Zajęcia 2 07.03.2022

**Notatnik 1**

Zadanie 1

Pobierz notatnik Daty.dbc, **użyj wszystkich funkcji** dat podanych w notatniku i dokonaj transformacji dat i czasu jak w przykładach. Rodzaj formatu daty i czasu zależy od Ciebie Chodzi o użycie i poznanie funkcji formatujących daty.

**Notatnik 2**

Zadanie 2.

Wybierz jeden z plików csv z poprzednich ćwiczeń (Mini Kurs, Actors.csv, Names.csv ect..) lub inny plik z **dbfs:/databricks-datasets/** i na jego podstawie stwórz ręcznie **schemat danych.** Przykład na wykładzie.

1. Użyj funckji **spark.read.** i stwórz DataFrame (DataFrameReader) wczytując plik z użyciem schematu, który stworzyłeś.

Zadanie 3

Celem tego zadanie jest stworzenie pliku **json**. Możesz to zrobić ręcznie do czego zachęcam, to pozwoli lepiej zrozumieć plik lub jakimś skryptem. Wykorzystaj jeden z pików z poprzedniego zadania wystarczy kilka rzędów danych z pliku csv i stworzenie/konwersję do json.

1. Stwórz schemat danych do tego pliku, tak jak w poprzednim zadaniu. Przydatny tool to sprawdzenie formatu danych. <https://jsonformatter.curiousconcept.com/>

Zadanie 4

Użycie Read Modes.

Wykorzystaj posiadane pliki bądź użyj nowe i użyj wszystkich typów read modes, zapisz co się dzieje. Poprawny plik nie wywoła żadnych efektów, więc popsuj dane tak aby każda z read modes zadziałała.

Zadanie 5

Użycie **DataFrameWriter.**

Zapisz jeden z wybranych plików do formatów (‘.parquet’, ‘.json’). Sprawdź, czy dane są zapisane poprawnie, użyj do tego spark.read (DataFrameReader).