Visualizing distances between the users

- Loading the data

```
csvReviews = Take[Drop[Import["/tmp/reviews_all.csv", "CSV"], 1], 100 000]
```

Out[ ]=

```
{{5, 213 719, 16 007 298}, {5, 198 763, 16 007 298}, {5, 239 184, 16 007 298},
 {5, 238 654, 16 007 298}, {5, 230 116, 16 007 298}, {5, 245 403, 16 007 298},
 {5, 12 141, 16 007 298}, {5, 23 032, 16 007 298}, {5, 269 891, 16 007 298},
 {5, 194 425, 16 007 298}, {5, 223 085, 16 007 298}, {5, 278 099, 16 007 298},
 {5, 238 162, 16 007 298}, {5, 86 331, 16 007 298}, {5, 230 281, 16 007 298},
 {5, 218 790, 16 007 298}, {5, 15 821, 16 007 298}, {5, 256 018, 16 007 298},
 {5, 67 789, 16 007 298}, {5, 15 931, 16 007 298}, {5, 255 335, 16 007 298},
 {5, 215 014, 16 007 298}, ··· 99 957 ···, {5, 121 411, 1 471 732}, {5, 216 190, 1 471 732},
 {5, 26 692, 1 471 732}, {5, 74 437, 1 471 732}, {5, 16 504, 1 471 732},
 {5, 13 206, 1 471 732}, {5, 25 203, 1 471 732}, {5, 44 839, 1 471 732},
 {5, 13 651, 1 471 732}, {5, 24 332, 1 471 732}, {2, 127 501, 1 471 732}, {3, 12 471, 359 494},
 {5, 143 069, 9 990 898}, {5, 24 069, 9 990 898}, {5, 229 508, 9 990 898},
 {5, 236 942, 2 447 592}, {5, 17 066, 11 403 286}, {5, 24 473, 11 403 286},
 {5, 26 257, 11 403 286}, {4, 203 769, 7 510 857}, {5, 25 451, 14 048 996}}
```

large output | show less | show more | show all | set size limit ...

- Getting a list of all the authors

In[ ]:=

```
authors = DeleteDuplicates [#[[3]] & /@ csvReviews ]
```

- Associating every author with its reviews

In[ ]:=
```
users = Table[author → DeleteCases[Table[If[review[[3]] == author, Drop[review, {3}]],
      {review, csvReviews}], Null], {author, authors}]
```

Out[ ]=
```
{16 007 298 → {{5, 213 719}, {5, 198 763}, {5, 239 184}, {5, 238 654}, {5, 230 116},
    {5, 245 403}, {5, 12 141}, {5, 23 032}, {5, 269 891}, {5, 194 425}, {5, 223 085},
    {5, 278 099}, {5, 238 162}, {5, 86 331}, {5, 230 281}, {5, 218 790}, {5, 15 821},
    {5, 256 018}, {5, 67 789}, {5, 15 931}, {5, 255 335}, {5, 215 014}, {5, 213 505},
    {5, 233 923}, {5, 76 278}, {5, 15 413}, {5, 235 589}, {5, 13 325}, {5, 55 451},
    {5, 218 180}, {5, 142 481}, {5, 25 331}, {5, 19 307}, {5, 255 362}, {5, 255 361},
    {5, 11 693}, {5, 209 990}, {5, 9115}, ··· 532 ···, {5, 56 927}, {5, 14 316}, {5, 143 082},
    {5, 214 488}, {5, 77 156}, {5, 8493}, {5, 9044}, {5, 51 301}, {5, 16 120}, {5, 152 239},
    {5, 10 638}, {5, 229 272}, {5, 98 390}, {5, 11 406}, {5, 14 556}, {5, 23 534},
    {5, 20 144}, {5, 25 971}, {5, 229 923}, {5, 146 100}, {5, 221 000}, {5, 9249},
    {5, 56 999}, {5, 139 580}, {5, 16 592}, {5, 221 988}, {5, 20 021}, {5, 15 057},
    {5, 24 132}, {5, 10 813}, {5, 17 165}, {5, 10 222}, {5, 19 247}, {5, 9471}, {5, 10 687},
    {5, 10 402}, {5, 31 965}, {5, 241 430}}, ··· 4938 ···, 14 048 996 → ··· 1 ···}
```

large output | show less | show more | show all | set size limit...

■ Fitting a dimension reducer

In[ ]:=
```
dimensionReducer = DimensionReduction[
  Flatten[Join[Values[#] // First, Table[review, {review, Values[#]}]]] & /@ users,
  Method → "Linear"]
```

Out[ ]=
```
DimensionReducerFunction [ ⊞  Input type: NumericalSequence
                                                  Output dimension: 2 ]
```

■ Applying the dimensionality reduction to our data

In[ ]:=
```
users2D = (Keys[#] → dimensionReducer [
        Flatten[Join[Values[#] // First , Table[review , {review , Values[#]}]]]]) & /@ users
```

