Relational Database Design

Module 8: Higher Normal Forms

Hugo Kornelis
hugo@perFact.info

pluralsight
hardcore developer training

Outline

- **Basic normal forms**
  - First Normal Form (1NF)
  - Second Normal Form (2NF)
  - Third Normal Form (3NF)
- **Higher normal forms**
  - EKNF, BCNF, 4NF, 5NF, DK/NF, ONF, 6NF
- **Denormalization**
- **Standard patterns for**
  - Referencing a range
  - Storing historic data

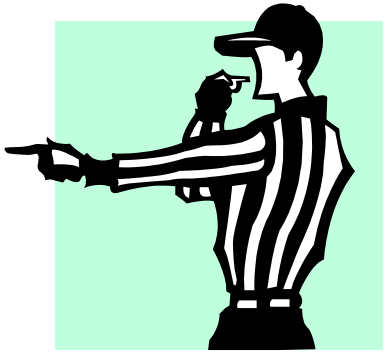
Boyce-Codd Normal Form

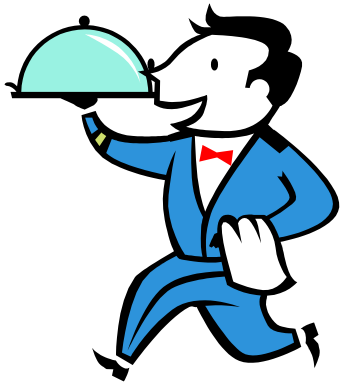
- **Third Normal Form (3NF)**
  - Every *non-key* attribute depends on each candidate key
    - (nothing more, nothing less)
  - But what about *key* attributes?
    - Every single attribute key **always** depends on each candidate key
    - But a *part of* a composite key **may** depend on other attributes as well!

- **Boyce-Codd Normal Form (BCNF)**
  - Every attribute (key **or** non-key!) depends on each candidate key
    - (nothing more, nothing less)

Boyce-Codd Normal Form

- Example



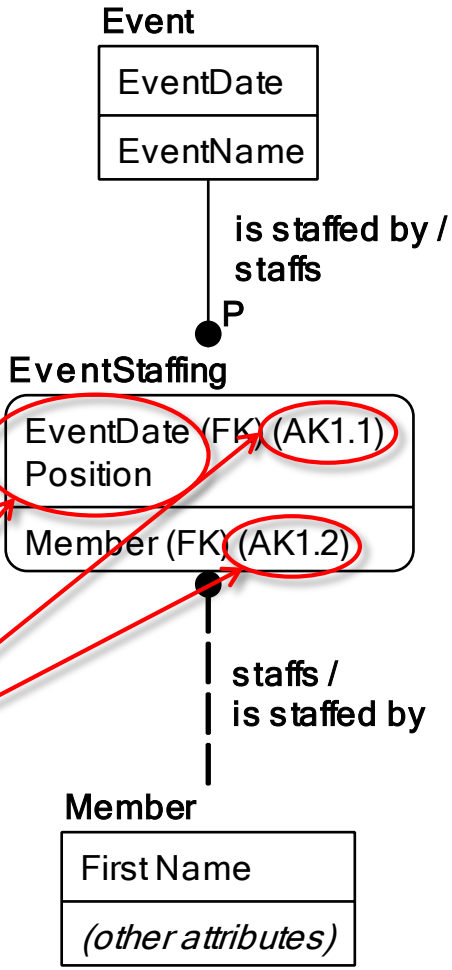


Boyce-Codd Normal Form

Event date	Position	Member
2013-06-04	Coach	Dave
2013-06-04	Referee	Mary
2013-06-04	Catering	John
2013-08-12	Referee	Hugo
2013-08-12	Catering	John

Functional dependencies:

- {Event date, Position} → Member
- {Event date, Member} → Position
- Member → Position

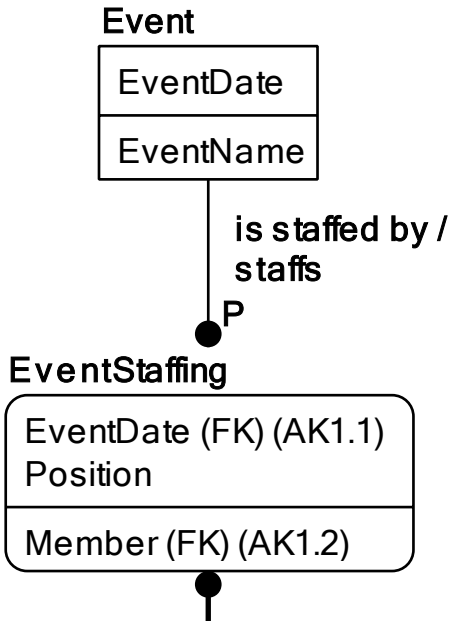


Boyce-Codd Normal Form

Event date	Position	Member
2013-06-04	Coach	Dave
2013-06-04	Referee	Mary
2013-06-04	Catering	John
2013-08-12	Referee	Hugo
2013-08-12	Catering	John
2013-08-12	Coach	Mary

Functional dependencies:

- {Event date, Position} → Member
- Member → Position



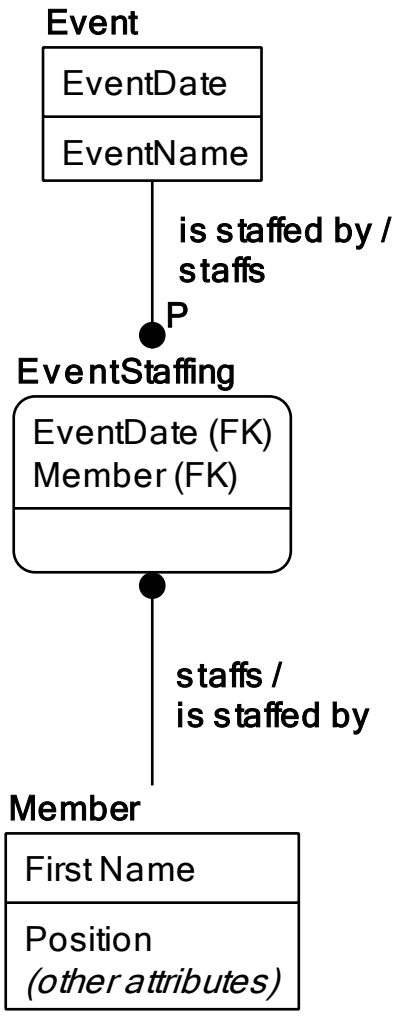
✓ 3NF

✗ BCNF

Boyce-Codd Normal Form

Event date	Member
2013-06-04	Hugo
2013-06-04	Mary
2013-06-04	John
2013-08-12	Hugo
2013-08-12	John

