instructions to run it are in the README.txt file. The Java file also depends on six jar files that are in the heatmap folder and a class file that is created when the program runs. The Java program also relies on the config.txt file that stores the Amazon Seller Central API Keys. It also includes the dates for the program to pull from. The data from the API is then store in the txt file AmazonData.txt. Running this program will also generate the Amazon_Response.txt file.

Cleaning Data

The next step after collecting the data from Amazon is to cache it with the latitude and longitude coordinates for the addresses. Running the ConvertAddresses.py python script takes all of the addresses in the AmazonData.txt file and annotates them with their latitude and longitude coordinates by running it through Google Maps' geocoder. This will output the new txt file AmazonDatawithLatLong.txt, which is AmazonData.txt, but now with the latitude and longitude coordinates attached at the end.

Collecting Specific Data

The next step of the program is to collect the data that is going to be visualized. This is done using the GenerateQuesryResults.py python script. The script will ask for a product's ASIN, the start date, and end dates of the sales that are going to be plotted. The script then

creates a new txt file called queryResults.txt, which is the latitude and longitude coordinates based on the inputs of the program.

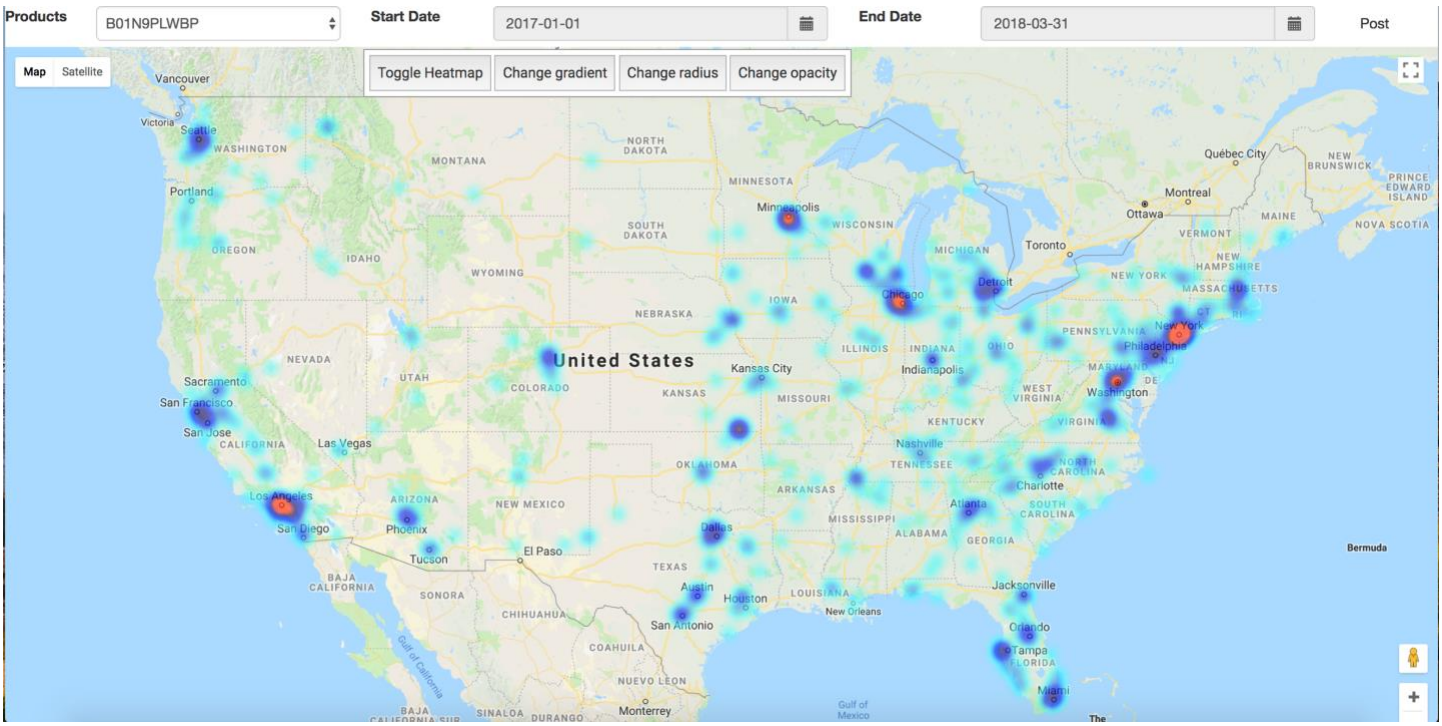
Visualizing the Heat Map

