



Moduł 10

Przypadek użycia 3

EDA dla COVID Data hub

iBigWorld:
Innovations for Big Data in a Real World

UBB team



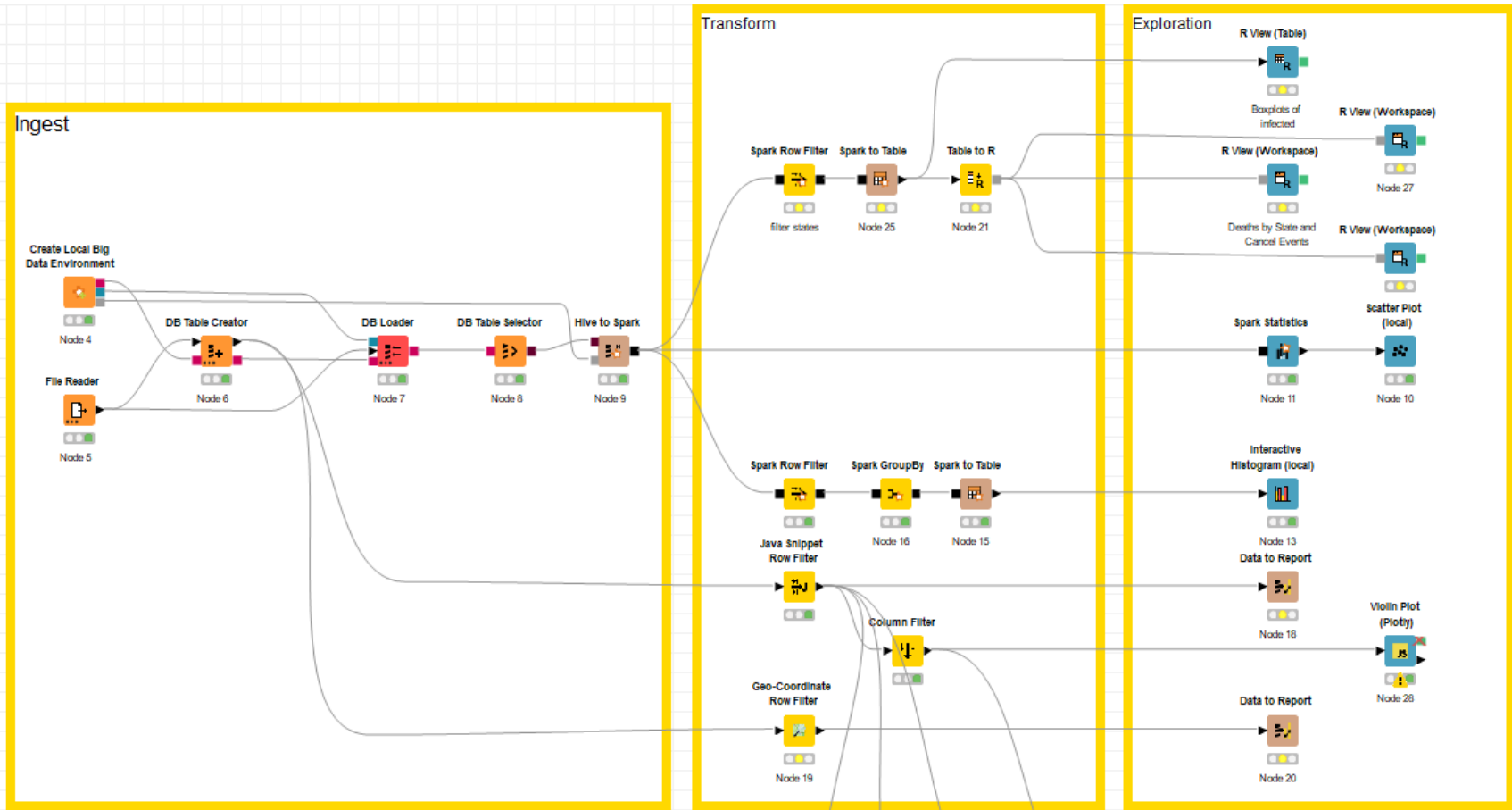
Disclaimer: Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency (NA). Neither the European Union nor NA can be held responsible for them.