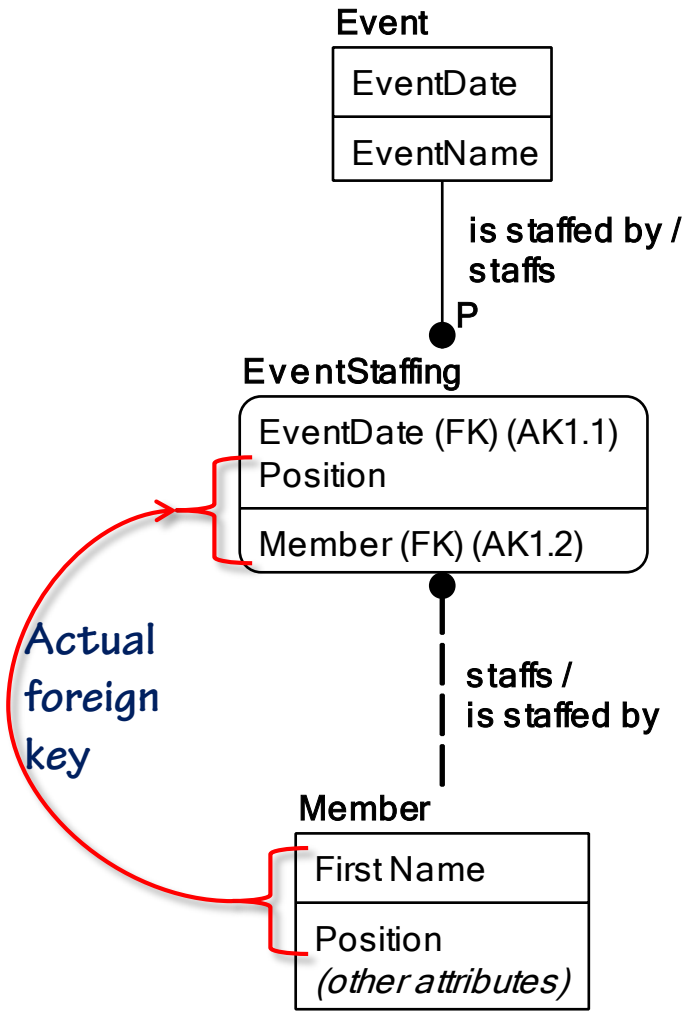
Member	Position
Dave	Coach
Mary	Referee
John	Catering
Hugo	Referee



Boyce-Codd Normal Form

Event date	Position	Member
2013-06-04	Coach	Dave
2013-06-04	Referee	Mary
2013-06-04	Catering	John
2013-08-12	Referee	Hugo
2013-08-12	Catering	John

Member	Position
Dave	Coach
Mary	Referee
John	Catering
Hugo	Referee



Boyce-Codd Normal Form

- **Boyce-Codd Normal Form (BCNF)**
  - Every attribute (key **or** non-key!) depends on each candidate key
    - (nothing more, nothing less)
  - Can't always be achieved
    - (But usually can!)

Elementary Key Normal Form

- **Boyce-Codd Normal Form (BCNF)**
  - Every attribute (key **or** non-key!) depends on each candidate key
    - (nothing more, nothing less)
  - Can't always be achieved
    - (But usually can!)

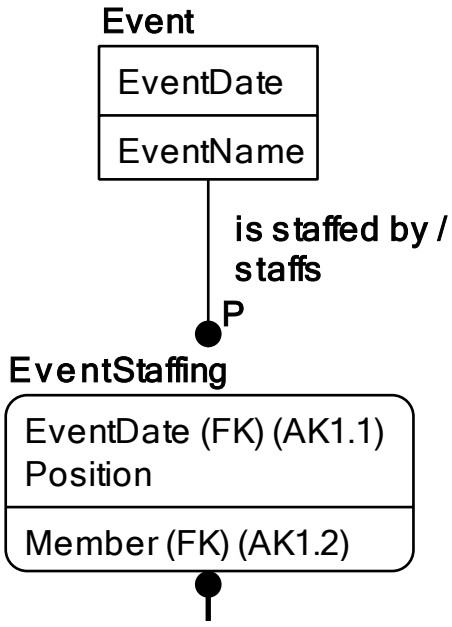
- **Elementary Key Normal Form (EKNF)**
  - Between 3NF and BCNF
    - Every non-key attribute depends on each candidate key
    - Every ***elementary*** key attribute depends on each candidate key
  - Does not solve all problems

Elementary Key Normal Form

Event date	Position	Member
2013-06-04	Coach	Dave
2013-06-04	Referee	Mary
2013-06-04	Catering	John
2013-08-12	Referee	Hugo
2013-08-12	Catering	John
2013-08-12	Coach	Mary

Functional dependencies:

- {Event date, Position} → Member
- Member → Position



✓ 3NF

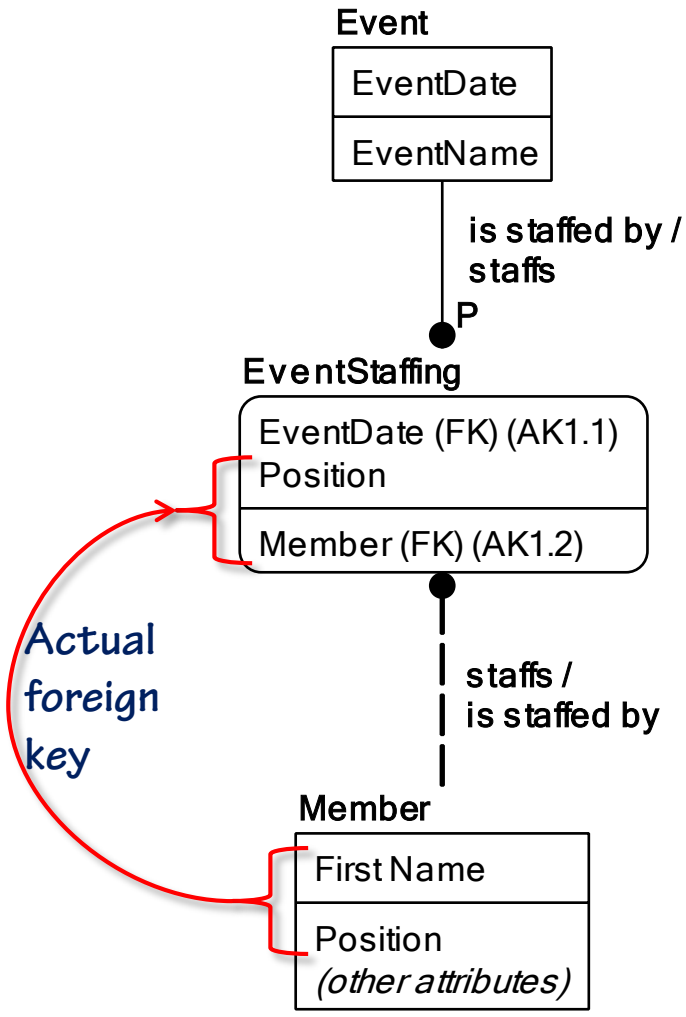
✗ BCNF

✓ EKNF

Boyce-Codd Normal Form

Event date	Position	Member
2013-06-04	Coach	Dave
2013-06-04	Referee	Mary
2013-06-04	Catering	John
2013-08-12	Referee	Hugo
2013-08-12	Catering	John

Member	Position
Dave	Coach
Mary	Referee
John	Catering
Hugo	Referee



Elementary Key Normal Form

- **Boyce-Codd Normal Form (BCNF)**
  - Every attribute (key **or** non-key!) depends on each candidate key
    - (nothing more, nothing less)
  - Can't always be achieved
    - (But usually can!)

