**Notebook for transforming raw cpdata to Mergable data**

**Filter cpdata.csv to MergeFileCrop.cv**

**Filter fertilizer.csv to MergerFileFert.csv**

**import** pandas **as** pd

In [2]:

*# Reading the data*

crop\_data\_path **=** '../Data-raw/cpdata.csv'

fertilizer\_data\_path **=** '../Data-raw/Fertilizer.csv'

crop **=** pd**.**read\_csv(crop\_data\_path)

fert **=** pd**.**read\_csv(fertilizer\_data\_path)

In [3]:

crop**.**head()

Out[3]:

|  | **temperature** | **humidity** | **ph** | **rainfall** | **label** |
| --- | --- | --- | --- | --- | --- |
| **0** | 20.879744 | 82.002744 | 6.502985 | 202.935536 | rice |
| **1** | 21.770462 | 80.319644 | 7.038096 | 226.655537 | rice |
| **2** | 23.004459 | 82.320763 | 7.840207 | 263.964248 | rice |
| **3** | 26.491096 | 80.158363 | 6.980401 | 242.864034 | rice |
| **4** | 20.130175 | 81.604873 | 7.628473 | 262.717340 | rice |

In [4]:

fert**.**head()

Out[4]:

|  | **Unnamed: 0** | **Crop** | **N** | **P** | **K** | **pH** |
| --- | --- | --- | --- | --- | --- | --- |
| **0** | 0 | Rice | 80 | 40 | 40 | 5.5 |
| **1** | 1 | Jowar(Sorghum) | 80 | 40 | 40 | 5.5 |
| **2** | 2 | Barley(JAV) | 70 | 40 | 45 | 5.5 |
| **3** | 3 | Maize | 80 | 40 | 20 | 5.5 |
| **4** | 4 | Ragi( naachnnii) | 50 | 40 | 20 | 5.5 |

In [5]:

*# Function for lowering the cases*

**def** change\_case(i):

i **=** i**.**replace(" ", "")

i **=** i**.**lower()

**return** i

In [6]:

fert['Crop'] **=** fert['Crop']**.**apply(change\_case)

crop['label'] **=** crop['label']**.**apply(change\_case)

In [7]:

*#make some changes in ferttilizer dataset*

fert['Crop'] **=** fert['Crop']**.**replace('mungbeans','mungbean')

fert['Crop'] **=** fert['Crop']**.**replace('lentils(masoordal)','lentil')

fert['Crop'] **=** fert['Crop']**.**replace('pigeonpeas(toordal)','pigeonpeas')

fert['Crop'] **=** fert['Crop']**.**replace('mothbean(matki)','mothbeans')

fert['Crop'] **=** fert['Crop']**.**replace('chickpeas(channa)','chickpea')

In [8]:

crop**.**head()

Out[8]:

|  | **temperature** | **humidity** | **ph** | **rainfall** | **label** |
| --- | --- | --- | --- | --- | --- |
| **0** | 20.879744 | 82.002744 | 6.502985 | 202.935536 | rice |
| **1** | 21.770462 | 80.319644 | 7.038096 | 226.655537 | rice |
| **2** | 23.004459 | 82.320763 | 7.840207 | 263.964248 | rice |
| **3** | 26.491096 | 80.158363 | 6.980401 | 242.864034 | rice |
| **4** | 20.130175 | 81.604873 | 7.628473 | 262.717340 | rice |

In [9]:

crop**.**tail()

Out[9]:

|  | **temperature** | **humidity** | **ph** | **rainfall** | **label** |
| --- | --- | --- | --- | --- | --- |
| **3095** | 25.287846 | 89.636679 | 6.765095 | 58.286977 | watermelon |
| **3096** | 26.638386 | 84.695469 | 6.189214 | 48.324286 | watermelon |
| **3097** | 25.331045 | 84.305338 | 6.904242 | 41.532187 | watermelon |
| **3098** | 26.897502 | 83.892415 | 6.463271 | 43.971937 | watermelon |
| **3099** | 26.986037 | 89.413849 | 6.260839 | 58.548767 | watermelon |

