Aggregation

Announcements

Aggregate Functions

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 "penguin" , 2 , 10 union
    select "t-rex" , 2 , 12000;
```

```
animals:

<table>
<thead>
<tr>
<th>kind</th>
<th>legs</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>cat</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>ferret</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>penguin</td>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>
```

```
create table animals as
    select max(legs), kind from animals;
```

```
animals:

<table>
<thead>
<tr>
<th>kind</th>
<th>legs</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>4</td>
</tr>
</tbody>
</table>
```

```
create table animals as
    select min(kind), kind from animals;
```

```
animals:

<table>
<thead>
<tr>
<th>kind</th>
<th>kind</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>dog</td>
</tr>
</tbody>
</table>
```

```
create table animals as
    select max(kind) as kind, legs from animals;
```

```
animals:

<table>
<thead>
<tr>
<th>kind</th>
<th>legs</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>4</td>
</tr>
</tbody>
</table>
```

```
create table animals as
    select avg(weight), kind from animals;
```

```
animals:

<table>
<thead>
<tr>
<th>kind</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>20</td>
</tr>
</tbody>
</table>
```

Discussion Question

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

```
select max(kind), legs from animals
```
Selecting Groups

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

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

A `having` clause filters the set of groups that are aggregated

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

Discussion Question

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