**Report**

Question 0

1. Another way to store data is to store them as txt files. It is an easy way to view the data on any platforms with numerous text-edit soft wares. However, it is more complicated to convert them into list of lists in python which is a simple form to process data.
2. A list of lists can represent rows and columns, then we can index a specific cell easily. The disadvantage is that a list of lists is not represented as rows and columns which may be hard to compare data to the original csv files.
3. Another data structure to represent is to use dictionary with the row indices as keys. It looks clear without many connected square brackets. However, dictionary has no specific orders which may cause more work of indexing.
4. <https://www.data.act.gov.au/Justice-Safety-and-Emergency/ACT-Fire-Danger-Rating-Meter/7zub-jyjj>. This file is based on the forecast weather conditions. Since it is more likely to get a fire in a dry condition and more unlikely to catch a fire in a wet condition, so this data can be used to model bushfire risk.

Question 1

1. All of my functions return a list of lists with the type string.
2. Blank values are represented as empty strings **‘’** because they have the same type as other data and they are easily to be converted.
3. I defined a temporary function to count the blank cells.

def sum\_blank(lst):

'''Count the number of blank values.

lst : a list of lists

'''

sum\_acc = 0

for i in range(len(lst)):

for j in range(len(lst[i])):

if lst[i][j] == '':

sum\_acc += 1

return sum\_acc

Anu: **‘wind.csv’** has **0** blank values.

Other files have **1181** blank values.

South: **‘wind.csv’** has **295** blank values.

Other files have **8348** blank values.

Act: **‘wind.csv’** has **86037** blank values.

Other files have **95313** blank values.

Question 2

1. anu/wind.csv : 7.29959

south/wind.csv : 7.0872

act/wind.csv : ﻿8.60718

1. m: the number of rows

n: the number of columns

time complexity: m\*n + m

Question 3

1. For **anu** dataset :

﻿Urban Vegetation: 315

Woodland: 125

Open Forest: 368

Forest: 370

Open Woodland: 50

Grassland: 65

Arboretum: 26

For **south** dataset:

﻿Open Woodland: 5475

Grassland: 2599

Woodland: 3182

Arboretum: 43

Urban Vegetation: 1866

Golf Course: 336

Open Forest: 252

Pine Forest: 325

Forest: 44

Shrubland: 30

For **act** dataset:

﻿Open Woodland: 23077

Woodland: 11934

Grassland: 17291

Forest: 3710

Open Forest: 3732

Urban Vegetation: 4023

Golf Course: 712

Vineyard: 95

Pine Forest: 6728

Shrubland: 705

Arboretum: 260

1. For **anu** dataset:

﻿Urban Vegetation: 0.00sq m

Woodland: 625000.00sq m

Open Forest: 2944000.00sq m

Forest: 3700000.00sq m

Open Woodland: 100000.00sq m

Grassland: 13000.00sq m

Arboretum: 0.00sq m

For **south** dataset:

﻿Open Woodland: 10950000.00sq m

Grassland: 519800.00sq m

Woodland: 15910000.00sq m

Arboretum: 0.00sq m

Urban Vegetation: 0.00sq m

Golf Course: 0.00sq m

Open Forest: 2016000.00sq m

Pine Forest: 0.00sq m

Forest: 440000.00sq m

Shrubland: 0.00sq m

For **act** dataset:

