

The Penguin Dataset EDA

Performing EDA on the dataset to gain valuable insights

Overview

This dataset is about the species of penguin separated on the basis of their physical features and place of living. They are divided into three main categories according to their species

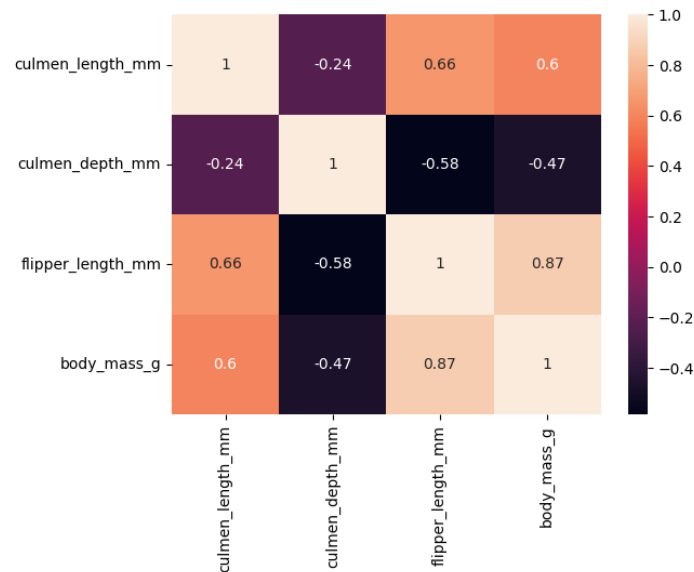
- Adelie
- Chinstrap
- Gentoo

Questions that need to be answered

- What particular species (in percentage) likes to live on a specific island —> Done
- In which species male have more culmen length and depth than females and vice versa (same for flipper length) —> Done
- Which island has the most population in total —> Done
- Ratio of females as compared to male on a particular island
- Relation of mass to the genders and species (there is a good chance that some species may be heavier than others) —> Done

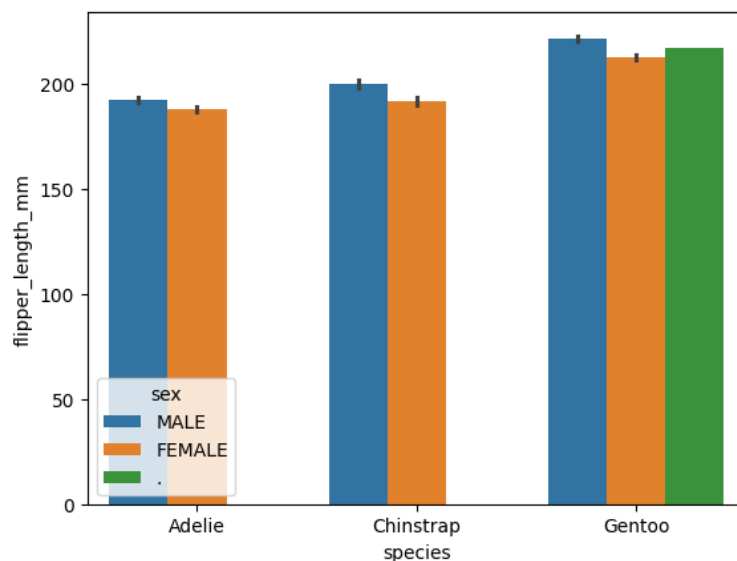
Bird's eye view of data

- The data contains 344 rows and 7 columns
- Species, island and sex are categorical columns whereas culmen_length_mm, culmen_depth_mm, flipper_length_mm and flipper_depth_mm are numerical columns
- The data contains total of 18 missing values
- The average culmen_length is about 43.92mm
- The average culmen_depth is about 17.15mm
- The average flipper_length is about 200.91mm
- The average body mass is about 4201.7 grams
- culmen_length has a moderate correlation to flipper_length and body_mass with a correlation of 0.66
- flipper_length is very strongly correlated with body_mass with a correlation of 0.87
there is a chance that penguins with bigger flipper tend to have more mass



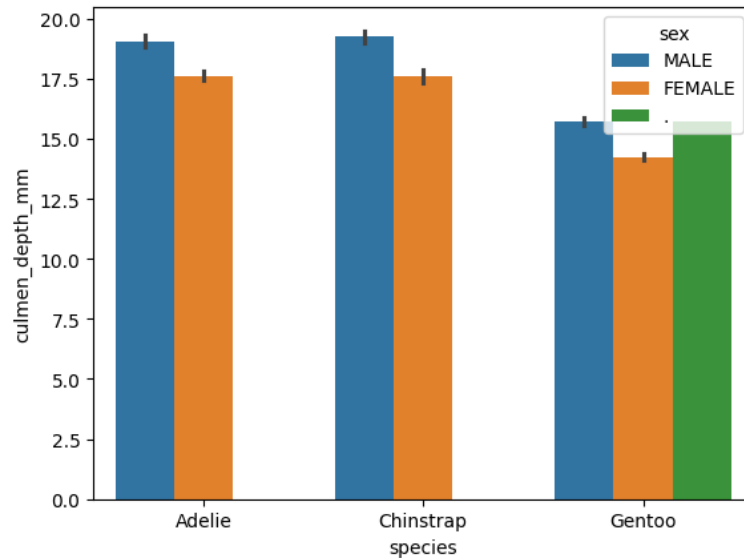
Species and sex vs flipper_length_mm

The Gentoo species have the largest flipper size followed by Chinstrap and Adelie. In all the species male have more flipper length as compared to females



