

Plan projektu zaliczeniowego kursu Analiza danych w Pythonie

1. Opis pliku/plików wejściowych

- Pochodzenie i charakterystyka inputu
- Podstawowe statystyki + próbka danych
- Cel analizy

2. Praca z danymi – operacje numpy/pandas

- Wybór danych ze zbioru
- Sortowanie
- Filtrowanie
- Grupowanie
- Brakujące dane - identyfikacja
- Czyszczenie zbioru
- Formatowanie danych
- Analiza outlierów

3. Wizualizacja danych

- Dane kategoryjne
- Dane ilościowe
- Statystyka opisowa
- Zastosowanie wybranych bibliotek do podstawowej wizualizacji
- Przedstawienie danych w postaci dashboardu

4. Podsumowanie projektu, wnioski biznesowe

- Wskazanie głównych wniosków wynikających z analizy danych
- Propozycja działań dla biznesu
- Wskazanie jakie dane byłyby potrzebne do pogłębionej analizy