- **Elementary Key Normal Form (EKNF)**
  - Between 3NF and BCNF
    - Every non-key attribute depends on each candidate key
    - Every ***elementary*** key attribute depends on each candidate key
  - Does not solve all problems
  - **Can** always be achieved
    - Using Bernstein's algorithm for synthesis of a Third Normal Form schema
      - (which actually generates a schema in Elementary Key Normal Form)

Fourth Normal Form

http://commons.wikimedia.org/wiki/File:Snooker_rest_crop.JPG

Fourth Normal Form

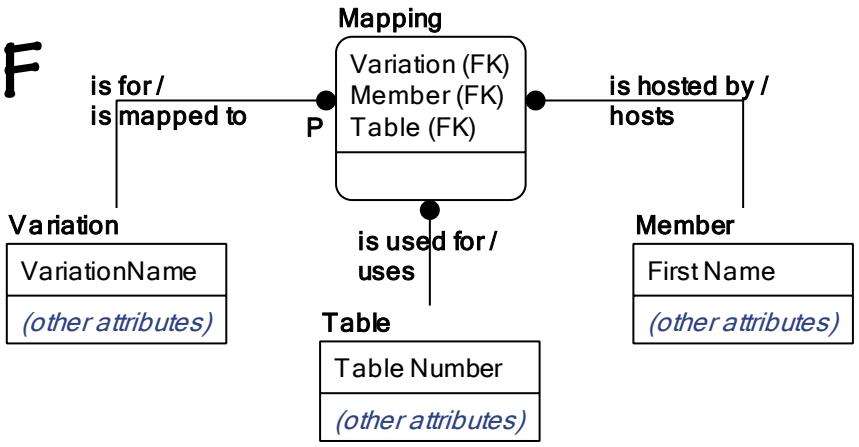
Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
Pooker	Mary	1
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3
Quick	Mary	4

Variation Pooker is hosted by Mary on table 2.

Variation Pooker is hosted by Mary.

Variation Pooker can be played on table 2.

✓BCNF



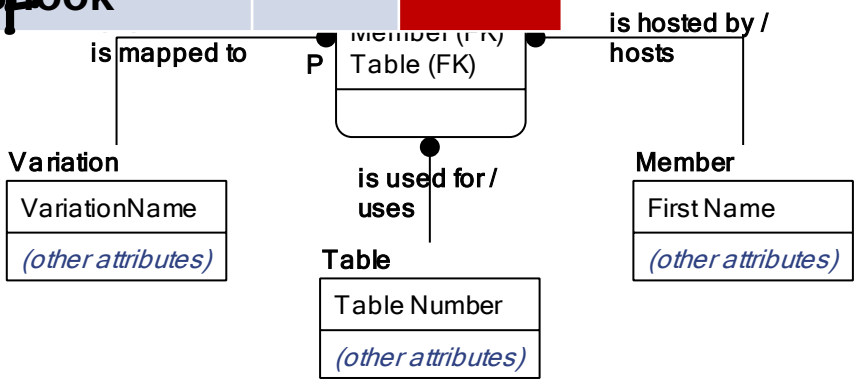
Fourth Normal Form

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
Pooker	Mary	1
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	1
Sneaker	John	1
Sneaker	Mary	2
Sneaker	?	3
Quick Snook	Mary	3
Quick Snook	Dave	?

✓ BCNF

✗ 4NF



Fourth Normal Form

- **Multivalued dependency**
  - Possible value**s** for dependent columns depend on determinant only
    - No influence from other columns
    - Special case: always one possible value – this is a functional dependency
  - Functional dependency:
    - {a, b} → {c, d} equivalent to {a, b} → c and {a, b} → d
  - Multivalued dependency:
    - {a, b} ↠ {c, d} **NOT** equivalent to {a, b} ↠ c and {a, b} ↠ d
  - Trivial multivalued dependency:
    - Every set of columns ↠ the combination of **ALL** other columns

Not relevant for 4NF

Fourth Normal Form

- **Requirements for Fourth Normal Form (4NF):**
  - Table must be in Boyce-Codd Normal Form
  - For every non-trivial dependency, the determinant is (a superset of) a key

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
Pooker	Mary	1
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3

Variation Pooker is hosted by Mary.

Variation Pooker can be played on table 2.

Pooker → Mary, Dave *(for each table)*

Variation ↠ Host

Pooker → Table 1, 2 *(for each host)*

Variation ↠ Table

Fourth Normal Form

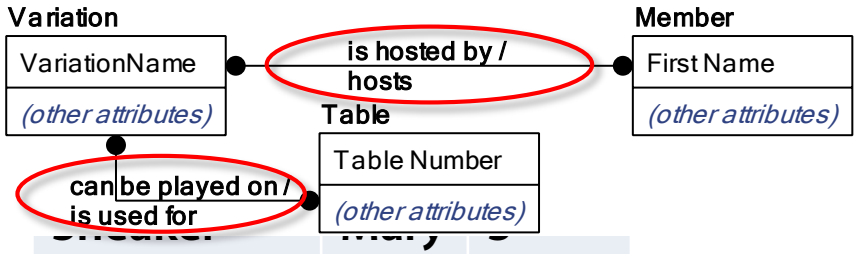
- **Requirements for Fourth Normal Form (4NF):**
  - Table must be in Boyce-Codd Normal Form
  - For every non-trivial dependency, the determinant is (a superset of) a key

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
Pooker	Mary	1
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3

Variation Pooker is hosted by Mary.

Variation Pooker can be played on table 2.

✓ 4NF



Variation	Table
Pooker	2
Pooker	1
Sneaker	1
Sneaker	2
Sneaker	3
Quick Snook	3

Fourth Normal Form

- **Requirements for Fourth Normal Form (4NF):**
  - Table must be in Boyce-Codd Normal Form
  - For every non-trivial dependency, the determinant is (a superset of) a key

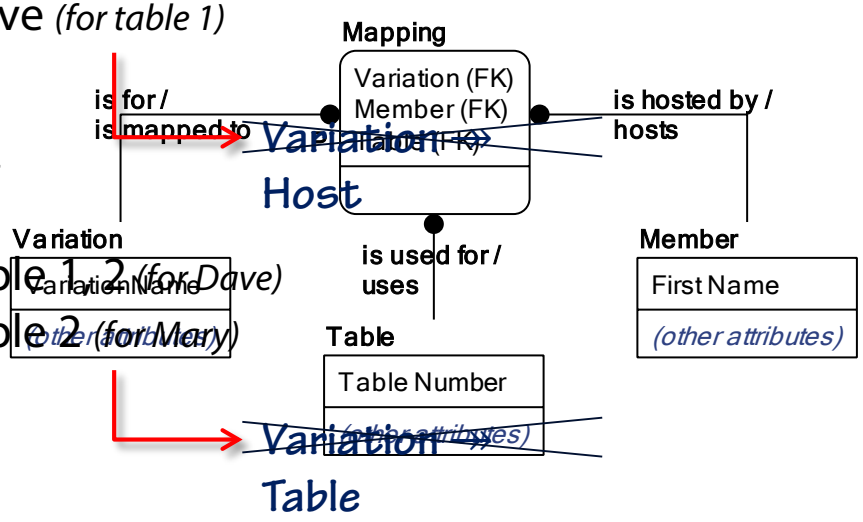
Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
~~Pooker~~	~~Mary~~	~~1~~
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3

Variation Pooker is hosted by Mary on table 2.

