Peter Mattoon

Overview, Key Words, and Subtopic Name:
　　A website is being created to analyze sets of data in order to draw conclusions from that data. The website will start by allowing the user to input one or more sets of data and then ask for the type of analysis desired. After analyzing the data the website will return its findings along with a graphical representation of the data in order to make it more easily understood.
Key words: Analysis, Big Data, Statistical Analysis, Graph, Website, Computer Science, Software
Subtopic: Information, Computer Science and Engineering (EA5)

Intellectual Merit:
　　This Small Business Innovation Research Phase I Project will be designing an algorithm to make multiple different analyses on a data set in a reasonable amount of time while still providing an accurate and comprehensive analysis of the entire data set. The creation of such an algorithm is technologically difficult because the software must be programed to sample enough data for an accurate analysis without wasting time sampling extraneous data. The algorithm must also be able to determine which data sets may be analyzed together, such as ones that were collected over time versus ones that cannot be compared against each other. The goal of this research project will be achieved through rigorous experimentation on a sample data set and when completed it will provide an abundance of tools to analyze data.

Broader/Commercial Impact:
　　The research project described above will be extremely important to smaller businesses in the coming years because of the increasing reliance of companies on big data. All businesses big and small collect huge amounts of data on their customers. Large companies are starting to utilize that data to better understand the needs and wants of their customers so that they may offer a better experience in the future. The currently large cost and time consuming nature of such data analysis means that small businesses are losing to their competition; however this online analysis tool will be able to help even the field at an affordable price and in a timely manner. Even for those without a commercial need will be able to use the website to answer questions and understand data on just about subject.

Elevator Pitch:
　　The website created from this project will be and do a host of different things for a wide variety of people in a fast, easy and understandable way. The website will be able to analyze large sets of data supplied by the user and will return an understandable analysis. However, while the website will only be able to analyzes data, the kind of data and analysis will differ from one user to another. The main focus of the website will be analysis for smaller businesses who cannot afford to hire someone but truly anyone will be able to use the website. Since it will easy to use and will return an easily understandable answer.
　　The value that the website will bring to its customers relies upon its ease of use, its affordability and its meaningful analysis. The website will be easy to use because it will be a

two-step process for any users. First they will have to upload their data in a clear format for the website to then read into a database. Then they need select the type of analysis they would like done on their data and that is it. All they need to do after that is wait for the data to be processed. Furthermore, the website will either be free with ads or have a small membership fee depending upon serve costs to run the website. Lastly, and most importantly the website will be able to supply the user with an easily understood and meaningful analysis of their data, which will include graphs that fit the type of analysis carried out on the user's data. The website will be a product that anyone will feel comfortable using and easily understand.

Moreover, while the website can be used by anyone with a dataset, the target customers are small businesses looking to improve the experience of their customers. The small businesses are the target consumer of the website because they have the data and a need to analyze and understand it at a low cost. Small businesses collect huge amounts of data on their customers in the normal course of business but mostly they do not do anything with that data either because of a lack of time, understanding or both. However, they all want to because they know it will lead to an improved customer experience. They cannot outsource the analysis because of cost and some just do not understand how to get anything meaningful from the data. That is why the website will be a perfect fit for small businesses trying to improve their customer experience and compete with larger businesses that can afford to pay for data analysis.

The website is not going to be the industry leader in data analysis but it is going to be extremely easy to use and understand. The data will be supplied by the user and the analysis selected and then after running it through our software the website will return the data in a meaningful and comprehensible format. That is the area in which the website is going to innovate.

In conclusion the website will be an affordable and valuable tool for anyone, but especially small businesses to turn their data into knowledge about their customers.

Commercial Opportunity:

The data analysis website will be competing, for most of its business, in the lower end of the consulting market. A growing market that as of 2013 had a value of around 40 billion dollars, however we are not going after the entire market but instead we are targeting the lower and less well served end. The ideal client for the our website is a small business looking to improve its ability to service its customers. As every business becomes more reliant on technology for every part of their activities they inevitably collect data on their own customers, as well as their suppliers and simply how the business operates. All of this data can be extremely useful to a company but it has to be analysed and properly understood before it can become anything more than a large amount of confusing numbers and letters. Large businesses can afford to higher consulting firms that will analysis everything and then provide recommendations on how to improve but smaller firms and businesses do not have the capital to afford a data analyst that can charge up to 250 dollars per hour. The consulting and data analytics marketplace is full of high end firms but the lower end is very underserved because a profit cannot be made with the current model.

The key risk involved in the data analysis website is that it cannot be protected from others who wish to copy it. The website is built to provide analysis of data but the way which the

data is analysed is not unique. The platform and ease of use are new, however, a common statistical algorithm cannot be patent protected and new statistical algorithms cannot be invented because that would lower the credibility of the website. That is why to be successful this website must be the first on the market in order that we secure a large market share and loyal customers. After we first draw in our customers we can insure they stay with the site by consistently increasing, the forms of analysis, ease of use and the clarity of the results.

However, while having a large market share is important for any business it is useless without any way to turn those customers into profit and the way that the website will do that is through ads and memberships. Ads for almost any website are an obvious choice but if the right ads are found a significant profit can still be made. Many websites today rely on google's ad service to get ads for their websites and while this works it is not the best option. First, because there may not be any synergy with our website and they are not the best paying. The best paying ads is when another company directly pays for space on a website and we will attract that type of ad by reaching out to other small business focused companies. The second source of revenue will be through memberships, which will provide extra services to our clients. Our clients will need those extra services to analyze any truly large data sets because the free version will have file size constraints, due to time concerns. Furthermore, customers will also be able to access more forms of analysis with the memberships that are too computationally intensive to run for free. Overall, our website is positioned to be a strong business in the future.

Societal Impact:
One of the goals of this project is to provide a way for small businesses that so far have been left behind by the increasing importance to big data to start to catch up to their larger competition. This product will allow the smaller players to continue to continue to compete in today's economy and in doing so help the economy as a whole since competition is the cornerstone upon which capitalism is built. If used wisely the data analytics website will allow small business to better understand how to serve their own customers with an easier and more enjoyable experience. And so if the product is widely used consumers everywhere will have more enjoyable experience throughout their normal lives.

Thankfully the data analysis website is very safe and will not do any harm if used correctly. Environmentally, other than normal electricity costs the website will not have any negative effects, furthermore, the website could lead others into being more efficient with their use of resources and in doing so help reduce the need for natural resources. When it comes to health effects the changes will only be positive due to increased customer experiences at health care providers. Moreover, in the case of children, if they use the website at all, it will only leave them more knowledgeable.

The only worrying aspect of the website is that it can be used by anyone with any data and so it is accessible to criminals as well. Fortunately, the system would require data sets from the criminals to be of any use and that requirement would later help authorities to convict them when they are caught. Overall, the website will do a lot of good for the world and its customers.