Social Media Database – Final Report

Authors: Matthew Jilk | [matthew.jilk@go.winona.edu](mailto:matthew.jilk@go.winona.edu),

Yuan Longge | [longge.yuan@go.winona.edu](mailto:yuan.longge@go.winona.edu)

Completed Dec. 1st, 2021

# Summary

Today, nearly three billion people use social networking sites regularly. The technology behind these ubiquitous services is under constant stress to store and retrieve data on an unprecedented scale. Numerous models have been devised to meet the demands of “big data”, but the scope of this project will be limited to one: the relational database. Limiting our scope even further, we will concern ourselves only with the most basic functionality: users, posts, tags, comments, and groups. Although contrived, this represents the core text-based functionality of any social media website. The primary value of the relational database model comes from the ability to retrieve highly specific subsets of data via complex queries in a manner that is both efficient and elegant.

Implemented in MySQL, our design could, in theory, be used in the backend of a real social media website. To test our design, we created a set of sample data for each table. In deployment, data would be added when a user completes the sign-up process, publishes a post, commenting on an existing post, or creates a new group.

# Schema

User(user\_id, email, username, password, bio, date\_signed\_up)

**P.K. is user\_id;**

PostGroup(group\_id, group\_title, group\_description)

**P.K. is group\_id;**

Post(post\_id, post\_author\_id, post\_group\_id, post\_title, post\_body\_text, date\_posted)

**P.K. is post\_id; F.K. are post\_author\_id and post\_group\_id;**

Tag(tag\_post\_id, tagged\_user\_id)

**F.K. are tag\_post\_id and tagged\_user\_id;**

Comment(comment\_id, comment\_author\_id, comment\_post\_id, comment\_text, data\_commented)

**P.K. is comment\_id; F.K. are comment\_author\_id and comment\_post\_id;**

# SQL Code to Define Tables

**create** **database** social\_media**;**

**use** social\_media**;**

**DROP** **TABLE** **IF** **EXISTS** `**User**`**;**

**CREATE** **TABLE** `**User**` **(**

`**user\_id**` int **NOT** **NULL,**

`email` varchar**(**30**)** **NOT** **NULL,**

`username` varchar**(**20**)** **NOT** **NULL,**

`password` varchar**(**50**)** **NOT** **NULL,**

`bio` varchar**(**500**),**

`date\_signed\_up` date **NOT** **NULL,**

**PRIMARY** **KEY** **(**`**user\_id**`**)**

**);**

**DROP** **TABLE** **IF** **EXISTS** `PostGroup`**;**

**CREATE** **TABLE** `PostGroup` **(**

`group\_id` int **NOT** **NULL,**

`group\_title` varchar**(**30**)** **NOT** **NULL,**

`group\_description` varchar**(**500**),**

**PRIMARY** **KEY** **(**`group\_id`**)**

**);**

**DROP** **TABLE** **IF** **EXISTS** `Post`**;**

**CREATE** **TABLE** `Post` **(**

`post\_id` int **NOT** **NULL,**

`post\_author\_id` int **NOT** **NULL,**

`post\_group\_id` int **NOT** **NULL,**

`post\_title` varchar**(**40**)** **NOT** **NULL,**

`post\_body\_text` varchar**(**500**),**

`date\_posted` date **NOT** **NULL,**

**FOREIGN** **KEY** **(**`post\_author\_id`**)** **REFERENCES** **User(**`**user\_id**`**),**

**FOREIGN** **KEY** **(**`post\_group\_id`**)** **REFERENCES** PostGroup**(**`group\_id`**),**

**PRIMARY** **KEY** **(**`post\_id`**)**

**);**

**DROP** **TABLE** **IF** **EXISTS** `Tag`**;**

**CREATE** **TABLE** `Tag` **(**

`tag\_post\_id` int **NOT** **NULL,**

`tagged\_user\_id` int **NOT** **NULL,**

**FOREIGN** **KEY** **(**`tag\_post\_id`**)** **REFERENCES** Post**(**`post\_id`**),**

**FOREIGN** **KEY** **(**`tagged\_user\_id`**)** **REFERENCES** **User(**`**user\_id**`**),**

**);**

**DROP** **TABLE** **IF** **EXISTS** `PostComment`**;**

**CREATE** **TABLE** `PostComment` **(**

`comment\_id` int **NOT** **NULL,**

`comment\_author\_id` int **NOT** **NULL,**

`comment\_post\_id` int **NOT** **NULL,**

`comment\_text` varchar**(**500**)** **NOT** **NULL,**

`date\_commented` date **NOT** **NULL,**

**FOREIGN** **KEY** **(**`comment\_author\_id`**)** **REFERENCES** **User(**`**user\_id**`**),**

**FOREIGN** **KEY** **(**`comment\_post\_id`**)** **REFERENCES** Post**(**`post\_id`**),**

**PRIMARY** **KEY** **(**`comment\_id`**)**

**);**

# SQL Code to Insert Sample Data

**use** social\_media**;**

**insert** **into** **user** **values**

**(**1**,** 'jkk12@yahoo.com'**,** 'JohnFKenny'**,** 'ghRuUypass'**,** 'What can I say, I\'m just a regular guy'**,** '2005-01-01'**),**

