CS 5530

Database Systems Spring 2020

Adv. Queries II

Project Teams

•Come to the front after class if you don't have a teammate yet

Database Server

- You will have access this week
 - Do not spam it with queries
 - Do not run expensive queries
 - Do not insert large quantities of data
 - Keep your total footprint < 10MB
 - Do not run queries during class

Midterm

- Next Wednesday
- •5 pages of notes

Find all students younger than everyone in Databases

Students

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

Enrolled

sID	cID	Grd
1	3500	Α
1	3810	A-
1	5530	Α
2	3810	Α
2	5530	В
3	3500	C
3	3810	В
4	3500	C

Courses

cID	Name
3500	SW Practice
3810	Architecture
5530	Databases

•Find all students younger than everyone in 'Databases'

```
select s1.sName from Students s1 where s1.DOB > all
```

```
(select s2.DOB from Students s2
natural join Enrolled natural join
Courses c where c.cName='Databases');
```

•Find all students younger than everyone in 'Databases'

```
select s1.sName from Students s1
where s1.DOB > all
    (select s2.DOB from Students s2 ...);
```

•Find all students younger than everyone in 'Databases'

```
select s1.sName from Students s1
where s1.DOB > all
   (select s2.DOB from Students s2 ...);
```

•Think of nested queries as nested for-loops

```
select sl.sName
                         (select s2.DOB
                          from Students s2
from Students s1
where s1.DOB > all
                          ...);
   foreach Student s1 in Students {
     foreach Student s2 in ... {
       if(s1.DOB \le s2.DOB)
         don't select s
```

•Think of nested queries as nested for-loops

```
select s1.sName
from Students s1
where s1.DOB > all
```

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

```
(select s2.DOB
from Students s2
...);
```

DOB	
1980	
1979	

•Think of nested queries as nested for-loops

select s1.sName
from Students s1
where s1.DOB > all

	sID	Name	DOB
s1	1	Hermione	1980
	2	Harry	1979
	3	Ron	1980
	4	Malfoy	1982

(select s2.DOB
from Students s2
...);

DOB \$2 1980 1979

•Think of nested queries as nested for-loops

select s1.sName
from Students s1
where s1.DOB > all

	sID	Name	DOB
s1	1	Hermione	1980
	2	Harry	1979
	3	Ron	1980
	4	Malfoy	1982

(select s2.DOB
from Students s2
...);

	DOB
	1980
s2	1979

•Think of nested queries as nested for-loops

select s1.sName
from Students s1
where s1.DOB > all

	sID	Name	DOB
	1	Hermione	1980
1	2	Harry	1979
	3	Ron	1980
	4	Malfoy	1982

(select s2.DOB
from Students s2
...);

DOB \$2 1980 1979

•Think of nested queries as nested for-loops

select s1.sName
from Students s1
where s1.DOB > all

	sID	Name	DOB
	1	Hermione	1980
s1	2	Harry	1979
	3	Ron	1980
	4	Malfoy	1982

(select s2.DOB
from Students s2
...);

	DOB
	1980
s2	1979

EXISTS

```
select x from y where EXISTS (select ...);
```

•If any rows exist in nested query, x is selected

NOT EXISTS

```
select x from y where
NOT EXISTS
(select ...);
```

•If nested query empty, x is selected

•Find students taking all classes

S

sID	Name
1	Hermione
2	Harry

 \mathbf{C}

cID	Name
3500	SW Practice
3810	Architecture

 ϵ

sID	cID
1	3500
1	3810
2	3810

•Find students taking all classes:

•Find students taking all classes:

```
select s.sName

from Students s

where not exists(select c.cID

from Courses c

where not exists(select e.cID
```

foreach Student s foreach Course c foreach Enrollment e

```
(select e.cID
  from Enrolled e
  where e.cID = c.cID
  and e.sID = s.sID));
```

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
from Enrolle
```

foreach Student s foreach Course c foreach Enrollment e

