# Zajęcia 12

**Zadanie 1**

Cel zadania – zapoznanie się z Spark UI zrób krótką notatkę z informacją co można znaleźć w poszczególnych elementach Spark UI Jobs > Stage > Storage > Executors > SQL/Dataframe, gdzie jest informacja o dystrybucji danych?

Żeby uzyskać informacje o przebiegu zadań w Spark UI uruchom istniejący kod notatnik z poprzednich ćwiczeń z przykładami (agregacja, szufle, join ect) i przejdź przez **Spark UI**, powinieneś wiedzieć, gdzie znaleźć informację o tym co przetwarza Spark.

**Zadanie 2**

**Bucketing**

Przygotuj notatnik z przykładem użycia bucketing i porównaj go z partitionBy.

Dla tych samych danych wykonaj obie operacje i sprawdź czy będą różnice w plikach?

**Zadanie 3**

Stwórz jedną tabelę i pobierz statystyki dla tabeli i wszystkich kolumn [ANALYZE TABLE - Spark 3.5.1 Documentation (apache.org)](https://spark.apache.org/docs/latest/sql-ref-syntax-aux-analyze-table.html).