Daniel Xiao

SI 201

Project 1

6 October 2025

Name of Dataset: penguins

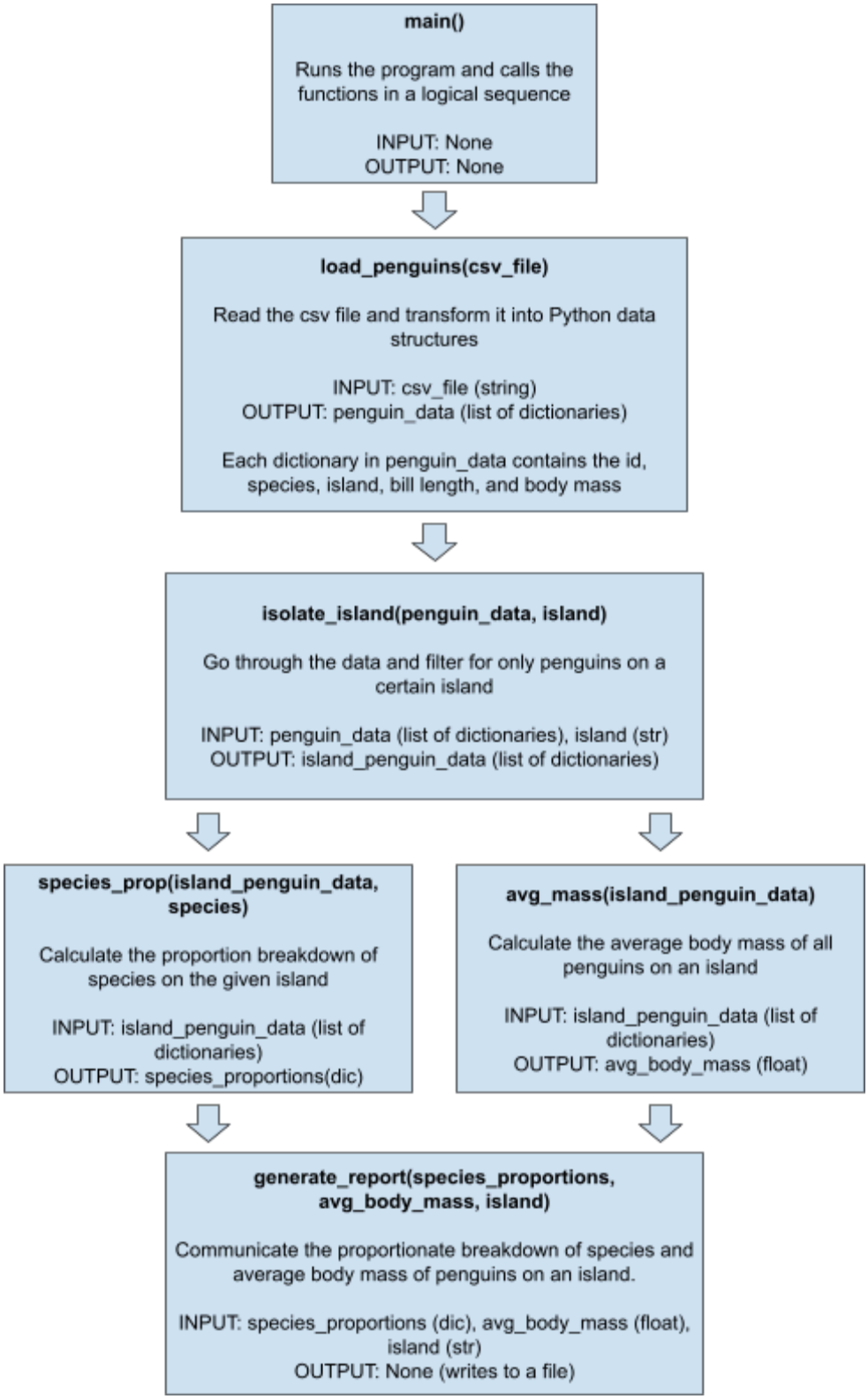
Columns Used: species, island, body_mass_g

Calculations Performed:

- What is the percentage breakdown of species of penguins on a certain island?
- What is the average body_mass_g of all penguins on a certain island?

Function Decomposition Diagram: (next page)

Names of Collaborators: Daniel Xiao (just me)


main()
Runs the program and calls the functions in a logical sequence
INPUT: None
OUTPUT: None
load_penguins(csv_file)
Read the csv file and transform it into Python data structures
INPUT: csv_file (string)
OUTPUT: penguin_data (list of dictionaries)
Each dictionary in penguin_data contains the id, species, island, bill length, and body mass
isolate_island(penguin_data, island)
Go through the data and filter for only penguins on a certain island
INPUT: penguin_data (list of dictionaries), island (str)
OUTPUT: island_penguin_data (list of dictionaries)
species_prop(island_penguin_data, species)
Calculate the proportion breakdown of species on the given island
INPUT: island_penguin_data (list of dictionaries)
OUTPUT: species_proportions(dic)
avg_mass(island_penguin_data)
Calculate the average body mass of all penguins on an island
INPUT: island_penguin_data (list of dictionaries)
OUTPUT: avg_body_mass (float)
generate_report(species_proportions, avg_body_mass, island)
Communicate the proportionate breakdown of species and average body mass of penguins on an island.
INPUT: species_proportions (dic), avg_body_mass (float), island (str)
OUTPUT: None (writes to a file)