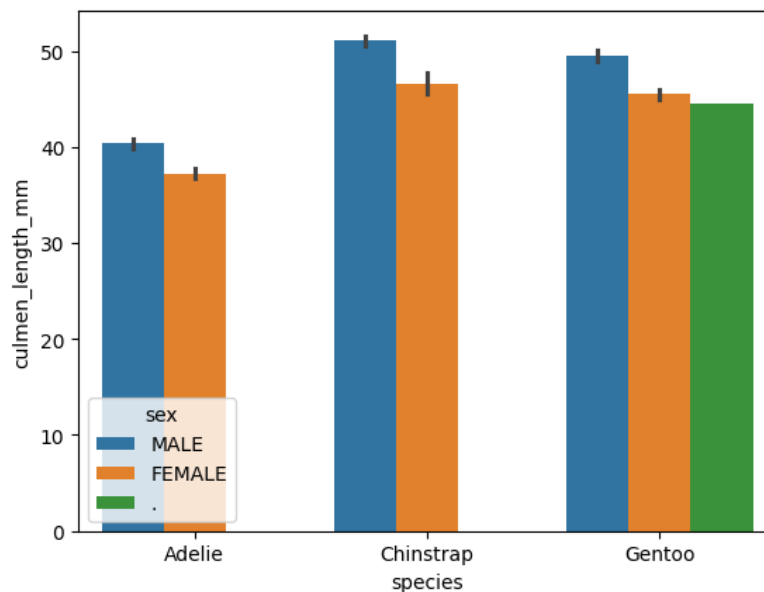
Species and sex vs culmen_depth_mm

The Adelie species have the deepest culmens followed by Chinstraps and then Gentoo. Here in every specie females tends to have deeper culmens as compared to males



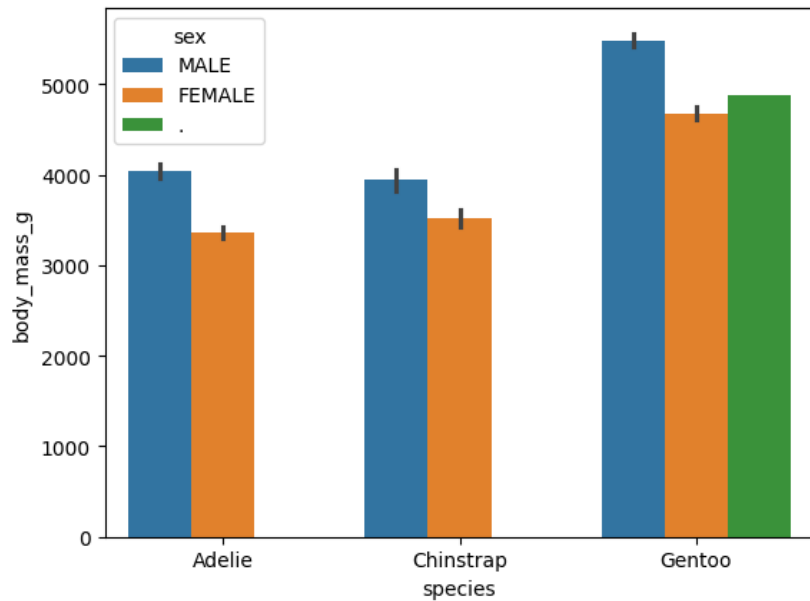
Species and sex vs culmen_length_mm

The Chinstrap species have the longest culmen followed by Gentoo and then Adelie. Every species males have more culmen length than the females



Species and sex vs body_mass

Gentoo penguins are the heaviest among the three species, while Adelie and Chinstrap penguins have similar masses, differing by only a few hundred grams. In all species, females have less mass compared to males.



