CREATING THE FOUNDATION FOR A STARTUP: DESIGN AND
IMPLEMENTATION OF A DATABASE SYSTEM FOR AN ELECTRICAL
BICYCLE RENTAL COMPANY

A Senior Project submitted to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree of

Bachelor of Science in Industrial Engineering

by

Alejandro Martinez

March 2017

ABSTRACT

CREATING THE FOUNDATION FOR A STARTUP: DESIGN AND
IMPLEMENTATION OF A DATABASE SYSTEM FOR AN ELECTRICAL
BICYCLE RENTAL COMPANY

Alejandro Martinez

Abstract.

BoltAbout was started by 4 Cal Poly students, and their COO is a fellow Industrial
Engineering student that I met during the first half of the senior project series class. He
presented project to me and I decided to work with him in developing this new database
for his startup company. BoltAbout's customer base is growing and current operations
are not fit to meet the demand of a growing company. More specifically, their data
management system can't be used by all current employees to perform essential tasks
such as rental processing, bicycle and equipment organization, and customer data
analysis. This is the origin and need for this project. The end goal was to design a
completely new database with an appropriate structure and user friendly interface that
would allow all employees to utilize it efficiently. The current and future needs of the
company were used as the main requirements for the database design, this decision would
ensure that the database is functional today but also five year down the line.

The final design meets the need for a simpler and more intuitive user friendly interface
with a main form that serves as a starting point for all functions or tasks employees might
need to perform. From this main form, a user can simply click a button to go to a form to
either register or update a customer, register new equipment, add equipment maintenance
details to the maintenance log, create, edit, or end a customer order, and view statistics
about their monthly performance. All these tasks are one click away. As a conclusion, the
design successfully meets BoltAbout's needs for a simpler and functional data
management system. In addition, it allows for flexibility if the company wants to add a
new product line or different additional equipment models in the future with minimal to
no changes required to the database structure or design. Future recommendations would
be related to data security. Currently, to be able to fully use the database and all its
operations, the user needs to have access to all the information. As the company and the
employee base grows, this availability to information might have to be limited or address
as a potential area of improvement.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

Page

LIST OF FIGURES

I. Introduction

BoltAbout is an electric bicycle rental company, founded by four Cal Poly students and located in downtown San Luis Obispo. The problem originated due to BoltAbout's expanding customer base and the current operations not being fit to meet the demand of a growing company. More specifically, there is no data management system that can be used by all current employees to perform essential tasks such as rental processing, bicycle and equipment organization, and customer data analysis. The idea for this project originated from IME 312, one of Cal Poly's Industrial Engineering (IE) course on data management systems. Peter Kean, one of the cofounders of BoltAbout, is a current IE senior student enrolled in the senior design course. This projected originated from him looking for another IE student who was knowledgeable about data management systems and could use techniques from IME 312 to design system that can solve all of the current needs and issues the company is going through. I am attempting to design a database management system that can tackle all the current needs mentioned above applying both, techniques I learned from IME 312 and additional knowledge and topics I find through my literature review. Overall, this report aims to solve or address the following objectives:

- Provide background information required to understand how database management systems work and their applications.

- Understand the current state of BoltAbout's operations and how a database management system would be the adequate solution for this problem.

- Provide a design that allows for data security and various user-friendly interfaces.

- Design of standardized data entry process for constant up-to-date and complete data records.

- Equipment, sales, and customer related metrics that allow for data analytics to be used in order to enhance company's performance and development.

As part of this project, I intend to complete the following deliverables:

- An extensive literature review that provides the background necessary to understand how database management systems work along with their use and application specifically in this scenario.

- The design process, methodology, and documentation utilized such as diagrams, tables, charts, etc.

- Microsoft Access Database file of finalized database design.

- Recommendations for optimal use, maintenance, and manipulation/adjustment of database system.

In order to reach the desired solution, the approach to be followed starts with brainstorming. To learn what others out in the industry have done and can be utilized as part of the solution for this problem, I will start with a literature review on methodology and examples of how databases have been used in different industries, companies, or scenarios. In my solution reaching approach, brainstorming will be followed by data gathering of the system's current state along with current input data and desired output information expected from the designed solution. Once these details are gathered, the next step in my approach is to combine and organize both, the results obtained from the literature review and the information gathered from the system's current state to begin the design of a suitable solution. The beginning of the design phase includes the planning of data tables necessary, structure of tables, relationship between entities, and data entry process. Then, moving to the design implementation phase, the approach would include the building of form interfaces, coding and query language writing required to produced desired outputs, and page to page system navigation to ensure that the database is user friendly, intuitive, and fluent. Lastly, to reach our solution, we need to incorporate the current data available into the database to test the design and ensure that it provides the desired results. Adjustments, if any, will be made in order fix possible bugs or design issues that might emerge. This is the approach I plan to follow in order to reach the desired objectives.

The key tasks that are essential to the success of this solution are:

- Detailed, complete, and accurate collection of initial data.

- Careful design of database basic components (tables, queries, forms)

- Careful design of database infrastructure (forms, data manipulation, code)


II. Background (includes **Literature Review**)

In order to understand the constrains in the design for this project, it is important to know the current state of the company. BoltAbout is a young startup that began its journey last year. As a young startup, the company is still working towards developing standard

operations for their specific company, market, and environment. With that in mind, there are pros and cons about this situation. For instance, the lack of standard operations makes it difficult to have a framework, criteria, or basis to build the design off of. On the other hand, not having a strict design criterion to satisfy takes away possible limitations that could prevent us from using innovative techniques and methods or certain design structures. Not having these limitations allow us to create the best design possible.

Some of the aspects we need to consider for this design are based on desired outputs specified by the BoltAbout, as well as required features based on current assumptions. The design must comply to the assumption that the company will grow and data will become an important company asset that will drive BoltAbout's performance and operations. This asset however could become a liability or be at risk. For that reason, an important aspect of the design is data security. BoltAbout deals with sensitive private information on their clients and the design must ensure that this information is secure form potential future employees or non-employee individuals that come across the database system. Next, we will go over general database systems background information to provide a basic understanding of what databases are and how they function. Also, I will provide detailed information on database security methodology, and applications.

Literature Review

The foundation of this research is based on the usefulness of databases and their potential for assisting companies with advancing their business strategies and operations. Data is abundant, and we as individuals generate and consume large amounts of data which can be stored and managed using databases. Figure 1.1 below is an example provided by Carlos Coronel and Steven Morris in their book "Database Systems: Design, Implementation, and Management" of how much data we as individuals come across on a daily basis.

FIGURE 1.1 THE PERVASIVE NATURE OF DATABASES

8
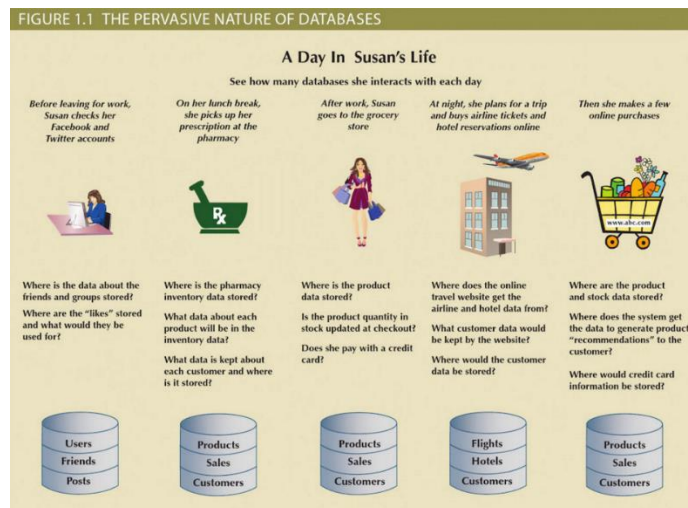
The size and nature of the data varies depending on the type of system or nature of business. For example, telecommunication companies such as AT&T store and manage data on trillions of calls with data input rate of up to 70,000 calls per second. This is an impressive number, but it helps us put into perspective the size of the data that we as consumers produce on a daily basis, and the potential and power behind understanding the trends and information that this massive amount of data has to offer. Figure 1.2 below shows how we can go from data to useful information that businesses can utilize.

