

Relational Databases

Enterprise systems need to store lots of data and use the *relational model* to achieve this. I have devised the following basketball league example to kick-start your understanding of relational models.

Although the example below seems unrelated to enterprise accounting systems, these systems use exactly the same concepts and understanding the structure of the example gets you well on your way to understanding how enterprise systems store data.

The beauty of this way of storing information is that when we store the information, we do not really need to worry about what kinds of retrieval requirements exist. We simply represent the information. After that, we can use (as we will see later) the SQL language to easily and quickly extract pretty much anything we want.

SBA

Here is some information about a basketball league. You can see that the information is organized in the form of tables. Each table stores information about a set of “things” of the same type with each row providing information on one object of the type.

For example, the “Coaches” table stores information about all the coaches in the league and each row has information about a coach. Each column represents an attribute of the type of object stored in a table. For example, for each coach, we are storing a coach_id and the coach’s first and last names.

Information between tables is sometimes related. For example, the table Teams has a field called coach_id to associate a coach with each table. For example, the first team has coach_id as 2, which means that Elliott Finch is the coach of this team.

Take a look at the tables Coaches, Venues, Players and Teams; you should find them to be almost self-explanatory.

Table: Coaches

coach_id	coach_first_name	coach_last_name
1	John	Smith
2	Elliott	Finch
3	Eddie	Murphy
4	Illya	Dragovich
5	Mad	Max
6	Arthur	Eddington
7	Paul	Brigley
8	Ivan	Dostov
9	Michael	Bloomberg
10	Joseph	Kennedy

Table: Venues

venue_id	venue_name	street_address	city	state	zip	capacity	year_built
----------	------------	----------------	------	-------	-----	----------	------------

1	Edison Arena	200 Main street	Edison	NJ	08820	5000	2000
2	Washington Nest	1600 Pennsylvanis Avenue	Washington	DC	20500	10000	1809
3	Mad Square Den	300 Fifth Avenue	New York San Francisco	NY	10011	20000	1968
4	Silicon Golly Hole Puch Arena	200 Market Street	Hollywood	CA	90210	48000	1975
5	Royal Arch Bowl	300 Main Street	St Louis	MO	63101	15000	1995
7	Peak Vista Arena	7500 Airport Blvd	Denver	CO	80010	23000	1965
8	Buffalo Yards	3240 Buffalo Bill Road	Buffalo	NY	14201	14000	1956
9	Lake View Arena	1000 Blagovich Drive	Chicago	IL	60601	35000	1986
10	Innovation Space	7000 Crooked Alley	Detroit	MI	48201	24000	1990

Players

player_id	player_first_name	player_last_name	birth_date	height	weight	team_id	home_state	birth_state
1	DARIUS	Brands	6/16/1991	76.2	197.5	1	MN	ND
2	BRAIN	Kunsch	3/28/1992	74.3	185.4	1	MD:	MD
3	KENDRICK	Kafton	6/23/1993	72.9	181	1	VT	VT
4	ARMAND	Lachat	3/23/1991	72.2	189.7	1	MT	MT
5	ALI	Borio	1/6/1993	74.3	180.6	1	FL	FL
6	JACKSON	Samay	5/29/1989	68.5	175	1	MN	MN
7	DANE	Talahytewa	12/13/1985	69.9	199.4	1	CT	CT
8	CLIFF	Casher	9/18/1987	73.9	201.7	1	ID	ID
9	MARCEL	Jacot	3/4/1985	76.5	183.5	1	ND	ND
10	MARCEL	Titze	6/22/1986	70	188.2	1	HI	HI
11	RAPHAEL	Sitko	9/21/1990	72.9	196.4	2	ND	ND
12	ALI	Aufderheide	11/5/1990	71.9	193.2	2	AL	AL
13	CAREY	Dense	3/22/1991	73.7	180.8	2	FL	FL
14	JACKSON	Vay	8/13/1988	73.7	186.3	2	NJ	NJ
15	BRAIN	Gattshall	9/8/1991	69.5	187	2	ND	ND
16	KENDRICK	Burde	10/8/1985	73.2	202.9	2	FL	FL
17	ALI	Dicapua	12/30/1992	70.8	183.1	2	ME	ME
18	JACKSON	Aitcheson	7/8/1987	71.3	203.9	2	NH	NH
19	MOISES	Nore	11/11/1999	72.1	175.3	2	ME	ME
20	CLIFF	Shely	1/27/1987	72.4	196.3	2	OK	OK
21	ALVARO	Billiott	8/17/1992	73.1	208.5	3	DE	DE
22	JEFFRY	Pacitto	3/7/1987	72.6	184.9	3	NH	NH
23	BRYON	Younie	2/25/1995	75.8	191.4	3	ME	ME
24	JACKSON	Sherick	3/29/1988	70.9	186.9	3	FL	FL
25	MOISES	Kopperman	5/16/1994	75.6	191.8	3	HI	HI
26	DARIUS	Mehtani	10/7/1986	76.7	188	3	AR	AR
27	RAPHAEL	Ramming	11/29/1990	66.3	204.4	3	ID	ID
28	ROB	Ducotey	8/30/1987	76.4	196.7	3	ND	ND
29	ROD	Janikowski	12/16/1999	69.3	183.9	3	CA	CA

