1. **Tytuł raportu**

„Analiza danych o piłce nożnej: Ocena wpływu cech zawodników na wyniki drużynowe”

1. **Wstęp - Motywacja, cele**

* **Motywacja**: Piłka nożna to jedna z najpopularniejszych dyscyplin sportowych na świecie. Analiza danych pozwala lepiej zrozumieć, jakie cechy piłkarzy przyczyniają się do sukcesu drużyny. Może to być użyteczne dla menedżerów sportowych i trenerów.
* **Cele**:
  + Przeanalizowanie wpływu różnych zmiennych, np. pracowitości w obronie piłkarzy na tracone bramki i pracowitości w ataku piłkarzy na strzelane bramki oraz wpływ ocen ogólnych z gry FIFA na wyniki meczów.
  + Opracowanie modelu przewidującego, które drużyny mają większe szanse na sukces w lidze.
  + Wyciągnięcie praktycznych wniosków dla analizy piłkarskiej.

1. **Opis danych - Struktura zbioru**

Dane dostępne w zbiorze zawierają szczegółowe informacje na temat piłkarzy, drużyn, lig, wyników meczów oraz atrybutów związanych z grą. Zbiór został podzielony na kilka kluczowych tabel, które umożliwiają analizę zarówno na poziomie zawodników, jak i całych drużyn czy lig.

**Tabele obejmują:**

**1. Country**

* Opis: Informacje o krajach, w których odbywają się rozgrywki piłkarskie.
* Kluczowe zmienne:
  + id: Unikalny identyfikator kraju.
  + name: Nazwa kraju.
* Zastosowanie: Łączenie z tabelami League w celu przypisania lig do konkretnych krajów.

**2. League**

* Opis: Informacje o ligach piłkarskich.
* Kluczowe zmienne:
  + id: Unikalny identyfikator ligi.
  + country\_id: Powiązanie z tabelą Country w celu określenia, w którym kraju znajduje się liga.
  + name: Nazwa ligi (np. “Spain LIGA BBVA”, ”Poland Ekstraklasa”, “England Premier League”).
* Zastosowanie: Segmentacja danych na poziomie lig, co pozwala na analizę różnic między rozgrywkami.

**3. Match**

* Opis: Dane o meczach piłkarskich, w tym wyniki i statystyki.
* Kluczowe zmienne:
  + id: Unikalny identyfikator meczu.
  + league\_id: Powiązanie z tabelą League.
  + season: Sezon rozgrywek.
  + date: Data rozegrania meczu.
  + home\_team\_api\_id, away\_team\_api\_id: Identyfikatory drużyn gospodarza i gości (łączone z tabelą Team).
  + home\_team\_goal, away\_team\_goal: Liczba zdobytych bramek przez gospodarzy i gości.
* Zastosowanie: Analiza wyników, przewidywanie wyników meczów, ocena formy drużyn.

**4. Player**

* Opis: Szczegółowe informacje o zawodnikach.
* Kluczowe zmienne:
  + id: Unikalny identyfikator zawodnika.
  + player\_api\_id: Klucz API zawodnika (łączone z tabelą Match).
  + player\_name: Imię i nazwisko zawodnika.
  + Player\_fifa\_api\_id: Klucz API zawodnika z gry FIFA (łączone z tabelą Player\_Attributes).
  + birthday: Data urodzenia zawodnika.
  + height: Wzrost zawodnika (w cm).
  + weight: Waga zawodnika (w funtach).
* Zastosowanie: Analiza cech demograficznych zawodników, np. wpływ wieku lub wzrostu na wyniki.

**5. Player\_Attributes**

* Opis: Atrybuty techniczne, fizyczne i mentalne zawodników, zebrane w czasie.
* Kluczowe zmienne:
  + id: Unikalny identyfikator rekordu.
  + player\_fifa\_api\_id: Powiązanie z tabelą Player.
  + date: Data aktualizacji danych.
  + Cechy zawodnika:
    - Techniczne: np. crossing (dośrodkowania), finishing (wykończenie akcji), dribbling (drybling).
    - Fizyczne: np. acceleration (przyspieszenie), stamina (wytrzymałość), strength (siła).
    - Mentalne: np. aggression (agresja), vision (wizja gry), positioning (ustawianie się).
* Zastosowanie: Analiza wpływu umiejętności zawodników na wyniki drużynowe, ocena wartości zawodnika.

**6. Team**

* Opis: Informacje o drużynach biorących udział w rozgrywkach.
* Kluczowe zmienne:
  + id: Unikalny identyfikator drużyny.
  + team\_api\_id: Klucz API drużyny (łączone z tabelą Match).
  + team\_fifa\_api\_id: Klucz API drużyny z gry FIFA (łączone z tabelą Team\_Attributes).
  + team\_long\_name, team\_short\_name: Pełna i skrócona nazwa drużyny.
* Zastosowanie: Łączenie danych meczowych z drużynami, analiza różnic między drużynami.

**7. Team\_Attributes**

* Opis: Atrybuty drużyn, związane z ich stylem gry i strategią.
* Kluczowe zmienne:
  + id: Unikalny identyfikator rekordu.
  + team\_fifa\_api\_id: Powiązanie z tabelą Team.
  + date: Data aktualizacji danych.
  + Cechy drużyny:
    - Styl gry: np. buildUpPlaySpeed (szybkość budowania akcji), chanceCreationPassing (tworzenie okazji poprzez podania).
    - Strategia: np. defenceAggression (agresywność obrony), defenceWidth (szerokość gry w obronie).
* Zastosowanie: Analiza wpływu strategii drużynowej na wyniki, budowa modeli przewidujących wyniki na podstawie stylu gry.

1. Opis procesu przygotowania danych

Etapy:

1. Eksploracja: Wczytanie danych, identyfikacja braków, analiza podstawowych statystyk.
2. Czyszczenie danych: Usunięcie brakujących danych.