

# CSCI-UA.60-1-REAL Midterm-B

Fenglei Gu

TOTAL POINTS

**89.5 / 90**

QUESTION 1

## 1 Problem 1 - Python output 15 / 15

✓ - 0 pts Correct

- 1 pts Output: either only one row or too many rows of stars
- 2 pts Output: First 2 rows of stars missing
- 3 pts Output: Years are not listed.
- 2 pts Output: years are not listed in chronological order
- 1 pts Output: one or several years missing
- 2 pts Output: rows with data use dashes
- 2 pts Output: 1953-1957 Two or more rows are incorrect.
- 3 pts Output: 1953-1957: missing the correct 4 values per line
- 3 pts Output: 1953-1957 - no values at all
- 2 pts Output: 1953-1957: two of the columns are incorrect
- 1 pts Output: 1953-1957: one of the columns is incorrect or there is an additional column
- 3 pts Output: 1950-1952,1958-1959 lines are not marked with 'n/a'
- 1 pts Output: 1950-1952,1958-1959 - only some lines are marked with 'n/a'
- 3 pts Output: Console output incorrect or missing.
- 0.5 pts Output: Console output number of records is not 12 or does not match output.
- 0.5 pts Output: Console does not include 'records written'.
- 12 pts Answer started but incomplete; partial credit
- 15 pts Answer incorrect or not provided.

QUESTION 2

## 2 Problem 2 - UNION 3 / 3

✓ - 0 pts Correct

- 1.5 pts Definition missing or incorrect.

- 1.5 pts Example missing or incorrect.

- 0.5 pts Example syntax or programming error or incomplete statement; partial credit.

- 3 pts No answer provided or incorrect answer.

QUESTION 3

## 3 Problem 2 - .mode csv 3 / 3

✓ - 0 pts Correct

- 1.5 pts Definition missing or incorrect.

- 1.5 pts Example missing or incorrect.

- 0.5 pts Example syntax or programming error or incomplete statement; partial credit.

- 3 pts No answer provided or incorrect answer.

QUESTION 4

## 4 Problem 2 - .import 3 / 3

✓ - 0 pts Correct

- 1.5 pts Definition missing or incorrect.

- 1.5 pts Example missing or incorrect.

- 0.5 pts Example syntax or programming error or incomplete statement; partial credit.

- 3 pts No answer provided or incorrect answer.

QUESTION 5

## 5 Problem 2 - LIKE 3 / 3

✓ - 0 pts Correct

- 1.5 pts Definition missing or incorrect.

- 1.5 pts Example missing or incorrect.

- 0.5 pts Example syntax or programming error or incomplete statement; partial credit.

- 3 pts No answer provided or incorrect answer.

QUESTION 6

## 6 Problem 2 - UPDATE 2.5 / 3

- 0 pts Correct

- **1.5 pts** Definition missing or incorrect.
- **1.5 pts** Example missing or incorrect.
- ✓ - **0.5 pts** Example syntax or programming error or incomplete statement; partial credit.
- **3 pts** No answer provided or incorrect answer.

#### QUESTION 7

### 7 Problem 3-1 3 / 3

- ✓ - **0 pts** Correct
- **1.5 pts** SELECT statement missing or incorrect.

#### QUESTION 8

### 8 Problem 3-2 3 / 3

- ✓ - **0 pts** Correct
- **1 pts** SELECT clause missing or incorrect; or incorrect or missing usage of COUNT()
- **0.5 pts** FROM clause missing or incorrect.
- **1 pts** GROUP BY clause missing or incorrect.
- **0.5 pts** ORDER BY clause missing or incorrect.
- **3 pts** Answer incorrect or not provided.

