

# Adatbázis rendszerek 1.

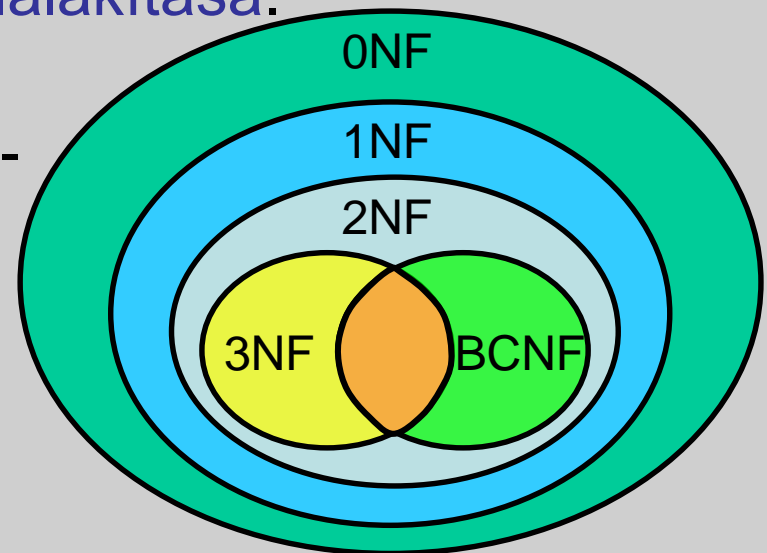
## 3. Ea: A „rendes állapot”

Normalizálás



# Normalizálás

- A normalizálás az adatbázis belső szerkezetének ellenőrzése, lépésenkénti átalakítása oly módon, hogy az adatbázis minden egyes lépésben egy-egy újabb kritériumnak, egymásra épülő **normálformának** feleljen meg.
- A műveletsorozat célja a **redundancia- és anomáliamentes adatbázisok kialakítása**.
- A **3NF**-re vagy **BCNF**-re alakított adatbázis redundancia- és anomáliamentes, a normalizálás célja ezen alakok elérése.



# Redundancia

**Redundancia:** Adatelemek többszörös, ismételt letárolása

<u>Rendszám</u>	Típus	Név	Cím	Telefon
GBC-765	Opel	Kis Géza	Eger Kő u. 4	20-9555666
JED-123	Nissan	Jó Éva	Miskolc Fa u.1.	30-8765432
AKT-392	Audi	Kék Béla	Ózd Retek u. 99	70-3355440
LOG-416	Volvo	Kis Géza	Eger Kő u.4	20-9555666

Következmény: **anomáliák**

- Beszúrási anomália
- Módosítási anomália
- Törlési anomália
- Inkonzisztencia

# Anomáliák

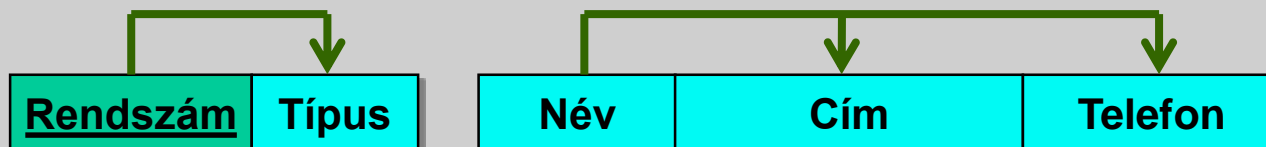
---

- **Beszúrási anomália:** egy rekord felvitelekor, felesleges, már letárolt információkat újra fel kell vinni (pl. Tulaj címe, telefonja); ha valahol véletlenül elgépeljük, új adatot hozunk létre.
- **Módosítási anomália:** egy adatmódosításhoz több helyen is módosítani kell az adatbázisban (pl. ha megváltozik a tulaj telefonszáma, mindenütt ki kell javítani).
- **Törlési anomália:** egy információelem megszűnésekor más, hozzá nem tartozó információk is elvesznek (ha pl. máshol nincs letárolva a tulaj telefonszáma, akkor az adott autó törlése után a tulajt többet nem érjük el).

# Az anomáliák oka

- Az anomáliák abból származnak, hogy nem az igazán összetartozó adatokat vesszük be egy relációba.
- Hogy mely mezők tartoznak igazán egy relációba, azt a mezők közötti összetartozási viszony, a mezők közötti **függőségek** határozzák meg.
- Legfontosabb függőségi típus a **funkcionális függőség**, **FD – Functional Dependency**, amikor egy tulajdonság értékét egy másik tulajdonság értéke egyértelműen meghatározza. (Egyirányú viszony!)

