Normalization

Database normalization was conceived by E.F. Codd. Prevent complexity from becoming unmanageable and keeping copies of things you don’t know you already had somewhere in storage. Permit data to be precisely found.

1. Reduce unnecessary insertions, updates and deletions
2. Reduce redundancy
3. Prevent restructuring the data as new information is added
4. Make the data more clear to users

Maximize atomicity (search accuracy) , minimize redundancy (efficiency.

Involves decomposing a table into less redundant (and smaller) tables without losing information, and then linking the data back together by defining foreign keys in the old table referencing the primary keys of the new ones. The objective is to isolate data so that additions, deletions, and modifications of an attribute can be made in just one table and then propagated through the rest of the database using the defined foreign keys.

Here is a table someone created in Excel.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Name | Age | Pet\_1 | Pet\_2 | Pet\_Name\_1 | Pet\_name\_2 |
| Heather | 10 | Dog | Cat | Leo | Thomas |
| Bobby | 12 | Dog |  | Leo |  |
| Rachel | 10 | Cat |  | Fluff |  |
| Jimmy | 11 | Dog |  | Kimba |  |
| Heather | 11 | Cat |  | Leo |  |
| Lola | 10 | Cat |  | Thomas |  |

1NF

Make values atomic. Make sure no table contains multiple columns for the same data. Each should have a primary key that distinguishes it as unique. The primary key is usually a single column, but sometimes more than one column can be combined to create a single primary key. Using the rules of first normal form, there may be redundant information across multiple rows, but each row will be unique.

STUDENTS

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Student\_ID** | Name | Age | Pet | Pet Name |
| **101** | **Heather** | **10** | Dog | Leo |
| **101** | **Heather** | **10** | Cat | Thomas |
| **465** | Bobby | 12 | Dog | Leo |
| **17** | Rachel | 10 | Cat | Fluff |
| **22** | Jimmy | 11 | Dog | Kimba |
| **156** | Heather | 11 | Cat | Leo |
| **24** | Lola | 10 | Cat | Thomas |

What columns would you have to use to uniquely identify one row?

2NF

1NF is somewhat searchable but had redundant data and therefore is an inefficient use of space. It would be helpful to split out the pets into an independent table, and match them up using the student names as foreign keys.

STUDENTS

|  |  |  |
| --- | --- | --- |
| **Student\_ID** | Name | Age |
| **101** | Heather | 10 |
| **465** | Bobby | 12 |
| **17** | Rachel | 10 |
| **22** | Jimmy | 11 |
| **156** | Heather | 11 |
| **24** | Lola | 10 |

PETS

|  |  |  |  |
| --- | --- | --- | --- |
| **Pet\_ID** | Pet Type | Pet Name | Owner |
| **923** | Dog | Leo | Heather |
| **923** | Dog | Leo | Bobby |
| **71** | Cat | Thomas | Heather |
| **8** | Cat | Fluffy | Rachel |
| **101** | Cat | Leo | Heather |
| **452** | Dog | Kimba | Jimmy |
| **359** | Cat | Thomas | Lola |

3NF

Avoids repeating the age of the student with two pets, so better. But dogs and cats are repeated several times in the pets table. Uncertainty arises when certain data is the same.

Table Pets Suggestion No. 1

Type and Owner

|  |  |  |
| --- | --- | --- |
| **Type** | Pet Name | **Owner** |
| **Dog** | Leo | **Heather** |
| **Dog** | Leo | **Bobby** |
| **Cat** | Thomas | **Heather** |
| **Cat** | Fluffy | **Rachel** |
| **Dog** | Rex | **Bobby** |
| **Cat** | Leo | **Heather** |
| **Dog** | Kimba | **Jimmy** |
| **Cat** | Thomas | **Lola** |

|  |  |
| --- | --- |
| **Suggestion** | **Problem** |
| 1 | 2 Heathers with a Cat |
| 2 | 2 Leos that are Dogs |
| 3 | 2 Heathers with a Leo |

Table Pets Suggestion No. 2

Type and Name

|  |  |  |
| --- | --- | --- |
| **Type** | **Pet Name** | Owner |
| **Dog** | **Leo** | Heather |
| **Dog** | **Leo** | Bobby |
| **Cat** | **Thomas** | Heather |
| **Cat** | **Fluffy** | Rachel |
| **Cat** | **Leo** | Heather |
| **Dog** | **Kimba** | Jimmy |
| **Cat** | **Thomas** | Lola |

|  |  |  |
| --- | --- | --- |
| Type | **Pet Name** | **Owner** |
| Dog | **Leo** | **Heather** |
| Dog | **Leo** | **Bobby** |
| Cat | **Thomas** | **Heather** |
| Cat | **Fluff** | **Rachel** |
| Cat | **Leo** | **Heather** |
| Dog | **Kimba** | **Jimmy** |
| Cat | **Thomas** | **Lola** |

Table Pets Suggestion No. 3

Name and Owner

Solution

Third normal form makes sure each non-key element in each row provides information about the key in the row. In order to establish an unambiguous unique identifier for each pet, it is useful to include a unique primary key that distinguishes each pet from all the others

STUDENTS: 3NF

|  |  |  |
| --- | --- | --- |
| **Student\_ID** | Name | Age |
| **101** | Heather | 10 |
| **465** | Bobby | 12 |
| **17** | Rachel | 10 |
| **22** | Jimmy | 11 |
| **89** | Heather | 11 |
| **24** | Lola | 10 |
| **156** | Heather | 11 |

|  |  |
| --- | --- |
| **Suggestion** | **Problem** |
| 1 | 2 Heathers with a Cat |
|  |  |
|  |  |

PETS: 3NF

|  |  |  |  |
| --- | --- | --- | --- |
| **Pet\_ID** | **Type** | Pet Name | **Owner** |
| **923** | Dog | Leo | Heather |
| **923** | Dog | Leo | Bobby |
| **71** | Cat | Thomas | Heather |
| **8** | Cat | Fluff | Rachel |
| **101** | Cat | Leo | Heather |
| **452** | Dog | Kimba | Jimmy |
| **339** | Cat | Thomas | Lola |

4NF

Which pet owned by which student is clear But what about Heather and Bobby co-owning Leo? We need one more table.

STUDENTS: 4NF

|  |  |  |
| --- | --- | --- |
| **Student\_ID** | Name | Age |
| **101** | Heather | 10 |
| **465** | Bobby | 12 |
| **17** | Rachel | 10 |
| **22** | Jimmy | 11 |
| **89** | Heather | 11 |
| **24** | Lola | 10 |
| **156** | Heather | 11 |

OWNERSHIP: 4NF

|  |  |
| --- | --- |
| **Pet\_ID** | **Own\_ID** |
| **923** | **101** |
| **923** | **465** |
| **71** | **101** |
| **8** | **17** |
| **101** | **89** |
| **452** | **22** |
| **339** | **24** |

PETS: 4Nf

|  |  |  |
| --- | --- | --- |
| **Pet\_ID** | Pet Type | Pet Name |
| **923** | Dog | Leo |
| **923** | Dog | Leo |
| **71** | Cat | Thomas |
| **8** | Cat | Fluffy |
| **101** | Cat | Leo |
| **452** | Dog | Kimba |
| **359** | Cat | Thomas |