

WYDZIAŁ MATEMATYKI I NAUK INFORMACYJNYCH  
POLITECHNIKA WARSZAWSKA

---

## Hurtownia danych NBA

---

HURTOWNIE DANYCH I SYSTEMY BUSINESS INTELIGENCE  
KAMIEŃ MIŁOWY I - DOKUMENTACJA

Filip Szlingiert, Bartłomiej Wójcik

9 czerwca 2025

# Spis treści

|          |                                     |           |
|----------|-------------------------------------|-----------|
| <b>1</b> | <b>Wprowadzenie</b>                 | <b>3</b>  |
| 1.1      | Cel projektu . . . . .              | 3         |
| 1.2      | Dane . . . . .                      | 3         |
| 1.3      | Motywacje biznesowe . . . . .       | 3         |
| 1.4      | Wykorzystane technologie . . . . .  | 3         |
| <b>2</b> | <b>Hurtownia danych</b>             | <b>4</b>  |
| 2.1      | Diagram hurtowni . . . . .          | 4         |
| 2.2      | Tabele faktów . . . . .             | 5         |
| 2.2.1    | FactPlayerSeason . . . . .          | 5         |
| 2.2.2    | FactTeamSeason . . . . .            | 7         |
| 2.2.3    | FactTeamGameStats . . . . .         | 9         |
| 2.3      | Tabele wymiarów . . . . .           | 11        |
| 2.3.1    | dimDate . . . . .                   | 11        |
| 2.3.2    | dimSeason . . . . .                 | 11        |
| 2.3.3    | dimAwards . . . . .                 | 11        |
| 2.3.4    | dimTeam . . . . .                   | 12        |
| 2.3.5    | dimGame . . . . .                   | 12        |
| 2.3.6    | dimPlayerInfo . . . . .             | 13        |
| 2.4      | Pozostałe tabele . . . . .          | 13        |
| 2.4.1    | PlayerSeasonAwards . . . . .        | 13        |
| <b>3</b> | <b>Proces ETL</b>                   | <b>14</b> |
| 3.1      | Opis procesu . . . . .              | 14        |
| 3.2      | Control Flow . . . . .              | 15        |
| 3.3      | Data Flow . . . . .                 | 15        |
| 3.3.1    | dimSeason . . . . .                 | 15        |
| 3.3.2    | dimAwards . . . . .                 | 16        |
| 3.3.3    | dimTeam . . . . .                   | 16        |
| 3.3.4    | dimGame . . . . .                   | 17        |
| 3.3.5    | dimPlayerInfo . . . . .             | 17        |
| 3.3.6    | PlayerSeasonAwards . . . . .        | 18        |
| 3.3.7    | FactPlayerSeason . . . . .          | 18        |
| 3.3.8    | FactTeamSeason . . . . .            | 19        |
| 3.3.9    | FactTeamGameStats . . . . .         | 19        |
| 3.4      | Testy . . . . .                     | 20        |
| <b>4</b> | <b>Raport Business Intelligence</b> | <b>20</b> |
| <b>5</b> | <b>Podsumowanie</b>                 | <b>23</b> |

# 1 Wprowadzenie

## 1.1 Cel projektu

Celem projektu jest zaprojektowanie i wdrożenie hurtowni danych umożliwiającej analizę danych związanych z ligą NBA. Projekt obejmuje przygotowanie modelu danych (w oparciu o schemat galaktyki), stworzenie procesu ETL odpowiedzialnego za zasilanie hurtowni oraz przygotowanie interaktywnego raportu BI. Finalnym efektem jest system analityczny pozwalający użytkownikom biznesowym na szybki dostęp do informacji dotyczących meczów, zawodników, drużyn i statystyk.

## 1.2 Dane

Nasza hurtownia zawiera ogromne ilości statystyk zarówno dla zawodników, jak i dla drużyn występujących w NBA, na przestrzeni wielu lat (od 1981 roku). Aby zrealizować nasz projekt, potrzebne były wiarygodne i bogate źródła informacji, które udało nam się znaleźć na [Kaggle](#). Oba zbiory zawierały dane w postaci plików `csv`. Poniżej zamieszczamy źródła oraz nazwy plików, które wykorzystaliśmy przygotowując nasze rozwiązanie:

### 1. NBA Stats (1947-present)

- *Advanced.csv*
- *All-Star Selections.csv*
- *End of Season Teams.csv*
- *Player Award Shares.csv*
- *Player Career Info.csv*
- *Player Directory.csv*
- *Player Shooting.csv*
- *Player Totals.csv*
- *Team Abbrev.csv*

- *Team Summaries.csv*

- *Team Totals.csv*

### 2. NBA Database

- *common\_player\_info.csv*
- *game.csv*
- *game\_info.csv*
- *game\_summary.csv*
- *team.csv*
- *team\_history.csv*

## 1.3 Motywacje biznesowe

NBA to jedna z najpopularniejszych lig sportowych na świecie, a dane z nią związane mają ogromną wartość analityczną. Hurtownia danych pozwala na:

- monitorowanie wyników drużyn i zawodników w przekroju sezonów,
- analizę wydajności graczy w kontekście decyzji transferowych lub kontraktowych,
- wspieranie działań marketingowych poprzez identyfikację trendów i zależności w zachowaniach kibiców,
- szybki dostęp do kluczowych wskaźników wydajności (KPI) dla menedżerów, analityków i trenerów.

## 1.4 Wykorzystane technologie

Do stworzenia hurtowni danych wykorzystaliśmy środowisko **SQL Server Management Studio**, umożliwiło nam ono projektowanie struktury tabel, przygotowanie diagramów oraz przeprowadzenie testów.

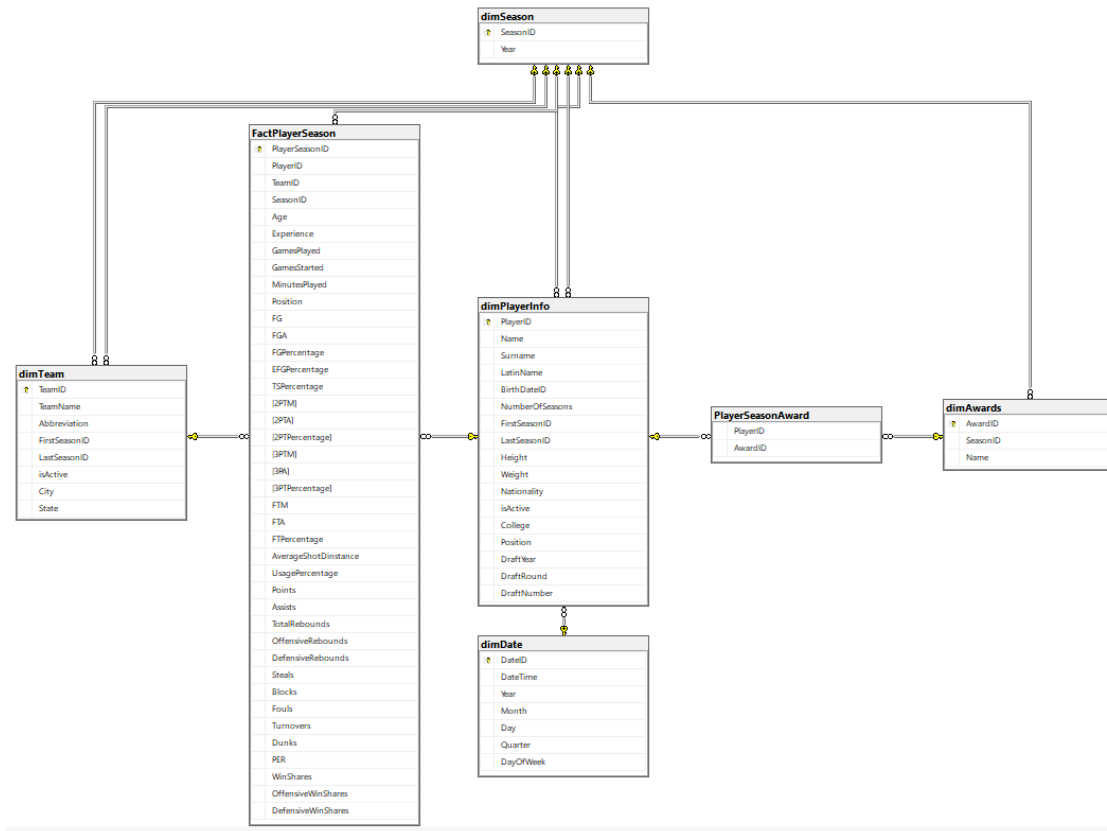
Proces ETL postanowiliśmy przeprowadzić w **SQL Server Integration Services (SSIS)**, czyli środowisku projektowym w **Visual Studio**. Narzędzie pozwoliło nam w efektywny sposób wczytać i przetworzyć dane z wielu plików oraz umieścić je w hurtowni, wykorzystanie parametrów projektu, pozwoliło nam na efektywną współpracę w grupie, ponieważ wyeliminowało potrzebę ciągłych zmian ścieżek plików w **Connection Manager**.