```
(select e.cID
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
```

If inner select is non-empty, then this is false

foreach Student s foreach Course c foreach Enrollment e

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

•Find students taking all classes:

If inner select is non-empty, then this is false

If false for all {s, c}, then this select is empty foreach Student s foreach Course c foreach Enrollment e

```
where not exists (select e.cID

from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));
elect is non-empty
```

•Find students taking all classes:

```
select s.sName
from Students s

where not exists(select c.cID
from Courses c
where not exists(select e.cID
which makes

from Enrolle
```

this true

If inner select is non-empty, then this is false

If false for all {s, c}, then this select is empty foreach Student s foreach Course c foreach Enrollment e

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
from Enrolle
```

foreach Student s foreach Course c foreach Enrollment e

```
(select e.cID
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

•Find students taking all classes:

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
```

If inner select is empty, then this is true

foreach Student s foreach Course c foreach Enrollment e

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

•Find students taking all classes:

If inner select is empty, then this is true

If ever true, then this select is non-empty

foreach Student s foreach Course c foreach Enrollment e

•Find students taking all classes:

```
select s.sName
from Students s

where not exists(select c.cID
from Courses c
where not exists(select e.cID
```

which makes this false

If inner select is empty, then this is true

If ever true, then this select is non-empty

foreach Student s foreach Course c foreach Enrollment e

```
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

foreach Student s foreach Course c foreach Enrollment e

```
where not exists(select e.cID

from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));
```

S

sID	Name
1	Hermione
2	Harry

C

cID	Name
3500	SW Practice
3810	Architecture

 ϵ

sID	cID
1	3500
1	3810
2	3810

foreach Student s foreach Course c foreach Enrollment e

where not exists (select e.cID

from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));

S

sID	Name
1	Hermione
2	Harry

C

cID	Name
3500	SW Practice
3810	Architecture

e

sID	cID
1	3500
1	3810
2	3810

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
```

foreach Student s foreach Course c foreach Enrollment e

```
where not exists(select e.cID

from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));
```

{3500}

3	
sID	Name
1	Hermione
2	Harry

cID	Name
3500	SW Practice
3810	Architecture

	•
sID	clD
1	3500
1	3810
2	3810

6

sID	Name
1	Hermione
2	Harry

clD	Name
3500	SW Practice
3810	Architecture

sID	cID
1	3500
1	3810
2	3810

foreach Student s

```
foreach Students

select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
from Enrolled e
where e.cID = c.cII
```

from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));

.	
sID	Name
1	Hermione
2	Harry

C		
	cID	Name
	3500	SW Practice
(3810	Architecture

sID	cID
	3500
1	3810
2	3810

select s.sName from Students s where not exists (select c.cID from Courses c where not exists (select e.cID

foreach Student s foreach Course c foreach Enrollment e

from Enrolled e where e.cID = c.cIDand e.sID = s.sID));