The first step to displaying the heat map is to launch a virtual cloud based server that allows a user to run client and server side programs. Philip used MAMP to launch the heat map while running PHP and Python. There are multiple programs involved in generating the heat map. The fist one is to run the Python Script products.py, which goes through the raw Amazon data to create the products.txt file, which is full of each unique ASIN. This txt file is used to create the product selection for the heat map. The heat map html file relies on three backend programs: products.php, query.php, and GenerateQueryResults.py. The products.php file is the one used to populate the product selection drop down menu on the heat map. The query.php file is a php script that is a loop used to take the latitude and longitude coordinates in queryResults.txt and populate the heat map. The GenerateQueryResults.py script is the python script that is built into the map and allows for the queryResults.txt file to be altered based on the three inputs at the top of the map. By clicking post at the top of the heat map the file will run. The heat map can be opened by opening the heatmap.html file on the cloud server. It should then run dynamically.

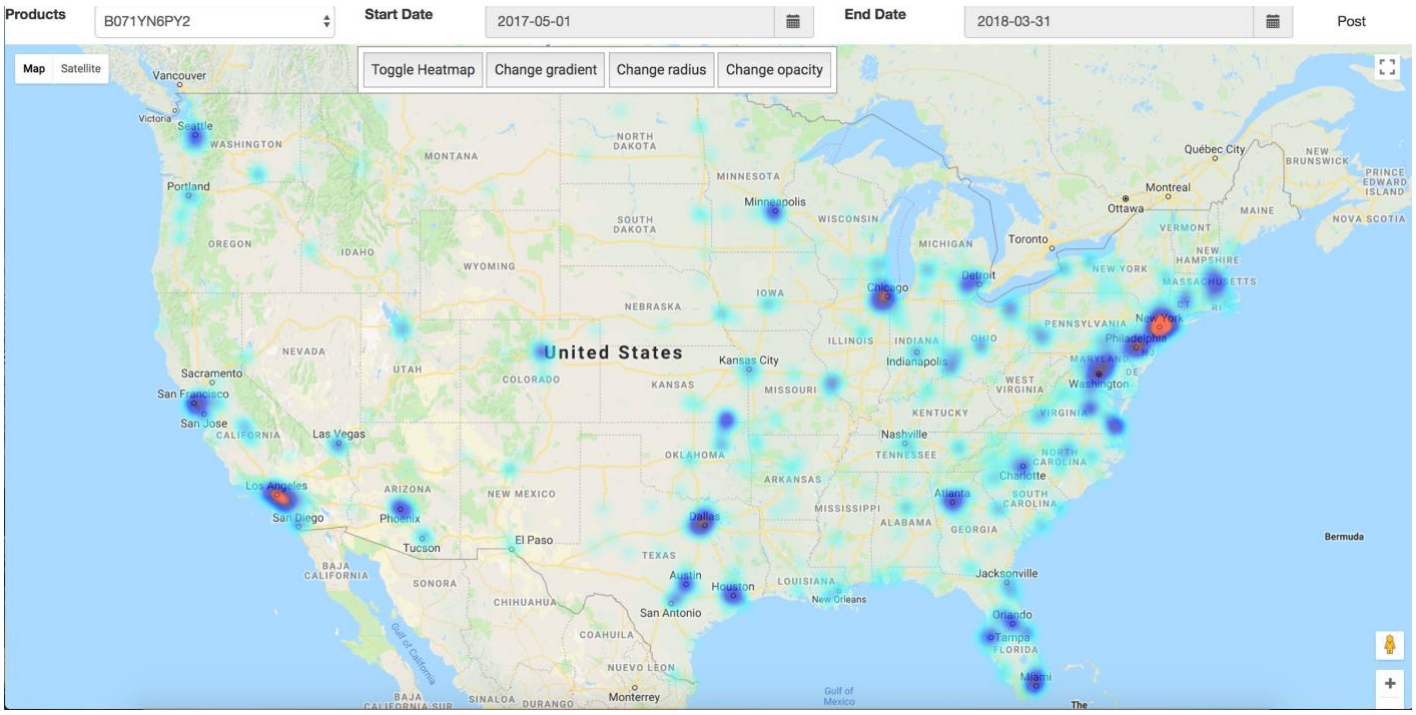
Final Product

The final product results in a heat map that can be generate dynamically and has interactive inputs. The user is also able to change the heat map by toggling it on and off, changing the gradient, the radius, or the opacity.

