Yelp Business Data Analysis

COP 5725

Spring 2019
Instructor: Dr. Markus Schneider

Project Phase - II

Group 24

Group Members:

Name:	UFID:	Email:
● Dashuai Qin	91493498	dqin@ufl.edu
● Shubham Kathuria	33939680	skathuria@ufl.edu
● Vikas Pathak	71984931	pathak.vikas@ufl.edu
● Pranav Mokal	68121781	mokalpranav@ufl.edu

INTRODUCTION:

The yelp business dataset has detailed information such as location, category, ratings, reviews, price, hours, and other attributes about various businesses. Using this database, we can recommend various businesses such as restaurants to users depending on their needs, i.e. restaurants good for late night dining, restaurants good for the family dinners, or even type of restaurants that be a good for starting as a business etc. Users can also use the data to do trip planning since we have queries like the total number of different category of restaurants and an average rating of restaurants. Even though our focus is mainly on the restaurants, we will also provide the same features to users in terms of hotels, shopping businesses etc.

OVERVIEW:

In the second phase of our project, we are trying to provide a brief idea about how the User Interface is going to look like and how it will work and also about the Conceptual Database Design. We are providing the entire flow of the User Interface which includes the linking between various types of web pages that are a part of the Interface and also a description regarding the working of each of these Web Pages. For the conceptual database design, we are providing an Entity-Relationship Diagram of the database and also describing the relationships between all the entities and also some of their attributes.

USER INTERFACE DESIGN:

We are trying to implement a User Interface that is easy to navigate by the users and it does not create any confusion. The Website will provide all the necessary information (both simple and complex) to the users regarding a particular business as well as cumulative data pertaining to various businesses that fit into a certain category. The interface allows users to view only the simple basic data at one time and they wish to dive in deeper, they can check out the complex analytical data too.

The User-Interface Design Diagram is provided below:

