



PROJECT SPECIFICATION

Udiddit, A Social News Aggregator (Part I)

SQL Syntax

CRITERIA	MEETS SPECIFICATIONS
Is the SQL code syntax-highlighted?	<p>All SQL code submitted should be properly syntax highlighted in order to make it easier to read.</p> <p>See this StackOverflow post for some tips on how to do so.</p>
Is the SQL code indented with proper spacing and newlines?	<p>Any SQL query that is more than about 80 characters should be multi-lined at the appropriate places (FROM, JOIN, WHERE, etc.) and indented to make it easier to read.</p>

Analysis of Provided ("bad") Data Schema

CRITERIA	MEETS SPECIFICATIONS

CRITERIA	MEETS SPECIFICATIONS
Does the analysis of the provided data schema point out at least three flaws related to data modeling best-practices?	<p>At least three flaws related to data modeling best-practices on the provided data schema are noted.</p> <p>The analysis of the provided data schema should contain bullet-point recommendations instead of only pointing out shortcomings. For example, instead of “such column shouldn’t be such data type”, you should say “such column should be this data type and have an index added to it”.</p>

Data Modeling and DDL

CRITERIA	MEETS SPECIFICATIONS
Does each of the DDL queries for creating a table execute properly in the Workspace?	Copy/pasting each DDL query in the workspace should produce a new table or index, and run without errors.

CRITERIA	MEETS SPECIFICATIONS
Does the provided data model satisfy all the features and queries outlined in Part 2's guidelines 1 and 2?	While the data model is allowed to contain extra fields not required to satisfy the features requested, like for example the addition of some audit fields, the data model should at least be able to answer all the features and queries outlined.
Is the provided data model normalized?	The data model provided should satisfy first, second, and third normal forms.
Is the provided data model consistent?	The data model provided should not allow any non-existent IDs to be inserted in related tables.
Is the provided data model performant?	The data model provided should have appropriate indexes in order for the queries in guidelines 1 and 2 to run quickly and efficiently. Note as some constraints automatically add indexes, it's important to make sure that you don't duplicate any indexes.

CRITERIA	MEETS SPECIFICATIONS
Does the new schema allow users to register? (Guideline 1)	<ul style="list-style-type: none">• Each username has to be unique• Usernames can be composed of at most 25 characters• Usernames can't be empty• We won't worry about user passwords for this project
Does the new schema allow registered users to create new topics? (Guideline 1)	<ul style="list-style-type: none">• Topic names have to be unique.• The topic's name is at most 30 characters• The topic's name can't be empty• Topics can have an optional description of at most 500 characters.
Does the new schema allow registered users to create new posts on existing topics? (Guideline 1)	<ul style="list-style-type: none">• Posts have a required title of at most 100 characters• The title of a post can't be empty.• Posts should contain either a URL or a text content, but not both.• If a topic gets deleted, all the posts associated with it should be automatically deleted too.• If the user who created the post gets deleted, then the post will remain, but it will become dissociated from that user.

CRITERIA	MEETS SPECIFICATIONS
Does the new schema allow registered users to comment on existing posts? (Guideline 1)	<ul style="list-style-type: none">• A comment's text content can't be empty.• Contrary to the current linear comments, the new structure should allow comment threads at arbitrary levels.• If a post gets deleted, all comments associated with it should be automatically deleted too.• If the user who created the comment gets deleted, then the comment will remain, but it will become dissociated from that user.• If a comment gets deleted, then all its descendants in the thread structure should be automatically deleted too.
Does the new schema make sure that a given user can only vote once on a given post? (Guideline 1)	<ul style="list-style-type: none">• A user can only cast one vote on a given post.• If the user who cast a vote gets deleted, then all their votes will remain, but will become dissociated from the user.• If a post gets deleted, then all the votes for that post should be automatically deleted too.

Data Migration and DML

CRITERIA	MEETS SPECIFICATIONS

CRITERIA	MEETS SPECIFICATIONS
Does the migrated data perfectly match the initial data provided in the workspace?	The migrated data will be in normalized form, with each table having its own primary key, and related tables using IDs and foreign keys. While the database will verify that the related IDs are consistent, it can't actually verify that they're correct. Thus, the DML used to migrate the data needs to make sure that the right IDs are associated together, not just any valid ID.
Does executing the DML provided run without errors?	Since the data has to be migrated following a certain order and logic — some data has dependencies — it's important that the DML that is provided can be literally copy/pasted as-is, and run without errors.

Suggestions to Make Your Project Stand Out!

- Outline more than three things that you find wrong with the initially provided, denormalized Udiddit schema
- Write clean and legible DDL with self-documenting table, column, constraint, and index names.
- Add a section containing all or some of the most challenging DQLs for the queries and features of Part 2
- Using WITH and JSON_AGG, devise a query that will retrieve a JSON document with three levels of nested comments for a given post
