



DALHOUSIE UNIVERSITY

CSCI-5410 SERVERLESS DATA PROCESSING

Project Design Document

Group 24

Benny Tharigopala	B00899629
Jaswanth Mandava	B00874871
Kandarp Parikh	B00873863
Prit Thakkar	B00890731
Ruchi Shinde	B00897784
Viren Malavia	B00895669

Course Instructor: Dr. Saurabh Dey

GitLab URL: [csci5410_Group24](#)

Table of Contents

1. Cloud Architecture	3
2. Cloud Services	4
2.1 User Management	4
2.2 Authentication	5
2.3 Online Support	5
2.4 Message Passing.....	5
2.5 Machine Learning	6
2.6 Web Application and Hosting	6
2.7 Report Generation	7
2.8 Visualization	7
2.9 Testing	7
3. User Flow.....	8
4. Task Distribution.....	11
5. Meeting Logs	11
6. References	12

1. Revised Cloud Architecture

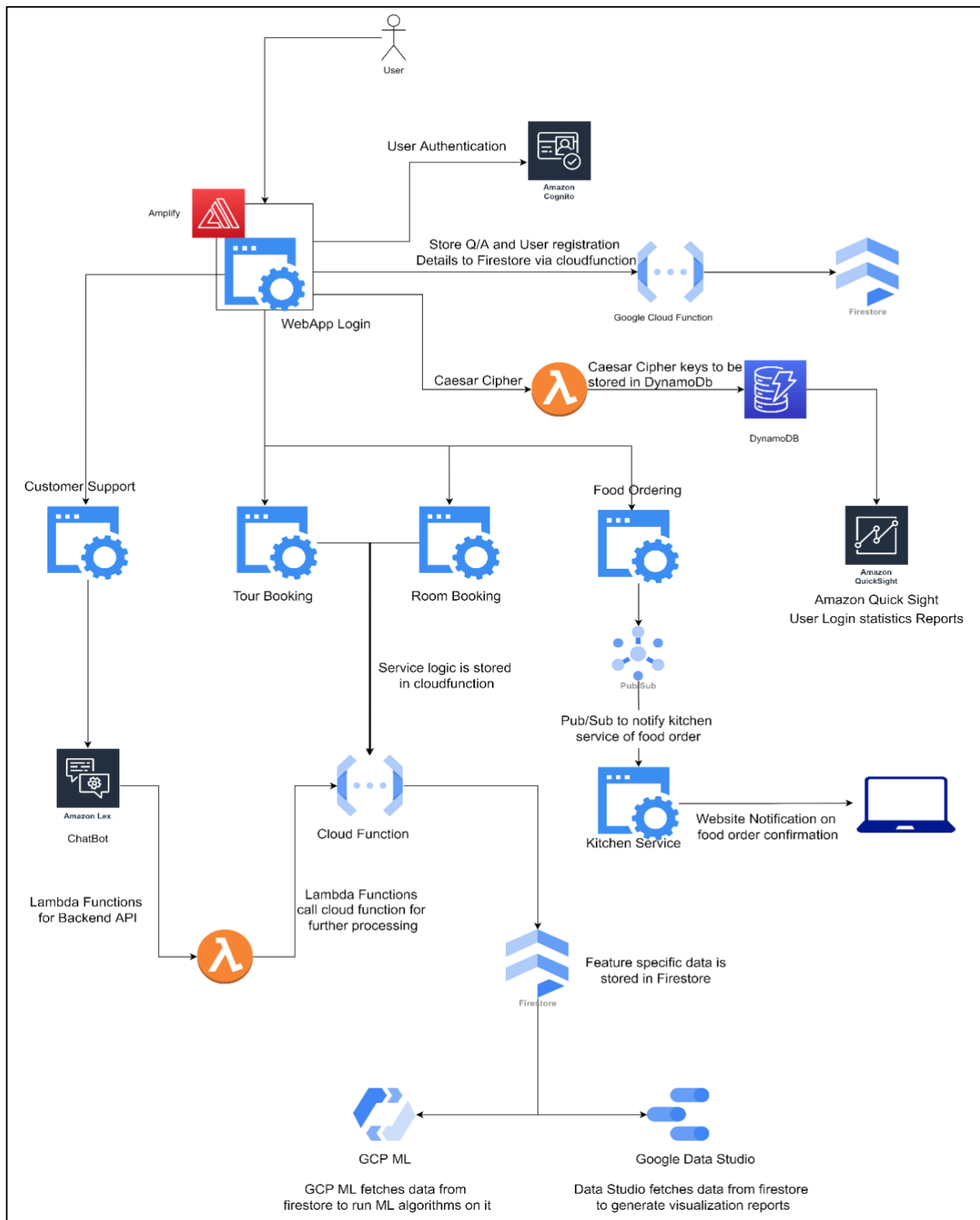


Figure 1: Cloud Architecture of our Application, diagram made with the use of draw.io [7]

Cloud Architecture diagram description:

The revised cloud architecture in figure 1, shows the interconnection and interoperability of cloud services of Google cloud and AWS. For developing our serverless “Bed and Breakfast” application our team will be leveraging multiple services of GCP and AWS for various features in our project.

2. Cloud Services

Module Name	Cloud Service to be Used
User Management - Sign up validation	Amazon Cognito
User Management - Managing and storing User details	Cloud Function and Firestore
Authentication - ID-password	AWS Cognito
Authentication - Question Answer	Firestore + CloudFunction
Authentication – Caesar Cipher	Lambda + DynamoDB
Online Support	AWS Lex
Message Passing	Pub/Sub
Machine Learning	GCP AutoML
Web Application Building and hosting	AWS Amplify
Report Generation Module	AWS QuickSight
Visualization Module	Google Data Studio

2.1 User Management

For User management, Amazon Cognito [1] would be used for user authentication and validation whereas combination of Cloud Function and Firestore would be used to store user specific data such as First name, Last name, Phone number, Email, and other user related details

At the time of registration, the user will be provided with a random key (1-26) which would be used for Caesar cipher during login and authentication.

2.2 Authentication

Authentication consists of three steps:

1. ID-password authentication will be done by Amazon Cognito.
2. For verification of security questions and answers, Google cloud function and firestore will be used.
3. Caesar Cipher: Once the above steps are completed, System will generate a plain text and ask the user to convert it into cipher text using the key they received at the time of registration. Once the user enters the cipher text, AWS Lambda function will compare the cipher text to the generated cipher text, if both the cipher texts match, Login would be successful.