{3810}

sID Name Hermione Harry

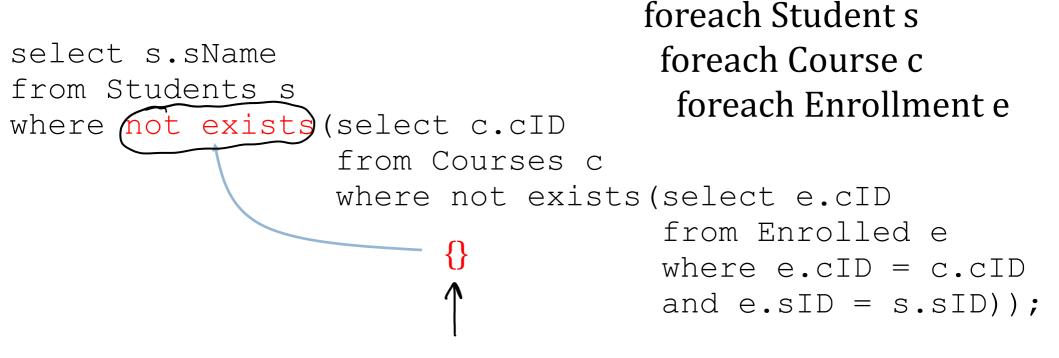
S

cID	Name
3500	SW Practice
3810	Architecture

 \mathbf{C}

sID cID 3500 3810 3810

e



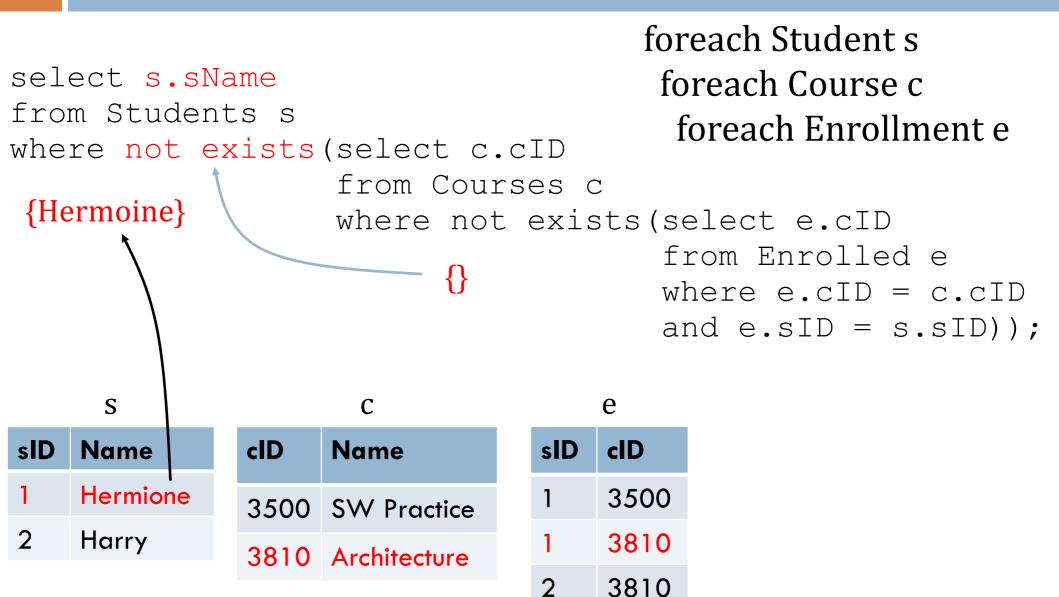
3	
sID	Name
1	Hermione
2	Harry

S

cID	Name
3500	SW Practice
3810	Architecture

 \mathbf{C}

sID	cID
1	3500
1	3810
2	3810



select s.sName

from Students s

where not exists(select c.cID

from Courses c

Where not exis

foreach Student s foreach Course c foreach Enrollment e

sID Name

1 Hermione

2 Harry

cID Name

3500 SW Practice

3810 Architecture

sID	cID
1	3500
1	3810
2	3810

e

```
select s.sName
from Students s
where not exists (select c.cID
```

foreach Student s foreach Course c foreach Enrollment e

{Hermoine}

from Courses c

where not exists (select e.cID from Enrolled e

where e.cID = c.cID

and e.sID = s.sID));

S

sID	Name
1	Hermione
2	Harry

 \mathbf{C}

cID	Name
3500	SW Practice
3810	Architecture

sID	cID
1	3500
1	3810
2	3810

```
select s.sName

from Students s

where not exists(select c.cID

from Courses c

Where not exis
```

foreach Student s foreach Course c foreach Enrollment e

S	
sID	Name
1	Hermione
2	Harry

C		
	:ID	Name
	3500	SW Practice
	3810	Architecture

	C
sID	cID
1	3500
1	3810
2	3810

```
select s.sName
from Students s
where not exists (select c.cID
```

foreach Student s foreach Course c foreach Enrollment e

{Hermoine}

from Courses c

where not exists (select e.cID

from Enrolled e where e.cID = c.cID

and e.sID = s.sID));

