Aggregation

So far, all SQL expressions have referred to the values in a single row at a time

```
select [columns] from [table] where [expression] order by [expression];
```

An aggregate function in the [columns] clause computes a value from a group of rows.

```
create table animals as
select "dog" as kind, 4 as legs, 20 as weight union
select "cat", 4, 10 union
select "ferret", 4, 10 union
select "parrot", 2, 6 union
select "penguin", 2, 10 union
select "t-rex", 2, 12000;
```

```
select max(legs) from animals;
```

```
animals:
kind  legs  weight
---  ----  ----
dog   4     20
```

Mixing Aggregate Functions and Single Values

An aggregate function also selects some row in the table to supply the values of columns that are not aggregated. In the case of max or min, this row is that of the max or min value. Otherwise, it is arbitrary.

```
select max(weight), kind from animals;
select min(kind), kind from animals;
select max(legs), kind from animals;
select avg(weight), kind from animals;
```

```
animals:
kind  legs  weight
---  ----  ----
dog   4     20
```

Discussion Question

What are all the kinds of animals that have the maximal number of legs?

```
create table animals as
select "dog" as kind, 4 as legs, 20 as weight union
select "cat", 4, 10 union
select "ferret", 4, 10 union
select "parrot", 2, 6 union
select "penguin", 2, 10 union
select "t-rex", 2, 12000;
```

Groups

Rows in a table can be grouped, and aggregation is performed on each group.

```
select [columns] from [table] group by [expression] having [expression];
```

The number of groups is the number of unique values of an expression.

```
select legs, max(weight) from animals group by legs;
```

```
animals:
legs  max(weight)
-----  -----------
4      12000
```
Rows in a table can be grouped, and aggregation is performed on each group

```
select [columns] from [table] group by [expression] having [expression];
```

A having clause filters the set of groups that are aggregated

```
select weight/legs, count(*) from animals group by weight/legs having count(*) > 1
```

<table>
<thead>
<tr>
<th>weight/legs</th>
<th>count(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>6000</td>
<td>1</td>
</tr>
</tbody>
</table>

**Discussion Question**
What's the maximum difference between leg count for two animals with the same weight?

---

**Example: Big Game**

[Text]

---

**Create Table and Drop Table**

**Create Table**

CREATE TABLE expression syntax:

```sql
CREATE TABLE animals (weight INT, legs INT);
```

**Drop Table**

```sql
DROP TABLE animals;
```
**Modifying Tables**

**Insert**

For a table $t$ with two columns...

To insert into one column:

```
INSERT INTO t(column) VALUES (value);
```

To insert into both columns:

```
INSERT INTO t VALUES (value0, value1);
```

**Update**

Update sets all entries in certain columns to new values, just for some subset of rows.

(Demo)

**Delete**

Delete removes some or all rows from a table.

(Demo)

**Python and SQL**

(Demo)

**SQL Injection Attack**

(Demo)
A Program Vulnerable to a SQL Injection Attack

```python
name = "Robert'); DROP TABLE Students; --"

# INSERT INTO Students VALUES ('name');
```

```sql
db.executescript(cmd)
```

![Database Connections](https://xkcd.com/327/)

Casino Blackjack

Player:

Dealer: