

RDBMS and SQL

Madhavan Mukund

<https://www.cmi.ac.in/~madhavan>

Lecture 6, 26 September 2024

Queries in SQL

- `select` columns
 `from` table
 `where` condition
- ```
select name
from instructor
where dept_name = 'Physics';
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL

- `select` columns  
    `from` table  
    `where` condition  
  
`select name`  
    `from instructor`  
    `where dept_name = 'Physics';`

- SQL `select` combines  $\sigma$  and  $\pi$  of relational algebra

```
select ID, name project
 from instructor
 where dept_name = 'Physics'
 and salary > 90000;
```

*select // condition*

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL

- `select` does not eliminate duplicates

```
select dept_name
 from instructor;
```



| <i>ID</i> | <i>name</i> | <i>dept_name</i>  | <i>salary</i> |
|-----------|-------------|-------------------|---------------|
| 10101     | Srinivasan  | <u>Comp. Sci.</u> | 65000         |
| 12121     | Wu          | Finance           | 90000         |
| 15151     | Mozart      | Music             | 40000         |
| 22222     | Einstein    | Physics           | 95000         |
| 32343     | El Said     | History           | 60000         |
| 33456     | Gold        | Physics           | 87000         |
| 45565     | Katz        | <u>Comp. Sci.</u> | 75000         |
| 58583     | Califieri   | History           | 62000         |
| 76543     | Singh       | Finance           | 80000         |
| 76766     | Crick       | Biology           | 72000         |
| 83821     | Brandt      | <u>Comp. Sci.</u> | 92000         |
| 98345     | Kim         | Elec. Eng.        | 80000         |

# Queries in SQL

- `select` does not eliminate duplicates

```
select dept_name
from instructor;
```

- Explicitly specify that duplicates should be removed — additional computation, so avoid if not needed

```
select distinct dept_name
from instructor;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL

- Select all attributes using \*

```
select *
 from instructor
 where dept_name = 'Physics';
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL

- Select all attributes using `*`

```
select *
 from instructor
 where dept_name = 'Physics';
```

- Can do arithmetic on attributes — get monthly salary

```
select ID,name,salary/12
 from instructor;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL

- Multiple tables in `from` — cartesian product

```
select *
 from instructor, teaches;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

| <i>ID</i> | <i>course_id</i> | <i>sec_id</i> | <i>semester</i> | <i>year</i> |
|-----------|------------------|---------------|-----------------|-------------|
| 10101     | CS-101           | 1             | Fall            | 2017        |
| 10101     | CS-315           | 1             | Spring          | 2018        |
| 10101     | CS-347           | 1             | Fall            | 2017        |
| 12121     | FIN-201          | 1             | Spring          | 2018        |
| 15151     | MU-199           | 1             | Spring          | 2018        |
| 22222     | PHY-101          | 1             | Fall            | 2017        |
| 32343     | HIS-351          | 1             | Spring          | 2018        |
| 45565     | CS-101           | 1             | Spring          | 2018        |
| 45565     | CS-319           | 1             | Spring          | 2018        |
| 76766     | BIO-101          | 1             | Summer          | 2017        |
| 76766     | BIO-301          | 1             | Summer          | 2018        |
| 83821     | CS-190           | 1             | Spring          | 2017        |
| 83821     | CS-190           | 2             | Spring          | 2017        |
| 83821     | CS-319           | 2             | Spring          | 2018        |
| 98345     | EE-181           | 1             | Spring          | 2017        |



# Queries in SQL

- Multiple tables in `from` — cartesian product

