# CS480 – Course project

Summer 2020

## Database: Twitter tweets about Covid19 with attached sentiment

**Description:**

Consider the design of the following database system for analyzing thoughts, opinions, and information regarding the current Covid19 pandemic. A stream of twitter tweets is collected via the twitter api. Each tweet has a unique ID, the text associated with it and the location. Sentiment analysis will be performed on each tweet. Each tweet after analysis receives a sentiment value that is held in a sentiment table. Each sentiment has a sentiment value of positive or negative and also another column stating whether it is high or low. To get a better understanding of the spread of information and opinion I will also be keeping track of the tweet's user followers. This table will hold the users' follower count, their opinion from the sentiment and the user ID.

# Project Part 2 – CRUD (Create, read, update, and delete)

## Deadline: July 18, 2020

**List of entities:**

1. Tweets
2. Sentiment
3. User's Followers

Based on the Demo (Part 1), implement the following functionality using Java and SQL with necessary GUI interfaces.

1. Insert/delete/update/read: a tweet can be inserted into the database and can be deleted and read. However, it can not be updated as I would like to keep the data intact as the user tweeted. Twitter ID will be used for uniquely identifying a tweet therefore it can not be altered. The tweet ID will be generated by Twitter.

2. Insert/delete/update/read: a sentiment value can be inserted and read. However, deletion and updates will not be allowed so that I can keep the integrity of the data and obtain accurate and non tempered results.

3. Insert/delete/update/read: a user follower record can be updated to reflect a more current piece of information. The fields that can be updated are the following count and location and opinion. The user ID can not change.

# Project Part 3 – Queries

## Deadline: August 1, 2020

Based on the Demo, implement the following functionality using Java and SQL with necessary GUI interfaces.

**Trivial Queries:**

1. List all Tweets
2. List all users follow count
3. List all Tweets sentiment

**Non-trivial Queries:**

1. List all Tweets with a high positive sentiment.

2. List all user with a follow count of # of higher

3. Find all tweets in which the location is from the US

4. List the user with the most number of followers if there is a tie list all users that are a part of the tie.

5. A tweet is considered highly influential if its sentiment is high and the user that tweeted it has a high number of followers. Define a VIEW called highly influential and then display the tweets that are highly influential