INSTAGRAM DATABASE

Rdbms Minor Project Report

SUBMITTED IN PARTIAL FULFILLMENT REQUIREMENT FOR

THE AWARD OF DEGREE OF

Bachelor of Technology

(Computer Science & Engineering)

SUBMITTED BY

Rahul

(UNIVERSITY ROLL No. 1606745)

December, 2018

I.K.G. PUNJAB TECHNICAL UNIVERSITY

JALANDHAR (PUNJAB), INDIA

GURU NANAK DEV ENGINEERING COLLEGE

CANDIDATE'S DECLARATION

I hereby certify that I have under 5th semester in Rdbms Lab and worked on project entitled, "**INSTAGRAM DATABASE**", in partial fulfillment of requirements for the award of Degree of **Bachelor of Technology** in Department of Computer Science & Engineering at **GURU NANAK DEV ENGINEERING COLLEGE** under **I.K.G. PUNJAB TECHNICAL UNIVERSITY, JALANDHAR**, having University Roll No. 1606745, is an authentic record of my own work carried out during a period from August, 2018 to November, 2018.

(**Rahul)**

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Geetika and Nidhi

(Lab Co-ordinator)

ABSTRACT

This is a project report on Instagram Database. During the development of this project we explored new ideas and functionalities behind the working of this Project. This project is the output of our planning, schedule, skill over various Languages and the hard work and this report reflects our steps taken at various levels of planning ,schedule and skill over various Languages. We have learnt a lot during this project and liked the improvement in our testing skills and deep concept related to these kinds of projects.

ACKNOWLEDGEMENT

I am highly grateful to **Dr. Sehijpal Singh**, Principal, Guru Nanak Dev Engineering College, Ludhiana, for providing this opportunity from August-November 2018.

Mrs.Geetika and Mrs.Nidhi (Rdbms Lab Co-ordinator) has provided great help in carrying out the my work and is acknowledged with reverential thanks. Without the wise counsel and able guidance, it would have been impossible to complete the Project in this manner.

I would like to express thanks profusely to **Dr. Parminder Singh**, Professor and Head (Computer Science & Engineering) for stimulating me time to time. I would also like to thank to entire faculty, staff of computer science and engineering. I also thanks my friends who devoted their valuable time and helped me in all possible ways towards successful completion of this work.

(Rahul)

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
GNDEC	Guru Nanak Dev Engineering College
RDBMS	Relational Database Management System
E-R	Entity Relationalship
SQL	Structured Query Language
FSS	File System Storage

TABLE OF CONTENTS

Chapter 1

Introduction to Project

1.1 Overview

This project sequentially applies a set of Mysql techniques to gain insights from the Instagram Database.Mysql Environoment analysis of this data will benefit the business processes of the Instagram.

This project deals with designing and implementing a system for handling the information of behavioral evaluations. An Analyst schedules subject evaluations and then analyses the recorded behaviors that occur during specified collection periods. The evaluations provide data that can be analyzed in order to develop plans that will help treat the subject as needed. The collection periods or appointments as they have been called in our project are scheduled by the analyst, conducted by the therapist and data during these appointments is collected by a collector present. Our system also implements an admin user who is required for user management and behavior data management.

1.2 Exitsing System

As day by day, the data used increases and therefore a better way of handling such a huge amount of data is becoming a hectic task. The traditional approach of data storage File System Storage.FFS is for unstructured data as well as structured data.It stores the average size of data.

When a size of data is too big for complex processing and storing or not easy to define the relationships between the data, then it becomes difficult to save the extracted

information in an FSS with a coherent relationship.

By the above comparison, we have come to know that MYSQL is the best technique for handling Big Data compared to that of FSS.

1.3 Functional Requirements

- Setup of MYSQL Environment.
- PC needs to have atleast 4 GB Ram and atleast 100GB of external usable memory.
- PC should have Ubuntu 16.4v or Windows 7 or above.

1.4 Feasibility Study

A feasibility study is used to determine the viability of an idea, such as ensuring a project is legally and technically feasible as well as economically justifiable. It tells us whether a project is worth the investmentin some cases, a project may not be doable. There can be many reasons for this, including requiring too many resources, which not only prevents those resources from performing other tasks but also may cost more than an organization would earn back by taking on a project that isnt profitable.