S

Name

Harry

Hermione

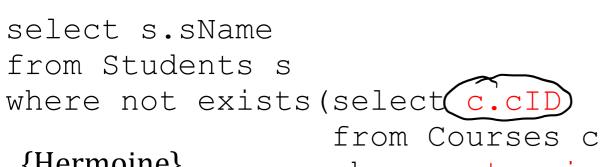
sID

 \mathbf{C}

cID	Name
3500	SW Practice
3810	Architecture

e

sID	cID
1	3500
1	3810
2	3810



foreach Student s foreach Course c foreach Enrollment e

{Hermoine}

where not exists (select e.cID

{3500}

from Enrolled	е
where e.cID =	c.cID
and $e.sID = s.$.sID));

sID Name Hermione Harry

S

c	
cID	Name
3500	SW Practice
3810	Architecture

sID	cID
1	3500
1	3810
2	3810

e

```
select s.sName
from Students s
where not exists(select c.cID
from Courses c
where not exists(select e.cID
from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));
```

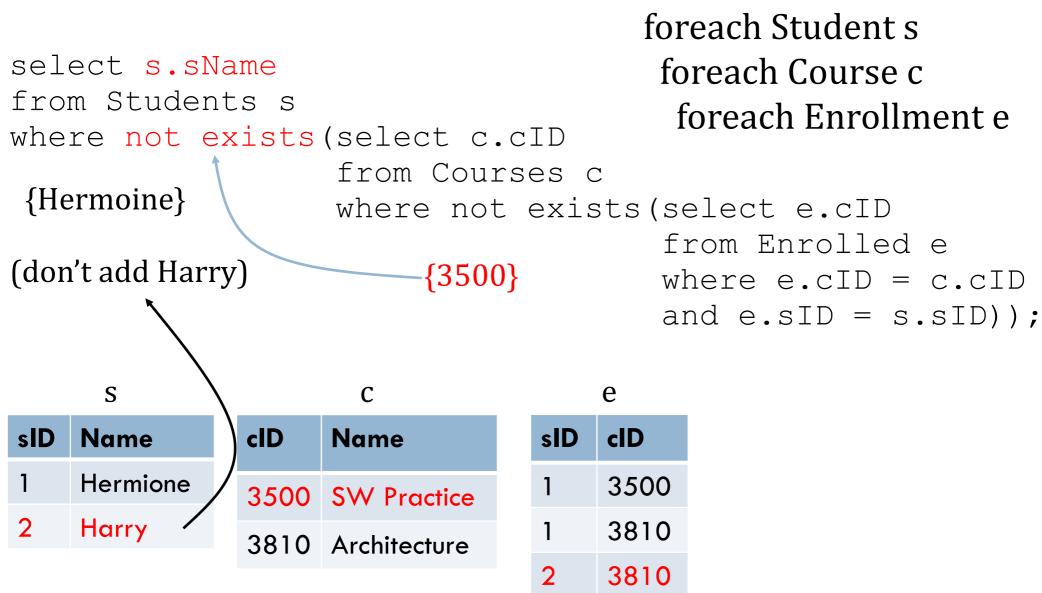
	_
sID	Name
1	Hermione
2	Harry

S

clD	Name
3500	SW Practice
3810	Architecture

 \boldsymbol{C}

sID	cID	
1	3500	
1	3810	
2	3810	



Ligt 8. ... 51

foreach x in SI foreach

52. Comtains

 foreach Student s foreach Course c foreach Enrollment e

{Hermoine}

rom Courses c

where not exists (select e.cID

Only student taking all classes

from Enrolled e
where e.cID = c.cID
and e.sID = s.sID));

S

sID	Name
1	Hermione
2	Harry

C

cID	Name
3500	SW Practice
3810	Architecture

e

sID	cID
1	3500
1	3810
2	3810

Strings

•SQL has pseudo regex:

```
... WHERE Title LIKE 'J_%m'
```

- '_' means any one character
- '%' means 0 or more arbitrary characters

Strings

•SQL has pseudo regex:

```
... WHERE Title LIKE 'J_%m'
```

• 'J' followed by at least one arbitrary character followed by 'm'