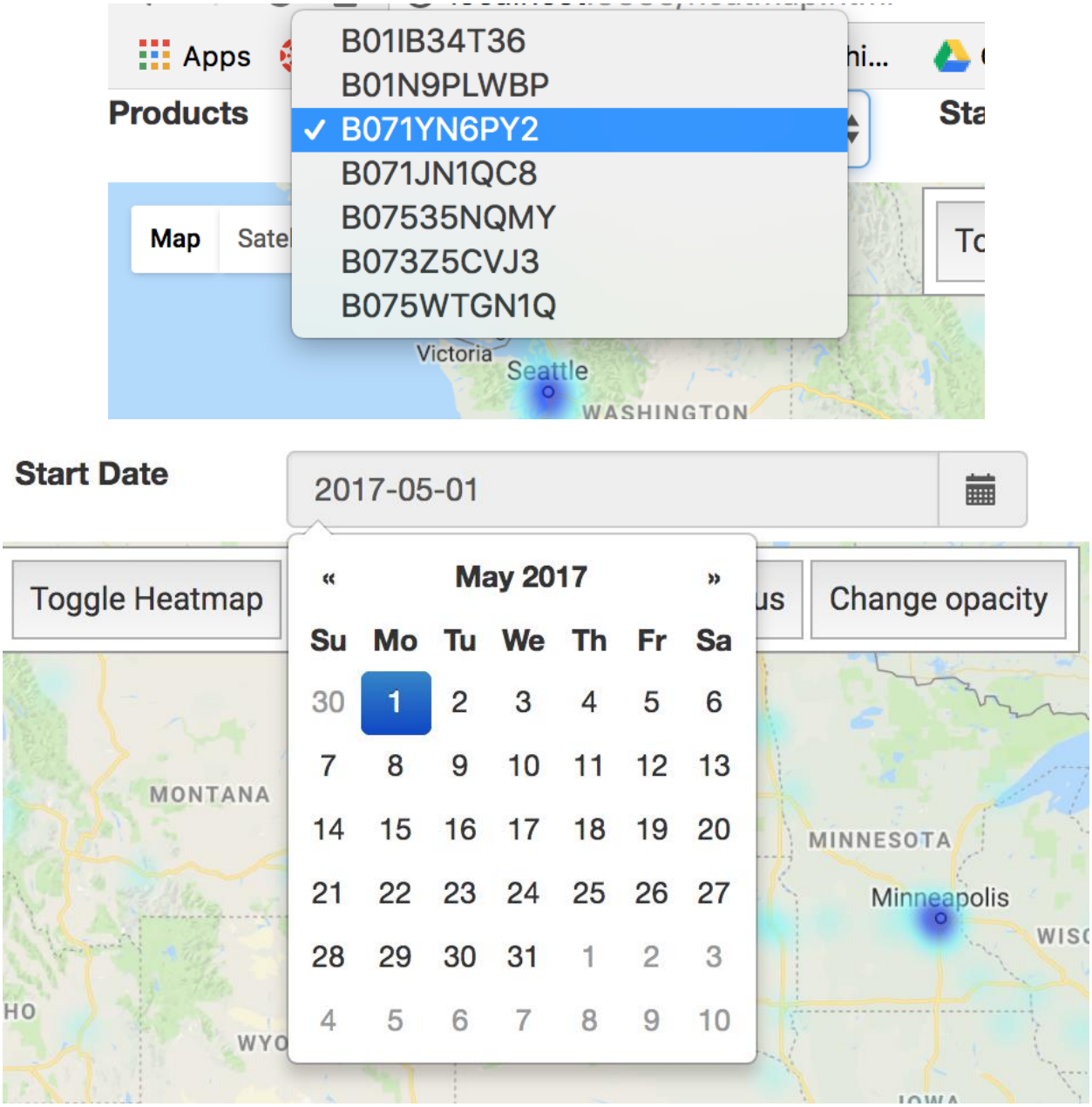
Sales Data of Philip's 8 Pack Wall Mount Sign Holder from 01/01/2017 to 03/31/2018



Sales Data of Philip's Acrylic Computer Keyboard Stand from 05/01/2017 to 03/31/2018



Interactive Controls of ASIN Selection and Dates



Next Steps

The next steps for this project would be to launch it as a fully complete automatic application for other Amazon Sellers called Streamline. Amazon Sellers would be able to log into the application and automatically be able to create a heat map of their sales data. In terms of making this into a business, the first step be allowing the people who were interviewed and provided input to use it as beta testers. Feedback would be collected from these users to update and improve the platform based on their responses.