The application is fully feasible. It just needs a working internet connection,PC have atleast 4 GB RAM. It is fully feasible if it is also deployed on a large scale.
The Project can also be upgraded further and can be deployed on large scale depending upon the need of business plan.

1. **Technical Feasibility :** The Project is fully feasible on technical terms. I have PC and required 4 GB Ram for development purpose.

I will use MYSQL Environment to deploy backend as it is frees.

The version control system is completely free and the website Github.com is also free for Open Source Projects.

2. **Economic Feasibility :** The Project is fully economically feasible as it has free and open source tools being used while developing the system.

3. **Legal Feasibility :** The Project doesn't violates any legal rights and will credit the author of open Source Library used while developing the project.

   The project will be available in open source under **GPLv3** license.

   What is GPLv3 License?

   - The source code must be made public whenever a distribution of the software is made.
   - Modifications of the software must be released under the same license.
   - Changes made to the source code must be documented.
   - If patented material was used in the creation of the Project, it grants the right for users to use it. If the user sues anyone over the use of the patented material, they lose the right to use the Project.

4. **Operational Feasibility :** As the Project satisfies the functional and non functional requirements, the Project will be fully operational once it releases.

5. **Scheduling Feasibility :** The project release targets for different versions are practical and have plenty of time develop and debug the Project before release.

1.5 Objectives of the Project

The main objective of this project is given below-

- finding 5 oldest users.
- what day of the weeks do most users register on?
- find the users who have never posted any photos.
- find most likes on single photo.
- HOW many times does the average user post ?
- What are the top 5 most commanly used hashtags ?

Chapter 2

Product Design

2.1 Product Perspective

This Project utilizes Data Classification to examine a dataset related with Instagram Database.Data Classification is the use of MYSQL techniques to organize datasets into related sub-populations, not previous specified in the dataset. This can uncover hidden characteristics within data, and identify hidden categories that new data belongs within.

2.2 Table Structure

The dataset examined by this Project was collected from a Instagram Databases.I use MYSQL Database for this project. Here we describe the all table-

users

```
mysql> desc users;
+------------+--------------+------+-----+-------------------+----------------+
| Field      | Type         | Null | Key | Default           | Extra          |
+------------+--------------+------+-----+-------------------+----------------+
| id         | int(11)      | NO   | PRI | NULL              | auto_increment |
| username   | varchar(255) | NO   | UNI | NULL              |                |
| created_at | timestamp    | NO   |     | CURRENT_TIMESTAMP |                |
+------------+--------------+------+-----+-------------------+----------------+
3 rows in set (0.06 sec)
```

FIGURE 2.1: users table

follows

```
mysql> desc follows;
+-------------+-----------+------+-----+-------------------+-------+
| Field       | Type      | Null | Key | Default           | Extra |
+-------------+-----------+------+-----+-------------------+-------+
| follower_id | int(11)   | NO   | PRI | NULL              |       |
| followee_id | int(11)   | NO   | PRI | NULL              |       |
| created_at  | timestamp | NO   |     | CURRENT_TIMESTAMP |       |
+-------------+-----------+------+-----+-------------------+-------+
3 rows in set (0.03 sec)

mysql>
```

FIGURE 2.2: follows table

likes

```
ysql> desc likes;
+------------+-----------+------+-----+-------------------+-------+
| Field      | Type      | Null | Key | Default           | Extra |
+------------+-----------+------+-----+-------------------+-------+
| user_id    | int(11)   | NO   | PRI | NULL              |       |
| photo_id   | int(11)   | NO   | PRI | NULL              |       |
| created_id | timestamp | NO   |     | CURRENT_TIMESTAMP |       |
+------------+-----------+------+-----+-------------------+-------+
rows in set (0.01 sec)

ysql>
```