FIGURE 1.2 TRANSFORMING RAW DATA INTO INFORMATION
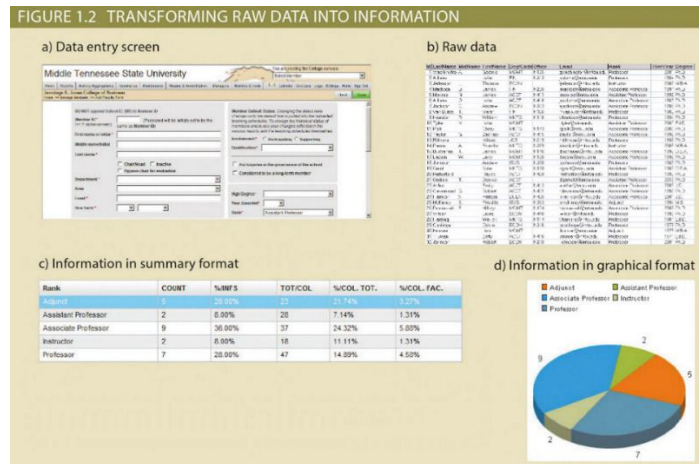
To summarize the topic of data and information importance and usefulness, there are some key points that Coronel and Morris pointed out for us to keep in mind.
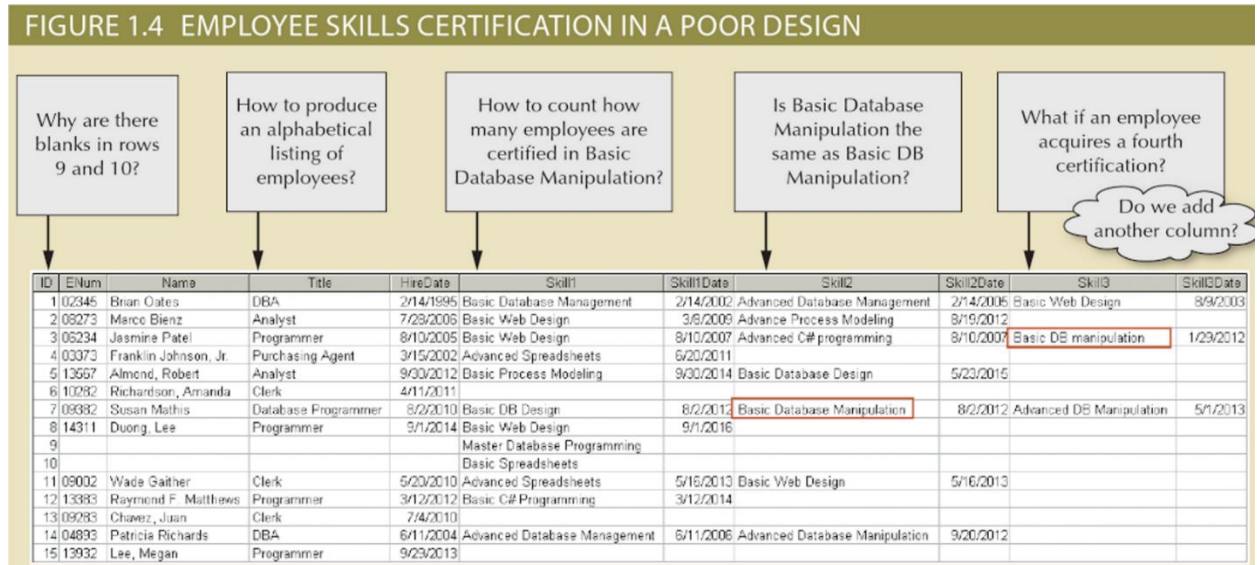
1. Data constitute the building blocks of information
2. Information is produced by analyzing data, therefore data collection is essential for learning, development and growth
3. Information is used to reveal the meaning of data
4. Accurate, relevant and timely information is key to good decision making
5. Good decision making is key to organizational survival in a global environment

This is the value that a database system can provide for Boltabout, as the company begins its journey in the world of business, consumers, and data.

Database Design

A current evolving problem is the typical lack of database design skills that lead to poor performing databases management systems. Let's first define database design as the activity of designing the database structure that will store and manage the end-user data. Different users require different data to be analyzed to produce information relevant to their business. In order to produce the required information requested by the user, the way in which the data is stored and analyzed has to be properly designed in order to

process and output the desired information. The following example helps us understand how design of database structure is crucial for appropriate database performance. Let's look at Figure 1.4 where all the data was compiled and mixed together.

FIGURE 1.4 EMPLOYEE SKILLS CERTIFICATION IN A POOR DESIGN

ID	ENum	Name	Title	HireDate	Skill1	Skill1Date	Skill2	Skill2Date	Skill3	Skill3Date
1	02345	Brian Oates	DBA	2/14/1995	Basic Database Management	2/14/2002	Advanced Database Management	2/14/2005	Basic Web Design	8/9/2003
2	08273	Marco Bienz	Analyst	7/28/2006	Basic Web Design	3/8/2009	Advance Process Modeling	8/19/2012		
3	06234	Jasmine Patel	Programmer	8/10/2005	Basic Web Design	8/10/2007	Advanced C# programming	8/10/2007	Basic DB manipulation	1/29/2012
4	03373	Franklin Johnson, Jr.	Purchasing Agent	3/15/2002	Advanced Spreadsheets	6/20/2011				
5	13567	Almond, Robert	Analyst	9/30/2012	Basic Process Modeling	9/30/2014	Basic Database Design	5/23/2015		
6	10282	Richardson, Amanda	Clerk	4/11/2011						
7	09382	Susan Mathis	Database Programmer	8/2/2010	Basic DB Design	8/2/2012	Basic Database Manipulation	8/2/2012	Advanced DB Manipulation	5/1/2013
8	14311	Duong, Lee	Programmer	9/1/2014	Basic Web Design	9/1/2016				
9					Master Database Programming					
10					Basic Spreadsheets					
11	09002	Wade Gaither	Clerk	5/20/2010	Advanced Spreadsheets	5/16/2013	Basic Web Design	5/16/2013		
12	13383	Raymond F. Matthews	Programmer	3/12/2012	Basic C# Programming	3/12/2014				
13	09283	Chavez, Juan	Clerk	7/4/2010						
14	04893	Patricia Richards	DBA	6/11/2004	Advanced Database Management	6/11/2006	Advanced Database Manipulation	9/20/2012		
15	13932	Lee, Megan	Programmer	9/23/2013						

The issues with this approach are the following:

1. It would be difficult, if not impossible, to alphabetical listings of employees based on their name
2. In order to determine how many employees are certified in Basic Database Manipulation, multiple variables would need to be created by programs and then added together.
3. You run the risk of misspelling employee names or skills creating redundancy and inaccurate records.
4. The structure of the database would have to change if a forth skill needs to be added, and again every time a new skill is added.

