



User Feedback Analysis System

08.19.2023

Nimra Sharnez
Personal Project

Overview

EduTech is an innovative educational application that offers a variety of online courses to students across the globe. As users engage with the app's content and features, they often leave feedback about their experiences. The EduTech team recognizes the importance of understanding user sentiments and preferences to continuously improve the app's offerings.

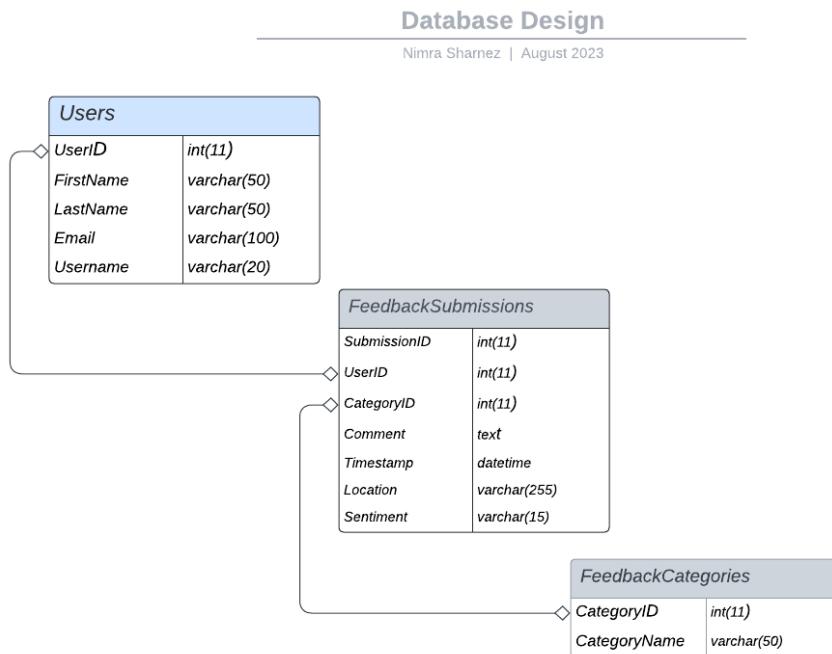
Goals

1. **Enhancing User Experience:** The primary goal of EduTech is to provide a seamless and enriching learning experience to its users. By analyzing user feedback, the company aims to identify pain points and areas of improvement within the app.
2. **Course Customization:** EduTech offers a wide range of courses, and the company's second goal is to tailor course offerings based on user preferences. By mining feedback data, they intend to gain insights into which courses resonate most with users and why.
3. **Engagement Optimization:** Keeping users engaged is crucial for the success of any app. EduTech's third goal is to optimize user engagement by understanding the factors that contribute to positive user experiences, as well as identifying patterns that lead to disengagement.



Database Design

- Create a SQL database with tables to store user feedback data. Create tables for users, feedback submissions, feedback categories, and timestamps.
- Define appropriate relationships between tables, such as a foreign key from feedback submissions to users, and from feedback submissions to categories.
- Include columns for sentiment scores, comments, and any relevant metadata.



Created using Lucid Software

Data:

[Dataset Inspiration](#)

Data Pipelining

The data will be extracted from the company's web app feedback database. Additionally, to gain insights on the users who are not leaving feedback, the data will also be extracted from the company's user database.

Data Sources:

- Company's Web App Feedback Database
- Company's User Database

DataGenerator [.ipynb file](#): additional data extracting, additional generating (for project scope purposes), transforming and cleaning (Tools utilized: *Python, Pandas*)

