

**ITC 6000**

**Database Management Systems**

**Final Project Report**

**2022 Summer Term A**

**BosBus Data Management System**

**Nsikak-Abasi Una**

[**Una.n@northeastern.edu**](mailto:Una.n@northeastern.edu)

**TABLE OF CONTENTS**

1. Introduction

1. Business Analysis
2. Table Design & Analysis
3. Database Implementation
4. Analytics, Report & Metrics
5. Security Concerns
6. Architecture
7. Future Considerations

**INTRODUCTION**

Transportation is essential to the growth of a city and the economy in general. A good transport system helps individuals to access jobs and opportunities, supports innovation as the city will attract companies, and leads to a greener and healthier city. Before coming to Northeastern for my master's program, I lived in Lagos, Nigeria, where the most reliable mode of transportation was a car. You either own a car, carpool with someone or take a taxi which could be expensive. As a result, there were many cars on the roads which led to heavy traffic and pollution. The train system in Nigeria is nonexistent, and the buses are unreliable.

Having stayed in Boston for about four months, I can see the difference between the 2 cities, with Boston having a vast network of buses and an organized system. Bosbus is a bus operator in Boston, Massachusetts. This database management system is a holistic system to track the bus that plies on the routes and maintain a repository of the ticket bookings made for each bus. I chose to work on a bus app because I was and still am fascinated by how the transport system in Boston operates. Considering a large number of bus services and passenger data, it is imperative to consolidate and store the data in a database.

The bus app will be used by teenagers and adults to see the real-time/updated schedule/route of buses and book & pay for a ride. Individuals need to get to their destination (offices, schools, etc.) promptly. Hence, the necessity of the app. The app is free to download, but customers will have to pay for their booked rides on the app via credit/debit card, apple.

**BUSINESS ANALYSIS**

**Primary Use Case**

For a business to remain competitive and relevant in the market, you must deliver a fantastic client experience. Over the past several years, people's expectations of what a positive customer experience should entail have significantly increased. Businesses are prepared to go above and beyond to satisfy their clients. Many businesses can build outstanding customer experiences because of data aggregation. They will have a better understanding of how their clients interact with their products as a result.

**Persona**

For BosBus service, there are 2 major user personas which include

1. **Employees**: people would work for Boston Bus, e.g., system administrator, station manager, bus operator, etc.
2. **Passengers:** these are the people who use Boston Bus as their means of transportation within Boston
   1. **Working class adults:** these users use the bus to commute to and from work. They must know the bus time to plan adequately for resumption and closing.
   2. **Students:** these users use the bus to commute to and from school and like to explore the city during their free time. They require a discount on the regular fee.
   3. **Senior Citizens:** these users use the bus to commute to the local grocery store, hospitals, and pharmacy weekly. They are time sensitive because of their strength. Hence, they require the buses to be a few minutes walk from the local grocery store, hospitals, pharmacy, etc.
   4. **Tourists:** these users are either new to Boston or are revisiting the city. They mostly use the bus to visit different landmarks, restaurants, and shopping malls. They require fair value for the bus expense.

**Business Rules/Logic**

* a **PASSENGER** cannot have more than 1 RIDE ongoing simultaneously.
* different **BUSES** can ply the same route
* a **PASSENGER** can’t BOOK a ride for another individual
* a **BUS** can only RIDE one route
* an **EMPLOYEE** can only WORK at one **STATION**
* a **STATION** can have more than one BUS CONNECTIONS
* a **BUS** will DEPART more than once a day

**TABLE DESIGN & ANALYSIS**

**Data Tables**

* **Passengers**: these are the people who pay to ride the bus. Their details include passenger id, name, Seat no, Gender, Phone number, and email.
* **Bus**: the means of conveying passengers from point A to point B. Bus details include bus id, bus name, bus frequency, number of stops, departure & destination station.
* **Station**: This is where buses stop to pick up and drop off passengers. Station details include station name, id, and the number of connections.
* **Ticket**: a document that shows a passenger has paid for a bus trip. Ticket details include ticket id, passenger id, date, departure/arrival time, fare amount, and departure/arrival station.
* **Fare**: the amount paid for a particular bus trip. Fare details include passenger id, receipt number, fare amount and ticket id.
* **Employee**: the people employed by BosBus for their services. Employee details include name, station, address, role, phone, email, etc.
* **Route**: the particular direction the bus takes from point A to B. Route details include route id, number of stops, and departure/arrival station

**ERD Diagram**

**Diagram

Description automatically generated**

**DATABASE IMPLEMENTATION**

**Queries**

1. The Strategic team requested the total employees currently working in Boston Bus and their respective stations to see if there is a deficit of employees at the various stations. The goal of the strategic team is to increase human capital across all stations to improve customer experience. In this query, I joined the employee table and the station table together.Graphical user interface, text, application