Once the user login is successful, a login timestamp along with User-Id will be stored in DynamoDB which will further be used by quick sight to generate reports about user statistics.

2.3 Online Support

The online support module is a virtual assistance or a chatbot which can help us to get answers to our queries easily and quickly. We will be using the AWS Lex [3] service to develop a chatbot. Lex is powered by the same technology as Alexa, it can support both text and voice features assistance. In our project, we will be implementing only textual assistance. It offers features like intent identification, it can extract entities like date, time, city, state, etc. from the text which is called slots. Training a chatbot using Lex is very easy. We have to add only a few utterances which would be the anticipated inputs entered by the user. In our BnB application, the chatbot will be accessible to logged in users as well as guest users. Logged in users can book rooms, order their meals, and perform other such transactions. The guest users, however, can access only generic information about the BnB services. If a guest user is trying to book a room for instance, they will be able to request the prices and availability of a room, but the chatbot will take them to the login page in order to proceed with the transaction.

2.4 Message Passing

The message passing module encompasses GCP (Google Cloud Platform) Pub/Sub [5] , which is an event-driven publisher subscriber asynchronous queue used for providing notifications to the end user.

Considering the food orders feature, when an authenticated user will place an order requesting some food item, it will trigger a cloud function with a request object of order details which will be later processed and stored in the Pub/Sub queue. As soon as Pub/Sub receives the message, it will trigger the Kitchen Service (another cloud function) which is responsible for preparing food. There will be a Firestore document having notifications from the hotel and kitchen services.

When the first cloud function (Hotel Management) receives an order, it will reflect as 'PENDING' in the notifications section, as soon as the second cloud function (Kitchen Service) picks an order for processing, the notification will be changed to 'APPROVED' and after a successful execution of the kitchen service, the notifications area would display it as 'READY.'

