**正規化について**

正規化とは、使用する情報を目的別に分類してテーブルを作成すること。

正規化技法：非正規形データを第一正規化にし、それを元に第二正規化をし、最後に第

　　　　　　三正規化をする。

　　　　（正規化にはボイスコット・第四・第五などあるが、基本情報では第三までが問われる。）

伝票NO　X1001

顧客NO　　K001

顧客名　　　情報太郎　様

売上伝票

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 商品番号 | 品名 | 単価 | 数量 | 金額 |
| A001 | プリンタ | 50,000 | 5 | 250,000 |
| A002 | スキャナ | 60,000 | 8 | 480,000 |
| A003 | パソコン | 300,000 | 3 | 900,000 |

非正規化：必要な情報をピックアップする。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 伝票ＮＯ | 顧客ＮＯ | 顧客名 | 商品番号 | 品名 | 単価 | 数量 | 金額 |
| X1001 | Ｋ００１ | 情報太郎 | A001  A002  A003 | プリンタ  スキャナ  パソコン | 50,000  60,000  300,000 | 5  8  3 | 250,000  480,000  900,000 |

第一正規化：繰り返しデータがない状態にする。

　　　　　　・1つの伝票Ｎｏに対し、複数個の商品が含まれている。これが、繰り返しデータである。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 伝票ＮＯ | 顧客ＮＯ | 顧客名 | 商品番号 | 品名 | 単価 | 数量 | 金額 |
| X1001 | Ｋ００１ | 情報太郎 | A001 | プリンタ | 50,000 | 5 | 250,000 |
| X1001 | Ｋ００１ | 情報太郎 | A002 | スキャナ | 60,000 | 8 | 480,000 |
| X1001 | Ｋ００１ | 情報太郎 | A003 | パソコン | 300,000 | 3 | 900,000 |

第二正規化：主キー以外の項目が主キーによって特定できる状態。（完全関数従属である表）

　　　　　　・主キーとは、行を特定するときに利用する項目であり、ユニークである。

　　　　　　　（複数の項目を合わせて主キーとする場合もある。）

　　　　　　　ユニークとは、重複しないということ。

　　　　　　　たとえば「売上伝票」が何十枚もある場合に、情報太郎さんの購入したスキャナに関する情報を

探したい場合を考えると、伝票ＮＯがＸ００１の商品番号がＡ００２で探すことになる。この場

合の「伝票ＮＯ」と「商品番号」が主キーとなる。

手順１：主キーを見つける。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 伝票ＮＯ | 顧客ＮＯ | 顧客名 | 商品番号 | 品名 | 単価 | 数量 | 金額 |
| X1001 | Ｋ００１ | 情報太郎 | A001 | プリンタ | 50,000 | 5 | 250,000 |
| X1001 | Ｋ００１ | 情報太郎 | A002 | スキャナ | 60,000 | 8 | 480,000 |
| X1001 | Ｋ００１ | 情報太郎 | A003 | パソコン | 300,000 | 3 | 900,000 |

手順２：主キーを個別に見た場合に、その項目によって特定されるものを区別する。

|  |  |  |
| --- | --- | --- |
| 商品番号 | 品名 | 単価 |
| A001 | プリンタ | 60,000 |
| A002 | スキャナ | 50,000 |
| A003 | パソコン | 300,000 |

|  |  |  |
| --- | --- | --- |
| 伝票ＮＯ | 顧客ＮＯ | 顧客名 |
| X1001 | Ｋ００１ | 情報太郎 |
| X1001 | Ｋ００１ | 情報太郎 |
| X1001 | Ｋ００１ | 情報太郎 |

手順３：手順２で使われていない項目を主キーと一緒にする。

|  |  |  |  |
| --- | --- | --- | --- |
| 伝票ＮＯ | 商品番号 | 数量 | 金額 |
| X1001 | A001 | 5 | 250,000 |
| X1001 | A002 | 8 | 480,000 |
| X1001 | A003 | 3 | 900,000 |

第三正規化：主キー以外の項目が主キーによってのみ特定できる状態。（推移的関数従属でない表）

　　　　　　考え方としては主キー以外の項目で、これがわかれば、これがわかるというものを見つけ分ける。

　　　　　　ここでの場合は、顧客ＮＯがわかれば、顧客名がわかる。このつながりがあるものは分離する。

　　　注）分離元と分離先を関係（リレーション）づけなければならないので、分離先のキーとなる項目を分離

元に付加する。

|  |  |
| --- | --- |
| 顧客ＮＯ | 顧客名 |
| Ｋ００１ | 情報太郎 |

|  |  |
| --- | --- |
| 伝票ＮＯ | 顧客ＮＯ |
| X1001 | Ｋ００１ |

|  |  |  |
| --- | --- | --- |
| 伝票ＮＯ | 商品番号 | 数量 |
| X1001 | A001 | 5 |
| X1001 | A002 | 8 |
| X1001 | A003 | 3 |

|  |  |  |
| --- | --- | --- |
| 商品番号 | 品名 | 単価 |
| A001 | プリンタ | 60,000 |
| A002 | スキャナ | 50,000 |
| A003 | パソコン | 300,000 |

　注）計算して求まるものは、テーブルから省く。