

Zadanie: Analiza sprzedaży samochodów z pliku CSV

Kontekst:

Jesteś analitykiem w firmie dealerskiej. Masz dane o sprzedaży samochodów z ostatnich 24 miesięcy. Twoim celem jest: oczyścić dane, przeanalizować trendy i wygenerować rekomendacje zwiększające marżę i wolumen. Możesz (i masz) wspierać się AI do pisania kodu – ale każdy fragment z AI skomentuj i zrozum.

Struktura danych (plik car_sales.csv):

sale_id – ID transakcji
sale_date – data sprzedaży (YYYY-MM-DD)
brand – marka (np. Toyota, BMW, Skoda...)
model – model
fuel_type – petrol/diesel/hybrid/ev
transmission – manual/automatic
engine_cc – pojemność silnika (cc)
year – rok produkcji
mileage_km – przebieg (km)
price_pln – cena sprzedaży (PLN)
cost_pln – koszt pozyskania (PLN)
dealer_city – miasto oddziału
customer_type – retail/fleet
warranty_months – gwarancja w miesiącach (0 jeśli brak)

Cele analityczne:

1. Opisać strukturę sprzedaży.
2. Zidentyfikować trendy miesięczne.
3. Wykryć anomalie i outliery.
4. Określić, co wpływa na cenę i marżę.
5. Sformułować rekomendacje dla zarządu.

Zadania krok po kroku:

1. Import i czyszczenie danych przy użyciu pandas.
2. Analiza eksploracyjna danych (EDA) z numpy i pandas.
3. Wizualizacje z matplotlib (linie, słupki, scatter, boxploty).
4. Prosta analiza wpływu cech na cenę i marżę.
5. Ranking miast i kanałów sprzedaży.
6. Raport końcowy z wnioskami i rekomendacjami.

Użycie promptowania (AI):

Wykorzystaj AI co najmniej 3 razy – przy czyszczeniu danych, tworzeniu wizualizacji i analizie insightów. Zapisz treść użytych promptów oraz komentarz, co zostało zmienione ręcznie.

Kryteria oceny:

- Poprawność techniczna kodu.
- Jakość wizualizacji.
- Trafność wniosków.
- Użycie i krytyczna analiza sugestii AI.
- Konkretnie rekomendacje biznesowe.

Rozszerzenia (opcjonalnie):

- Analiza rok-do-roku (YOY).
- Segmentacja cenowa.
- Prosty forecasting sprzedaży.
- Eksport wyników i wykresów do plików.