Dane jakie wykorzystamy w naszym tutorialu to zbiór danych wszystkich pasażerów Tytanica (titanic.csv), z takimi informacjami jak:

* **id** – numer (id) rekordu w tabeli
* **survived**– czy pasażer przeżył (1-tak, 0-nie)
* **pclass** – klasa, którą podróżował pasażer (1-najlepsza)
* **name** – imię, nazwisko pasażera
* **sex** – płeć pasażera (female-kobieta, male-mężczyzna)
* **age** – wiek pasażera
* **sibsp** – liczba rodzeństwa na pokładzie
* **parch** – ilość dzieci lub rodziców pasażera na pokładzie
* **ticket** – numer biletu pasażera
* **fare** – cena biletu pasażera
* **cabin** – numer kajuty pasażera
* **embarked** – Port, w którym osoba wsiadła na pokład (C = Cherbourg; Q = Queenstown; S = Southampton)
* **boat** – numer łódki, którą ewakułował się pasażer
* **body** – numer identyfikacyjny znalezionego ciała, jeśli pasażer zginął i jego ciało zostało zidentyfikowane
* **home.dest** – pochodzenie pasażera