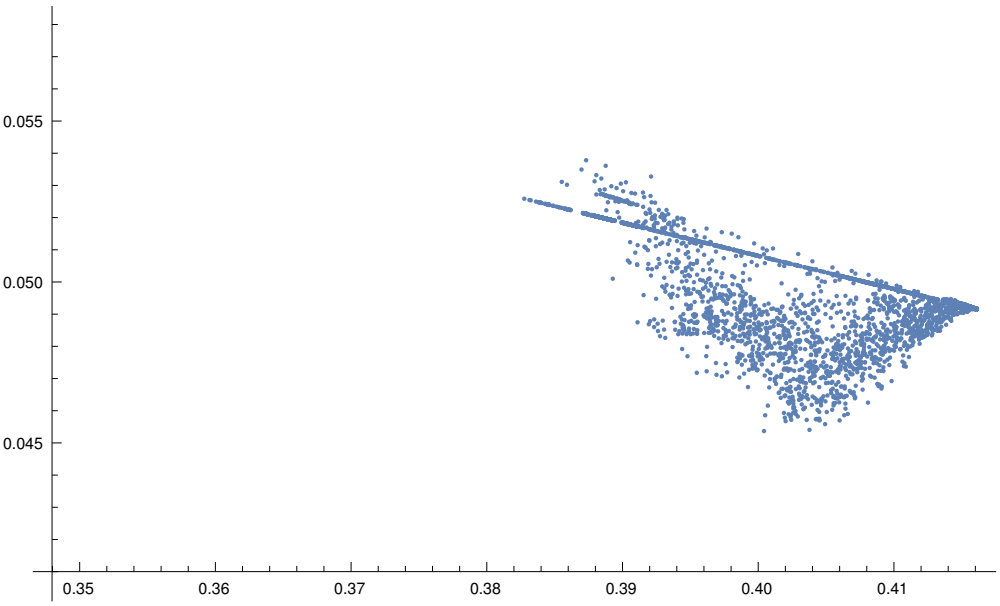
Out[ ]=
```
{16 007 298 → {0.155663 , -0.16671}, 2 012 871 → {-0.0731174 , -0.337524},
 9 020 456 → {0.413375 , 0.0487773}, 1 922 840 → {0.411156 , 0.0484272},
 286 648 → {-0.966066 , -0.740453}, 3 672 150 → {0.390562 , 0.0512373},
 731 327 → {0.401412 , 0.0482864}, 430 608 → {0.408628 , 0.0482791},
 594 098 → {0.40627 , 0.0501646}, 775 411 → {-4.95718 , -0.28721}, ··· 4920 ··· ,
 15 471 065 → {0.391145 , 0.0517223}, 15 777 604 → {0.400428 , 0.0453685},
 5 901 381 → {0.415527 , 0.0492111}, 1 471 732 → {0.403905 , 0.0479741},
 359 494 → {0.415405 , 0.0492234}, 9 990 898 → {0.399426 , 0.0495313},
 2 447 592 → {0.388339 , 0.0520115}, 11 403 286 → {0.414395 , 0.0493765},
 7 510 857 → {0.392339 , 0.0515993}, 14 048 996 → {0.41384 , 0.0493848}}
```

large output | show less | show more | show all | set size limit ...

- Applying the dimensionality reduction to our data

In[ ]:=
```
ListPlot[Values[#] & /@ users2D]
```

Out[ ]=



Visualizing data about the users

- Loading the data

```
csvProfiles = Drop[Import["/tmp/profiles_all .csv", "CSV"], 1]
```

Out[ ]=

```
{{3 771 717 , cheri , , [none  selected], , },
 {9 020 456 , Mariwanna , , united  states , Maryland , Baltimore},
 {13 145 275 , Tyler  Bradley , , [none  selected], , },
 {6 613 609 , Terrie  Melcher  Murphy , , , , }, {16 194 924 , cheri , , [none  selected], , },
 {21 140 057 , cheri , , [none  selected], , }, {7 222 641 , toto858 , , [none  selected], , },
 {8 919 808 , 70sdonna , , [none  selected], , }, {6 294 607 , Margo  Pfeifer , , , , },
 {15 700 044 , Barb  Fredricks , , , , }, ··· 2 412 931 ··· , {17 473 495 , Ashlee  Jones , , , , },
 {16 560 257 , Gaby  Sunderland , , , , }, {17 340 542 , Austin  Leifi , , , , },
 {16 043 773 , Nate  James , , , , }, {20 477 386 , Rhianon  Solomon , , , , },
 {7 460 924 , Lora  Marie , , , , }, {21 829 435 , Laura  Kicker , , , , },
 {17 904 570 , Kelly  Dalessio , , , , }, {14 773 527 , Rhonda  Roberts , , , , }}
```

large output | show less | show more | show all | set size limit ...