Raport Business Intelligence zrealizowaliśmy przy pomocy **Power BI**, który pozwala w łatwy sposób przeprowadzać analizy oraz przygotowywać interaktywne wizualizacje.



## 2.2 Tabele faktów

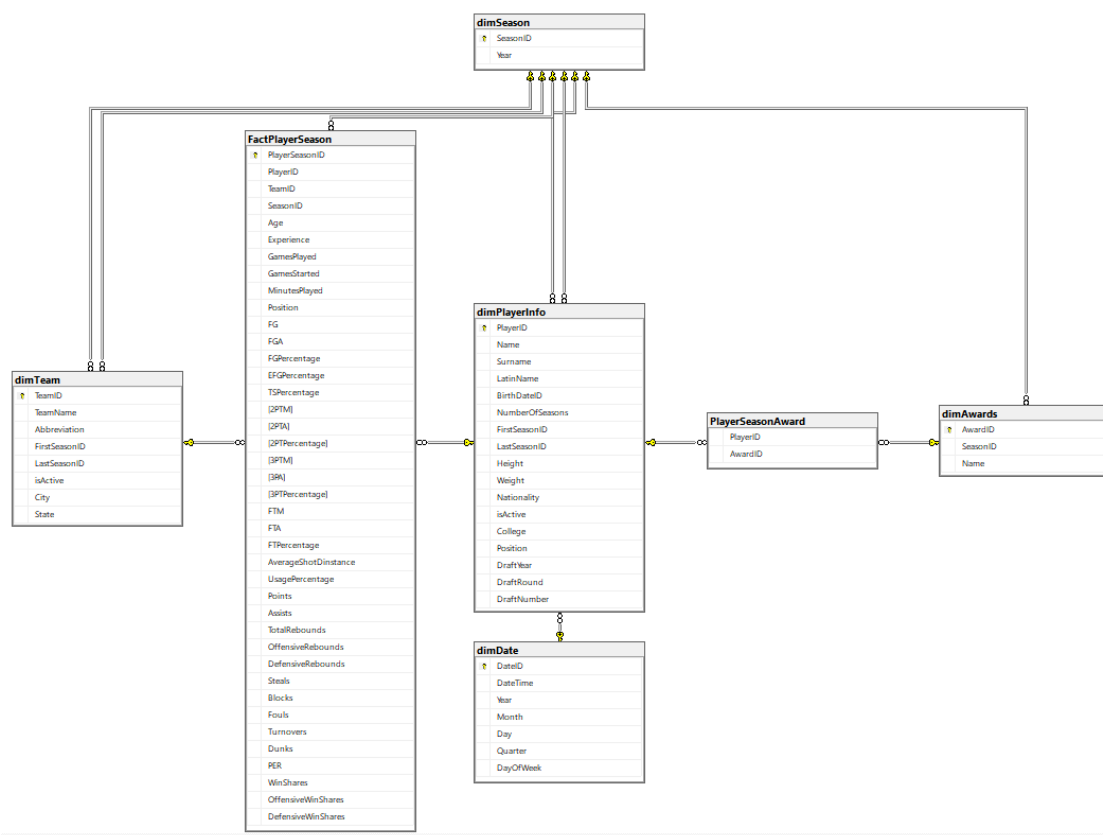
### 2.2.1 FactPlayerSeason



Rysunek 2: Diagram tabel powiązanych z FactPlayerSeason

| Struktura tabeli FactPlayerSeason |                      |  |
|-----------------------------------|----------------------|--|
| Nazwa kolumny                     | Typ danych           | Opis   |
| PlayerSeasonID                    | BIGINT IDENTITY(1,1) | Klucz podstawowy tabeli                                      |
| PlayerID                          | BIGINT               | Klucz obcy do tabeli dimPlayerInfo (identyfikator zawodnika) |
| TeamID                            | BIGINT               | Klucz obcy do tabeli dimTeam (identyfikator drużyny)         |
| SeasonID                          | BIGINT               | Klucz obcy do tabeli dimSeason (identyfikator sezonu)        |
| Age                               | TINYINT              | Wiek zawodnika w latach                                      |
| Experience                        | TINYINT              | Doświadczenie (liczba rozegranych sezonów)                   |
| GamesPlayed                       | SMALLINT             | Łączna liczba rozegranych meczów                             |
| GamesStarted                      | SMALLINT             | Liczba meczów w wyjściowym składzie                          |
| MinutesPlayed                     | SMALLINT             | Łączna liczba minut na boisku                                |
| Position                          | NVARCHAR(50)         | Pozycja na boisku (np. PG, SG, SF, PF, C)                    |
| FG                                | SMALLINT             | Celne rzuty z gry (Field Goals Made)                         |
| FGA                               | SMALLINT             | Rzuty z gry ogółem (Field Goals Attempted)                   |
| FGPercentage                      | DECIMAL(6,4)         | Skuteczność rzutów z gry (w procentach)                      |
| EFGPercentage                     | DECIMAL(6,4)         | Skuteczność rzutów z gry z uwzględnieniem rzutów za 3 punkty |
| TSPercentage                      | DECIMAL(6,4)         | Skuteczność rzutów ogółem (True Shooting Percentage)         |
| 2PTM                              | SMALLINT             | Celne rzuty za 2 punkty                                      |
| 2PTA                              | SMALLINT             | Rzuty za 2 punkty ogółem                                     |
| 2PTPercentage                     | DECIMAL(6,4)         | Skuteczność rzutów za 2 punkty                               |
| 3PTM                              | SMALLINT             | Celne rzuty za 3 punkty                                      |
| 3PA                               | SMALLINT             | Rzuty za 3 punkty ogółem                                     |
| 3PTPercentage                     | DECIMAL(6,4)         | Skuteczność rzutów za 3 punkty                               |
| FTM                               | SMALLINT             | Celne rzuty wolne  |
| FTA                               | SMALLINT             | Rzuty wolne ogółem   |
| FTPercentage                      | DECIMAL(6,4)         | Skuteczność rzutów wolnych                                   |
| AverageShotDinstance              | DECIMAL(6,3)         | Średnia odległość rzutów (w stopach)                         |
| UsagePercentage                   | DECIMAL(6,4)         | Procent akcji zespołu zakończonych przez zawodnika           |
| Points                            | SMALLINT             | Łączna liczba punktów  |
| Assists                           | SMALLINT             | Łączna liczba asyst  |
| TotalRebounds                     | SMALLINT             | Łączna liczba zbiórek  |
| OffensiveRebounds                 | SMALLINT             | Liczba zbiórek ofensywnych                                   |
| DefensiveRebounds                 | SMALLINT             | Liczba zbiórek defensywnych                                  |
| Steals                            | SMALLINT             | Liczba przechwyty  |
| Blocks                            | SMALLINT             | Liczba bloków  |
| Fouls                             | SMALLINT             | Łączna liczba fauli  |
| Turnovers                         | SMALLINT             | Łączna liczba strat  |
| Dunks                             | SMALLINT             | Liczba wsadów  |
| PER                               | DECIMAL(4,1)         | Wskaźnik efektywności zawodnika (Player Efficiency Rating)   |
| WinShares                         | DECIMAL(5,2)         | Łączna liczba zwycięstw przypisanych zawodnikowi             |
| OffensiveWinShares                | DECIMAL(5,2)         | Zwycięstwa przypisane ofensywie zawodnika                    |
| DefensiveWinShares                | DECIMAL(5,2)         | Zwycięstwa przypisane defensywie zawodnika                   |