<u>Rendszám</u>	Típus	Név	Cím	Telefon
GBC-765	Opel	Kis Géza	Eger Kő u. 4	20-9555666



# Funkcionális függőségek felderítése

---

- A funkcionális függőségek felderítése nagy segítséget nyújt a helyes séma kialakításában.
- A funkcionális függőségekkel magyarázható a redundancia is. Ha ugyanis egy ismétlődő értékű mezőből FD indul ki, akkor az FD definíciója alapján a függőmezőnek is ismétlődnie kell.
- Tehát **a redundancia oka a nem megfelelő, felesleges FD a relációsémán belül.**
- Az FD-k felderítését, és az adatbázis FD-mentessé alakítását (dekompozícióját) nevezzük **normalizálásnak**.
- A normalizálás néhány rögzített irányelven alapszik, amelyek iránymutatást adnak a tervezéshez, helyes mederbe terelve a modellezés menetét.

# Dekompozíció

---

- A **dekompozíció** során az induló séma felbontásával emeljük ki a nem kívánt FD-ket külön relációkba. Azaz a felesleges FD-t tartalmazó sémát dekompozícióval hozzuk normalizált alakra.
- A dekompozíciós módszerben felállított tervezési irányelveket több, egymásra épülő követelmény alakjában adják meg. Az egyes követelményeket szokás **normálformák**nak is nevezni.
- A normalizálás erről az oldalról nézve nem más, mint a megadott normálformák teljesülésének ellenőrzése, illetve az adatsémák átalakítása olyan alakra, hogy azok kielégítsék a megadott normálformákat.

# Normalizálás – Példa: Bolti számla

## ■ Kiinduló adatok, ún. **nulladik normálforma (0NF)**

SzSzáma	Dátum	Vnév	Vcím	Dkód	Dnév	Tkód	Tnév	Menny	EgysÁr	Összeg	Összár
112	2010. 12.23	Kis Éva	Mc. Jég u.1	14	Hó Ida	A673	Izzósor	2	2500	5000	14500
						C193	Gyertya	4	500	2000	
						F003	Fenyő	2,5	3000	7500	

Számla

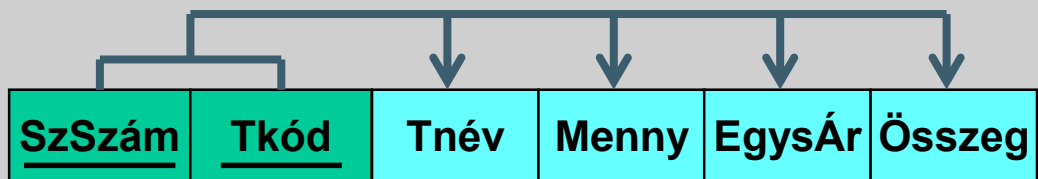
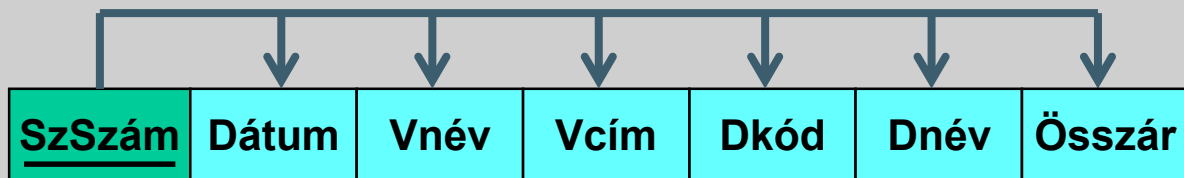
Vevő adatai

Pénztáros

Termékek adatai

Szla.

- Egy vevő általában több terméket vásárol, emiatt a Tkód...Összeg mezők többértékűek.
- Keressük meg a mezők közötti összefüggéseket:





# Normalizálás 2.

**1NF – Első normálforma:** *Egy reláció 1NF-ben van, ha minden attribútum egyértékű, és a relációban van kulcs.*  
(A kulcs miatt a sorok különböznek egymástól.)