In [10]:

crop\_names **=** crop['label']**.**unique()

crop\_names

Out[10]:

array(['rice', 'wheat', 'mungbean', 'tea', 'millet', 'maize', 'lentil',

'jute', 'coffee', 'cotton', 'groundnut', 'peas', 'rubber',

'sugarcane', 'tobacco', 'kidneybeans', 'mothbeans', 'coconut',

'blackgram', 'adzukibeans', 'pigeonpeas', 'chickpea', 'banana',

'grapes', 'apple', 'mango', 'muskmelon', 'orange', 'papaya',

'pomegranate', 'watermelon'], dtype=object)

In [11]:

fert**.**head()

Out[11]:

|  | **Unnamed: 0** | **Crop** | **N** | **P** | **K** | **pH** |
| --- | --- | --- | --- | --- | --- | --- |
| **0** | 0 | rice | 80 | 40 | 40 | 5.5 |
| **1** | 1 | jowar(sorghum) | 80 | 40 | 40 | 5.5 |
| **2** | 2 | barley(jav) | 70 | 40 | 45 | 5.5 |
| **3** | 3 | maize | 80 | 40 | 20 | 5.5 |
| **4** | 4 | ragi(naachnnii) | 50 | 40 | 20 | 5.5 |

In [12]:

**del** fert['Unnamed: 0']

In [13]:

crop\_names\_from\_fert **=** fert['Crop']**.**unique()

crop\_names\_from\_fert

Out[13]:

array(['rice', 'jowar(sorghum)', 'barley(jav)', 'maize',

'ragi(naachnnii)', 'chickpea', 'frenchbeans(farasbi)',

'favabeans(papdi-val)', 'limabeans(pavta)', 'clusterbeans(gavar)',

'soyabean', 'blackeyedbeans(chawli)', 'kidneybeans', 'pigeonpeas',

'mothbeans', 'mungbean', 'greenpeas', 'horsegram(kulthi)',

'blackgram', 'rapeseed(mohri)', 'corianderseeds', 'mustardseeds',

'sesameseed', 'cuminseeds', 'lentil', 'brinjal', 'beetroot',

'bittergourd', 'bottlegourd', 'capsicum', 'cabbage', 'carrot',

'cauliflower', 'cucumber', 'corianderleaves', 'curryleaves',

'drumstick–moringa', 'chili', 'ladyfinger', 'mushroom', 'onion',

'potato', 'pumpkin', 'radish', 'olive', 'sweetpotato',

'fenugreekleaf(methi)', 'spinach', 'ridgegourd',

'gooseberry(amla)', 'jambun(syzygiumcumini)',

'ziziphusmauritiana(bor)', 'garciniaindica(kokam)', 'tamarind',

'tapioca(suran)', 'garlic', 'lemon', 'tomato', 'ashgourd',

'pineapple', 'pomegranate', 'banana', 'mango', 'grapes',

'jackfruit', 'guava', 'watermelon', 'muskmelon', 'apricot',

'apple', 'chickoo', 'custardapple', 'dates', 'figs', 'orange',

'papaya', 'aniseed', 'asafoetida', 'bayleaf', 'blackpepper',

'cardamom', 'cinnamon', 'cloves', 'jaiphal(nutmeg)', 'ginger',

'turmeric', 'cashewnuts', 'raisins', 'coconut', 'almondnut',

'arecanut', 'pistachionut', 'lemongrass', 'cotton', 'jute',

'coffee', 'sunflower'], dtype=object)

In [14]:

**for** i **in** crop\_names\_from\_fert:

print(crop[crop['label'] **==** i])

temperature humidity ph rainfall label

0 20.879744 82.002744 6.502985 202.935536 rice

1 21.770462 80.319644 7.038096 226.655537 rice

2 23.004459 82.320763 7.840207 263.964248 rice

3 26.491096 80.158363 6.980401 242.864034 rice

4 20.130175 81.604873 7.628473 262.717340 rice

.. ... ... ... ... ...