Pooker → Mary, Dave *(for table 2)*
Pooker → Dave *(for table 1)*

✓ 4NF

Pooker → Table 1, 2 *(for Dave)*
Pooker → Table 2 *(for Mary)*



Fifth Normal Form

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
Pooker	Mary	1
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3

Variation Pooker is hosted by Mary.

Variation Pooker can be played on table 2.

X 4NF

Fifth Normal Form

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
~~Pooker~~	~~Mary~~	~~1~~
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3

Variation Pooker is hosted by Mary.

Variation Pooker can be played on table 2.

Variation Pooker is hosted by Mary on table 2.

Fifth Normal Form

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
~~Pooker~~	~~Mary~~	~~1~~
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
~~Sneaker~~	~~Mary~~	~~1~~
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3

Variation Pooker is hosted by Mary.

Variation Pooker can be played on table 2.

Variation Pooker is hosted by Mary on table 2.

Mary can host games on table 2.
Mary can host games on table 3.

✓4NF

~~Variation ↠ Host~~
~~Variation ↠ Table~~
~~Host ↠ Variation~~
~~Host ↠ Table~~
Table ↠ Variation

Fifth Normal Form

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
Pooker	Mary	1
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3

Variation Pooker is hosted by Mary.

→ <variation> is hosted by <host>.

Variation Pooker can be played on table 2.

→ <variation> can be played on <table>.

Fifth Normal Form

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
~~Pooker~~	~~Mary~~	~~1~~
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3

Variation Pooker is hosted by Mary.

→ <variation> is hosted by <host>.

Variation Pooker can be played on table 2.

→ <variation> can be played on <table>.

Variation Pooker is hosted by Mary on table 2.

→ <variation> is hosted by <host> on <table>.

Fifth Normal Form

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
~~Pooker~~	~~Mary~~	~~1~~
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
~~Sneaker~~	~~Mary~~	~~1~~
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3

Variation Pooker is hosted by

Variatic

2.

Variatic

2.

Mary ca

<host> can host games on <t

Variation	Table
Pooker	2
Pooker	1
Sneaker	1
Sneaker	2
Sneaker	3
Quick Snook	3

yed

pla

by Mary on table

ted by <hos

able 2.

Variation	Host
Pooker	Mary
Pooker	Dave
Sneaker	John
Sneaker	Mary
Quick Snook	Mary
Quick Snook	Dave

Host	Table
Mary	2
Dave	2
Dave	1
John	1
John	2
John	3
Mary	3
Dave	3

JOIN

Fifth Normal Form

Variation	Host	Table
Pooker	Mary	2
Pooker	Dave	2
~~Pooker~~	~~Mary~~	~~1~~
Pooker	Dave	1
Sneaker	John	1
Sneaker	John	2
Sneaker	John	3
Sneaker	Mary	1
Sneaker	Mary	2
Sneaker	Mary	3
Quick Snook	Mary	3
Quick Snook	Dave	3

Variation Pooker is hosted by

Variatic

2.

Variatic

2.

Mary ca

<host> can host games on <t

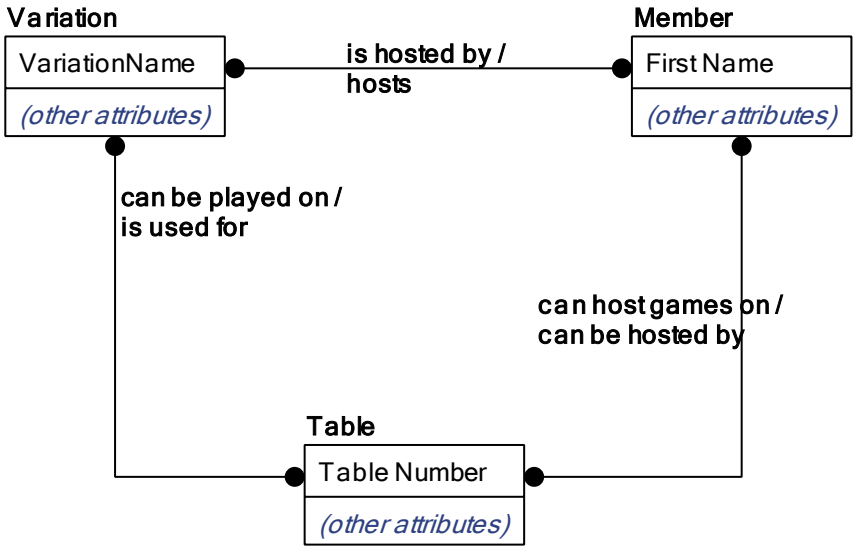
Variation	Table
Pooker	2
Pooker	1
Sneaker	1
Sneaker	2
Sneaker	3
Quick Snook	3

Variation	Host
Pooker	Mary
Pooker	Dave
Sneaker	John
Sneaker	Mary
Quick Snook	Mary
Quick Snook	Dave

Host	Table
Mary	2
Dave	2
Dave	1
John	1
John	2
John	3
Mary	3
Dave	3

JOIN

Fifth Normal Form



Variation	Table
Pooker	2
Pooker	1
Sneaker	1
Sneaker	2
Sneaker	3
Quick Snook	3

Variation	Host
Pooker	Mary
Pooker	Dave
Sneaker	John
Sneaker	Mary
Quick Snook	Mary
Quick Snook	Dave

Host	Table
Mary	2
Dave	2
Dave	1
John	1
John	2
John	3
Mary	3
Dave	3

Fifth Normal Form

- **Join dependency**
  - Possible to split table in multiple new tables, …
  - … such that joining them reconstructs original table
  - Attributes represent independent facts
  - Trivial join dependency:
    - One of the new tables is equal to original table
  - Join dependency implied by candidate key:
    - Candidate key of original table included in every new table
  - Join dependency with two new tables:
    - Equivalent to multivalued dependency
    - Shared columns ↠ distributed columns

