Collaborators: Jackson Miller & Vittorio Centore + GenAI

GenAI Statement:
Infrastructure (Including Flowchart Ideas) and slight debugging of program functions.

Kaggle Data set: Penguins

Columns to work with:

- Species
- Body mass
- Bill length

Calculations to perform:

- Average body mass for each species and how they compare
- Average bill length from penguins over a 3500G
- Number of penguins who are in the upper 25% AND have a bill length over 42?
- What amount of penguins from the Gentoo species are above 4000G?

Decomposition program:

main()

Calls the other functions
in a logical sequence

INPUT:
None

load_data (csv_file)

Reads CSV as dataFrame

INPUT: csv_file)
OUTPUT: df (DataFrame)

calculate_average body_mass (df)

Calculates average miss
for each penguin specl

INPUT: df (DataFrame)
OUTPUT: avg_body_mass)

select_heavy_bills

Filters penguins with a
body mass over 3500 g
and calculates average bill
length

INPUT: df (DataFrame)
OUTPUT: avg_biill_length

find_upper_quartile long_bills (df)

Identifies penguins in upper 25%
of body mass with a bill length over
42 mm

INPUT: df (DataFrame)
OUTPUT: num_upper_quartile_long
bills tint)

find_heavy_quartile long_bills (df)

Counts penguins
above 4000 g¬ft

find_heavy_gentoo_count (df)

Counts: total cantoo peng-
uins above 4 000 g df