Raw Data (Before)										Cleaned Data (After)											
Comment	Sentiment	CategoryID	Timestamp	UserID	Location	SubmissionID	Categories	First Name	Last Name	Email	Comment	Sentiment	CategoryID	Timestamp	UserID	Location	SubmissionID	Categories	First Name	Last Name	Email
0 "I love this product!"	Positive	Twitter	2023-09-15 09:23:14	@user123	New York	1	General Comments	Winnie	Levinson	W.Levinson23@yahoo.com	0 "I love this product!"	Positive	Twitter	2023-09-15 09:23:14	@user123	New York	1	General Comments	Winnie	Levinson	W.Levinson23@yahoo.com
1 "The service was terrible."	Negative	Yelp Review	2023-09-15 11:45:32	user456	Los Angeles	2	User Experience Comments	Katherine	Gray	K.Gray43@yahoo.com	1 "The service was terrible."	Negative	Yelp Review	2023-09-15 11:45:32	user456	Los Angeles	2	User Experience Comments	Katherine	Gray	K.Gray43@yahoo.com
2 "This movie is amazing!"	Positive	IMDb	2023-09-15 12:00:00	moviereview789	London	3	Algebra	Ruth	Blume	R.Blume68@gmail.com	2 "This movie is amazing!"	Positive	IMDb	2023-09-15 12:00:00	moviereview789	London	3	Algebra	Ruth	Blume	R.Blume68@gmail.com
3 "I'm so disappointed with their customer support."	Negative	Customer Support	2023-09-15 17:25:11	forumuser1	Toronto	4	English	George	Cardenas	G.Cardenas87@outlook.com	3 "I'm so disappointed with their customer support."	Negative	Customer Support	2023-09-15 17:25:11	forumuser1	Toronto	4	English	George	Cardenas	G.Cardenas87@outlook.com
4 "Just had the best meal of my life!"	Positive	Foodie	2023-09-16 08:10:59	foodie22	Paris	5	Chemistry	Jessie	Harrison	J.Harrison8@gmail.com	4 "Just had the best meal of my life!"	Positive	Foodie	2023-09-16 08:10:59	foodie22	Paris	5	Chemistry	Jessie	Harrison	J.Harrison8@gmail.com
5 "The quality of this product is subpar."	Negative	Customer Support	2023-09-16 10:15:27	shopper123	San Francisco	6	Biology	Patricia	Mero	P.Mero31@yahoo.com	5 "The quality of this product is subpar."	Negative	Customer Support	2023-09-16 10:15:27	shopper123	San Francisco	6	Biology	Patricia	Mero	P.Mero31@yahoo.com
6 "I can't stop listening to this song. It's incredible!"	Positive	Music Lover	2023-09-16 11:30:00	musiclover456	Berlin	7	Computer Science	Betty	Mercado	B.Mercado89@gmail.com	6 "I can't stop listening to this song. It's incredible!"	Positive	Music Lover	2023-09-16 11:30:00	musiclover456	Berlin	7	Computer Science	Betty	Mercado	B.Mercado89@gmail.com
7 "Their website is so user-friendly. Love it!"	Positive	Customer Support	2023-09-16 16:00:36	testimonials1	Sydney	8	Writing	David	Williams	D.Williams63@outlook.com	7 "Their website is so user-friendly. Love it!"	Positive	Customer Support	2023-09-16 16:00:36	testimonials1	Sydney	8	Writing	David	Williams	D.Williams63@outlook.com
8 "I loved the movie! It was fantastic!"	Positive	Customer Support	2023-09-17 10:45:21	user123	New York	9	Algebra	Mary	Montana	M.Montana88@gmail.com	8 "I loved the movie! It was fantastic!"	Positive	Customer Support	2023-09-17 10:45:21	user123	New York	9	Algebra	Mary	Montana	M.Montana88@gmail.com
9 "The customer service was terrible."	Negative	Customer Support	2023-09-17 12:30:00	customer222	Los Angeles	10	User Experience Comments	Katherine	Gray	K.Gray43@yahoo.com	9 "The customer service was terrible."	Negative	Customer Support	2023-09-17 12:30:00	customer222	Los Angeles	10	User Experience Comments	Katherine	Gray	K.Gray43@yahoo.com
10 "This book made me feel inspired. Highly recommended."	Positive	Bookworm	2023-09-17 12:30:00	bookworm789	London	11	Statistics	Denver	Johnson	D.Johnson32@outlook.com	10 "This book made me feel inspired. Highly recommended."	Positive	Bookworm	2023-09-17 12:30:00	bookworm789	London	11	Statistics	Denver	Johnson	D.Johnson32@outlook.com
11 "I'm extremely disappointed with their product."	Negative	Customer Support	2023-09-18 15:21:43	shopper123	San Francisco	12	Geometry	John	Harb	J.Harb1@yahoo.com	11 "I'm extremely disappointed with their product."	Negative	Customer Support	2023-09-18 15:21:43	shopper123	San Francisco	12	Geometry	John	Harb	J.Harb1@yahoo.com
12 "Just had the most amazing vacation! I can't wait to go back!"	Positive	Travel Enthusiast	2023-09-18 16:01:23	travelenthusiast1	Sydney	13	Chemistry	Peter	Clausen	P.Clausen57@gmail.com	12 "Just had the most amazing vacation! I can't wait to go back!"	Positive	Travel Enthusiast	2023-09-18 16:01:23	travelenthusiast1	Sydney	13	Chemistry	Peter	Clausen	P.Clausen57@gmail.com

