SQL Writing Overview

```
#
import pymysql import
sqlalchemy

import numpy

# You have installed and configured ipython-sql for previous assignments.
#  https://pypi.org/project/ipython-sql/  #
%load_ext sql

# This is a hack to fix a version problem/incompatibility              with some
of the packages and magics.
#
%config SqlMagic.style = '_DEPRECATED_DEFAULT'

# Make sure that you set these values to the correct values for your installation
and
# configuration of MySQL #
db_user = "root" db_password =
"shuxintang"

import pymysql



#
db_url = "mysql+pymysql://root:shuxintang@localhost/db_book? local_infile=1"

from sqlalchemy import create_engine default_engine =
create_engine(db_url)

import pandas as pd result_df =
pandas.read_sql(
        "show tables from db_book", con=default_engine
)
result_df
```

	Tables_in_db_book
0	advisor
1	classroom
2	course
3	department
4	instructor
5	prereq

Set Operations in SQL

Question

Using the sample data associated with the recommended textbook,

1. What is wrong with the query below.
2. Write and execute a query that produces accurate results that contains all of the information.

Answer:

1. Error: Union in select statement need have the same numbers of columns.
2. See below code;

```
select * from student where dept_name='Comp. Sci.'
union
select * from instructor where dept_name='Comp. Sci.'
```

Answer

1. Error: Union in select statement need have the same numbers of columns.
2. See below code;

```
# Load SQL extension
%reload_ext sql

# Connect to MySQL database
%sql mysql+pymysql://root:shuxintang@localhost/db_book?local_infile=1

%%sql
SELECT id, name, dept_name FROM student WHERE dept_name = 'Comp. Sci.'
UNION
SELECT id, name, dept_name FROM instructor WHERE dept_name = 'Comp.
Sci.';

 * mysql+pymysql://root:***@localhost/db_book?local_infile=1
7 rows affected.

[('00128', 'Zhang', 'Comp. Sci.'),
 ('12345', 'Shankar', 'Comp. Sci.'),
 ('54321', 'Williams', 'Comp. Sci.'),
 ('76543', 'Brown', 'Comp. Sci.'),
 ('10101', 'Srinivasan', 'Comp. Sci.'),
 ('45565', 'Katz', 'Comp. Sci.'),
 ('83821', 'Brandt', 'Comp. Sci.')]
```

Set Operations in Relational Algebra

Question

The query below produces information about instructors that are not advisors. You must write an equivalent relational algebra expression that contains only set operators and project. Replace the query and screen capture below with you answer.

Answer

S1 ← π ID, name (instructor) S2 ← π i_ID (advisor) S3 ← S1 - S2 Result ← π ID, name (S3) from

IPython.display import display from PIL import Image

```
%%sql
SELECT ID, name
FROM instructor
WHERE ID NOT IN (SELECT i_ID FROM advisor);
```

```
 * mysql+pymysql://root:***@localhost/db_book?local_infile=1
6 rows affected.
[('12121', 'Wu'),
 ('15151', 'Mozart'),
 ('32343', 'El Said'),
 ('33456', 'Gold'),
 ('58583', 'Califieri'),
 ('83821', 'Brandt')]
```

```
!pip install Pillow
from IPython.display import display from PIL import
Image

image_path = "/Users/nini/Downloads/FA1F97EF-15AB-4790-BB37- B681054BD4B2.jpeg"
img = Image.open(image_path) display(img)
```

```
%reload_ext sql

# Connect to MySQL database
%sql mysql+pymysql://root:shuxintang@localhost/db_book?local_infile=1
```

✓ [56] < 10 ms

```
%%sql
SELECT id, name, dept_name FROM student WHERE dept_name = 'Comp. Sci.'
UNION
SELECT id, name, dept_name FROM instructor WHERE dept_name = 'Comp. Sci.';
```

✓ [57] < 10 ms

```
 * mysql+pymysql://root:***@localhost/db_book?local_infile=1
7 rows affected.
```

id	name	dept_name
00128	Zhang	Comp. Sci.
12345	Shankar	Comp. Sci.
54321	Williams	Comp. Sci.
76543	Brown	Comp. Sci.
10101	Srinivasan	Comp. Sci.
45565	Katz	Comp. Sci.
83821	Brandt	Comp. Sci.

ER-Modeling

Question

Using Crow's Foot Notation and a tool like Lucidchart, draw a logical ER diagram modeling the relationship. You may add notes/comments that explain decisions you make.