On the other side, let's turn our attention to Figure 1.5 where the relational database model is used.

FIGURE 1.5 EMPLOYEE SKILL CERTIFICATIONS IN A GOOD DESIGN

Database name: Ch01_Text

Table name: EMPLOYEE

Employee_ID	Employee_FName	Employee_LName	Employee_HireDate	Employee_Title
02345	Johnny	Jones	2/14/1995	DBA
03373	Franklin	Johnson	3/15/2002	Purchasing Agent
04093	Patricia	Richards	6/11/2004	DBA
06234	Jasmine	Patel	6/10/2005	Programmer
08273	Marco	Bienz	7/28/2006	Analyst
09002	Ben	Joiner	5/20/2010	Clerk
09283	Juan	Chavez	7/4/2010	Clerk
09382	Jessica	Johnson	8/2/2010	Database Programmer
10282	Amanda	Richardson	4/11/2011	Clerk
13383	Raymond	Matthews	3/12/2012	Programmer
13567	Robert	Almond	9/30/2012	Analyst
13932	Megan	Lee	9/29/2013	Programmer
14311	Lee	Duong	9/1/2014	Programmer

Table name: CERTIFIED

Employee_ID	Skill_ID	Certified_Date
02345	100	2/14/2002
02345	110	8/9/2003
02345	180	2/14/2006
03373	120	6/20/2011
04093	180	6/11/2006
04093	220	9/20/2012
06234	110	8/10/2007
06234	200	8/10/2007
06234	210	1/29/2012
08273	110	3/8/2009
08273	190	8/19/2012
09002	110	5/16/2013
09002	120	5/16/2013
09382	140	8/2/2012
09382	210	8/2/2012
09382	220	5/1/2013
13383	170	3/12/2014
13567	130	9/30/2014
13567	140	5/23/2015
14311	110	9/1/2016

Table name: SKILL

Skill_ID	Skill_Name	Skill_Description
100	Basic Database Management	Create and manage database user accounts.
110	Basic Web Design	Create and maintain HTML and CSS documents.
120	Advanced Spreadsheets	Use of advanced functions, user-defined functions, and macroing.
130	Basic Process Modeling	Create core business process models using standard libraries.
140	Basic Database Design	Create simple data models.
150	Master Database Programming	Create integrated trigger and procedure packages for a distributed environment.
160	Basic Spreadsheets	Create single tab worksheets with basic formulas.
170	Basic C# Programming	Create single-tier data aware modules.
180	Advanced Database Management	Manage Database Server Clusters.
190	Advance Process Modeling	Evaluate and Redesign cross-functional internal and external business processes.
200	Advanced C# Programming	Create multi-tier applications using multi-threading.
210	Basic Database Manipulation	Create simple data retrieval and manipulation statements in SQL.
220	Advanced Database Manipulation	Use of advanced data manipulation methods for multi-table inserts, set operations, and correlated subqueries.

The entity-relationship model finds its strength in simplicity. Figure 2.2 below graphically explains how this model works.

FIGURE 2.2 A RELATIONAL DIAGRAM

AGENT
- AGENT_CODE
- AGENT_LNAME
- AGENT_FNAME
- AGENT_INITIAL
- AGENT_AREACODE
- AGENT_PHONE
- AGENT_ADDRESS
- AGENT_CITY
- AGENT_STATE
- AGENT_ZIP
- AGENT_DATE_HIRED
- AGENT_YTD_PAY
- AGENT_YTD_FIT
- AGENT_YTD_FICA
- AGENT_YTD_SLS
- AGENT_DEP

CUSTOMER
- CUS_CODE
- CUS_LNAME
- CUS_FNAME
- CUS_INITIAL
- CUS_AREACODE
- CUS_PHONE
- CUS_INSURE_TYPE
- CUS_INSURE_AMT
- CUS_RENEW_DATE
- AGENT_CODE
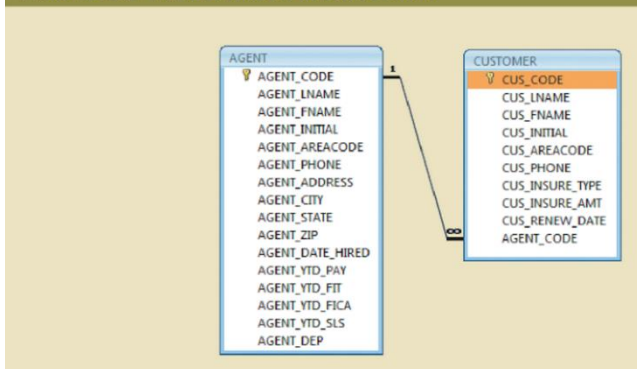
The relational diagram is a graphical representation of a relational database's entities such as AGENT and CUSTOMER, the attributes within each entity, and the connection between entities. Notice that the connection between entities that is key to the relational model is thanks to the overlap or commonality of an attribute between the two entities. The relationship between agents and customers is possible thanks to the AGENT_CODE, an attribute that is unique to the AGENT, but also part of CUSTOMER. This way we have the capability of analyzing the any possible correlations between agents and customers. For example, what type of insurance do customers have based on the agent's state or zip code. This information can potentially help the company better assign agents to customers based on the agents' location.

Another concept required for the correct implementation of a relational database model is relational algebra and how query language is used to process data. As explained in the book "The Structure of the Relational Database Model", relational algebra is the initial tool presented to express queries in a relational database. Relational algebra is a set of operations carried out by a query resulting in an answer. For instance, using query language we can make a query that uses the relational algebra "sum" and adds all of the customers assigned to "agent A" this year, then divide by 12 and result in the average number of customers assigned to "agent A" per month. In general, we create queries using relational algebra to process data and produced answers to our business questions.

Using the concepts of relational database models, relational algebra, query language, as well as other concepts I learned through Cal Poly's data management course as my design foundation, I will have a solid ground on which to build a strong, useful, and high performing database system for BoltAbout that will help them adapt, develop, and reach their business goals.

Task-role-base access control model in smart healthcare system

One aspect of the database design to consider is security and access control to sensitive data. There are different models to Access control (AC), such as Mandatory AC (MAC), Discretionary AC (DAC), and Role-based AC (RBAC). Professor Ravi S. Sandhu, has provided many articles describing Role-based Access Controls as well as their functionality and implementation. Professor Sandhu is an expert in the field of database and information cyber security. He teaches at the University of Texas at San Antonio and has published dozens of articles that have been cited thousands of times on Google Scholar. In addition, he holds many patents on cyber security systems and methods. The information he provides on access controls, specifically RBAC, is well defined, useful, extremely valuable, and a key component I would like to incorporate into the design of this database.

For DAC models, access for each subject in the system is specific, and all the subjects and objects are enumerated. Examples of subjects can be users, groups, or processes that act on behalf of other subjects. Role-based AC (RBAC) models are gaining popularity thanks to the several well-recognized advantages this model provides. Considering how roles represent organizational responsibilities and functions, a role-based model supports arbitrary, organization specific security policies. Another advantage is the simplification of security administration to organize access privileges. In simple terms, it is easier to say or determine who gets access to what data using RBAC. A clear example of how this access control methodology is used is presented in the article Role-Based Access Control Models, where the authors provides the following example. An operator role could access all resources but not change access permissions, a security-officer role could change

permissions but have no access to resources, and an auditor role could access audit trails. This administrative use of roles is also found in modern network operating system (Sandhu, pg. 2).

