**Temat 1: Klasyfikacja**

Zbiór Danych: Dane **creditcard.csv** dotyczą transakcji dokonywanych kartą kredytową. Zmienna **Class** informuje, czy transakcja jest podejrzana, czy nie.

Zadanie:

* Przygotuj raport opisujący danę oraz ich strukturę. Pamiętaj o wyłączeniu zmiennych, które mogą być nieadekwatne do zadania.
* Przeprowadź klasyfikację z użyciem wybranych algorytmów uczenia maszynowego.
* W oparciu o poznane miary wyznacz najlepszy model.

Wyniki mogą być przedstawione w formie notatnika w Jupyter, albo w formie raporty z użyciem python Dash.

**Temat 2: Analiza danych statystycznych, regresja**

Zbiór Danych: **kc\_house\_data.csv** zawiera informacje o cenach i różnych właściwościach nieruchomości.

* Przygotuj raport opisujący danę oraz ich strukturę. Pamiętaj o wyłączeniu zmiennych, które mogą być nieadekwatne do zadania.
* Przeprowadź regresję liniową, gdzie zmienną zależną jest cena nieruchomości, oceń jej jakość.
* Wskaż najważniejsze zmienne przekładające się na cenę nieruchomości

Wyniki mogą być przedstawione w formie notatnika w Jupyter, albo w formie raporty z użyciem python Dash

**Temat 3: Text Mining**

Zbiór Danych: **sentyment\_movies.csv** zawiera recenzję filmów z portalu branżowego, wraz ze zmienną zawierającą informację, czy recenzja jest pozytywna, czy negatywna.

* Wskaż, które słowa są najbardziej charakterystyczne dla recenzji pozytywnych, a które dla negatywnych.
* Pamiętaj o filtrowaniu słów stanowiących szum, oraz o wybieraniu tych kategorii słów, które mogą być adekwatne do zadania.
* Za pomocą znanych Ci metod uczenia maszynowego sprawdź, czy istnieją jakieś zgrupowania współwystępujących ze sobą słów.

Wyniki mogą być przedstawione w formie notatnika w Jupyter, albo w formie raporty z użyciem python Dash