

# INFSCI 2710 Database Management, Fall 2022

## Homework 2: SQL

100 pts

Due Date: 10/13. Please submit a pdf to the Canvas assignment.

Note: is no actual data provided in this assignment, you can assume there are plenty row of data been inserted to the database, **you just need to provide MYSQL statements**

Consider the following relational database from Yelp

### Business

- Id
- Name
- Address
- City
- State
- Category: business category
- Stars: average stars from all reviewed users
- Review\_count: total number of reviews
- Is\_open: a Boolean indicate if it's still open or not

### User

- Id
- Name
- Review\_count number of reviews this user wrote
- Yelping\_since: the date that user registered

- Fans: total number of fans

## Review

- Id
- Stars
- Date
- Text
- Useful: total number of 'useful' rated by other users
- Funny: total number of 'funny' rated by other users
- Cool: total number of 'cool' rated by other users
- Business\_id
- User\_id

## Tip

- User\_id
- Business\_id
- Text
- Date
- amount

**Q1 [5 pt]** write a SQL query to find name of all the business in the 'fast food' category

**Q2 [10 pt]** write a SQL query to find the top 3 users based on total number of fans

**Q3 [10 pt]** write a SQL query to find the number of users registered after 2020

**Q4 [10 pt]** write a SQL query to list the cities with the most reviews in descending order

**Q5 [10 pt]** write SQL to update stars and review\_counts for business named " yami sushi"

**Q6 [10 pt]** write a SQL query to find the distribution of star ratings to the business in the Pittsburgh (return stars and corresponding counts)

**Q7 [10 pt]** for the restaurant named “yami sushi” write a SQL query to find if there are more reviews with the word ‘love’ or with the word ‘hate’

**Q8 [10 pt]** for each user, calculate average star, average number of “useful”, “funny” and “cool” for all the reviews this user wrote

**Q9 [10 pt]** Find the user who has the largest number of fans, return the name and amount for the top 5 businesses he tipped

**Q10 [15 pt]** Group business based on the one that are open and the ones that are closed, what differences you can find between those two groups (you can use different SQL to answer different sub questions)

1. Are there more reviews written for the business that are still open
2. Are the average star rating higher for business that are open than business that are closed
3. Are there more tips for business that are open than business that are closed