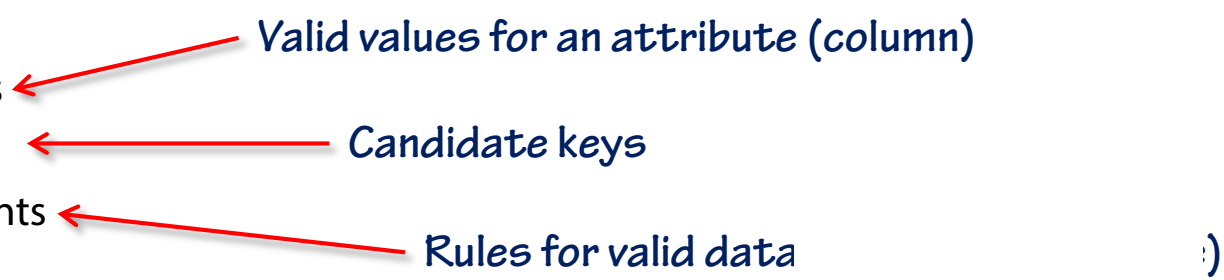
Fifth Normal Form

- **Requirements for Fifth Normal Form (5NF):**
  - Table must be in Fourth Normal Form
  - For every non-trivial dependency, the determinant is (a superset of) a key

- **Fifth Normal Form is also known as PJ/NF**
  - Project-Join Normal Form
  - Projection-Join Normal Form

Domain-Key Normal Form

- **Requirements for Domain-Key Normal Form (DK/NF):**
  - **NOT** based on dependencies
  - Based on:
    - Domains
    - Keys
    - Constraints
  - Every constraint must be implied by the keys and domains
    - (Implies Fifth –and lower– Normal Form)



Domain-Key Normal Form

- **Relevance of Domain-Key Normal Form**
  - Domains → declared, enforced (no code needed)
  - Keys → declared, enforced (no code needed)
  - Other constraints → code needed to enforce
    - Code = cost factor:
      - Time to write
      - Time to test and debug
      - Future maintenance

More code for constraints, less code for queries

- **Achievability of Domain-Key Normal Form**
  - Sometimes impossible
  - Otherwise often requires extra tables (subtypes)
    - Introduces need for more (and more complex) code

Less code for constraints, more code for querying

Domain-Key Normal Form

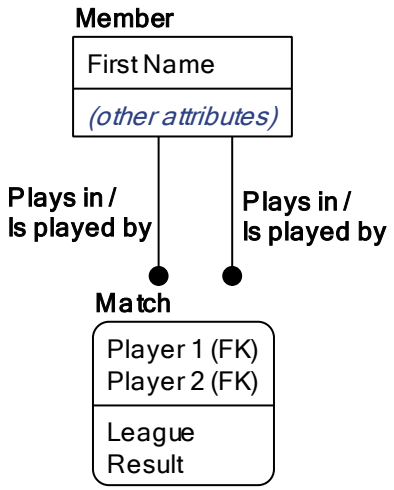
{3-0, 2-1, 1-2, 0-3,
5-0, 4-1, 3-2, 2-3, 1-4, 0-5}

(B and C league only)
(A league only)

{A, B, C}

✗ DK/NF

Player 1	Player 2	League	Result
Mary	Dave	C	1-2
Tony	Jack	A	3-2
Hugo	Mary	C	0-3
Antonio	Anne	B	1-2



Domain-Key Normal Form

{5-0, 4-1, 3-2, 2-3, 1-4, 0-5}

✓ DK/NF

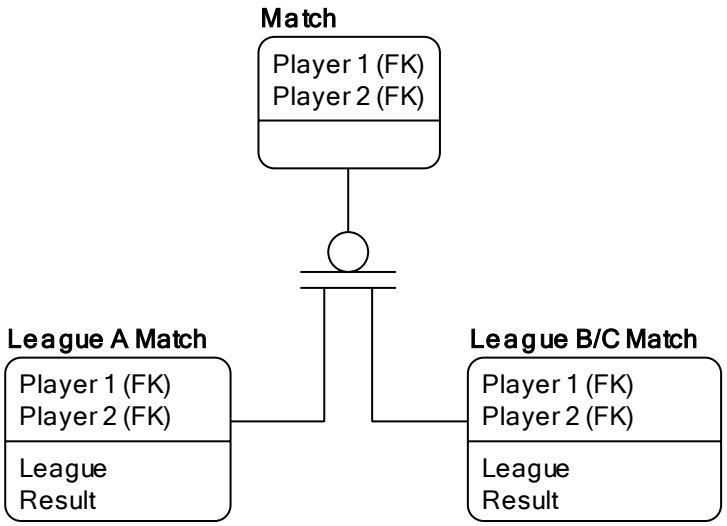
{A}

Player 1	Player 2	League	Result
Tony	Jack	A	3-2

{3-0, 2-1, 1-2, 0-3}

{B, C}

Player 1	Player 2	League	Result
Mary	Dave	C	1-2
Hugo	Mary	C	0-3
Antonio	Anne	B	1-2



Sixth Normal Form

- **Sixth Normal Form (6NF)**
  - Has historically sometimes been used for DK/NF
- **Fifth Normal Form:**
  - Non-trivial join dependency not allowed, unless implied by a candidate key
- **Sixth Normal Form:**
  - Non-trivial join dependency not allowed~~, unless implied by a candidate key~~
  - Any 6NF table has candidate key + at most one other column

Sixth Normal Form

- **Relevance of Sixth Normal Form**
  - No NULLs needed in storage

Fifth Normal Form

Member	Birthdate	League	Board position
Mary	*NULL*	C	Secretary
Jack	15 Dec 1960	A	*NULL*
Anthony	8 Apr 1975	B	*NULL*

Sixth Normal Form

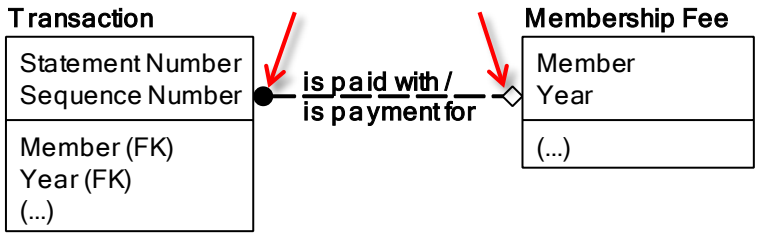
Member	League
Mary	C
Jack	A
Anthony	B

Member	Birthdate
Jack	15 Dec 1960
Anthony	8 Apr 1975

Member	Board position
Mary	Secretary

Sixth Normal Form

- **Relevance of Sixth Normal Form**
  - No NULLs needed in storage



- **Consequence of Sixth Normal Form**
  - Huge number of tables
  - Many joins required for most queries
  - Simple constraints can be difficult to enforce

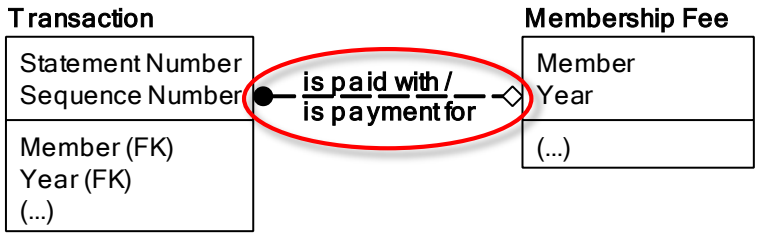
Bank Statement # 396

(April 16 through April 23, 2013)

