



Akademia Górniczo-Hutnicza  
Im. Stanisława Staszica w Krakowie

Sprawozdanie z przedmiotu:

**Analiza Danych**

Temat:

**Analiza ras psów**

Wydział: Odlewnictwa

Kierunek: Technologie Przemysłu 4.0

Rok: III

Studia stacjonarne, rok akademicki 2024/2025

Mikołaj Mikołajczyk

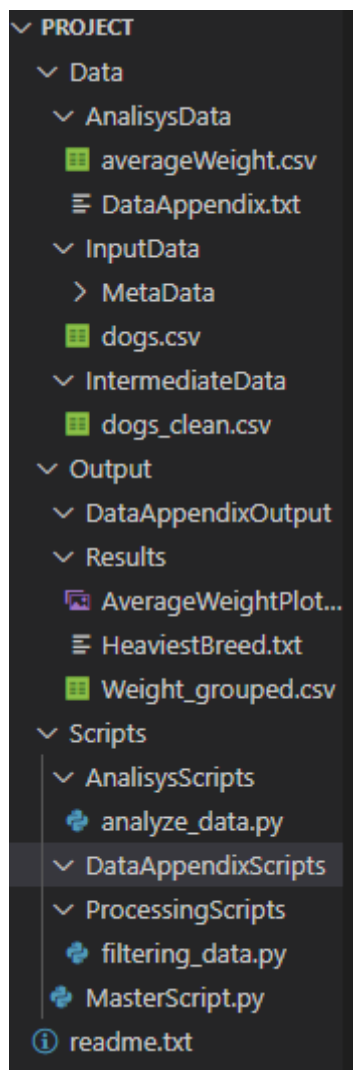
## 1. Założenia projektu

Celem projektu jest rozwiązanie problemu, która rasa psa jest najcięższa. Przeprowadzono szczegółową analizę na podstawie otrzymanego zestawu danych *dogs.csv* przez platformę UPeL.

W przygotowaniu struktury projektu użyty został protokół [TIER 4.0](#).

## 2. Struktura projektu

Tak jak wspomniano, projekt został ułożony strukturalnie przy pomocy protokołu TIER 4.0. Protokół TIER określa zawartość i organizację dokumentacji dla projektu zajmującego się statystyką oraz analizą danych. Zasady struktury protokołu są ściśle określone.



### 3. Przebieg projektu

- Pierwszym punktem projektu było pobranie i zapoznanie się z surowymi danymi z pliku *dogs.csv* pobranego z platformy UPeL
- Pobrano do odczytu dany zestaw danych, naprawiając jego format:

```
b = pd.read_csv("dogs.csv", sep=";")

b.columns = b.columns.str.strip()

print(b)
```

- Pozbyto się pustych wierszy oraz 'niepotrzebnych' do dalszej analizy kolumn:

```
new_b = b.dropna()

print(new_b)
```

```
new_b = new_b.drop(columns=['Color'], errors='ignore')

print(new_b)
```

- Dodano nową kolumnę, określającą wiek danego psa:

```
today = pd.to_datetime('today')

new_b.loc[:, 'Age'] = today.year - new_b['Date of Birth'].dt.year
```

- Zgrupowano dane potrzebne do analizy wagi ras psów i umieszczono w osobnym pliku csv

```
breed_grouped = new_b.groupby('Breed')[['Weight (kg)', 'Height (cm)']].mean()

breed_grouped.to_csv('Weight_grouped.csv', index=True, sep=";")
```

- Sprawdzono przy pomocy prostego algorytmu, która rasa ma największą średnią wagę. Następnie wprowadzono odpowiedź do pliku tekstowego *HeaviestBreed.txt*

```
heaviest_one = breed_grouped["Weight (kg)"].idxmax()
heaviest_weight = breed_grouped["Weight (kg)"].max()

result = f"Heaviest breed is {heaviest_one} with an average weight of {heaviest_weight}"

print(result)

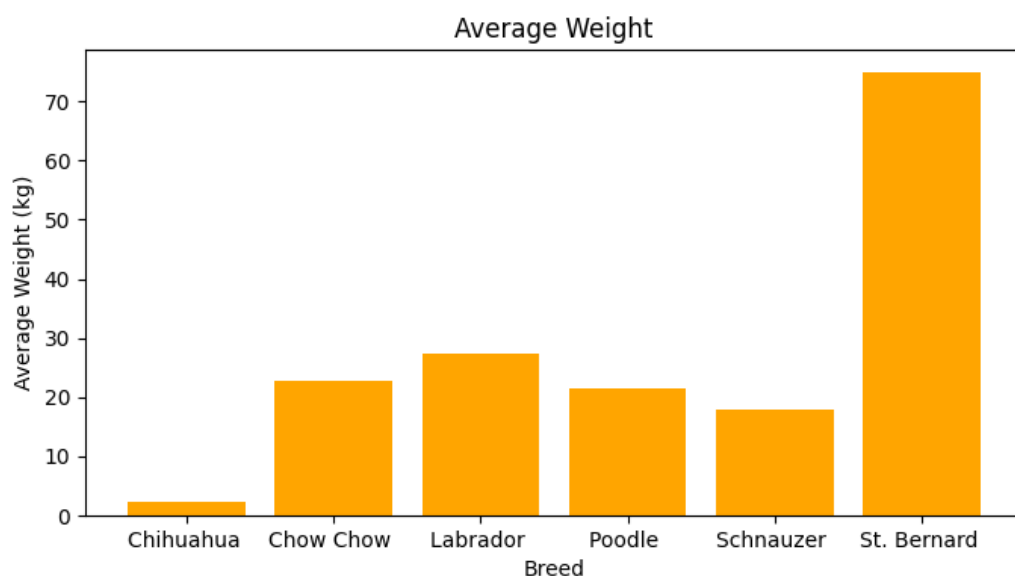
with open('./Output/Results/HeaviestBreed.txt', 'w') as f:
    f.write(result)
```

- Na potrzeby wizualizacji utworzono wykres za pomocą biblioteki Matplotlib:

```
plt.figure(figsize=(8, 4))
plt.bar(breed_grouped.index, breed_grouped['Weight (kg)'], color='orange')
plt.title('Average Weight')
plt.xlabel('Breed')
plt.ylabel('Average Weight (kg)')

plt.savefig("AverageWeightPlot")
```

#### 4. Wyniki



Rys 1. Średnia waga różnych ras psów

Według przedstawionego wyżej wykresu wynika, że najcięższą wagą jest **St. Bernard** ze średnim wynikiem **75 kg**.

#### 5. Wnioski

Analiza danych okazała się skuteczna, prawidłowo została określona najcięższa rasa psa. Protokół TIER 4.0 znacząco ułatwia analizę przygotowanego projektu.