FIGURE 2.3: likes table

comments

```
mysql> desc comments;
+--------------+--------------+------+-----+-------------------+----------------+
| Field        | Type         | Null | Key | Default           | Extra          |
+--------------+--------------+------+-----+-------------------+----------------+
| id           | int(11)      | NO   | PRI | NULL              | auto_increment |
| comment_text | varchar(255) | NO   |     | NULL              |                |
| user_id      | int(11)      | NO   | MUL | NULL              |                |
| photo_id     | int(11)      | NO   | MUL | NULL              |                |
| created_at   | timestamp    | NO   |     | CURRENT_TIMESTAMP |                |
+--------------+--------------+------+-----+-------------------+----------------+
5 rows in set (0.03 sec)

mysql>
```

FIGURE 2.4: comments table

photos

```
MySQL 5.6 Command Line Client
mysql> desc photos;
+------------+--------------+------+-----+-------------------+----------------+
| Field      | Type         | Null | Key | Default           | Extra          |
+------------+--------------+------+-----+-------------------+----------------+
| id         | int(11)      | NO   | PRI | NULL              | auto_increment |
| image_url  | varchar(255) | NO   | UNI | NULL              |                |
| user_id    | int(11)      | NO   | MUL | NULL              |                |
| created_at | timestamp    | NO   |     | CURRENT_TIMESTAMP |                |
+------------+--------------+------+-----+-------------------+----------------+
4 rows in set (0.03 sec)

mysql>
```

FIGURE 2.5: photos table

photos-tags

tags

```
mysql> desc photo_tags;
+----------+---------+------+-----+---------+-------+
| Field    | Type    | Null | Key | Default | Extra |
+----------+---------+------+-----+---------+-------+
| photo_id | int(11) | NO   | PRI | NULL    |       |
| tag_id   | int(11) | NO   | PRI | NULL    |       |
+----------+---------+------+-----+---------+-------+
2 rows in set (0.13 sec)

mysql>
```

FIGURE 2.6: photo tags table

```
mysql> desc tags;
+------------+--------------+------+-----+-------------------+----------------+
| Field      | Type         | Null | Key | Default           | Extra          |
+------------+--------------+------+-----+-------------------+----------------+
| id         | int(11)      | NO   | PRI | NULL              | auto_increment |
| tag_name   | varchar(255) | YES  | UNI | NULL              |                |
| created_at | timestamp    | NO   |     | CURRENT_TIMESTAMP |                |
+------------+--------------+------+-----+-------------------+----------------+
3 rows in set (0.03 sec)

mysql>
```

FIGURE 2.7: tags table

2.3 E-R MOdel

Entity Realationship model of table is given below-

2.4 Specific Requirements

- PC have ubuntu 16.4,windows 7.0 or above os.
- Working internet connection.
- Minimum 4 GB Ram and 100 GB external memory.