Data Retriever Client:

- Data Extraction pt.2:
 - ~~Retrieve from Company's Web Application Databases~~
- Data Transforming pt.2:
 - ~~Organize data to fit the database schema~~
- Data Loading:
 - ~~Upload new user data into the MySQL database utilizing prepared statements. Done daily.~~
 - ~~Upload new feedback data into the MySQL database referencing the last timestamp data was collected. As per Data Analyst request, weekly.~~
 - ~~Upload new feedback categories into the MySQL database. Done only if the application adds a new course (rare).~~

Data Base

```
+-----+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
| CategoryID | int(11) | NO | PRI | NULL | auto_increment |
| CategoryName | varchar(50) | YES | | NULL |
+-----+-----+-----+-----+-----+
2 rows in set (0.01 sec)

[mysql]> DESCRIBE FeedbackSubmissions;
+-----+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
| SubmissionID | int(11) | NO | PRI | NULL | auto_increment |
| UserID | int(11) | YES | MUL | NULL |
| CategoryID | int(11) | YES | MUL | NULL |
| Comment | text | YES | | NULL |
| Timestamp | datetime | YES | | NULL |
| Location | varchar(255) | YES | | NULL |
| Sentiment | varchar(15) | YES | | NULL |
+-----+-----+-----+-----+-----+
7 rows in set (0.01 sec)

[mysql]> DESCRIBE Users;
+-----+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
| UserID | int(11) | NO | PRI | NULL | auto_increment |
| FirstName | varchar(50) | YES | | NULL |
| LastName | varchar(50) | YES | | NULL |
| Email | varchar(100) | YES | UNI | NULL |
| Username | varchar(20) | YES | | NULL |
+-----+-----+-----+-----+-----+
5 rows in set (0.00 sec)
```

Data Mining

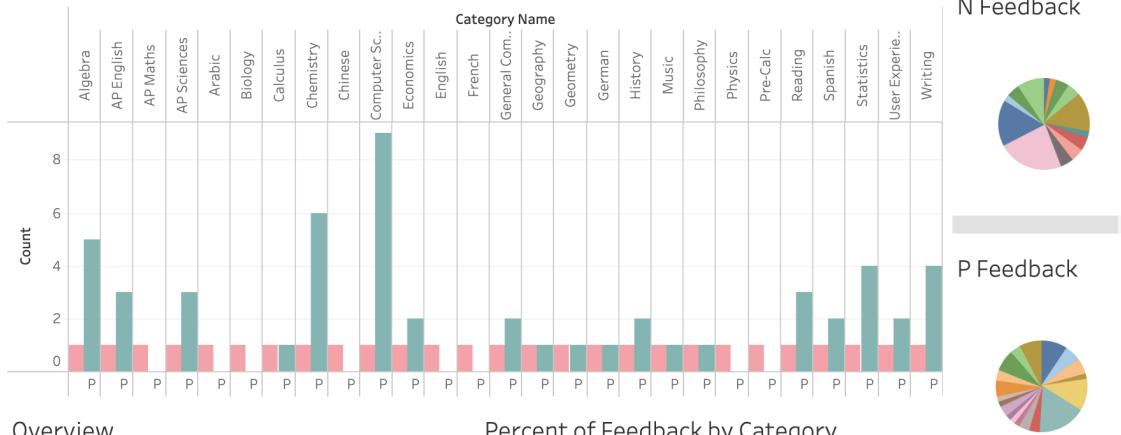
Write SQL queries and create stored procedures to retrieve insights from the data. Please refer [here](#) for the queries.

- Retrieve the most recent feedback
- Find the amount of positive and negative feedback from the latest week.
- Discover the prevailing sentiment within each feedback category by determining the most frequently occurring sentiment.

Visuals (Made with Tableau)

Sentiment Distribution Across Different Categories ([Interactive on Tableau](#))

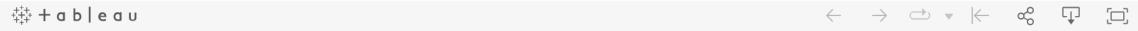
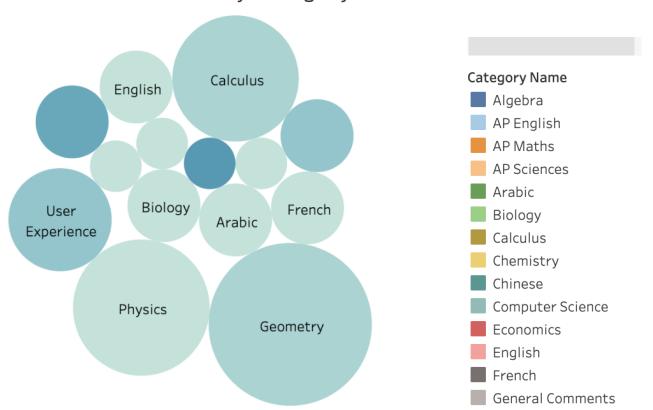
Count of N:P Feedback by Category



Overview

Category Name	Count o...	Negativ...
Algebra	1.00	1.00
AP English	1.00	0.00
AP Maths	1.00	1.00
AP Sciences	1.00	0.00
Arabic	1.00	2.00
Biology	1.00	2.00
Calculus	1.00	6.00
Chemistry	1.00	0.00
Chinese	1.00	1.00
Computer Science	1.00	0.00
Economics	1.00	2.00
English	1.00	2.00
French	1.00	2.00
General Comments	1.00	0.00
Geography	1.00	0.00
Geometry	1.00	10.00

Percent of Feedback by Category



Additional Notes and Edits

Image of database design

SQL queries used

```
[mysql> CREATE DATABASE UserFeedbackDB;
Query OK, 1 row affected (0.07 sec)

[mysql> USE UserFeedbackDB;
Database changed
[mysql> CREATE TABLE Users (
[|     -> UserID INT PRIMARY KEY AUTO_INCREMENT,
[|     -> FirstName VARCHAR(50),
[|     -> LastName VARCHAR(50),
[|     -> Email VARCHAR(100) UNIQUE);
Query OK, 0 rows affected (0.08 sec)

[mysql> CREATE TABLE FeedbackCategories (
[|     -> CategoryID INT PRIMARY KEY AUTO_INCREMENT,
[|     -> CategoryName VARCHAR(50));
Query OK, 0 rows affected (0.07 sec)

[mysql> CREATE TABLE FeedbackSubmissions (
[|     -> SubmissionID INT PRIMARY KEY AUTO_INCREMENT,
[|     -> UserID INT,
[|     -> CategoryID INT,
[|     -> SentimentScore DECIMAL(5, 2),
[|     -> Comment TEXT,
[|     -> Timestamp DATETIME,
[|     -> FOREIGN KEY (UserID) REFERENCES Users(UserID),
[|     -> FOREIGN KEY (CategoryID) REFERENCES FeedbackCategories(CategoryID));
Query OK, 0 rows affected (0.12 sec)
```