#### QUESTION 9

### 9 Problem 3-3 3 / 3

- ✓ - **0 pts** Correct
- **1.5 pts** SELECT clause missing or incorrect; or incorrect or missing usage of concatenation
- **0.5 pts** FROM clause missing or incorrect.
- **1.5 pts** WHERE clause missing or incorrect.
- **1 pts** Additional clause e.g. GROUP BY not needed or incorrectly used.
- **1 pts** Clauses in the wrong order or other syntax error.
- **3 pts** Answer incorrect or not provided.

#### QUESTION 10

### 10 Problem 3-4 3 / 3

- ✓ - **0 pts** Correct
- **1 pts** SELECT clause missing or incorrect; or incorrect or missing usage of COUNT() or missing Borough
- **0.5 pts** FROM clause missing or incorrect.
- **1 pts** GROUP BY clause missing or incorrect.

- **0.5 pts** ORDER BY clause missing or incorrect.
- **1 pts** LIMIT clause incorrect or not provided.
- **1 pts** Additional clause incorrect e.g. HAVING
- **3 pts** Answer incorrect or not provided.

#### QUESTION 11

### 11 Problem 3-5 3 / 3

- ✓ - **0 pts** Correct
- **1 pts** SELECT clause missing or incorrect
- **0.5 pts** FROM clause missing or incorrect.
- **1 pts** WHERE clause missing or incorrect.
- **0.5 pts** ORDER BY clause missing or incorrect.
- **1 pts** Additional clause that is not necessary (e.g. GROUP BY) or clauses are in the wrong order.
- **3 pts** Answer incorrect or not provided.

#### QUESTION 12

### 12 Problem 3-6 3 / 3

- ✓ - **0 pts** Correct
- **0.5 pts** SELECT clause missing or incorrect; or incorrect or missing usage of COUNT()
- **0.5 pts** FROM clause missing or incorrect.
- **1 pts** WHERE clause missing or incorrect; or incorrectly uses HAVING
- **1 pts** GROUP BY clause missing or incorrect.
- **0.5 pts** ORDER BY clause missing or incorrect.
- **1 pts** Order of the clauses is incorrect
- **3 pts** Answer incorrect or not provided.

#### QUESTION 13

### 13 Problem 3-7 3 / 3

- ✓ - **0 pts** Correct
- **1 pts** SELECT clause missing or incorrect; or incorrect or missing usage of COUNT() or COUNT() without GROUP BY
- **0.5 pts** FROM clause missing or incorrect.
- **1.5 pts** WHERE clause missing or incorrect; includes missing or incorrect parentheses.
- **0.5 pts** Syntax error
- **1 pts** Additional clause or clauses in the wrong order
- **3 pts** Answer incorrect or not provided.

#### QUESTION 14

#### 14 Problem 3-8 3 / 3

##### ✓ - 0 pts Correct

- 1 pts SELECT clause missing or incorrect; or incorrect or missing usage of COUNT()
- 0.5 pts FROM clause missing or incorrect.
- 0.5 pts GROUP BY clause missing or incorrect.
- 1.5 pts HAVING clause missing or incorrect
- 0.5 pts ORDER BY clause missing or incorrect.
- 1 pts The order of the clauses is incorrect.
- 3 pts Answer incorrect or not provided.

#### QUESTION 15

#### 15 Problem 3-9 4 / 4

##### ✓ - 0 pts Correct

- 1 pts SELECT clause missing or incorrect; or incorrect or missing usage of COUNT()
- 0.5 pts FROM clause missing or incorrect.
- 1.5 pts WHERE clause missing, incomplete or incorrect.
- 1 pts Additional clause e.g. GROUP BY, ORDER BY which is not needed
- 1 pts Incorrect order of the clauses or additional clauses that are not needed or incorrect.
- 0.5 pts Syntax: SQLite uses parentheses with aggregate functions, not square brackets.
- 3 pts Answer incorrect or not provided.