95 22.683191 83.463583 6.604993 194.265172 rice

96 21.533463 82.140041 6.500343 295.924880 rice

97 21.408658 83.329319 5.935745 287.576694 rice

98 26.543481 84.673536 7.072656 183.622266 rice

99 23.359054 83.595123 5.333323 188.413665 rice

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

500 22.613600 63.690706 5.749914 87.759539 maize

501 26.100184 71.574769 6.931757 102.266244 maize

502 23.558821 71.593514 6.657965 66.719955 maize

503 19.972160 57.682729 6.596061 60.651715 maize

504 18.478913 62.695039 5.970458 65.438354 maize

.. ... ... ... ... ...

595 18.928519 72.800861 6.158860 82.341629 maize

596 23.305468 63.246480 6.385684 108.760300 maize

597 18.748267 62.498785 6.417820 70.234016 maize

598 19.742133 59.662631 6.381202 65.508614 maize

599 25.730444 70.747393 6.877869 98.737713 maize

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

2100 17.024985 16.988612 7.485996 88.551231 chickpea

2101 19.020613 17.131591 6.920251 79.926981 chickpea

2102 17.887765 15.405897 5.996932 68.549329 chickpea

2103 18.868056 15.658092 6.391174 88.510490 chickpea

2104 18.369526 19.563810 7.152811 79.263577 chickpea

... ... ... ... ... ...

2195 17.341502 18.756263 8.861480 67.954543 chickpea

2196 17.437327 14.338474 7.861128 73.092670 chickpea

2197 18.897802 19.761829 7.452671 69.095125 chickpea

2198 18.591908 14.779596 7.168096 89.609825 chickpea

2199 18.315615 15.361435 7.263119 81.787105 chickpea

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

1500 17.136928 20.595417 5.685972 128.256862 kidneybeans

1501 19.634743 18.907056 5.759237 106.359818 kidneybeans

1502 22.913502 21.339531 5.873172 109.225556 kidneybeans

1503 16.433403 24.240459 5.926677 140.371781 kidneybeans

1504 22.139747 23.022511 5.955617 76.641283 kidneybeans

... ... ... ... ... ...

1595 20.109938 23.223238 5.595032 73.363865 kidneybeans

1596 23.605066 21.905396 5.525905 100.597873 kidneybeans

1597 19.731369 24.894874 5.819404 84.063541 kidneybeans

1598 20.934099 21.189301 5.562202 133.191442 kidneybeans

1599 18.782263 20.247683 5.630665 104.257072 kidneybeans

[100 rows x 5 columns]

temperature humidity ph rainfall label

2000 36.512684 57.928872 6.031608 122.653969 pigeonpeas

2001 36.891637 62.731782 5.269085 163.726655 pigeonpeas

2002 29.235405 59.389676 5.985793 103.330180 pigeonpeas

2003 27.335349 43.357960 6.091863 142.330368 pigeonpeas

2004 21.064368 55.469859 5.624731 184.622671 pigeonpeas

... ... ... ... ... ...

2095 29.892866 66.353751 6.931925 198.140300 pigeonpeas

2096 29.377356 44.822946 6.842744 172.401680 pigeonpeas

2097 29.650529 42.898332 6.876573 186.922605 pigeonpeas

2098 19.542849 66.347773 6.151029 173.110698 pigeonpeas

2099 20.046118 48.939056 4.567446 122.456420 pigeonpeas

[100 rows x 5 columns]

temperature humidity ph rainfall label

1600 27.910952 64.709306 3.692864 32.678919 mothbeans

1601 27.322206 51.278688 4.371746 36.503791 mothbeans

1602 28.660242 59.318912 8.399136 36.926297 mothbeans

1603 29.029553 61.093875 8.840656 72.980166 mothbeans

1604 27.780315 54.650300 8.153023 32.050253 mothbeans

... ... ... ... ... ...

