# Normalisering

### Uppgift 1: Skivbolag

Vi har följande tabell där titelnr är primärnyckel:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| titelnr | titelnamn | artistnamn | årtal | skivbolag | huvudkontor |
| 1 | Abbey Road | Beatles | 1969 | EMI | London |
| 2 | Rubber Soul | Beatles | 1966 | EMI | London |
| 3 | Nevermind | Nirvana | 1991 | Sub Pop | Seattle |
| 4 | Nevermind | Eminem | 2012 | Yoda Records | Seattle |
| 5 | Joyride | Roxette | 1991 | EMI | London |
| 6 | Tourism | Roxette | 1991 | EMI | London |
| 7 | Nevermind | Soda | 2012 | Yoda Records | Seattle |

Enbart utifrån de rader som är inlagda i tabellen (mer data kan inte tillkomma) ovan ska du:

1. Identifiera och redogöra för de fullständigt funktionella beroenden som existerar. De beroenden du identifierat ska motiveras. Svara enligt:  
     
   ffd1: attribut1 --> attribut2, attribut3  
   ffd2: attribut4, attribut5 --> attribut6, attribut7, attribut8  
   etc.  
     
   Svar:

ffd1: titelnr --> titelnamn, artistnamn, årtal, skivbolag, huvudkontor

ffd2: artistnamn --> skivbolag, huvudkontor

ffd3: skivbolag --> huvudkontor

ffd4: titelnamn --> huvudkontor

ffd5: titelnamn, artistnamn --> skivbolag, huvudkontor

1. Ange vilken den högsta normalform tabellen uppfyller och varför.   
     
   Svar: Den lägsta formen som den erhåller är 1NF då endast ett värde finns i varje kolumn och tuple, det finns endast ett värde i varje cell. 2NF uppfylls även det då? Titelnr (primärnyckeln) endast kan determinera samtliga attribut, trots att titelnamn och artistnamn som kanditatnyckel determinerar skivbolag och huvudkontor så kommer fortfarande titelnr determinera värdet. ~~Inget fullständigt funktionellt beroende finns mellan primärnyckeln och andra attribut vilket gör att kraven för 2NF uppfylls. 3NF kommer däremot inte att uppfyllas då huvudkontor är beroende av skivbolag, samt att titelnamn är beroende av artistnamn. Den högsta normal formen som uppfylls är 2NF.~~ Endast 1NF uppfylls då sammansatta kandidatnycklar som i ffd5 bryter mot dess krav.

1. Om tabellen inte uppfyller BCNF ska du, steg för steg, visa hur tabellen normaliseras till BCNF.  
     
   Svar: Enligt BCNF ska ALLA determinanter i en tabell vara en kandidatnyckel.

BCNF uppfyller samma krav som 3NF, däremot får det inte finnas några beroenden av attribut utan primärnyckeln. Då vår tabell uppfyller 2NF men inte 3NF så måste vi först göra om den till 3NF.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| titelnr | titelnamn | artistnamn | årtal | skivbolag | huvudkontor |
| 1 | Abbey Road | Beatles | 1969 | EMI | London |
| 2 | Rubber Soul | Beatles | 1966 | EMI | London |
| 3 | Nevermind | Nirvana | 1991 | Sub Pop | Seattle |
| 4 | Nevermind | Eminem | 2012 | Yoda Records | Seattle |
| 5 | Joyride | Roxette | 1991 | EMI | London |
| 6 | Tourism | Roxette | 1991 | EMI | London |
| 7 | Nevermind | Soda | 2012 | Yoda Records | Seattle |

Artistnamn och skivbolag behöver delas upp i en ny tabell.

|  |  |
| --- | --- |
| artistnamn | skivbolag |
| Beatles | EMI |
| Roxette | EMI |
| Eminem | Yoda Records |
| Soda | Yoda Records |
| Nirvana | Sub Pop |

### artist(artistnamn, skivbolag\*)

Lika så behöver skivbolag och huvudkontor delas upp i en ny tabell, för att undvika redundans och beroenden.

|  |  |
| --- | --- |
| skivbolag | huvudkontor |
| EMI | London |
| Sub Pop | Seattle |
| Yoda Records | Seattle |

### bolag(skivbolag, huvudkontor)

Den nya tabellen blir som följande, med BCNF uppfyllt.

