

Assignment 2

SQL Statements (10%) – INDIVIDUAL PROJECT

The Case

The *NetFilms* company aims to be a new player in the online film streaming space. The central component of their software is a database which supports two broad functions: (1) to make available to customers a searchable list of available films, (2) to track the customers' viewing and rating of those films.

The database stores the title, genre, classification and release year of a large number of films, allowing customers to search for films using these attributes. The database stores also the actors, writers and directors of the films so that customers can search for films starring, or written or directed by, cast and crew members of interest. We must also store the wholesale price at which *NetFilms* bought the rights to each film, and the price-per-view we will charge to customers – these two amounts can be different for each film, and allow us to calculate the profit being made on particular films.

For each customer we store their name, username and password, and email address, as well as their postcode so that we can generate statistics like “customers who live near you liked these films...”. Customers are charged each time they watch a film, so we need to track views. When customers watch a film, they may rate it if they wish (out of ten, where ten is the highest). Ratings are stored so we can display rankings which guide customer choices.

Your Task

The *NetFilms* database has already been designed and implemented using MySQL. The data model is shown below. Your job is to write SQL code that the company's application software will execute. Below you will find the specifications for 15 scripts that you must write.

Setup scripts

To help you program your scripts you are provided with a test version of the database, which contains part of the film library, and a small amount of ‘dummy’ customer data. To set up the database in your MySQL server, download the file *NetFilms.sql* from LMS and run it in Workbench. This script creates the schema and database tables and populates them with data.

Submission

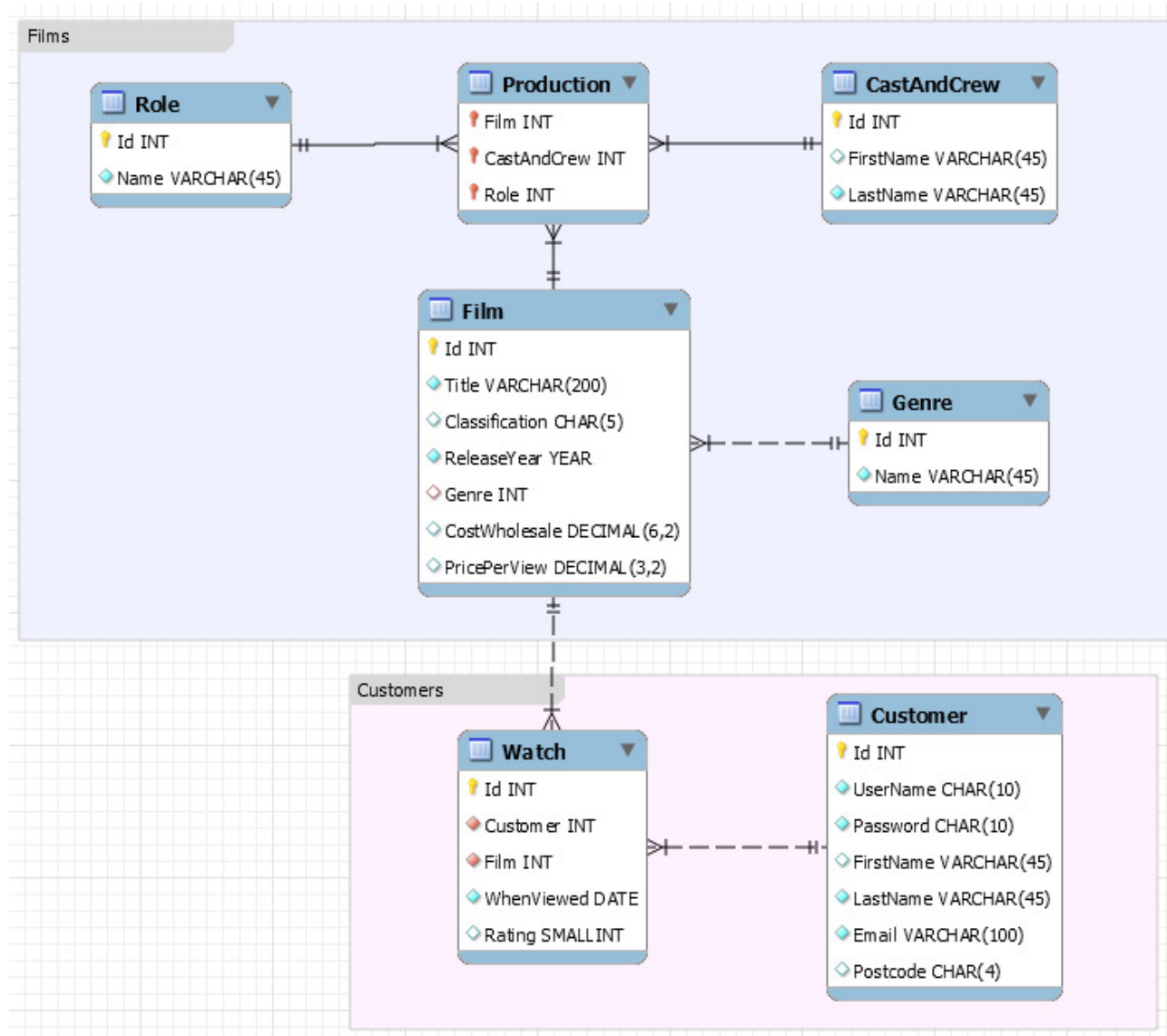
This assignment is worth 10% of your marks in the subject. Your work will be assessed on the correctness and simplicity of the SQL that you write. (A correct answer that is more complex than it needs to be may not achieve full marks, even though it produces the correct results.)