- Ha az FD-knek megfelelően az eredeti relációt két új relációra bontjuk, akkor egy olyan adatbázist kapunk, amely 1NF-ben van:

R1	<u>SzSzá</u> m	Dátum	Vnév	Vcím	Dkód	Dnév	Összár
R2	<u>SzSzá</u> m	<u>Tkód</u>	Tnév	Menny	EgysÁr	Összeg	

SzSzá	Dátum	Vnév	Vcím	Dkód	Dnév	Összár
112	2010.12.23	Kis Éva	Mc. Jég u.1	14	Hó Ida	14500

R1

R2

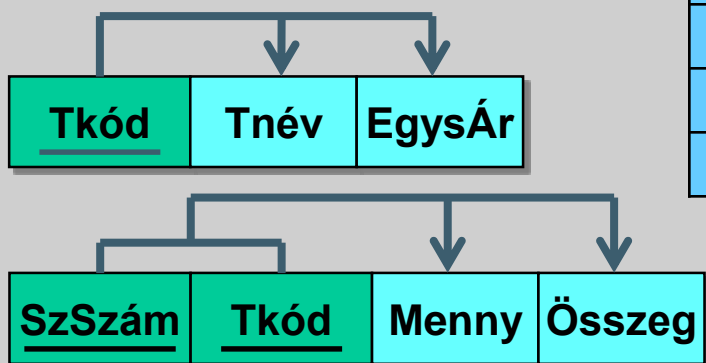
SzSzá	Tkód	Tnév	Menny	EgysÁr	Összeg
112	A673	Izzó	2	2500	5000
112	C193	Gyertya	4	500	2000
112	F003	Fenyő	2,5	3000	7500

# Normalizálás 3.

**2NF – Második normálforma:** *Egy reláció 2NF-ben van, ha 1NF teljesül, és minden attribútum a teljes kulcstól függ.*

- Ha a kulcs egyetlen mező, a 2NF automatikusan teljesül!
- Vizsgáljuk az R2 relációt, hiszen ott összetett a kulcs.

Keressünk FD-eket!



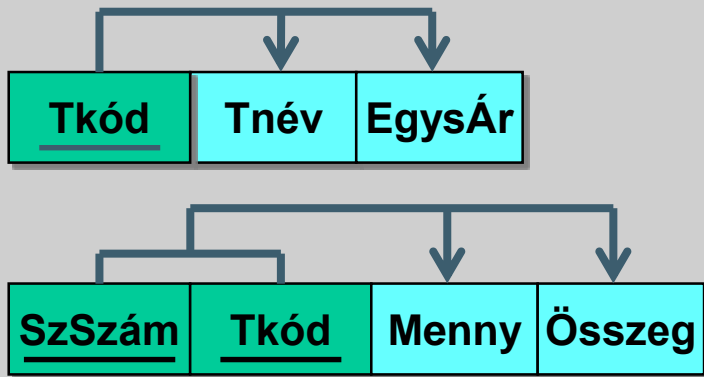
SzSzám	Tkód	Tnév	Menny	EgysÁr	Összeg
112	A673	Izzósor	2	2500	5000
112	C193	Gyertya	4	500	2000
112	F003	Fenyő	2,5	3000	7500

- Most már minden a kiválasztott kulcsoktól függ, így az ennek megfelelően felbontott relációk 2NF-ben lévő adatbázist adnak.

# Normalizálás 4. R1

SzSzám	Dátum	Vnév	Vcím	Dkód	Dnév	Összár
112	2010.12.23	Kis Éva	Mc. Jég u.1	14	Hó Ida	14500

## ■ FD-k:



**R2**

SzSzám	Tkód	Tnév	Menny	EgysÁr	Összeg
112	A673	Izzósor	2	2500	5000
112	C193	Gyertya	4	500	2000
112	F003	Fenyő	2,5	3000	7500

## Termék

Tkód	Tnév	EgysÁr
A673	Izzósor	2500
C193	Gyertya	500
F003	Fenyő	3000

## 2NF relációk:

R1

<u>SzSzám</u>	Dátum	Vnév	Vcím	Dkód	Dnév	Összár
---------------	-------	------	------	------	------	--------

Termék

<u>Tkód</u>	Tnév	EgysÁr
-------------	------	--------

Tétel

<u>SzSzám</u>	<u>Tkód</u>	Menny	Összeg
---------------	-------------	-------	--------

## Tétel

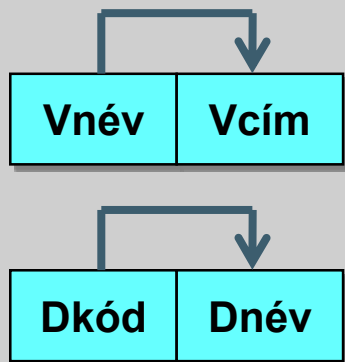
SzSzám	Tkód	Menny	Összeg
112	A673	2	5000
112	C193	4	2000
112	F003	2,5	7500

# Normalizálás 5.

**3NF – Harmadik normálforma:** *Egy reláció 3NF-ben van, ha 2NF teljesül, és a nem kulcs attribútumok nem függenek tranzitíven a kulcstól.* Ez azt jelenti, hogy a nem kulcs attribútumok között nem találunk FD-t!

