

- Table ở dạng chưa chuẩn hóa gọi là Unnormalized Form (UNF)
Table ở dạng này có thể gặp phải các bất thường khi cập nhật dữ liệu
VD:

Customer Number	Customer	Terms	Ship Via	Order Date	Product No.	Description	Quantity	Unit Price	Extended Amount
1454837	W.Coyote	Net 30	USPS	12/01/2008	SPR-2290	Super Strength Springs	2	24.00	\$48.00
1454837	W.Coyote	Net 30	USPS	12/01/2008	SRT-67	Foot Straps, leather	2	2.50	\$5.00
1454837	W.Coyote	Net 30	USPS	12/01/2008	HLM-45	Deluxe Crash Helmet	1	67.88	\$67.88
1454837	W.Coyote	Net 30	USPS	12/01/2008	SFR-1	Rocket, Solid fuel	1	128,200.4	\$128,200.4
1454837	W.Coyote	Net 30	USPS	12/01/2008	ELT-1	Emergency Location Transmitter	1	79.88	\$79.88

Table chưa chuẩn hóa – UNF

Các bất thường gặp phải khi cập nhật dữ liệu

- Khi thêm dữ liệu
 - Nếu khách hàng không mua hàng thì sẽ không thể thêm khách hàng đó vào CSDL
- Khi cập nhật dữ liệu
 - Muốn thay đổi địa chỉ của khách hàng thì phải thay đổi địa chỉ trên tất cả các hóa đơn → trên hóa đơn đã lưu trữ dư thừa địa chỉ của khách hàng
- Khi xóa dữ liệu
 - Nếu xóa hóa đơn của khách hàng thì tất cả dữ liệu về khách hàng đó cũng bị mất

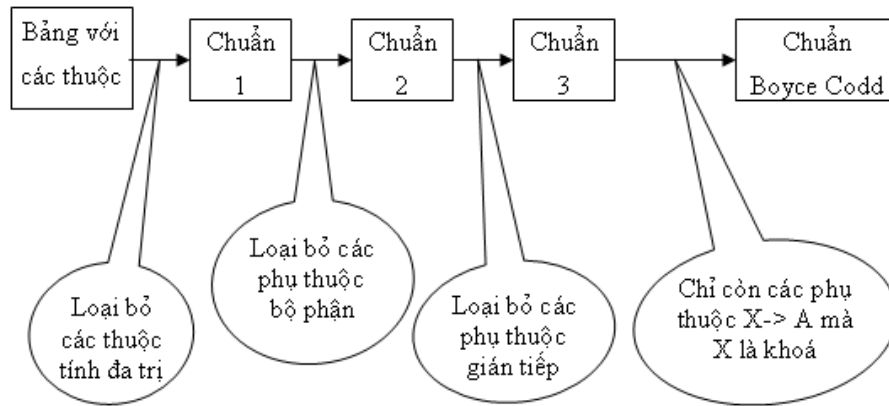
⇒ Cần phải chuẩn hóa dữ liệu

- Chuẩn hóa dữ liệu là quá trình tách table thành các table nhỏ hơn dựa vào các phụ thuộc hàm
- Mục đích chuẩn hóa:
 - Loại bỏ các dữ liệu dư thừa (nhưng không được làm mất thông tin)
 - Loại bỏ các bất thường khi cập nhật CSDL (insert, delete, update)
- Nhược điểm:
 - Làm tăng thời gian truy vấn

➤ Phụ thuộc hàm

Có 2 thuộc tính A và B

B được gọi là phụ thuộc hàm vào A nếu có không quá 1 giá trị của B liên kết với 1 giá trị của A



Các dạng chuẩn hóa (Normal Form)

VD: Trong database cần lưu các field

Mã hóa đơn, Ngày bán, Tổng tiền Hóa đơn,

Mã khách hàng, Tên Khách hàng, Số ĐT Khách hàng,

Mã Nhân viên, Số CMT của NV, Tên NV, Địa chỉ NV, Số ĐT Nhân viên

Mã thuốc, Tên thuốc, Công dụng, Ngày SX, Hạn SD, Số lượng,

Thành tiền

⇒ Các phụ thuộc hàm xảy ra

- Mã hóa đơn xác định duy nhất 1 hóa đơn
- Mã khách hàng xác định duy nhất 1 khách hàng
- Mã nhân viên xác định duy nhất 1 nhân viên
- Mã thuốc xác định duy nhất 1 loại thuốc
- 1 hóa đơn có thể có nhiều thuốc nhưng chỉ do 1 nhân viên bán và 1 người mua
- Số CMT của NV → Tên NV, địa chỉ NV

➤ **Chuẩn hóa 1NF** - Loại bỏ nhóm lặp và loại bỏ các thuộc tính tính toán

Không có thuộc tính lặp

Có 2 thuộc tính cần tính toán: tổng tiền hóa đơn và thành tiền → đây là dữ liệu dư thừa

Chuẩn hóa lại như sau:

Mã hóa đơn, Ngày bán,

Mã khách hàng, Tên Khách hàng, Số ĐT Khách hàng,

Mã Nhân viên, Số CMT của NV, Tên NV, Địa chỉ NV, Số ĐT Nhân viên

Mã thuốc, Tên thuốc, Công dụng, Ngày SX, Hạn SD, Số lượng,

bỏ 2 field tính toán đi

➤ **Chuẩn hóa 2NF** - Loại bỏ các phụ thuộc hàm không hoàn toàn vào khóa chính

- **Chuẩn hóa 3NF** - Loại bỏ các phụ thuộc hàm bậc cầu vào khóa chính