Analiza danych sprzedażowych sklepu Cosmic Electronics

Dokumentacja dla dewelopera

Raport przygotowany przez **Monikę Kaczan** ([m.kaczan2@student.uw.edu.pl](mailto:m.kaczan2@student.uw.edu.pl)) w ramach projektu zaliczeniowego na zajęcia **Dashboardy i narzędzia Power Excel**

# OPIS RAPORTU

Opis raportu zawarty jest w samym raporcie, w pierwszym arkuszu Opis raportu.

# BAZA DANYCH

Ze względu na bardzo niską dostępność danych, które można byłoby wykorzystać w tym projekcie, zdecydowałam się wygenerować losowo własne dane opierając się częściowo na danych rzeczywistych. Wykorzystałam do tego operacje w Excelu oraz w R. Na tym etapie nie dokumentowałam jednak poszczególnych kroków, ponieważ było ich bardzo dużo i przeprowadzałam je metodą prób i błędów dostosowując dane tak, aby jak najbardziej przypominały realne.

Ostatecznie utworzona przeze mnie baza składa się z 3 tabel:

|  |  |  |
| --- | --- | --- |
| **Tabela** | **Zmienna** | **Dodatkowe informacje** |
| **Klienci**  **(6000 rekordów)** | **IDKlienta** | Liczba naturalna od 1 do 6000 |
| **ImięKlienta** | Pobrane ze strony Otwarte Dane[[1]](#footnote-1) i przypisywane losowo, część wykorzystana również z bazy Klienci z zajęć |
| **NazwiskoKlienta** |
| **Wiek** | Losowany |
| **Kraj** | Wygenerowany losowo kraj wśród wybranych 12 krajów Unii Europejskiej |
| **Produkty (1000 rekordów)** | **IDProduktu** | Liczba naturalna od 1 do 1000 |
| **Cena sprzedaży** | Losowane ze zbioru Electronic Products and Pricing Datafinity[[2]](#footnote-2) |
| **Producent** |
| **Produkt** | Losowana kategoria produktu |
| **Cena zakupu** | Losowana jako ułamek od 0,4 do 0,98 ceny sprzedaży. |
| **Transakcje (15000 rekordów)** | **IDTransakcji** | Liczba naturalna od 1 do 15000 |
| **IDKlienta** | j.w. |
| **Data** | Losowana z dat ze zbioru Electronic Products and Pricing Datafinity + 4 lata[[3]](#footnote-3) |
| **Liczba sztuk** | Losowana z od 1 do 10 z malejącym prawdopodobieństwem dla kolejnych sztuk |

# IMPORT DANYCH – POWER QUERY

Na początku wszystkie tabele zostały wczytane ze skoroszytów Excela i przekształcane za pomocą Power Query, a następnie załadowane do modelu danych.

Ponieważ dane pochodzą z plików, a nie ze stron internetowych, nie jest konieczne ich odświeżanie.

Klienci

Ujednolicono zapis imion i nazwisk (poprzednio niektóre z nich były np. pisane wielkimi literami). Usunięto zbędne odstępny przekształceniem Przytnij. Na podstawie ostatniej litery imienia określono **Płeć** klienta, a na podstawie wieku przypisano **GrupaWiekowa** i dodano je jako osobne kolumny. Sprawdzono formatowanie kolumn, zmieniono nazwy.

Produkty

Zamieniono kropki na przecinki przy cenach. Zmieniono formatowanie kolumn. Po odpowiedniej analizie ujednolicono nazwy producentów. Sprawdzono formatowanie kolumn, zmieniono nazwy.

Transakcje

Po sprawdzeniu poprawności danych w tabeli, formatowania itp. zmieniono jedynie nazwy.

# MODEL DANYCH – POWER PIWOT

W narzędzi PowerPivot utworzono odpowiednie relacje między wczytanymi tabelami posługując się kluczami IDKlienta oraz IDProduktu.

