source: @allison_horst https://github.com/allisonhorst/penguins ⧉

You have been asked to support a team of researchers who have been collecting data about penguins in Antartica! The data is available in csv-Format as `penguins.csv`
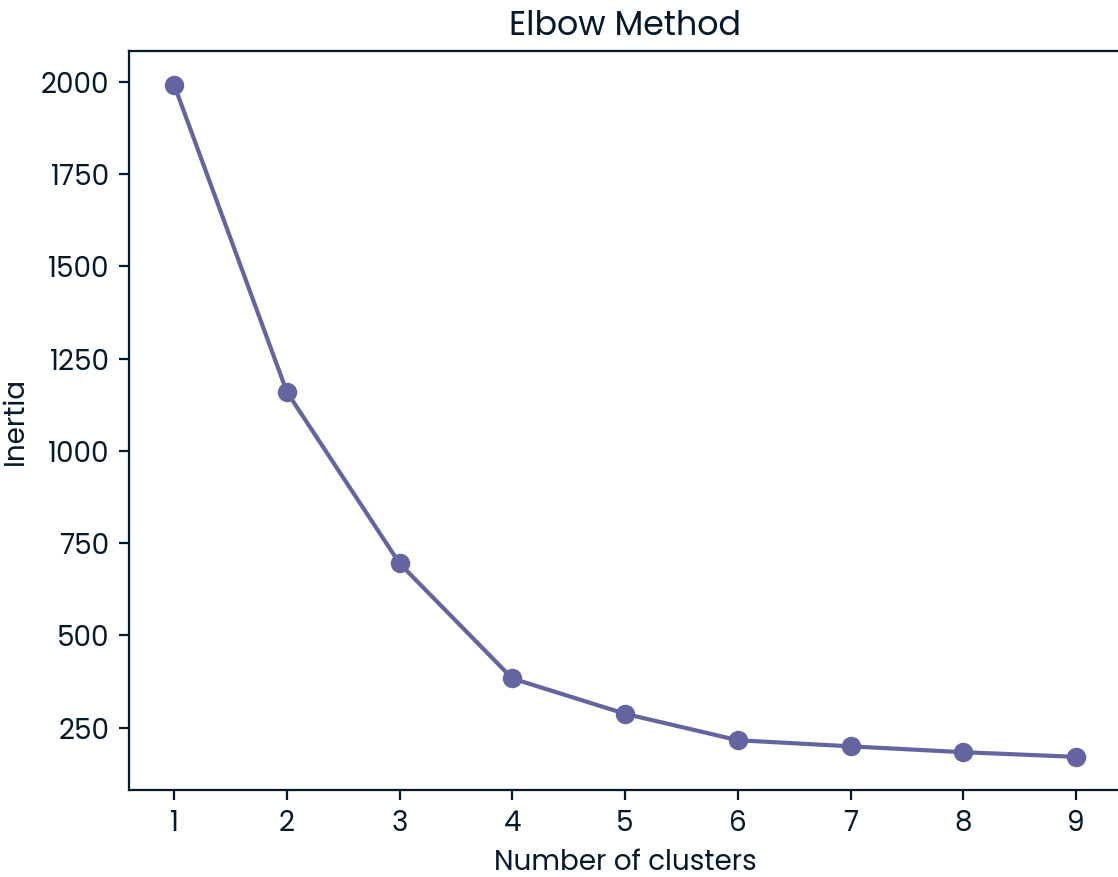
Origin of this data : Data were collected and made available by Dr. Kristen Gorman and the Palmer Station, Antarctica LTER, a member of the Long Term Ecological Research Network.

The dataset consists of 5 columns.

Column	Description
culmen_length_mm	culmen length (mm)
culmen_depth_mm	culmen depth (mm)
flipper_length_mm	flipper length (mm)
body_mass_g	body mass (g)
sex	penguin sex

Unfortunately, they have not been able to record the species of penguin, but they know that there are **at least three** species that are native to the region: **Adelie, Chinstrap,** and **Gentoo.** Your task is to apply your data science skills to help them identify groups in the dataset!

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 332 entries, 0 to 331
Data columns (total 5 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   culmen_length_mm   332 non-null    float64
 1   culmen_depth_mm    332 non-null    float64
 2   flipper_length_mm  332 non-null    float64
 3   body_mass_g        332 non-null    float64
 4   sex                332 non-null    object
dtypes: float64(4), object(1)
memory usage: 13.1+ KB
```

Elbow Method

K-means Clustering (K=4)