**CS 5402 Intro to DataMining**

**Cole Davis**

**Homework #2**

**Question #1:**

Table

Description automatically generated

**Question #2:** Graphical user interface, table

Description automatically generated

Diagram

Description automatically generated

**Question #3:**

Country\_likelyhood = 2/3 \* 1/3 \* 1/3 \* 3/8

Not\_Country\_likelyhood = (1/3 \* 2/3 \* 0/3 \* 3/8) \* (0/2 \* ½ \* ½ \* 2/8)

Convert to probabilities by normalizing so they sum to 1:

Final Answer Probability = Country\_likelyhood / (Country\_likelyhood + Not\_Country\_likelyhood)

**Question #4:**

Graphical user interface, application

Description automatically generated

Diagram

Description automatically generated

**Question #5a:**

entropyBeforeSplit = -3/8\*log(3/8) – ¼\*log(1/4) – 3/8\*log(3/8)

**Question #5b:**

entropyMystery = -1/4 \* log(1/4) – ¼ \* log(1/4) – ½ \* log(1/2)

**Question #5c:**

informationGain = X – (1/2 \* Y + ½ \* Z)

**Question #6a:**

**Question #6b:**

**Question #6c:**