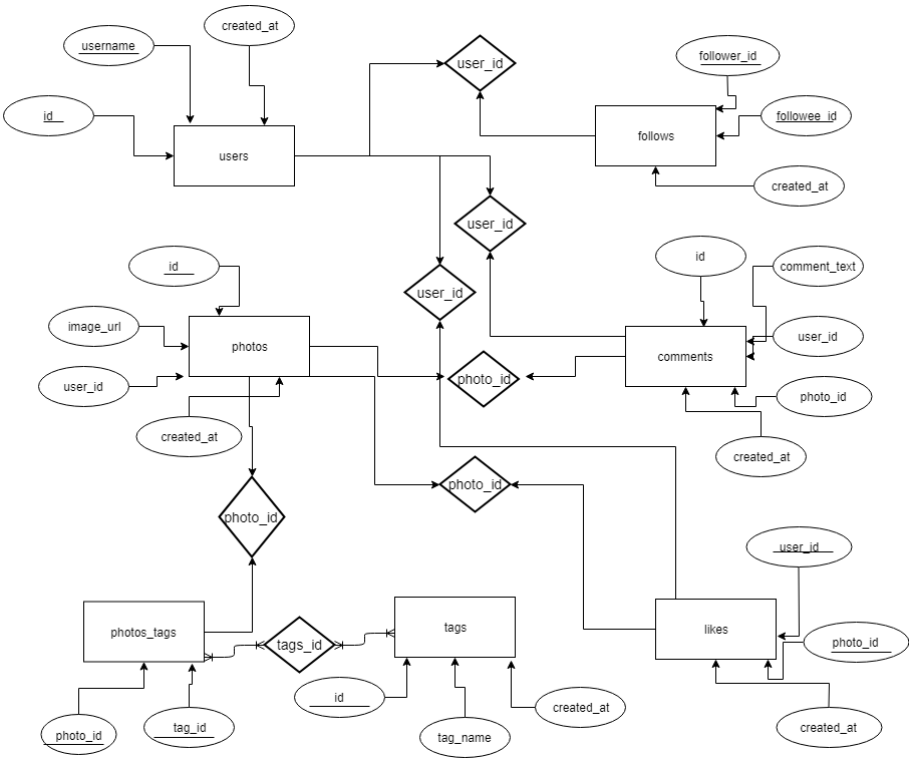

FIGURE 2.8: E-R Model

Chapter 3

Development and Implementation

3.1 Data Importing

In order to begin processing the Instagram Datasets are imported into the MySql environment and stored in MySql. After import of the data, we find the datasets contains lots of records.

3.2 Introduction to Language

SQL

SQL is a standard language for storing, manipulating and retrieving data in databases.SQL can execute queries against a database.It can retrieve data from a database.It can insert records in a database.It can update records in a database.It can delete records from a database.It can create new databases.SQL can create new tables in a database.It can create stored procedures in a database.SQL can create views in a database.SQL can set permissions on tables, procedures, and views.

