CSE 414 Database Systems

Section 1: More SQL practice,

Intro to Homework 2

TA: Daseul Lee (dslee@cs)

Homework 2: SQL

- You will get to work on a large dataset: imdb
  - Note that imdb dataset is not perfectly clean
- Goals:
  - Create tables and import imdb database
  - Write SQL queries that answer each question
    - Avoid subqueries
  - Create up to 10 indexes
    - Which index? Up to you, but aim for ~1 min runtime

So, what is an Index?

- SQL queries run fast on small tables, but not on large tables or files.

- To improve performance on a large data set, we can use index.

- Index is one or more attributes chosen as a search key for a faster look-up on records.
  - Here, search key does not necessarily mean primary key.

- More on Friday's lecture