```
select *
 from instructor, teaches;
```

- Use `where` to implement join

```
select *
 from instructor, teaches
 where instructor.ID = teaches.ID;
```

$$\sigma_{\theta} (T_1 \times T_2)$$

$T_1 \bowtie_{\theta} T_2$

| ID    | name       | dept_name  | salary |
|-------|------------|------------|--------|
| 10101 | Srinivasan | Comp. Sci. | 65000  |
| 12121 | Wu         | Finance    | 90000  |
| 15151 | Mozart     | Music      | 40000  |
| 22222 | Einstein   | Physics    | 95000  |
| 32343 | El Said    | History    | 60000  |
| 33456 | Gold       | Physics    | 87000  |
| 45565 | Katz       | Comp. Sci. | 75000  |
| 58583 | Califieri  | History    | 62000  |
| 76543 | Singh      | Finance    | 80000  |
| 76766 | Crick      | Biology    | 72000  |
| 83821 | Brandt     | Comp. Sci. | 92000  |
| 98345 | Kim        | Elec. Eng. | 80000  |

| ID    | course_id | sec_id | semester | year |
|-------|-----------|--------|----------|------|
| 10101 | CS-101    | 1      | Fall     | 2017 |
| 10101 | CS-315    | 1      | Spring   | 2018 |
| 10101 | CS-347    | 1      | Fall     | 2017 |
| 12121 | FIN-201   | 1      | Spring   | 2018 |
| 15151 | MU-199    | 1      | Spring   | 2018 |
| 22222 | PHY-101   | 1      | Fall     | 2017 |
| 32343 | HIS-351   | 1      | Spring   | 2018 |
| 45565 | CS-101    | 1      | Spring   | 2018 |
| 45565 | CS-319    | 1      | Spring   | 2018 |
| 76766 | BIO-101   | 1      | Summer   | 2017 |
| 76766 | BIO-301   | 1      | Summer   | 2018 |
| 83821 | CS-190    | 1      | Spring   | 2017 |
| 83821 | CS-190    | 2      | Spring   | 2017 |
| 83821 | CS-319    | 2      | Spring   | 2018 |
| 98345 | EE-181    | 1      | Spring   | 2017 |

# Queries in SQL

- Multiple tables in `from` — cartesian product

```
select *
 from instructor, teaches;
```

- Use `where` to implement join

```
select *
 from instructor, teaches
 where instructor.ID = teaches.ID;
```

- Choose output columns

```
select name, course_id
 from instructor, teaches
 where instructor.ID = teaches.ID;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

| <i>ID</i> | <i>course_id</i> | <i>sec_id</i> | <i>semester</i> | <i>year</i> |
|-----------|------------------|---------------|-----------------|-------------|
| 10101     | CS-101           | 1             | Fall            | 2017        |
| 10101     | CS-315           | 1             | Spring          | 2018        |
| 10101     | CS-347           | 1             | Fall            | 2017        |
| 12121     | FIN-201          | 1             | Spring          | 2018        |
| 15151     | MU-199           | 1             | Spring          | 2018        |
| 22222     | PHY-101          | 1             | Fall            | 2017        |
| 32343     | HIS-351          | 1             | Spring          | 2018        |
| 45565     | CS-101           | 1             | Spring          | 2018        |
| 45565     | CS-319           | 1             | Spring          | 2018        |
| 76766     | BIO-101          | 1             | Summer          | 2017        |
| 76766     | BIO-301          | 1             | Summer          | 2018        |
| 83821     | CS-190           | 1             | Spring          | 2017        |
| 83821     | CS-190           | 2             | Spring          | 2017        |
| 83821     | CS-319           | 2             | Spring          | 2018        |
| 98345     | EE-181           | 1             | Spring          | 2017        |

- Renaming for self join

| <i>ID</i> | <i>course_id</i> | <i>sec_id</i> | <i>semester</i> | <i>year</i> |
|-----------|------------------|---------------|-----------------|-------------|
| 10101     | CS-101           | 1             | Fall            | 2017        |
| 10101     | CS-315           | 1             | Spring          | 2018        |
| 10101     | CS-347           | 1             | Fall            | 2017        |
| 12121     | FIN-201          | 1             | Spring          | 2018        |
| 15151     | MU-199           | 1             | Spring          | 2018        |
| 22222     | PHY-101          | 1             | Fall            | 2017        |
| 32343     | HIS-351          | 1             | Spring          | 2018        |
| 45565     | CS-101           | 1             | Spring          | 2018        |
| 45565     | CS-319           | 1             | Spring          | 2018        |
| 76766     | BIO-101          | 1             | Summer          | 2017        |
| 76766     | BIO-301          | 1             | Summer          | 2018        |
| 83821     | CS-190           | 1             | Spring          | 2017        |
| 83821     | CS-190           | 2             | Spring          | 2017        |
| 83821     | CS-319           | 2             | Spring          | 2018        |
| 98345     | EE-181           | 1             | Spring          | 2017        |