1695 29.337434 49.003231 8.914075 42.440543 mothbeans

1696 27.965837 61.349001 8.639586 70.104721 mothbeans

1697 24.868040 48.275320 8.621514 63.918765 mothbeans

1698 25.876823 45.963419 5.838509 38.532547 mothbeans

1699 31.019636 49.976752 3.532009 32.812965 mothbeans

[100 rows x 5 columns]

temperature humidity ph rainfall label

200 27.433294 87.805077 7.185301 54.733676 mungbean

201 28.334043 80.772760 7.034214 38.797641 mungbean

202 27.014704 84.342627 6.635969 55.296354 mungbean

203 28.174327 81.045548 6.828187 36.357207 mungbean

204 29.878881 87.327612 6.890780 44.752159 mungbean

.. ... ... ... ... ...

295 28.727527 89.127604 7.069748 58.529743 mungbean

296 27.956397 83.527060 6.921994 43.257268 mungbean

297 28.174587 83.696593 6.770955 37.246465 mungbean

298 28.776535 86.691340 6.983130 56.124432 mungbean

299 28.438097 83.489914 6.267684 52.554700 mungbean

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

1800 29.484400 63.199153 7.454532 71.890907 blackgram

1801 26.734340 68.139997 7.040056 67.150964 blackgram

1802 26.272744 62.288149 7.418651 70.232076 blackgram

1803 34.036792 67.211138 6.501869 73.235736 blackgram

1804 28.036441 65.066017 6.814411 72.495077 blackgram

... ... ... ... ... ...

1895 33.369844 65.677182 6.874142 64.895175 blackgram

1896 31.434506 62.993035 7.760618 64.776515 blackgram

1897 27.716783 63.291034 6.781842 68.565080 blackgram

1898 32.639187 61.300905 7.326980 61.838761 blackgram

1899 32.747739 67.779546 7.453975 63.377844 blackgram

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

600 28.051536 63.498022 7.604110 43.357954 lentil

601 19.440843 63.277715 7.728832 46.831301 lentil

602 29.848231 60.638726 7.491217 46.804526 lentil

603 21.363838 69.923759 6.633865 46.635286 lentil

604 26.286639 68.519667 7.324863 46.138330 lentil

.. ... ... ... ... ...

695 23.052764 60.424786 7.011121 52.602853 lentil

696 21.658458 63.583371 6.280726 38.076594 lentil

697 26.250703 67.627797 7.621495 40.810630 lentil

698 20.971953 63.831799 7.630424 53.102079 lentil

699 23.897364 66.321020 7.802212 40.745368 lentil

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

2900 24.559816 91.635362 5.922936 111.968462 pomegranate

2901 19.656901 89.937010 5.937650 108.045893 pomegranate

2902 18.783596 87.402477 6.804781 102.518476 pomegranate

2903 24.146963 94.511066 6.424671 110.231663 pomegranate

2904 22.445813 89.901470 6.738016 109.390600 pomegranate

... ... ... ... ... ...

2995 20.002190 85.836182 7.116539 112.337046 pomegranate

2996 19.851393 89.807323 6.430163 102.818636 pomegranate

2997 21.254336 92.650589 7.159521 106.278467 pomegranate

2998 23.653741 93.326575 6.431266 109.807618 pomegranate

2999 23.884048 86.206138 6.082572 108.312179 pomegranate

[100 rows x 5 columns]

temperature humidity ph rainfall label

2200 29.367924 76.249001 6.149934 92.828409 banana

2201 27.333690 83.676752 5.849076 101.049479 banana

2202 27.400536 82.962213 6.276800 104.937800 banana

2203 29.315908 80.115857 5.926825 90.109781 banana

2204 26.054330 79.396545 5.519088 113.229737 banana

... ... ... ... ... ...

2295 27.359116 84.546250 6.387431 90.812505 banana

2296 28.010680 76.528081 5.891414 103.704078 banana

2297 28.672089 82.207936 5.725419 94.379875 banana

2298 27.345851 78.487383 6.281070 92.155243 banana

2299 29.507046 78.205856 5.507642 98.125658 banana

[100 rows x 5 columns]

temperature humidity ph rainfall label

2500 29.737700 47.548852 5.954627 90.095869 mango

2501 33.556956 53.729798 4.757115 98.675276 mango

2502 27.003155 47.675254 5.699587 95.851183 mango

2503 33.561502 45.535566 5.977414 95.705259 mango

2504 35.898556 54.259642 6.430139 92.197217 mango

... ... ... ... ... ...