   Description automatically generated

Graphical user interface, text

Description automatically generated with medium confidence

1. The customer experience agent got a call from a customer Niha concerning her lost ticket ID. To get the ticket ID, Niha was referring to; the agent placed a quick call to the database team to get a query for all the ticket IDs for a customer called Niha. For this query, I joined the passenger and ticket table.Graphical user interface, text, application

   Description automatically generated

Graphical user interface, application

Description automatically generated

1. A passenger who took the bus on February 4th, 2022, was seen at a crime scene after the trip ended. To get the identity, the police department requested the buses that rode that day to get the passengers that rode in the buses. To get this query, I joined the bus and ticket table.Graphical user interface

   Description automatically generated with medium confidence

Table

Description automatically generated

1. The Business Intelligence team analyzed data and discovered that Haymarket is the most popular passenger destination. Hence, there is a plan to provide buses from more departure stations so passengers can have more departure options. The database team is asked to provide the list of buses that end their journey at Haymarket ie to find the possible route/starting point to get to Haymarket. To get this query, I joined the bus and route table.

Graphical user interface

Description automatically generated with low confidence Table

Description automatically generated

1. The Sales team requested the total sales made from tickets on each day so far.



A picture containing table

Description automatically generated

**ANALYTICS, REPORT & METRICS**

1. The chart below shows the different bus routes and the number of stops they have. We can determine the route with the fewest stops. According to the graph, Route 675 (Sulivan Square - Ruggles) has the highest number of stops, and Route 682 (Union Square - Oak Groove) and route 678 (Park Street – Lechmere) have the fewest number of stops.



1. The chart shows the age range and gender demographic of the passengers. Boston Bus wants to create initiatives targeted at specific age groups and gender. From the chart, we can see that most passengers are in the 18-29 age range. Hence, BosBus can target this age group with initiatives e.g., discounted rates and increasing the bus frequency in notable areas.



**SECURITY CONCERNS**

Boston Bus database management system will be deployed in Amazon Web Services (AWS) database like RDS to store employees’ and passengers’ sensitive information like numbers, addresses, payment details, etc. Because this database system is stored in the cloud, it is vulnerable to hackers, and there is a need to implement security policies, technologies, and controls like

* Encrypting the users’ data before uploading it to the cloud
* Implementing identity and access controls in the cloud database
* Implementing network and infrastructure security like virtual private networks (VPN), firewalls, etc, that will protect the network.

**ARCHITECTURE**

The Boston Bus database management system will be hosted in the cloud like AWS RDS for the following benefits:

* It supports different databases like MySQL, Oracle, etc.,
* It creates several instances to give high availability.
* It has a pay-as-you-go model.
* It offers horizontal & vertical scaling.

**LESSONS LEARNED & FUTURE CONSIDERATIONS**

**Lessons Learned**

Working on this management system has stretched my thinking and capabilities. I have gained knowledge in the following areas:

* 1. **Domain Knowledge**: I gained theoretical and practical knowledge on database management systems, like the importance of database systems, entities, and their different relationships, how to create an ERD diagram, and how to navigate SQLite for the creation and querying of databases.
  2. **Design Thinking**: in order to create the database for Boston Bus, I had to deeply think about what the database should be like, the various tables/entities, and the kind of relationships the entities will have with each other. Design thinking enabled me to proffer the appropriate database system for Boston Bus.
  3. **Business Knowledge**: creating a database management system is always for a business, in this case, a transportation business. Hence, I had to research the transportation business and industry in the United States of America for the kind of personas, entities, and relationships the entities would have in the database management system.

**Future Considerations**

Going forward, I need to increase my skills and knowledge of database management systems as there is still much to be covered. I will research the possible tables a transportation database can have, learn how to add more entities to the system, and how to query multiple tables.

**REFERENCES**

* Smallcombe, M. (n.d.). Amazon RDS: What Is It and How Does It Work? Retrieved August 21, 2022, from Integrate.io website: <https://www.integrate.io/blog/amazon-rds-what-is-it-and-how-does-it-work/#architecture>
* What is Network Infrastructure Security? | VMware Glossary. (2022, January 12). Retrieved from VMware website: <https://www.vmware.com/topics/glossary/content/network-infrastructure-security.html>
* Pena, B. de la. (2019, December 13). Traveler personas and their information needs. Retrieved August 21, 2022, from Medium website: <https://medium.com/@benjiedlp/traveler-personas-and-their-information-needs-92a0131e063a>
* 5 Different Travel Personas Every Hotelier Interacts With. (n.d.). Retrieved from Hospitality Net website: <https://www.hospitalitynet.org/news/4103515.html>