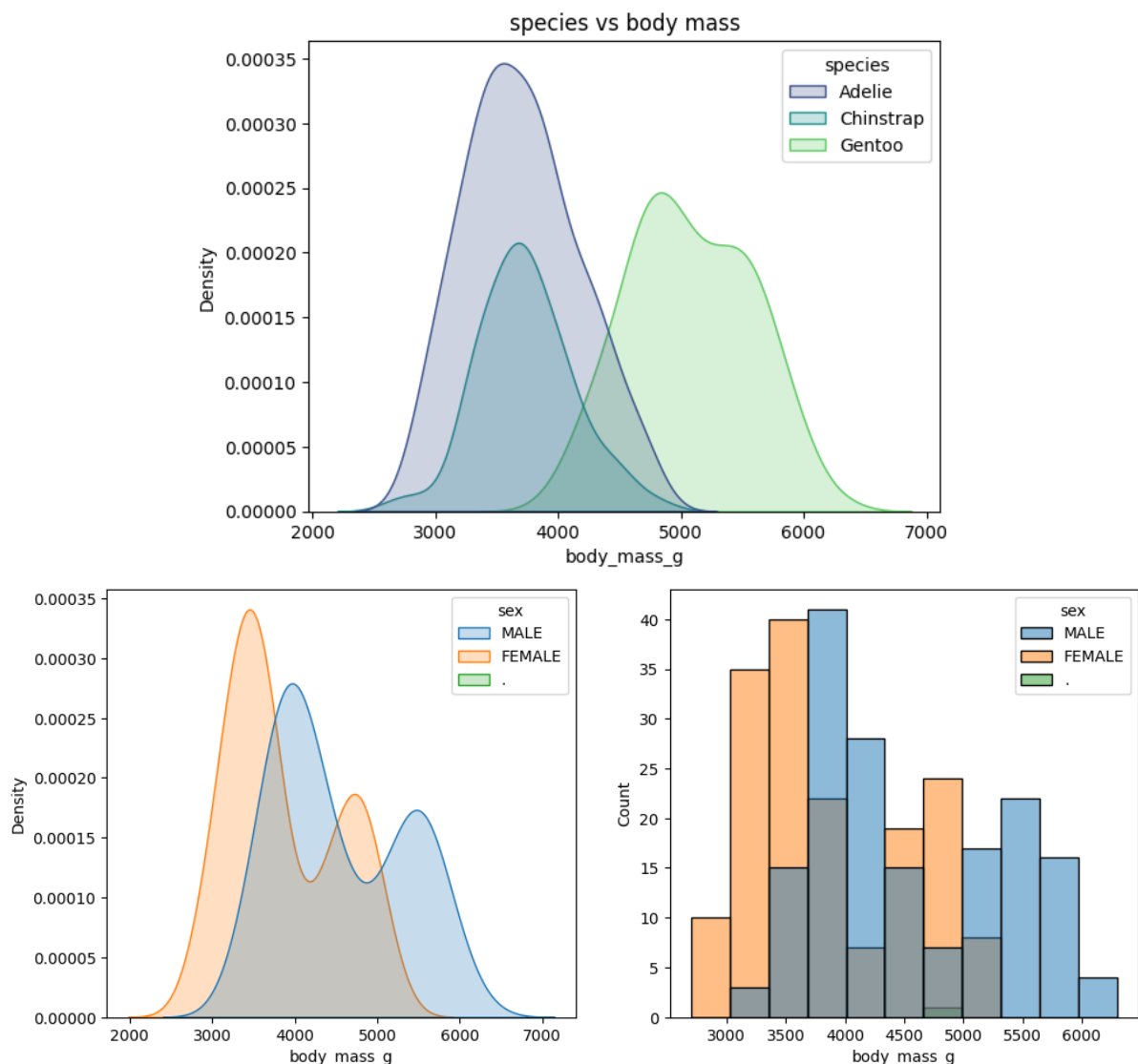
The data shows that most females weigh between 2,000 and 3,500 grams. Few females weigh around 4,000 grams, though there is a significant group weighing between 4,500 and 5,000 grams.

Males display a similar weight distribution pattern.

Gentoo penguins show the highest mass variance at approximately 504 grams.

Chinstrap penguins show the lowest mass variance at approximately 384 grams.

| | species | island | mean | std | count |
|---|-----------|-----------|-------------|------------|-------|
| 0 | Adelie | Biscoe | 3709.659091 | 487.733722 | 44 |
| 1 | Adelie | Dream | 3688.392857 | 455.146437 | 56 |
| 2 | Adelie | Torgersen | 3706.372549 | 445.107940 | 51 |
| 3 | Chinstrap | Dream | 3733.088235 | 384.335081 | 68 |
| 4 | Gentoo | Biscoe | 5076.016260 | 504.116237 | 123 |



Species vs Islands

- Adelie: They are found on all the islands probably showing that they are not strictly bound to a particular environment (36.84% on Dream island, 34.21% on Torgersen island and 28.94% on Biscoe island)
- Chinstrap: They are only found on Dream island indicating that this specie can only survive in its favourable environment which was there on Dream island (100% on Dream island)
- Gentoo: Similar case as Chinstrap they are only found on Biscoe island (100% on Biscoe island)