2.5 Machine Learning

In this module, we will be recommending tour packages based on the similarity of previous stay durations. Besides this, we also need to assign relevant scores based on the polarity of customer feedback. To perform the operations of this module, we will be utilizing the services provided by GCP AutoML [4]. GCP AutoML is a service provided by Google Cloud Platform for developers who are novice in the field of machine learning. This service enables developers to train, build and deploy high-quality custom machine learning models that can analyze documents, categorize them, identify entities, or assess attitudes [4].

As we are supposed to recommend tour packages to customers, we will be using the past data stored in firestore as input. The existing data will then be ingested into Google's machine learning service called GCP AutoML that will process, analyze, and recommend appropriate tour packages based on the set of business rules we choose at the time of configuring the model. In the second part of the module where we need to assign appropriate scores based on the polarity of customer feedback, we will take customer's feedback stored in firestore as input and Vertex AI AutoML service will inspect the provided input to identify the sentiments (positive, negative, or neutral) and assign a relevant score.

2.6 Web Application and Hosting

Our team has decided to host the application on AWS Amplify, amplify is a serverless framework for frontend developers; it offers frontend libraries for JavaScript, iOS, Android, and React Native and a CLI that helps to create serverless backend services for different use cases [6]. We would be setting up a pipeline from our GitLab repository such that whenever a code is pushed to master branch, amplify would fetch the latest code and build the app. The frontend of our application would be developed with ReactJS which in turn would call lambda functions and cloud functions for the backend functionalities of the application.

The CI/CD pipeline that AWS Amplify integrates with your chosen Git repository is one of the main advantages it offers. Without further manual intervention, updates to your code base that are produced and pushed to the master branch will be automatically deployed to AWS Amplify [7].

2.7 Report Generation

The Report Generation module, in our application, will facilitate comprehension of data related to user interaction and access statistics. The module will utilize Amazon QuickSight [2] to summarize user statistics.

Activities in the module include, collection of data generated from interactions between the user and the application, uploads of Data to QuickSight [2] and analysis of user access statistics.

We plan to embed interactive dashboards within the application so that business stakeholders can access vital information in an accessible manner. With accurate information and derivable insights stakeholders can perform anomaly detection and even apply forecasting to make predictions and tailor a bespoke application to their target customer demographics.

2.8 Visualizations

The visualization module is responsible for rendering charts which represent various forms of data in intuitive formats. Stakeholders can utilize the visualizations to extract easy-to-understand insights and perform informed business decision-making.

Data for various requirements will be collected from our storage system and uploaded to Google Data Studio [9]. These datasets will then subsequently be used to populate charts in Google Data Studio. Stakeholders can use these charts to analyze business data in a visually appealing manner. For instance, business owners can analyze room booking trends, tour request trends, popular orders from the restaurants and act according to the requirements of consumers.

We will set up a pipeline such that the most recent data is always available in Google Data Studio based on which the users of the charts can make accurate decisions.

2.9 Testing

The testing module is used to test all the functionalities of our application. We will be doing two kinds of testing here. One is functional testing using code. Here, we will use lambda functions to test the functionality of other lambda functions used on the project. Other than this we will do some manual Testing to check the workflows overall.

3. User Flow

- **Registration:** Figure 2 represents the path taken by users to register an account with the application.

