**CS 4500 Software Engineering**

**Small Group 3 (SG3)**

Anthony Burrows

Zach Brown

Diamon Daniels

Khristian East

**Due: May 9th, 2025**

**Major Revisions Log**

1. **Threshold Calculation Order Swapped**  
Date: 2025‑04‑25  
**Issue:** When testing the ASCII heat map on sample data, low/medium/high labels were reversed (M and L swapped).  
**Details:** For a species with min=0 and max=90, cut‑points A=30, B=60 were misassigned; code used B for low/medium split and A for medium/high.  
**Fixes Implemented:**

* **Corrected assignment:** computed t1 = min + (max-min)/3 and t2 = min + 2\*(max-min)/3 in proper order.
* Added unit check logging to verify A < B.  
  Outcome: The L/M/H mapping now matches expected bins (e.g. values ≤30→L, ≤60→M, >60→H).

2. **ASCII HeatMap.txt Formatting Error**  
Date: 2025‑04‑25  
**Issue:** The generated HeatMap.txt file lacked comma delimiters, producing lines like 01/05/2025 LMH instead of 01/05/2025,L,M,H.  
**Fixes Implemented:**

* Updated code to join codes with ",".join(codes).
* Ensured header date + "," + codes\_string.  
  Outcome: HeatMap.txt now contains correctly formatted CSV lines, and downstream parsers can read it.

3. **Date‑Grouping Not Detecting Dupes**  
Date: 2025‑04‑25  
**Issue:** Identical H/M/L patterns for 02/22/2006 and 11/12/2016 were not reported.  
**Details:** Code used separate threshold lists for species and dates and a mismatched thresholds variable, causing lookup errors.  
**Fixes Implemented:**

* Unified thresholds computation at one place and reused it for both date- and species-grouping.
* Consolidated grouping logic to use the same thresholds list.  
  Outcome: Duplicate dates are now detected and printed correctly.

2025-04-27 - Initial Design

Date - Updates to Pseudocode and Call Graph

START

Print program description.

Prompt user repeatedly for valid .CSV filename and file existence.

Open file and read species names (count N).

Validate file: ensure 1–999 data lines.

FOR each data line:

Validate date format and abundance counts (N non-negative numbers).

Store valid dates and counts.

Print species and dates count; wait for ENTER.

Write Species.txt (species names) and DatedData.txt (dates).

Create presence/absence matrix:

Map abundance counts > 0 to 1; else 0.

Write PresentAbsent.txt.

FOR each date:

Find and print max abundance and species sharing it.

Group dates with identical presence/absence vectors and display.

Create heat map:

For each species, classify abundance as Low, Medium, or High.

Display and save HeatMap.txt.

Find and report species sharing identical L/M/H patterns.

Print completion message; wait for ENTER.

END

main()

|

|-- print\_program\_description()

|

|-- filename = get\_valid\_filename()

| |-- validate\_csv\_extension(filename)

| |-- check\_file\_exists(filename)

|

|-- species\_list = read\_species\_names(file)

|

|-- validate\_file\_data(file)

| |-- validate\_date\_format(date)

| |-- validate\_abundance\_counts(numbers)

|

|-- display\_species\_and\_dates\_count(species\_list, dates\_list)

|

|-- write\_species\_file(species\_list) | |-- write\_dated\_data\_file(dates\_list)

|

|-- presence\_absence\_matrix = generate\_presence\_absence\_matrix(abundance\_data)

| |-- for each abundance value |

|-- map value to 0 or 1

|

|-- write\_present\_absent\_file(presence\_absence\_matrix)

|

|-- print\_max\_abundance\_per\_date(abundance\_data)

|

|-- group\_dates\_by\_presence\_absence(presence\_absence\_matrix)

|

|-- heatmap\_matrix = generate\_heat\_map(abundance\_data)

| |-- calculate\_low\_mid\_high\_ranges(species\_data)

| |-- map abundance values to L/M/H

|

|-- write\_heatmap\_file(heatmap\_matrix)

|

|-- print\_heat\_map(heatmap\_matrix)

|

|-- find\_species\_with\_same\_heatmap\_patterns(heatmap\_matrix)

|

|-- display\_exit\_message()

**Test Plans:**

**Test Plan 1: Threshold Bin Order**  
**Scenario:** Verify that low/medium/high cut-points are assigned in the correct order.  
**Input:** A CSV with one species whose min abundance = 0 and max abundance = 90.

,SpeciesX

2025-01-01,0

2025-01-02,90

**Expected Output in HeatMap.txt (ASCII):**

2025-01-01,L

2025-01-02,H

**Steps:**

1. Run python sg3.py.
2. Enter the above file.
3. Open HeatMap.txt and check that the first line is 2025-01-01,L and the second 2025-01-02,H.

* **Before Revision Result:**  
  HeatMap.txt showed 2025-01-01,H and 2025-01-02,L (bins reversed).
* **After Revision Result:**  
  Correctly shows L for 0 and H for 90.

**Test Plan 2: ASCII Formatting**  
**Scenario:** Ensure comma separators appear in the ASCII heat map.  
**Input:** A CSV with one date and three species where labels should be L, M, H:

,SpA,SpB,SpC

2020-01-01,10,40,80

**Expected Output in HeatMap.txt:**

2020-01-01,L,M,H

**Steps:**

1. Run python sg3.py.
2. Enter the above file.
3. Open HeatMap.txt and confirm the line contains commas between the letters.

* **Before Revision Result:**  
  HeatMap.txt showed 2020-01-01 LMH (no commas).
* **After Revision Result:**  
  HeatMap.txt shows 2020-01-01,L,M,H.

**Test Plan 3: Duplicate‐Date Detection**  
**Scenario:** Detect two dates sharing the same L/M/H pattern.  
**Input:** A CSV where two rows yield identical L/L/L codes:

,Organism1,Organism2

2006-02-22,0,0

2016-11-12,0,0

**Expected Console Output:**

Dates 2006-02-22, 2016-11-12 share the H/M/L pattern: L,L

**Steps:**

1. Run python sg3.py.
2. Enter the above file.
3. Observe console messages for duplicate-date grouping.

* **Before Revision Result:**  
  No duplicate-date message.
* **After Revision Result:**  
  Correct message appears.

**Test Plan 4: Duplicate‐Species Detection**  
**Scenario:** Detect two species sharing the same H/M/L sequence across all dates.  
**Input:** A CSV with two species whose counts track identically over three dates:

,SpA,SpB

2020-01-01,5,5

2020-02-01,10,10

2020-03-01,15,15

**Expected Console Output:**

Species SpA, SpB share the H/M/L pattern: L,M,H

**Steps:**

1. Run python sg3.py.
2. Enter the above file.
3. Observe console messages for duplicate-species grouping.

* **Before Revision Result:**  
  No duplicate-species message.
* **After Revision Result:**  
  Correct message appears.