The initial community and users would be built by utilizing Facebook groups. Philip is constantly active in various Amazon Seller Facebook groups and forums, such as Reddit, to establish his own community and network. Philip would leverage these resources and his connections to grow the user base. Philip would continue to market the platform through content marketing by driving Amazon Sellers to his website by writing then positing not only beneficial, but also educational material for other Amazon Sellers. Philip would be using his success from creating a business that generates $30,000 a month in revenue to create a story for Streamline.

The visual heat map would be the first tool produced for Streamline. Many more would follow though, such as accounting, product tracking, and seasonality software. The main goal and mission of Streamline will be to develop a

OneStopShop of all of Amazon Seller's needs. The initial platform will cost either $4.99 or $9.99 a month, more data needs to be collected before an accurate price point can be established.

Appendix

Appendix A: Initial Proposal

Action Based Learning Project – Winter Semester 2018

Student: Philip Ruffini
Class: Technology and Operations 399
Academic Sponsor: Professor Sanjeev Kumar
Title: Developing Business Tools for Amazon Sellers to Draw Key Insights
Proposed Number of Credits: 3

Background Information on Project:

Philip is currently a seller on Amazon who specializes in private label office products. His company, Ruffini LLC, generates approximately $30,000 in sales per month. Philip is a Ross School of Business Junior with an emphasis in Technical Operations and a minor in Computer Science. He is an active participant in the Amazon Seller community. He regularly answers questions for other sellers in Facebook groups, keeps up to date through various newsletters, and has been featured on Amazon Seller stories hosted by Greg Mercer. Additionally, Philip participated in Ross' Dare to Dream program where he interviewed over 25 Amazon Sellers to find an unmet need in data analysis tools for seller data.

Philip believes that there are unmet needs for Amazon Sellers in terms of data analysis tools as the Amazon Seller Central platform has a weak user interface and lacks many basic needs such as accounting and customer relationship management. Amazon does not supply their sellers with product specific information in terms of customer location, seasonality, projected inventory management, or accounting. These four areas all need separate tools for data analysis. Ultimately, the development of data analysis tools will improve Amazon Sellers' capacity to conduct market research and draw key insights from Amazon's raw data.

Goal of Action Based Learning Project:

The main goal of this action based learning project is for the student, Philip Ruffini, to develop data analysis tools for not only his personal Amazon Business, but also the businesses of other Amazon Sellers. These tools will allow Philip to gain key market insights from the raw data that Amazon supplies. This data will primarily focus on customer locations, product seasonality, and product profitability. Thus this project will give Philip a significant advantage for both his business and his future career as an entrepreneur as he will gain first-hand experience building and using his own data analysis tools to draw insights.

Executive Summary:

The main focus of the independent study will be the business objective of improving the data analysis and market research capabilities of Amazon Sellers. This will be done by utilizing newly built information technology tools to organize raw data. Amazon supplies their sellers with raw data that is difficult to categorize, collect, and understand. Philip will be utilizing technology to create data analysis tools to analyze and draw insights from the raw data. The initial tool that will be focused on will be a heat mapping software for customer locations based on the product

type. This will allow Amazon Sellers to visualize the raw data collected from Amazon's platform. For example, Amazon Sellers are able to see each of their orders, which specifies the product purchased and the address of the customer. However, those orders are individually listed and not grouped by location or even product type. The seller receives the orders based on the time they are placed. A heat map will allow Amazon Sellers to visualize the geographic locations of their Customers across the United States sorted by product type. The heat map will allow Amazon Sellers to visualize the location of their customers for a specific product and identify if there is a correlation between location and product type. With this information, Amazon Sellers will be able to optimize their social media Ads, such as Facebook Ads, by focusing on specific geographic location such as select zip codes that have a high purchasing rate for a specific product.