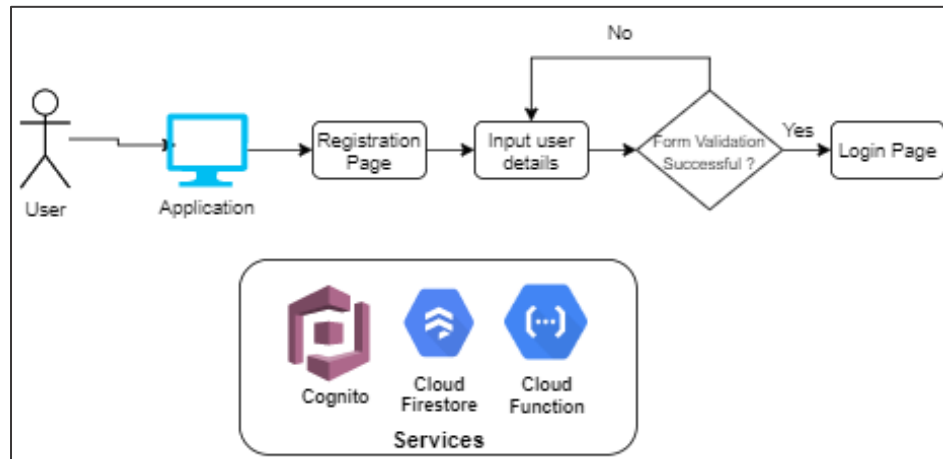


Figure 2: User Flow Diagram for Registration

- **Login:** Figure 3 represents the path taken by users to login to the application.

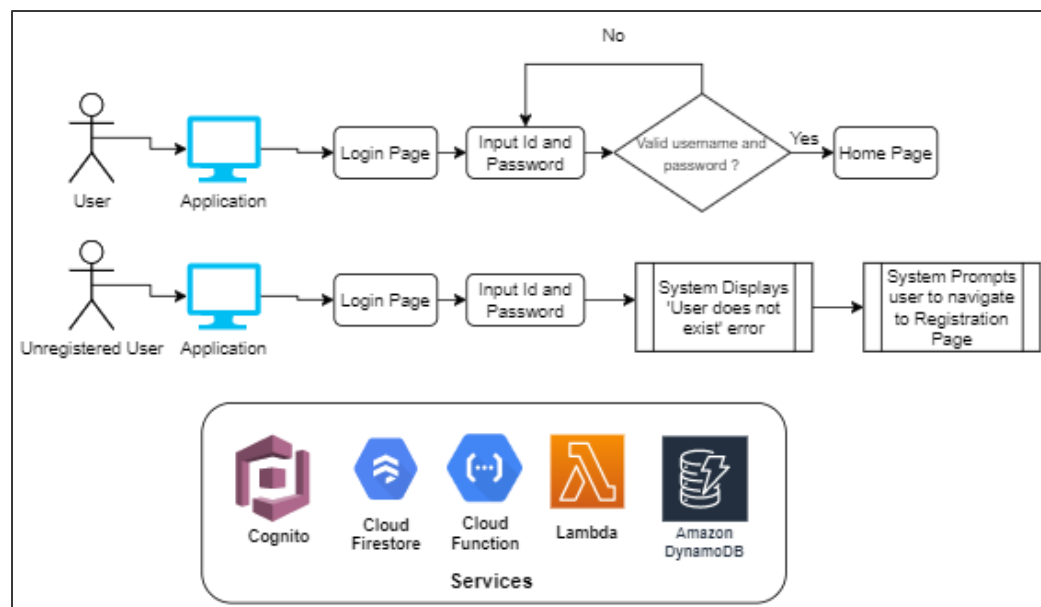


Figure 3: User Flow Diagram for Login

- **Online Support Feature:** Figure 4 represents the path taken by users to submit queries or instructions to avail services, to the chatbot. Note that only registered users can request services through the chatbot whereas unregistered users can only ask questions.

