

Exploratory Data Analysis: Penguin Size Dataset

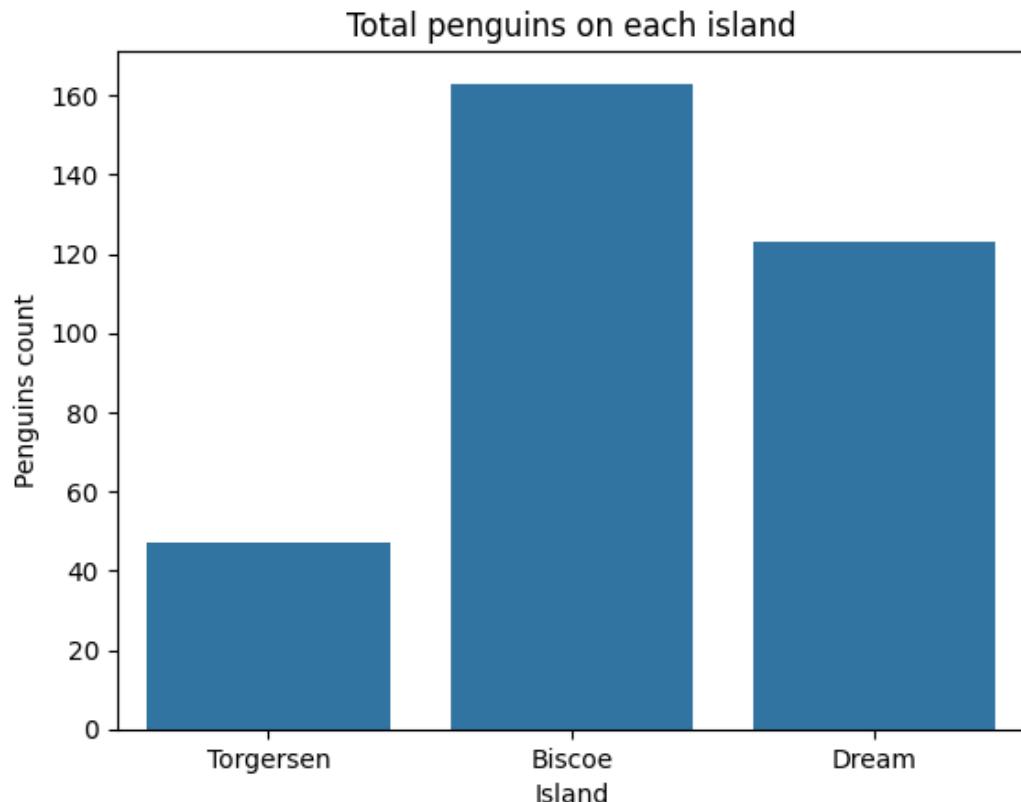
INTRODUCTION

This dataset contains measurements for penguins, including species, island, culmen length and depth, flipper length, body mass, and sex. The goal of this analysis is to explore physical differences across species and between males and females.

DATA CLEANING AND PREPARATION

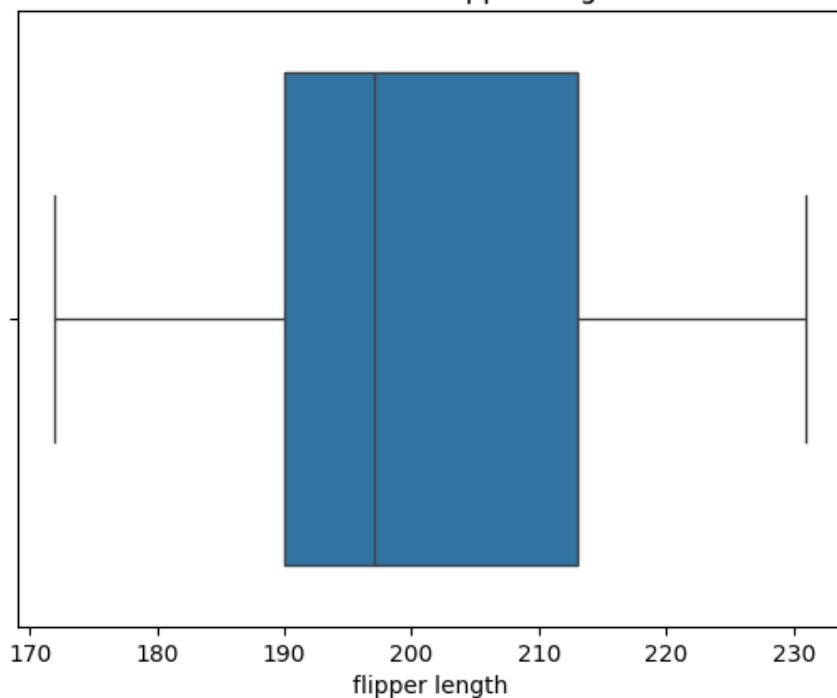
Columns `flipper_length_mm` and `body_mass_g` were converted to integers for consistency. Duplicate rows were removed. Rows containing missing or invalid values (e.g., '?' in the `sex` column) were discarded.