## 2.2.2 FactTeamSeason



Rysunek 3: Diagram tabel powiązanych z FactTeamSeason

| Struktura tabeli FactTeamSeason |                      |   |
|---------------------------------|----------------------|---|
| Nazwa kolumny                   | Typ danych           | Opis  |
| TeamSeasonID                    | BIGINT IDENTITY(1,1) | Klucz podstawowy tabeli   |
| TeamID                          | BIGINT               | Klucz obcy do tabeli dimTeam (identyfikator drużyny)                  |
| SeasonID                        | BIGINT               | Klucz obcy do tabeli dimSeason (identyfikator sezonu)                 |
| AverageAge                      | DECIMAL(3,1)         | Średni wiek zawodników w drużynie                                     |
| MinutesPlayed                   | SMALLINT             | Łączna liczba minut rozegranych przez drużynę                         |
| FG                              | SMALLINT             | Celne rzuty z gry (Field Goals)                                       |
| FGA                             | SMALLINT             | Rzuty z gry ogółem (Field Goal Attempts)                              |
| FGPercentage                    | DECIMAL(5,3)         | Skuteczność rzutów z gry (w procentach)                               |
| FT                              | SMALLINT             | Celne rzuty wolne (Free Throws)                                       |
| FTA                             | SMALLINT             | Rzuty wolne ogółem (Free Throw Attempts)                              |
| FTPercentage                    | DECIMAL(5,3)         | Skuteczność rzutów wolnych (w procentach)                             |
| 3PTM                            | SMALLINT             | Celne rzuty za 3 punkty (3-Pointers Made)                             |
| 3PTA                            | SMALLINT             | Rzuty za 3 punkty ogółem (3-Point Attempts)                           |
| 3PTPercentage                   | DECIMAL(5,3)         | Skuteczność rzutów za 3 punkty (w procentach)                         |
| Points                          | SMALLINT             | Łączna liczba punktów   |
| Assists                         | SMALLINT             | Łączna liczba asyst   |
| TotalRebounds                   | SMALLINT             | Łączna liczba zbiórek   |
| OffensiveRebounds               | SMALLINT             | Liczba zbiórek ofensywnych  |
| DefensiveRebounds               | SMALLINT             | Liczba zbiórek defensywnych   |
| Steals                          | SMALLINT             | Liczba przechwyty   |
| Blocks                          | SMALLINT             | Liczba bloków   |
| Fouls                           | SMALLINT             | Łączna liczba fauli   |
| Turnovers                       | SMALLINT             | Łączna liczba strat   |
| Playoffs                        | NVARCHAR(3)          | Czy drużyna grała w play-off (Yes/No)                                 |
| Wins                            | SMALLINT             | Liczba zwycięstw w sezonie  |
| Losses                          | SMALLINT             | Liczba porażek w sezonie  |
| OffensiveRating                 | DECIMAL(5,1)         | Wskaźnik skuteczności ofensywnej                                      |
| DefensiveRating                 | DECIMAL(5,1)         | Wskaźnik skuteczności defensywnej                                     |
| NetRating                       | DECIMAL(5,1)         | Wskaźnik efektywności netto (różnica między ofensywnym a defensywnym) |
| Pace                            | DECIMAL(5,1)         | Tempo gry (liczba posiadów na 48 minut)                               |
| AverageAttendance               | INT                  | Średnia frekwencja na meczach domowych                                |
| GamesPlayed                     | SMALLINT             | Łączna liczba rozegranych meczów                                      |



### 2.2.3 FactTeamGameStats



Rysunek 4: Diagram tabel powiązanych z FactTeamGameStats

| Struktura tabeli FactTeamGameStats |                      |   |
|------------------------------------|----------------------|---|
| Nazwa kolumny                      | Typ danych           | Opis  |
| TeamGameStatsID                    | BIGINT IDENTITY(1,1) | Klucz podstawowy                                |
| TeamID                             | BIGINT               | Klucz obcy do dimTeam (identyfikator drużyny)   |
| GameID                             | BIGINT               | Klucz obcy do dimGame (identyfikator meczu)     |
| SeasonID                           | BIGINT               | Klucz obcy do dimSeason (identyfikator sezonu)  |
| PointsScored                       | SMALLINT             | Punkty zdobyte przez drużynę                    |
| PointsAllowed                      | SMALLINT             | Punkty stracone przez drużynę                   |
| FGM                                | SMALLINT             | Celne rzuty z gry (Field Goals Made)            |
| FGA                                | SMALLINT             | Rzuty z gry ogółem (Field Goals Attempted)      |
| FGPercentage                       | DECIMAL(5,3)         | Skuteczność rzutów z gry (w procentach)         |
| 3PTM                               | SMALLINT             | Celne rzuty za 3 punkty (3-Pointers Made)       |
| 3PTA                               | SMALLINT             | Rzuty za 3 punkty ogółem (3-Pointers Attempted) |
| 3PTPercentage                      | DECIMAL(5,3)         | Skuteczność rzutów za 3 punkty (w procentach)   |
| FTM                                | SMALLINT             | Celne rzuty wolne (Free Throws Made)            |
| FTA                                | SMALLINT             | Rzuty wolne ogółem (Free Throws Attempted)      |
| FTPercentage                       | DECIMAL(5,3)         | Skuteczność rzutów wolnych (w procentach)       |
| Rebounds                           | SMALLINT             | Łączna liczba zbiórek                           |
| Assists                            | SMALLINT             | Liczba asyst                                    |
| Steals                             | SMALLINT             | Liczba przechwytów                              |
| Turnovers                          | SMALLINT             | Liczba strat                                    |
| Blocks                             | SMALLINT             | Liczba bloków                                   |
| Fouls                              | SMALLINT             | Liczba fauli                                    |
| Result                             | NVARCHAR(4)          | Wynik meczu (Win/Loss)                          |
| GameType                           | NVARCHAR(50)         | Typ meczu (np. regular season, playoff)         |

## 2.3 Tabele wymiarów

### 2.3.1 dimDate

| Struktura tabeli dimDate |                      |  |
|--------------------------|----------------------|--|
| Nazwa kolumny            | Typ danych           | Opis   |
| DateID                   | BIGINT IDENTITY(1,1) | Klucz podstawowy tabeli, unikalny identyfikator daty |
| DateTime                 | DATETIME             | Pełna data   |
| Year                     | SMALLINT             | Rok  |
| Month                    | SMALLINT             | Miesiąc  |
| Day                      | SMALLINT             | Dzień  |
| Quarter                  | TINYINT              | Kwartał  |
| DayOfWeek                | TINYINT              | Dzień tygodnia                                       |

### 2.3.2 dimSeason

| Struktura tabeli dimSeason |                      |  |
|----------------------------|----------------------|--|
| Nazwa kolumny              | Typ danych           | Opis   |
| SeasonID                   | BIGINT IDENTITY(1,1) | Klucz podstawowy tabeli, unikalny identyfikator sezonu                                     |
| Year                       | SMALLINT             | Rok w którym odbywały się rozgrywki (np. dla sezonu 2024/2025 wartość tej kolumny to 2025) |

### 2.3.3 dimAwards

| Struktura tabeli dimAwards |                        |   |
|----------------------------|------------------------|---|
| Nazwa kolumny              | Typ danych             | Opis  |
| AwardID                    | SMALLINT IDENTITY(1,1) | Klucz podstawowy tabeli, unikalny identyfikator nagrody |
| SeasonID                   | BIGINT                 | Edycja nagrody, rok za który została ona przyznana      |
| Name                       | NVARCHAR(70)           | Nazwa nagrody   |

## 2.3.4 dimTeam

| Struktura tabeli dimTeam |                      |   |
|--------------------------|----------------------|---|
| Nazwa kolumny            | Typ danych           | Opis  |
| TeamID                   | BIGINT IDENTITY(1,1) | Klucz podstawowy tabeli, unikalny identyfikator drużyny |
| TeamName                 | NVARCHAR(50)         | Nazwa drużyny   |
| Abbreviation             | NVARCHAR(3)          | Skrót nazwy drużyny                                     |
| FirstSeasonID            | BIGINT               | ID pierwszego sezonu, w którym występowała dana drużyna |
| LastSeasonID             | BIGINT               | ID ostatniego sezonu, w którym występowała dana drużyna |
| isActive                 | NVARCHAR(3)          | Czy drużyna jest aktywna                                |
| City                     | NVARCHAR(50)         | Nazwa miasta, z którego pochodzi drużyna                |
| Stata                    | NVARCHAR(50)         | Nazwa stanu, z którego pochodzi drużyna                 |