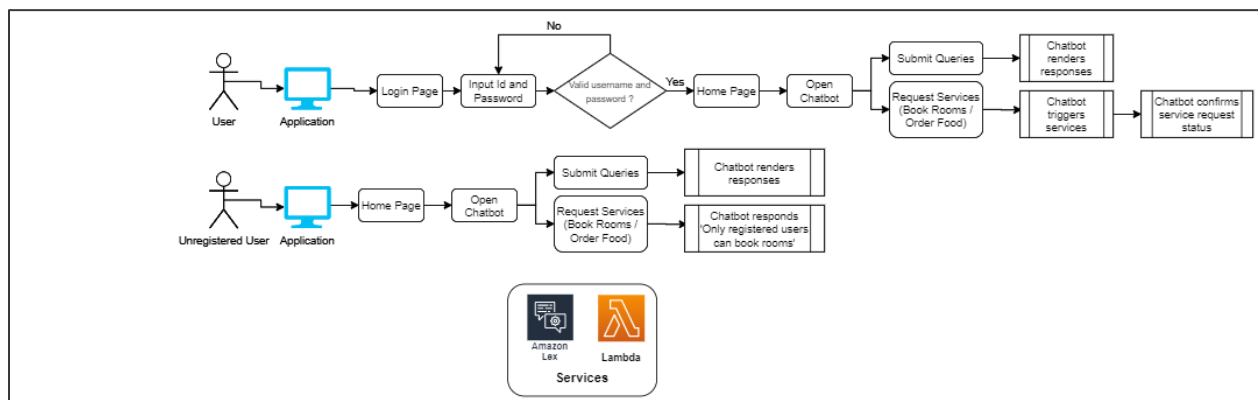


Figure 4: User Flow Diagram for Online Support

- Notification Feature:** Figure 5 represents the path taken by users to avail services such as food and tours. Note that only registered users can request services whereas unregistered users are prompted to login before availing the features offered by the application.

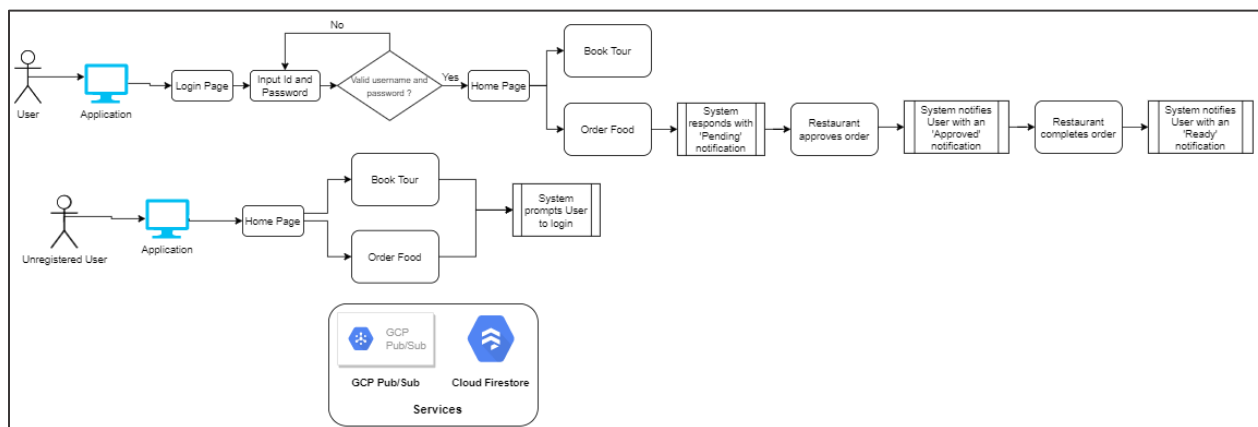


Figure 5: User Flow Diagram for Services and Notifications

- Machine Learning Feature:** Figure 6 represents the path taken by administrators to predict and propose appropriate tour packages to users and predict the polarity of customers' feedback.

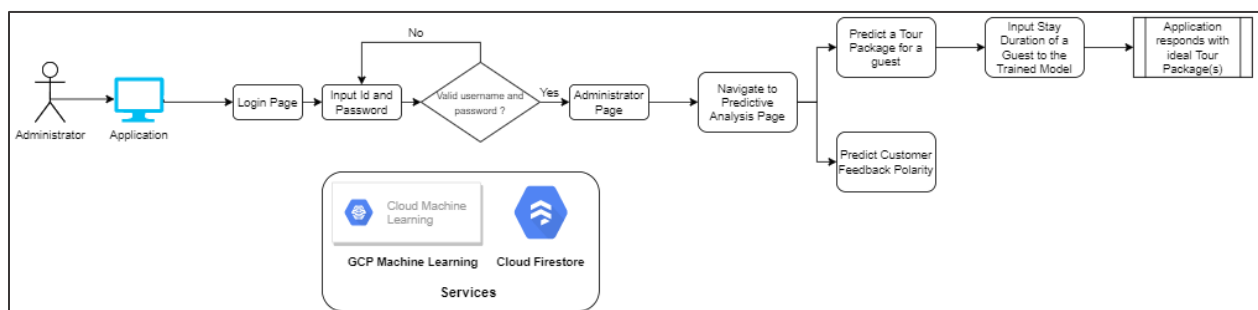


Figure 6: User Flow Diagram for Machine Learning and analysis

- **Report Generation:** Figure 7 represents the path taken by administrators to analyze reports on user statistics.