player_id	player_first_name	player_last_name	birth_date	height	weight	team_id	home_state	birth_state
			2					
30	ROB	Franzi	10/30/1990	77.9	179.9	3	AR	AR
31	HAL	Barrack	12/8/1993	71.1	185.2	4	DE	DE
32	RAPHAEL	Sherretts	1/18/1986	76.2	187	4	ID	ID
33	MARCEL	Entsminger	5/24/1990	78.3	176.9	4	NH	NH
34	CLIFF	Eggink	5/24/1990	68.7	177.1	4	KS	KS
35	DANE	Rindone	6/8/1994	73.7	190.9	4	NV	NV
36	RAPHAEL	Murdough	12/13/1987	73.1	196.5	4	DE	DE
37	ALVARO	Langman	1/27/1985	73.9	190.3	4	PA	PA
38	CLIFF	Kolinski	6/15/1990	71.6	195.6	4	IL	IL
39	ELWOOD	Briskey	8/4/1985	75.7	172.8	4	AZ	AZ
40	JACKSON	Barness	5/28/1991	78.7	186.4	4	HI	HI
41	MARCEL	Montee	5/2/1991	74	194.8	5	NC	NC
42	THADDEUS	Shadler	7/26/1985	75.6	196.7	5	IA	IA
43	HAL	Dellano	10/19/1985	76.3	187.2	5	NV	NV
44	ARMAND	Stempien	1/18/1991	70.5	194.4	5	RI	RI
45	ARMAND	Borkoski	10/8/1985	71.4	192	5	PA	PA
46	RAPHAEL	Henagan	10/11/1985	72.3	213.6	5	HI	HI
47	DANE	Pincince	5/6/1988	70.8	192.6	5	NY	NY
48	ROD	Fogelson	5/20/1994	75.3	205.6	5	NY	NY
49	JACKSON	Mccarson	2/27/1990	69	198.1	5	NY	NY
50	CLIFF	Sanburg	4/26/1991	74.3	216.3	5	CO	CO
51	KENDRICK	Colbenson	8/13/1988	76.5	192.3	6	MO	MO
52	ROD	Dalphonse	7/12/1995	72.7	183.1	6	DE	DE
53	MOISES	Dohn	3/18/1988	72.9	183.1	6	KY	KY
54	JACKSON	Ausland	6/19/1986	72.3	219.7	6	LA	LA
55	CLIFF	Khamo	4/14/1992	70.1	172.2	6	AZ	AZ
56	FIDEL	Maretti	3/11/1985	71.2	200.2	6	IA	IA
57	MARCEL	Gambler	2/16/1985	71.8	195.2	6	NJ	NJ
58	JACKSON	Fullagar	12/9/1986	74.5	193.4	6	NY	NY
59	BRAIN	Kritzer	7/4/1992	72.3	201.9	6	HI	HI
60	JEFFRY	Cintora	7/13/1993	71.2	203.4	6	AR	AR
61	Jack	Fasheh	6/18/1994	71.3	181.5	7	ME	ME
62	NOLAN	Benthin	6/4/1990	76.3	192.9	7	ND	ND
63	MOISES	Newville	3/27/1994	72.7	201.7	7	AL	AL
64	HAL	Carotenuto	2/10/1987	73.7	207.1	7	VT	VT
65	MARCEL	Eison	3/15/1986	69.4	186.3	7	NH	NH
66	ELWOOD	Ahrends	12/1/1987	71.9	175.3	7	KS	KS
67	DANE	Roskelley	11/14/1995	81	183.6	7	ID	ID
68	MARCEL	Colom	5/14/1990	78.8	207.3	7	ME	ME
69	MOISES	Arave	6/6/1987	76.5	195.1	7	NH	NH
70	DARIUS	Bokman	7/22/1993	74.6	193.1	7	ID	ID
71	CAREY	Mccament	5/8/1990	78	176.4	8	LA	LA
72	FIDEL	Blondell	9/19/1988	72.6	189.5	8	AL	AL
73	ELWOOD	Farb	1/24/1985	72.8	200.1	8	NM	NM
74	ARMAND	Delprincipe	4/16/1990	75.1	204.8	8	AL	AL
75	MARCEL	Texiera	10/5/1989	73.8	182.8	8	GA	GA
76	ALI	Perrenoud	7/11/1989	71	197.5	8	NJ	NJ