Answer

!pip install Pillow from IPython.display import display from PIL import Image

image_path = "/Users/nini/Downloads/Database ER diagram (crow's foot).png" img = Image.open(image_path) display(img)

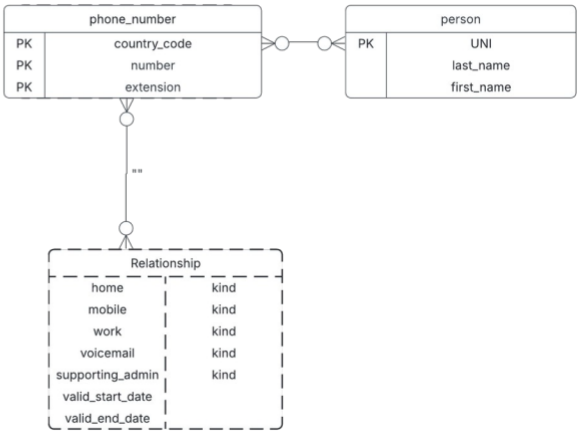
```
!pip install Pillow

from IPython.display import display
from PIL import Image

# Corrected file path
image path = r"/Users/nini/Downloads/Database ER diagram (crow's
foot).png"

# Open and display the image
img = Image.open(image_path)
display(img)
```

```
Requirement already satisfied: Pillow in /Volumes/Columbia /2025
Spring/Datebase 4111W/S25-W4111-HW0-sunday/.venv1/lib/python3.13/site-
packages (11.1.0)
```



ER Diagram to DDL

Question

ER Diagram to DDL

Consider the preceding, **approximate** ER logical model diagram. The diagram is approximate because the definition below of the model may require minor changes in the implemented DDL relative to the diagram. For example, you may have to add constraints, columns not shown, etc.

The semantics/requirements are below.

A sample personrecord for me in personwould be in the form

```
{dff9, Ferguson, Donald, Faculty, donald.ferguson@cs.columbia.edu, dff9@columbia.edu}
```

- The default email is always of the form uni@columbia.edu.
- Preferred email is always UNIQUEbut a person *may not have* a preferred email.
- The possible values for kindare one of {Student, Faculty, Staff}. A sample course

record for our *Intro. to Databases* course would be in the form

```
{COMS, W, 4111, Introduction to Databases, OMG! This class is terrifying., COMSW4111}
```

- dept_codeis always 4 characters and will not contain a digit, space, -, or _
- faculty_codeis one of {W, C, E, B, G}.
- course_nois always 4 digits and cannot begin with a 0.
- full_course_nois the concatenation of dept_code, faculty_code, course_no.

A sample sectionfor our session of COMSW4111 would be

{11969, COMSW4111, 002, 1, 2025, COMSW4111_002_1_2025}

- call_nois always 5 digits and may begin with 0.
- course_nois the same as full_course_noin course.
- section_nois always 3 characters. It can be 3 digits and may start with 0. Or, it can be of the form V02, that is starts with Vand has two digits.
- yearhas the obvious meaning and constraints.
- section_keyis the concatenation of the fields with the _ delimiter. A sample

person_sectionfor me would be

{dff9, 11969, instructor, 20250125, 20250502}

- The roleis one of {instructor, student, TA, auditor}. A person may have nore than one rolein a course.
- The start_datemust be before the end_date.

Put, execute and test your DDL in the code cells below. You can explain assumptions and changes in the markdown cell that precedes the code cells.

Answer

Place you explanatory notes on design choices and assumption in this markdown cell.

```
%%sql

-- 1 Drop the table if it exists to prevent conflicts DROP TABLE IF EXISTS person;

-- 2 Create the `person` table CREATE TABLE
person (
        UNI VARCHAR(10) PRIMARY KEY,
        last_name VARCHAR(50) NOT NULL,
        first_name VARCHAR(50) NOT NULL,
        kind ENUM('Student', 'Faculty', 'Staff') NOT NULL, preferred_email VARCHAR(100)
        UNIQUE,
        default_email VARCHAR(100) UNIQUE NOT NULL
);

-- 3 Drop the trigger if it already exists DROP TRIGGER IF
EXISTS set_default_email;

-- 4 Create the trigger **without DELIMITER** CREATE
TRIGGER set_default_email
BEFORE INSERT ON person FOR
EACH ROW
BEGIN
        IF NEW.default_email IS NULL OR NEW.default_email = '' THEN SET
                NEW.default_email = CONCAT(NEW.UNI, '@columbia.edu');
```

```
        ENDIF; END;
```

```
 * mysql+pymysql://root:***@localhost/db_book?local_infile=1
16    rows  affected.
16    rows  affected.
16    rows  affected.
16    rows affected.
```