#### QUESTION 16

#### 16 Problem 3-10 4 / 4

##### ✓ - 0 pts Correct

- 1 pts SELECT clause missing or incorrect; or incorrect or missing usage of COUNT()
- 0.5 pts FROM clause missing or incorrect.
- 1 pts WHERE clause missing or incorrect.
- 1 pts Additional clause e.g. ORDER BY which is not needed or incorrect use of GROUP BY
- 0.5 pts Syntax: SQLite uses parentheses with aggregate functions, not square brackets.
- 4 pts Answer incorrect or not provided.

#### QUESTION 17

#### 17 Problem 3-11 4 / 4

##### ✓ - 0 pts Correct

- 1.5 pts SELECT clause missing or incorrect; or incorrect usage of DISTINCT
- 0.5 pts FROM clause missing or incorrect.
- 1 pts ORDER BY clause missing or incorrect
- 1 pts Additional clause e.g. WHERE or UNION used incorrectly or not needed
- 1 pts GROUP BY used incorrectly in place of DISTINCT.
- 0.5 pts Syntax: SQLite uses parentheses with aggregate functions, not square brackets.
- 4 pts Answer incorrect or not provided.

#### QUESTION 18

#### 18 Problem 3-12 4 / 4

##### ✓ - 0 pts Correct

- 2 pts Description missing
- 2 pts SQLite query missing or incorrect
- 1 pts Syntax or other errors in the query; partial credit.
- 2 pts SQLite query does not match the description.
- 4 pts No solution provided.

#### QUESTION 19

#### 19 Problem 4-a 1NF definition 5 / 5

##### ✓ - 0 pts Correct

- 1 pts Partial credit ; definition incomplete.
- 2 pts Partial credit ; definition and examples incomplete.
- 3 pts Partial credit; valid table not described and only one rule cited.
- 5 pts Answer incorrect or not provided.

#### QUESTION 20

#### 20 Problem 4-a 1NF examples 5 / 5

##### ✓ - 0 pts Correct

- 1 pts Answer incomplete; partial credit.
- 5 pts Answer missing or incorrect. For example, while some of the data demonstrates a valid table, there are examples where the data do not.

QUESTION 21

21 Problem 4-b 2NF definition 5 / 5

✓ - 0 pts Correct

- 1 pts Partial credit ; definition incomplete.
- 2 pts Partial credit ; definition and examples incomplete.
- 3 pts Partial credit; valid table not described and only one rule cited.
- 5 pts Answer incorrect or not provided.

QUESTION 22

22 Problem 4-b 2NF examples 5 / 5

✓ - 0 pts Correct

- 1 pts Answer incomplete; partial credit.
- 4 pts Answer addresses only 1NF but not 2NF.
- 5 pts Answer missing or incorrect. For example, while some of the data demonstrates a valid table, there are examples where the data do not.

Your Name: Fenglei Gu

CSCI-UA.60-1 – Database Design and Web Implementation  
Midterm Exam  
Fall, 2019 – Version “B”  
Deena Engel

Your Name:

Fenglei Gu

Your Netid / NYU email address:

fg1121

Which row are you in?

1

Who is sitting to your left?

Zi chuan Wang

Who is sitting to your right?

Iva Porfiriova (Sophia)

CSCI-UA.60-1 Midterm Exam – Fall, 2019 – Version “B”

Your Name: \_\_\_\_\_

1. Python and data files (15 points) [Source: [https://data.giss.nasa.gov/csi-bin/gistemp/stdata\\_show.cgi?id=425003058010&ds=5&dt=1](https://data.giss.nasa.gov/csi-bin/gistemp/stdata_show.cgi?id=425003058010&ds=5&dt=1) ]

Given the following input as **data\_B.txt** and the following python script:

