Zajęcia 2 07.03.2022

Zadanie 1.

Wybierz jeden z plików csv z poprzednich ćwiczeń (Mini Kurs) i stwórz ręcznie schemat danych. Przykład na wykładzie. Stwórz DataFrame (DataFrameReader) wczytując plik z użyciem schematu.

Zadanie 2

Użyj kilku rzędów danych z jednego z plików csv i stwórz plik json. Stwórz schemat danych do tego pliku. Przydatny tool to sprawdzenie formatu danych. <https://jsonformatter.curiousconcept.com/>

Zadanie 3

Pobierz notatnik Daty.dbc, użyj funkcji i dokonaj transformacji dat i czasu jak w przykładach.

Zadanie 4

Użycie Read Modes.

Wykorzystaj posiadane pliki bądź dodaj nowe i użyj wszystkich typów oraz ‘badRecordsPath’, zapisz co się dzieje. Jeśli jedna z opcji nie da żadnych efektów, trzeba popsuć dane.

Zadanie 5

Użycie **DataFrameWriter.**

Zapisz jeden z wybranych plików do formatów (‘.parquet’, ‘.json’). Sprawdź, czy dane są zapisane poprawnie, użyj do tego DataFrameReader. Opisz co widzisz w docelowej ścieżce i otwórz używając DataFrameReader.