# Queries in SQL

## ■ Renaming for self join

```
select *
 from teaches as T1, teaches as T2
 where T1.ID = T2.ID and
 T1.course_id <> T2.course_id;
```

| <i>ID</i> | <i>course_id</i> | <i>sec_id</i> | <i>semester</i> | <i>year</i> |
|-----------|------------------|---------------|-----------------|-------------|
| 10101     | CS-101           | 1             | Fall            | 2017        |
| 10101     | CS-315           | 1             | Spring          | 2018        |
| 10101     | CS-347           | 1             | Fall            | 2017        |
| 12121     | FIN-201          | 1             | Spring          | 2018        |
| 15151     | MU-199           | 1             | Spring          | 2018        |
| 22222     | PHY-101          | 1             | Fall            | 2017        |
| 32343     | HIS-351          | 1             | Spring          | 2018        |
| 45565     | CS-101           | 1             | Spring          | 2018        |
| 45565     | CS-319           | 1             | Spring          | 2018        |
| 76766     | BIO-101          | 1             | Summer          | 2017        |
| 76766     | BIO-301          | 1             | Summer          | 2018        |
| 83821     | CS-190           | 1             | Spring          | 2017        |
| 83821     | CS-190           | 2             | Spring          | 2017        |
| 83821     | CS-319           | 2             | Spring          | 2018        |
| 98345     | EE-181           | 1             | Spring          | 2017        |

# Queries in SQL

## ■ Renaming for self join

```
select *
 from teaches as T1, teaches as T2
 where T1.ID = T2.ID and
 T1.course_id <> T2.course_id;
```

## ■ Filter output

```
select distinct T1.ID
 from teaches as T1, teaches as T2
 where T1.ID = T2.ID and
 T1.course_id <> T2.course_id;
```

| <i>ID</i> | <i>course_id</i> | <i>sec_id</i> | <i>semester</i> | <i>year</i> |
|-----------|------------------|---------------|-----------------|-------------|
| 10101     | CS-101           | 1             | Fall            | 2017        |
| 10101     | CS-315           | 1             | Spring          | 2018        |
| 10101     | CS-347           | 1             | Fall            | 2017        |
| 12121     | FIN-201          | 1             | Spring          | 2018        |
| 15151     | MU-199           | 1             | Spring          | 2018        |
| 22222     | PHY-101          | 1             | Fall            | 2017        |
| 32343     | HIS-351          | 1             | Spring          | 2018        |
| 45565     | CS-101           | 1             | Spring          | 2018        |
| 45565     | CS-319           | 1             | Spring          | 2018        |
| 76766     | BIO-101          | 1             | Summer          | 2017        |
| 76766     | BIO-301          | 1             | Summer          | 2018        |
| 83821     | CS-190           | 1             | Spring          | 2017        |
| 83821     | CS-190           | 2             | Spring          | 2017        |
| 83821     | CS-319           | 2             | Spring          | 2018        |
| 98345     | EE-181           | 1             | Spring          | 2017        |

## ■ Set operations

| <i>course_id</i> | <i>sec_id</i> | <i>semester</i> | <i>year</i> | <i>building</i> | <i>room_number</i> | <i>time_slot_id</i> |
|------------------|---------------|-----------------|-------------|-----------------|--------------------|---------------------|
| BIO-101          | 1             | Summer          | 2017        | Painter         | 514                | B                   |
| BIO-301          | 1             | Summer          | 2018        | Painter         | 514                | A                   |
| CS-101           | 1             | Fall            | 2017        | Packard         | 101                | H                   |
| CS-101           | 1             | Spring          | 2018        | Packard         | 101                | F                   |
| CS-190           | 1             | Spring          | 2017        | Taylor          | 3128               | E                   |
| CS-190           | 2             | Spring          | 2017        | Taylor          | 3128               | A                   |
| CS-315           | 1             | Spring          | 2018        | Watson          | 120                | D                   |
| CS-319           | 1             | Spring          | 2018        | Watson          | 100                | B                   |
| CS-319           | 2             | Spring          | 2018        | Taylor          | 3128               | C                   |
| CS-347           | 1             | Fall            | 2017        | Taylor          | 3128               | A                   |
| EE-181           | 1             | Spring          | 2017        | Taylor          | 3128               | C                   |
| FIN-201          | 1             | Spring          | 2018        | Packard         | 101                | B                   |
| HIS-351          | 1             | Spring          | 2018        | Painter         | 514                | C                   |
| MU-199           | 1             | Spring          | 2018        | Packard         | 101                | D                   |
| PHY-101          | 1             | Fall            | 2017        | Watson          | 100                | A                   |