player_id	player_first_name	player_last_name	birth_date	height	weight	team_id	home_state	birth_state
77	THADDEUS	Boatfield	3/8/1992	72.8	200.9	8	WI	WI
78	FIDEL	Bushart	3/30/1987	72.1	199.8	8	NH	NH
79	JACKSON	Joncas	11/8/1991	78.5	193.3	8	MA	MA
80	SON	Mellow	6/29/1989	74.9	183.1	8	DE	DE
81	THADDEUS	Joliet	5/19/1987	69.2	189.8	9	IA	IA
82	BRAIN	Plane	1/27/1988	74.9	196.5	9	PA	PA
83	ALVARO	Thyberg	2/27/1987	76.5	195.4	9	HI	HI
84	ELWOOD	Elward	9/6/1990	71.7	181.1	9	MD:	MD:
85	MARCEL	Barbarin	8/8/1994	76.8	192.4	9	NJ	NJ
86	JEFFRY	Callum	4/6/1990	70.5	187.4	9	MO	MO
87	ALVARO	Hillegas	3/25/1990	72.4	194	9	PA	PA
88	CLIFF	Sibilia	1/20/1994	70.9	179.7	9	NY	NY
89	JEFFRY	Grissinger	6/29/1993	70.1	192.1	9	AR	AR
90	ALI	Gehler	9/2/1990	71.7	197.5	9	MD:	MD
91	ROBERT	Mcteer	4/15/1994	69.3	189.2	10	MD:	MD
92	MARLIN	Hummingbird	12/24/1986	70.2	208.9	10	NC	NC
93	KENDRICK	Grates	12/31/1991	74.7	176.8	10	FL	FL
94	ALI	Boike	9/13/1995	75.1	179	10	NV	NV
95	THADDEUS	Requena	7/20/1986	73.2	185.4	10	FL	FL
96	CLIFF	Cuna	7/20/1985	72.6	195	10	PA	PA
97	MARLIN	Waldhauser	10/31/1994	75.8	195.7	10	TN	TN
98	CLIFF	Pratka	5/20/1991	71.1	149.9	10	MN	MN
99	SON	Mertel	9/11/1991	74.7	185.9	10	MO	MO
100	DANE	Bellanca	4/18/1995	78.2	188.8	10	NM	NM

In the table Teams, the column coach_id tells us who is the coach of a team (we can use the coach_id and from the Coaches table find the name of the coach. The column home_venue_id tells us the venue at which each team plays its home games. We can use the home_venue_id and from the Venue table find the name and other characteristics of the venue. Similarly, we can use the Players table in conjunction with the captain_id and find which player is the captain of the team.

