Relational Calculus & Relational Algebra

1. **Give one reason why MRN is a poor choice for the primary & foreign keys?**
   One single patient may come to see clinicians for more than one time, so the same MRN could appear on multiple tuples in the relation VISIT.

2. **Write Relational Calculus expressions for the following:**

   (a) Who is 25 years old?

   $$\{pa.\,FIRSTNAME, pa.\,LASTNAME \mid PATIENT(pa) \wedge pa.\,AGE = '25'\}$$

   (b) Who had a medical visit in December, 2017?

   $$\{\, pa.\,FIRSTNAME, pa.\,LASTNAME \mid PATIENT(pa) \wedge \exists(v)(VISIT(v) \wedge v.\,DATETIME = 'December, 2017' \wedge pa.\,MRN = v.\,MRN)\}$$

   (c) Who got a 'flu shot'?

   $$\{pa.\,FIRSTNAME, pa.\,LASTNAME \mid PATIENT(pa) \wedge \exists(v, pr)(VISIT(v) \wedge PROCEDURE(pr) \wedge pr.\,NAME = 'flu\ shot' \wedge v.\,VISIT\_ID = pr.\,VISIT\_ID \wedge pa.\,MRN = v.\,MRN)\}$$

   (d) Who did NOT get a 'flu shot'?

   $$\{pa.\,FIRSTNAME, pa.\,LASTNAME \mid PATIENT(pa) \wedge \neg\exists(v, pr)(VISIT(v) \wedge PROCEDURE(pr) \wedge pr.\,NAME = 'flu\ shot' \wedge v.\,VISIT\_ID = pr.\,VISIT\_ID \wedge pa.\,MRN = v.\,MRN)\}$$

   (e) What is the first and last name of all patients who have seen MD Paula Jones?

   $$\{pa.\,FIRSTNAME, pa.\,LASTNAME \mid PATIENT(pa) \wedge \exists(v, pr, c)(VISIT(v) \wedge PROCEDURE(pr) \wedge CLINICIAN(c) \wedge c.\,CERT = 'MD' \wedge c.\,FIRSTNAME = 'Paula' \wedge c.\,LASTNAME = 'Jones' \wedge pr.\,CLIN\_ID = c.\,CLIN\_ID \wedge v.\,VISIT\_ID = pr.\,VISIT\_ID \wedge pa.\,MRN = v.\,MRN)\}$$

3. **Write Relational Algebra expressions for the following questions. Return the relation primary key to identify the tuples, unless otherwise specified.**

   (a) Who is 25 years old?

   $$\pi_{FIRSTNAME,\ LASTNAME}(\sigma_{AGE\ =\ '25'}(PATIENT))$$

   (b) What is the first and last name of all patients who have seen a Physician's Assistant (cert is 'PA')

   $$C \leftarrow \rho_{FIRSTNAME\_C/FIRSTNAME,\ LASTNAME\_C/LASTNAME}(CLINICIAN)$$

   $$\pi_{FIRSTNAME,\ LASTNAME}\ (\sigma_{CERT\ =\ 'PA'}(PATIENT * VISIT * PROCEDURE * C))$$

(c) Which patients have the same names and age?

$\pi_{\text{FIRSTNAME, LASTNAME}}(\sigma_{\text{P.MRN != P1.MRN}}(\rho_{\text{P(.....)}}(\text{PATIENT}) \bowtie_{\text{P.FIRSTNAME=P1.FIRSTNAME} \wedge \text{P.LASTNAME=P1.LASTNAME} \wedge \text{P.AGE=P1.AGE}} \rho_{\text{P1(.....)}}(\text{PATIENT}))$)

(d) Which patients who got a flu shot also got a measles immunization during the same visit?

$R \leftarrow \pi_{\text{VISIT\_ID, NAME1}}(\rho_{\text{NAME1/NAME}}(\sigma_{\text{NAME = 'flu shot'}}(\text{PROCEDURE})))$

$S \leftarrow \pi_{\text{VISIT\_ID, NAME2}}(\rho_{\text{NAME2/NAME}}(\sigma_{\text{NAME = 'measles immunization'}}(\text{PROCEDURE})))$

$\pi_{\text{FIRSTNAME, LASTNAME}}(R*S*VISIT*PATIENT)$

(e) Which patients who have seen an MD have not seen a PA?

$C \leftarrow \rho_{\text{FIRSTNAME\_C/FIRSTNAME, LASTNAME\_C/LASTNAME}}(\text{CLINICIAN})$

$R \leftarrow \pi_{\text{FIRSTNAME, LASTNAME}}(\sigma_{\text{CERT = 'MD'}}(\text{PATIENT} * \text{VISIT} * \text{PROCEDURE} * C))$

$S \leftarrow \pi_{\text{FIRSTNAME, LASTNAME}}(\sigma_{\text{CERT = 'PA'}}(\text{PATIENT} * \text{VISIT} * \text{PROCEDURE} * C))$

$\pi_{\text{FIRSTNAME, LASTNAME}}((\pi_{\text{FIRSTNAME, LASTNAME}}(\text{PATIENT}) - S) \cap R)$

Queries

1. **List the nct id and study type from the study whose brief title is "Autologous Cell Therapy After Stroke".**

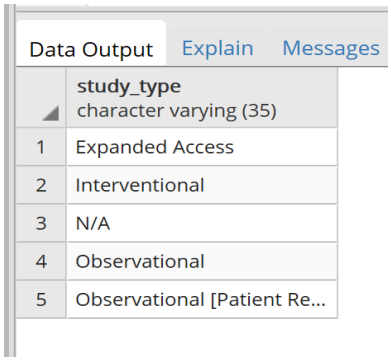
```
SELECT nct_id, study_type
FROM studies
WHERE brief_title = 'Autologous Cell Theory After Stroke';
```

Data Output | Explain | Messages | Notifications

	nct_id character (11)	study_type character varying (35)
1	NCT00908856	Interventional

2. List the different values for study type, in alphabetical order.

```
SELECT DISTINCT study_type
FROM studies
ORDER BY study_type ASC;
```

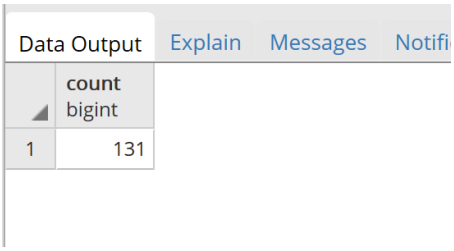


Data Output | Explain | Messages

	study_type character varying (35)
1	Expanded Access
2	Interventional
3	N/A
4	Observational
5	Observational [Patient Re...

3. How many terminated studies that started and completed in 2016 have reported events?

```
SELECT COUNT (DISTINCT studies.nct_id)
FROM studies, reported_events
WHERE studies.nct_id = reported_events.nct_id
        AND start_date >= '2016-01-01'
        AND completion_date < '2017-01-01';
```

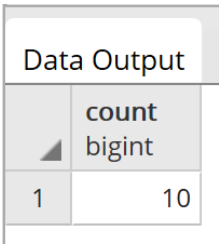


Data Output | Explain | Messages | Notifi

	count bigint
1	131

4. How many of the studies that started in February 2016, but on or after the 15th, are expected to complete (or have completed) within 6 months of their start date?

```
SELECT COUNT(DISTINCT nct_id)
FROM studies
WHERE start_date >= '2016-02-15'
        AND start_date < '2016-03-01'
        AND completion_date <= start_date + INTERVAL '6 MONTH';
```



Data Output

	count bigint
1	10