This does not mean that roles and permissions are strictly exclusive, in reality, well designed RBAC can allow for interactions between roles, permissions, and responsibilities. As claimed by the authors, these role-role relations can be used to enforce security policies that include separation of duties and delegation of authority (Sandhu, pg. 2). The authors of "Role-Based Access Control Models" provides us with Definition 1 below, which serves as a model or set of requirements that provide guidance for the design of a role-based access control model.

Definition 1 The $RBAC_0$ model has the following components:

- U, R, P, and S (users, roles, permissions and sessions respectively),

- $PA \subseteq P \times R$, a many-to-many permission to role assignment relation,

- $UA \subseteq U \times R$, a many-to-many user to role assignment relation,

- $user : S \to U$, a function mapping each session s_i to the single user $user(s_i)$ (constant for the session's lifetime), and

- $roles : S \to 2^R$, a function mapping each session s_i to a set of roles $roles(s_i) \subseteq \{r \mid (user(s_i), r) \in UA\}$ (which can change with time) and session s_i has the permissions $\cup_{r \in roles(s_i)}\{p \mid (p, r) \in PA\}$. $\square$

As a start-up, BoltAbout might not encounter this issue just yet since all four current members need to be fully involved with all of the company's operations, however it is essential to think ahead of time and design under the assumption that the company will grow. With growth, data becomes a stronger driver for improving the business, it becomes a handy tool, but also a valuable asset that needs to be secured.

Role-Based Access Control Features in Commercial Database Management Systems

As mentioned previously, one aspect to consider is expandability, and how the database security and access granting authority will be designed. The reason why expandability is important is that as mentioned currently BoltAbout has only four employees who are very involved with all the operations and they all need open access to all available information whether it is for operational purposes, marketing, or data analysis. However, we need to contemplate the possibility that as the company grows, more employees will be hired to most likely performed clerical tasks such as completing purchases and registering new clients, which means these new employees will need to access the database which also

contains valuable and private information. With this in mind, I hope Microsoft Access allows for a design that uses a role and permission granting system. Chandramouli Ramaswamy and Ravi Sandhu present an interesting methodology used by commercial database packages in which roles can be assigned to one or more people, and it has to be assigned by the "System Security Officer." The article by Ramaswamy and Sandhu discuss many of the functionalities these commercial packages are capable of, however we are restricted to Microsoft Access capabilities, and we are mainly interested in borrowing the concept of having a System Security Officer that can grant roles and permissions without providing the options for users to propagate the given role and permission to other users. The design must allow for an interface where the System Security Officer can perform this tasks.

Data mining security and privacy

Data mining is the process by which database users utilize data collected in the database to discover information we didn't expect to find. This could create a security or privacy issue, having important information we did not intend to have. Data mining is a powerful tool that provides lots of business benefits by analyzing corporate data which can be used to improve efficiency. Another advantage of this technology is information sharing throughout an organization or publicly. The article "Security and Privacy Implications of Data Mining" mentions the sharing of corporate telephone books to the public, and how this can decrease the need for telephone operators by offloading this task to whoever is calling. On the other hand, when we combined these two advantages, a potential disadvantage is created. Clifton and Marks use the example of a supplier, Dedtrees, who gets access to and mines a retail store's database to then alter their consumers' tendencies through coupon advertising. Customers' purchases changed and affected competition between products and as a result supplier relations and deals. Some solutions are proposed and these can be directly related and used for the design of my database tables and data accessibility to ensure only certain data can be access, seen, or analyzed.

Methods to secure databases against vulnerabilities

The main component of a relational database is a table. Tables represent entities that are part of the system. For example, in a company there are employees, departments, clients, and other entities that are modelled by a table in the database system. Taking employees as an example for an entity, the columns of the table would describe common attributes that all employees share, such as name, address, and salary. Tables can have as many attributes as needed or allowed by the system, but the key characteristic of the entity is that all employees share the same attributes. One of the attributes is chosen as the primary key, and this attribute must be unique for each employee so that it can uniquely identify and serve as reference for that employee. In this example, the primary key may be the

employee's identification number. What makes relational database systems unique and broadly accepted is how this model allows for entities to have relationships with each other. For example, a department, which is one entity, can work with many different employees, which is another entity. Those employees can work with many different clients, which is yet another entity. We can also have a more romantic and closer one to one relationship such as employee to cubicle, which could be another entity. In the previous example, you can also see how these relationships can be of two types, one to many (one department to many employees), many to many (many employees to many clients), or one to one (one employee to one cubicle). These relationships may also be represented by tables, and again each column would represent an attribute that all relationships have in common. Therefore, in this database model, all data is very structured. We need to note that these relationship cardinalities as well as other additional rules governing these relationships are what the database management system uses to enforce consistency within the database.

The reason why consistency is so important is that it is a key component of ACID (Atomicity, Consistency, Isolation, and Durability). Relational Database Management Systems (RDBMS) provide ACID to those who choose to use this model. Let's go more in depth into what each one of these attributes represent so we can have a better understanding of the qualities RDBMS can provide for us. As defined by Jonathan P. Sloan in his Master's thesis: Atomicity means that a RDBMS will complete all elements of a transaction or will complete no elements of the transaction in the event of an error (Sloan, pg. 6). In this case, the RDBMS will try again to complete the entire transaction. Consistency means that a RDBMS will ensure that after every write to the database, any subsequent read to the database will reflect the new data and not the old data (Sloan, pg. 6). Isolation means the RDBMS will ensure that individual transactions will not interfere with other transactions (Sloan, pg. 6). Durability means the RDBMS will ensure that any changes to the database are persistent (Sloan, pg. 6). A write to the database will not suddenly roll back to a previous state. Examples of RDBMS are MySQL, Microsoft Access, and Microsoft SQL Server. I will be using Microsoft Access as the RDBMS platform for my design, making consistency as well as the other attributes previously mentioned a key asset of my design.

Key Success Factors for Eco-Businesses

Making a profit is the principal focus of any business. In order to make money, companies need to constantly analyze their operations and processes to identify room for improvement. The following are lessons that provide ideas on how to increase profit and reduce operational costs. These can be summarized as the eight rules for increasing profit and eight rules for reducing costs for eco-businesses but many of them are directly applied to the bicycle rental company we are focusing on.

General rules for increasing profit

Servicing: This refers to selling service or functionality rather than a product. While the product is still owned by an eco-business provider, customers pay for use or maintenance. E-learning and videoconference substituting for transportation are examples of servicing.

Timesharing: By encouraging users to abandon individual ownership, more intensive utilization of products can be realized. This can reduce users' procurement costs and risks for disposal of products. Leasing and rental schemes are examples of this rule.

General rules for reducing costs

Management of life cycles: Proper management and control of product life cycles can reduce costs. This can be directly implemented into the design of the database by recording data on the date products were received, maintenance done on bicycles, and resell dates. This inputs can help BoltAbout keep track of all products' life cycles, and total return on investment as well as potential resell salvage value.

Utilization of knowledge and information: High utilization and sharing of knowledge and information of business processes can increase efficiency of material/product usage, labor, and capacity of facilities. This is another rule that suggests the implementation of a database as a mean for cost reduction through the usage and management of data.