﻿Open Woodland: 46154000.00sq m

Woodland: 59670000.00sq m

Grassland: 3458200.00sq m

Forest: 37100000.00sq m

Open Forest: 29856000.00sq m

Urban Vegetation: 0.00sq m

Golf Course: 0.00sq m

Vineyard: 0.00sq m

Pine Forest: 0.00sq m

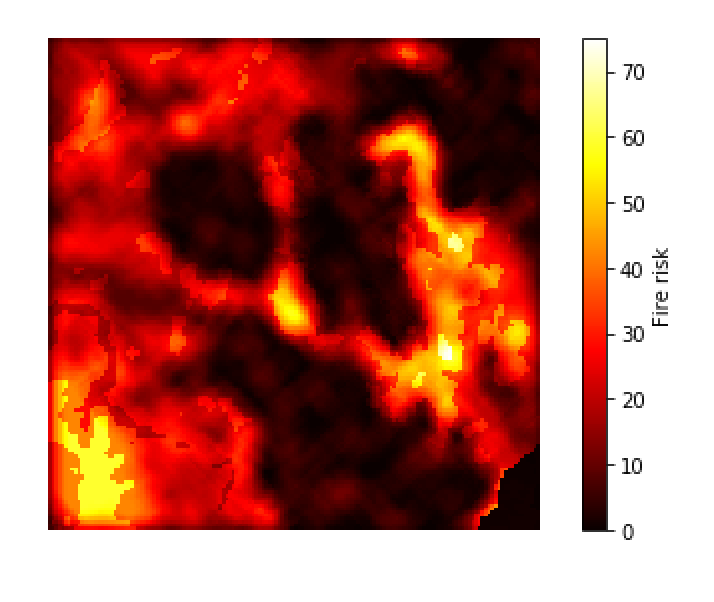
Shrubland: 0.00sq m

Arboretum: 0.00sq m

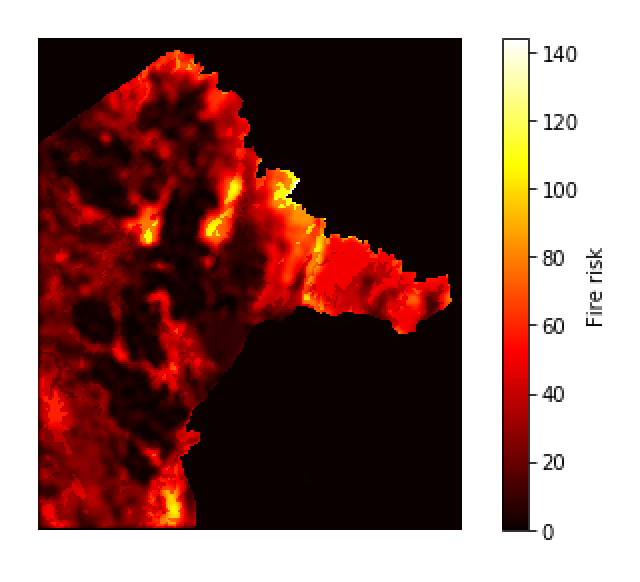
1. n\*n times since I need to traverse all cells

Question 4

1. For **south** dataset, the similarities are that the general shapes are similar as well as the level of the fire risk. However, in the bottom part of my image where the fire risk is almost 0, but in the same area of BPA map, some parts still have rather high risk. The reason I think we use a rather simple function to calculate the factor and the distant of nearby we define is not very accurate.



For **act** dataset, the shape and general fire risk are similar to BPA map. However, in the left part of my image, there exists some areas whose fire risk are 0. The same areas in the BPA all have high fire risk. The reason caused this problem I think is regarding the nearby cells and the algorithm we use for calculating fire risk factors.



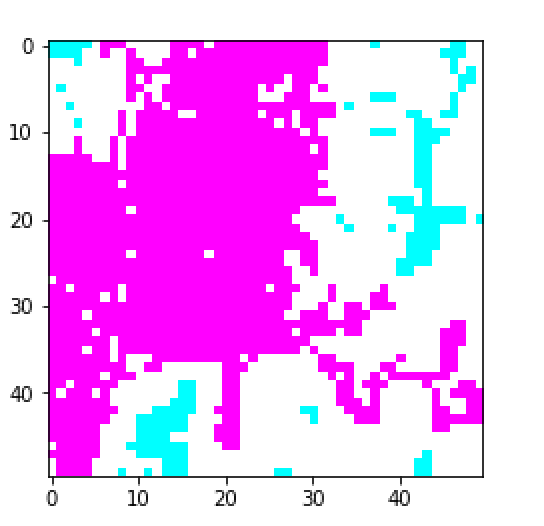
1. If the index range of the cell reaches the edges, we use **upper bound** and **lower bound** for the index ranges. **Upper bound** takes the minimum of **the length of outer lists** and (**x+nearby+1**) for outer lists, and **length of inner lists** and **(y+nearby+1)** for the inner lists. Lower bound takes the maximum of **0** and **(x-nearby)** for outer lists, and **0** and **(y-nearby)** for inner lists.