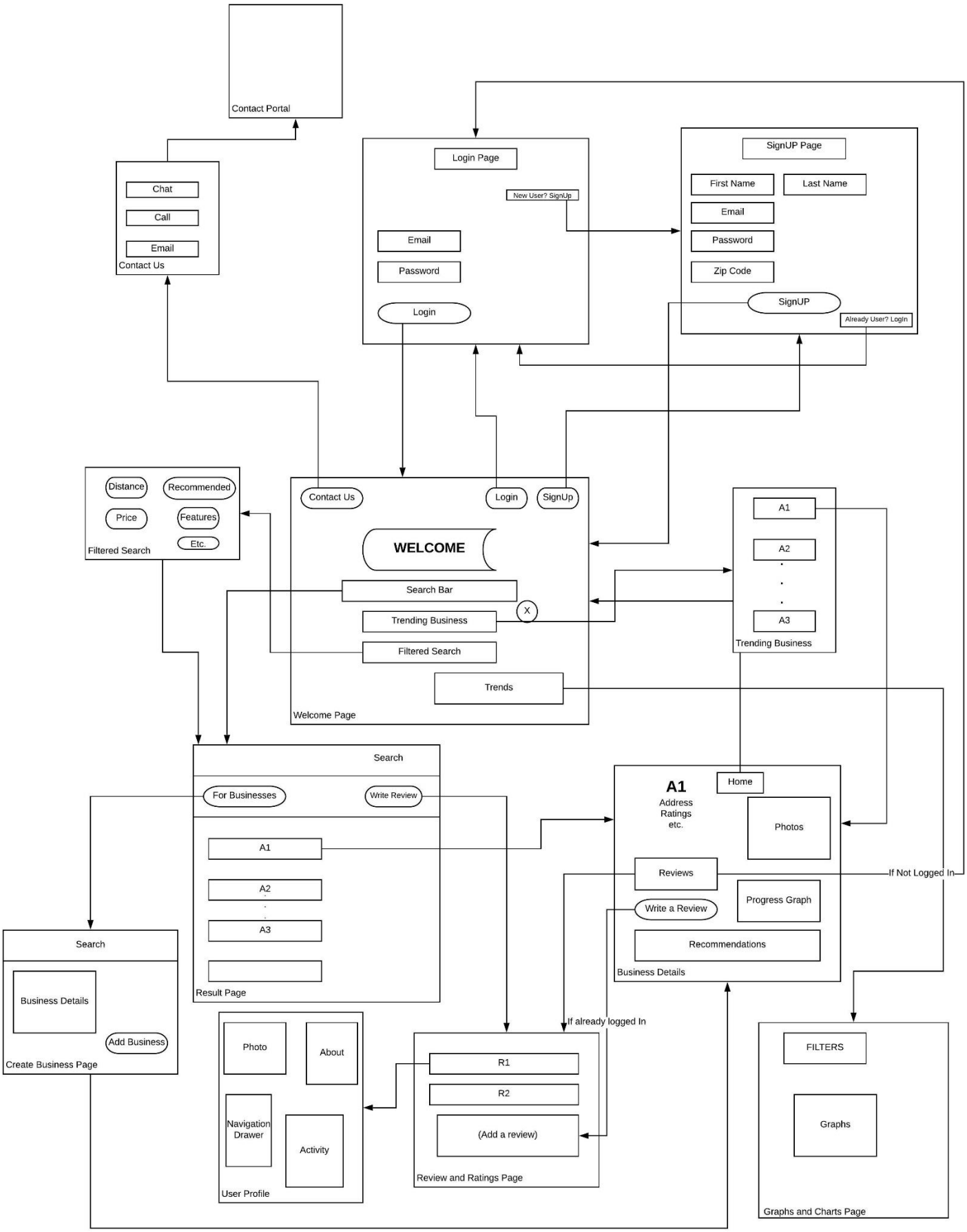

Contact Portal
Chat
Call
Email
Contact Us
Login Page
New User? SignUp
Email
Password
Login
SignUP Page
First Name
Last Name
Email
Password
Zip Code
SignUP
Already User? LogIn
Distance
Recommended
Price
Features
Etc.
Filtered Search
Contact Us
Login
SignUp
WELCOME
Search Bar
Trending Business
Filtered Search
Trends
X
Welcome Page
A1
A2
A3
Trending Business
Search
For Businesses
Write Review
A1
A2
A3
Result Page
A1
Address
Ratings
etc.
Home
Photos
Reviews
Write a Review
Progress Graph
Recommendations
Business Details
If Not Logged In
Search
Business Details
Add Business
Create Business Page
Photo
About
Navigation Drawer
Activity
User Profile
If already logged In
R1
R2
(Add a review)
Review and Ratings Page
FILTERS
Graphs
Graphs and Charts Page