- Keressünk a relációkban a mezők között összefüggéseket!

R1	SzSzám	Dátum	Vnév	Vcím	Dkód	Dnév	Összár
	112	2010. 12.23	Kis Éva	Mc. Jég u.1	14	Hó Ida	14500



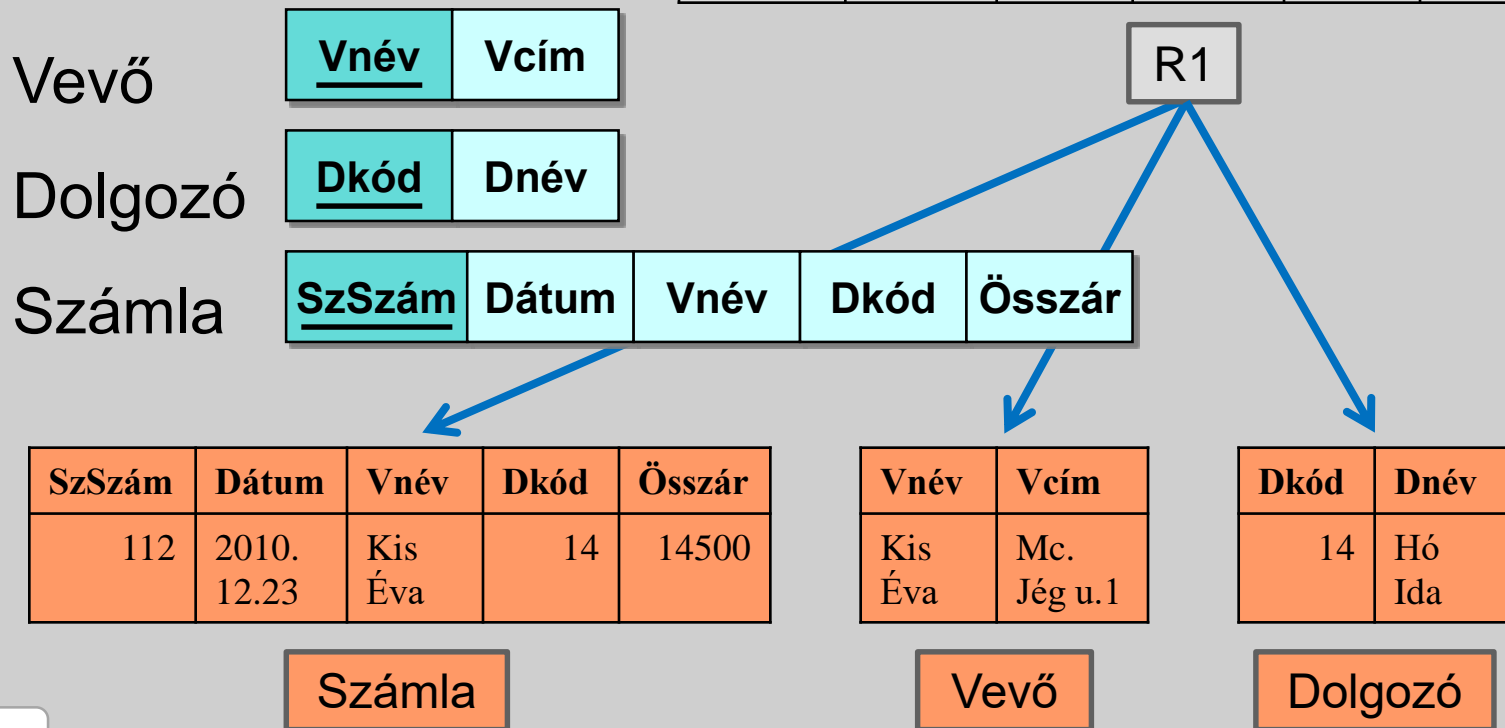
A vevő neve egyértelműen meghatározza a vevő címét!

A dolgozó kódja egyértelműen meghatározza a dolgozó nevét!

# Normalizálás 6.

- A megtalált FD-knek megfelelően újból átalakítjuk a relációkat, és így egy 3NF-ben lévő adatbázishoz jutunk.