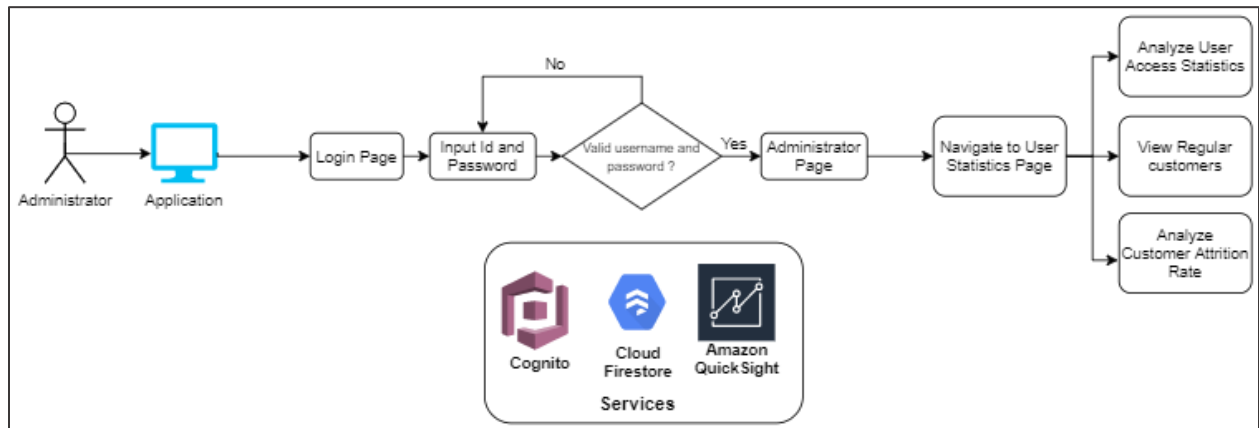


Figure 7: User Flow Diagram for Report Generation

- **Visual Analytics:** Figure 8 represents the path taken by administrators to analyze charts relevant to profits, losses, room booking trends and other topics of interest, to derive actionable insights.