■ Creating a list of all the countries

In[ ]:=

```
countries = Capitalize [
  DeleteCases [If[#⟦4⟧ == "[none  selected]" || #⟦4⟧ == "", "other/not selected ", #⟦4⟧] & /@
    csvProfiles , None], "TitleCase "]
```

Out[ ]=

```
{Other/Not  Selected , United  States , Other/Not  Selected , Other/Not  Selected ,
 Other/Not  Selected , Other/Not  Selected , Other/Not  Selected ,
 Other/Not  Selected , Other/Not  Selected , Other/Not  Selected , United  States ,
 United  States , Other/Not  Selected , Other/Not  Selected , Other/Not  Selected ,
 Other/Not  Selected , Other/Not  Selected , Other/Not  Selected , United  States ,
 Other/Not  Selected , ··· 2 412 910 ··· , Other/Not  Selected , Other/Not  Selected ,
 Other/Not  Selected , Other/Not  Selected , Other/Not  Selected ,
 Other/Not  Selected , Other/Not  Selected , Other/Not  Selected ,
 Other/Not  Selected , Other/Not  Selected , Other/Not  Selected , Other/Not  Selected ,
 Other/Not  Selected , Other/Not  Selected , Other/Not  Selected , Other/Not  Selected ,
 Other/Not  Selected , Other/Not  Selected , Other/Not  Selected , Other/Not  Selected}
```

large output | show less | show more | show all | set size limit ...

■ Extracting countries data

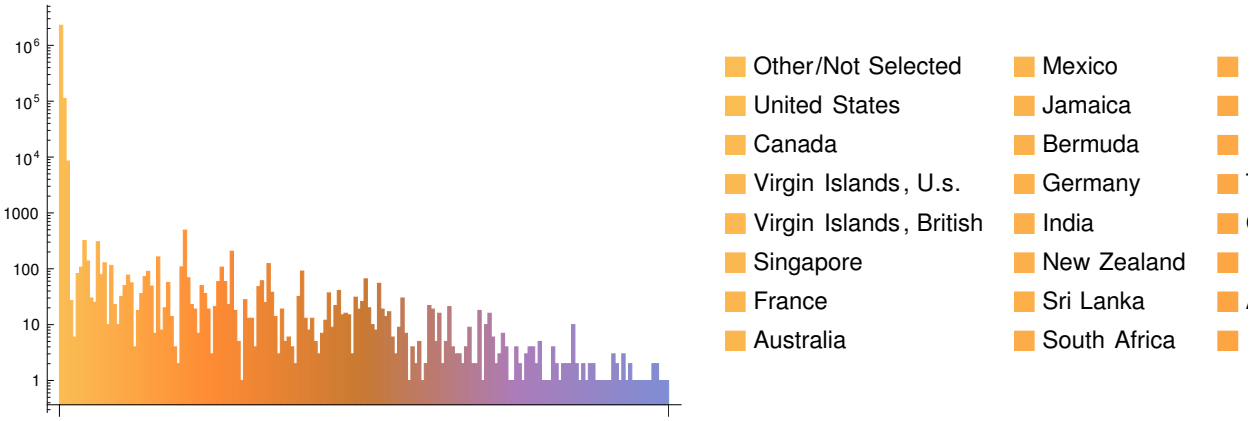
In[ ]:=

```
countriesFrequence  = (# → Count[countries , #]) & /@ DeleteDuplicates [countries ]
```

■ Plotting the regions

In[ ]:=
```
BarChart[{Values[countriesFrequence ]},
 ChartLegends → Keys[countriesFrequence ], ScalingFunctions → "Log"]
```

Out[ ]=



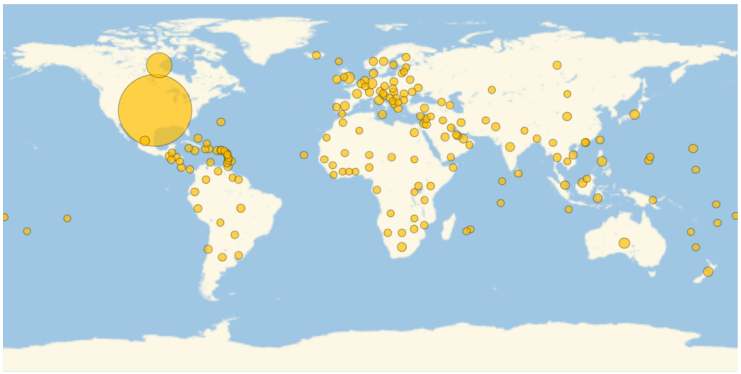
■ Calculate the percentages

```
percentages = DeleteCases[
  (Interpreter["Country"][#] → N[Association[countriesFrequence ][#] * 100 / Total[Values[countriesFrequence ]]]) & /@
   DeleteCases[Keys[countriesFrequence ], "Other/Not Selected"] Null]
```

■ Plot on a world map

```
GeoBubbleChart[percentages ]
```

Out[ ]=



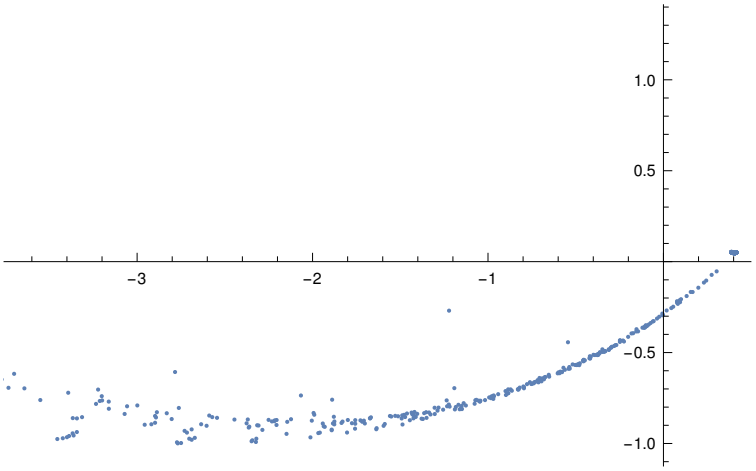
■ Visualizing the tastes of U.S. users

In[ ]:=
```
usTastes = DeleteCases[
  If[#[[4]] ≠ "united states" || Not[MemberQ[Keys[users2D], #[[1]]], None, users2D[#[[1]]] & /@
   csvProfiles, None]
```

