

# Dokumentacja projektu Power BI – Analiza danych sprzedaży gier wideo

---

## 1. Wprowadzenie do projektu

Celem projektu jest analiza danych sprzedaży gier wideo w różnych regionach świata w celu zidentyfikowania trendów rynkowych, najbardziej popularnych gatunków i wydawców oraz regionalnych preferencji. Projekt ma na celu stworzenie interaktywnego raportu, który ułatwia podejmowanie decyzji biznesowych i zrozumienie dynamiki rynku.

---

## 2. Źródła danych

Dane zostały zaimportowane z pliku CSV zawierającego informacje o sprzedaży gier wideo. Kluczowe kolumny to:

- **Name:** Nazwa gry.
- **Platform:** Platforma (np. PS4, Xbox One, PC).
- **Year:** Rok wydania.
- **Genre:** Gatunek gry.
- **Publisher:** Wydawca.
- **Regional\_Sales:** Sprzedaż w różnych regionach (USA, Europa, Japonia).

Global\_Sales zostało wyliczone na podstawie sumy sprzedaży regionalnej, aby upewnić się, że wartości są spójne. Skrypt do pobrania danych jest dostępny na stronie [GitHub](#). Skrypt został napisany w Pythonie z wykorzystaniem biblioteki BeautifulSoup. Ten zestaw danych zawiera listę gier wideo, które sprzedały się w liczbie większej niż 100 000 kopii. Dane zostały zebrane za pomocą web-scrapingu strony vgchartz.com. Zestaw danych zawiera **16 598 rekordów**. Dwa rekordy zostały usunięte z powodu niekompletnych informacji.

---

## 3. Model danych

Model danych opiera się na jednej tabeli faktów oraz jednej tabeli wymiarów:

- **Tabela Faktów:** Dane o sprzedaży gier.
  - **Tabela Wymiarów:** Regiony, którą utworzono na potrzeby analizy regionalnej.
  - **Relacje:** Kluczowe relacje łączą dane o grach z wymiarem regionów.
- 

## 4. Proces analizy

### 1. Przygotowanie danych:

- Zaimportowano dane z pliku CSV do Power BI.
- W Power Query wykonano następujące kroki:

- Usunięto zbędne kolumny.
- Braki danych opisano jako 1700, a przy wyświetlaniu zastosowano funkcję zamieniającą wartości na „No Value”.

## 2. Kalkulacje DAX:

- Stworzono miary, takie jak:
  - Całkowita sprzedaż: `SUM(Sales[Global_Sales])`
  - Sprzedaż w regionach: `SUM(Sales[Regional_Sales])`
- Wykorzystano funkcje, takie jak SWITCH i IF, do tworzenia dynamicznych miar.

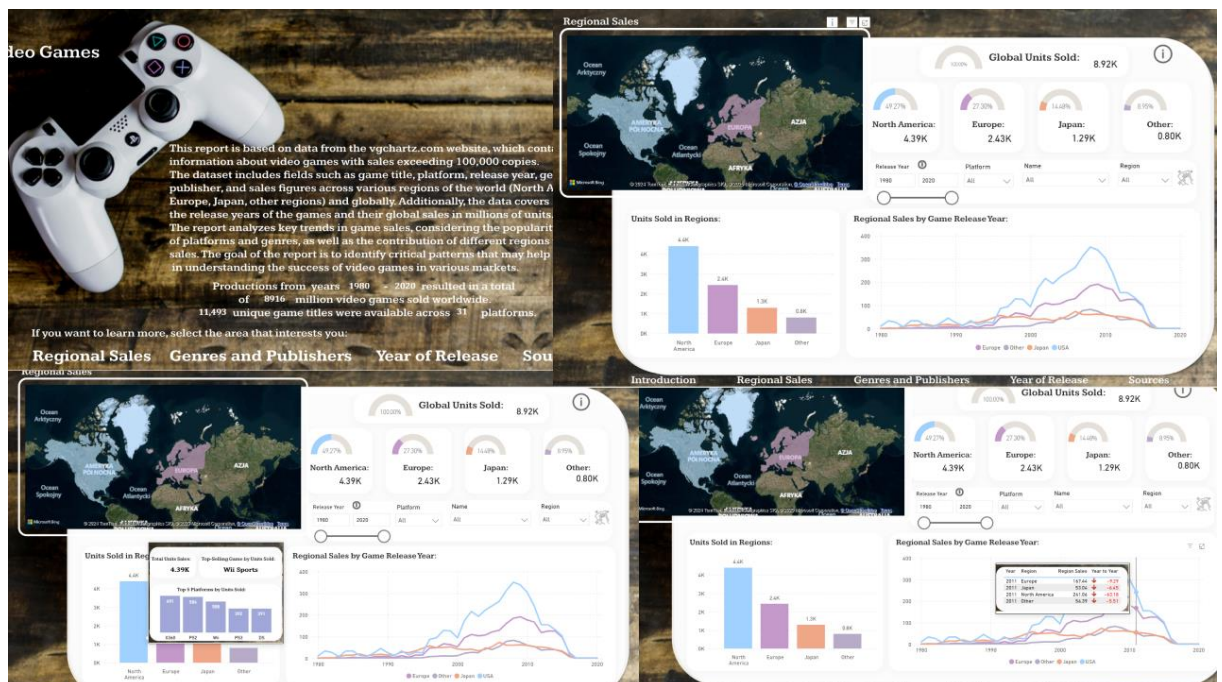
## 3. Wizualizacje:

- Stworzono interaktywne wykresy i fragmentatory dla różnych wymiarów (np. region, platforma, wydawca).

# 5. Wizualizacje

Raport składa się z kilku stron, każda skupiona na innym aspekcie analizy:

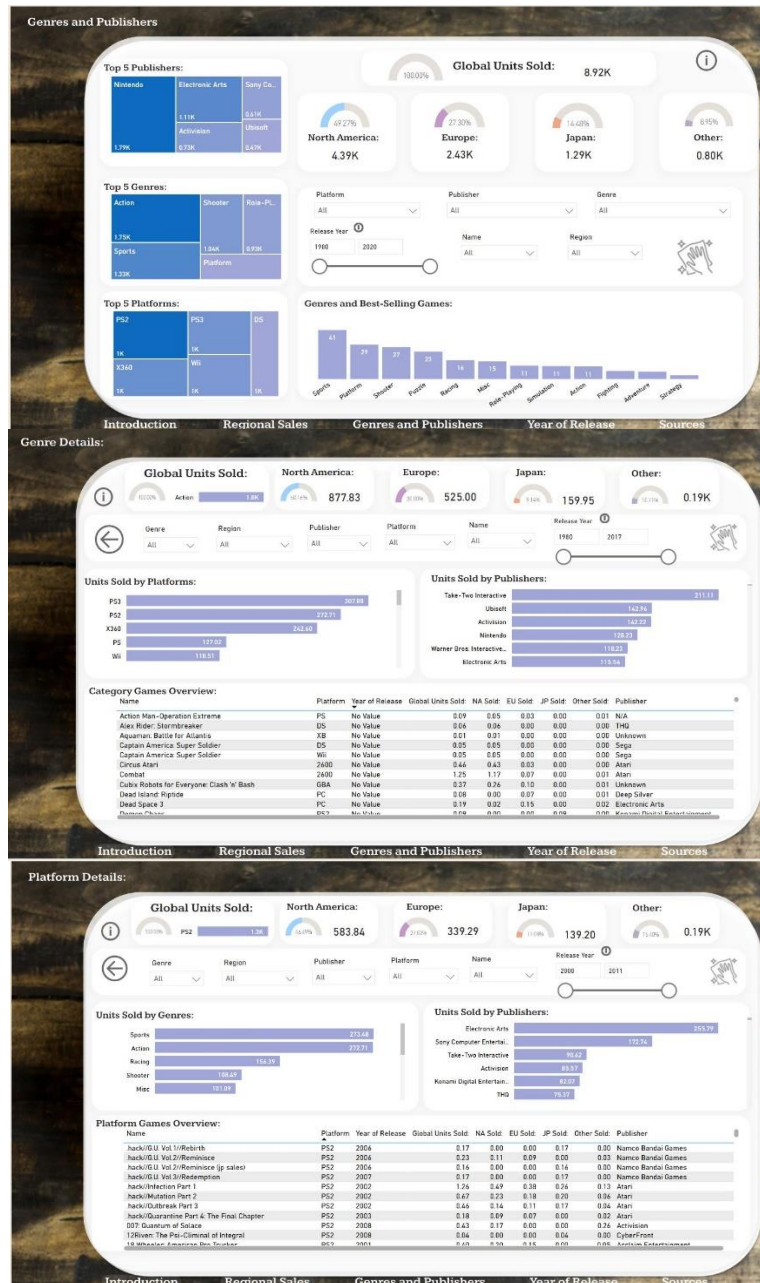
## 1. Sprzedaż globalna:



- Wykresy słupkowe pokazujące najwyższą sprzedaż w regionach, największych wydawców i najbardziej dochodowe gry.
- Wykres liniowy publikacji w czasie i ich sprzedaż całkowitą w regionach, wraz z informacjami odnośnie relacji wartości sprzedaży do roku poprzedniego.
- Dodatkowo wytępują karty informacyjne o udziale sprzedaży w poszczególnych regionach

- Fragmentatory umożliwiające filtrowanie według regionów, platform, nazwy gry i czasu .

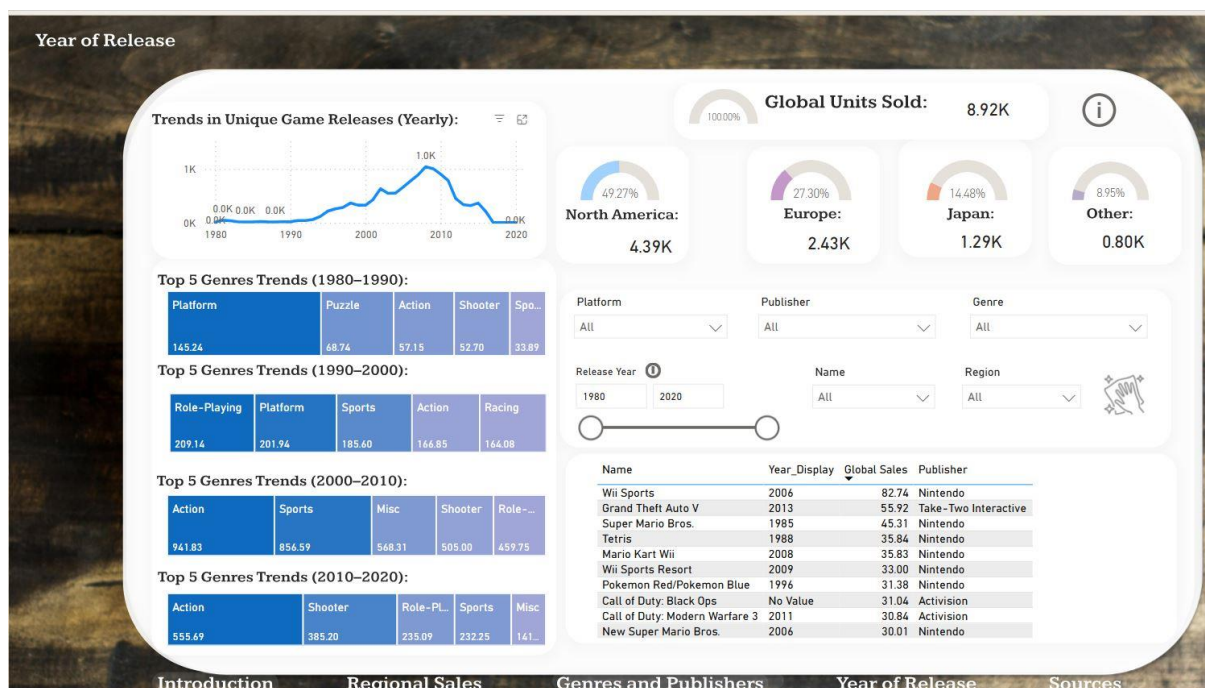
## 2. Popularność gatunków i Producentów:



- Treemap przedstawiający 5 najpopularniejszych gatunków, producentów oraz platform.
- Funkcja drill-through umożliwia wyświetlenie pełnej listy gier w zakresie oraz popularności gatunku wśród wydawców i platform.
- Dodatkowo występują karty informacyjne o udziale sprzedaży w poszczególnych regionach

- Fragmentatory umożliwiające filtrowanie według regionów, platform, nazwy gry i czasu .