## 2.3.5 dimGame

| Struktura tabeli dimGame |                      |   |
|--------------------------|----------------------|---|
| Nazwa kolumny            | Typ danych           | Opis  |
| GameID                   | BIGINT IDENTITY(1,1) | Klucz podstawowy tabeli, unikalny identyfikator gry |
| GameDateID               | INT                  | Klucz obcy do dimDate (data meczu)                  |
| SeasonID                 | BIGINT               | Klucz obcy do dimSeason (sezon)                     |
| HomeTeamID               | BIGINT               | Klucz obcy do dimTeam (gospodarz)                   |
| AwayTeamID               | BIGINT               | Klucz obcy do dimTeam (gość)                        |
| Attendance               | INT                  | Frekwencja  |
| GameType                 | NVARCHAR(50)         | Typ meczu (regular/playoff)                         |

### 2.3.6 dimPlayerInfo

| Struktura tabeli dimPlayerInfo |                      |   |
|--------------------------------|----------------------|---|
| Nazwa kolumny                  | Typ danych           | Opis  |
| PlayerID                       | BIGINT IDENTITY(1,1) | Klucz podstawowy tabeli, unikalny identyfikator zawodnika |
| Name                           | NVARCHAR(50)         | Imię zawodnika  |
| Surname                        | NVARCHAR(50)         | Nazwisko zawodnika  |
| LatinName                      | NVARCHAR(50)         | Imię i nazwisko w formacie łacińskim                      |
| BirthDateID                    | INT                  | Klucz obcy do tabeli dimDate (data urodzenia)             |
| NumberOfSeasons                | TINYINT              | Liczba sezonów w karierze                                 |
| FirstSeasonID                  | BIGINT               | Klucz obcy do dimSeason (pierwszy sezon)                  |
| LastSeasonID                   | BIGINT               | Klucz obcy do dimSeason (ostatni sezon)                   |
| Height                         | SMALLINT             | Wzrost w stopach  |
| Weight                         | SMALLINT             | Waga w funtach  |
| Nationality                    | NVARCHAR(50)         | Narodowość  |
| isActive                       | NVARCHAR(50)         | Czy zawodnik jest aktywny                                 |
| College                        | NVARCHAR(100)        | Uczelnia  |
| Position                       | NVARCHAR(50)         | Pozycja na boisku   |
| DraftYear                      | NVARCHAR(50)         | Rok draftu  |
| DraftRound                     | NVARCHAR(50)         | Runda draftu  |
| DraftNumber                    | NVARCHAR(50)         | Numer draftu  |

## 2.4 Pozostałe tabele

### 2.4.1 PlayerSeasonAwards

| Struktura tabeli PlayerSeasonAward |            |                             |
|------------------------------------|------------|-----------------------------|
| Nazwa kolumny                      | Typ danych | Opis                        |
| PlayerID                           | BIGINT     | Klucz obcy do dimPlayerInfo |
| AwardID                            | SMALLINT   | Klucz obcy do dimAwards     |

## 3 Proces ETL

### 3.1 Opis procesu

Proces *ETL* (Extract, Transform, Load) w naszym projekcie został zaprojektowany w sposób modularny i przejrzysty, co pozwala na łatwe zarządzanie przepływem danych oraz potencjalną rozbudowę hurtowni w przyszłości.

#### Extract – ekstrakcja danych

Dane źródłowe pochodzą z przygotowanych plików CSV z dwóch różnych zbiorów danych o NBA opisanych na początku raportu. Ze względu na różnice w nazewnictwach czy enkodowaniu dane musiały zostać odpowiednio sformatowane i ujednolicone, aby możliwe było ich dalsze przetwarzanie.

#### Transform – transformacja danych

W kroku transformacji przeprowadzono:

- czyszczenie danych i usuwanie duplikatów,
- ujednolicenie formatów (np. dat, nazw pozycji),
- Łączenie wielu tabel bez jednolitego ID
- filtracja danych (np. braliśmy tylko sezony od 1980, ponieważ wtedy wprowadzono linie rzutów za 3 punkty)
- tworzenie kluczy głównych i pochodnych tabel
- obliczenie wskaźników pochodnych (np. punkty na mecz, Offensive/Defensive Rating, Pace),
- wykrywanie i obsługę zmian historycznych (SCD2) w danych drużyn,
- mapowanie danych na strukturę tabel faktów i wymiarów zgodnie z przyjętym modelem hurtowni danych.

Transformacje zostały zaimplementowane w postaci pipeline'u w środowisku **SSIS** (SQL Server Integration Services), a parametryzacja ścieżek umożliwia łatwą adaptację do środowisk lokalnych.

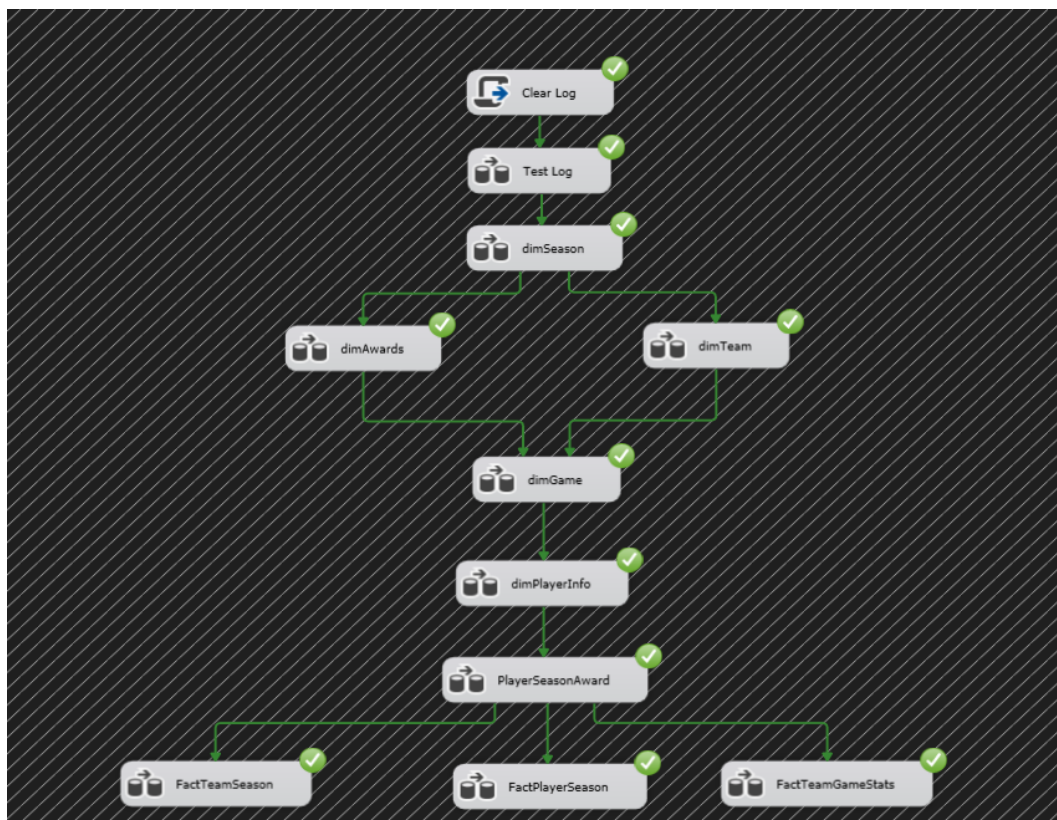
#### Load – załadunek danych

Załadunek danych odbywa się do wcześniej przygotowanego modelu hurtowni danych w **SQL Server Management Studio (SSMS)**. Proces ten jest podzielony na etapy:

1. Uruchomienie skryptu `schema.sql` w SSMS w celu utworzenia struktury bazy danych.
2. Wypełnienie tabeli `dimDate` przy pomocy skryptu `populate_date.sql`.
3. Uruchomienie pipeline'u SSIS w projekcie `prepare_warehouse.sln`, po wcześniejszym ustawieniu parametru `BaseDir` na lokalną ścieżkę repozytorium.