Figure 1

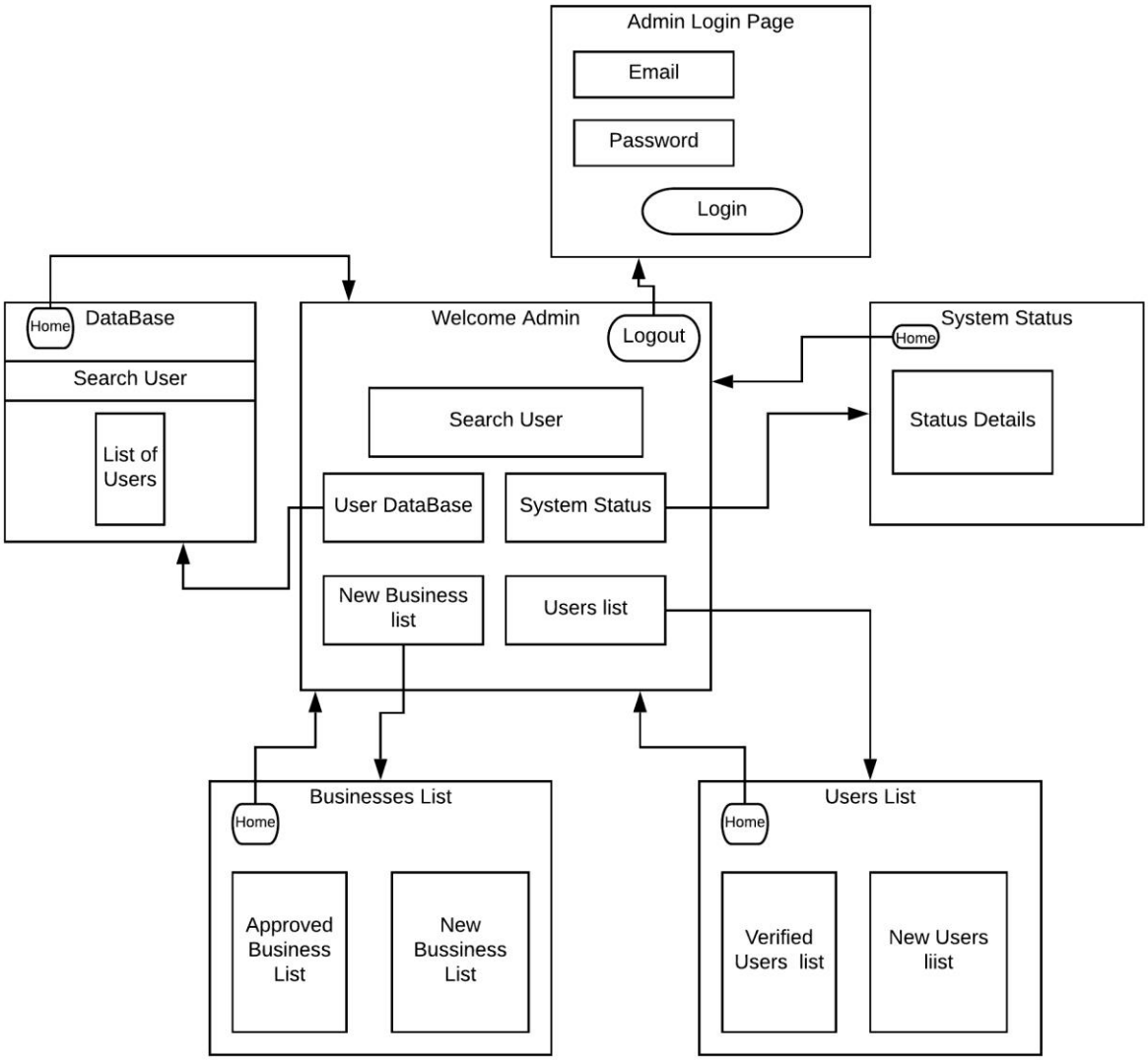


Figure 2

User-Interface Description:

Figure 1 demonstrates the part of the Interface that is accessible to the Users of our Application whereas Figure 2 represents the part of the Interface that accessible to the Admins/ Managers of this Database Management System.

- The Application starts at the **HOME** or the **WELCOME Page** from where the user can navigate to various other pages. The HOME page enables the users to perform various actions like Login, Sign Up, Search etc.

- The **Login Page** allows the user to Sign In so that the other actions become easier and a certain number of privileges are given to the user. The Login Page can also direct the user to the Sign Up Page if the user does not have an account already present.

- The **Sign Up Page** allows the user to create an where various details are asked to fill in so that the user can Login in the future.

- The **Contact Us Page** and the **Contact Portal** are for the users to post complaints, inquiries or issues with the Website.

- The **Filtered Search Page** is for users who want to dive deep into the website and search for things pertaining to their specific needs. It is also used to address certain complex queries.

- The **Results Page** is displayed when the user searches for something in the Search Bar of the Home Page or in the Filtered Search page. It displays all the results/list of the respective search query.

- The **Business Details Page** is displayed when a user selected any particular item from the list in the Results Page. It displays almost all the details regarding a particular business and also enables users to write a review.

- The **Review and Ratings Page** is displayed when the user wants to read the reviews from other users or post a review. This page can also direct the user to the User Profile Page of the reviewer in order to check for authenticity of the review.

- The **User Profile Page** displays the history of reviews posted by that user.