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	J-D	D-N	DJF	MAM	JJA	SON	Year
1950	-29	-29	-7	-19	-11	-10	-15	-22	-12	-17	-33	-19	-19	-25	-12	-16	-21	1950	
1951	-33	-43	-20	-9	3	-5	0	7	5	10	-2	12	-7	-32	-9	0	4	1951	
1952	13	11	5	0	-2	7	4	6	-2	-15	-2	1	3	12	-1	3	-4	1952	
1953	11	20	13	19	10	0	6	3	7	-3	9	8	10	14	5	4	1953		
1954	-22	-10	13	-15	-19	-14	-17	-14	-7	-1	9	-11	-10	-8	-16	-15	1	1954	
1955	16	-16	32	-18	-20	-10	-8	8	-10	-3	-26	-31	-12	-23	-3	-13	1955		
1956	-16	-23	19	-24	-25	-14	-10	-27	-15	-22	-15	-9	-18	-20	-24	-23	-17	1956	
1957	-12	-7	-8	-1	5	16	0	19	8	4	0	16	4	-9	-1	12	5	1957	
1958	37	24	9	2	5	-12	0	-8	-7	2	1	-1	4	6	26	-7	-1	1958	
1959	8	6	17	14	2	2	4	-1	-5	-9	1	3	2	5	11	2	-7	1959	

```

def midterm():
    source_file = "data_B.txt"
    target_file = "data_B_out.txt"
    s = open(source_file, 'r')
    t = open(target_file, 'w')
    c = 0
    for line in s:
        a = line.rstrip('\n')
        if ( a[0:4].isdigit() ):
            v = a.split()
            if ( int(a[3:4]) > 2 and int(a[3:4]) < 8 ):
                v = a.split()
                n = v[0] + ' ' + '19' + v[1]
                for i in range(3,13,3):
                    n += v[i] + '-'
                n += '\n'
            else:
                n = v[0] + ': ' + n/a + '\n'
        else:
            n = '*' * 4 + '\n'
        t.write(n)
        c+=1
    print(c, "records written.")
    s.close()
    t.close()
midterm()

```

... Write all of the output to **data\_B\_out.txt** and/or the console as appropriate: **data\_B\_out.txt**:

```

*****
1950: n/a
1951: n/a
1952: n/a
1953: (3-6-6-7-
1954: -13--14--7--15-
1955: -32--10--10--31-
1956: -19--14--15--9-
1957: -8-16-9-16-
1958: n/a
1959: n/a

```

Console : >>> 12 records written.

CSCI-UA.60-1 Midterm Exam -- Fall, 2019 – Version “B”

Your Name: \_\_\_\_\_

2. SQLite keywords (15 points) - Define each of the following SQLite terms, dot commands or operators and then use each one correctly in an SQLite query or statement.

Term	Definition and example of how to use it
UNION	Put the results of two sub-queries together to form one final result. <pre>SELECT (SELECT * FROM books ORDER BY price DESC LIMIT 1) UNION (SELECT * FROM books ORDER BY price ASC LIMIT 1);</pre>
.mode csv	Set the output mode to comma separated vectors. <pre>,mode CSV</pre> <p>(we can then use .show to check mode)</p>
.import	import a file into a table <pre>.import FILE-NAME TABLE-NAME</pre>
LIKE	Give a boolean result on whether a string item comply with the regular expression. Can be used in WHERE or HAVING sentence. <pre>SELECT * FROM books WHERE author_name LIKE '%open Name%';</pre>
UPDATE	Set new values to the current table values. <pre>UPDATE price = price * 2 FROM books;</pre>

**CSCI-UA.60-1 Midterm Exam – Fall, 2019 – Version “B”**

Your Name: \_\_\_\_\_

4. Given the following table schema for film permits in New York City: (40 points)

```
/* Source: https://data.cityofnewyork.us/City-Government/Film-Permits/tg4x-b46p/data*/
DROP TABLE IF EXISTS permits;
CREATE TABLE IF NOT EXISTS permits (
    EventID      INTEGER,
    Borough      TEXT,
    CommunityBoard TEXT,
    Precinct     TEXT,
    Category     TEXT,
    SubCategory  TEXT,
    ZipCode      TEXT
);
```

Notes:

EventID	Permit ID (unique)
Borough	Bronx, Brooklyn, Queens, Manhattan, Staten Island
CommunityBoard	NYC community board(s) who approved the permit
Precinct	NYC police precinct(s) associated with the permit
Category	permit category (television, film, commercial, etc.)
SubCategory	sub-categories within each category*
ZipCode	zip code where the filming primarily takes place

\* For example, a “feature” is a sub-category of films.