■ Making a plot

In[ ]:=
```
ListPlot[usTastes]
```

Out[ ]=



Visualizing distances between the recipes

- Loading the data

```
csvRecipes = Drop[Import["/tmp/recipes.csv", "CSV"], 1]
```

Out[ ]=

{{6663, [{"id":16238,"text":"1/2 cup Parmesan
cheese","quantity":40.0},{"id":16406,"text":"3/4 teaspoon ground black
pepper","quantity":1.57499993},{"id":16396,"text":"1/2 teaspoon garlic
powder","quantity":1.38833332},{"id":2329,"text":"1 (17.5 ounce) package
frozen puff pastry, thawed","quantity":490.0},{"id":16318,"text":"1
egg white","quantity":33.4}]}, ··· 95 129 ···,
{207 304, [{"id":0,"text":"Batter:","quantity":0.0},{"id":1336,"text":"4
(1 ounce) squares unswe
... ","quantity":29.75},{"id":1527,"text":"2 cups sifted
powdered sugar","quantity":250.8}]}}

large output | show less | show more | show all | set size limit ...

- Getting all the recipes ingredients

In[ ]:=
```
recipes = ((# // First) →
    Table[Association[currentIngredient][["id"]], {currentIngredient, ImportString[
      StringJoin[Select[Characters[#[[2]]], PrintableASCIIQ]], "JSON"]}]) & /@ csvRecipes
```

Out[ ]=
```
{6663 → {16 238, 16 406, 16 396, 2329, 16 318},
 6664 → {16 339, 4397, 1631, 1684, 1526, 2356,
   16 421, 2359, 16 287, 16 317, 443, 4317, 12 338, 16 159},
 6665 → {2496, 6311, 1526, 16 421, 2496, 2362, 1684, 16 317},
 6666 → {1526, 6379, 16 317, 1684, 2359, 16 421, 16 386, 16 401, 2496, 4557, 3810},
 6667 → {2362, 16 278, 1526, 1767, 16 421, 16 157},
 6668 → {16 157, 1526, 16 278, 2362, 1686, 1526, 16 317, 5140, 5145, 16 421},
 ··· 95 120 ···, 195 176 → {4175, 18 864}, 195 263 → {3103, 4397, 26 852, 1718},
 123 362 → {1631, 1684, 1526, 2356, 16 421, 16 317, 16 278, 6300},
 206 912 → {21 429, 1536, 21 316},
 207 304 → {0, 1336, 16 157, 16 406, 16 317, 1526, 16 424, 8314, 1684,
   16 421, 2356, 3819, 0, 0, 1336, 1338, 16 157, 8314, 16 258, 1527}}
```

large output | show less | show more | show all | set size limit ...

■ Applying the dimensionality reduction to our data

In[ ]:=
```
recipes2D =
 AssociationThread[Keys[recipes], DimensionReduce[Values[recipes], Method → "Linear"]]
```