Teams

team_id	team_name	coach_id	home_venue_id	captain_id
1	Orange Peels	2	3	3
2	Rockets	3	4	14
3	Raging Bulls	4	5	28
4	Royal Tuskers	5	6	31
5	Towering Dunkers	6	7	46
6	Hilbillies	7	8	57
7	Martyrs	8	9	61
8	Patriots	9	10	78
9	Marlins	10	1	86
10	Electrics	1	2	91

The table Games contains information on all the games played during the league, with one row per game. For each game we have information about the teams that played in the game (the ids of the corresponding teams are shown), where the game was played and the points scored by each team.

Games

game_id	Team1_id	team2_id	venue	home team	team1 pts	team 2 pts	start	end
1	3	7	9	7	75	99	1/1/11 8:00 PM	1/1/11 9:34 PM
2	2	6	8	6	76	82	1/2/11 8:04 PM	1/2/11 9:42 PM
3	1	8	3	1	74	75	1/3/11 8:07 PM	1/3/11 9:40 PM
4	7	9	1	9	70	63	1/4/11 8:03 PM	1/4/11 9:37 PM
5	2	3	4	2	84	96	1/5/11 8:06 PM	1/5/11 9:32 PM
6	4	8	6	4	68	89	1/6/11 8:05 PM	1/6/11 9:35 PM
7	5	10	2	10	61	93	1/7/11 8:03 PM	1/7/11 9:26 PM
8	9	10	2	10	86	93	1/8/11 8:05 PM	1/8/11 9:34 PM
9	4	10	6	4	74	75	1/9/11 8:07 PM	1/9/11 9:41 PM
10	4	5	7	5	79	85	1/10/11 8:02 PM	1/10/11 9:28 PM
11	1	4	3	1	65	92	1/11/11 8:05 PM	1/11/11 9:41 PM
12	1	6	3	1	78	85	1/12/11 8:08 PM	1/12/11 9:37 PM
13	5	7	7	5	96	75	1/13/11 8:08 PM	1/13/11 9:45 PM
14	4	9	1	9	76	72	1/14/11 8:04 PM	1/14/11 9:34 PM
15	1	9	1	9	80	73	1/15/11 8:04 PM	1/15/11 9:35 PM
16	7	10	9	7	80	90	1/16/11 8:03 PM	1/16/11 9:30 PM
17	2	9	4	2	81	76	1/17/11 8:08 PM	1/17/11 9:38 PM
18	8	10	10	8	96	89	1/18/11 8:03 PM	1/18/11 9:25 PM
19	3	8	5	3	90	96	1/19/11 8:03 PM	1/19/11 9:40 PM
20	3	6	5	3	80	83	1/20/11 8:06 PM	1/20/11 9:41 PM
21	3	10	5	3	94	93	1/21/11 8:05 PM	1/21/11 9:31 PM
22	3	9	1	9	83	87	1/22/11 8:04 PM	1/22/11 9:37 PM
23	2	10	2	10	95	96	1/23/11 8:06 PM	1/23/11 9:41 PM
24	1	5	7	5	100	88	1/24/11 8:07 PM	1/24/11 9:33 PM
25	6	8	10	8	96	77	1/25/11 8:10 PM	1/25/11 9:33 PM
26	1	7	9	7	90	90	1/26/11 8:01 PM	1/26/11 9:26 PM
27	5	8	10	8	74	81	1/27/11 8:05 PM	1/27/11 9:35 PM
28	5	9	7	5	90	93	1/28/11 8:05 PM	1/28/11 9:38 PM
29	4	6	6	4	94	73	1/29/11 8:05 PM	1/29/11 9:32 PM
30	3	4	5	3	76	73	1/30/11 8:04 PM	1/30/11 9:27 PM