|  |  |  |  |
| --- | --- | --- | --- |
| titelnr | titelnamn | artistnamn | årtal |
| 1 | Abbey Road | Beatles | 1969 |
| 2 | Rubber Soul | Beatles | 1966 |
| 3 | Nevermind | Nirvana | 1991 |
| 4 | Nevermind | Eminem | 2012 |
| 5 | Joyride | Roxette | 1991 |
| 6 | Tourism | Roxette | 1991 |
| 7 | Nevermind | Soda | 2012 |

### titlar(titelnr, titelnamn, artistnamn\*, årtal)

### Uppgift 2: Betyg

Nedan har vi en tabell, Betyg, som innehåller information om olika kurser, vilka studenter som läst vilka kurser, samt vilket betyg studenten fick på kursen. Primärnyckeln i tabellen är sammansatt av studentnr och kurskod.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| studentnr | namn | kursnamn | kurskod | betyg | benämning |
| 1 | Kalle Svensson | Java | DT006G | C | Godkänd |
| 2 | Ida Larsson | Ekonomi | FÖ094G | B | Väl godkänd |
| 3 | Anna Andersson | Databaser | DT022G | A | Väl godkänd |
| 2 | Ida Larsson | Java | DT006G | F | Underkänd |
| 1 | Kalle Svensson | Ekonomi | FÖ094G | C | Godkänd |
| 4 | Kalle Svensson | Databaser | DT022G | B | Väl godkänd |
| 3 | Anna Andersson | Linux | DT029G | A | Väl godkänd |
| 1 | Kalle Svensson | Databaser | DT076G | Fx | Underkänd |

Utifrån de rader som är inlagda i tabellen ovan ska du (tänk på att fler rader kan tillkomma i tabellen):

1. Identifiera och redogöra för de fullständigt funktionella beroenden som existerar. De beroenden du identifierat ska motiveras. Svara enligt:  
     
   ffd1: attribut1 --> attribut2, attribut3  
   ffd2: attribut4, attribut5 --> attribut6, attribut7, attribut8  
   etc.  
     
   Svar:

ffd1: studentnr --> namn

ffd2: kurskod --> kursnamn

ffd3: studentnr, kursnamn -> kurskod, betyg, benämning

ffd4: namn, kursnamn --> betyg

ffd5: namn, kurskod --> betyg

ffd6: betyg --> benämning

1. Ange vilken den högsta normalform tabellen uppfyller och varför.  
     
   Svar: 1NF uppfylls få endast ett värde per cell finns, tuple är även unika. Den kommer däremot inte att uppfylla 2NF då det finns fullständiga funktionella beroenden mellan primärnyckeln och attribut som till exempel syns på ”studentnr 1” tillsammans med ”kursnamn”. Endast ”studentnr 1” kommer inte determinera varken kursnamn, kurskod, betyg eller benämning och behöver därför använda sig av ”kursnamn”. Den högsta normalform som tabeller uppfyller är därför 1NF.
2. Om tabellen inte uppfyller BCNF ska du, steg för steg, visa hur tabellen normaliseras till BCNF.  
     
   Svar:

|  |  |  |
| --- | --- | --- |
| studentnr | kurskod | betyg |
| 1 | DT006G | C |
| 2 | FÖ094G | B |
| 3 | DT022G | A |
| 2 | DT006G | F |
| 1 | FÖ094G | C |
| 4 | DT022G | B |
| 3 | DT029G | A |
| 1 | DT076G | Fx |

### student(studentnr\*, kurskod\*, betyg\*)

|  |  |
| --- | --- |
| studentnr | namn |
| 1 | Kalle Svensson |
| 2 | Ida Larsson |
| 3 | Anna Andersson |
| 4 | Kalle Svensson |

### namn(studentnr, namn)

|  |  |
| --- | --- |
| kurskod | kursnamn |
| DT006G | Java |
| FÖ094G | Ekonomi |
| DT022G | Databaser |
| DT029G | Linux |
| DT076G | Databaser |

### kurser(kurskod, kursnamn)

|  |  |
| --- | --- |
| betyg | benämning |
| A | Väl godkänd |
| B | Väl godkänd |
| C | Godkänd |
| F | Underkänd |
| Fx | Underkänd |

### bedömning(betyg, benämning)

Bättre är såhär får jag inte tabellerna, helst skulle jag vilja ha studentnr unikt men vet inte vart jag ska ta mig härifrån, om det ens går att bearbeta det vidare. Studentnr är ingen primärnyckel då olika värden i tuplen finns för vardera, 1 ger tex java, ekonomi och databaser.