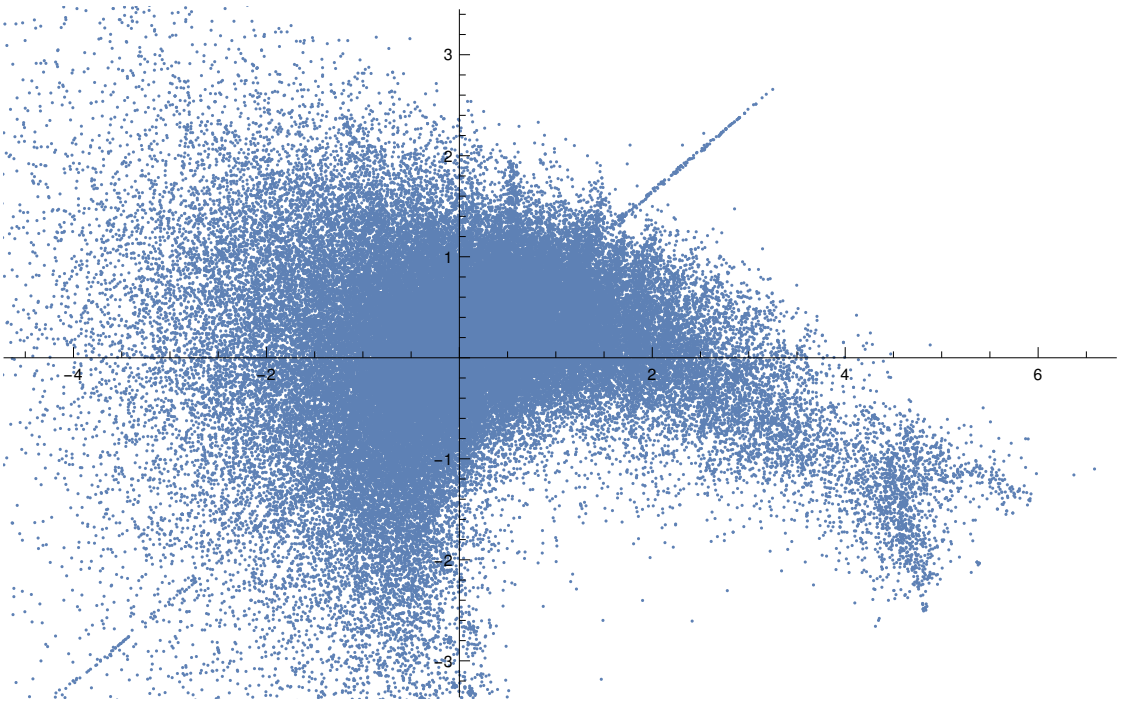
Out[ ]=
```
⟨| 6663 → {-0.932952, -0.707267}, 6664 → {0.985418, 0.68257},
 6665 → {1.90016, 0.755415}, 6666 → {0.581578, 0.372414},
 6667 → {-0.0830507, 0.859058}, 6668 → {0.998631, 0.700727},
 6669 → {1.29309, 0.790004}, 6670 → {1.47406, 0.695382}, 6671 → {2.362, 0.610974},
 6672 → {0.673806, 0.413596}, 6673 → {-0.226909, 0.0106009}, ··· 95 110 ···,
 247 222 → {3.13947, 0.052905}, 247 224 → {0.658275, -0.0845572},
 194 440 → {0.0967534, 0.0974819}, 195 173 → {1.01517, 0.330533},
 195 175 → {0.958892, 0.64575}, 195 176 → {-0.47511, -0.429728},
 195 263 → {-3.35229, 3.41717}, 123 362 → {0.336475, 0.802206},
 206 912 → {-3.63879, 3.5842}, 207 304 → {0.759299, 0.807547} |⟩
```

large output | show less | show more | show all | set size limit ...

■ Plot all the recipes

In[ ]:=
```
ListPlot[Values[recipes2D]]
```

Out[ ]=



Using Deep Learning to predict wether an users might or might not like a food

- Creating associations for users and recipes

In[ ]:=
```
users2DAssociation = Association[users2D]
```

Out[ ]=

```
⟨| 16 007 298 → {0.155663, -0.16671}, 2 012 871 → {-0.0731174, -0.337524},
 9 020 456 → {0.413375, 0.0487773}, 1 922 840 → {0.411156, 0.0484272},
 286 648 → {-0.966066, -0.740453}, 3 672 150 → {0.390562, 0.0512373},
 731 327 → {0.401412, 0.0482864}, 430 608 → {0.408628, 0.0482791},
 594 098 → {0.40627, 0.0501646}, 775 411 → {-4.95718, -0.28721}, ··· 4920 ···,
 15 471 065 → {0.391145, 0.0517223}, 15 777 604 → {0.400428, 0.0453685},
 5 901 381 → {0.415527, 0.0492111}, 1 471 732 → {0.403905, 0.0479741},
 359 494 → {0.415405, 0.0492234}, 9 990 898 → {0.399426, 0.0495313},
 2 447 592 → {0.388339, 0.0520115}, 11 403 286 → {0.414395, 0.0493765},
 7 510 857 → {0.392339, 0.0515993}, 14 048 996 → {0.41384, 0.0493848} |⟩
```

large output | show less | show more | show all | set size limit ...

In[ ]:=
```
recipes2DAssociation = Association[recipes2D]
```

Out[ ]=
```
⟨| 6663 → {-0.932952 , -0.707267}, 6664 → {0.985418 , 0.68257},
 6665 → {1.90016 , 0.755415}, 6666 → {0.581578 , 0.372414},
 6667 → {-0.0830507 , 0.859058}, 6668 → {0.998631 , 0.700727},
 6669 → {1.29309 , 0.790004}, 6670 → {1.47406 , 0.695382}, 6671 → {2.362 , 0.610974},
 6672 → {0.673806 , 0.413596}, 6673 → {-0.226909 , 0.0106009}, ··· 95 110 ··· ,
 247 222 → {3.13947 , 0.052905}, 247 224 → {0.658275 , -0.0845572},
 194 440 → {0.0967534 , 0.0974819}, 195 173 → {1.01517 , 0.330533},
 195 175 → {0.958892 , 0.64575}, 195 176 → {-0.47511 , -0.429728},
 195 263 → {-3.35229 , 3.41717}, 123 362 → {0.336475 , 0.802206},
 206 912 → {-3.63879 , 3.5842}, 207 304 → {0.759299 , 0.807547} |⟩
```

large output | show less | show more | show all | set size limit ...