Components of a bicycle rental company

Another part of this research is to identify the components of a bicycle rental company. Bicycle -sharing systems are typically share the following components:

- Shared bicycles need to be easy to use, adaptable to users of different sizes, mechanically reliable, resistant to vandalism or theft and distinctive in appearance. The exceptional high quality electrical bicycles offered by BoltAbout meet this component.
- Most bicycle -sharing systems are equipped with bicycles that weigh between 16kg to 22kg and are heavier than typical personal bicycles. They are sturdy and designed to be used between 10 and 15 times a day in all weathers.
- Most bicycles come equipped with a Global Position System (GPS) unit, a Radio Frequency Identification (RFID) tag, or other type of tracking mechanism. This function is typically used in fleet management and retrieval of lost or stolen bicycles. The benefit of having a database management system for this component is the potential implementation of anti-theft GPS tracking methods at BoltAbout. In addition, to make them unattractive to potential thieves, they are made in such a way that special tools are required to disassemble them and their components

are incompatible with other bicycles. BoltAbout bicycles do meet this component as they have special keys and special equipment needed to recharge them.

Conceptual Knowledge Discovery in Databases Using Formal Concept Analysis Methods

In this publication, the author introduces the concept of Knowledge Discovery in Databases (KDD). KDD is aimed at the development of methods, techniques, and tools that support human analysts in the process of discovering useful information and knowledge in databases. This is as we as we have learned essential for growing companies that want to use the available data to advance their businesses. In addition, the author introduces the idea of human-centered KDD which refers to the constitutive character of human interpretation for the discovery of knowledge, and stresses the complex, interactive process of KDD as being led by human thought. An interesting point made by the author is how an indispensable part of the discovery process is that the analyst explores the data and sifts through the raw data to become familiar with it and to get a feel for what the data may cover. As a designer of the database analytics and data mining processes, I must pay close attention to the data and the potential ground it could cover.

Data Mining and Knowledge Discovery: Making Sense out of Data

Databases provide the necessary infrastructure to store, access, and manipulate the raw data. Databases provide a structure that helps break down the analysis for and easily digest the details of how the data is stored and retrieved.
An emerging area in databases is data warehousing. Data warehousing primarily involves storage, data selection, data cleaning, and infrastructure for updating databases after new knowledge or representations are developed. Once the clean and processed data is appropriately warehoused, we can get ready to utilize it and that's were statistics come in. Statistics plays an important role primarily in data selection and sampling, data mining, and evaluation of extracted knowledge steps.

Current Solutions to Tracking Systems

A big part of the bicycle rental business model is being able to provide security for the company's equipment specially the bicycles currently rented out by customers. Part of my research is to find what has been done to address this topic. The use of databases and tracking systems go hand in hand and the following are some examples of patented tracking systems.

- Bicycle theft monitoring and recovery devices
    - Pub. No.: US 2013/0150028 A1
    - Pub. Date: Jun. 13, 2013
    - Bicycle theft monitoring and recovery systems include a tracking device with geolocation capabilities that is mount able to or integrated into a bicycle or other assets.

Current solution to marketing support database management method, system and program product

This invention relates to database management and usage. There is a special focus on collecting, analyzing, and presenting data by extracting input data from an input database. The inventor explains that as enterprises grow larger, and more complex, their supply and buying needs, as a customer, grow even more complex. This situation is similar to what BoltAbout is currently going through with their expanding customer base. An example provided is how a representative servicing one line at a given customer may have need for information about other lines sold to that customer, or about similar lines sold to a different customer.

Furthermore, the inventor address the need that exists for rapid collection, analysis, and presentation of mission critical data located across several enterprise-wide databases. In addition, there exists a need for rapid extraction of input data from one or more input databases, especially with transformation into a suitable schema, that is, suitably specialized, tailored, and engineered schema, for subsequent analysis. This means that a schema especially engineered to fit BoltAbout needs must be designed. A further need exists beyond analysis of the extracted and transformed data, for presentation of the analyzed, transformed, extracted data in a useful and understandable manner. This final point can be met by utilizing statistics, graphs, and charts that can clearly express what the data gathered is telling us.

III. Design (or Theory)

This project includes the development of a Microsoft Access database system. There will not be any other interface or website used as part of the development other than the MS Access file. The MS Access database will store and keep track of the following data required by BoltAbout.

Customers
- Name
- Phone number
- Email
- Date of birth
- Driver's license ID number
- Address
- Occupation
- Reference
- Rental start date
- Rental end date
- Packaged purchased ($79, $99, $129)

Bicycle Information
- ID (Label)
- Model
- Serial number (manufacturer)
- Bicycle key ID
- Date Received

Battery Information
- ID (Label)
- Model
- Date Received

Lock Information
- Serial Number (manufacturer)
- Model
- Date Received

The previously mentioned data are inputs that BoltAbout will be required to enter in order to process their customers, and perform their inventory management tasks. This input data will also create the output data necessary to help BoltAbout make better decisions based on their actual data collected and utilize it to drive their business. Another requirement of this database is functionality and user interface.

The MS Access database will have interfaces that meet at least the following requirements:
- Item search capability: this will allow BoltAbout to look up any of the items that are registered in the tables, know their location, or what user has them if they are rented out.

- User search capability: this will allow for the search of a user by either user ID or user name (first or last). The result of the search will provide basic contact information for that user (email address and phone number). Also, the search will provide information on the items currently checked out by this user.
- Statistics page: this form will be designed to provide useful statistics that BoltAbout can visit monthly to assess company monthly performance and drive improvement.
- Registration page: this form will provide the functionality of registering new users and recording their information into the user database table. Also, the registration form will allow for the entry of new equipment into the database tables.

The chosen platform was Microsoft Access, the decision to use this database system was the solid knowledge foundation, I have using this system. The course I took on database management systems, IME 312, heavily focused on Microsoft Access as this is an easy to use, free entry level software, that has all the capabilities needed to satisfy any possible need of a small company. It would be ideal to utilize the skills and knowledge from my curriculum, rather than purchasing and learning a new software package that might be more advanced, though underused. The design for the tables will be based on relational database techniques and normalization rules. The course, IME 312, and my literature review research support the use of these methods as essential techniques for a strong and successful database design. Database interface (forms) will be based on the collaboration of both, myself as an industrial engineer considering knowledge on human factors engineering to create an intuitive and easy to navigate interface, and BoltAbout management who will provide input on their processes and how the database interface should be laid out to match their processes.

The overall approach was based upon the outputs desired out of the database system, and the assumption of growth and the need for security associated to the company's growth.

The initial step in my design was to create the appropriate tables required to store all the data BoltAbout needs to keep track of in addition to any other functional tables the database requires to perform functions such as creating reports, searches, and statistics.