**(**2**,** 'bsmish@gmail.com'**,** 'BobbyS'**,** '4ugf8eR1'**,** 'Hi fellas, my name is robert'**,** '2005-02-05'**),**

**(**3**,** 'fredbundy@aol.com'**,** 'fb\_man'**,** '48gU43feQR'**,** 'I have never commited a crime in my entire life...'**,** '2006-01-07'**),**

**(**4**,** 'freedomguy@gmail.com'**,** 'free\_dude'**,** '5jruFD9IUU'**,** 'We live in a society'**,** '2006-08-14'**),**

**(**5**,** 'yourmomlol@gmail.com'**,** 'urmother'**,** '80vbn3jf'**,** 'Your mother is a very fine lady...'**,** '2007-11-28'**),**

**(**6**,** 'yourdadlol@yahoo.com'**,** 'urdad'**,** '9vvnewouiA'**,** 'Your dad is a solid gentrleman.'**,** '2007-10-19'**),**

**(**7**,** 'raidshadowlegends@raid.com'**,** 'play\_raid\_now'**,** '98831023'**,** 'Play Raid Shadow Legends now!!!!!'**,** '2007-06-08'**),**

**(**8**,** 'epicness@integra.net'**,** 'i\_am\_epic'**,** 'secret\_password'**,** **NULL,** '2008-09-09'**);**

**insert** **into** postgroup **values**

**(**1**,** 'Environmentalism'**,** 'the environment is rad, man'**),**

**(**2**,** 'Cars'**,** 'cars are rad, man'**),**

**(**3**,** 'Electronics'**,** 'electronics are rad, man'**),**

**(**4**,** 'Politics'**,** 'politics are rad, man'**),**

**(**5**,** 'Science'**,** 'science is rad, man'**),**

**(**6**,** 'Religion'**,** 'religion'**);**

**insert** **into** post **values**

**(**1**,** 1**,** 1**,** 'save the environment!!!'**,** 'the environment is being wrecked, we need to save it'**,** '2005-02-01'**),**

**(**2**,** 1**,** 1**,** 'Screw trees'**,** 'I like the environment and all, but I would be perfectly fine if all trees died right now...'**,** '2005-02-01'**),**

**(**3**,** 1**,** 1**,** 'Green'**,** 'Green is my favorite color'**,** '2009-02-03'**),**

**(**4**,** 3**,** 1**,** 'Blue'**,** 'Blue is my favorite color'**,** '2010-05-01'**),**

**(**5**,** 4**,** 4**,** 'Science is da beset'**,** 'I like studying and reading and talking about and reading about and learning about science'**,** '2011-01-01'**),**

**(**6**,** 7**,** 5**,** 'Science sucks'**,** 'the title says it all'**,** '2015-09-25'**);**

**insert** **into** postcomment **values**

**(**1**,** 1**,** 4**,** 'this is a certified hood classic'**,** '2010-05-01'**),**

**(**2**,** 1**,** 5**,** 'bruh'**,** '2011-1-1'**),**

**(**3**,** 1**,** 6**,** 'lets go (dababby)'**,** '2015-09-25'**),**

**(**4**,** 6**,** 6**,** 'ur rarded'**,** '2015-09-25'**),**

**(**5**,** 6**,** 6**,** 'only a spoonful...'**,** '2015-10-16'**),**

**(**6**,** 8**,** 2**,** 'i hate you'**,** '2006-01-01'**);**

**insert** **into** tag **values**

**(**1**,** 5**),**

**(**2**,** 5**),**

**(**3**,** 5**),**

**(**4**,** 6**),**

**(**5**,** 6**),**

**(**5**,** 2**);**

# Entity-relationship Diagram

# Diagram Description automatically generatedDiagram Description automatically generated

# Example Queries

|  |  |  |
| --- | --- | --- |
| Description | SQL Code | Result |
| Get the title of every post in the group titled “Environmentalism” | **select** post\_title  **from** post **inner** **join** postgroup  **on** post\_group\_id **=** group\_id **and** group\_title **=** "Environmentalism"**;** | |  | | --- | | post\_title | | save the environment!!! | | Screw trees | | Green | | Blue | |
| Get the id and date of every post by the user with the username “i\_am\_epic” | **select** post\_id, date\_posted  **from** post **inner** **join** `**user**` **on** post\_author\_id **=** **user\_id**  **where** username **=** 'JohnFKenny'**;** | |  |  | | --- | --- | | post\_id | date\_posted | | 1 | 2005-02-01 | | 2 | 2005-02-01 | | 3 | 2009-02-03 | |
| Rank users based on how often they are tagged | **select** username**,** **count(**tagged\_user\_id**)** **as** tag\_count  **from** tag **inner** **join** **user** **on** tagged\_user\_id **=** **user\_id**  **group** **by** **user\_id**  **order** **by** tag\_count **desc;** | |  |  | | --- | --- | | username | tag\_count | | urmother | 3 | | urdad | 2 | | BobbyS | 1 | |

# Contributions

|  |  |  |
| --- | --- | --- |
|  | Matthew Jilk | Longge Yuan |
| Database design | X | X |
| Database programming | X | X\* |
| Sample data generation | X |  |
| Sample Queries | X |  |
| Acompanying Slideshow | X | X |
| Final Report | X | X |

“ \* ” signifies who did most of the work