LIMIT

- •Limits the number of rows returned
 - Select the first 5 rows

SELECT ... LIMIT 5;

• Select rows 3-7

SELECT ... LIMIT 5 OFFSET 2;

LIMIT

- •Useful in combination with ORDER BY
 - Find ... with smallest CardNum
 SELECT ... LIMIT 1 ORDER BY CardNum;
 - Find ... with biggest CardNum

 SELECT ... LIMIT 1 ORDER BY CardNum DESC;

Aggregate Functions

- •An aggregate function returns a single number from a multiset
- •Usually used on a single column

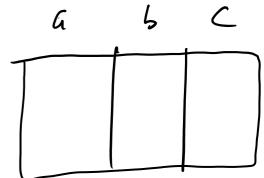
```
•COUNT (...)
```

- •MAX (...)
- •MIN (...)
- •SUM (...)
- •AVG (...)

•...

COUNT

- •Returns count instead of rows
- •How many copies of 'Harry Potter'?



SELECT COUNT(*)

FROM Titles NATURAL JOIN Inventory
WHERE Title='Harry Potter';

Inventory

Serial	ISBN
1001	978-0590353427
1002	978-0590353427
1003	978-0679732242

Titles

ISBN	Title	Author
978-0590353427	Harry Potter	Rowling
978-0679732242	The Sound and the Fury	Faulkner

COUNT

- •Returns count instead of rows
- •How many copies of 'Harry Potter'?

```
SELECT COUNT(*)
FROM Titles NATURAL JOIN Inventory
WHERE Title='Harry Potter';
```

•Demo

•Gives one representative row for each group in the specified column

select Major from Students;

ART
CS
CS
PHYS

select Major from Students group by Major;

CS PHYS

•Isn't this...

```
select Major from Students group by Major;
```

•the same as this?

```
select DISTINCT Major from Students;
```

•Gives one representative row for each group in the specified column

select * from Students group by Major;

	υ0123456	ART	Sandy
_ 	υ0654321	CS	Jim
	υ0555555	PHYS	Ray

•GROUP BY usually used in conjunction with an aggregate function

```
SELECT Major, COUNT(*) FROM Students GROUP BY Major;
```

•COUNT (*) is applied to each group separately

•GROUP BY usually used in conjunction with an aggregate function

SELECT Major, COUNT(*) FROM Students GROUP BY Major;

count of each group u0123456 Sandy **ART ART** u0654321 CS Jim Groups CS u0999999 CS Danny PHYS u0555555 Ray **PHYS**

•GROUP BY usually used in conjunction with an aggregate function

SELECT Major, COUNT(*) FROM Students GROUP BY Major;

ART	1
CS	2
PHYS	1

SELECT Major, AVG(gpa) FROM Students GROUP BY Major;

ART	4.0
CS	3.4
PHYS	3.6

Quiz

- •Find number of students in each class
- •Find number of classes for each student

Students

sID	Name	DOB
1	Hermione	1980
2	Harry	1979
3	Ron	1980
4	Malfoy	1982

Enrolled

sID	cID	Grd
1	3500	Α
1	3810	A-
1	5530	Α
2	3810	Α
2	5530	В
3	3500	С
3	3810	В
4	3500	С

Courses

cID	Name
3500	SW Practice
3810	Architecture
5530	Databases

HAVING

•Filter on groups (instead of rows)

```
select Major, count(*) from Students
group by Major
HAVING count(*) = 2;
```

CS 2

HAVING

•Filter on groups (instead of rows)

```
select Major, count(*) from Students
group by Major
HAVING count(*) = 2;
```

•Better

```
select Major, count(*) as c
from Students group by Major
HAVING(c) = 2;
```

WHERE VS. HAVING

- •Where is used before GROUP BY to filter rows
- •Having is used after GROUP BY to filter groups
 - Or on the result of an aggregate function

WHERE VS. HAVING

•Find students with 2 or more A's

```
select sID, count(*) as c
from Enrolled e
where e.Grade='A'
group by sID
having c >= 2;
```

