Individual Project 4

**DS160-01**

**Introduction to Data Science**

**Spring 2022**

**Writing SQL Queries(50 points)**

**Goal:** The goal of this project is to provide you practice for writing simple SQL queries.

**Instructions:** Create a new .sql file titled **IP4\_XXX.sql**, where **XXX** are your initials; and rename this file to **IP4\_XXX.docx**. Also create a GitHub repository titled **IP4\_XXX** to which you can push both your .sql file and this document.

Attached to the assignment are two .sql files. HenryBooksCreate.sql which will create the database on you MySQL server and HenryBooksFill.sql which will fill it with data.

You will form the following queries for the Henry Books database and provide the statement and the output in this document. All of your SQL statements should also be in the .sql file. Use a comment (# mark) in your .sql file to number the statements according to this document. Each problem is marked with its point worth.

1. Retrieve the first name and last name of each author in the author relation. Order does not matter. **(1 points)**

**SQL Statement:**

SELECT firstName, lastName

FROM author;

**Output:**

'Toni', 'Morrison'

'Paul', 'Solotaroff'

'Vernor', 'Vintage'

'Dick', 'Francis'

'Peter', 'Straub'

'Stephen', 'King'

1. Retrieve the title and book type in the book relation. Order does not matter. **(1 points)**

**SQL Statement:**

SELECT title, TYPE

From book;

**Output:**

'A Deepness in the Sky', 'SFI'

'Magic Terror', 'HOR'

'The Stranger', 'FIC'

'Venice', 'ART'

'Second Wind', 'MYS'

'The Edge', 'MYS'

1. Retrieve the publisherCode in the book relation. List each publisherCode only once in the result. Order does not matter. **(1 points)**

**SQL Statement:**

SELECT publisherCode

FROM book

GROUP BY publisherCode;

**Output:**

'BA'

'BP'

'BY'

'CT'

'FA'

'FS'

'HC'

'JP'

'LB'

'PE'

'PL'

'PU'

'RH'

'SC'

'SS'

'ST'

'TA'

'TB'

'TO'

'VB'

'WP'

1. Retrieve the title and price of each book in the book relation. Further add a calculated column named ‘discount’ that shows the price the book with 25% discount. Order does not matter. Show the first five rows of the result (**LIMIT 5**). **(3 points)**

**SQL Statement:**

SELECT title,price, price\*.75 as discount

FROM book

LIMIT 5;

**Output:**

'A Deepness in the Sky', '7.19', '5.3925'

'Magic Terror', '7.99', '5.9925'

'The Stranger', '8.00', '6.0000'

'Venice', '24.50', '18.3750'

'Second Wind', '24.95', '18.7125'

1. Retrieve the title and price for any book whose price is higher than $20.00 in the book relation. Show the full result. **(3 points)**

**SQL Statement:**

SELECT title,price

FROM book

WHERE price>20;

**Output:**

'Venice', '24.50'

'Second Wind', '24.95'

'Treasure Chests', '24.46'

'Van Gogh and Gauguin', '21.00'

'A Guide to SQL', '37.95'

1. Retrieve the publisherName of all publishers that are in New York only in the publisher relation. Order does not matter. **(2 points)**

**SQL Statement:**

SELECT publisherName

FROM publisher

WHERE city='New York';

**Output:**

'Arcade Publishing'

'Back Bay books'

'Fawcett books'

'Farrar Straus and Giroux'

'HarperCollins publishers'

'Jove Publications'

'Lb books'

'Penguin USA'

'Plume'

'Putnam Publishing Group'

'Random House'

'Schoken books'

'Scribner'

'Simon and Schuster'

'SchoLASTic Trade'

'Tor books'

'Thames and Hudson'

'Vintage books'

'W.W. Norton'

1. Retrieve the publisherName of all publishers that are not in New York in the publisher relation. (use != for inequality). Order does not matter. Show the full result. **(3 points)**

**SQL Statement:**

SELECT publisherName

FROM publisher

WHERE city!='New York';

**Output:**

'Arkham House'

'Basic books'

'Berkley Publishing'

'Course Technology'

'Jeremy P. Tarcher'

'McPherson and Co.'