Dzięki zastosowaniu modularnego podejścia i wykorzystaniu SSIS, cały proces ETL może być łatwo kontrolowany, modyfikowany i ponownie wykonywany w razie potrzeby.

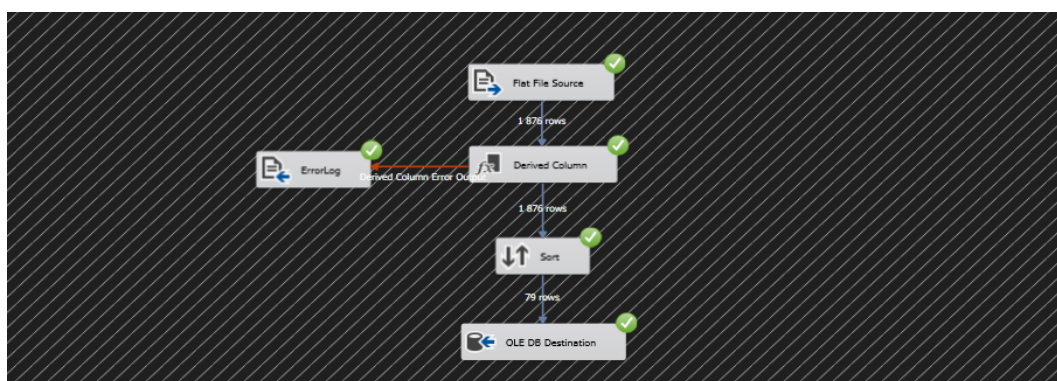
## 3.2 Control Flow



Rysunek 5: Schemat procesu ETL

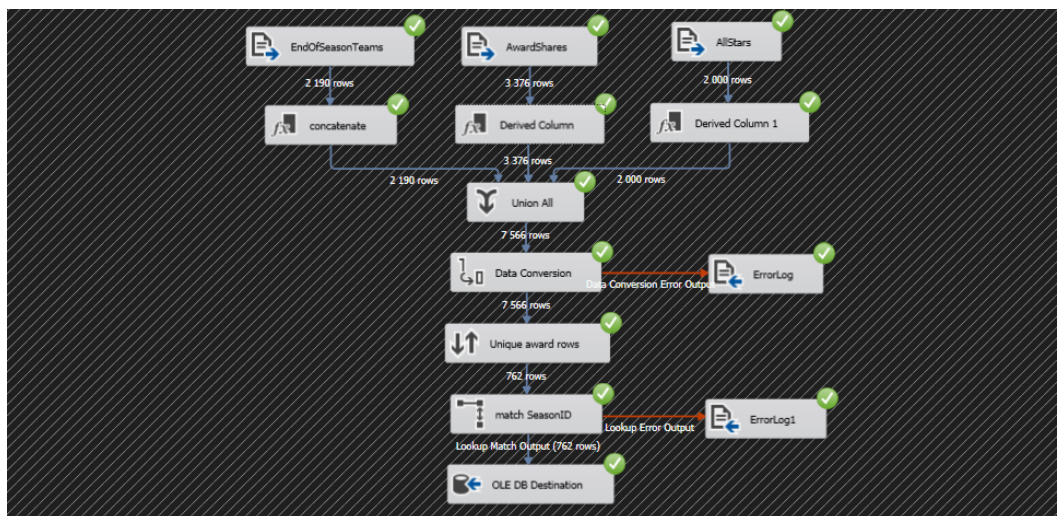
## 3.3 Data Flow

### 3.3.1 dimSeason



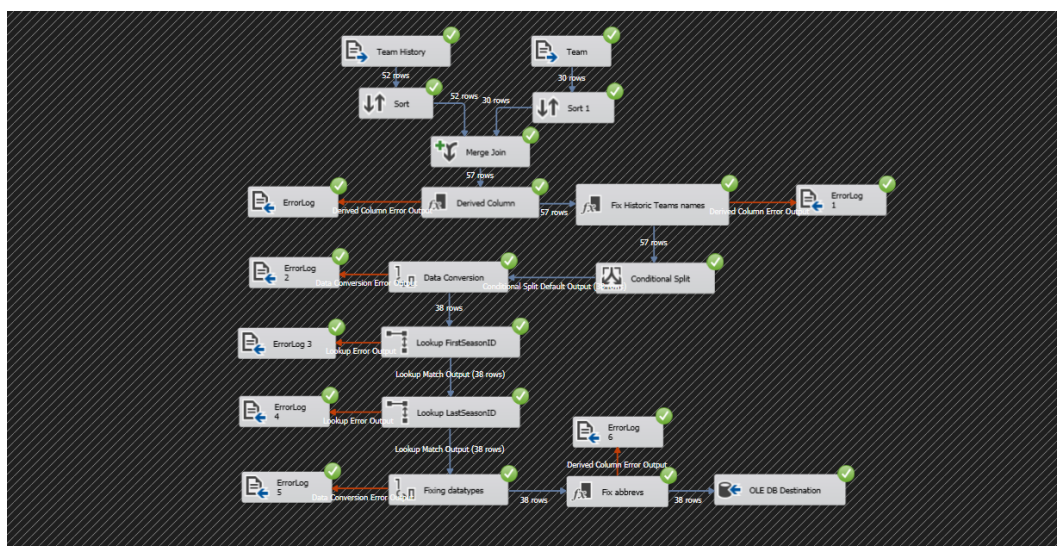
Rysunek 6: Schemat procesu ETL dla tabeli dimSeason