3.3 Implementation with ScreenShots

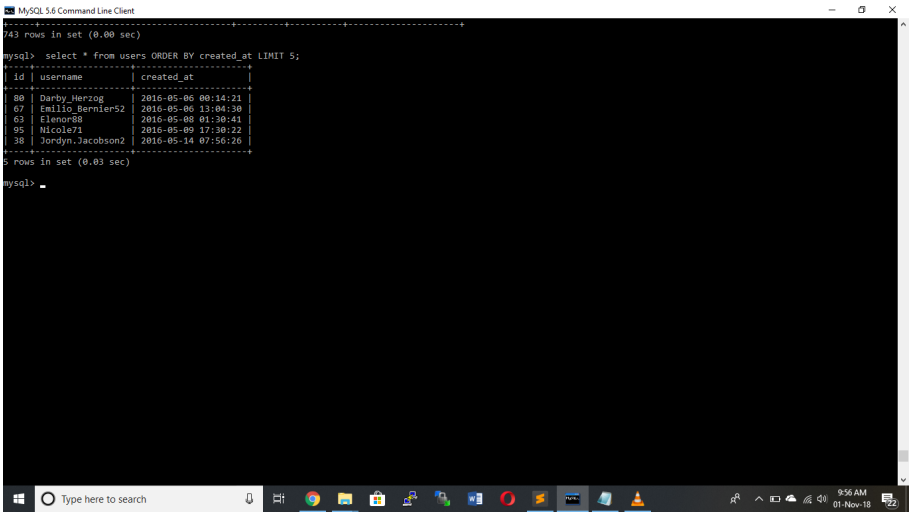

FIGURE 3.1: 5 oldest users

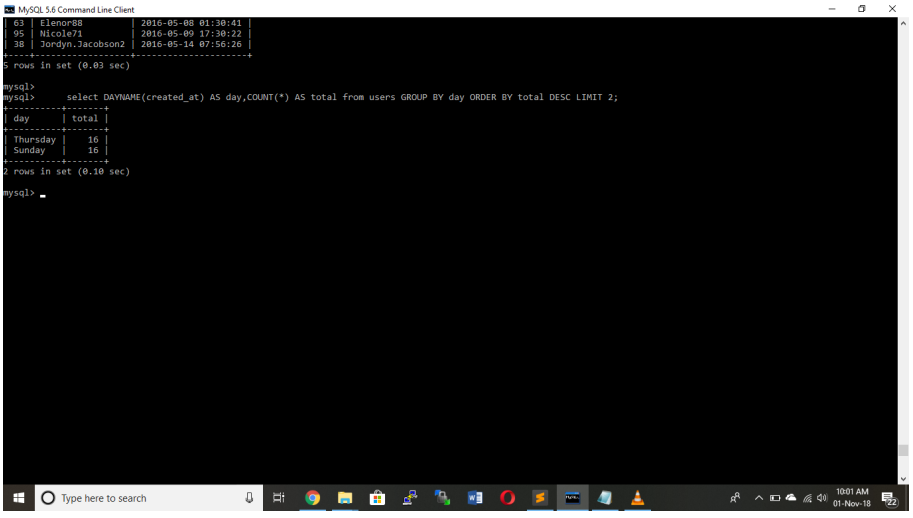

FIGURE 3.2: day of weeks do most users register

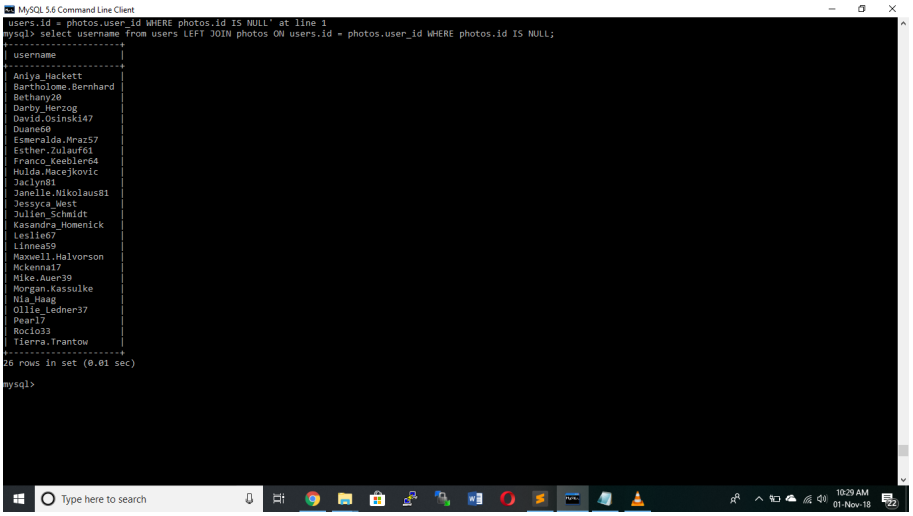

FIGURE 3.3: user who have never posted any photos

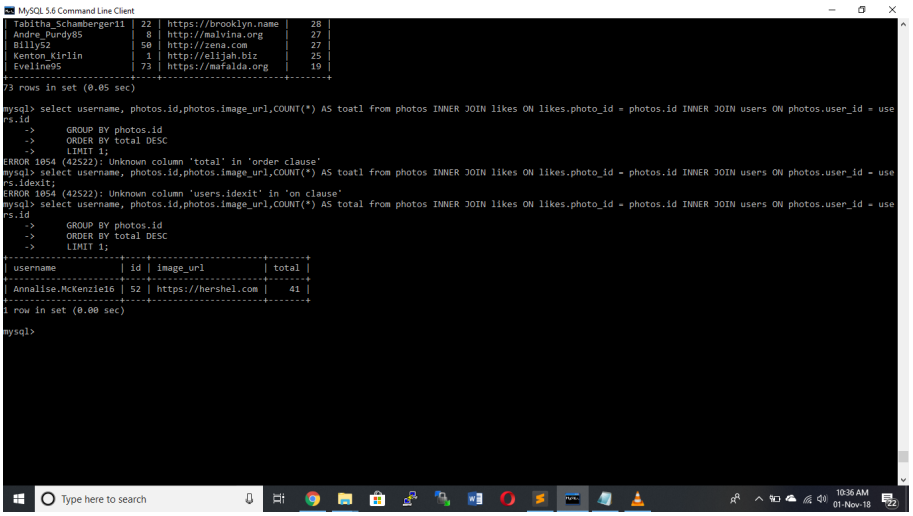

FIGURE 3.4: most likes on single photo

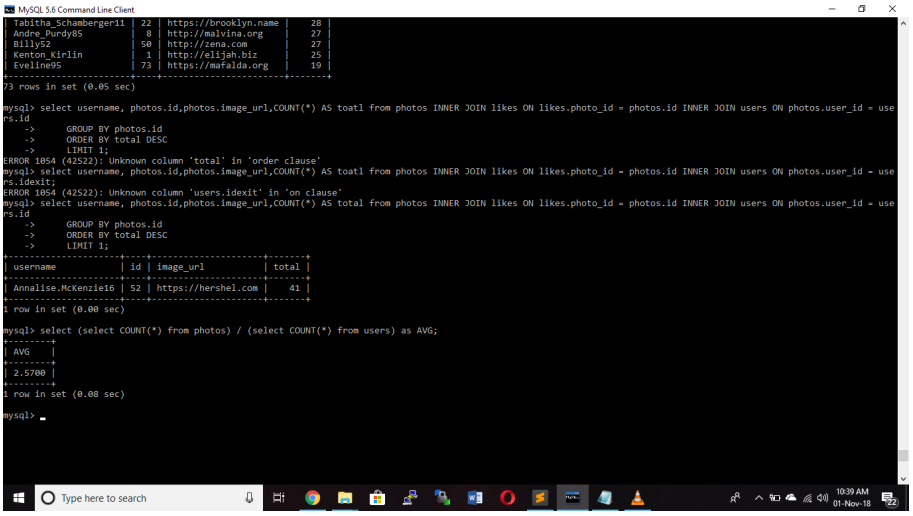

FIGURE 3.5: average user post

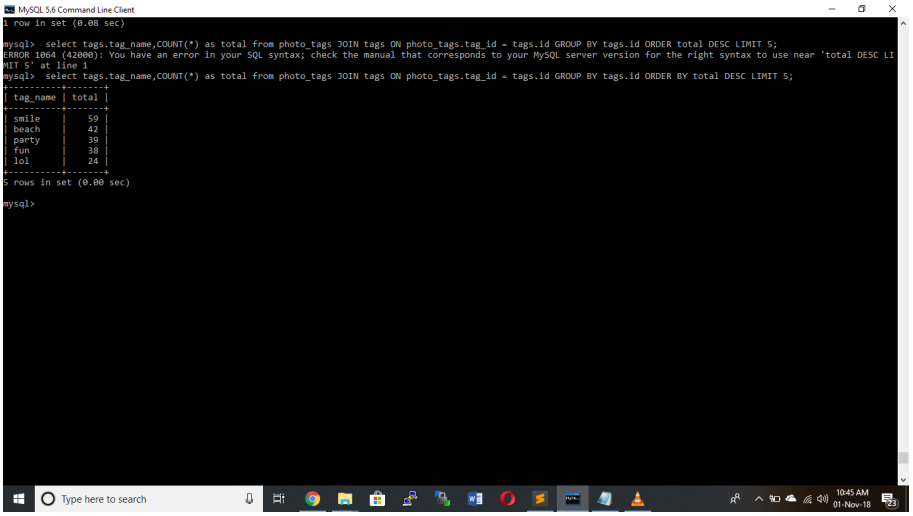

FIGURE 3.6: 5 commanly used hashtags

Chapter 4

Conclusion and Future Scope

4.1 Conclusion

This project has implemented all the features of Rdbms to handle Instagram Database and analyze to increases Instagram business.

4.2 Future Scope

The Project has been keeping in mind of future feature additions.

Due to Digitalization and other factors have made things accessible while simultaneously making it difficult to keep data structured and well-managed.

The Project has main moto to increase revenue of Instagram and make them increases marketing. I hope,this Project will achieve its aim of development and will generate a great revenue for Instagram.

Chapter 5

Refrences

1. https://www.mysql.com/ [accesed on 02/10//018]

2. https://www.w3schools.com/sql/default.asp [accesed on 08/10//018]

3. https://www.tutorialspoint.com/sql/ [accesed on 20/10//018]

4. https://stackoverflow.com/questions/tagged/sql-server [accesed on 25/10//018]

5. https://repository.genmymodel.com/JamesonReden/InstagramSchema [accesed on 30/10//018]

6. https://www.draw.io/ [accesed on 03/11//018]