Tables_in_userfeedbackdb
FeedbackCategories
FeedbackSubmissions
Users

Upon further consideration, I have added the available Location data as well:

```
mysql> ALTER TABLE FeedbackSubmissions ADD Location VARCHAR(255);
Query OK, 0 rows affected (0.07 sec)
Records: 0  Duplicates: 0  Warnings: 0
```

Additionally ended up adding in a Username field.

```
mysql> DESCRIBE Users;
+-----+-----+-----+-----+-----+
| Field | Type   | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
| UserID | int(11) | NO   | PRI | NULL    | auto_increment |
| FirstName | varchar(50) | YES  |     | NULL    |
| LastName | varchar(50) | YES  |     | NULL    |
| Email | varchar(100) | YES  | UNI | NULL    |
| Username | varchar(20) | YES  |     | NULL    |
+-----+-----+-----+-----+-----+
```

The Data Generation process can be seen on this online [python notebook](#).

```
mysql> describe FeedbackSubmissions;
+-----+-----+-----+-----+-----+
| Field | Type   | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
| SubmissionID | int(11) | NO   | PRI | NULL    | auto_increment |
| UserID | int(11) | YES  | MUL | NULL    |
| CategoryID | int(11) | YES  | MUL | NULL    |
| SentimentScore | decimal(5,2) | YES  |     | NULL    |
| Comment | text   | YES  |     | NULL    |
| Timestamp | datetime | YES  |     | NULL    |
| Location | varchar(255) | YES  |     | NULL    |
+-----+-----+-----+-----+-----+
7 rows in set (0.01 sec)

mysql> SELECT * FROM FeedbackSubmissions;
Empty set (0.00 sec)

mysql> ALTER TABLE FeedbackSubmissions
[   -> ADD Sentiment VARCHAR(15);
Query OK, 0 rows affected (0.08 sec)
Records: 0  Duplicates: 0  Warnings: 0

mysql> UPDATE FeedbackSubmissions
[   -> SET Sentiment = CAST(SentimentScore AS CHAR(15));
Query OK, 0 rows affected (0.01 sec)
Rows matched: 0  Changed: 0  Warnings: 0

mysql> ALTER TABLE FeedbackSubmissions
[   -> DROP COLUMN SentimentScore;
Query OK, 0 rows affected (0.09 sec)
Records: 0  Duplicates: 0  Warnings: 0

mysql> DESCRIBE FeedbackSubmissions;
+-----+-----+-----+-----+-----+
| Field | Type   | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+
| SubmissionID | int(11) | NO   | PRI | NULL    | auto_increment |
| UserID | int(11) | YES  | MUL | NULL    |
| CategoryID | int(11) | YES  | MUL | NULL    |
| Comment | text   | YES  |     | NULL    |
| Timestamp | datetime | YES  |     | NULL    |
| Location | varchar(255) | YES  |     | NULL    |
| Sentiment | varchar(15) | YES  |     | NULL    |
+-----+-----+-----+-----+-----+
7 rows in set (0.00 sec)
```

Additional edits can be seen in the screenshot here.

I initially added the new attribute Sentiment and replaced SentimentScore. I then re-adjusted its type from VARCHAR(15) to CHAR(15).

