**ZADANIE DOMOWE – SAMODZIELNY ANALITYK EOBUWIE**

**ZADANIE 1**

Do zadania załączone są 3 pliki CSV:

* Zamówienia
* Cena bazowa
* Wysłane

W pliku zamówienia są informacje o zamówieniach klientów w wybranym okresie. W pliku znajdują się kolumny:

* order\_id – unikalny numer zamówienia
* client\_id – unikalny numer klienta
* order\_date – data złożenia zamówienia
* kupon – informacja o tym, czy podczas składania zamówienia klient użył kuponu rabatowego obniżającego cenę bazową; wartość kuponu to procent zniżki od ceny bazowej

W pliku Cena bazowa znajdują się informacje o cenie bazowej każdego z produktów.

W pliku Wysłane znajdują się informacje o produktach wysłanych ostatecznie do każdego zamówienia. W jednym zamówieniu mogło się znaleźć więcej niż 1 produkt.

Na podstawie załączonych danych odpowiedz na pytania:

1. W jakim dniu roku klienci złożyli najwięcej zamówień?
2. Ilu klientów skorzystało z kuponu rabatowego w trakcie zakupów?
3. Ilu klientów zrobiło w analizowanym okresie więcej niż jedno zamówienie?
4. Który z produktów cieszył się największym powodzeniem? Ilu klientów kupiło go ze zniżką?
5. Który produkt był najczęściej kupowany ze zniżką?
6. Jaka była końcowa wartość wszystkich zamówień w badanym okresie?

Proszę o przesłanie odpowiedzi na pytania i kodu R/ Python, który posłużył do ich wygenerowania.

**ZADANIE 2**

W pliku predykcja.csv znajdują się dane o liczbie zamówień w okresie 01.01.2018 – 01.01.2020. w kolumnie wyprzedaż oznaczono okresy wyprzedaży. Użyj wybranego modelu, żeby zaprognozować sprzedaż w styczniu i lutym 2020.

Jak zmieniłaby się Twoja predykcja, gdybyś wiedział(a), że w dniach 01.02.2020 – 07.02.2020 planowana jest wyprzedaż?

Jako rozwiązanie proszę o przesłanie szacowanej liczby zamówień za styczeń i luty 2020 przy założeniu wyprzedaży w lutym i bez wyprzedaży oraz informacje o krokach analizy i wybranym modelu.