2595 31.484517 48.779263 4.525722 93.172220 mango

2596 27.698193 51.415932 5.403908 100.772070 mango

2597 30.412358 52.481006 6.621624 93.923759 mango

2598 32.177520 54.013527 6.207496 91.887661 mango

2599 32.611261 47.749165 5.418475 91.101908 mango

[100 rows x 5 columns]

temperature humidity ph rainfall label

2300 29.996772 81.541566 6.112306 67.125345 grapes

2301 30.728040 82.426141 6.092242 68.381355 grapes

2302 32.445778 83.885049 5.896343 68.739325 grapes

2303 37.465668 80.659687 6.155261 66.838723 grapes

2304 22.032962 83.743728 5.732454 65.344408 grapes

... ... ... ... ... ...

2395 9.851243 80.226317 5.965379 68.428024 grapes

2396 24.972561 82.728287 6.476758 66.700163 grapes

2397 27.237083 82.945733 6.224543 70.425089 grapes

2398 18.706791 83.479529 6.209928 66.596449 grapes

2399 9.949929 82.551390 5.841138 66.008176 grapes

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

3000 26.473302 80.922544 6.283818 53.657426 watermelon

3001 25.187800 83.446217 6.818261 46.874209 watermelon

3002 25.299547 81.775276 6.376201 57.041471 watermelon

3003 24.746313 88.308663 6.581588 57.958261 watermelon

3004 26.587407 81.325632 6.932740 41.875400 watermelon

... ... ... ... ... ...

3095 25.287846 89.636679 6.765095 58.286977 watermelon

3096 26.638386 84.695469 6.189214 48.324286 watermelon

3097 25.331045 84.305338 6.904242 41.532187 watermelon

3098 26.897502 83.892415 6.463271 43.971937 watermelon

3099 26.986037 89.413849 6.260839 58.548767 watermelon

[100 rows x 5 columns]

temperature humidity ph rainfall label

2600 27.578269 94.118782 6.776533 28.082532 muskmelon

2601 27.820548 93.035552 6.528404 26.324055 muskmelon

2602 29.099104 94.222378 6.750146 22.524973 muskmelon

2603 28.049436 90.831307 6.562833 20.762230 muskmelon

2604 29.916906 94.556956 6.117530 28.160572 muskmelon

... ... ... ... ... ...

2695 29.527531 94.574594 6.700338 21.135457 muskmelon

2696 28.504164 93.468065 6.565313 24.200072 muskmelon

2697 28.895786 94.789930 6.286515 23.036250 muskmelon

2698 27.049275 91.382173 6.448062 23.657475 muskmelon

2699 28.960179 91.695322 6.585873 24.745820 muskmelon

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

2400 22.750888 90.694892 5.521467 110.431786 apple

2401 23.849401 94.348150 6.133221 114.051250 apple

2402 22.608010 94.589006 6.226290 116.039659 apple

2403 21.186674 91.134357 6.321152 122.233323 apple

2404 23.410447 91.699133 5.587906 116.077793 apple

... ... ... ... ... ...

2495 23.805938 92.488795 5.889481 119.633555 apple

2496 22.319441 90.851744 5.732758 100.117344 apple

2497 22.144641 93.825674 6.400321 120.631078 apple

2498 23.651676 94.505288 6.496934 115.361127 apple

2499 22.169395 90.271856 6.229499 124.468311 apple

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

2700 15.781442 92.510777 6.354007 119.035002 orange

2701 26.030973 91.508193 7.511755 101.284774 orange

2702 13.360506 91.356082 7.335158 111.226688 orange

2703 18.879577 92.043045 7.813917 114.665951 orange

2704 29.477417 91.578029 7.129137 111.172750 orange

... ... ... ... ... ...