Co-funded by the
Erasmus+ Programme
of the European Union

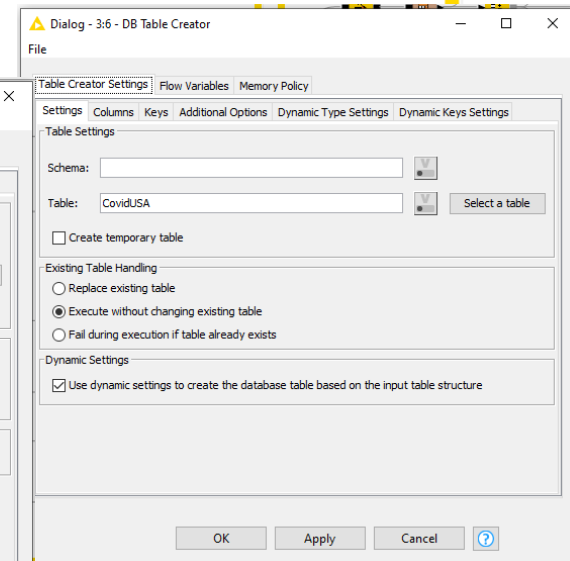
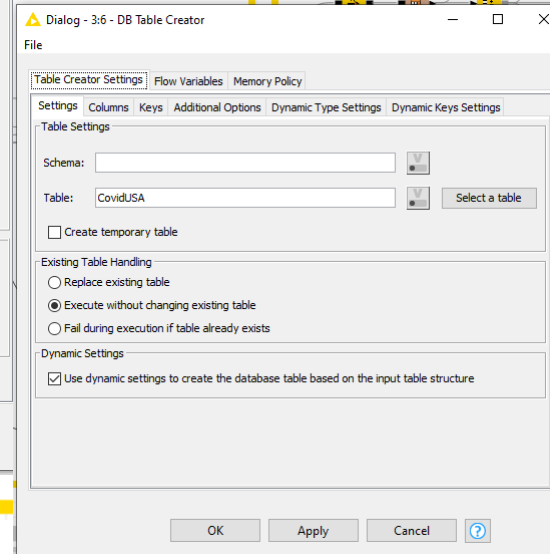
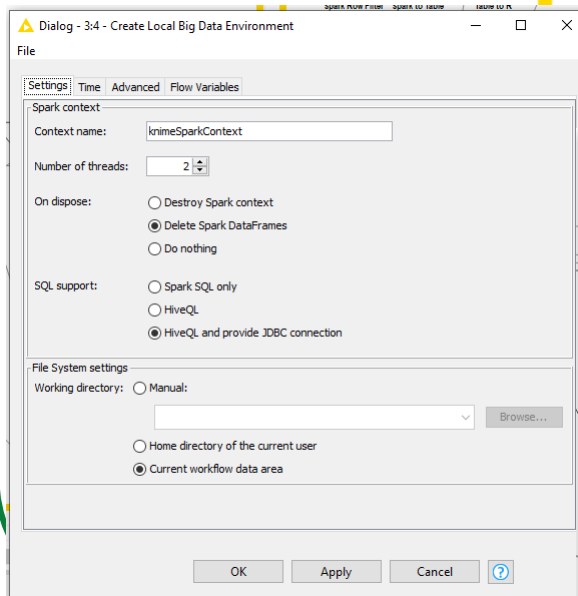
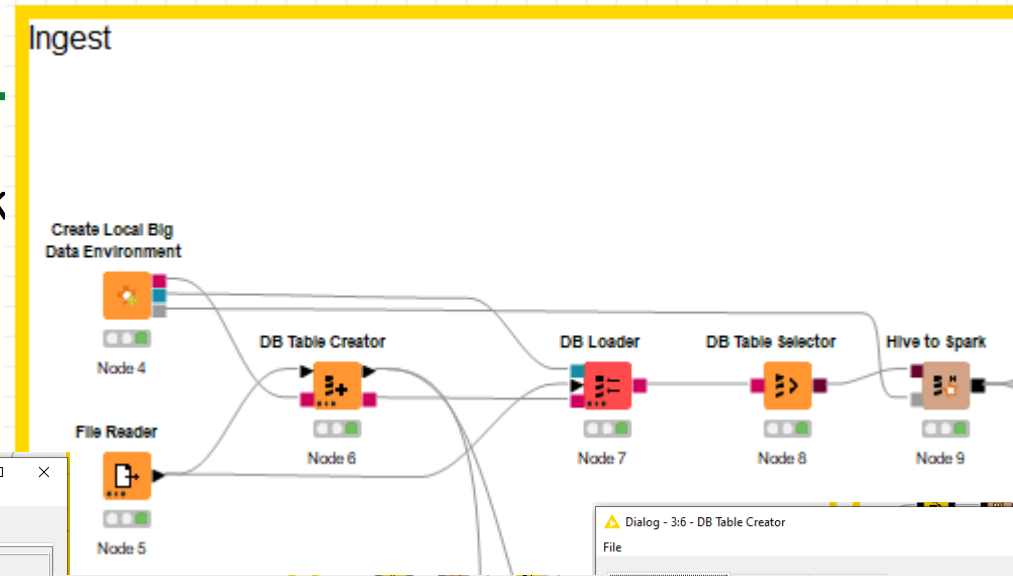