Stored Procedures:

latest_feedback(): Retrieves the most recent feedback(s) from the database

```
mysql> DELIMITER //
mysql> CREATE PROCEDURE latest_feedback()
-> BEGIN
->     SELECT fs.*
->         FROM FeedbackSubmissions fs
->         WHERE Timestamp = (SELECT MAX(Timestamp) FROM FeedbackSubmissions);
-> END //
Query OK, 0 rows affected (0.09 sec)
```

Output:

```
mysql> CALL latest_feedback();
+-----+-----+-----+-----+-----+-----+
| SubmissionID | UserID | CategoryID | Comment | Timestamp | Location | Sentiment |
+-----+-----+-----+-----+-----+-----+
|       68 |    59 |        12 | "The website loading speed is frustratingly slow. Needs improvement." | 2023-07-17 17:15:22 | Toronto | Negative |
|       79 |    59 |        12 | "The website loading speed is frustratingly slow. Needs improvement." | 2023-07-17 17:15:22 | Toronto | Negative |
+-----+-----+-----+-----+-----+-----+
2 rows in set (0.00 sec)
```

latestweeks_sentiment(): Retrieves the amount of positive and negative feedback from the latest week.

```
mysql> SET @weekAgo = (SELECT MAX(Timestamp) - INTERVAL 7 DAY FROM FeedbackSubmissions);
Query OK, 0 rows affected (0.00 sec)

mysql> SELECT @weekAgo;
+-----+
| @weekAgo |
+-----+
| 2023-07-10 17:15:22 |
+-----+
1 row in set (0.00 sec)
```

```
mysql> DELIMITER //
mysql> CREATE PROCEDURE latestweeks_sentiment(IN weekAgo DATETIME)
-> BEGIN
->     SELECT Positive, Negative
->         FROM (
->             SELECT
->                 (SELECT COUNT(FS.Sentiment) FROM FeedbackSubmissions FS WHERE FS.Sentiment = 'Positive' AND FS.Timestamp > weekAgo) AS Positive,
->                 (SELECT COUNT(FS.Sentiment) FROM FeedbackSubmissions FS WHERE FS.Sentiment = 'Negative' AND FS.Timestamp > weekAgo) AS Negative
->         ) AS Sentiments;
-> END //
Query OK, 0 rows affected (0.05 sec)
```

Output:

```
mysql> CALL latestweek_sentiment();
+-----+-----+
| Positive | Negative |
+-----+-----+
|      19 |       18 |
+-----+-----+
```

Negative Sentiment Feedback Count by Categories

```
mysql> SELECT COUNT(FS.Sentiment) AS NegativeSentimentCount, FC.CategoryName
[   -> FROM FeedbackSubmissions FS JOIN FeedbackCategories FC ON FS.CategoryID = FC.CategoryID
[   -> WHERE FS.Sentiment='Negative'
-> GROUP BY FS.CategoryID;
+-----+-----+
| NegativeSentimentCount | CategoryName
+-----+-----+
|          4 | User Experience Comments
|          2 | English
|          2 | Biology
|         10 | Geometry
|          6 | Calculus
|          7 | Physics
|          2 | Economics
|          2 | Statistics
|          1 | AP Maths
|          2 | French
|          1 | Algebra
|          1 | Chinese
|          2 | Arabic
|          1 | Pre-Calc
+-----+-----+
14 rows in set (0.00 sec)
```

Positive Sentiment Feedback Count by Categories

```
mysql> SELECT COUNT(FS.Sentiment) AS PositiveSentimentCount, FC.CategoryName
[   -> FROM FeedbackSubmissions FS JOIN FeedbackCategories FC ON FS.CategoryID = FC.CategoryID
[   -> WHERE FS.Sentiment='Positive' GROUP BY FS.CategoryID;
+-----+-----+
| PositiveSentimentCount | CategoryName
+-----+-----+
|          2 | General Comments
|          5 | Algebra
|          6 | Chemistry
|          9 | Computer Science
|          4 | Writing
|          4 | Statistics
|          2 | History
|          3 | Reading
|          2 | User Experience Comments
|          1 | Calculus
|          3 | AP Sciences
|          3 | AP English
|          2 | Economics
|          2 | Spanish
|          1 | German
|          1 | Geometry
|          1 | Geography
|          1 | Philosophy
|          1 | Music
+-----+-----+
19 rows in set (0.00 sec)
```