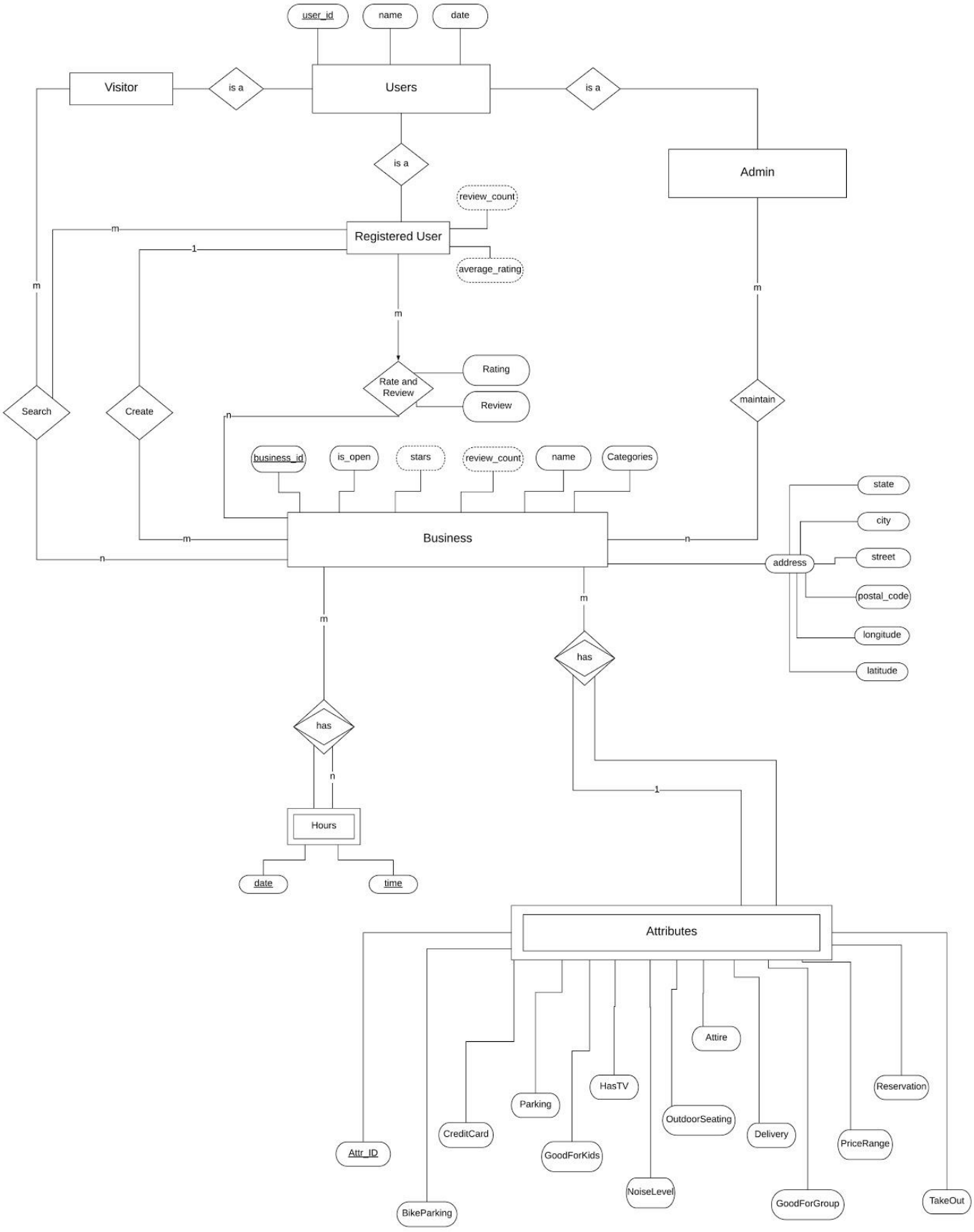
- The user can also register their business through the **Create Business Page.** The Business is first approved by the Administrator and then entered into the database.

- The **Trending Business Page** displays the list of businesses who have received the highest average rating in a particular category.

- The **Graphs and Charts Page** provide data regarding various trends to the user so the user can analyze the data such as ratings over the months to come up certain conclusions regarding the business and come up with certain ideas.

- The Administrator Side of the Interface allows the admins to verify/delete the suspicious reviews, approve business requests etc.

CONCEPTUAL DATABASE DESIGN:
Entity Relationship Diagram Design:



ER Diagram Design Motivation:
We designed this model based on the layout of data in the original datasets. Also, based on the types of users interacting with the data, it could be pure searching, creating their own business, rating and reviewing any business, or maintaining the data as admin.

Important concepts of ER Diagram:
ER Model is used to model the logical view of the system from data perspective.
Following are some important aspects of the ER Diagram:

- The above ER Diagram has six Entities: Four strong entities and Two weak entities.
  Strong Entities are **Users**, **Visitor, Admin and Business.**
  Weak Entities are **Hours** and **Attributes**.
  This is because 'Hours' and 'Attributes' entities are dependent on Business Entity.

- **Attributes:** The above ER Diagram has three type of attributes.
  - **Derived Attribute –** There are total four derived attribute in the ER Diagram.All are derived from rating and review.
    They are : average_rating and review_count for the registered users entity, and stars and review_count for business entity.
  - **Composite Attribute –** The address is a composite attribute in the ER because it is composed of many other attributes like city state etc.
  - **Key Attributes –** The user_id and bussiness_id are the key attributes which uniquely identifies users and business entities in entity set respectively. Likewise date and time are key attributes of hours entity and Attr_Id is key attribute of attributes entity. Although these are partial keys for this ER because attributes and hours entities are weak entities.

- **Participation Constraints:** In the ER Diagram 'hours' and 'attributes' entities has total participation with the business entity because they are weak entities. The participation of weak entity type is always total.

- **Cardinalities:**

  - **(Registered User, Business): (m, n)**
    One user can rate and review multiple businesses and create multiple business. One business can be rated and reviewed by many users and created by only one registered user.

  - **(Visitor, Business): (m, n)**
    One visitor can search many business and one business can be searched by many visitors.

  - **(Admin, Business): (m, n)**
    One admin can maintain many businesses and one business can be maintained by many admins

  - **(Attributes, Business): (1, m)**
    Each business has one attributes and each attributes can belong to multiple business.

  - **(Business, Hours): (m, n)**
    Each business has multiple opening hours and each opening hour can belong to multiple business.

References:

1. www.lucidchart.com for ER diagram and UI design diagram.