### 3.3.2 dimAwards



Rysunek 7: Schemat procesu ETL dla tabeli dimAwards

### 3.3.3 dimTeam

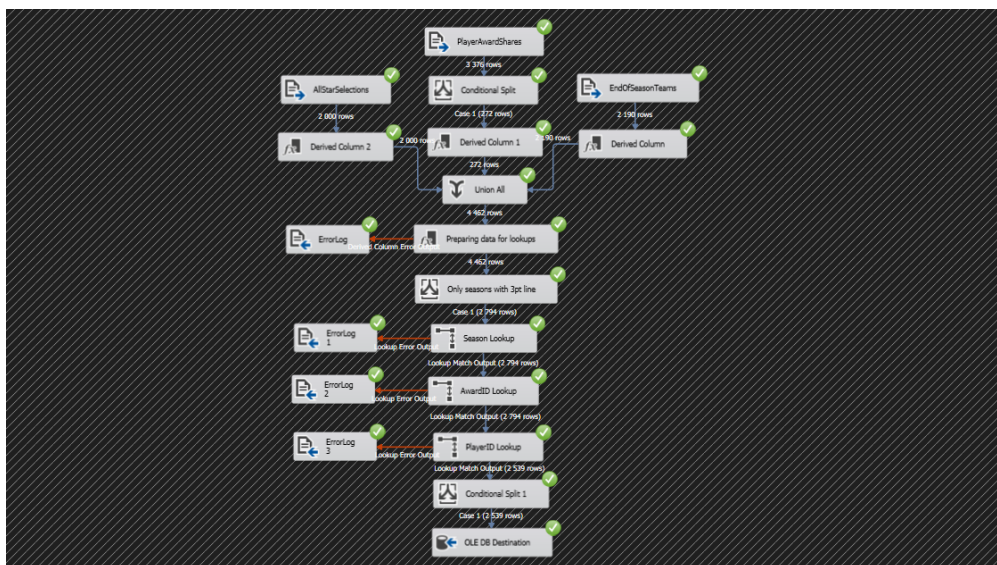


Rysunek 8: Schemat procesu ETL dla tabeli dimTeam



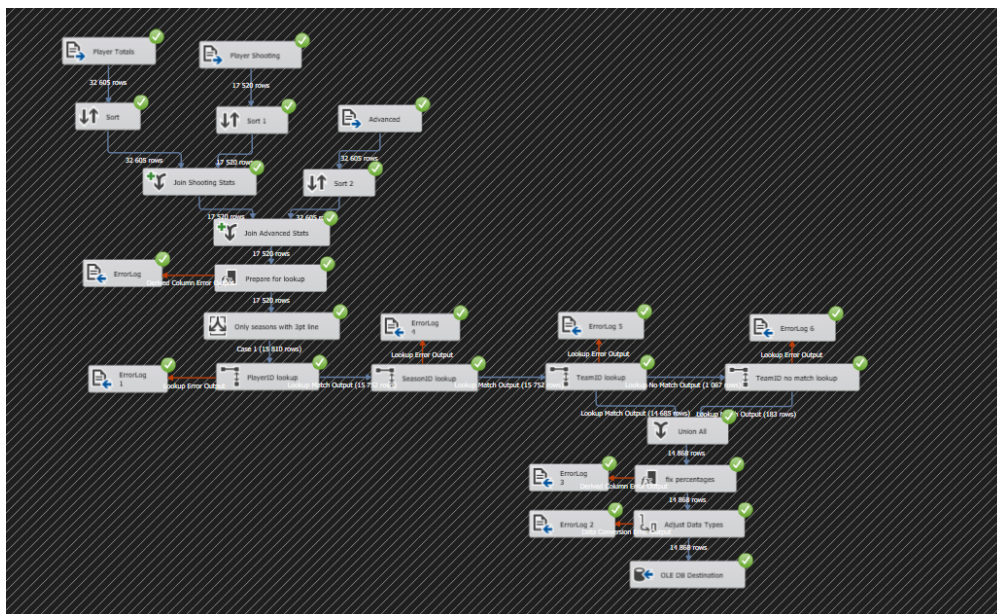


### 3.3.6 PlayerSeasonAwards



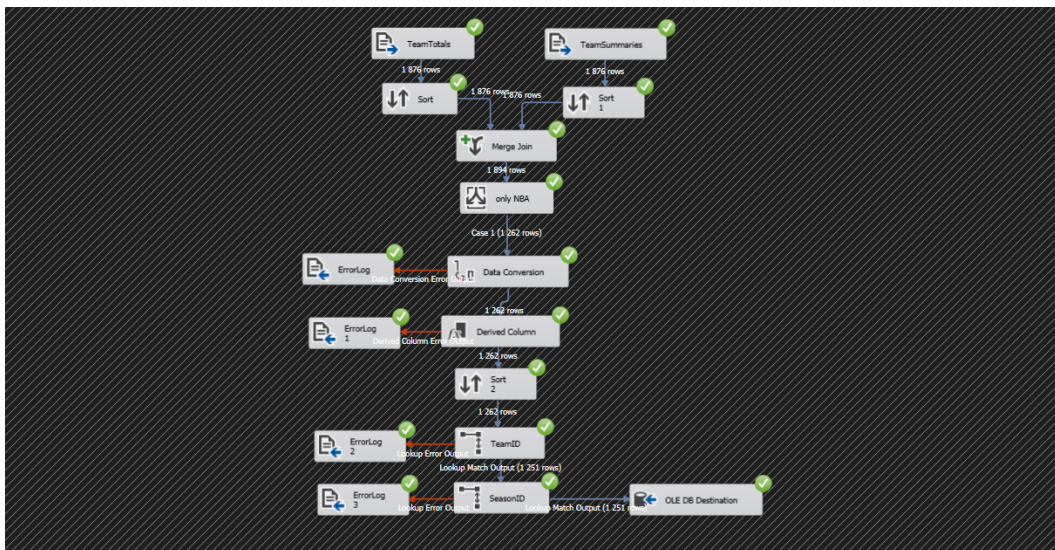
Rysunek 11: Schemat procesu ETL dla tabeli PlayerSeasonAwards

### 3.3.7 FactPlayerSeason



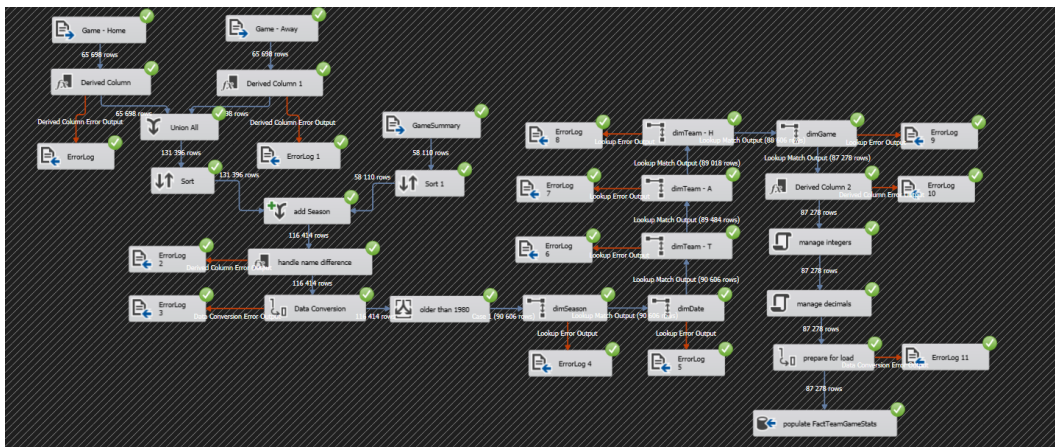
Rysunek 12: Schemat procesu ETL dla tabeli FactPlayerSeason

### 3.3.8 FactTeamSeason



Rysunek 13: Schemat procesu ETL dla tabeli FactTeamSeason

### 3.3.9 FactTeamGameStats



Rysunek 14: Schemat procesu ETL dla tabeli FactTeamGameStats

### 3.4 Testy

W ramach testów, wykonaliśmy wiele zapytań do naszej hurtowni, z których dwa zamieściliśmy poniżej, dzięki nim można też sprawdzić poprawność wizualizacji. Warto też zauważyć, że proces ETL przebiegł bezproblemowo i działanie wszystkich komponentów zakończyło się sukcesem, co widać na zdjęciach zamieszczonych w poprzednim akapicie.

```

use NBA
SELECT
  dpi.Name AS first_name,
  dpi.Surname AS last_name,
  COUNT(*) AS all_star_appearances
FROM
  PlayerSeasonAward p
JOIN
  dimAwards a ON p.awardID = a.awardID
JOIN
  dimPlayerInfo dpi ON p.playerID = dpi.playerID
WHERE
  a.name = 'All Star'
GROUP BY
  p.playerID, dpi.Name, dpi.Surname
ORDER BY
  all_star_appearances DESC;

```

|    | first_name | last_name | all_star_appearances |
|----|------------|-----------|----------------------|
| 1  | LeBron     | James     | 20                   |
| 2  | Kobe       | Bryant    | 18                   |
| 3  | Shaquille  | O'Neal    | 15                   |
| 4  | Tim        | Duncan    | 15                   |
| 5  | Kevin      | Garnett   | 15                   |
| 6  | Kevin      | Durant    | 14                   |
| 7  | Karl       | Malone    | 14                   |
| 8  | Dirk       | Nowitzki  | 14                   |
| 9  | Michael    | Jordan    | 14                   |
| 10 | Dwyane     | Wade      | 13                   |

Rysunek 15: Zapytanie zliczające liczbę udziałów zawodników w meczu gwiazd NBA.

```

SELECT
  p.LatinName,
  s.Year,
  fps.GamesPlayed,
  fps.MinutesPlayed,
  fps.Points,
  fps.Assists,
  fps.TotalRebounds,
  fps.PER,
  fps.FGPercentage,
  fps.[3PTPercentage],
  fps.WinShares
FROM FactPlayerSeason fps
JOIN dimPlayerInfo p ON fps.PlayerID = p.PlayerID
join dimSeason s on fps.SeasonID=s.SeasonID
WHERE p.LatinName='Zach LaVine'
ORDER BY fps.SeasonID;

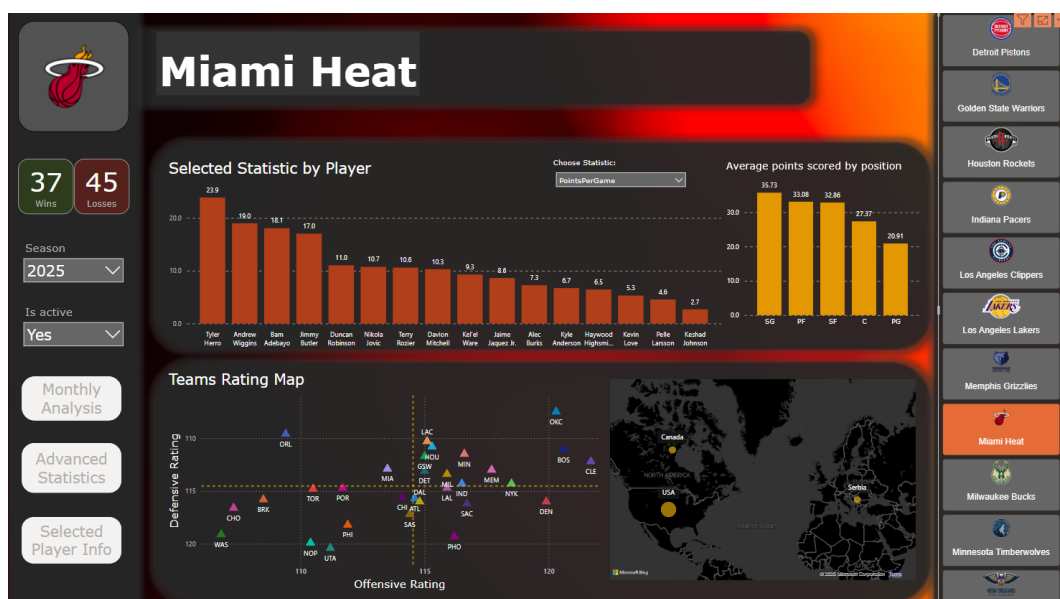
```