Przeptyw pracy EDA dla COVID Data hub



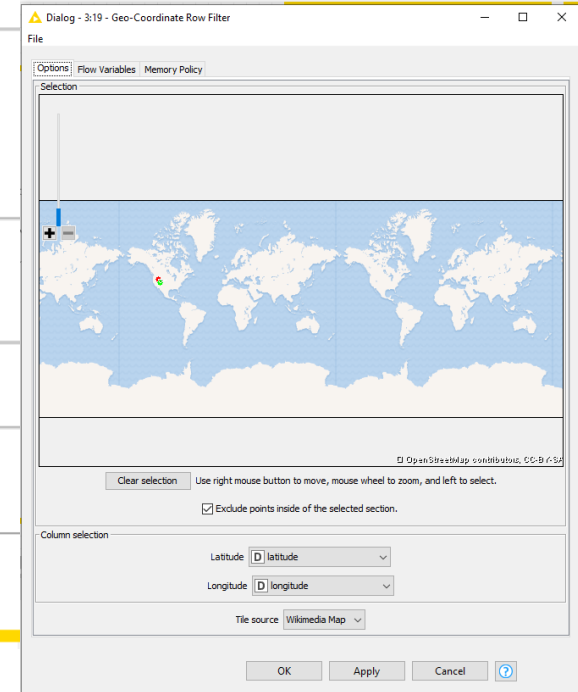
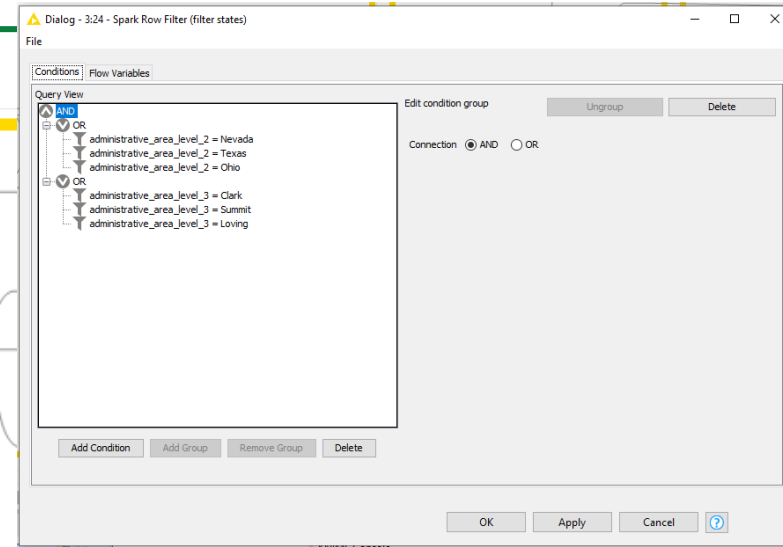
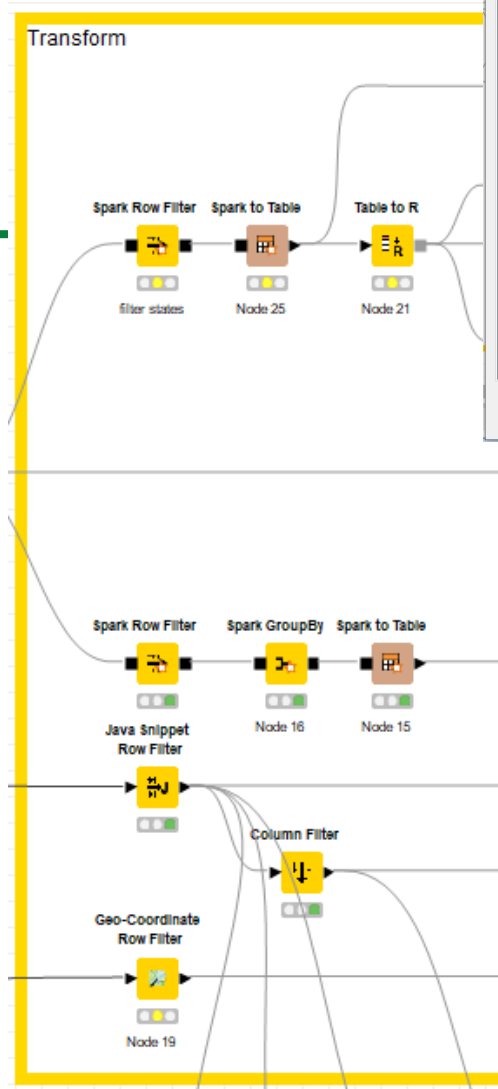
Przygotowanie danych