No	Date	D/C	Amount	Description
1	April 17, 2013	C	420.00	Membership Fee for Dave, 2013
2	April 17, 2013	D	74.95	Replacement Cloth for Table 3
3	April 22, 2013	C	105.00	Membership Fee for Mary, 2013 (April - June)
4	April 23, 2013	C	35.00	Membership Fee for Hugo, 2013 (May)

Sixth Normal Form

- **Relevance of Sixth Normal Form**
  - No NULLs needed in storage



- **Consequence of Sixth Normal Form**
  - Huge number of tables
  - Many joins required for most queries
  - Simple constraints can be difficult to enforce

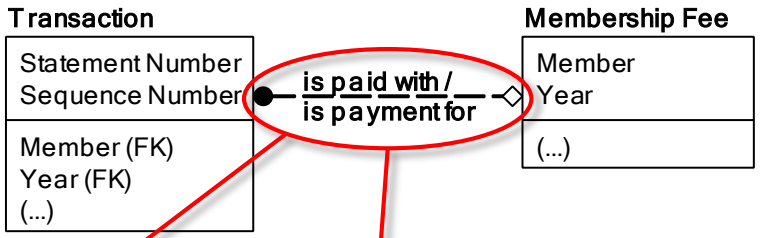
```
CONSTRAINT IsPaymentFor
 FOREIGN KEY (Member, Year)
 REFERENCES MembershipFees (Member, Year)
```

Fifth Normal Form

StmtNbr	SeqNbr	Member	Year	(...)
396	1	Dave	2013	(...)
396	2	*NULL*	*NULL*	(...)
396	3	Mary	2013	(...)
396	4	Hugo	2013	(...)

Sixth Normal Form

- **Relevance of Sixth Normal Form**
  - No NULLs needed in storage
- **Consequence of Sixth Normal Form**
  - Huge number of tables
  - Many joins required for most queries
  - Simple constraints can be difficult to enforce



Sixth Normal Form

StmtNbr	SeqNbr	Member
396	1	Dave
396	3	Mary
396	4	Hugo

StmtNbr	SeqNbr	Year
396	1	2013
396	3	2013
396	4	2013

(...)

Optimal Normal Form

- **Optimal Normal Form (ONF)**
  - Based on fact-based modeling methods (e.g. ORM, NIAM)
  - Every *fact type* becomes a table

StmtNbr	SeqNbr	Amount
396	1	420.00
396	2	74.95
396	3	105.00
396	4	35.00

StmtNbr	SeqNbr	TranDate
396	1	2013-04-17
396	2	2013-04-17
396	3	2013-04-22
396	4	2013-04-23

StmtNbr	SeqNbr	Member	Year
396	1	Dave	2013
396	3	Mary	2013
396	4	Hugo	2013

```
CONSTRAINT IsPaymentFor
  FOREIGN KEY (Member, Year)
  REFERENCES MembershipFees (Member, Year)
```

Denormalization

- **OLTP (Online Transaction Processing)**
  - Normalization is always good

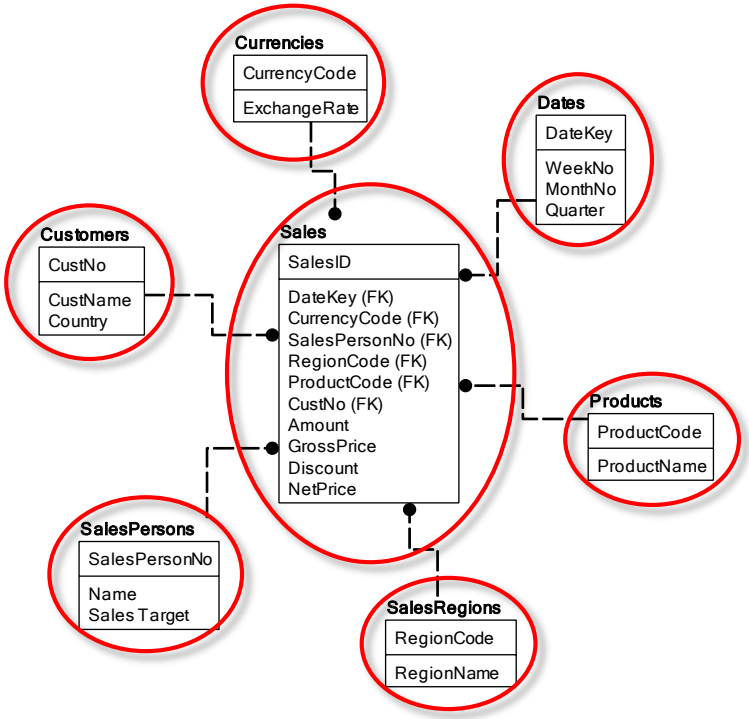
- **OLAP (Online Analytical Processing) and reporting**
  - Normalization may harm performance
  - Solution: denormalize
    - **NOTE:** Unnormalized **is not** denormalized!
    - Denormalization is done **after** normalization

Denormalization

- **Storing the result of a query**
  - "Do it yourself"
    - Store query results in table
    - Add application logic to update stored results when underlying data changes
  - Feature in some RDBMS's
    - Materialized view
    - Indexed view
  - Typical use cases
    - Data stored in separate tables but often queries together
    - Pre-aggregating data

Denormalization

- **Star schema**
  - Fact table
  - Dimension tables
- **Snowflake schema**



Denormalization

- **OLAP cubes**
  - Specifically designed to support Online Analytical Processing
  - Pre-aggregates data on several levels
  - Usually built on top of a star or snowflake schema

Denormalization

- **Non-First Normal Form (N1NF or NF^2)**
  - Attribute can be a single value (1NF)
  - **OR** attribute can be a complete table (violates 1NF)
    - Results in nested tables
  - Rigid design still applies
  - Multi-valued databases comes close

Variation	Hosting
Pooker	Host: Mary, Table: 2 Host: Dave, Table: 1 Host: Dave, Table: 2
Sneaker	Host: John, Table: 1 Host: John, Table: 2 Host: Mary, Table: 2 Host: Mary, Table: 3
Quick Snook	Host: Mary, Table: 3 Host: Dave, Table: 3

Non-standard patterns

- Referencing a range

Ranking points	League
0 – 500	C
500 – 1500	B
1500+	A

+ 37 ranking points

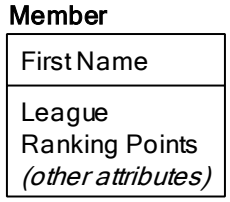
Total ranking points:
1,496

League: B

Functional dependency:
- Ranking Points → League

Non-standard patterns

- Referencing a range

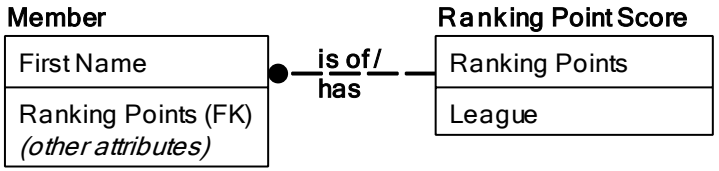


FirstName	League	RankingPoints	(…)
Dave	C	478	(…)
Hugo	C	271	(…)
Jack	A	2080	(…)
Mary	B	772	(…)