- •Remember: aggregate functions don't return a row
 - They aggregate a group into one value
- •Find the name of the cheapest product from each Dept

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

•Find the name of the cheapest product from each Dept

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

```
select Dept, min(Price), Name from Items group by Dept;
```

•Find the name of the cheapest product from each Dept

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

select Dept, min(Price), Name from Items group by Dept;

Dept	min(Price)	Name
produce	.70	Carrots
bakery ((2)	Cake

•Find the name of the cheapest product from each Dept

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

select Dept, min(Price), Name from Items group by Dept;

Dept	min(Price)	Name
produce	.70	Carrots
bakery	2 ← →	Cake

•Which tuple would we associate with AVG?

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

select Dept, AVG(Price), Name from Items group by Dept;

Dept	avg(Price)	Name
produce	.925	š šš
bakery	6	\$\$\$

•Which tuple would we associate with AVG?

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

Dept	avg(Price)	Name
produce	.925	Carrots
bakery	6	Cake

Just pick some representative from the group

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

select * from Items group by Dept;

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

select Dept, min(Price), Name from Items group by Dept;

Dept	Price	Name	Dept	min(P)	Name
produce	.70	Carrots	produce	.70	Carrots
produce	1.15	Peaches	produce	1.15	Peaches
bakery	10	Cake	bakery (2	Cake
bakery	2	Bread	bakery	2	Bread

•Find the Name of the cheapest item from each Dept

Dept	Price	Name
produce	.70	Carrots
produce	1.15	Peaches
bakery	10	Cake
bakery	2	Bread

```
select Items.Name, mins.Dept from Items join (select Dept, min(Price) as p from Items group by Dept) as mins
```

```
where Items.Price = mins.p and
Items.Dept=mins.Dept;
```

Aggregate of Aggregates?

•What if we want the average number of students per major

SELECT Major, COUNT(*) FROM Students GROUP BY Major;

ART	1
CS	2
PHYS	1

Aggregate of Aggregates?

•What if we want the average number of students per major

SELECT Major, COUNT(*) FROM Students GROUP BY Major;

ART	1		AVG (count(1)
CS	2	ς χ le (⁺	
PHYS	1		

SELECT AVG(counts.c) FROM (SELECT COUNT(*) AS c FROM Students GROUP BY Major) AS counts;

- •DATE is a compound type, containing:
 - Year
 - Month
 - Day

- •DATE is a compound type, containing:
 - Year
 - Month
 - Day
- •But we can still do comparisons on them:

- •DATE is a compound type, containing:
 - Year
 - Month
 - Day

11:15;20

•But we can still do comparisons on them:

Automatic conversion from string

•We can extract one component from a date:

```
J
```

```
select YEAR(DOB) from Students;
select MONTH(DOB) from Students;
select DAY(DOB) from Students;
```

Date Arithmetic?

```
select ("2020-02-10" - DOB) ...;
```

Date Arithmetic?

```
select (DATE("2019-02-11") - DOB) ...;
```

- •Doesn't really do what you want...
 - And what about leap years?

Example

•Find age in years of all students

Example

•Find age in years of all students

```
select (2019 - year(DOB)) from Students;
```

- Better, but still some problems
 - (2019) doesn't account for today's month and day
 - Would have to re-write query every day

TIMESTAMPDIFF

•Gives number of years, months, days, or seconds between two dates

SELECT TIMESTAMPDIFF (unit, date1, date2) ...

• ... and it accounts for leap years

CURDATE ()

•We don't want to have to rewrite this query every day:

```
SELECT TIMESTAMPDIFF (unit, date1, DATE ("2019-02-11")) ...
```

•CURDATE gives us the server's date

```
SELECT TIMESTAMPDIFF (unit, date1, CURDATE()) ...
```

•Lots of built-in date functions:

```
dayname(date)
monthname(date)
dayofmonth(date)
dayofweek(date)
quarter(date)
from_unixtime(uint)
```

•...many more