2795 32.717485 90.546083 7.656978 113.328978 orange

2796 25.162966 92.547360 7.105905 114.311720 orange

2797 27.681673 94.473169 7.199106 113.999515 orange

2798 21.350934 90.949297 7.871063 107.086209 orange

2799 11.698946 93.256389 7.566166 103.200599 orange

[100 rows x 5 columns]

temperature humidity ph rainfall label

2800 35.214628 91.497251 6.793245 243.074507 papaya

2801 42.394134 90.790281 6.576261 88.466075 papaya

2802 38.419163 91.142204 6.751453 119.265388 papaya

2803 35.332949 92.115086 6.560743 235.613359 papaya

2804 42.923253 90.076005 6.938313 196.240824 papaya

... ... ... ... ... ...

2895 40.102077 94.351102 6.979102 149.119999 papaya

2896 38.589545 91.580765 6.825665 102.270823 papaya

2897 41.313301 91.150880 6.617067 239.742755 papaya

2898 37.035519 91.794302 6.551893 188.518142 papaya

2899 23.012402 91.073555 6.598860 208.335798 papaya

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

1700 26.762749 92.860569 6.420019 224.590366 coconut

1701 25.612944 94.313884 5.740055 224.320676 coconut

1702 28.130115 95.648076 5.686973 151.076190 coconut

1703 25.028872 91.537209 6.293662 179.824894 coconut

1704 27.797977 99.645730 6.381975 181.694228 coconut

... ... ... ... ... ...

1795 28.435729 95.884041 5.665785 203.928371 coconut

1796 28.940997 93.001090 5.764615 191.772309 coconut

1797 26.454887 93.450426 5.901496 149.222026 coconut

1798 25.794905 93.841506 5.779033 152.423871 coconut

1799 26.931419 98.803136 5.671549 166.571288 coconut

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

temperature humidity ph rainfall label

900 24.402289 79.197320 7.231325 90.802236 cotton

901 23.095956 84.862757 6.925412 71.295811 cotton

902 23.965635 76.976967 7.633437 90.756167 cotton

903 24.887381 75.621372 6.827355 89.760504 cotton

904 25.362438 83.632761 6.176716 88.436189 cotton

.. ... ... ... ... ...

995 22.107190 78.583201 6.364730 74.941366 cotton

996 23.038140 76.110215 6.913679 91.496975 cotton

997 24.547953 75.397527 7.766260 63.880799 cotton

998 23.738680 75.775038 7.556064 76.636692 cotton

999 22.318719 83.861300 7.288377 65.357470 cotton

[100 rows x 5 columns]

temperature humidity ph rainfall label

700 25.524690 72.248508 6.002525 151.886997 jute

701 26.591050 82.941641 6.033485 161.247000 jute

702 25.297818 86.887054 7.121934 196.624951 jute

703 25.721009 88.165136 6.207460 175.608670 jute

704 23.584193 72.004608 6.090060 190.424216 jute

.. ... ... ... ... ...

795 23.874845 86.792613 6.718725 177.514731 jute

796 23.928879 88.071123 6.880205 154.660874 jute

797 24.814412 81.686889 6.861069 190.788639 jute

798 24.447439 82.286484 6.769346 190.968489 jute

799 26.574217 73.819949 7.261581 159.322307 jute

[100 rows x 5 columns]

temperature humidity ph rainfall label

800 26.333780 57.364700 7.261314 191.654941 coffee

801 26.452885 55.322227 7.235070 144.686134 coffee

802 25.708227 52.886671 7.189156 136.732509 coffee

803 24.128325 56.181077 6.431900 147.275782 coffee

804 23.443723 60.395233 6.423211 122.210325 coffee

.. ... ... ... ... ...

895 26.774637 66.413269 6.780064 177.774507 coffee

896 27.417112 56.636362 6.086922 127.924610 coffee

897 24.131797 67.225123 6.362608 173.322839 coffee

898 26.272418 52.127394 6.758793 127.175293 coffee

899 23.603016 60.396475 6.779833 140.937041 coffee

[100 rows x 5 columns]

Empty DataFrame

Columns: [temperature, humidity, ph, rainfall, label]

Index: []

In [15]:

crop['label']

Out[15]:

0 rice

1 rice

2 rice

3 rice

4 rice

...