- Preparing the reviews (getting data from the CSV file and associating them with a rating)

In[ ]:=
```
usersReviews =
  Association[Table[author → DeleteCases[Table[If[review[[3]] == author, review[[2]]],
      {review, csvReviews}], Null], {author, authors}]]
```

Out[ ]=
```
<| 16 007 298 → {213 719, 198 763, 239 184, 238 654, 230 116, 245 403, 12 141, 23 032,
   269 891, 194 425, 223 085, 278 099, 238 162, 86 331, 230 281, 218 790, 15 821,
   256 018, 67 789, 15 931, 255 335, 215 014, 213 505, 233 923, 76 278, 15 413,
   235 589, 13 325, 55 451, 218 180, 142 481, 25 331, 19 307, 255 362, 255 361,
   11 693, 209 990, 9115, 94 570, 230 049, 230 122, 15 213, 255 359, 255 360,
   255 356, 240 245, 23 813, 9253, 51 283, 47 724, 231 351, 240 751, 230 749,
   235 278, 232 363, 235 272, 15 994, 12 538, 255 357, 12 293, 244 458, 73 423,
   217 975, 255 358, 223 297, 18 830, 8495, 239 047, 244 876, 254 690, 254 686,
   254 685, 254 687, 254 689, 241 614, 245 267, 246 655, 244 867, 254 694, 245 221,
   234 664, 73 634, 228 859, 215 309, 186 965, 236 776, 254 692, 254 691, 254 693,
   254 698, 254 695, 254 699, 254 696, 254 697, 21 134, 77 797, 229 423, 239 808,
   234 909, 235 814, 228 319, 88 086, 219 963, 89 334, ··· 400 ···, 232 096, 171 407,
   223 284, 24 094, 15 836, 12 322, 12 235, 218 791, 87 438, 145 194, 214 644,
   222 850, 39 748, 223 302, 91 353, 20 725, 14 581, 14 565, 17 048, 214 709,
   222 512, 149 591, 12 340, 21 694, 170 876, 8546, 215 622, 20 019, 18 443, 13 436,
   146 819, 216 459, 92 359, 217 944, 153 660, 6897, 89 403, 244 597, 218 311,
   131 107, 13 079, 51 147, 240 064, 132 814, 19 490, 25 642, 9247, 100 663,
   220 468, 10 549, 24 609, 75 133, 15 070, 23 176, 15 436, 192 666, 189 930, 79 805,
   77 194, 222 243, 23 600, 235 270, 20 156, 8652, 212 381, 15 206, 56 927, 14 316,
   143 082, 214 488, 77 156, 8493, 9044, 51 301, 16 120, 152 239, 10 638, 229 272,
   98 390, 11 406, 14 556, 23 534, 20 144, 25 971, 229 923, 146 100, 221 000,
   9249, 56 999, 139 580, 16 592, 221 988, 20 021, 15 057, 24 132, 10 813, 17 165,
   10 222, 19 247, 9471, 10 687, 10 402, 31 965, 241 430}, ··· 4938 ···, ··· 1 ··· |>
```

large output | show less | show more | show all | set size limit ...

In[ ]:=
```
reviews = Association [((#[[2]]) → (# // First)) & /@ csvReviews ]
```

Out[ ]=
```
⟨| 213 719 → 5, 198 763 → 5, 239 184 → 5, 238 654 → 3, 230 116 → 5, 245 403 → 5,
 12 141 → 1, 23 032 → 5, 269 891 → 5, 194 425 → 2, 223 085 → 5, 278 099 → 5,
 238 162 → 5, 86 331 → 5, 230 281 → 5, 218 790 → 5, 15 821 → 5, 256 018 → 5,
 67 789 → 3, 15 931 → 4, 255 335 → 5, 215 014 → 5, 213 505 → 5, 233 923 → 5,
 76 278 → 4, 15 413 → 5, 235 589 → 5, 13 325 → 5, 55 451 → 5, 218 180 → 5, 142 481 → 5,
 25 331 → 4, 19 307 → 5, 255 362 → 5, 255 361 → 5, 11 693 → 5, 209 990 → 5, 9115 → 5,
 94 570 → 2, 230 049 → 5, 230 122 → 5, 15 213 → 4, 255 359 → 5, 255 360 → 5,
 255 356 → 5, 240 245 → 5, 23 813 → 3, 9253 → 5, 51 283 → 1, 47 724 → 5, 231 351 → 5,
 240 751 → 5, 230 749 → 4, 235 278 → 5, 232 363 → 4, 235 272 → 5, 15 994 → 4,
 12 538 → 5, ··· 42 191 ···, 104 869 → 5, 64 077 067 → 4, 64 077 075 → 4, 100 839 → 3,
 63 848 876 → 5, 223 518 → 5, 151 538 → 4, 63 113 441 → 5, 62 793 218 → 5, 55 249 928 → 5,
 83 652 → 4, 123 229 → 2, 14 594 → 4, 63 686 → 4, 216 309 → 5, 62 207 634 → 5,
 121 299 → 4, 270 594 → 5, 64 346 178 → 5, 228 015 → 5, 62 147 001 → 5, 62 144 851 → 5,
 62 143 736 → 5, 77 013 → 4, 241 372 → 5, 234 186 → 5, 72 845 → 4, 64 251 → 4,
 23 245 → 5, 10 020 → 5, 237 427 → 4, 162 202 → 5, 122 621 → 5, 121 137 → 5,
 187 155 → 5, 64 468 055 → 5, 82 203 → 4, 33 290 → 4, 13 266 → 5, 11 650 → 4,
 46 960 → 5, 237 888 → 5, 219 569 → 4, 217 094 → 5, 232 790 → 5, 215 739 → 5,
 64 149 547 → 5, 62 364 662 → 5, 62 364 648 → 5, 234 654 → 5, 229 094 → 4, 17 648 → 3,
 200 340 → 5, 267 700 → 5, 121 411 → 5, 127 501 → 2, 12 471 → 3, 203 769 → 4 |⟩
```