'Taunton Press'

'Touchstone books'

'Westview Press'

1. Retrieve the bookCode and onHand for each book for which a branch has between 2 and 4 copies in the inventory relation. **Use the BETWEEN keyword in this query.** Order does not matter. Show the full result. **(3 points)**

**SQL Statement:**

SELECT bookCode, onHand

FROM inventory

WHERE onHand between 2 and 4;

1. Retrieve a count of the number of books published by Penguin USA. Name the column ‘Penguin Books’. Order does not matter. Show your full result. **(3 points)**

**SQL Statement:**

SELECT count(title) AS PenguinBooks

FROM book

where publisherCode='PE';

**Output:**

'4'

1. Retrieve the number of books in the book relation whose prices is $20.00 or lower. Order does not matter. Show your full result. **(3 points)**

**SQL Statement:**

SELECT count(title)

FROM book

where price<=20;

**Output:**

‘28’

1. Retrieve all of the columns from the book and publisher relations in one result. Use aliases in your query and use the simple JOIN syntax (WHERE clause). Order does not matter. **(5 points)**

**SQL Statement:**

SELECT \*

FROM book, publisher

WHERE book.publisherCode=publisher.publisherCode;

**Output:**

'8092', 'Godel, Escher, Bach', 'BA', 'PHI', '14.00', 'Y', 'BA', 'Basic books', 'Boulder CO'

'3350', 'Group: Six People in Search of a Life', 'BP', 'PSY', '10.40', 'Y', 'BP', 'Berkley Publishing', 'Boston'

'3906', 'The Soul of a New Machine', 'BY', 'SCI', '11.16', 'Y', 'BY', 'Back Bay books', 'New York'

'669X', 'A Guide to SQL', 'CT', 'CMP', '37.95', 'Y', 'CT', 'Course Technology', 'Boston'

'0189', 'Magic Terror', 'FA', 'HOR', '7.99', 'Y', 'FA', 'Fawcett books', 'New York'

'2908', 'Electric Light', 'FS', 'POE', '14.00', 'N', 'FS', 'Farrar Straus and Giroux', 'New York'

'9931', 'To Kill a Mockingbird', 'HC', 'FIC', '18.00', 'N', 'HC', 'HarperCollins publishers', 'New York'

'0808', 'The Edge', 'JP', 'MYS', '6.99', 'Y', 'JP', 'Jove Publications', 'New York'

'8720', 'When Rabbit Howls', 'JP', 'PSY', '6.29', 'Y', 'JP', 'Jove Publications', 'New York'

'9882', 'Slay Ride', 'JP', 'MYS', '6.99', 'Y', 'JP', 'Jove Publications', 'New York'

'3743', 'Nine Stories', 'LB', 'FIC', '5.99', 'Y', 'LB', 'Lb books', 'New York'

'6908', 'Franny and Zooey', 'LB', 'FIC', '5.99', 'Y', 'LB', 'Lb books', 'New York'

'9883', 'The Catcher in the Rye', 'LB', 'FIC', '5.99', 'Y', 'LB', 'Lb books', 'New York'

'2766', 'Of Mice and Men', 'PE', 'FIC', '6.95', 'Y', 'PE', 'Penguin USA', 'New York'

'5163', 'Travels with Charley', 'PE', 'TRA', '7.95', 'Y', 'PE', 'Penguin USA', 'New York'

'7405', 'East of Eden', 'PE', 'FIC', '12.95', 'Y', 'PE', 'Penguin USA', 'New York'

'9701', 'The Grapes of Wrath', 'PE', 'FIC', '13.00', 'Y', 'PE', 'Penguin USA', 'New York'

'138X', 'Beloved', 'PL', 'FIC', '12.95', 'Y', 'PL', 'Plume', 'New York'

'6128', 'Jazz', 'PL', 'FIC', '12.95', 'Y', 'PL', 'Plume', 'New York'

'9627', 'Song of Solomon', 'PL', 'FIC', '14.00', 'Y', 'PL', 'Plume', 'New York'

'079X', 'Second Wind', 'PU', 'MYS', '24.95', 'N', 'PU', 'Putnam Publishing Group', 'New York'

'9611', 'Black House', 'RH', 'HOR', '18.81', 'N', 'RH', 'Random House', 'New York'

'1351', 'Dreamcatcher: A Novel', 'SC', 'HOR', '19.60', 'N', 'SC', 'Scribner', 'New York'

'5790', 'Catch-22', 'SC', 'FIC', '12.00', 'Y', 'SC', 'Scribner', 'New York'

'0378', 'Venice', 'SS', 'ART', '24.50', 'N', 'SS', 'Simon and Schuster', 'New York'

'2226', 'Harry Potter and the Prisoner of Azkaban', 'ST', 'SFI', '13.96', 'N', 'ST', 'SchoLASTic Trade', 'New York'

'7443', 'Harry Potter and the Goblet of Fire', 'ST', 'SFI', '18.16', 'N', 'ST', 'SchoLASTic Trade', 'New York'

'1382', 'Treasure Chests', 'TA', 'ART', '24.46', 'N', 'TA', 'Taunton Press', 'Newtown CT'

'0180', 'A Deepness in the Sky', 'TB', 'SFI', '7.19', 'Y', 'TB', 'Tor books', 'New York'

'6328', 'Band of Brothers', 'TO', 'HIS', '9.60', 'Y', 'TO', 'Touchstone books', 'Westport CT'

'0200', 'The Stranger', 'VB', 'FIC', '8.00', 'Y', 'VB', 'Vintage books', 'New York'

