

## I. Dokumentacja Biznesowa

### 1. Cele projektu

Raport dostarcza kompleksowej wiedzy na temat bazy **2240 Klientów** oraz ich zachowań zakupowych. Narzędzie umożliwia identyfikację najbardziej dochodowych segmentów oraz ocenę zwrotu z inwestycji w kampanie marketingowe.

### 2. Kluczowe Wnioski

- **Profil Klienta:** Największą aktywność zakupową wykazują **Seniorzy** oraz osoby w **Wieku średnim**. Grupy te dominują we wszystkich kanałach sprzedaży.
- **Struktura Sprzedaży:** Sprzedaż dominuje **wino (35,8%)** oraz **produkty mięsne (25,8%)**. Pozostałe kategorie stanowią uzupełnienie oferty.
- **Preferencje Kanałów:** Mimo dostępności zakupów online i katalogowych, **sklepy stacjonarne** pozostają głównym miejscem kontaktu z klientem dla każdego segmentu wiekowego.
- **Marketing:** Kampanie 3 i 4 odnotowały najwyższy wskaźnik odpowiedzi, co sugeruje ich wysoką efektywność w porównaniu do pozostałych działań.

### 3. Rekomendacje

- Utrzymanie szerokiego asortymentu win w sklepach stacjonarnych ze względu na preferencje kluczowej grupy (Seniorów).
- Przeanalizowanie przyczyn niskiej aktywności grupy "Młodzi dorośli" w celu dopasowania przyszłych kampanii.

---

## II. Dokumentacja Techniczna

### 1. Architektura Danych

#### I: Architektura danych i Modelowanie Relacyjne

Projekt opiera się na **zoptymalizowanym modelu relacyjnym**, poddany procesowi **normalizacji**. W bazie danych zdefiniowano klucze, wyeliminowano nadmiarowość danych i zapewniono pełną **integralność referencyjną** pomiędzy profilem klienta a historią zdarzeń transakcyjnych.

#### II: Warstwa analityczna

Surowe dane zostały przetransformowane w gotową do analizy warstwę widoków logicznych (`v_customer_golden_record`, `v_customer_engagement`)

## 2. Przepływ Danych (ETL)

**Cel:** przekształcenie surowych, nieustrukturyzowanych danych z formatu CSV na zoptymalizowany model relacyjny.

### I: Utworzenie tabeli tymczasowej

Szybki import danych bez walidacji typów. Wszystkie kolumny są typu `TEXT`, co zapobiega błędom podczas ładowania plików źródłowych. (`tmp_marketing`)

### II: Transformacja i Ładowanie

- Konwersja z `TEXT` na `INT`, `VARCHAR` oraz `DATE`.
- **Standaryzacja dat:** Użycie funkcji `TO_DATE(dt_customer, 'DD-MM-YYYY')` w celu ujednolicenia formatu czasowego.
- Dodanie segmentacji klientów przy użyciu instrukcji `CASE`: segmentacja wieku i dochodu

### III: PowerBi i wizualizacja danych

#### 1. Transformacja danych

- **Lokalizacja (Locale):** Ustawiono język polski dla interfejsu i formatowania dat.
- **Tłumaczenia:** Zastosowano niestandardowe mapowanie nazw grup wiekowych oraz poziomów dochodu z języka angielskiego na polski.

#### 2. Implementacja DAX

- **Hierarchia Czasu:** Nazwy miesięcy zostały zmapowane na język polski i posortowane według liczby porządkowej miesiąca (1-12) w celu uniknięcia sortowania alfabetycznego.
- **SWITCH:** użyto funkcji w celu zmiany niektórych slicerach z polskiego na poplsi

#### 3. Wizualizacje i Interakcje

- **Slicery (Fragmentatory):** Wykorzystano fragmentatory typu "Lista" dla grup wiekowych i dochodowych oraz "Kafelki" dla selekcji miesięcy.
- **Spójność danych:** Wszystkie obiekty wizualne posiadają włączone filtrowanie wzajemne.
- **KPI Cards:** Wyświetlają sumę klientów oraz transakcji promocyjnych.

**Źródło danych:**

<https://www.kaggle.com/datasets/imakash3011/customer-personality-analysis>