DATA STORIES AND VISUALISATIONS



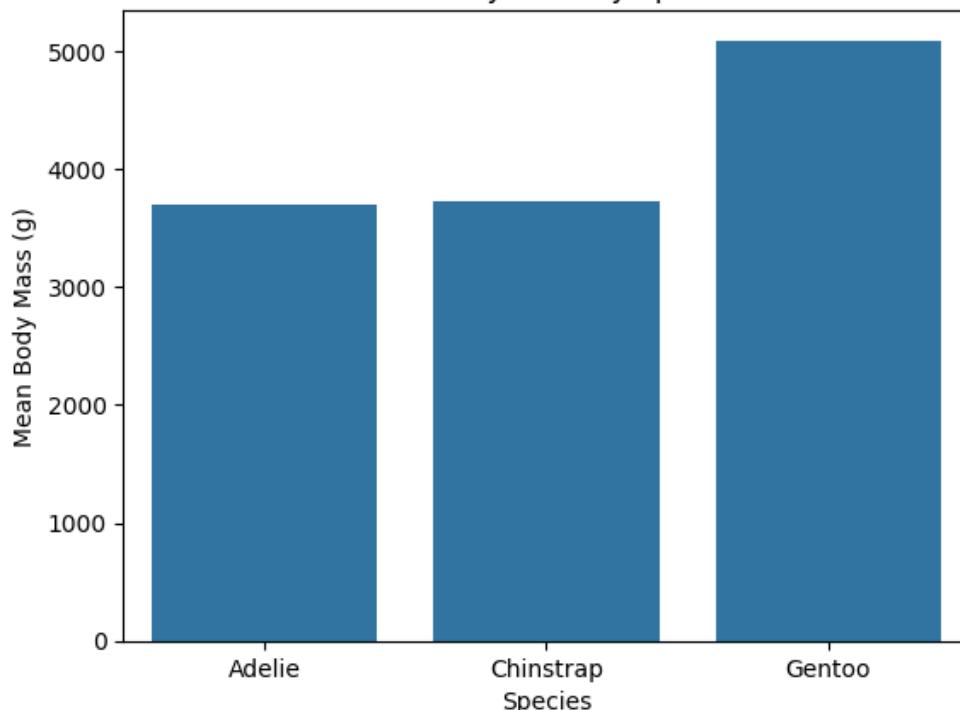
Biscoe Island has the highest penguin population (+/-160), followed by Dream (+/-120) and Torgersen (+/-50).

Distribution of flipper length

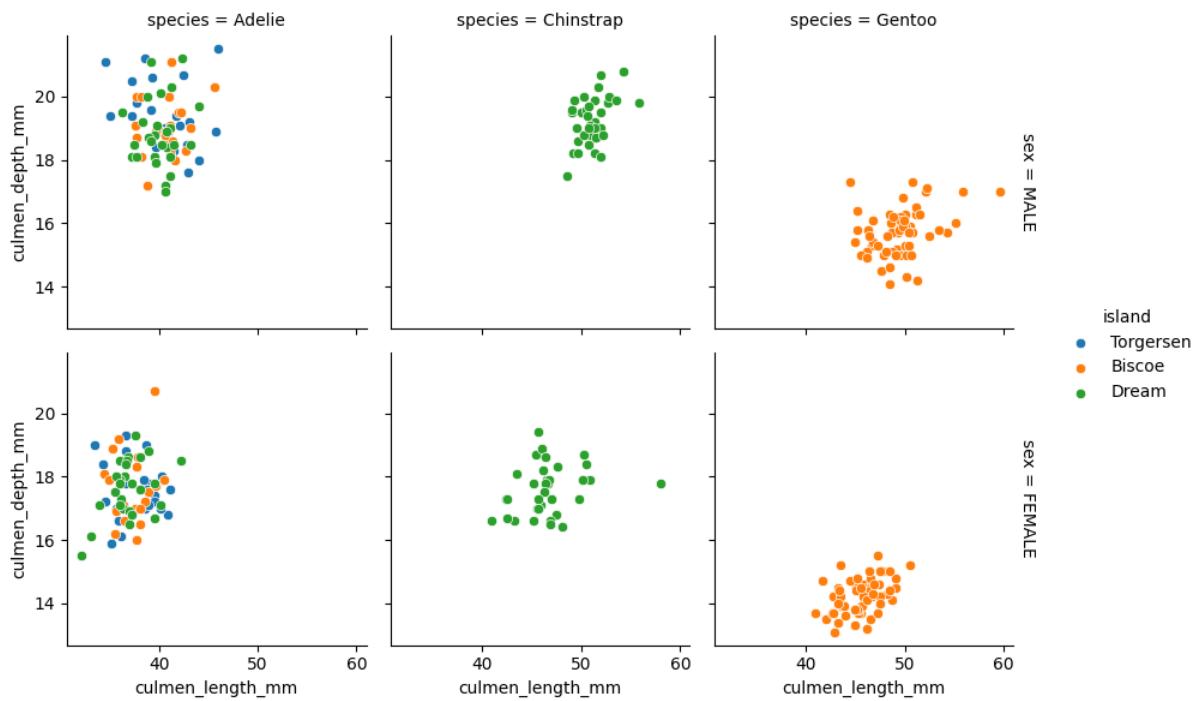


Most penguins have flipper lengths between roughly 190–215 mm.

Mean Body Mass by Species



Gentoo penguins are the heaviest (+/-500g on average), while Adelie and Chinstrap are slightly lighter, just under 4000g.



Males generally have larger culmen lengths and depths than females.

Adelie penguins inhabit all three islands, Chinstrap penguins only Dream, and Gentoo penguins only Biscoe.

Across all species, culmen length and depth show a positive correlation.

Prepared by: Christiaan Cronje | Data Analytics Portfolio