Spark UI to interfejs użytkownika, który pozwala na monitorowanie i analizowanie aplikacji uruchamianych w Apache Spark. Poniżej znajduje się krótki opis tego, co można znaleźć w poszczególnych sekcjach Spark UI:

**Jobs**

W sekcji "Jobs" można znaleźć szczegółowe informacje na temat wszystkich uruchomionych, ukończonych i nieudanych zadań (jobs). Można tu zobaczyć m.in.:

* Identyfikatory zadań (Job IDs)
* Status zadań (np. Running, Succeeded, Failed)
* Czas rozpoczęcia i zakończenia każdego zadania
* Ilość etapów (stages) w ramach danego zadania

**Stages**

Sekcja "Stages" zawiera informacje o poszczególnych etapach przetwarzania danych w ramach zadań. Można tu znaleźć:

* Identyfikatory etapów (Stage IDs)
* Status etapów (np. Pending, Running, Completed)
* Ilość zadań (tasks) w ramach każdego etapu
* Informacje o czasie wykonania, liczbie przetworzonych rekordów i ilości przetworzonych danych
* Informacje o błędach w przypadku nieudanych etapów

**Storage**

Sekcja "Storage" pokazuje informacje na temat danych przechowywanych w pamięci (cache) lub na dysku. Można tu znaleźć:

* RDDs (Resilient Distributed Datasets) i DataFrames przechowywane w pamięci
* Liczbę partycji dla każdego RDD/DataFrame
* Rozmiar danych w pamięci i na dysku
* Informacje o tym, które węzły (executors) przechowują poszczególne partycje

**Executors**

W sekcji "Executors" można znaleźć szczegółowe informacje na temat wszystkich wykonawców (executors) działających w klastrze. Informacje te obejmują:

* Identyfikatory wykonawców
* Stan wykonawców (np. Active, Dead, Removed)
* Ilość użytej pamięci
* Liczbę przetworzonych zadań (tasks)
* Informacje o zasobach (CPU, pamięć) zużywanych przez poszczególnych wykonawców

**SQL/Dataframe**

Sekcja "SQL/Dataframe" zawiera informacje na temat zapytań SQL i operacji na DataFrames. Można tu znaleźć:

* Listę uruchomionych zapytań SQL
* Fizyczne plany wykonania (execution plans) zapytań
* Statystyki dotyczące czasu wykonania, ilości przetworzonych danych i liczby operacji I/O
* Informacje o optymalizacjach wykonanych przez Catalyst Optimizer

**Dystrybucja danych**

Informacje o dystrybucji danych można znaleźć w kilku miejscach Spark UI, w szczególności:

* W sekcji "Storage" – gdzie można zobaczyć, które węzły przechowują poszczególne partycje danych.
* W sekcji "Executors" – gdzie można zobaczyć, jak dane są przetwarzane przez poszczególnych wykonawców oraz ile zadań jest przypisanych do każdego z nich.
* W sekcji "SQL/Dataframe" – gdzie można znaleźć szczegóły dotyczące przetwarzania zapytań i dystrybucji operacji na partycje w klastrze.

Aby uzyskać te informacje, należy uruchomić istniejący kod z wcześniejszych ćwiczeń (np. agregacje, shuffle, joiny), a następnie przejść przez Spark UI, aby zidentyfikować, co dokładnie przetwarza Spark i jak dane są dystrybuowane w klastrze.