## 3NF relációk:



# Normalizálás 7.

## ■ A 3NF-ben lévő teljes reláció:

Számla	<u>SzSzám</u>	Dátum	Vnév	Dkód	Összár
Vevő	<u>Vnév</u>	Vcím			
Dolgozó	<u>Dkód</u>	Dnév			
Termék	<u>Tkód</u>	Tnév	Egys Ár		
Tétel	SzSzám	Tkód	Menny	Összeg	

SzSzám	Dátum	Vnév	Dkód	Összár
112	2010.12.23	Kis Éva	14	14500

Vnév	Vcím
Kis Éva	Mc. Jég u.1

Dkód	Dnév
14	Hó Ida

Számla

Vevő

Dolgozó

Termék

Tkód	Tnév	Egys Ár
A673	Izzósor	2500
C193	Gyertya	500
F003	Fenyő	3000

SzSzám	Tkód	Menny	Összeg
112	A673	2	5000
112	C193	4	2000
112	F003	2,5	7500

Tétel

# Normalizálás 8.

---

**BCNF – Boyce-Codd normálforma:** *Egy reláció BCNF-ben van, ha minden függőség csak jelölt kulcsból indul ki.*

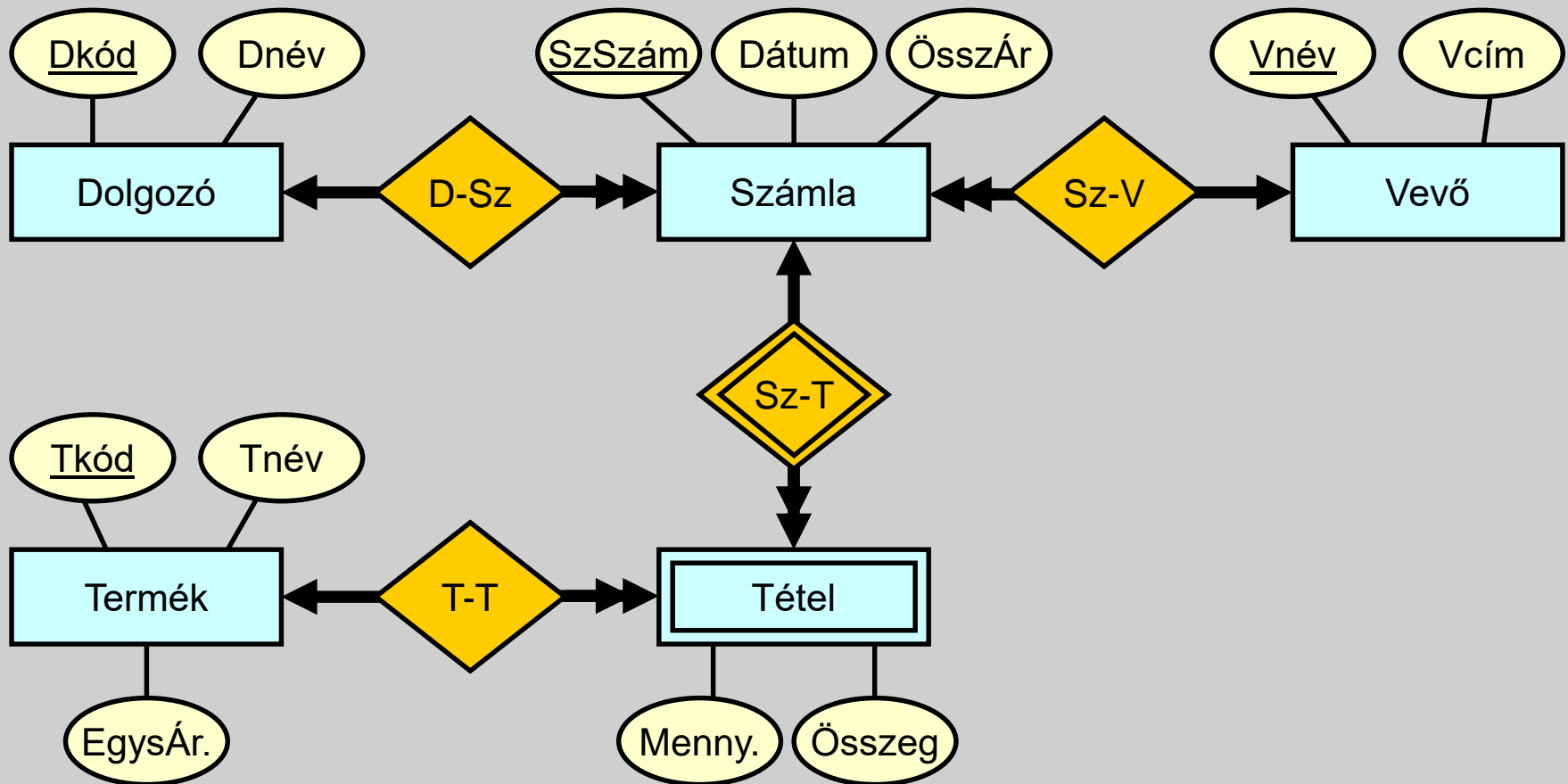
Jellemzői:

- BCNF átfogja 2NF-et.
- 3NF-et célszerű ezzel helyettesíteni.
- A 3NF nem foglalja magába a BCNF-et.
- A BCNF nem foglalja magába a 3NF-et.

A példa reláció teljesíti ezt a kritériumot, így BCNF-ben van!

# Normalizálás 9.

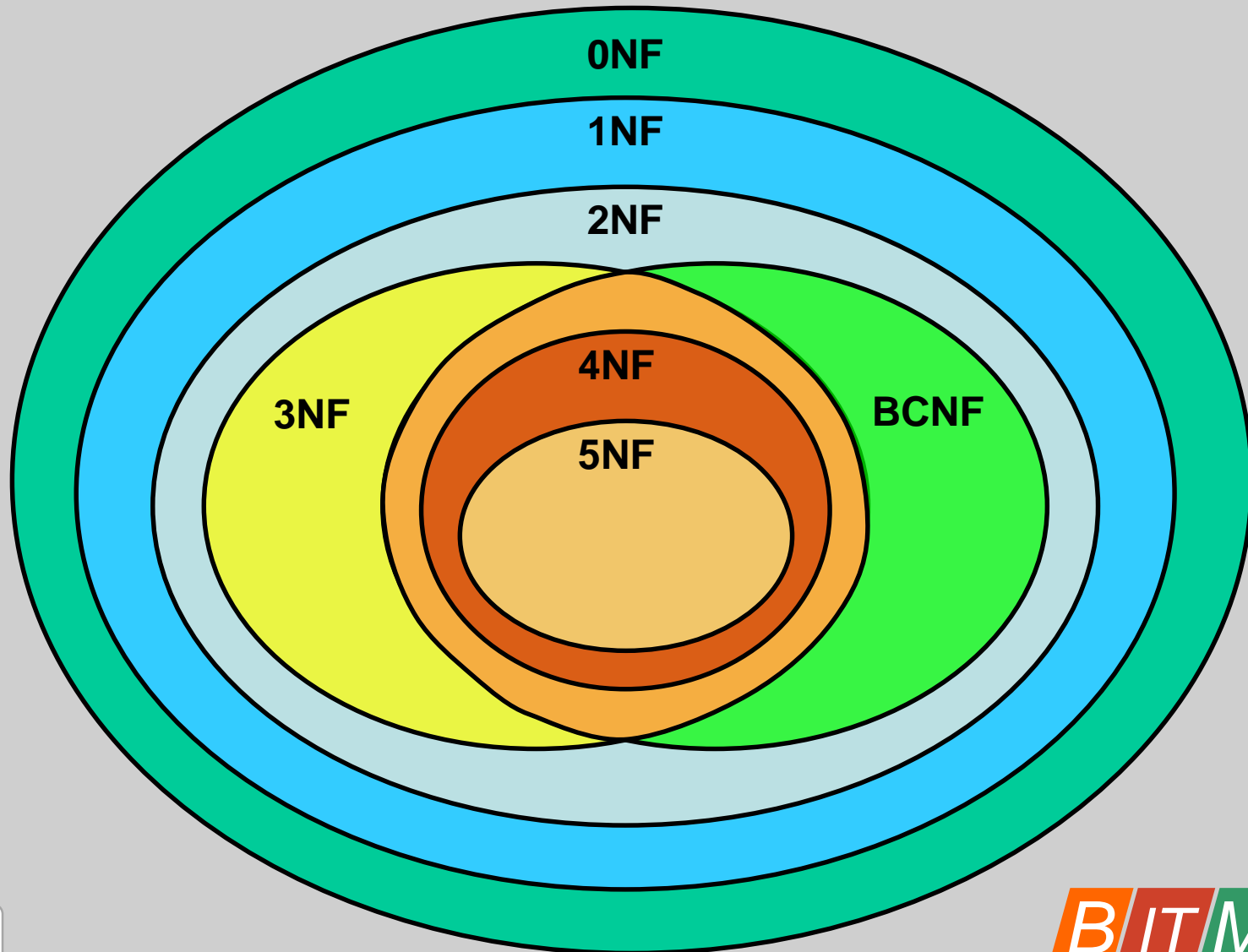
- A normalizálás eredményeként kapott ER-modell:





# Normalizálás 10.

---



# Kártyavár



Ez most a  
legnép-  
szerűbb  
modellünk.