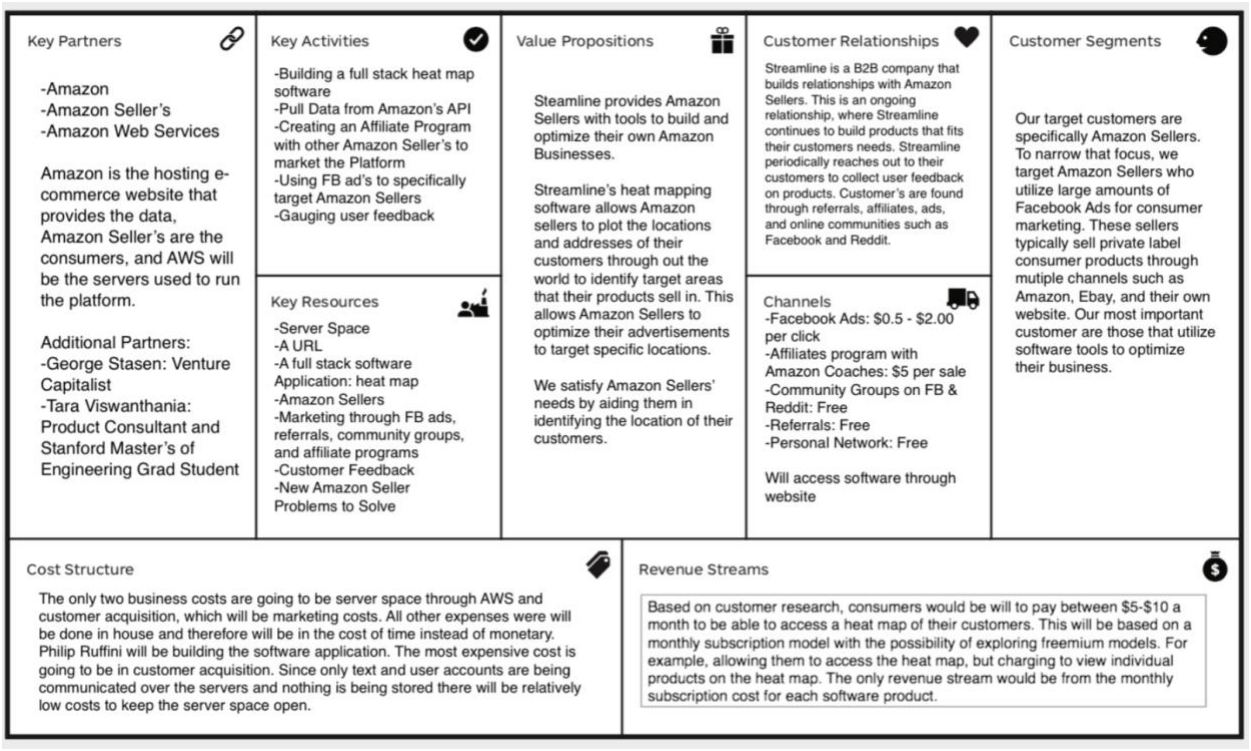
After the completion of a Heat Mapping tool, two other tools to work on developing to benefit the data analysis capabilities of Amazon Sellers would be analyzing seasonality of products and automatic accounting. Currently, Amazon Sellers have no way of checking if a potential product is seasonal or not. Amazon Sellers are only aware of the seasonality of a product if they are presently selling it and notice the fluctuation in their sales. However, by utilizing Amazon's API developers are able to pull past sales rank of products they do not sell. Therefore, a software tool will allow Amazon Sellers to put in the SKU number of select products they plan to sell, then visualize the change in the product's sales rank on a line graph. This will allow Amazon Sellers to see a product's change in sales over time, even if they do not sell it. Utilizing this data, along with Google Trends, will allow Amazon Sellers to create a seasonality index for their products to not only visualize, but also quantify a product's seasonality. Moving forward, a later stage product to develop for Amazon Sellers would be accounting software specific for Amazon Sellers. Amazon's dashboard only projects sales and does not manually calculate profit based on product cost and shipping. This is difficult to do for most Amazon Seller's as they regularly change the price of their products, which impacts the listing fee Amazon charges. Utilizing automatic accounting would save hundreds of hours for Amazon Sellers and provide them improved profitability analysis for their products.

Key Deliverables:

- Market Analysis of Current Amazon Software Platforms
    - The student will complete a full market analysis of current software applications to determine the benefits and costs of each platform. This will allow the student to identify what current customers do and do not like as well as identify where unmet needs are for Amazon Sellers.
- Data Analysis Tool: Online SaaS platform – Heat Mapping Amazon Customer Locations
    - The student will produce a full stack online software application that allows users to register and heat map their products. The product will integrate user registration and payment processing, which will be two more key learning concepts for the student.
- Insights from Data Analysis Tool: Applying the Heat Map
    - The student will utilize his heat mapping program to draw insights from the raw data to see where he should be focusing his marketing strategy for his consumer products such as a scratch off map of the world and a camera lens attachment for phones. The student will optimize his online advertisements to focus on the key areas that consumers are located.
- Future Projects

- The student will continue to follow these steps to develop other tools for Amazon Sellers. Two of these two tools will be the seasonality tracking and accounting software programs.

Appendix B: Business Model Canvas



Appendix C: Files

See attached folder for all Files