## Intern Data Challenge 2025

**Purpose:** To understand your problem-solving and coding background.

**Expected time:** 1hour

**Exercise:**  The enclosed csv file contains three columns:

1. RegionID – an integer 1 – 10 corresponding to the geographic regions shown below
2. Population – an integer indicating the current population of each region
3. MoveProb – a double between 0 and 1 indicating the probability that an individual in the region will move to some **adjacent** region the following year.

A region is adjacent to another if they share a portion of their boundary, excluding corners (so 1 and 9 are not adjacent for this exercise).

A map of the state of pennsylvania

Description automatically generated

With the goal of estimating the population of each block in the future, assuming no one moves in or out of the area (i.e., moving is only occurring between the regions shown), complete the following:

1. Explain how you decided the number of residents moving from a region to another adjacent region. Note that the probabilities only indicate the likelihood of residents moving out, but do not indicate which adjacent region they move to. \*\* Note: There is no “correct” answer here. This is a necessary step to make the problem well-posed so you can program possible solutions. Also, this is not how population moves in real life.
2. Using R or Python (your preference):
   1. Import “population\_data.csv”
   2. Use an appropriate data structure to encode region adjacency. (There are several options here – use what you are comfortable with)
   3. For each region, identify how many people will move out and how many regions they could choose among.
   4. Create a function to estimate the number of people moving TO each region.
   5. Compute the population of each region after one year.