# Queries in SQL

- Set operations

- Union

```
(select course_id from section
 where sem = 'Fall'
 and year = '2017')
```

```
union
```

```
(select course_id from section
 where sem = 'Spring'
 and year = '2018');
```

| <i>course_id</i> | <i>sec_id</i> | <i>semester</i> | <i>year</i> | <i>building</i> | <i>room_number</i> | <i>time_slot_id</i> |
|------------------|---------------|-----------------|-------------|-----------------|--------------------|---------------------|
| BIO-101          | 1             | Summer          | 2017        | Painter         | 514                | B                   |
| BIO-301          | 1             | Summer          | 2018        | Painter         | 514                | A                   |
| CS-101           | 1             | Fall            | 2017        | Packard         | 101                | H                   |
| CS-101           | 1             | Spring          | 2018        | Packard         | 101                | F                   |
| CS-190           | 1             | Spring          | 2017        | Taylor          | 3128               | E                   |
| CS-190           | 2             | Spring          | 2017        | Taylor          | 3128               | A                   |
| CS-315           | 1             | Spring          | 2018        | Watson          | 120                | D                   |
| CS-319           | 1             | Spring          | 2018        | Watson          | 100                | B                   |
| CS-319           | 2             | Spring          | 2018        | Taylor          | 3128               | C                   |
| CS-347           | 1             | Fall            | 2017        | Taylor          | 3128               | A                   |
| EE-181           | 1             | Spring          | 2017        | Taylor          | 3128               | C                   |
| FIN-201          | 1             | Spring          | 2018        | Packard         | 101                | B                   |
| HIS-351          | 1             | Spring          | 2018        | Painter         | 514                | C                   |
| MU-199           | 1             | Spring          | 2018        | Packard         | 101                | D                   |
| PHY-101          | 1             | Fall            | 2017        | Watson          | 100                | A                   |

# Queries in SQL

- Set operations
- Union
- Intersection

```
(select course_id from section
 where sem = 'Fall'
 and year = '2017')
intersect
(select course_id from section
 where sem = 'Spring'
 and year = '2018');
```

| <i>course_id</i> | <i>sec_id</i> | <i>semester</i> | <i>year</i> | <i>building</i> | <i>room_number</i> | <i>time_slot_id</i> |
|------------------|---------------|-----------------|-------------|-----------------|--------------------|---------------------|
| BIO-101          | 1             | Summer          | 2017        | Painter         | 514                | B                   |
| BIO-301          | 1             | Summer          | 2018        | Painter         | 514                | A                   |
| CS-101           | 1             | Fall            | 2017        | Packard         | 101                | H                   |
| CS-101           | 1             | Spring          | 2018        | Packard         | 101                | F                   |
| CS-190           | 1             | Spring          | 2017        | Taylor          | 3128               | E                   |
| CS-190           | 2             | Spring          | 2017        | Taylor          | 3128               | A                   |
| CS-315           | 1             | Spring          | 2018        | Watson          | 120                | D                   |
| CS-319           | 1             | Spring          | 2018        | Watson          | 100                | B                   |
| CS-319           | 2             | Spring          | 2018        | Taylor          | 3128               | C                   |
| CS-347           | 1             | Fall            | 2017        | Taylor          | 3128               | A                   |
| EE-181           | 1             | Spring          | 2017        | Taylor          | 3128               | C                   |
| FIN-201          | 1             | Spring          | 2018        | Packard         | 101                | B                   |
| HIS-351          | 1             | Spring          | 2018        | Painter         | 514                | C                   |
| MU-199           | 1             | Spring          | 2018        | Packard         | 101                | D                   |
| PHY-101          | 1             | Fall            | 2017        | Watson          | 100                | A                   |



# Queries in SQL

- Set operations
- Union
- Intersection
- Set difference

```
(select course_id from section
 where sem = 'Fall'
 and year = '2017')
```

