

# CS 327E Class 7

March 11, 2019

# Data Integration Milestone

	<b><i>Dataset1</i> Examples</b>	<b><i>Dataset2</i> Examples</b>
<b>Transportation</b>	Airline on-time performance (source: Bureau of Transportation Statistics)	Storm events (source: NOAA)
<b>Housing</b>	Short-term rentals in various cities (source: Airbnb)	Long-term rentals nationwide (source: Zillow)
<b>Political Campaigns</b>	Federal campaign finance (source: Federal Election Commission)	State campaign finance (source: TX Ethics Commission)
<b>Movies</b>	Hollywood movies, directors, actors (source: IMDB)	Bollywood movies, actors and songs (source: Cinemalytics)
<b>Music</b>	Artists and songs (source: MusicBrainz)	Artists, labels, recordings on vinyl and other formats (source: Discog)

1) Which is not an aggregate function?

- A. SUM()
- B. COUNT(\*)
- C. AVG()
- D. MIN()
- E. None of the above

2) Consider the `Women_Basketball_Players` table shown below. What is the output from Q1 when run on this table?

```
Q1: SELECT COUNT(*) FROM Women_Basketball_Players
```

`Women_Basketball_Players`

<u>player_id</u>	player_name	height	position	points
40	Jatarie White	6-4	Center	24
5	Jordan Hosey	6-1	Forward	13
31	Audrey-Ann Caron Goudreau	6-4	Forward	21
14	Olamide Aborowa	6-3	Forward	11
20	Brianna Tayler	5-9	Guard	19
30	Khaleann Caron-Goudreau	6-4	Forward	11
12	Jada Underwood	6-0	Forward	19

- A. 7
- B. 4
- C. 3
- D. 0
- E. NULL

3) Consider the `Women_Basketball_Players` table shown below. What is the output from Q2 when run on this table?

```
Q2: SELECT MIN(points) FROM Women_Basketball_Players
```

`Women_Basketball_Players`

<u>player_id</u>	player_name	height	position	points
40	Jatarie White	6-4	Center	24
5	Jordan Hosey	6-1	Forward	13
31	Audrey-Ann Caron Goudreau	6-4	Forward	21
14	Olamide Aborowa	6-3	Forward	11
20	Brianna Tayler	5-9	Guard	19
30	Khaleann Caron-Goudreau	6-4	Forward	11
12	Jada Underwood	6-0	Forward	19

- A. 0
- B. 11
- C. 22
- D. 24
- E. NULL

4) Consider the `Women_Basketball_Players` table shown below. What is the output from Q3 when run on this table?

```
Q3: SELECT MAX(points) FROM Women_Basketball_Players
```

`Women_Basketball_Players`

<u>player_id</u>	player_name	height	position	points
40	Jatarie White	6-4	Center	24
5	Jordan Hosey	6-1	Forward	13
31	Audrey-Ann Caron Goudreau	6-4	Forward	21
14	Olamide Aborowa	6-3	Forward	11
20	Brianna Tayler	5-9	Guard	19
30	Khaleann Caron-Goudreau	6-4	Forward	11
12	Jada Underwood	6-0	Forward	19

- A. 0
- B. 11
- C. 22
- D. 24
- E. NULL

5) Consider the `Women_Basketball_Players` table shown below. What is the output from Q4 when run on this table?

```
Q4: SELECT SUM(points) FROM Women_Basketball_Players
WHERE position = 'Center' OR position = 'Guard'
```

`Women_Basketball_Players`

<code>player_id</code>	<code>player_name</code>	<code>height</code>	<code>position</code>	<code>points</code>
40	Jatarie White	6-4	Center	24
5	Jordan Hosey	6-1	Forward	13
31	Audrey-Ann Caron Goudreau	6-4	Forward	21
14	Olamide Aborowa	6-3	Forward	11
20	Brianna Tayler	5-9	Guard	19
30	Khaleann Caron-Goudreau	6-4	Forward	11
12	Jada Underwood	6-0	Forward	19

- A. 10
- B. 34
- C. 19
- D. 43
- E. NULL

# Syntax of Global Aggregate Queries

```
SELECT <aggregate function>  
FROM <single table>  
JOIN <single table> ON <join condition>  
WHERE <boolean condition>
```



# Syntax of Aggregate Queries with Groups

```
SELECT <unaggregated field>, <aggregate function>  
FROM <single table>  
JOIN <single table> ON <join condition>  
WHERE <boolean conditions>  
GROUP BY <unaggregated field>  
ORDER BY <list of fields to sort on>
```

# Syntax of Aggregate Queries with Groups

```
SELECT <unaggregated field>, <aggregate function>  
FROM <single table>  
JOIN <single table> ON <join condition>  
WHERE <boolean condition>  
GROUP BY <unaggregated field>  
HAVING <boolean condition>  
ORDER BY <fields to sort on>
```

# How COUNT () works

1) `SELECT COUNT (*)  
FROM Employee`

2) `SELECT COUNT (emp_dept)  
FROM Employee`

3) `SELECT COUNT (DISTINCT emp_dept)  
FROM Employee`

## Employee

Row	empid	emp_name	emp_dept
1	6	Sunil	1
2	2	Mike	1
3	23	Dave	2
4	5	Jim	4
5	37	Morgan	4
6	3	Sarah	<i>null</i>

# First Question

*How many students are taking each class?*

Student(sid, fname, lname, dob)

Class(cno, cname, credits)

Teacher(tid, fname, lname, dept)

Takes(sid, cno, grade)

Teaches(tid, cno)

## Second Question

*For each class with at least two students, how many students are taking such a class?*

Student(sid, fname, lname, dob)

Class(cno, cname, credits)

Teacher(tid, fname, lname, dept)

Takes(sid, cno, grade)

Teaches(tid, cno)

# iClicker Question

*For each class with at least two students, how many students are taking such a class?*

Student(sid, fname, lname, dob)

Class(cno, cname, credits)

Teacher(tid, fname, lname, dept)

Takes(sid, cno, grade)

Teaches(tid, cno)

Does this query require a HAVING clause?

- A. Yes
- B. No

# Third Question

*For each student who is 19-years old or older and is earning at least 3 class credits, how many total class credits are such students earning?*

Student(sid, fname, lname, dob)

Class(cno, cname, credits)

Teacher(tid, fname, lname, dept)

Takes(sid, cno, grade)

Teaches(tid, cno)

# iClicker Question

*For each student who is 19-years old or above and is earning at least 3 class credits, how many total class credits are such students earning?*

Student(sid, fname, lname, dob)  
Class(cno, cname, credits)  
Teacher(tid, fname, lname, dept)  
Takes(sid, cno, grade)  
Teaches(tid, cno)

Does this query require a WHERE clause?

A. Yes      B. No



# Fourth Question

*Who takes exactly 3 classes?*

Show the answer as a sorted list of sids.

Student(sid, fname, lname, dob)

Class(cno, cname, credits)

Teacher(tid, fname, lname, dept)

Takes(sid, cno, grade)

Teaches(tid, cno)

# iClicker Question

*Who takes exactly 3 classes?*

Show the answer as a sorted list of sids.

Student(sid, fname, lname, dob)

Class(cno, cname, credits)

Teacher(tid, fname, lname, dept)

Takes(sid, cno, grade)

Teaches(tid, cno)

Does this query contain an aggregate function in the `SELECT` clause?

A. Yes      B. No

# Database Views and Data Studio Demo

```
git clone https://github.com/cs327e-spring2019/snippets.git
```

# Milestone 7

<http://www.cs.utexas.edu/~scohen/milestones/Milestone7.pdf>