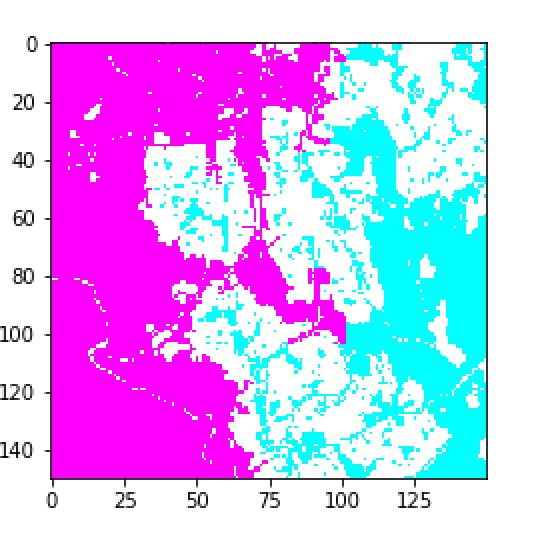
Question 5

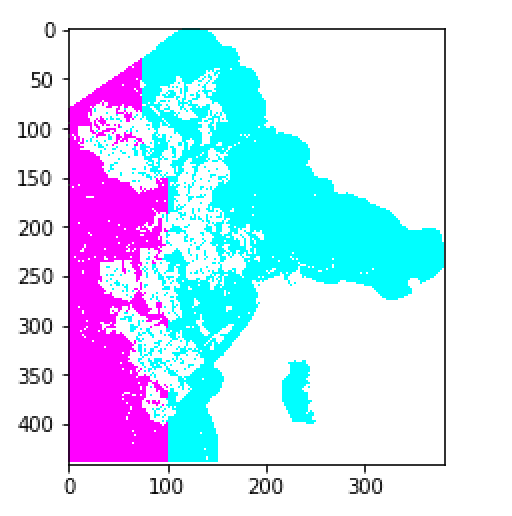
1. For **anu** dataset: 60 steps

For **south** dataset: 100 steps

For **act** dataset: 100 steps

ANU

South

ACT

1. The spread of fire I simulate covers much more space than the real 2013 bushfire, and the areas the real 2013 bushfire cover are all covered by my map.

Question 6

1. ﻿For **anu** dataset: ﻿0.12206216830932524

For **south** dataset : ﻿0.9074335782928208

For **act** dataset: ﻿0.8606832994312756

1. For **anu** dataset, the real bush fire is just the same as the initial one, so the percentage of similarity is pretty low.

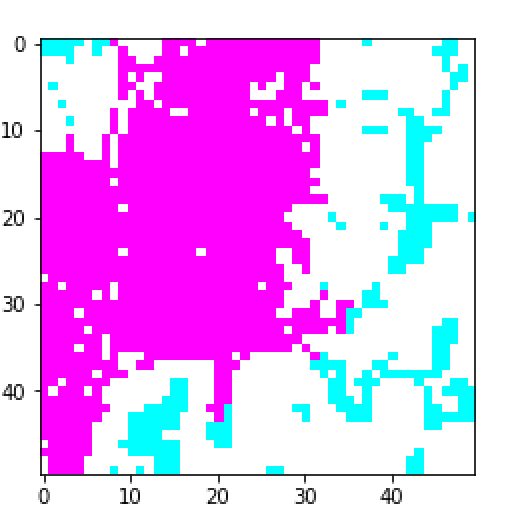
For the other datasets, the areas of fire spread I simulate are just a little larger than the real bush fire. Thus, the percentage of similarity is high.

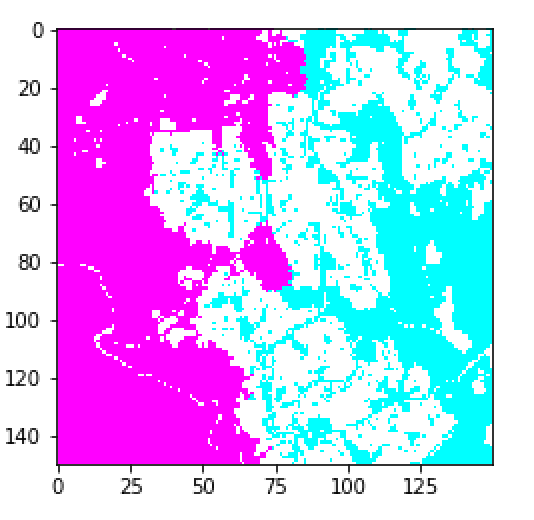
Question 7

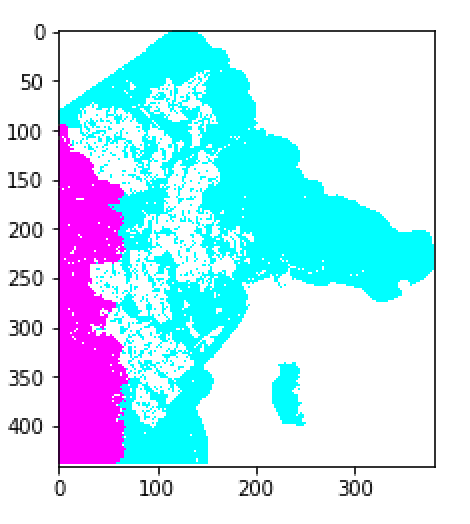
1. For **anu** dataset: 60 steps

For **south** dataset: 100 steps

For **act** dataset: 100 steps

Anu

South

ACT

1. By qualitatively, for **anu** dataset, the simulation is quite different because the fire spreading I simulate coves almost half of the map.

For **south** dataset, most parts of the map are the same except for the top and the centre part. On the top of my simulation. The bush fire I spread covers more areas than the real one, while in the centre part, the real bush fire extends more areas than my simulation.

For **act** dataset, the similarity is that both of the maps cover the left part of the whole map. However, my simulation covers more areas vertically while the real bush fire covers more areas horizontally.

By quantitatively, for **anu** dataset, the percentage of similarity is ﻿﻿**0.19636087945413191**, which is still pretty low.

For **south** dataset, ﻿the percentage is ﻿**0.926088185415489** which is very close to 1.

For **act** dataset, the percentage is **﻿﻿0.9270483069727539** which is high.

1. I think my simulation is realistic. Compared with the simple simulation with same steps of fire spreading, the similarity of percentage is higher.