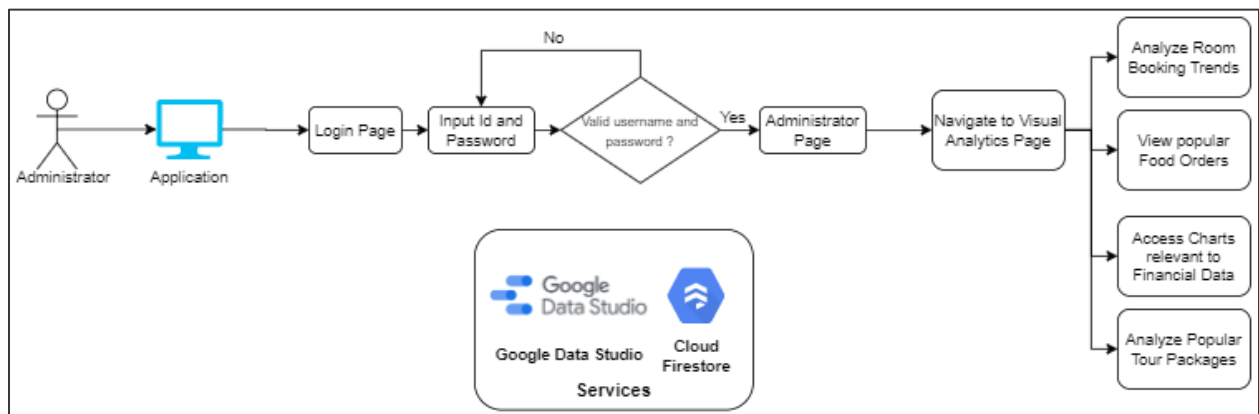


Figure 8: User Flow Diagram for Visual Analytics

4. Task Distribution

Team Member Name	Assigned Module
Kandarp Parikh	User Management, Authentication, Report Generation
Prit Thakkar	Food Ordering, Message Passing
Ruchi Shinde	Online Support module
Benny Tharigopala	Room booking, Visualization
Viren Malavia	Toor booking, Machine learning, Feedback
Jaswanth Mandava	Toor booking, Machine learning, Feedback
Team Effort	Web application Hosting and Documentation

5. Meeting Logs

Date	Place	Attendees	Discussion Agenda	Duration
18/5/22	MS Teams	Benny, Kandarp, Ruchi, Viren, Prit	Understanding the requirements, cloud services and serverless architecture	27m
28/5/22	MS Teams	Benny, Jaswant, Kandarp, Ruchi, Viren, Prit	Discussion on Cloud services	30m
29/5/22	Wallace McCain Learning Commons	Jaswant, Kandarp, Ruchi, Viren, Prit	Database Schema, BnB features and assignment of features to devs, conceptual document task division	1h20m
31/5/22	MS Teams	Benny, Jaswant, Kandarp, Ruchi, Viren, Prit	Finalized the conceptual document contents and formatting the doc	40m
30/6/22	MS Teams	Benny, Jaswant, Kandarp, Ruchi, Viren, Prit	Discussed on the conceptual report feedback and worked o	30m
3/7/2022	MS Teams	Benny, Kandarp, Ruchi, Prit	Clarified doubts and reviewed the architectural diagram.	48m

Figure 9 Meeting Logs

6. References

- [1] "Amazon Cognito - Simple and Secure User Sign Up & Sign In | Amazon Web Services (AWS)," *Amazon Web Services, Inc.*, 2022. [Online]. Available: <https://aws.amazon.com/cognito/>. [Accessed: Jul. 05, 2022]
- [2] "Amazon QuickSight - Business Intelligence Service - Amazon Web Services", Amazon Web Services, Inc., 2022. [Online]. Available: <https://aws.amazon.com/quicksight/> . [Accessed: 04- Jul- 2022].
- [3] "Conversational AI and Chatbots - Amazon Lex - Amazon Web Services", Amazon Web Services, Inc., 2022. [Online]. Available: <https://aws.amazon.com/lex/> . [Accessed: 04-Jul- 2022].
- [4] "Features and capabilities of AutoML Natural Language | AutoML Natural Language Documentation | Google Cloud," Google [Online].
Available:<https://cloud.google.com/natural-language/automl/docs/features>. [Accessed: 04-Jul-2022]
- [5] "What is Pub/Sub?," *Google Cloud*. [Online]. Available:
<https://cloud.google.com/pubsub/docs/overview> . [Accessed: 05-Jul-2022].
- [6] Kay, "Serverless Authentication with AWS Amplify: A Practical Guide - IOD - The Content Engineers," *IOD - The Content Engineers*, Apr. 30, 2020. [Online]. Available: <https://iamondemand.com/blog/serverless-authentication-with-aws-amplify-a-practical-guide/#:~:text=Amplify%20is%20a%20serverless%20framework,services%20for%20different%20use%20cases..> [Accessed: Jul. 05, 2022]
- [7] C. McCrory, "Hosting a serverless web app using AWS Amplify - Towards AWS," *Medium*, Jan. 30, 2022. [Online]. Available: <https://towardsaws.com/hosting-a-serverless-web-app-using-aws-amplify-3f8c4ffd5a90>. [Accessed: Jul. 05, 2022]
- [8] "Flowchart Maker & Online Diagram Software," *Diagrams.net*, 2022. [Online].
Available: <https://app.diagrams.net/>. [Accessed: Jul. 05, 2022]
- [9] "Google Data Studio," *Google.com*, 2014. [Online]. Available: <https://datastudio.google.com/u/0/>. [Accessed: Jul. 05, 2022]