... and sample data: Note: This is an excerpt. The complete table contains approximately 50,000 records and 14 columns. You may assume valid data although some data have been changed for this example.

EventID	Borough	Community Board(s)	Police Precinct(s)	Category	SubCategoryName	ZipCode(s)
300134	Manhattan	7	20	Television	Morning Show	10023
299797	Queens	2	108	Television	Cable-episodic	11101
300179	Queens	1, 2	108, 94	Television	Episodic series	11101
300630	Queens	14	100	Commercial	Commercial	11692
300127	Manhattan	4	10	Television	Reality	10001
43547	Brooklyn	1, 2	108, 94	Television	Episodic series	11101
300435	Brooklyn	1	94	Television	Episodic series	11222
300417	Queens	2	108	Television	Episodic series	11101
300441	Brooklyn	14	70	WEB	Not Applicable	11218
300122	Manhattan	10	28	Theater	Theater	10027
297653	Brooklyn	1	90	Film	Short	11211
300432	Brooklyn	16	73	Film	Feature	11212
299740	Brooklyn	1	90	Commercial	Commercial	11206
299639	Brooklyn	6	76	Television	Cable-episodic	11231

Write SQLite queries to answer the following questions.

Note: Be sure to write each **QUERY** ... not the ANSWER.

CSCI-UA.60-1 Midterm Exam – Fall, 2019 – Version “B”

Your Name: \_\_\_\_\_

1. How many permits are on file?

```
SELECT COUNT(*)  
FROM permits;
```

2. How many permits are there in each category? List the categories in alphabetical order.

```
SELECT Category, COUNT(*)  
FROM permits  
GROUP BY Category  
ORDER BY Category ASC;
```

3. Display all of the permits in Brooklyn in this format including the Event ID, Category, Sub-category and zip code:

300435 Television/Episodic series (11222)

```
SELECT EventID || ' ' || Category || ' / ' || SubCategory || ' (' || ZipCode || ')'  
FROM permits  
WHERE Borough = 'Brooklyn';
```

4. Which borough has the most film permits on file?

```
SELECT Borough, COUNT(*) AS c  
FROM permits  
GROUP BY Borough  
ORDER BY c DESC  
LIMIT 1;
```

CSCI-UA.60-1 Midterm Exam – Fall, 2019 – Version “B”  
Your Name: \_\_\_\_\_

5. List the Event ID, borough and precinct for all Queens commercial permits in order by EventID.

```
SELECT EventID, Borough, Precinct  
FROM permits  
WHERE Borough == 'Queens'  
AND Category == 'Commercial'  
ORDER BY EventID ASC;
```

6. How many permits are issued for each SubCategory within Television in increasing order?

```
SELECT SubCategory, COUNT(*) AS c  
FROM permits  
WHERE Category == 'Television'  
GROUP BY SubCategory  
ORDER BY c ASC;
```

7. How many film permits were issued in zip codes 11211 and 11212?

```
SELECT COUNT(*)  
FROM permits  
WHERE ZipCode IN ('11211', '11212')  
AND Category == 'Film';
```

8. Which Borough(s) in alphabetical order - if any - have over 10,000 film permits on file?

```
SELECT Borough, COUNT(*) AS c  
FROM permits  
WHERE Category == 'Film'  
GROUP BY Borough  
HAVING c > 10000  
ORDER BY Borough ASC;
```

