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 "parrot" , 2 , 6 union
select "penguin" , 2 , 10 union
select "t-rex" , 2 , 12000;
select max(legs) from animals;
```

```
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>parrot</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>penguin</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>t-rex</td>
<td>2</td>
<td>12000</td>
</tr>
</tbody>
</table>

max(legs) (Demo) 4
```
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.

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

(Demo)

```sql
create table animals as
  select "dog" as kind, 4 as legs, 20 as weight union
  select "cat" as kind, 4 as legs, 10 union
  select "ferret" as kind, 4 as legs, 10 union
  select "parrot" as kind, 2 as legs, 6 union
  select "penguin" as kind, 2 as legs, 10 union
  select "t-rex" as kind, 2 as legs, 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>parrot</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>penguin</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>t-rex</td>
<td>2</td>
<td>12000</td>
</tr>
</tbody>
</table>

Discussion Question

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

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];
```

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

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

Groups

animals:

<table>
<thead>
<tr>
<th>legs</th>
<th>max(weight)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>12000</td>
</tr>
</tbody>
</table>

(Demo)
Selecting Groups

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

```
[expression] as [name], [expression] as [name], ...

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;
```

Discussion Question

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