Submit your assignment to the Assessment page on LMS as a single PDF, by midnight on the due date of Monday 11th May.

Please ensure that you place your username and student number at the top of every page of your submission. Other formatting requirements are listed below.

Data model

The following schema describes the database which has been implemented.



The SQL queries required

In this section are listed 15 questions for you to answer. Write one SQL Select statement per question.

Beside each question is a maximum mark which reflects the difficulty of that question. The marks total to 60: this will be scaled down to a mark out of 10.

Where it would improve the readability of your output, use 'order by', and column aliases. A small proportion of the marks will be devoted to this.

1. How many films does *NetFilms* offer? (2 marks)
2. List all films of genre 'Documentary' (2 marks)
3. How many films in our list were released during the 1990s? (2 marks)
4. List all of our R-rated comedy films. Show the title and year of release of each film, and order by title. (3 marks)
5. List all films directed by Steven Spielberg (4 marks)
6. List all of the directors and how many films each has made. (5 marks)
7. Who has acted in the most films? (6 marks)
8. Which are the 5 most-watched films? (4 marks)
9. Which customers have not viewed any films yet? (2 marks)
10. On which day of the week do most people watch films? (3 marks)
11. How many different films have been watched by customers in postcode 3053? (4 marks)
12. List the top 8 films as ranked by our customers. For each, show the title, and average rank, and order the list by rank. Only include films that have been ranked by at least 5 customers. (5 marks)
13. Which films has customer Steven Kloves watched? List the films along with how many times he has watched them. (5 marks)
14. Which genres of films have been watched the most in postcode 3053? List the top 3 genres for that postcode. (6 marks)
15. Which films are making a profit for *NetFilms*? (Profit is the net takings from viewings minus the wholesale cost of the film.) (7 marks)

Formatting requirements for your submission

Submit a single PDF showing your answers to the 15 questions.

For each question, present an answer in the following format:

- Show the question number and question.
- Show your answer (the SQL statement) in **blue text (not a screen shot)**
- Show a screenshot from Workbench showing output of 10 or less lines.
- Show how many rows were returned, in **red text**

For example:

7. List all cast and crew with the last name 'Altman'

```
select * from CastAndCrew  
where LastName = 'Altman';
```

Id	FirstName	LastName
1376	Robert	Altman
2880	Mark	Altman
10308	Robert	Altman
21472	Harry	Altman
21722	Robert	Altman

5 Rows Returned

Please ensure that you place your username and student number at the top of every page of your submission.

Academic Honesty

A reminder that this assignment should be your own work. Plagiarism - the copying of another's work without proper acknowledgment - is not permitted. Nor is allowing another person to copy your work. Work submitted for assessment purposes must be the independent work of the student concerned (or where joint work is permitted, of the students concerned). Please refer to <http://academichonesty.unimelb.edu.au/> for details about academic honesty.

Academic misconduct occurs when students portray someone else's work as their own. There are many ways in which academic misconduct can occur. Some of these are:

- Sham Paraphrasing: Material copied verbatim from text and source acknowledged in-line but represented as paraphrased.
- Illicit Paraphrasing: Material paraphrased from text without in-line acknowledgement of source.
- Other Plagiarism: Material copied from another student's assignment with the knowledge of the other student.
- Verbatim Copying: Material copied verbatim from text without in-line acknowledgement of the source.
- Recycling: Same assignment submitted more than once for different subjects.
- Ghost Writing: Assignment written by third party and represented as own work.
- Purloining: Assignment copied from another student's assignment or other person's papers without that person's knowledge.

The University is committed to graduating students with "a profound respect for truth, and for the ethics of scholarship... we want our graduates to be capable of independent thought, to be able to do their own work, and to know how to acknowledge the work of others" (Prof. Peter McPhee). As such, the university takes a dim view of students who are not able to correctly acknowledge the work of others, or who try to pass this work off as their own. All students should be aware of the following web site: <http://academichonesty.unimelb.edu.au/advice.html> which provides practical advice to students about how not to be involved in academic misconduct. In particular look at the Quick Checklist section (paraphrased below).

Quick checklist (Reproduced with permission of Dr Stephen Morgan, Faculty of Economics and Commerce)

1. To be certain to acknowledge sources fairly and avoid plagiarising, review this checklist before beginning to write your essay and again after you have completed your first draft.
2. What type of source are you using: your own independent material, common knowledge, or someone else's independent material?
3. If you are quoting someone else's material, is the quotation exact? Have you used quotation marks for quotations run into the text? Have you set off block quotes with an extra space before and after the quote, single spacing within the quote, and left indenting of all lines of the block quote? Are omissions shown with ellipses and additions with square brackets?
4. If you are paraphrasing someone else's material, have you rewritten it in your own words and sentence structures? Does your paraphrase employ quotation marks when you resort to the author's exact language? Have you represented the author's meaning without distortion?
5. Have you acknowledged each use of someone else's material?
6. Do all references contain complete and accurate information on the sources you have cited?