large output | show less | show more | show all | set size limit ...

■ Creating the dataset

In[ ]:=
```
dataset = Table[Table[Join[users2DAssociation [user], recipes2DAssociation [recipe]] →
      If[reviews[recipe] ≥ 4, True , False],
    {recipe , usersReviews [user]}], {user , Keys[usersReviews ]}] // First
```

■ Cleaning up the dataset

In[ ]:=
```
cleanedDataset = DeleteCases [(If[ListQ[Keys[#]], #, Null]) & /@ dataset , Null]
```

■ Creating the Deep Neural Network

```
In[ ]:= chain = NetChain[{
          LinearLayer[4], Ramp,

          LinearLayer[128], Ramp, DropoutLayer[0.2],
          LinearLayer[128], Ramp, DropoutLayer[0.2],
          LinearLayer[128], Ramp, DropoutLayer[0.2],
          LinearLayer[128], Ramp, DropoutLayer[0.2],
          LinearLayer[128], Ramp, DropoutLayer[0.2],
          LinearLayer[128], Ramp, DropoutLayer[0.2],

          LinearLayer[2], SoftmaxLayer[]
        }, "Input" → 4, "Output" → NetDecoder[{"Class", {True, False}}]]
```

Out[ ]= NetChain[uninitialized | Input port: vector (size: 4) | Output port: class]

■ Training the Deep Neural Network

```
In[ ]:= model = NetTrain[chain, cleanedDataset, ValidationSet → Scaled[0.2], BatchSize → 128]
```

Out[ ]= NetChain[Input port: vector (size: 4) | Output port: class]

■ Measuring Neural Network's accuracy

```
In[ ]:= ClassifierMeasurements[model, cleanedDataset, "Accuracy"]
```

```
Out[ ]= 0.932343
```

Finding similar recipes

■ Fitting a dimensionality reduction function on the recipes and on the users

```
In[ ]:= recipesDimensionalityReduction =
         DimensionReduction[Values[recipes], Method → "Linear"]
```

Out[ ]= DimensionReducerFunction [Input type: NumericalSequence | Output dimension: 2]

■ Getting 3 sample recipes

```
In[ ]:= sampleRecipes = RandomSample[recipes, 3]
```

```
Out[ ]= {14 510 → {16 157, 16 420, 7428, 18 845, 16 428, 4311},
         146 478 → {16 157, 4397, 4432, 4292, 5489,
           16 278, 16 215, 16 407, 7428, 16 404, 25 387, 21 267, 16 238},
         183 288 → {1525, 16 386, 16 223, 16 312, 16 424, 4978, 3819, 11 188, 3819}}
```

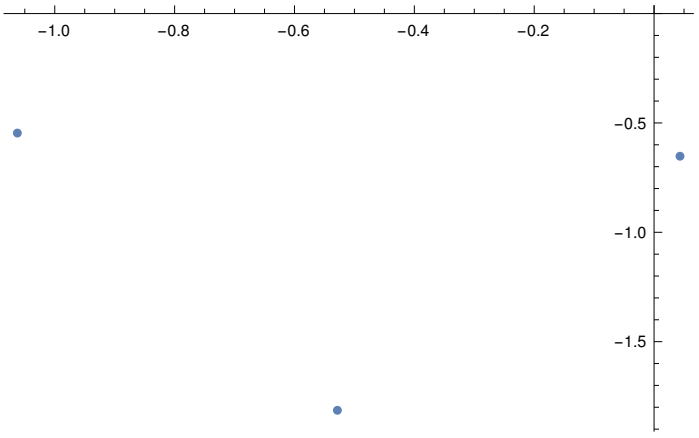
■ Compute the 2D position of each recipe and plotting them

In[ ]:=
```
sampleRecipes2D = AssociationThread [
  recipiesDimensionalityReduction [Values[sampleRecipes ]], Keys[sampleRecipes ]]
```

Out[ ]=
```
⟨|{-0.528366 , -1.81385} → 14 510 ,
 {-1.0624 , -0.546329} → 146 478 , {0.0432639 , -0.652252} → 183 288 |⟩
```

In[ ]:=
```
ListPlot[Keys[sampleRecipes2D ]]
```

Out[ ]=



■ Creating a recipe index

In[ ]:=
```
recipesIndex = AssociationThread [Values[recipes2D ], Keys[recipes2D ]]
```