Figure 2.3 below shows the initial proposed ER (Entity-Relationship) diagram for the database. This network had the potential of working appropriately for the required functions. However, this solution was changed after analyzing potential future complications due to having multiple tables, one for each type of equipment. An alternative would be to have one master inventory table that includes every single piece of equipment, bicycles, batteries, and locks. The reason this approach was initially disregarded was that each type of equipment has unique qualities or properties that are not shared with the other equipment types. For example, bicycles would have a bicycle key ID which needs to be recorded whereas, locks and batteries do not have this feature. Once the benefits of consolidating those tables into one master equipment inventory table

were compared versus the cost of having multiple separate tables for each equipment type, which is to have null values for some table fields, it was concluded that the best approach would be to change the structure of the database to a single master inventory table. Overall, all three equipment types and potential new equipment will share most of the inventory master table fields such as serial number, model, and date received. Also, the fields where null values are present are not used or referenced at all which means they will not ever affect database usage or cause any database malfunctioning.
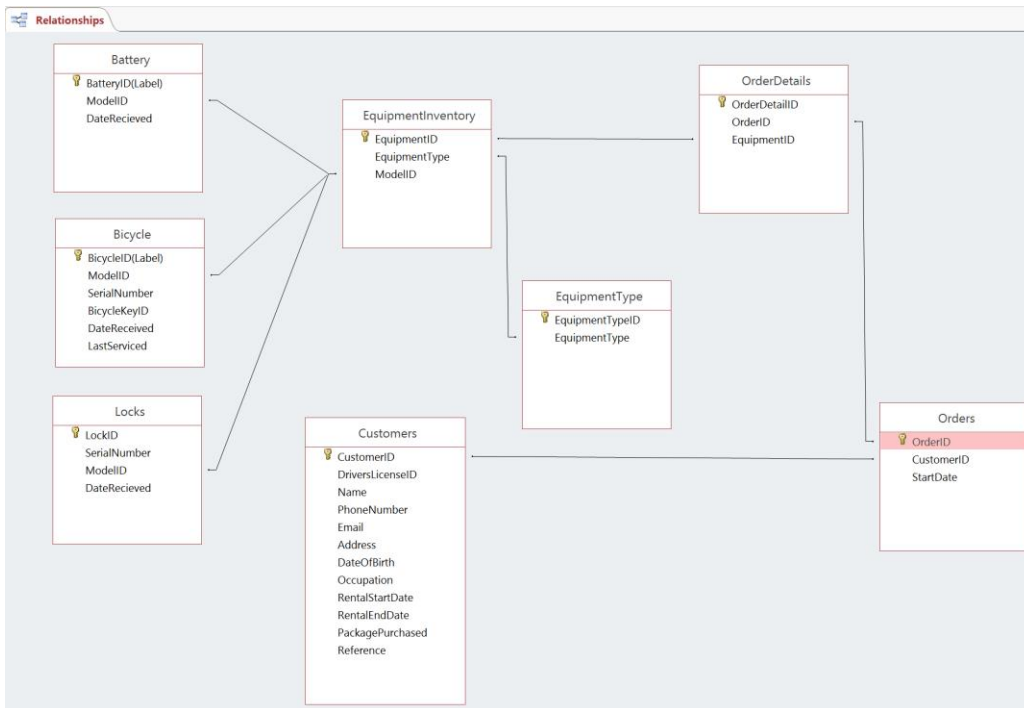
FIGURE 2.3 ALTERNATIVE ENTITY RELATIONSHIP DIAGRAM

Also, another change that was made to the initial design of the ER diagram was the adjustment of order tracking. Initially, the database was designed to have one orders table where every detail of the order was going to be recorded. The issue with this approach was that the table would be too complex and would increase the room for technical malfunctioning. A suggestion made by this senior project's technical advisor was to add an orders details table and split up the complexity. The original orders form would create unique orders that keep track of the order ID, customer associated to that order, and the order start data. The additional order details table would hold the rest of the order's information such as the equipment associated to that order which will be connected to the master inventory table. This way customers, orders, and inventory are all connected and can be clearly managed. The previous approach would have multiple orders with similar users each with similar inventory items. Figure 2.4 below shows the structure of the new ER diagram after the previously discussed changes were made.
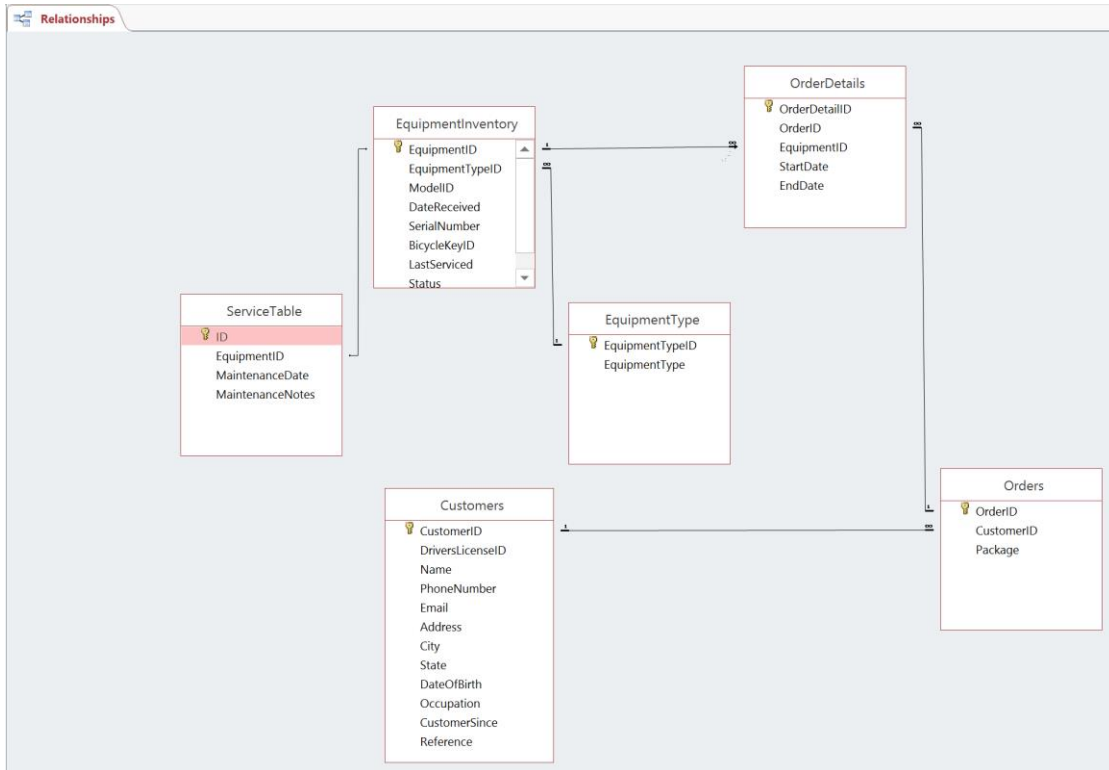
FIGURE 2.4 SECOND ALTERNATIVE ENTITY RELATIONSHIP DIAGRAM

