Raport z analizy danych zakupów online

# 1. Wprowadzenie

W dobie rosnącego znaczenia e-commerce analiza danych zakupowych staje się kluczowym elementem zrozumienia potrzeb klientów i optymalizacji procesów sprzedażowych. Niniejszy raport stanowi wynik eksploracyjnej analizy danych transakcyjnych pochodzących ze sklepu internetowego, zrealizowanych w latach 2009–2010. Analiza obejmowała czyszczenie danych, wyznaczenie kluczowych statystyk, a także wykorzystanie algorytmu Apriori do wykrywania reguł asocjacyjnych między produktami.

# 2. Charakterystyka danych

Zakres czasowy danych: 2009-12-01 - 2010-12-09. Rekordy: 407695, Transakcje: 19215, Klienci: 4314, Produkty: 4444, Kraje: 37,   
Średnia wartość transakcji: 459.64 GBP.

# 3. Najpopularniejsze produkty i rynki

Top 10 produktów wg ilości:

|  |  |
| --- | --- |
| **Produkt** | **Suma ilości** |
| WHITE HANGING HEART T-LIGHT HOLDER | 56915 |
| WORLD WAR 2 GLIDERS ASSTD DESIGNS | 54754 |
| BROCADE RING PURSE | 48166 |
| PACK OF 72 RETRO SPOT CAKE CASES | 45156 |
| ASSORTED COLOUR BIRD ORNAMENT | 44551 |
| 60 TEATIME FAIRY CAKE CASES | 35806 |
| PACK OF 60 PINK PAISLEY CAKE CASES | 31006 |
| JUMBO BAG RED RETROSPOT | 29578 |
| SMALL POPCORN HOLDER | 25718 |
| BLACK AND WHITE PAISLEY FLOWER MUG | 25685 |

Top 10 produktów wg przychodu:

|  |  |
| --- | --- |
| **Produkt** | **Przychód [GBP]** |
| WHITE HANGING HEART T-LIGHT HOLDER | 151,624.31 |
| REGENCY CAKESTAND 3 TIER | 143,893.35 |
| Manual | 98,560.64 |
| ASSORTED COLOUR BIRD ORNAMENT | 70,493.83 |
| JUMBO BAG RED RETROSPOT | 51,759.30 |
| POSTAGE | 48,741.08 |
| ROTATING SILVER ANGELS T-LIGHT HLDR | 40,186.65 |
| PAPER CHAIN KIT 50'S CHRISTMAS | 36,933.50 |
| PARTY BUNTING | 35,035.90 |
| EDWARDIAN PARASOL NATURAL | 34,044.75 |

Top 10 krajów wg przychodu:

|  |  |
| --- | --- |
| **Kraj** | **Przychód [GBP]** |
| United Kingdom | 7,414,755.96 |
| EIRE | 356,085.21 |
| Netherlands | 268,786.00 |
| Germany | 202,395.32 |
| France | 146,215.42 |
| Sweden | 53,171.39 |
| Denmark | 50,906.85 |
| Spain | 47,601.42 |
| Switzerland | 43,921.39 |
| Australia | 31,446.80 |

Top 10 krajów wg liczby transakcji:

|  |  |
| --- | --- |
| **Kraj** | **Liczba transakcji** |
| United Kingdom | 17614 |
| Germany | 347 |
| EIRE | 316 |
| France | 236 |
| Netherlands | 135 |
| Sweden | 68 |
| Spain | 66 |
| Belgium | 52 |
| Australia | 40 |
| Switzerland | 40 |

# 4. WykresyD

# 5. Reguły asocjacyjne (analiza koszykowa)

Belgium: 9 zbiorów częstych, 3 reguł  
Channel Islands: 7 zbiorów częstych, 0 reguł  
Denmark: 4 zbiorów częstych, 0 reguł  
EIRE: 1 zbiorów częstych, 0 reguł  
France: 4 zbiorów częstych, 1 reguł  
Germany: 4 zbiorów częstych, 1 reguł  
Netherlands: 1 zbiorów częstych, 0 reguł  
Spain: 1 zbiorów częstych, 0 reguł  
Sweden: 2 zbiorów częstych, 0 reguł  
Switzerland: 6 zbiorów częstych, 0 reguł

# 6. Wykorzystane narzędzia i podejście techniczne

Analizę wykonano w języku Python z użyciem bibliotek: pandas, matplotlib, seaborn, mlxtend, openpyxl oraz python-docx. Dane wczytano z Excela, oczyszczono i poddano analizie koszykowej metodą Apriori. Raport i wykresy wygenerowano automatycznie.

# 7. Wnioski i rekomendacje

Przeprowadzona analiza ujawniła wyraźną dominację Zjednoczonego Królestwa jako kluczowego rynku sprzedażowego — zarówno pod względem liczby transakcji, jak i generowanego przychodu. Zdecydowana większość zamówień pochodziła właśnie z tego kraju, co może wskazywać na silnie ugruntowaną bazę klientów oraz skuteczne działania marketingowe prowadzone lokalnie.

Wśród produktów największą popularnością cieszyły się przedmioty dekoracyjne, takie jak świeczniki czy ozdobne akcesoria kuchenne. Świadczy to o popycie na estetykę oraz personalizację przestrzeni domowej wśród klientów e-commerce. Z kolei w analizie przychodów, najwyższe wartości osiągały również produkty o charakterze użytkowym, lecz premium, co sugeruje istnienie segmentu klientów gotowych do ponoszenia większych wydatków za wyższą jakość lub estetykę.

Oczyszczenie danych z błędnych lub niepełnych rekordów (np. zwrotów i niezapłaconych zamówień) okazało się kluczowe dla jakości analizy i poprawności wygenerowanych modeli. Bez tego etapu istniałoby ryzyko błędnych wniosków, zwłaszcza przy analizie reguł asocjacyjnych.

Analiza koszykowa przeprowadzona na 12 krajach o najwyższym przychodzie wykazała istnienie wielu zbiorów częstych, z których część prowadziła do znaczących reguł asocjacyjnych. Choć nie wszystkie kraje wykazały wystarczające powiązania między produktami, uzyskane wyniki mogą zostać z powodzeniem wykorzystane w implementacji mechanizmów rekomendacyjnych (np. systemów typu "często kupowane razem"), a także w działaniach cross-sellingowych i tworzeniu spersonalizowanych ofert produktowych.

Zebrane dane oraz przeprowadzone analizy mogą stanowić solidną podstawę do wdrożenia bardziej zaawansowanych technik predykcyjnych, takich jak segmentacja klientów, prognozowanie popytu, czy też dynamiczne ustalanie cen. Rekomenduje się kontynuację badań w kierunku analizy sezonowości zakupów, a także integracji danych z kanałów marketingowych w celu pełniejszego zrozumienia ścieżki zakupowej klienta.