Out[ ]=
```
⟨|{-0.932952 , -0.707267} → 6663 , {0.985418 , 0.68257} → 6664 ,
 {1.90016 , 0.755415} → 6665 , {0.581578 , 0.372414} → 6666 ,
 {-0.0830507 , 0.859058} → 6667 , {0.998631 , 0.700727} → 6668 ,
 {1.29309 , 0.790004} → 6669 , {1.47406 , 0.695382} → 6670 ,
 {2.362 , 0.610974} → 6671 , {0.673806 , 0.413596} → 6672 ,
 {-0.226909 , 0.0106009} → 6673 , {-0.280432 , 0.430453} → 6674 ,
 {0.350019 , 0.585328} → 6675 , {1.5431 , 0.810775} → 6676 , ··· 94 100 ··· ,
 {1.84672 , 0.392131} → 121 646 , {-2.5214 , -1.1018} → 246 789 ,
 {0.272744 , 0.914196} → 246 918 , {-0.536669 , -1.09069} → 247 059 ,
 {5.92626 , -1.34733} → 247 221 , {3.13947 , 0.052905} → 247 222 ,
 {0.658275 , -0.0845572} → 247 224 , {1.01517 , 0.330533} → 195 173 ,
 {0.958892 , 0.64575} → 195 175 , {-0.47511 , -0.429728} → 195 176 ,
 {-3.35229 , 3.41717} → 195 263 , {0.336475 , 0.802206} → 123 362 ,
 {-3.63879 , 3.5842} → 206 912 , {0.759299 , 0.807547} → 207 304 |⟩
```

large output | show less | show more | show all | set size limit ...

■ Removing samples from index if they exists

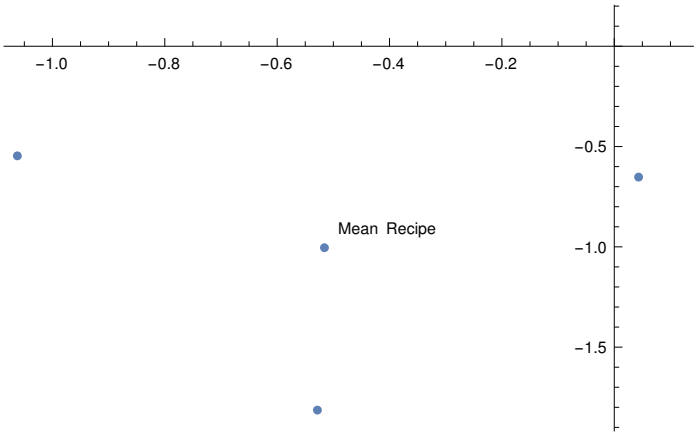
In[ ]:=
```
Do[KeyDropFrom [recipesIndex , recipe], {recipe , sampleRecipes2D }]
```

■ Calculating the mean of the sample recipes and plotting it

```
In[ ]:= meanRecipe = N[Mean[Keys[sampleRecipes2D ]]]
Out[ ]= {-0.515834 , -1.00415}
```

```
In[ ]:= ListPlot[Join[Keys[sampleRecipes2D ], {Labeled[meanRecipe , "Mean Recipe"]}]]
```

Out[ ]=

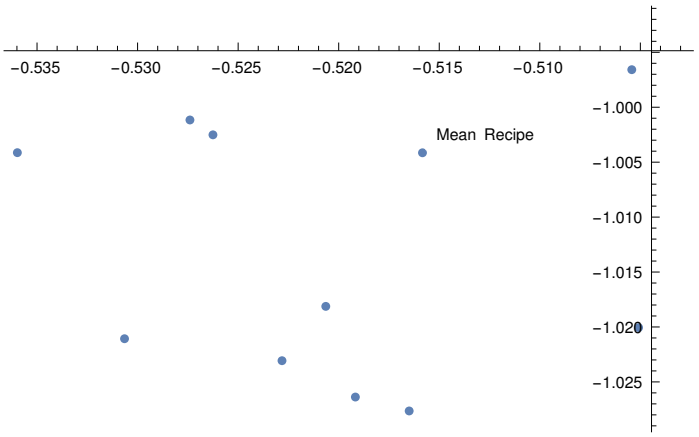


■ Getting the nearest recipes and plotting them

```
In[ ]:= nearestRecipes = Nearest[Keys[recipesIndex ], meanRecipe , 10]
Out[ ]= {{-0.526253 , -1.00251}, {-0.527397 , -1.00116},
  {-0.505426 , -0.996582}, {-0.520642 , -1.01813},
  {-0.505088 , -1.02006}, {-0.535978 , -1.00413}, {-0.522819 , -1.02307},
  {-0.51917 , -1.02638}, {-0.530649 , -1.02107}, {-0.516495 , -1.02765}}
```

```
In[ ]:= ListPlot[Join[nearestRecipes , {Labeled[meanRecipe , "Mean Recipe"]}]]
```

Out[ ]=



■ Getting the IDs

```
In[ ]:= nearestRecipesIDs = recipesIndex [#] & /@ nearestRecipes
Out[ ]= {274 672 , 149 702 , 121 608 , 80 929 , 256 154 , 213 774 , 229 942 , 278 305 , 38 402 , 245 651}
```