'7559', 'The Fall', 'VB', 'FIC', '8.00', 'Y', 'VB', 'Vintage books', 'New York'

'2281', 'Van Gogh and Gauguin', 'WP', 'ART', '21.00', 'N', 'WP', 'Westview Press', 'Boulder CO'

1. Rewrite the previous query using the ON keyword. **(5 points)**

**SQL Statement:**

SELECT \*

FROM book

INNER JOIN publisher ON book.publisherCode=publisher.publisherCode;

**Output:**

'8092', 'Godel, Escher, Bach', 'BA', 'PHI', '14.00', 'Y', 'BA', 'Basic books', 'Boulder CO'

'3350', 'Group: Six People in Search of a Life', 'BP', 'PSY', '10.40', 'Y', 'BP', 'Berkley Publishing', 'Boston'

'3906', 'The Soul of a New Machine', 'BY', 'SCI', '11.16', 'Y', 'BY', 'Back Bay books', 'New York'

'669X', 'A Guide to SQL', 'CT', 'CMP', '37.95', 'Y', 'CT', 'Course Technology', 'Boston'

1. Retrieve the title from the book relation and the city from the publisher relation using a JOIN query. Use aliases in your query. Order the result by title. **(5 points)**

**SQL Statement:**

SELECT title, city

FROM book

join publisher on book.publisherCode=publisher.publisherCode;

**Output:**

'Godel, Escher, Bach', 'Boulder CO'

'Group: Six People in Search of a Life', 'Boston'

'The Soul of a New Machine', 'New York'

'A Guide to SQL', 'Boston'

'Magic Terror', 'New York'

'Electric Light', 'New York'

'To Kill a Mockingbird', 'New York'

'The Edge', 'New York'

1. Retrieve the title from the book relation and branchNum and onHand from the inventory relation. Use aliases in your query. Order the result by title. **(5 points)**

**SQL Statement:**

SELECT title as Title, branchNum as Number\_On\_Branch, onHand as Amount\_Availible

FROM book

join inventory on book.bookCode=inventory.bookCode;

**Output:**

'A Deepness in the Sky', '1', '2'

'Magic Terror', '2', '2'

'The Stranger', '1', '1'

'The Stranger', '2', '3'

'Venice', '3', '2'

'Second Wind', '2', '1'

'Second Wind', '3', '2'

'Second Wind', '4', '3'

1. Retrieve the title from the book relation and compute the number of copies of the title that all branches have on hand. Name this computed column ‘Inventory’ **Hint: You will need to join book and inventory and do an aggregate query.** Use aliases in your query. Order the result by the total number of copies of the book in descending order. Show the first two rows of your result. **(5 points)**

**SQL Statement:**

SELECT title as Title, sum(onHand) as total\_inventory

FROM book

join inventory on book.bookCode=inventory.bookCode

group by title;

**Output:**

'A Deepness in the Sky', '2'

'Magic Terror', '2'

'The Stranger', '4'

'Venice', '2'

'Second Wind', '6'

'The Edge', '1'

'Dreamcatcher: A Novel', '6'

1. Retrieve the first name and last name from the author relation and the title from the book relation for all paperback books in the book relation. Order the result by the author last name and title. **(5 points)**

**SQL Statement:**

SELECT firstName, lastName, title

FROM author,book

WHERE paperback='Y'

ORDER BY lastName,title;

**Output:**

'Stephen E.', 'Ambrose', 'A Deepness in the Sky'

'Stephen E.', 'Ambrose', 'A Guide to SQL'

'Stephen E.', 'Ambrose', 'Band of Brothers'

'Stephen E.', 'Ambrose', 'Beloved'

'Stephen E.', 'Ambrose', 'Catch-22'

'Stephen E.', 'Ambrose', 'East of Eden'

'Stephen E.', 'Ambrose', 'Franny and Zooey'

'Stephen E.', 'Ambrose', 'Godel, Escher, Bach'

'Stephen E.', 'Ambrose', 'Group: Six People in Search of a Life'

**BONUS (5 points):**

Retrieve the title from the book relation and the author lastName from the author relation. Order by author lastName. Use aliases in your query. **This will involve JOINING the book, author and wrote relations.** **(5 points)**

**SQL Statement:**

SELECT title as Title, lastName as Author\_Last\_Name

FROM book

JOIN wrote on book.bookCode=wrote.bookCode

JOIN author on wrote.authorNum=author.authorNum

ORDER BY lastName;

**Output:**

'Band of Brothers', 'Ambrose'

'The Stranger', 'Camus'

'The Fall', 'Camus'

'When Rabbit Howls', 'Chase'

'Van Gogh and Gauguin', 'Collins'

'Van Gogh and Gauguin', 'Collins, Jr.'

'Second Wind', 'Francis'

'The Edge', 'Francis'

'Slay Ride', 'Francis'