**Deliverable 2**

**Memo**

Data mining is the process of identifying information from vast amounts of data. Data mining utilizes employees and automated programs to analyze an organization’s data looking for trends, patterns, anomalies, or outliers, with the hope of gleaning actionable insights from the evaluation of the data. Data mining seeks to transform volumes of data that a business collects into information that businesses can use in ways such as better understanding their customers, evaluating product sales performance, or even improving processes. With contributions from statistics, artificial intelligence, and data visualization and leveraging business analysis and technical development, there a myriad of ways to conduct data mining and the field of data mining continues to grow today.

Regarding data mining, Thomas Davenport in the Harvard Business Review details, “The latest strategic weapon for companies is analytical decision making.” [1] Davenport further explain companies “have used analytics to better understand their customers and optimize their extended supply chains to maximize their returns on investment while providing the best customer service.” [1] Most importantly, Davenport concludes “the level of success is highly dependent on a company understand its customers, vendors, business processes, and the extended supply chain,” [1] which can be done using data mining.

A great, easy to comprehend example of data mining is data mining used in retail stores on customer and transaction data. Upon analyzing this data, data mining can reveal unknown insights regarding customer shopping habits, transaction size or frequency, and even specific product sales details. A excellent real-world application of data mining comes from Target. As describe by Forbes “Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy.” [2] One of the ways Target uses data mining is to try “to hook parents-to-be at that crucial moment before they turn into rampant -- and loyal -- buyers of all things pastel, plastic, and miniature.” [2]

As such, Target applies data mining techniques to their customer details and buying habits to identify customers on the verge of becoming parents, sometimes before the customers even realize it themselves. Target created a ‘pregnancy score’ based on customer buying habits of key items such as unscented lotion, supplements, and big bags of cotton balls. Based on the pregnancy score, Target began sending automatic mail coupons for nursery furniture and maternity clothing to customers exceeding a certain score. The article further sites an instance where Target new one man’s high school age daughter was pregnant before he did. The man found out about the pregnancy after seeing Target baby clothes coupons in the mail. Target used data mining to better understand their customers and in this case, because of data mining, Target advantageously knows more about their customers than their customers know about themselves.

The chief benefit of data mining is identifying actionable information that is advantageous to the organization that may not have been discovered otherwise. Companies collect troves of data through a myriad of sensor, device, and input network each day. Unfortunately, much of that data is never used. Data mining puts this previously untapped data to use by uncovering hidden value, unknown insights, and quite unexpected realizations within the data. In describing the benefits of data mining, Dr. Arno Penzias explained, “If you’re not doing this, you’re out of business.” [1] Dr. Penzias frames data mining not exactly as a benefit, but more so a necessary means of survival in today’s hypercompetitive business marketplace.

Data mining is often hypothesis-driven, which allows the company to ask and seek answers a plethora of critical questions. By using hypothesis testing, companies can use the data they already have to answer the outstanding questions necessary to improve their business. Asking good questions are key to hypothesis-driven data mining and included are several questions SuiteSpot may attempt to answer using data mining:

1. What are the most popular dates for dinner at each hotel location?
2. As storing inventory, particularly perishable items, for an extended period of time can be costly, what appetizers, entrees, and desserts are the least popular?
3. What desserts are most often selected with what entrees?
4. What appetizers and entrees sell the most wine along with the dinner?
5. How many tables does each restaurant turn over in a night and how fast on average does each table turn over?
6. What is the average table size in each restaurant and what is the average party size at each restaurant?
7. Do certain menu items, either food or drinks, sell better in certain locations?
8. What is the most common payment method used and does payment method have any relation to tip amount?
9. Does the day of the week of the dinner date correlate to any food or drink purchases?
10. Which restaurants receive the highest amount in tips and which food and drink items correlate with the highest amount in tips?

To even perform data mining, SuiteSpot will first need to have adequate information technology infrastructure on hand. On the software side, SuiteSpot will need the appropriate data mining software. Many big players in the software world publish their own data mining software such as IBM SPSS, Statistica, and SAS. SuiteSpot would also be able to use business intelligence software such as SAP, IBM Cognos, or Teradata for data mining although their capabilities are limited when compared to data mining software. SuiteSpot will also likely want to have a software platform and analytics code language such as R, SQL, or Python so employees can conduct their own, ad-hoc type analyses. Additionally, SuiteSpot will need to have their data stored in a data warehouse or a collection of data marts so these software tools will be able to analyze the data.

On the hardware side, SuiteSpot will need server space to store the data and hold the software necessary for data mining. SuiteSpot has plenty of options in this arena and they can choose from a physical/on-premise hardware architecture, a cloud-based architecture such as Amazon Web Services (AWS), or a mix of both. SuiteSpot should decide on their hardware strategy first as without any hardware to back it up, software is useless.

**Data Mining**

**Data Analysis**

Overall, the data from each location is deeply uncorrelated. While mining data through each of the location, steak stands out as a top seller for entrees but is closely followed by chicken as well as vegetarian options. A conclusion that can be drawn by this data is that steak is prepared well, and based on typical meat price points, customers are willing to spend money for a well-prepared entrée. To improve the order rates of the other entrees, recipes can be improved and advertising can be increased. However, under the assumption that steak is the highest priced entrée, each location is doing well with maintaining the highest profitability possible. In order to lower costs, each location could build partnerships with vendors providing this type of meat in hopes of receiving lower prices and turning a higher profit.

For Appetizers at each location, there are a lot of people not choosing an appetizer before their salad or entrée. This can be concluded that appetizers are not well advertised on the menu or within a restaurant. Because the restaurant needs to be prepared for someone to order an appetizer, there is a risk that ingredients are being wasted- therefore lowering the potential profits of the restaurant. A conclusion that can be drawn from this data is to increase the pairings of the appetizers and entrees and encourage customers to purchase an appetizer through “paired” deals.

For desserts, cheesecake is the definite choice for the customers at each location across the board. The conclusion of this data can be drawn from the cheesecake well-prepared dish as well as perhaps “word of mouth” advertising. Most restaurants experience a “chain” reaction for customer orders. Once one person speaks about the quality of a dish, more people are willing to try it. Chardonnay is also a popular pairing at each location with the cheesecake. In order to increase profit, the restaurant could do a cheesecake and chardonnay pairing to increase customer dessert activity.

As stated before, there is not a specific correlation of data between each location. There are definitely top sellers and lower sellers, but each entrée, desserts, and appetizer has a decent showing at each location. I think advertising and entrée and appetizer pairings are a good idea to increase the profit of each restaurant. Tipping is sufficient at each restaurant for the most part which shows that customer service is doing well, I simply thing the menu and restaurant appeal would help to increase profit and lower costs.

**References**

[1] Sharda, R., Delen, D., & Turban, E. (2015). Business Intelligence and Analytics (10th ed.). Upper

Saddle River, NJ: Pearson.

[2] Forbes. (2017). Forbes.com. Retrieved 28 September 2017, from <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#2c85e77b6668>