- Załadowanie danych do Spark



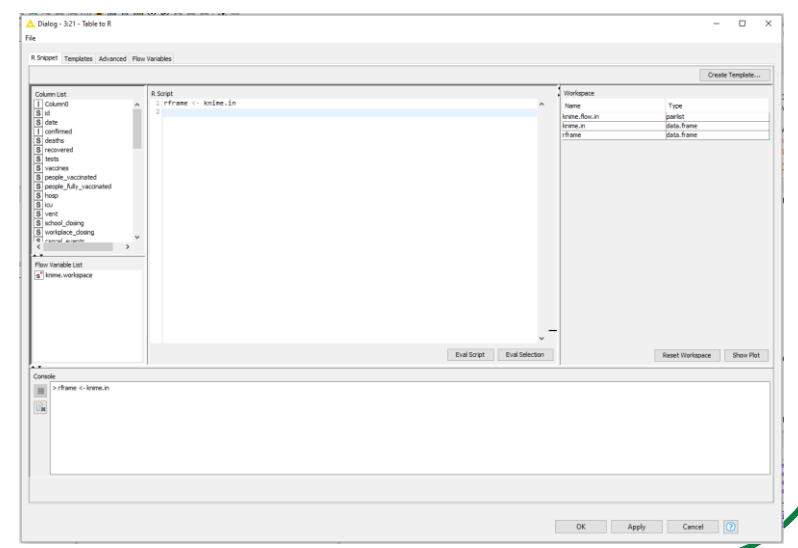
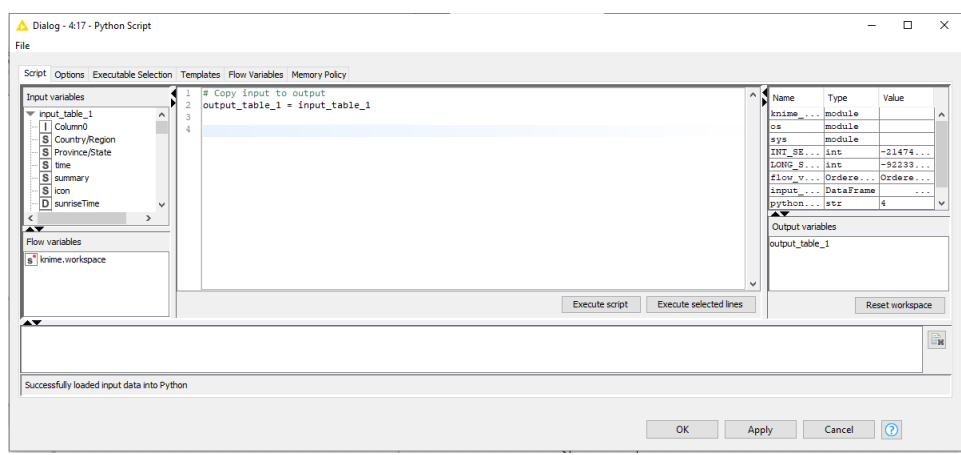
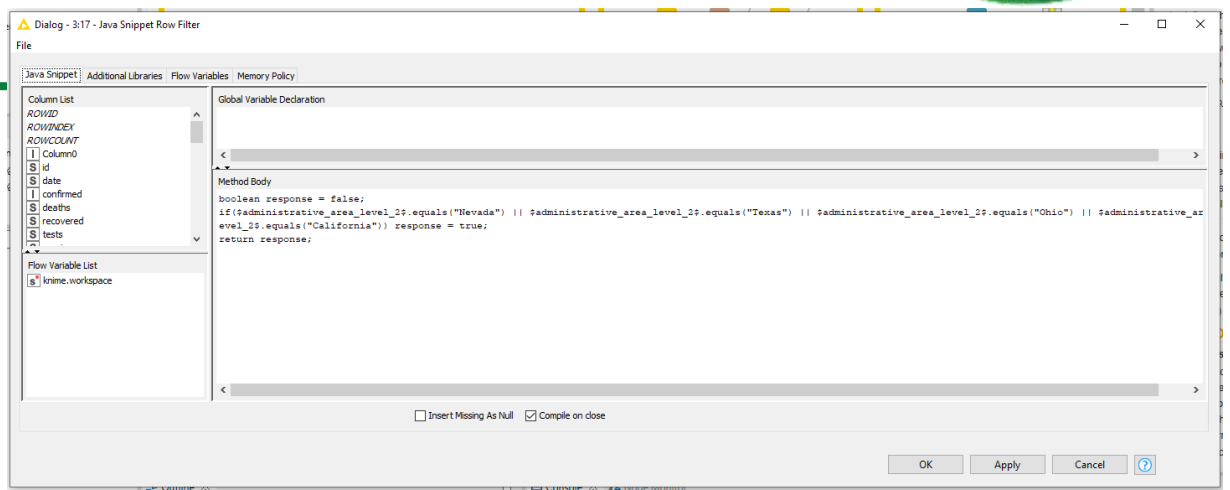
Transformacja danych

- Kwerendy dla środowiska Spark
- Transformacja do obiektu Tabela używanej do analizy i wizualizacji (Eksploracji)
- Zaawansowane filtrowanie danych (geo-koordynaty, itp.)



