main()

Runs the program and calls the functions in a logical sequence.

INPUT: NONE OUTPUT: NONE

load_penguin_data(csv_filepath):

Reads the Kaggle penguins.csv file and converts it into a dictionary where each index serves as a key, and the corresponding value is a nested dictionary containing variable names and their respective values.

INPUT: csv_filepath (string)
OUTPUT: penguin_data (dictionary)

group_weight_by_year_island_species_sex(penguin_data)

Organizes penguin weight data into nested groups based on year, island, species, and sex. For each group, it collects a list of body weights to enable calculations like mean, median, or mode.

INPUT: penguin_data (dictionary)
OUTPUT: grouped_penguin_data (dictionary)

mean_weight_by_sex(grouped_penguin_data)

Calculates the average body weight of penguins for each sex (Male and Female), across all years, islands, and species.

INPUT: grouped_penguin_data (dictionary)
OUTPUT: means (dictionary)

median_weight_by_sex(grouped_penguin_data)

Calculates the median body weight of penguins for each sex (Male and Female), across all years, islands, and species.

INPUT: grouped_penguin_data (dictionary)
OUTPUT: medians (dictionary)

mode_weight_by_sex(grouped_penguin_data)

Calculates the mode body weight of penguins for each sex (Male and Female), across all years, islands, and species.

INPUT: grouped_penguin_data (dictionary)
OUTPUT: modes (dictionary)

distribution_shapes(means, medians, modes)

Analyzes the relationship between the mean, median, and mode for each group to infer the shape of the distribution: symmetric, right-skewed, or left-skewed.

INPUT: means, medians, modes (dictionaries)
OUTPUT: distribution_shapes (dictionary)

generate_report(distribution_shapes)

It creates a report summarizing the inferred distribution shape, e.g., symmetric, right-skewed, left-skewed, for each group and writes it to a file in CSV format.

INPUT: distribution_shapes (dictionary)
OUTPUT: NONE (outputs to a csv file)