



Northeastern University, Khoury College of Computer Science

## CS 6220 Data Mining | Assignment 1

Due: September 19, 2023(100 points)

Chun Wei Tseng

<https://github.com/Chun-Wei-Tseng/CS6220-Assignment-1>

### Question X [20 pts]

1. Cardinality in basket\_data.csv file: 21

2. [['ham', 'cheese'], ['ham', 'bread'], ['cheese', 'bread']]

3a. Total records of movie ratings in the entire dataset: 100480507

3b. Total number of unique users 480189

3c. Range of years that this data is valid over: 1999 ~ 2005

4a. Number of movies with unique name: 17359

4b. Number of movie names referring to four different movies: 5

5a. Number of users rated exactly 200 movies: 605

5b. Names of the movies that they (User ID: 508) liked the most: ['High Fidelity', 'Monty Python's The Meaning of Life: Special Edition', 'American Beauty', 'Roger & Me', 'Eternal Sunshine of the Spotless Mind', 'Being John Malkovich', 'Vietnam: A Television History', 'Super Size Me', 'Lord of the Rings: The Fellowship of the Ring', 'This Is Spinal Tap', 'The Pianist', 'The Silence of the Lambs', 'Sideways', 'Whale Rider', 'Garden State', 'Bowling for Columbine', 'Gandhi', 'Apocalypse Now Redux', 'To Die For', 'Monty Python's Life of Brian', 'The Manchurian Candidate', 'Memento', 'Amelie', 'Apocalypse Now', 'The Usual Suspects', 'Lord of the Rings: The Two Towers: Extended Edition', 'The Lord of the Rings: The Fellowship of the Ring: Extended Edition', 'Touching the Void', 'Minority Report', 'The Royal Tenenbaums', 'Election', 'Good Will Hunting', 'L.A. Confidential', 'Taxi Driver', 'Lord of the Rings: The Two Towers', 'Cabaret', 'Adaptation', 'The Accused', 'Lost in Translation', 'Boys Don't Cry', 'To Be and To Have', 'Schindler's List', 'Raging Bull', 'Lord of the Rings: The Return of the King', 'Monty Python and the Holy Grail', 'Raising Arizona', 'The Shawshank Redemption: Special Edition', 'Harold and Maude', 'Downfall', 'Lord of the Rings: The Return of the King: Extended Edition', 'Monster', 'Band of Brothers', 'Three Kings', 'Unforgiven', 'Maria Full of Grace', 'Days of Wine and Roses', 'Shakespeare in Love']