# Armstrong-axiómák

---

- Az Armstrong-axiómák segítségével a már meglévő függőségekből új függőségek származtathatók. Az Armstrong-axiómák a következők:
  - Transzitivitás
  - Reflexivitás
  - Augmentivitas – Bővíthetőség
- Az Armstrong-axiómákból levezethető szabályok
  - Additivitás – Egyesítési szabály
  - Dekompozíció – Szétvághatósági szabály
  - Pszeudotranzitivitás

# Armstrong-axiómák

---

## ■ Tranzitivitás:

Az egyik legegyszerűbb következtetési szabály azt mondja ki, hogy ha az  $R(A,B,C)$  attribútum-halmazokra teljesülnek az  $A \rightarrow B$  és  $B \rightarrow C$  függőségek, akkor teljesül az  $A \rightarrow C$  függőség is. Ezt a tulajdonságot tranzitivitásnak nevezzük.

## ■ Reflexivitás (triviális függőség):

Nyilvánvalóan igaz az a szabály is, hogy a  $B$  halmaz minden  $B'$  részhalmazára is teljesül a  $B \rightarrow B'$  függőség. Ezt nevezzük reflexivitásnak vagy triviális függőségnek.

# Armstrong-axiómák

---

- A **bővíthetőség** azt mondja ki, hogy ha egy funkcionális függés mindkét oldalát ugyanazzal az attribútum-halmazzal bővítjük, akkor a függés továbbra is megmarad a két halmaz között.
- Formálisan, ha  $A \rightarrow B$  és  $C$  tetszőleges attribútum-halmaz, akkor  $A \cup C \rightarrow B \cup C$ .

# Armstrong-axiómákból következő szabályok

---

A következő két szabály szorosan kapcsolódik egymáshoz, mivel lényegében egymás ellentétjei.

- **Egyesítési szabály** azt jelenti, hogy ha teljesül egy  $A \rightarrow B$  és egy  $A \rightarrow C$  függőség, akkor teljesül az  $A \rightarrow B \cup C$  függőség is. Szavakkal kifejezve: azonos bal oldalú függőségek jobb oldalán szereplő attribútum-halmazait egyesíthetjük.
- **Szétvághatósági szabály.** A funkcionális függőség jobb oldalán szereplő attribútum-halmazra teljesül az, hogy annak minden részhalmaza is függ a baloldaltól. Vagyis ha  $A \rightarrow B$  és  $C \subseteq B$ , akkor  $A \rightarrow C$  is teljesül, és ezt nevezzük szétvághatósági szabálynak.

# Armstrong-axiómákból következő szabályok

---

## ■ Pszeudotranzitivitás:

ha  $A \rightarrow B$  és  $C \cup B \rightarrow D$ , akkor  $A \cup C \rightarrow D$

Jelentése: ha egy részhalmaz-t kicserélek egy olyan halmazra, amelytől funkcionálisan függ, akkor a függés megmarad.

# Gyakorló feladat (1)

---

- Normalizáljuk az alábbi sémát 3NF-ig, alkalmazzuk a tanult szabályokat és axiómákat:

$R(A,B,C,D,E,F)$  ahol:

- $A \rightarrow C$
- $C \rightarrow E$
- $(A,B) \rightarrow F$
- $B \rightarrow D$



# Gyakorló feladat (1)

R(A,B,C,D,E,F)	
▪	$A \rightarrow C$
▪	$C \rightarrow E$
▪	$(A,B) \rightarrow F$
▪	$B \rightarrow D$