3095 watermelon

3096 watermelon

3097 watermelon

3098 watermelon

3099 watermelon

Name: label, Length: 3100, dtype: object

In [16]:

extract\_labels **=** []

**for** i **in** crop\_names\_from\_fert:

**if** i **in** crop\_names:

extract\_labels**.**append(i)

In [17]:

*# using extract labesl on crop to get all the data related to those labels*

new\_crop **=** pd**.**DataFrame(columns **=** crop**.**columns)

new\_fert **=** pd**.**DataFrame(columns **=** fert**.**columns)

In [18]:

**for** label **in** extract\_labels:

new\_crop **=** new\_crop**.**append(crop[crop['label'] **==** label])

In [20]:

**for** label **in** extract\_labels:

new\_fert **=** new\_fert**.**append(fert[fert['Crop'] **==** label]**.**iloc[0])

In [21]:

new\_crop

Out[21]:

|  | **temperature** | **humidity** | **ph** | **rainfall** | **label** |
| --- | --- | --- | --- | --- | --- |
| **0** | 20.879744 | 82.002744 | 6.502985 | 202.935536 | rice |
| **1** | 21.770462 | 80.319644 | 7.038096 | 226.655537 | rice |
| **2** | 23.004459 | 82.320763 | 7.840207 | 263.964248 | rice |
| **3** | 26.491096 | 80.158363 | 6.980401 | 242.864034 | rice |
| **4** | 20.130175 | 81.604873 | 7.628473 | 262.717340 | rice |
| **...** | ... | ... | ... | ... | ... |
| **895** | 26.774637 | 66.413269 | 6.780064 | 177.774507 | coffee |
| **896** | 27.417112 | 56.636362 | 6.086922 | 127.924610 | coffee |
| **897** | 24.131797 | 67.225123 | 6.362608 | 173.322839 | coffee |
| **898** | 26.272418 | 52.127394 | 6.758793 | 127.175293 | coffee |
| **899** | 23.603016 | 60.396475 | 6.779833 | 140.937041 | coffee |

2200 rows × 5 columns

In [22]:

new\_fert

Out[22]:

|  | **Crop** | **N** | **P** | **K** | **pH** |
| --- | --- | --- | --- | --- | --- |
| **0** | rice | 80 | 40 | 40 | 5.5 |
| **3** | maize | 80 | 40 | 20 | 5.5 |
| **5** | chickpea | 40 | 60 | 80 | 5.5 |
| **12** | kidneybeans | 20 | 60 | 20 | 5.5 |
| **13** | pigeonpeas | 20 | 60 | 20 | 5.5 |
| **14** | mothbeans | 20 | 40 | 20 | 5.5 |
| **15** | mungbean | 20 | 40 | 20 | 5.5 |
| **18** | blackgram | 40 | 60 | 20 | 5.0 |
| **24** | lentil | 20 | 60 | 20 | 5.5 |
| **60** | pomegranate | 20 | 10 | 40 | 5.5 |
| **61** | banana | 100 | 75 | 50 | 6.5 |
| **62** | mango | 20 | 20 | 30 | 5.0 |
| **63** | grapes | 20 | 125 | 200 | 4.0 |
| **66** | watermelon | 100 | 10 | 50 | 5.5 |
| **67** | muskmelon | 100 | 10 | 50 | 5.5 |
| **69** | apple | 20 | 125 | 200 | 6.5 |
| **74** | orange | 20 | 10 | 10 | 4.0 |
| **75** | papaya | 50 | 50 | 50 | 6.0 |
| **88** | coconut | 20 | 10 | 30 | 5.0 |
| **93** | cotton | 120 | 40 | 20 | 5.5 |
| **94** | jute | 80 | 40 | 40 | 5.5 |
| **95** | coffee | 100 | 20 | 30 | 5.5 |

In [23]:

new\_crop**.**to\_csv('../Data-raw/MergeFileCrop.csv')

new\_fert**.**to\_csv('../Data-raw/Fertilizer