### 3. Analiza Lat wydania gier:



- Treemap przedstawiające zmiany w preferencji gatunkowej gier na przestrzeni lat 1990/2000/2010/2020.
- Lista najpopularniejszych 10 gier.
- Fragmentatory umożliwiające filtrowanie według regionów, platform, nazwy gry i czasu .

## 6. Wnioski

### Popularność gatunków (1990–2000):

Gatunki takie jak **Action**, **Sports** i **Role-Playing** dominowały w latach 1990–2000, przyciągając największe zainteresowanie graczy. Gry z tych kategorii generowały najwyższą sprzedaż, co wskazuje na ich silną pozycję na rynku w tym okresie.

### Liczba unikalnych gier na przestrzeni lat:

Najwięcej nowych gier wydano w latach [największy rok, np. 2000 lub inny wierzchołek], co pokazuje dynamiczny rozwój rynku w tym czasie. Wzrost liczby gier był szczególnie widoczny w drugiej połowie lat 90., co można przypisać postępowi technologicznemu i rosnącej liczbie platform.

### Regionalne różnice w preferencjach:

- **Japonia:** Dominacja gier z gatunku **Role-Playing** wskazuje na silne zainteresowanie narracyjnymi tytułami.
- **Ameryka Północna:** Gry akcji i sportowe były najbardziej popularne, odzwierciedlając dynamiczny styl rozgrywki preferowany w tym regionie.
- **Europa:** Gatunki takie jak **Action** i **Racing** zyskały znaczną popularność, sugerując różnorodność w preferencjach.

### Wpływ starszych gier:

Gry wydane przed rokiem 2000, takie jak **Super Mario Bros.** czy **Pokémon Red/Blue**, nadal odgrywają istotną rolę w globalnej sprzedaży. Wskazuje to na trwałość klasyków, które pomimo wieku wciąż generują dochody.

### Zróżnicowanie rynku po 2000 roku:

Po roku 2000 obserwujemy większą fragmentację rynku – wzrosła liczba gier, ale sprzedaż rozkłada się bardziej równomiernie między gatunki. To wskazuje na większą konkurencję i bardziej zróżnicowane preferencje graczy.

### Preferencje gatunkowe zmieniają się w czasie:

W latach 90. dominowały prostsze gatunki, jak platformowe i sportowe, natomiast w późniejszych latach rosła popularność bardziej złożonych gatunków, takich jak **Shooter** czy **Role-Playing**, odzwierciedlając zmieniające się oczekiwania graczy.

---

## 7. Zakres i ograniczenia

- Dane obejmują tylko wybrane lata i mogą nie odzwierciedlać pełnego obrazu rynku gier.
- Braki danych w kolumnie „Year” mogą wpływać na dokładność analizy trendów.
- Wyniki analizy zależą od poprawności zaimportowanych danych.

---

## Załączniki:

1. **Przykładowe funkcje w Dax**
2. **Plik projektu Power BI (.pbix):** gamesProject.pbix
3. **Plik danych (.csv):** vgsales.csv

## Przykładowe funkcje w Dax:

Sprzedaż Globalna:

```
1 Global Sales = COALESCE(SUM('vg-sales'[Sales]),0)
```

Ranking według sprzedaży:

```
1 Rank by Sales = RANKX(All(vg-sales[Name]),  
2 | [Global Sales],  
3 | ,DESC,Dense  
4 )
```

Ranking według sprzedaży w platformach:

```
1 Rank by Platform = RANKX(All(vg-sales[Platform]),  
2 | [Global Sales],  
3 | ,DESC,Dense  
4 )
```

Top 5 platform:

```
1 Top 5 Platform =  
2 IF(  
3 | [Rank by Platform] <= 5,  
4 | SUM(vg-sales[Sales]),  
5 | BLANK()  
6 )
```

Top w sprzedaży gry:

---

```
1 Top Selling Game =  
2     Var maxResale = MAXX(  
3         'vgsales',  
4         vgsales[Sales]  
5     )  
6     RETURN  
7     CONCATENATEX(  
8     FILTER(  
9         ALLSELECTED('vgsales'),  
10        'vgsales'[Sales] = maxResale  
11    ),  
12    'vgsales'[Name],  
13    ", "  
14 )  
15  
16
```

---