- Armstrong 2. axiómája alapján (az egész meghatározza a részét):
  - $(A,B) \rightarrow A$  és  $(A,B) \rightarrow B$
- Armstrong 1. axiómája alapján (a függőségek tranzitívek):
  - $(A,B) \rightarrow A$  és  $A \rightarrow C \Rightarrow (A,B) \rightarrow C$
  - $(A,B) \rightarrow C$  és  $C \rightarrow E \Rightarrow (A,B) \rightarrow E$
  - $(A,B) \rightarrow B$  és  $B \rightarrow D \Rightarrow (A,B) \rightarrow D$
- A fentiekből következik, hogy  $(A,B)$  meghatározza az összes többi attribútumot, így 1NF-ben ez lesz a kulcs:
  - **1NF:**  $R(\underline{A}, \underline{B}, C, D, E, F)$

# Gyakorló feladat (1)

R(A,B,C,D,E,F)					
▪	A	→	C		
▪	C	→	E		
▪	(A,B)	→	F		
▪	B	→	D		

- 2N-hez úgy daraboljuk a relációkat, hogy a relációban maradó teljes kulcstól függjenek:
  - $A \rightarrow C, C \rightarrow E \Rightarrow R1(\underline{A}, C, E)$
  - $(A, B) \rightarrow F \Rightarrow R2(\underline{A}, \underline{B}, F)$
  - $B \rightarrow D \Rightarrow R3(\underline{B}, D)$
- Vagyis 2NF:
  - $R1(\underline{A}, C, E) \quad R2(\underline{A}, \underline{B}, F) \quad R3(\underline{B}, D)$

# Gyakorló feladat (1)

R(A,B,C,D,E,F)					
▪ $A \rightarrow C$					
▪ $C \rightarrow E$					
▪ $(A,B) \rightarrow F$					
▪ $B \rightarrow D$					

- 3NF-hez nem kulcsból kiinduló FD-ket keresünk (most:  $C \rightarrow E$ ), és ezek alapján daraboljuk az adott relációkat, vagyis:
- $R1(\underline{A}, C, E) \Rightarrow R1(\underline{A}, C) \quad R2(\underline{C}, E)$
- Vagyis 3NF:  $R1(\underline{A}, C) \quad R2(\underline{C}, E) \quad R3(\underline{A}, \underline{B}, F) \quad R4(\underline{B}, D)$

# Gyakorló feladat (2)

---

- Normalizálja az alábbi sémát 3NF-ig:

$R(X,Y,Z,Q,W)$  ahol:

- $Y \rightarrow W$
- $X \rightarrow (Q,Z)$
- $Z \rightarrow Y$

- A szétvághatósági szabály alapján:

- $X \rightarrow (Q,Z) \leftrightarrow X \rightarrow Q \text{ és } X \rightarrow Z$

- Armstrong 1. axiómája alapján:

- $X \rightarrow Z \text{ és } Z \rightarrow Y \leftrightarrow X \rightarrow Y$
- $X \rightarrow Y \text{ és } Y \rightarrow W \leftrightarrow X \rightarrow W$

- A mezők atomiságát feltételezve:

- 1NF:  $R(\underline{X},Y,Z,Q,W)$

- 2NF: = 1NF

- 3NF:  $R1(\underline{X},Q,Z) \quad R2(\underline{Z},Y) \quad R3(\underline{Y},W)$

# Gyakorló feladat (3)

---

- Normalizálja az alábbi sémát BCNF-ig:

$R(A,B,C,D,E)$  ahol:

- $C \rightarrow E$
- $A \rightarrow D$
- $E \rightarrow B$
- $(A,E) \rightarrow A$

- Armstrong 2. axiómája alapján:

–  $(A,E) \rightarrow A$  és  $(A,E) \rightarrow E$

- Armstrong 1. axiómája alapján:

–  $(A,E) \rightarrow A$  és  $A \rightarrow D \leftrightarrow (A,E) \rightarrow D$

–  $(A,E) \rightarrow E$  és  $E \rightarrow B \leftrightarrow (A,E) \rightarrow B$

29/31 – De  $C \rightarrow E$ , ezért  $(A,C)$  a kulcs.

# Gyakorló feladat (3)

R(A,B,C,D,E)	
▪	$C \rightarrow E$
▪	$A \rightarrow D$
▪	$E \rightarrow B$
▪	$(A,E) \rightarrow A$

- A mezők atomiságát feltételezve:
  - 1NF:  $R(\underline{A}, \underline{C}, B, D, E)$
  - 2NF:  $R1(\underline{A}, \underline{C}) \quad R2(\underline{A}, D) \quad R3(\underline{C}, E, B)$
  - 3NF:  $R1(\underline{A}, \underline{C}) \quad R2(\underline{A}, D) \quad R3(\underline{C}, E) \quad R4(\underline{E}, B)$
  - BCNF: = 3NF

# VÉGE

