

Final Project Rubric

Overall Point Breakdown

- Milestone 1 (3/24) - 50 pts
- Milestone 2 (3/31) - 100 pts
- Milestone 3 (4/7) - 100 pts
- Milestone 4 (4/14) - 100 pts
- Milestone 5 (4/21) - 100 pts
- Milestone 6 (4/28) - 100 pts
- Final Presentation - 100 pts

Milestone 1 [50]

Environment Setup [35]

- **Screenshots** (Screenshots from both members from connection tests) [35]

Organization [15]

- **Stache Entry** (Correct JSON format in Stache. The credentials in this entry should correspond with the Redshift instance that you are running your program with) [5]
- **Github File Structure** (Folder and files must be named exactly as specifications. **Must be in folder named 'final-project'**) [5]
- **Submission JSON** (Correct JSON format and content) [5]

Milestone 2 [100]

Staging Schemas [30]

- **Physical Diagrams** (Schemas correctly represents data set contents) [20]
- **Minimum Requirments** (Minimum of 30 files from 2 datasets and good prominent attributes) [10]

Unified Schema [55]

- **English Queries** (10-12 queries that use the data from your staging schemas, should have some queries that use both data sets) [10]
- **Conceptual Diagram** (Valid and well designed, information can be found in staging tables, should show bridging of data between data sets) [20]
- **Physical Diagram** (Consistent with conceptual, normalized) [15]
- **Data Dictionary** (Consistent with diagrams) [10]
- Will be looking for naming conventions. Also table names should be renamed to prefix the table name with the dataset name (e.g. discog_releases instead of releases).

Organization [15]

- **Stache Entry** (Correct JSON format in Stache. The credentials in this entry should correspond with the Redshift instance that you are running your program with) [3]
- **Github File Structure** (Folder and files must be named exactly as specifications. **Must be in folder named 'final-project'**) [3]
- **Github Commits** (Both members must commit) [3]
- **Github Issues** (At least 3 issues) [3]
- **Submission JSON** (Correct JSON format and content) [3]

Milestone 3 [100]

DDL Scripts [40]

- **DDL Scripts** (2-3 DDL Scripts, scripts are able to run and copy data to the database instance) [30]
- **Script Content** (Scripts contain the tables that were being used in the diagrams) [10]

Instance [40]

- **Instance Content** (All the data from chosen the datasets are imported into the database) [20]
- **Valid Schemas** (Schemas can be found in the database with all the respective relations within the correct schemas) [20]

Organization [20]

- **Stache Entry** (Correct JSON format in Stache. The credentials in this entry should correspond with the Redshift instance. Make sure your IAM account access works by trying to login and change your security settings from there) [10]
- **Github File Structure** (Folder and files must be named exactly as specifications. You should have 2-3 main DDL .sql files, 2-3 copy .sql files, and 2-3 check .txt files. Do not upload csv files to Github. **Must be**

in folder named 'final-project') [3]

- **Github Commits** (Both members must commit) [1]
- **Github Issues** (At least 3 issues) [1]
- **Submission JSON** (Correct JSON format and content) [5]

Milestone 4 [100]

Data Transformation [40]

- **Join Column** (Punctuation removed and first letters capitalized and using actual max number of bytes) [20]
- **Datatypes** (Boolean and Date and Number datatypes are actually their respective datatype and not varchar.) [20]

Unified Schema [20]

- **Unified Table** (Create tables based off of a subset of existing and cleansed columns) [20]

Queries [20]

- **Queries** (Use unified tables to answer queries and give meaningful reports) [20]

Organization [20]

- **Stache Entry** (Correct JSON format in Stache. The credentials in this entry should correspond with the Redshift instance. Make sure your IAM account access works by trying to login and change your security settings from there) [10]
- **Github File Structure** (Folder and files must be named exactly as specifications. You should have 2-3 transform sql files, a create_unified.sql file, and a queries.sql file. **Must be in folder named 'final-project'**) [3]
- **Github Commits** (Both members must commit) [1]
- **Github Issues** (At least 3 issues) [1]
- **Submission JSON** (Correct JSON format and content. Please name file with lowercase .json) [5]

Milestone 5 [100]

Virtual Views [40]

- **Virtual Views** (Views created based off select statements to answer the queries) [30]
- **Timing Constraint** (View does not take too long to run) [10]

Reworked Queries [10]

- **Queries** Adjustments made to the analysis queries for visualization purposes (e.g. changing the output from one column to multiple columns and from one scalar value to a list of values, etc.) [10]

Data Visualization [30]

- **QuickSight** (10 visuals, one corresponding to each view. Should show information in an interesting way) [30]

Organization [20]

- **Stache Entry** (Correct JSON format in Stache. The credentials in this entry should correspond with the Redshift instance. Make sure your IAM account access works by trying to login and change your security settings from there) [10]
- **Github File Structure** (Folder and files must be named exactly as specifications. You should have one create_views.sql file and one quicksight folder with 10 screenshots. **Must be in folder named 'final-project'**) [3]
- **Github Commits** [1]
- **Github Issues** (At least 3 issues) [1]
- **Submission JSON** (Correct JSON format and content. Please name file with lowercase .json) [5]

Milestone 6 [100]

Technical Report [70]

- **Documentation of Experiences** (Should include documenting experience with all parts of milestones, also challenges encountered. You can choose how you want to organize the paper) [40]
- **Minimum Length** (At least 10 pages double-spaced of words) [20]
- **English Grammer** (Shows generally good writing skills, should not be difficult to read) [10]

Reworkings [10]

- **Corrected Work** (Will be looking for corrected diagrams to reflect the final outcome of the project) [10]

Organization [20]

- **Github File Structure** (Folder and files must be named exactly as specifications. final_report.pdf file under the final-project directory. **Must be in folder named 'final-project'**) [4]
- **Github Commits** [3]
- **Github Issues** (At least 3 issues) [3]
- **Submission JSON** (Correct JSON format and content. Please name file with lowercase .json) [10]

Milestone 7 [100]

Presentation [100]

- **Group Prepared Content** (Presentation that shows notable parts of the project. Must be 4-5 minutes) [40]
- **Questions** (Ability to answer questions in a way that shows understanding of the assignment) [40]
- **Clarity and Understanding** (Overall shows understanding of assignment and is conveyed clearly) [20]