2NF and 3NF

Comprehensive Version

- Takes account of all keys
- Primary key
- Candidate keys
- Prime Attribute: member of some key
- Non-Prime Attribute: member of no key
The Story of the LOTS Relation

- A LOT is a piece of land that someone owns.
- Years ago there was a State of the United States which had 2 counties: Earp and Kidd
- These two kept records about each lot within their boundaries.
- Each county gave its lots unique Lot Numbers (LOT#).

Story (cont.)

- Later the State took over database duties.
- It combined the county DBs and gave each lot a state ID#.
- Now LOTS has primary key ID#
- but (CNTY,LOT#) is a candidate key
- Each county still sets one tax rate for the whole county.
Fields in LOTS:

- **ID#**: The state's unique number for the lot
- **CNTY**: The county in which the lot is located.
- **LOT#**: The number given the lot by the county. Unique within the county.
- **AREA**: The area of the lot, in acres.
- **VAL**: The current value of the lot.
- **RATE**: The tax rate in percent on the lot.

<table>
<thead>
<tr>
<th>ID#</th>
<th>CNTY</th>
<th>LOT#</th>
<th>AREA</th>
<th>VAL</th>
<th>RATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>Ear p</td>
<td>100</td>
<td>0.5</td>
<td>5K</td>
<td>3.1%</td>
</tr>
<tr>
<td>222</td>
<td>Ki dd</td>
<td>100</td>
<td>1.1</td>
<td>11K</td>
<td>2.9%</td>
</tr>
<tr>
<td>333</td>
<td>Ear p</td>
<td>101</td>
<td>0.6</td>
<td>6K</td>
<td>3.1%</td>
</tr>
<tr>
<td>444</td>
<td>Ki dd</td>
<td>101</td>
<td>1.1</td>
<td>11K</td>
<td>2.9%</td>
</tr>
<tr>
<td>555</td>
<td>Ear p</td>
<td>200</td>
<td>0.7</td>
<td>7K</td>
<td>3.1%</td>
</tr>
<tr>
<td>666</td>
<td>Ki dd</td>
<td>200</td>
<td>1.3</td>
<td>13K</td>
<td>2.9%</td>
</tr>
<tr>
<td>777</td>
<td>Ear p</td>
<td>300</td>
<td>0.7</td>
<td>7K</td>
<td>3.1%</td>
</tr>
<tr>
<td>888</td>
<td>Ki dd</td>
<td>300</td>
<td>1.1</td>
<td>11K</td>
<td>2.9%</td>
</tr>
<tr>
<td>999</td>
<td>Ki dd</td>
<td>400</td>
<td>1.3</td>
<td>13K</td>
<td>2.9%</td>
</tr>
</tbody>
</table>

- In PK version, Lots is 2nf: key is one field.
- In full version, must take account of (CNTY,LOT#)
2NF in Comprehensive

- A relation R is in 2NF
  - if every nonprime attribute
  - is fully functionally dependent
  - on EVERY key of R
- Lots is not 2nf in comprehensive version.

2NF Problem

- Rate is partially dependent on Cnty.
- In PK version this is 3nf problem -- do you see why?
3NF in Comprehensive

- A relation R with functional dependencies F is in 3NF if
  - For each FD in F
    - LHS is a key (primary or candidate), or
    - RHS is prime (part of a key)
  
- New Problem with LOTS1 -- new FD revealed:
  - Area → Val
3NF Problem

- For Area $\rightarrow$ Val
  - Area not a key either PK or candidate
  - Val not part of any key
- So Lots1 is not 3NF
- Area $\rightarrow$ Val is the problem

<table>
<thead>
<tr>
<th>ID#</th>
<th>CNTY</th>
<th>LOT#</th>
<th>AREA</th>
<th>VAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>Earp</td>
<td>100</td>
<td>0.5</td>
<td>5K</td>
</tr>
<tr>
<td>222</td>
<td>Kidd</td>
<td>100</td>
<td>1.1</td>
<td>11K</td>
</tr>
<tr>
<td>333</td>
<td>Earp</td>
<td>101</td>
<td>0.6</td>
<td>6K</td>
</tr>
<tr>
<td>444</td>
<td>Kidd</td>
<td>101</td>
<td>1.1</td>
<td>11K</td>
</tr>
<tr>
<td>555</td>
<td>Earp</td>
<td>200</td>
<td>0.7</td>
<td>7K</td>
</tr>
<tr>
<td>666</td>
<td>Kidd</td>
<td>200</td>
<td>1.3</td>
<td>13K</td>
</tr>
<tr>
<td>777</td>
<td>Earp</td>
<td>300</td>
<td>0.7</td>
<td>7K</td>
</tr>
<tr>
<td>888</td>
<td>Kidd</td>
<td>300</td>
<td>1.1</td>
<td>11K</td>
</tr>
<tr>
<td>999</td>
<td>Kidd</td>
<td>400</td>
<td>1.3</td>
<td>13K</td>
</tr>
</tbody>
</table>

The usual solution

\[ \pi_{\text{area, val} (\text{LOTS})} \]

\[ \pi_{\text{all except val} (\text{LOTS1})} \]
All dependencies preserved