|    | LatinName   | Year | GamesPlayed | MinutesPlayed | Points | Assists | TotalRebounds | PER  | FGPercentage | 3PTPercentage | WinShares |
|----|-------------|------|-------------|---------------|--------|---------|---------------|------|--------------|---------------|-----------|
| 1  | Zach LaVine | 2015 | 77          | 1902          | 778    | 276     | 214           | 11.3 | 0.4219       | 0.3410        | -0.70     |
| 2  | Zach LaVine | 2016 | 82          | 2294          | 1150   | 251     | 228           | 14.3 | 0.4520       | 0.3890        | 2.60      |
| 3  | Zach LaVine | 2017 | 47          | 1749          | 889    | 139     | 160           | 14.6 | 0.4590       | 0.3870        | 3.00      |
| 4  | Zach LaVine | 2018 | 24          | 656           | 401    | 72      | 94            | 14.6 | 0.3830       | 0.3410        | 0.30      |
| 5  | Zach LaVine | 2019 | 63          | 2171          | 1492   | 283     | 294           | 18.6 | 0.4670       | 0.3740        | 2.80      |
| 6  | Zach LaVine | 2020 | 60          | 2085          | 1530   | 254     | 289           | 19.3 | 0.4500       | 0.3800        | 4.00      |
| 7  | Zach LaVine | 2021 | 58          | 2034          | 1591   | 282     | 289           | 21.5 | 0.5070       | 0.4189        | 5.90      |
| 8  | Zach LaVine | 2022 | 67          | 2328          | 1635   | 303     | 308           | 20.0 | 0.4759       | 0.3890        | 5.80      |
| 9  | Zach LaVine | 2023 | 77          | 2768          | 1913   | 327     | 345           | 19.0 | 0.4849       | 0.3750        | 7.10      |
| 10 | Zach LaVine | 2024 | 25          | 872           | 487    | 98      | 129           | 15.1 | 0.4520       | 0.3489        | 1.50      |
| 11 | Zach LaVine | 2025 | 42          | 1432          | 1007   | 189     | 202           | 18.3 | 0.5110       | 0.4460        | 3.00      |
| 12 | Zach LaVine | 2025 | 32          | 1170          | 717    | 121     | 113           | 15.9 | 0.5110       | 0.4460        | 2.10      |

Rysunek 16: Zapytanie z bazy zbierające statystyki zawodnika na przestrzeni sezonów.

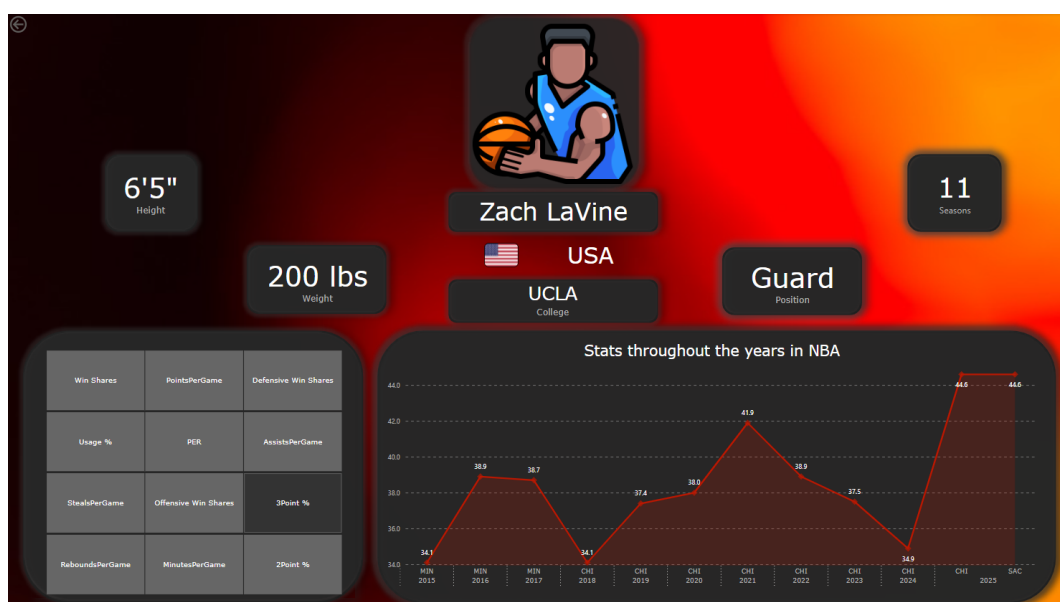
## 4 Raport Business Intelligence

Nasza główna strona raportu pozwala użytkownikowi wybrać dowolną drużynę zawartą w naszej hurtowni zarówno spośród tych aktywnych obecnie, jak i historycznych. Jest również możliwość wyboru sezonu wyświetlanych na tej stronie danych czy statystyki na podstawie której zostanie stworzony wykres zawodników danej drużyny. Jeśli chodzi o wizualizacje to mamy tutaj statystyki zawodników wybranej drużyny, ukazuje się jedna z opcji wybrana przez użytkownika (np. punkty na mecz, minuty na mecz). Widzimy tutaj również wykres narodowości obecnych w danej drużynie czy słupkowy wykres dystrybucji punktów w drużynie na poszczególne pozycje. Jednym z najciekawszych do analizy wykresów, można zaliczyć scatter w lewym dolnym rogu. Na osi X mamy OffensiveRating czyli metrykę mówiącą ile zespół średnio zdobywa punktów na 100 posiadanych piłki, natomiast na osi Y mamy DefensiveRating, który mówi ile zespół średnio traci punktów na 100 posiadanych. Drużyna może zobaczyć tutaj swoje dokonania zarówno na atakowanej, jak i bronionej części parkietu w odniesieniu do innych zespołów NBA.



Rysunek 17: Przykładowy widok naszej strony głównej raportu.

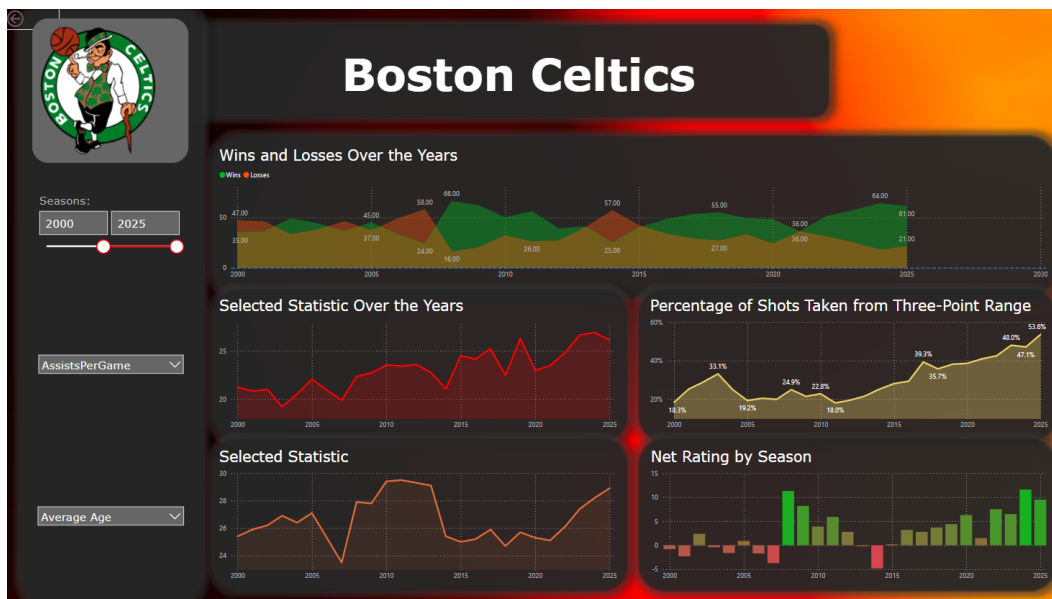
Ze strony głównej możemy przejść do strony ze statystykami zawodnika poprzez zaznaczenie go na stronie a następnie kliknięcie odpowiedniego przycisku. Na tej stronie zebrane są podstawowe informacje o zawodniku, takie jak wiek, wzrost, waga, narodowość czy główna pozycja. Mamy również możliwość prześledzić wszelkie statystyki indywidualne zawodnika oraz ich zmiany na przestrzeni kolejnych sezonów.