**except**

```
(select course_id from section
 where sem = 'Spring'
 and year = '2018');
```

| course_id | sec_id | semester | year | building | room_number | time_slot_id |
|-----------|--------|----------|------|----------|-------------|--------------|
| BIO-101   | 1      | Summer   | 2017 | Painter  | 514         | B            |
| BIO-301   | 1      | Summer   | 2018 | Painter  | 514         | A            |
| CS-101    | 1      | Fall     | 2017 | Packard  | 101         | H            |
| CS-101    | 1      | Spring   | 2018 | Packard  | 101         | F            |
| CS-190    | 1      | Spring   | 2017 | Taylor   | 3128        | E            |
| CS-190    | 2      | Spring   | 2017 | Taylor   | 3128        | A            |
| CS-315    | 1      | Spring   | 2018 | Watson   | 120         | D            |
| CS-319    | 1      | Spring   | 2018 | Watson   | 100         | B            |
| CS-319    | 2      | Spring   | 2018 | Taylor   | 3128        | C            |
| CS-347    | 1      | Fall     | 2017 | Taylor   | 3128        | A            |
| EE-181    | 1      | Spring   | 2017 | Taylor   | 3128        | C            |
| FIN-201   | 1      | Spring   | 2018 | Packard  | 101         | B            |
| HIS-351   | 1      | Spring   | 2018 | Painter  | 514         | C            |
| MU-199    | 1      | Spring   | 2018 | Packard  | 101         | D            |
| PHY-101   | 1      | Fall     | 2017 | Watson   | 100         | A            |

# Queries in SQL — aggregate operations

- SQL allows arithmetic across rows

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — aggregate operations

- SQL allows arithmetic across rows
- Count the number of rows

```
select count(name) count(*)
 from instructor;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — aggregate operations

- SQL allows arithmetic across rows

- Count the number of rows

```
select count(name)
 from instructor;
```

- Count departments?

```
select count(dept_name)
 from instructor;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — aggregate operations

- SQL allows arithmetic across rows

- Count the number of rows

```
select count(name)
 from instructor;
```

- Count departments?

```
select count(dept_name)
 from instructor;
```

- Avoid duplicates. Instead

```
select count(distinct dept_name)
 from instructor;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — aggregate operations

## ■ Other functions

- `min`
- `max`
- `sum`
- `avg`

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — aggregate operations

- Other functions

- `min`

- `max`

- `sum`

- `avg`

- Average salary in Comp Sc dept

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — aggregate operations

- Other functions

- min
- max
- sum
- avg

- Average salary in Comp Sc dept

```
select avg(salary)
from instructor
where dept_name = 'Comp. Sci.';
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |



# Queries in SQL — grouping

- Filter and calculate
- Extract the average value in each department

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — grouping

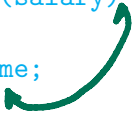
- Filter and calculate
- Extract the average value in each department
  - Group rows by department name
  - Report average in each group of rows

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — grouping

- Filter and calculate
- Extract the average value in each department
  - Group rows by department name
  - Report average in each group of rows

```
select dept_name, avg(salary)
from instructor
group by dept_name;
```



| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — grouping

- Filter and calculate
- Extract the average value in each department
  - Group rows by department name
  - Report average in each group of rows

```
select dept_name, avg(salary)
from instructor
group by dept_name;
```

- Attributes in `select` must appear in `group by`
  - Should be the same across the entire group

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

## Queries in SQL — filtering groups

- Use `having` to specify a condition on groups

```
select dept_name, avg(salary)
 from instructor
 group by dept_name
 having max(salary) > 80000;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — filtering groups

- Use `having` to specify a condition on groups

```
select dept_name, avg(salary)
 from instructor
 group by dept_name
 having max(salary) > 80000;
```

- Condition is evaluated with respect to groups

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — sorting the output

- Sort in ascending order

```
select name
 from instructor
 where dept_name = 'Physics'
 order by name;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |

# Queries in SQL — sorting the output

- Sort in ascending order

```
select name
 from instructor
 where dept_name = 'Physics'
 order by name;
```

- Add `desc` for descending order

```
select name
 from instructor
 where dept_name = 'Physics'
 order by name desc;
```

| <i>ID</i> | <i>name</i> | <i>dept_name</i> | <i>salary</i> |
|-----------|-------------|------------------|---------------|
| 10101     | Srinivasan  | Comp. Sci.       | 65000         |
| 12121     | Wu          | Finance          | 90000         |
| 15151     | Mozart      | Music            | 40000         |
| 22222     | Einstein    | Physics          | 95000         |
| 32343     | El Said     | History          | 60000         |
| 33456     | Gold        | Physics          | 87000         |
| 45565     | Katz        | Comp. Sci.       | 75000         |
| 58583     | Califieri   | History          | 62000         |
| 76543     | Singh       | Finance          | 80000         |
| 76766     | Crick       | Biology          | 72000         |
| 83821     | Brandt      | Comp. Sci.       | 92000         |
| 98345     | Kim         | Elec. Eng.       | 80000         |