Functional dependency:
- Ranking Points → League

Non-standard patterns

- Referencing a range



FirstName	RankingPoints	(…)
Dave	478	(…)
Hugo	271	(…)
Jack	2080	(…)
Mary	772	(…)

RankingPoints	League
0	C
1	C
…	…
500	C
501	B
…	…
1500	B
1501	A
…	…

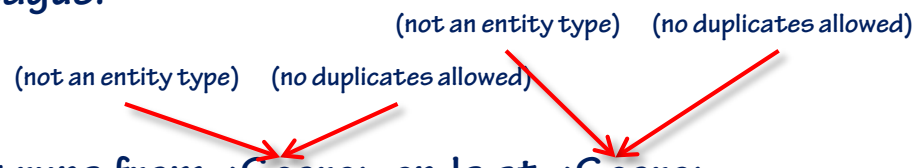
Functional dependency:
- Ranking Points → League

Non-standard patterns

- Referencing a range

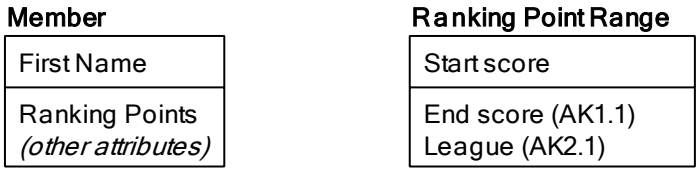
Ranking points	League
0 – 500	C
500 – 1500	B
1500+	A

The ranking point range that runs from 0 to 500 corresponds to the C league.
The ranking point range that runs from 500 to 1,500 corresponds to the B league.
The ranking point range that runs from 1,500 corresponds to the C league.



Non-standard patterns

- Referencing a range



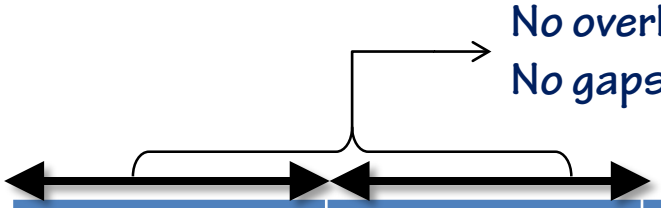
FirstName	RankingPoints	(…)
Dave	478	(…)
Hugo	271	(…)
Jack	2080	(…)
Mary	772	(…)

Start score	End score	League
0	500	C
500	1500	B
1500	*NULL*	A

Non-standard patterns

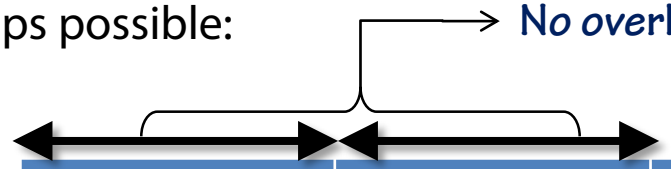
- **Referencing a range**
  - No gaps:

Lowerbound	(…)
0	(…)
500	(…)
1500	(…)



Lowerbound	Upperbound	(…)
0	500	(…)
500	1500	(…)
1500	*NULL*	(…)

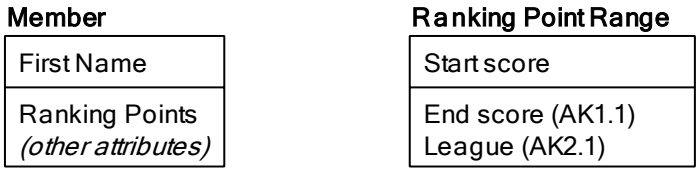
  - Gaps possible:



Lowerbound	Upperbound	(…)
0	400	(…)
600	1500	(…)
1500	*NULL*	(…)

Non-standard patterns

- Referencing a range

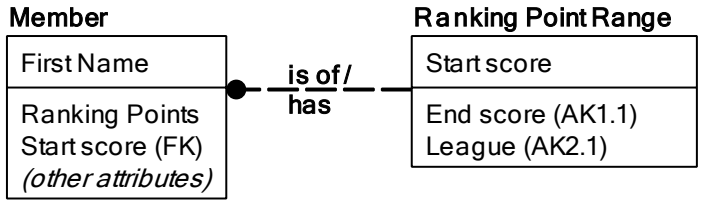


FirstName	RankingPts	(…)
Dave	478	(…)
Hugo	-271	(…)
Jack	2080	(…)
Mary	772	(…)

StartScore	EndScore	League
0	500	C
500	1500	B
1500	*NULL*	A

Non-standard patterns

- Referencing a range

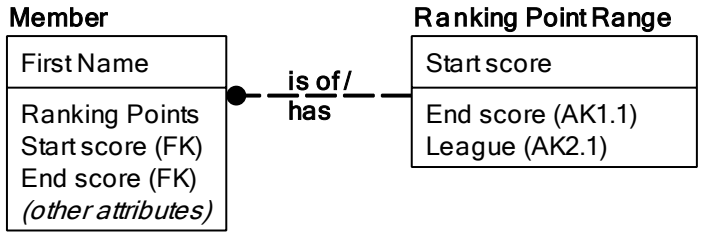


FirstName	RankingPts	StartScore	(…)
Dave	478	0	(…)
Hugo	271	0	(…)
Jack	2080	500	(…)
Mary	772	500	(…)

StartScore	EndScore	League
0	500	C
500	1500	B
1500	*NULL*	A

Non-standard patterns

- Referencing a range



FirstName	RankingPts	StartScore	EndScore	
Dave	478	0	500	
Hugo	271	0	500	
Jack	2080	1500	*NULL*	(…)
Mary	772	500	1500	(…)

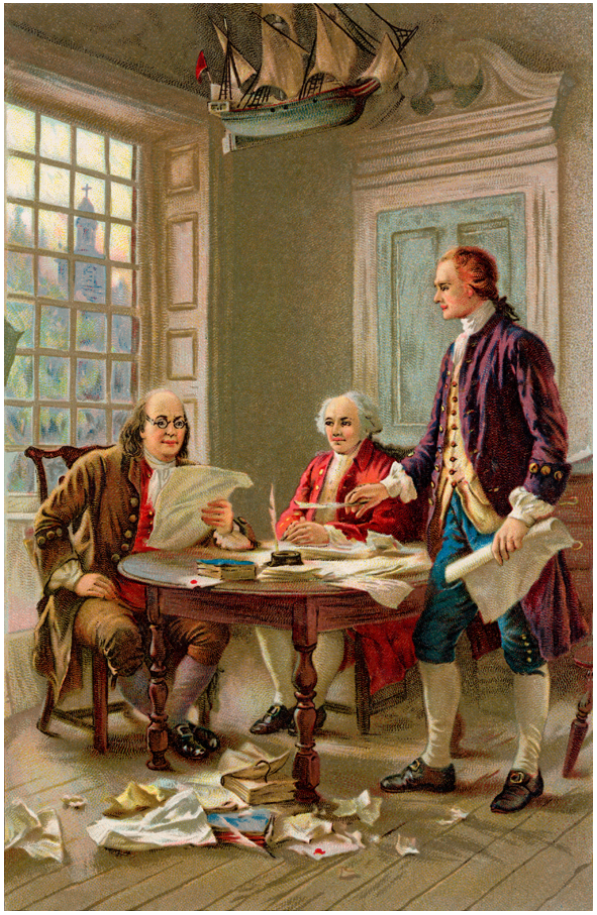
StartScore	EndScore	League
0	500	C
500	1500	B
1500	*NULL*	A

```
FOREIGN KEY (StartScore, EndScore)
REFERENCES RankingPointRange (StartScore, EndScore)

CHECK (RankingPts BETWEEN StartScore AND EndScore)
```