This second revised ER diagram was utilized as the framework to design most of the database forms. This alternative met all the current requirements for BoltAbout's current needs. However, flexibility was not considered at that point and the ability to adapt to changes was limited and difficult as the database structure and design would require major changes. In order to add new equipment models or a different product line that might be purchased in the future, new tables and forms would have to be created and the design would be heavily constrained by the current structure of the database. Knowing this, additions were made to the ER diagram with additional tables. A table for order packages and equipment models was added and linked to the existing ER structure by the model ID field which can be linked to the model ID field in the master equipment inventory table. Following this approach, new promotional packages can be added and linked to different models allowing for customized orders. In addition, a table for models was added and linked to the existing equipment type table allowing new product lines and models to be seamlessly added to the current operations without the need to change or alter the structure of the database. The final design for the database structure can be seen on the ER diagram below shown by figure 2.5.
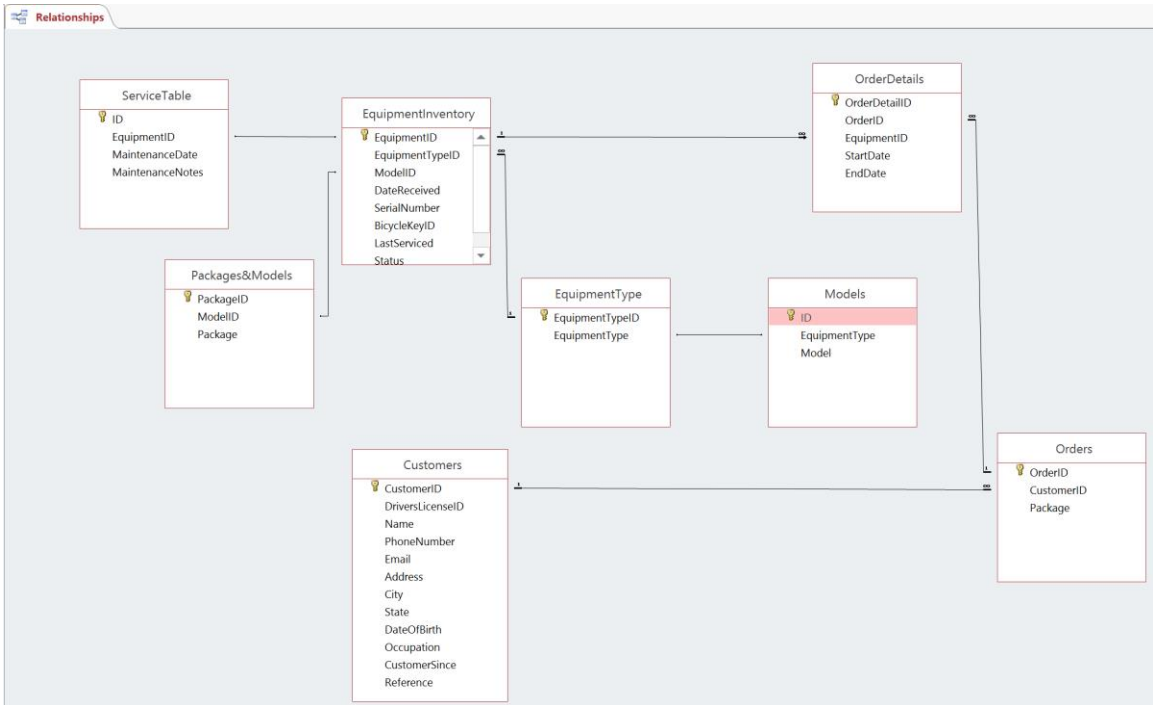
FIGURE 2.5 FINAL ENTITY RELATIONSHIP DIAGRAM

This final database structure meets not only the current needs of the company but also considers future potential needs of expansion and change making this database functional today and tomorrow.

The following figures show the final design of some of the different database user interfaces used to perform the tasks required by BoltAbout.

FIGURE 3.1 MAIN FORM USER INTERFACE DESIGN

The main form design seen above in figure 3.1 shows what was previously described. This form is a simpler and straight forward interface that allows user to perform all the tasks they might need by clicking just one button. They can search and view rental and maintance records for any piece of equipment by selecting it and clicking the "SEARCH" button, or if they get new equipment, they can go ahead and add it to their inventory by clicking the "NEW" button under inventory. Without the feel of a confusing and unorganized form, customer management is conviniently located right next to inventory management options. The options provided for customer management include registration of new customers by clicking the "NEW" button, viewing any selected customer's activity by clicking the "VIEW" button, updating customers' details and contact information by clicking the "UPDATE" button, or placing a new order by selecting a customer for that order and clicking the "PLACE ORDER" button.

FIGURE 3.2 INVENTORY EQUIPMENT DETAILS USER INTERFACE DESIGN

When an inventory item is selected, in this case a bicycle, and the user clicks on the "SEARCH" button, the interface shown in figure 3.2 above is displayed. This interface clearly shows all records associated to an inventory item organized by rental and maintenance records with the ability to add new service entries by clicking "Add Service Entry" or go back to the main form by clicking "BACK".

FIGURE 3.3 NEW ORDER PLACEMENT USER INTERFACE DESIGN

Another example of the database user interface design is shown in figure 3.3 above. On the left-side, we can see how this form looks when the place order button is clicked on the main form. Then, the user can select a bicycle, a battery, and a lock based on the different models BoltAbout has. Then a package is selected determining the monthly cost and user privileges the customer wishes to pay for, to then click the "Complete Order" button and finalize the order placement process as seen on the right-side in figure 3.3.

FIGURE 3.4 CUSTOMER INFO UPDATE USER INTERFACE DESIGN

The above form is what's seen by the user after selecting a customer (Maicol Petrulus in this case) and clicking the update button on the main form. All fields display the customer's current information. The user simply replaces the field that needs to be updated and clicks the "UPDATE" button to finalize updating that customer's information.

Each button on the main form opens a different form to either display or enter the information needed. This simply but efficient central form navigation system will improve user efficiency and minimize erroneous and or missing information as the forms are coded to ensure all required fields are completed before completing the task. This database design has been successfully designed and created, now it is important to discuss how to test and ensure the database works properly and assimilates to BoltAbout's current operations and data.

IV. Methods (or **Experimentation**)

The previous database which contains all the current available information for customers and equipment was provided by BoltAbout. The best way to test whether the database can be directly implemented into BoltAbout operations is to use the current data to simulate customers coming in and registering them as you would if a normal customer was to walk in. The same process would be followed to test the equipment registration form.

The testing phase is only required the current information BoltAbout has, and the new database. All testing was done electronically as the database is a Microsoft Access system software system. In addition to the usage of BoltAbout's current information, the database was tested by simulating different scenarios the user can run into. An example of such scenario includes selecting in a combo box that changes the options for another combo box such is the case for equipment type and model options which depend on the previous equipment type selection. Attempting to "break" the code or complete an action incorrectly was the second phase for testing. Human error is inevitable, and solving issues that might be caused or can potentially occur due to user error is also the responsibility of the designer. Testing was completed once each form was tested by completing several trials of the operations it is designed to perform and passing the "breaking" secondary test.

Most issues that occurred and needed to be fixed occurred during the secondary test, when attempts to break the form were successful. Some examples are as previously mentioned, making a selecting that would later be changed and ensuring other options would change accordingly.

V. Results and Discussion

The result for this project was exceeded the initial expectations. The final design for the database has a rational flow as the ER diagram was carefully designed, reviewed, and revised multiple times as improvements were made and issues were addressed. The final structure is straight forward and allows for future modifications in areas where BoltAbout might later expand. These areas include new product lines and new equipment models. The way this was accomplished was by separating the source of equipment and models to a single table and referencing this table from all other forms. This means that if BoltAbout wants to add two more models for their bicycles, they can simply type them into the "Models" tables and they will instantly and automatically be available under the model options for any form where bicycle models are an option.

The most important customer requirements which are the customer and equipment search capabilities are available and function as intended. These capabilities also make other tasks easier, for example the customer search capability makes creating new orders for existing customer quicker as the search capability allows the user to find any user faster to then move on to creating an order by simply selecting the customer and clicking the "Place Order" button.

