NBA Database Data Dictionary

Applications of Artificial Intelligence | Medina County Career Center

Database: `nba_5seasons.db`
Description: NBA statistics database containing player and team data across multiple seasons

Table: players

Description: Contains basic information about NBA players

Column Name	Data Type	Description	Example
`player_id`	Integer	Unique player identifier	2544
`full_name`	Text	Player's full name	"LeBron James"

Usage Notes:

- Use for getting player names

Table: teams

Description: Contains information about NBA teams including location and history

Column Name	Data Type	Description	Example
`team_id`	Integer	Unique team identifier	1610612739
`full_name`	Text	Team's full name	"Cleveland Cavaliers"
`abbreviation`	Text	Team abbreviation (3 letters)	"CLE"
`nickname`	Text	Team nickname	"Cavaliers"
`city`	Text	City where team is located	"Cleveland"
`state`	Text	State where team is located	"Ohio"
`year_founded`	Integer	Year the team was founded	1970

Usage Notes:

- Use for finding teams by location
- Use for analyzing team history
- Join with `player_season_stats` or `team_game_stats` using `team_id`

Table: player_season_stats

Description: Season-level statistics for players (one row per player per season)

Column Name	Data Type	Description	Example	Notes
season	Text	Season identifier	"2021-22"	Format: "YYYY-YY"
player_id	Integer	Player ID (foreign key)	2544	Links to players table
team_id	Integer	Team ID (foreign key)	1610612739	Links to teams table
gp	Integer	**Games played**	56	**Important for filtering!**
min	Float	Minutes per game	37.2	
pts	Float	**Points per game**	30.3	Key offensive statistic
reb	Float	**Rebounds per game**	8.2	Total rebounds
ast	Float	**Assists per game**	6.2	Key passing statistic
stl	Float	Steals per game	1.3	Defensive statistic
blk	Float	Blocks per game	1.1	Defensive statistic
tov	Float	Turnovers per game	3.5	
fg_pct	Float	Field goal percentage	50.4	Scale: 0-100
fg3_pct	Float	**3-point percentage**	35.9	Scale: 0-100
ft_pct	Float	Free throw percentage	75.6	Scale: 0-100

Example Filters:

```
WHERE gp >= 40          -- Players who played most of season
WHERE pts >= 20         -- Star scorers
WHERE fg3_pct >= 40     -- Elite shooters
```

Usage Notes:

- Each row represents one player's statistics for one season
- All per-game statistics are averages over the season

Table: team_game_stats

Description: Game-level statistics for teams (one row per team per game)

Column Name	Data Type	Description	Example	Notes
season	Text	Season identifier	"2021-22"	Format: "YYYY-YY"
game_id	Integer	Unique game identifier	22100001	

Column Name	Data Type	Description	Example	Notes
team_id	Integer	Team ID (foreign key)	1610612739	Links to teams table
game_date	Text	Date of game	"2021-10-19"	Format: "YYYY-MM-DD"
matchup	Text	Matchup description	"CLE vs. BOS" or "CLE @ BOS"	Home team listed second
wl	Text	**Win/Loss indicator**	"W" or "L"	
pts	Integer	**Points scored**	121	Total points for the game
fgm	Integer	Field goals made	44	
fga	Integer	Field goals attempted	88	
fg3m	Integer	3-pointers made	14	
fg3a	Integer	3-pointers attempted	35	
ftm	Integer	Free throws made	19	
fta	Integer	Free throws attempted	23	
oreb	Integer	Offensive rebounds	8	
dreb	Integer	Defensive rebounds	35	
reb	Integer	**Total rebounds**	43	offensive reb + defensive reb
ast	Integer	**Assists**	25	
stl	Integer	**Steals**	8	
blk	Integer	**Blocks**	6	
tov	Integer	**Turnovers**	14	
plus_minus	Integer	Point differential	+12	Team's score minus opponent's score

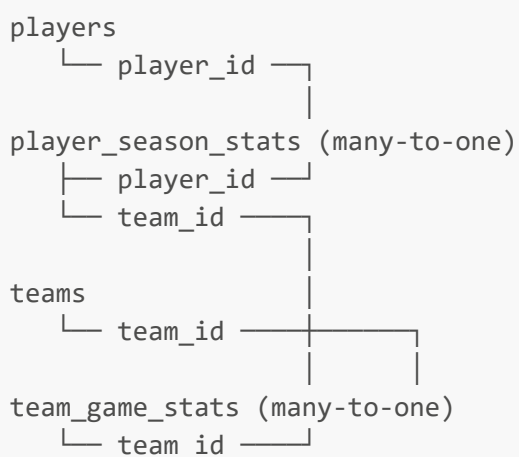**Example Filters:**

```
  WHERE wl = 'W'          -- Wins only
  WHERE wl = 'L'          -- Losses only
```

```
    WHERE pts >= 120         -- High-scoring games
    WHERE season = '2021-22' -- Specific season
```

Usage Notes:

- Each row represents one team's statistics for one game
- Each game appears twice in the table (once for each team)
- Use `game_id` to identify games
- Home/away indicated by matchup format ("vs." = home, "@" = away)

Table Relationships

```
players
    └── player_id ──┐
                    │
player_season_stats (many-to-one)
    ├── player_id ──┘
    └── team_id ───┐
                   │
teams              │
    └── team_id ───┼────────┐
                   │        │
team_game_stats (many-to-one)
    └── team_id ───┘
```

Key Relationships:

- One player can have multiple season records (one per season)
- One team can have multiple players in a season
- One team can have multiple game records
- Each game has two team records (one for each team)

Data Notes

Seasons Included:

- Database contains 5 seasons of data
- Season format: "YYYY-YY" (e.g., "2021-22")
- Always filter by season for accurate analysis

Statistical Ranges:

- Shooting percentages: 0-100 scale
- Per-game stats: Averages (can have decimals)
- Game stats: Integers (actual counts)