Non-standard patterns

- Storing historic data



Non-standard patterns

- **Storing historic data**
  - OLTP:
    - Mostly current data
  - OLAP:
    - Lots of historic data
    - Standard design pattern (“Slowly changing dimensions”)

Non-standard patterns

- **Storing historic data**
  - OLTP:
    - Mostly current data
    - Some historic data needed as well

ProductID	Product Name	Price
123	Snooker cue with extension	208.50
456	Standard snooker ball set	65.00
789	Replacement felt	195.00

67.50

Non-standard patterns

- **Storing historic data**
  - OLTP:
    - Mostly current data
    - Some historic data needed as well

ProductID	Product Name	Price	OldPrice	DatePriceChange
123	Snooker cue with extension	208.50	*NULL*	*NULL*
456	Standard snooker ball set	65.00	67.50	Oct 5, 2013
789	Replacement felt	195.00	*NULL*	*NULL*

Non-standard patterns

- **Storing historic data**
  - OLTP:
    - Mostly current data
    - Some historic data needed as well

ProductID	Product Name	Price	ValidTo
123	Snooker cue with extension	208.50	*NULL*
456	Standard snooker ball set	65.00	Oct 5, 2013
456	Standard snooker ball set	67.50	*NULL*
789	Replacement felt	195.00	*NULL*

Non-standard patterns

- **Storing historic data**
  - OLTP:
    - Mostly current data
    - Some historic data needed as well

Sold "Standard snooker ball set" on August 12, 2013

ProductID	Product Name	Price	ValidTo
123	Snooker cue with extension	208.50	Dec 31, 9999
456	Standard snooker ball set	65.00	Oct 5, 2013
456	Standard snooker ball set	67.50	Dec 31, 9999
789	Replacement felt	195.00	Dec 31, 9999

Non-standard patterns

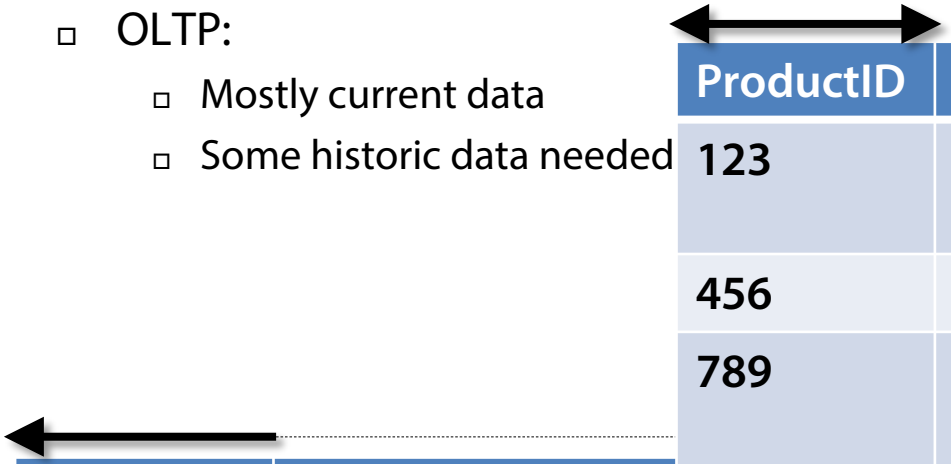
- **Storing historic data**
  - OLTP:
    - Mostly current data
    - Some historic data needed

ProductID	Product Name	Price
123	Snooker cue with extension	208.50
456	Standard snooker ball set	67.50
789	Replacement felt	195.00

ProductID	Product Name	Price	ValidTo
123	Snooker cue with extension	208.50	Dec 31, 9999
456	Standard snooker ball set	65.00	Oct 5, 2013
456	Standard snooker ball set	67.50	Dec 31, 9999
789	Replacement felt	195.00	Dec 31, 9999

Non-standard patterns

- **Storing historic data**
  - OLTP:
    - Mostly current data
    - Some historic data needed



ProductID	Product Name	Price
123	Snooker cue with extension	208.50
456	Standard snooker ball set	67.50
789	Replacement felt	195.00

ProductID	Product Name	Price	ValidTo
456	Standard snooker ball set	65.00	Oct 5, 2013

ProductID	Product Name	Price	ValidTo
123	Snooker cue with extension	208.50	*NULL*
456	Standard snooker ball set	65.00	Oct 5, 2013
456	Standard snooker ball set	67.50	*NULL*
789	Replacement felt	195.0	*NULL*

Non-standard patterns

- **Storing historic data**
  - OLTP:
    - Mostly current data
    - Some historic data needed

ProductID	Product Name	Price
123	Snooker cue with extension	208.50
456	Standard snooker ball set	67.50
789	Replacement felt	190.00

ProductID	Product Name	Price	ValidTo
456	Standard snooker ball set	65.00	Oct 5, 2013
789	Replacement felt	195.00	Oct 27, 2013

			To
123	Snooker cue with extension	208.50	*NULL*
456	Standard snooker ball set	65.00	Oct 5, 2013
456	Standard snooker ball set	67.50	*NULL*
789	Replacement felt	195.0	Oct 27, 2013

Non-standard patterns

- **Storing historic data**
  - OLTP:
    - Mostly current data
    - Some historic data needed as well

ProductID	ValidTo	Price
123	Dec 31, 9999	208.50
456	Oct 5, 2013	65.00
456	Dec 31, 9999	67.50
789	Oct 27, 2013	195.00
789	Dec 31, 9999	190.00

ProductID	ValidTo	Product Name
123	Dec 31, 9999	Snooker cue with extension
456	Dec 31, 9999	Standard snooker ball set
789	Dec 31, 9999	Replacement felt

Summary

- **Higher normal forms**
  - Elementary Key Normal Form (EKNF)
  - Boyce-Codd Normal Form (BCNF)
  - Fourth Normal Form (4NF)
  - Fifth Normal Form (5NF)
  - Domain/Key Normal Form (DK/NF)
  - Sixth Normal Form (6NF)
    - Optimal Normal Form (ONF)
- **Denormalization**
- **Non-standard problems**
  - Referencing a range
  - Storing historic data

References

- **Further reading:**
  - Database normalization on Wikipedia:
    http://en.wikipedia.org/wiki/Database_normalization
    or
    http://tinyurl.com/Norm-DB
  - Pluralsight Course Library
    - Business Intelligence section