Rysunek 18: Przykładowy widok strony zawodnika.



Z panelu głównego użytkownik ma również możliwość przejścia do zakładki *TeamPage*, która zawiera bardziej szczegółowe statystyki dla danej drużyny. Możliwe jest dostosowanie przedziału sezonów na, którym wizualizowane są dane. Użytkownik może przeanalizować trendy i zmiany w organizacji na przełomie lat na podstawie statystyk takich jak średni wiek, zapelnienie stadionu i wiele innych.



Rysunek 19: Przykładowy widok strony drużyny.

Do analizy statystyk drużynowych posłużyć może również *DetailedTeamPage*. Na tej stronie raportu mamy do czynienia z różnymi statystykami uśrednionymi w zależności od miesiąca czy dnia tygodnia danego sezonu.



Rysunek 20: Przykładowy widok strony drużyny.

Ostatnią ze stron naszego raportu jest *PlayerFinderPage*, który byłby świetnym narzędziem zarówno dla analityków, skautów, jak i zagorzałych fanów. Użytkownik może na niej filtrować zawodników i ich statystyki na podstawie sezonu, pozycji, jak i podstawowych statystyk indywidualnych. W tabeli ze statystykami przefiltrowanych zawodników, również znajduje się możliwość sortowania po wszystkich zawartych w niej informacjach.

The screenshot shows the PlayerFinderPage interface. At the top, there are filters for Season (2025), Points per Game (8.00 to 20.00), Assists per Game (2.00 to 10.00), Rebounds per Game (3.00 to 18.00), and Minutes per Game (0.00 to 24.00). Below the filters is a 'Go to Player Page' button. The main table displays player statistics for the 2025 season. The table has columns for Player Name, Age, Games Played, Experience, and various statistical categories (PPG, APG, AST/TOV, RPG, ORB, DRB, STL, BLK, TOV, FG%, 2PT%, 3PT%, FT%, EFG%, TS%, USG%, WinShares, DefWinShares, OffWinShares). The table lists 10 players: Scotty Pippen Jr., Ausar Thompson, Jaime Jaquez Jr., Cole Anthony, Khris Middleton, Zach Collins, Malcolm Brogdon, Trendon Watford, and Khris Middleton.

| LatinName         | LogoURL | Age | GamesPlayed | Experience | PPG   | APG  | AST/TOV | RPG  | ORb  | DRb  | STL  | BLK  | TOV  | FG%   | 2PT%  | 3PT%  | FT%   | EFG%  | TS%   | USG% | WinShares | DefWinShares | OffWinShares |
|-------------------|---------|-----|-------------|------------|-------|------|---------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|------|-----------|--------------|--------------|
| Scotty Pippen Jr. |         | 24  | 79          | 3          | 9.86  | 4.39 | 2.51    | 3.28 | 0.82 | 2.46 | 1.30 | 0.38 | 1.75 | 48.0% | 52.8% | 39.7% | 71.3% | 55.3% | 58.2% | 19.3 | 4.7       | 2.3          | 2.4          |
| Ausar Thompson    |         | 22  | 59          | 2          | 10.07 | 2.27 | 1.68    | 5.14 | 1.88 | 3.25 | 1.66 | 0.68 | 1.36 | 53.5% | 57.2% | 22.4% | 64.1% | 54.7% | 56.8% | 19.1 | 3.8       | 2.4          | 1.5          |
| Jaime Jaquez Jr.  |         | 23  | 66          | 2          | 8.64  | 2.52 | 1.69    | 4.39 | 1.27 | 3.12 | 0.92 | 0.21 | 1.48 | 46.1% | 51.3% | 31.1% | 75.4% | 50.1% | 54.3% | 20.2 | 2.3       | 1.8          | 0.5          |
| Cole Anthony      |         | 24  | 67          | 5          | 9.40  | 2.85 | 1.85    | 3.04 | 0.69 | 2.36 | 0.67 | 0.48 | 1.54 | 42.4% | 47.1% | 35.3% | 82.3% | 49.4% | 53.1% | 24.6 | 2.0       | 2.0          | 0.0          |
| Khris Middleton   |         | 33  | 23          | 13         | 12.61 | 4.43 | 2.76    | 3.70 | 0.22 | 3.48 | 0.74 | 0.22 | 1.61 | 51.2% | 58.4% | 40.7% | 84.8% | 59.5% | 62.7% | 22.0 | 1.7       | 0.6          | 1.1          |
| Zach Collins      |         | 27  | 28          | 7          | 8.57  | 2.11 | 1.59    | 6.68 | 2.29 | 4.39 | 0.54 | 0.50 | 1.32 | 54.1% | 67.3% | 30.0% | 88.4% | 59.4% | 63.5% | 17.1 | 1.7       | 0.6          | 1.1          |
| Malcolm Brogdon   |         | 32  | 24          | 9          | 12.71 | 4.08 | 2.58    | 3.79 | 0.88 | 2.92 | 0.54 | 0.21 | 1.58 | 43.3% | 47.8% | 28.6% | 88.0% | 46.7% | 54.4% | 23.7 | 0.7       | 0.1          | 0.6          |
| Trendon Watford   |         | 24  | 44          | 4          | 10.23 | 2.61 | 1.39    | 3.64 | 0.84 | 2.80 | 0.61 | 0.27 | 1.89 | 46.9% | 51.5% | 33.0% | 76.2% | 51.0% | 55.7% | 23.3 | 0.7       | 0.8          | 0.0          |
| Khris Middleton   |         | 33  | 14          | 13         | 10.71 | 3.43 | 2.53    | 3.71 | 0.57 | 3.14 | 1.29 | 0.21 | 1.36 | 41.3% | 49.4% | 27.7% | 86.8% | 46.4% | 52.6% | 22.0 | 0.4       | 0.2          | 0.1          |

Rysunek 21: Przykładowy widok strony do wyszukiwania zawodników.

## 5 Podsumowanie

Stworzona przez nas hurtownia danych NBA, wraz z towarzyszącym raportem Business Intelligence, stanowi solidną podstawę do prowadzenia zarówno klasycznych, jak i bardziej zaawansowanych analiz statystycznych związanych z zawodnikami oraz drużynami ligi NBA. Dzięki zastosowaniu wielu tabel faktów, obejmujących statystyki sezonowe i meczowe, oraz odpowiednio zaprojektowanym wymiarom, możliwe jest śledzenie długofalowych trendów i zmian zachodzących w lidze.

Zaprojektowany raport BI oferuje przejrzysty i interaktywny interfejs, który umożliwia użytkownikom szybkie uzyskanie potrzebnych informacji, zarówno na poziomie całych drużyn, jak i indywidualnych zawodników. Intuicyjna nawigacja, dynamiczne filtry oraz bogaty zestaw wizualizacji czynią go przydatnym narzędziem analitycznym dla menedżerów, trenerów, skautów, jak również fanów koszykówki. Nasze rozwiązanie jest elastyczne i może być w łatwy sposób rozwijane o nowe źródła danych, miary czy perspektywy analityczne. Stanowi ono kompleksowe środowisko do analizy danych NBA, wspierające podejmowanie decyzji oraz eksplorację ciekawych zależności i trendów w świecie profesjonalnej koszykówki.