9. How many permits have been approved by community board # 2?

```
SELECT COUNT(*) FROM permits  
WHERE CommunityBoard == '2' OR CommunityBoard LIKE '2,%'  
OR CommunityBoard LIKE '%, 2' OR CommunityBoard LIKE  
'%, 2,%';
```

CSCI-UA.60-1 Midterm Exam – Fall, 2019 – Version “B”

Your Name: \_\_\_\_\_

10. New York City is famous for feature films that show off the city. How many feature film permits are on file?

```
SELECT COUNT(*) FROM permits  
WHERE Category = 'Film' AND SubCategory = 'Feature';
```

11. Display a list of categories and sub-categories in alphabetical order by category and then sub-Category so that each sub-category appears only once.

```
SELECT Category, Subcategory  
FROM permits  
GROUP BY Category, Subcategory  
ORDER BY Category ASC, Subcategory ASC;
```

12. Write an original query. First write out the question that you are asking using English and then write an appropriate SQLite query to answer your question.

How many permits<sup>in total</sup> are issued in Brooklyn or Queens for short film?

```
SELECT COUNT(*)  
FROM permits  
WHERE Category = 'Film' AND SubCategory = 'Short'  
AND Borough IN ('Brooklyn', 'Queens');
```

4. Database Normalization (20 points)

Title	Language	Author	Year	Edition	Price in \$
A Room of One's Own	English	Woolf, Virginia (1882-1941)	1922	Penguin; Signet Classics	29.00
Anna Karenina	RU - Russian	Tolstoy, Leo	1875	Penguin [London]	13.50
	English	Dickens, Charles	1870	Random House	none
Der Steppenwolf	German	Hesse, Herman (died in Switzerland in 1923)	1929	Fischer Verlag [Berlin]	95.00
Don Quixote	SP - Spanish	Cervantes	1605	Penguin [New York]	25.00
Hamlet, Prince of Denmark	Shakespearean English	Shakespeare	1603	Signet Classics	7.95
Harry Potter	English	Rowling, J.K.	2000	Harcourt Brace	19.95
Huckleberry Finn	American English	Twain, Mark	1865	Penguin [London]	5.76
Lord of the Rings	English	Tolkien, J.R.	1937	Penguin [NY]	27.45
Mrs. Dalloway	English	Woolf, Virginia	1925	Harcourt Brace	\$25 + tax
One Hundred Years of Solitude	Spanish	Márquez, Gabriel García (born in Colombia)	1964	Harper [Toronto]	14.00
Sense and Sensibility	English	Austen, Jane (died in the U.K.)	1814	Penguin [NYC]	€18.20
The Great Gatsby	American English	Fitzgerald, F. Scott	1925	Penguin [London]	12.00
The Hours	English	Cunningham, Michael (1952- )	1999	Harcourt Brace	12.34
Tom Sawyer	American English	Twain, Mark	1862	Random House	25.00
War and Peace	RU - Russian	Tolstoy, Leo	1865	Penguin	€12.70

Your Name: \_\_\_\_\_

a.) First Normal Form

Define First Normal Form

All columns are atomic, i.e., each cell has only one single value,

Does this excerpt satisfy 1NF? Why or Not? Be specific using the data above and cite at least one example from each column to justify your view.

No. Some cells contain more than one value,

e.g., "Woolf, Virginia (1882-1941)"

it should only contain name of author, rather than birth/death year in this cell to be 1NF.

e.g. ~~Penguin~~ "Penguin [London]"

the location London should be separated from edition Penguin.

Your Name: \_\_\_\_\_

b.) Second Normal Form

Define Second Normal Form

Satisfy 1NF, and;

No partial dependency (Values depend on the whole of every candidate key).

Which tables and/or fields would you need to normalize this database if it does not currently satisfy 2NF?

In the current table, bio info of authors are dependent on author (name). To satisfy 2NF, we can create a new table storing bio info of author (with name of author as primary key of this newly created table). And name of author would become a foreign key in the original table.