game_id	Team1_id	team2_id	venue	home team	team1 pts	team 2 pts	start	end
31	5	6	8	6	72	80	1/31/11 8:07 PM	1/31/11 9:34 PM
32	6	7	8	6	71	70	2/1/11 8:08 PM	2/1/11 9:39 PM
33	2	8	10	8	89	58	2/2/11 8:05 PM	2/2/11 9:33 PM
34	7	8	9	7	80	80	2/3/11 8:03 PM	2/3/11 9:35 PM
35	6	10	2	10	77	93	2/4/11 8:01 PM	2/4/11 9:23 PM
36	1	3	5	3	67	93	2/5/11 8:03 PM	2/5/11 9:29 PM
37	8	9	1	9	87	86	2/6/11 8:05 PM	2/6/11 9:30 PM
38	1	10	3	1	82	79	2/7/11 8:04 PM	2/7/11 9:33 PM
39	6	9	8	6	74	74	2/8/11 8:07 PM	2/8/11 9:32 PM
40	4	7	9	7	83	86	2/9/11 8:05 PM	2/9/11 9:37 PM
41	1	2	3	1	86	68	2/10/11 8:06 PM	2/10/11 9:32 PM
42	2	4	6	4	75	83	2/11/11 8:05 PM	2/11/11 9:28 PM
43	3	5	7	5	96	86	2/12/11 8:04 PM	2/12/11 9:31 PM
44	2	5	4	2	91	74	1/0/00 8:05 PM	2/13/11 9:39 PM
45	2	7	4	2	73	80	2/14/11 8:08 PM	2/14/11 9:46 PM

The table Participations needs very careful examination. It shows, for each occasion when a player played in a game, the details of the player's performance. Since we have 45 games and at least 10 players would have played in each game (often more than 10), this table would have to show at least 450 rows of information – one row each for each time a player played in a game. We have shown below only a small extract from the table.

For example, the first row tells us information about 5 players who played in game 1. The first of these was player 21 who played for 20 minutes in the game, scored 10 points, made 3 assists and committed 3 fouls.

Participations (partial information only – real data has many more rows)

game_id	player_id	starter	total_time	points	assists	fouls
1	21	TRUE	20	10	3	3
1	24	TRUE	40	12	0	4
1	27	FALSE	10	2	1	0
1	63	TRUE	60	35	5	3
1	66	TRUE	45	13	2	2
1	68	TRUE	20	4	0	3
2	11	FALSE	5	3	1	1
2	12	TRUE	45	30	4	4
2	15	TRUE	40	18	3	4
2	52	TRUE	34	12	2	4
2	55	TRUE	30	22	4	2
2	59	FALSE	13	2	0	0
3	5	TRUE	30	21	3	5
3	6	FALSE	12	3	2	1
3	7	TRUE	40	15	2	4
3	71	FALSE	15	3	0	2
3	75	TRUE	25	12	3	2

game_id	player_id	starter	total_time	points	assists	fouls
3	79	FALSE	18	5	1	1

Answer the following questions. For the highlighted questions, you only need to briefly explain how you would find the answer, because finding the actual answer might be cumbersome.

1. How many teams does the league have?
2. How many players are in team Hillbillies?
3. How many home games have the Marlins played?
4. For each game played at venue with id 5, show the names of the teams that played there and the points each one scored in each game.
5. Which team(s) has played the maximum number of home games?
6. Which was the longest game (time-wise)?
7. Can the same team number appear in the Game table more than once? Explain.
8. Which team won the maximum number of games?
9. Assuming that the league has ended and that complete participation information is available, how would you find the total number of points scored by a particular player?
10. How would you find the total number of points scored by a team?
11. Totaling all the points scored in each venue, which venue had the maximum points?
12. Do you see any redundant information in any of the tables? A piece of information would be considered redundant if it can be derived from some other piece of information that is already present. (Hint: Look carefully at the Games and the Participations tables.)
13. If you can answer a query (question) using information from just a single table, then we say that the query uses just a single table. For example, to find the number of teams we need to use just the Teams table. To find which players are from NJ, we would need only the Players table. Write down three questions in plain English, answering each of which would involve just a single table – any table from the above set.
14. Write down three queries, answering which would involve the use of two tables. For example, given the name of a team, finding the name of its coach would require us to use both the Teams and Coaches tables.
15. Write down one query, answering which would involve the use of three tables. For example, if we are given a game id and asked to find the names of the teams that played and the name of the coach of the winning team, then we would need to use the tables Games, Coaches and teams.