Another customer requirement was the simplification of the database as most users at BoltAbout are not familiar with Access. The result is a main form from where all operations can be performed. The design is also intuitive and follows a sequence where operations are likely to be performed as a series. An example of a series of operations is registering a new customer and placing an order for that customer. Considering that this is very likely to be case for most new customers, the user is automatically taken to the order placement form right after registering a new customer and associates this new order to the customer that was just registered. If this customer does not want to place an order yet, the user can just exit out of the order placement form and go back to the main form. This approach is intended to minimize user confusion. In addition, throughout the database, code enforces the user to not miss or leave fields of information empty which will ensure orders and information is complete and accurate.

BoltAbout is less than two years old as a company. Standard processes are still being designed, which means that the criteria and requirements for the database must also be designed or considered by the designer taking into account future needs that might emerge. This was one of the major challenges of this project as lacking a direction and customer input made difficult to narrow down the scope and decide what the best design would look like.

Ethics Research and Impact

After consulting the NCEES, I analyzed my senior project's potential impact considering my responsibility to the public welfare as my upmost priority. The analysis of my senior project's potential ethical impacts led to two major points of focus.

First, it is important to carefully design this database considering organizational impacts. This database will deal with sensitive data composed of users' personal information that relates to knowledge retention and information security. For that reason, it is key to take information availability and security into account to ensure users' information is properly used and secured.

The second factor to consider is societal impacts primarily effects on local economy and community health. The implantation of this design can directly lead to the company's growth. Part of the planned design for this database is the implementation of statistics that BoltAbout can use to determine how to better target customers, and this can potentially result in an increase in business for BoltAbout. The company's growth would subsequently lead to more members of our community utilizing BoltAbout services. This translates to an increase in the local economy as well as the community's health.

First, the community can benefit economically by saving on transportation costs. For example, a college student leaving off campus who not only pays for gas as he commutes daily, but also pays for parking permits, or a local business owner driving to work daily. As mentioned, there would be health impacts for our community since there would be an increase in users' physical activity using this bicycles as a main form of transportation.

Economic Impact

The company, BoltAbout, does not require an economic justification for this database for two reasons. First, currently data management is nonexistent and this is a problem that is stopping the development of their business. As their customer base increases, equipment can potentially get lost and/or damaged, and keeping track of the equipment and customers who came in contact with all equipment is essential for accountability purposes. This can potentially create cost savings for damaged and lost equipment. Another source of monetary savings comes from optimal performance of tasks. Employees are currently not able to perform any tasks, and can spend hours tracking incorrectly recorded or missing information. With the implementation of this database those hours are eliminated which represents direct cost saving for the company. Assuming 15 hours are spent on average per week correcting and searching for information, at a rate of $10/hour, the cost savings per week can be up to $150 USD. Overall, the first reason is that BoltAbout needs a data management tool to run their business regardless of the cost.

The second reason is related to the conclusion of the first reason. BoltAbout needs a properly functioning database regardless of the cost. In this case, there is no cost of implementation for my project. This database is currently operational and can be immediately implanted at no cost. For that reason, the economic justification not only comes from the fact that there is no cost for implementation, but all potential opportunity for savings from damaged equipment and optimal task performance.

VI. Conclusions (or **Summary and Conclusions**)

The problem originated due to BoltAbout's expanding customer base and the fact that current operations are not fit to meet the demand of a growing company. More specifically, there is no data management system that can be used by all current employees to perform essential tasks such as rental processing, bicycle and equipment organization, and customer data analysis.

What this means is that the current method used to process orders, register customers, and manage equipment is not appropriate and does not provide the functionality required by BoltAbout. The database cannot be operated by other employees other than the original database creator. Orders are often not registered, there is data missing for multiple customers and orders and data management is currently nonexistent, meaning there is no way to search for and or connect equipment with orders and customers. This creates plenty of room for issues such as lost equipment, erroneous order information, and lack of accountability as records are not accurate.

The ultimate goal for this senior project was to improve the current state of operations at BoltAbout. The most important results are listed below.

○ The main customer requirements included search capabilities and overall simplification of the database interface so that other users who are not familiar with Access can also use the database. These requirements were considered throughout the design and met by the final design.

○ Other requirements such as enforcement of accurate information entry were discussed between the senior project academic advisor and the designer and implemented making this database design professional and more user friendly.

○ The design allows for adaptability and addition of future equipment with minimal to no changes needed to the forms or structure of the database.

○ Economic justification analysis resulted in potential savings of up to $150 USD per week (with given assumptions), as well as other potential sources of savings including equipment damage accountability.

○ Lastly, as a confirmation for the successful completion of the project. The initial objectives listed in the introduction were addressed and met successfully. However, due to the current state of the company and the need for information accessibility, objectives related to data security and statistics have been discarded and suggested as future areas of improvement for BoltAbout. Currently, data security can't be achieved as full access to information is required to perform essential tasks (updating customer information, placing orders, etc.).

Through this senior project, I learned the importance of research. I had limited knowledge of Microsoft Access, coding and appropriate form design. The end product I was capable of producing with my initial knowledge would have not met the requirements for BoltAbout nor this project. Learning and gathering ideas from the client, the academic advisor, and the research assignment is a must for any senior project that attempt to not only repeat what has been already learned in class, but take one's academic and professional career to the next level, one step further.

REFERENCES

Ramaswamy, Chandramouli, and Ravi Sandhu. "Role-based access control features in commercial database management systems." *Proc. 21st Nat'l Information Systems Security Conf*. 1998.

Sandhu, Ravi S., et al. "Role-based access control models." *Computer* 29.2 (1996): 38-47.

Sloan, Jonathan P. *Methods to secure databases against vulnerabilities*. Diss. Monterey, California: Naval Postgraduate School, 2015.

Sandhu, Ravi. "Relational database access controls." *Handbook of Information Security Management, Auerbach Publisher, Philadelphia, PA* (1994): 145-160.

Sandhu, Ravi S., and Hal Feinstein. "A three tier architecture for role-based access control." *Proc. of the 17th NIST-NCSC National Computer Security Conference*. 1994.

Sandhu, Ravi S., and Pierangela Samarati. "Access control: principle and practice." *IEEE communications magazine* 32.9 (1994): 40-48.

Coronel, Carlos, and Steven Morris. *Database systems: design, implementation, & management*. Cengage Learning, 2016.

Paredaens, Jan, et al. *The structure of the relational database model*. Vol. 17. Springer Science & Business Media, 2012.

Kondoh, Shinsuke, et al. "Analysis of key success factors for eco-business through case studies in Japan." *Proceedings of 13th CIRP International conference on Life Cycle Engineering*. 2006.

Midgley, Peter. "Bicycle-sharing schemes: enhancing sustainable mobility in urban areas." *United Nations, Department of Economic and Social Affairs* (2011): 1-12.

Akins, Kristine, et al. "Bicycle theft monitoring and recovery devices." U.S. Patent Application No. 13/712,831.

Fayyad, Usama, Gregory Piatetsky-Shapiro, and Padhraic Smyth. "From data mining to knowledge discovery in databases." *AI magazine* 17.3 (1996): 37.

Lee, Michael, and Ashley Stirrup. "Marketing support database management method, system and program product." U.S. Patent Application No. 09/261,773.

Stumme, Gerd, Rudolf Wille, and Uta Wille. "Conceptual knowledge discovery in databases using formal concept analysis methods." *European Symposium on Principles of Data Mining and Knowledge Discovery*. Springer Berlin Heidelberg, 1998.

Fayyad, Usama M. "Data mining and knowledge discovery: Making sense out of data." *IEEE Expert: Intelligent Systems and Their Applications* 11.5 (1996): 20-25.