Transformacja danych

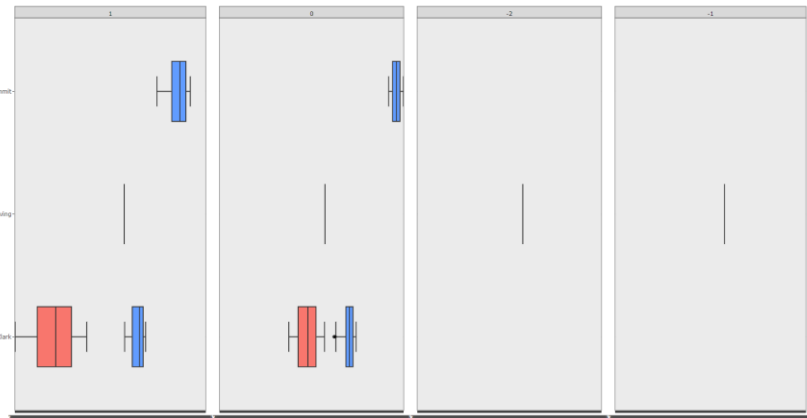
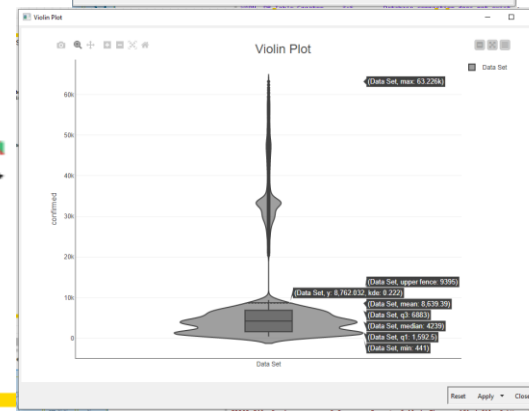
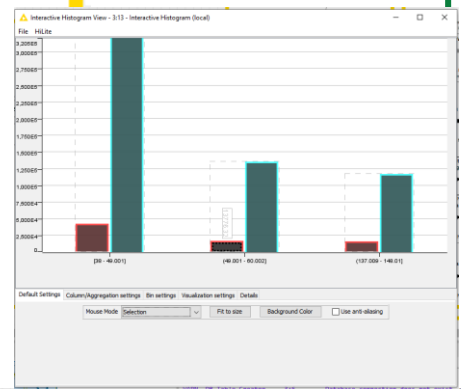
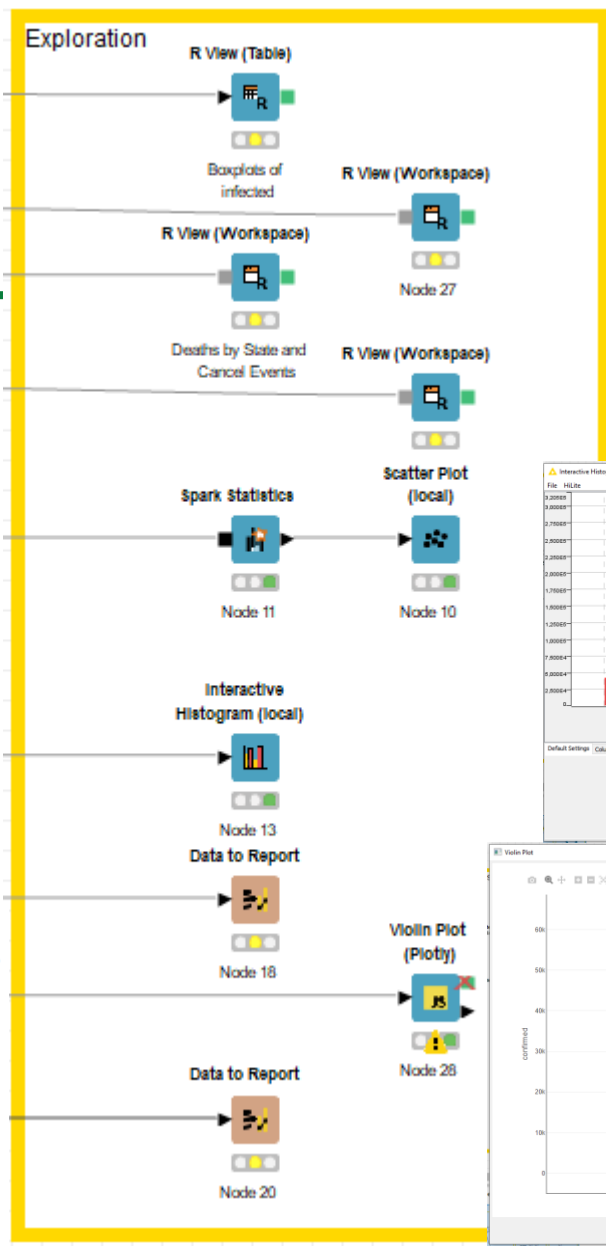
- Kwerendy przy użyciu „Java Snippet”
- Skryptów R
- Skryptów Python



Eksploracja danych



- Agregujące statystyki
- Tabele przestawne
- raporty
- Wykresy rozrzutu
- Wykresy pudełkowe
- Histogramy



Eksploracja danych c.d.

- Agregujące statystyki
- Tabele przestawne
- raporty
- Wykresy rozrzutu
- Wykresy pudełkowe
- Histogramy