Usunięto zbędne spacje w kolumnach Imię i Nazwisko klienta, które powstały przy generowaniu danych.

# ARKUSZ Analiza Finansowa

W tym arkuszu zostały przedstawione głównie dane finansowe, związane z dochodami, kosztami i zyskiem.

Aby policzyć przychód (wartość zamówienia), koszt i zysk z pojedynczej transakcji utworzyłam nowe kolumny za pomocą Power Pivot. Dodatkowo przyjęłam, że jeżeli wartość zamówienia przekracza 100 Euro, koszt sklepu wzrasta o 5 Euro za pokrycie darmowej dostawy. Aby móc ładniej zwizualizować dochody i koszty na wykresie, utworzyłam również kolumnę z kosztem wyrażonym jako liczba ujemna.

Aby użytkownik mógł ograniczyć analizę do interesującego go okresu, krajów i kategorii, wstawiłam odpowiednie fragmentatory.

Na Wykresie 1. zaprezentowałam rozłożenie przychodów i kosztów w poszczególnych miesiącach, natomiast na Wykresie 2. rozłożenie zysków. Na Wykresie 2. dodałam również liniową linię trendu, która wskazuje, czy w analizowanym okresie zyski miały tendencję rosnącą czy malejącą. Obok wykresów zamieściłam również mapę wykonaną w Power Map obrazującą łączny zysk i łączną liczbę transakcji dla klientów z poszczególnych państw.

# ARKUSZ Produkty i producenci

W tym arkuszu zawarłam analizę zainteresowania poszczególnymi kategoriami produktów i producentami.

Z boku umieściłam fragmentatory. Wykres na górze po prawej prezentuje 10 producentów, z których sklep osiągnął najwyższy zysk. Natomiast wykresy na dole odnoszą się popularności poszczególnych kategorii produktów pod względem udziału w łącznym zysku i liczby sprzedanych sztuk.

# ARKUSZ Klienci – łącznie

W tym arkuszu zamieściłam krótkie podsumowanie liczby klientów w podziale na płeć i wiek, a także (utworzoną w Power Map) mapę ich rozmieszczenia. Istnieje również możliwość sprawdzenia średniej liczby transakcji i średniej łącznej wartości transakcji przypadającej na jednego klienta danej grupy.

# ARKUSZ Klienci - szczegóły

Ten arkusz zawiera szczegółowe informacje o każdym kliencie w formie tabeli przestawnej. Listę klientów można również filtrować za pomocą fragmentatorów. Ma to na celu analizę zachowań zakupowych poszczególnych klientów.

W tabeli Klienci za pomocą Power Pivota dodałam nową kolumnę zawierającą Imię, Nazwisko i ID klienta (samo Imię i Nazwisko nie było unikalne). Ponadto utworzyłam następujące miary:

* Data pierwszego zamówienia, na podstawie której później wyliczyłam Liczbę miesięcy, które od niego minęły zakładając obecną datę 01.01.2021
* Wartość złożonych zamówień łącznie, czyli sumę wszystkich zamówień. Na tej podstawie stworzyłam wskaźnik KPI, który pokazuje, czy dana osoba zamówiła w sumie więcej niż średnia wybranych osób na liście (oznaczona na zielono), czy mniej niż 50% średniej (oznaczona na czerwono).

W tabeli przestawnej zamieściłam również Liczbę zamówionych produktów łącznie oraz Średnią Wartość Pojedynczej Transakcji.

# Arkusz pomocniczy

W tym arkuszu znajdują się jedynie tabele przestawne wykorzystane przy tworzeniu raportu, dlatego jest on domyślnie ukryty.

1. Np. https://dane.gov.pl/pl/dataset/1681,nazwiska-osob-zyjacych-wystepujace-w-rejestrze-pesel [↑](#footnote-ref-1)
2. https://data.world/datafiniti/electronic-products-and-pricing-data [↑](#footnote-ref-2)
3. https://data.world/datafiniti/electronic-products-and-pricing-data [↑](#footnote-ref-3)