

# Bottle Dataset

## import libraries

```
In [1]:  
import pandas as pd  
import numpy as np  
import matplotlib.pyplot as pp
```

## import dataset

```
In [2]:  
data=pd.read_csv(r"E:\154\9_bottle.csv")
```

C:\ProgramData\Anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3165: DtypeWarning: Columns (47,73) have mixed types. Specify dtype option on import or set low\_memory=False.

```
    has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

```
In [3]:  
display(data)
```

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R
0	1	1	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0000A-3	0	10.500	33.4400	NaN	25.64900	NaN	...	
1	1	2	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0008A-3	8	10.460	33.4400	NaN	25.65600	NaN	...	
2	1	3	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0010A-7	10	10.460	33.4370	NaN	25.65400	NaN	...	
3	1	4	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0019A-3	19	10.450	33.4200	NaN	25.64300	NaN	...	
4	1	5	054.0 056.0	19- 4903CR-	20	10.450	33.4210	NaN	25.64300	NaN	...	

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R
					HY-060- 0930- 05400560- 0020A-7							
	...	...	...	...	...	...	...	...	...	...	...	...
<b>864858</b>	34404	864859	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0000A-7	0	18.744	33.4083	5.805	23.87055	108.74	...	
<b>864859</b>	34404	864860	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0002A-3	2	18.744	33.4083	5.805	23.87072	108.74	...	
<b>864860</b>	34404	864861	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0005A-3	5	18.692	33.4150	5.796	23.88911	108.46	...	
<b>864861</b>	34404	864862	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0010A-3	10	18.161	33.4062	5.816	24.01426	107.74	...	
<b>864862</b>	34404	864863	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0015A-3	15	17.533	33.3880	5.774	24.15297	105.66	...	

864863 rows × 74 columns

## To display top 10 rows

In [4]:

data.head(10)

Out[4]:

Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_PHAE0
0	1	1	054.0 056.0	19- 4903CR- HY-060-	0	10.50	33.440	NaN	25.649	NaN	...

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_PHAE0
				0930- 05400560- 0000A-3								
1	1	2	054.0 056.0	HY-060- 0930- 05400560- 0008A-3	19- 4903CR-	8	10.46	33.440	NaN	25.656	NaN	...
2	1	3	054.0 056.0	HY-060- 0930- 05400560- 0010A-7	19- 4903CR-	10	10.46	33.437	NaN	25.654	NaN	...
3	1	4	054.0 056.0	HY-060- 0930- 05400560- 0019A-3	19- 4903CR-	19	10.45	33.420	NaN	25.643	NaN	...
4	1	5	054.0 056.0	HY-060- 0930- 05400560- 0020A-7	19- 4903CR-	20	10.45	33.421	NaN	25.643	NaN	...
5	1	6	054.0 056.0	HY-060- 0930- 05400560- 0030A-7	19- 4903CR-	30	10.45	33.431	NaN	25.651	NaN	...
6	1	7	054.0 056.0	HY-060- 0930- 05400560- 0039A-3	19- 4903CR-	39	10.45	33.440	NaN	25.658	NaN	...
7	1	8	054.0 056.0	HY-060- 0930- 05400560- 0050A-7	19- 4903CR-	50	10.24	33.424	NaN	25.682	NaN	...
8	1	9	054.0 056.0	HY-060- 0930- 05400560- 0058A-3	19- 4903CR-	58	10.06	33.420	NaN	25.710	NaN	...

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_PHAEO
9	1	10	054.0	19- 4903CR- HY-060- 0930-	75	9.86	33.494	NaN	25.801	NaN	...	NaN

10 rows × 74 columns

## To display last 5 rows

In [5]:

```
data.tail()
```

Out[5]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R
864858	34404	864859	093.4	20- 1611SR- MX-310- 2239- 09340264- 0000A-7	0	18.744	33.4083	5.805	23.87055	108.74	...	
864859	34404	864860	093.4	20- 1611SR- MX-310- 2239- 09340264- 0002A-3	2	18.744	33.4083	5.805	23.87072	108.74	...	
864860	34404	864861	093.4	20- 1611SR- MX-310- 2239- 09340264- 0005A-3	5	18.692	33.4150	5.796	23.88911	108.46	...	
864861	34404	864862	093.4	20- 1611SR- MX-310- 2239- 09340264- 0010A-3	10	18.161	33.4062	5.816	24.01426	107.74	...	
864862	34404	864863	093.4	20- 1611SR- MX-310- 2239- 09340264- 0015A-3	15	17.533	33.3880	5.774	24.15297	105.66	...	

5 rows × 74 columns

In [6]: `data.dtypes`

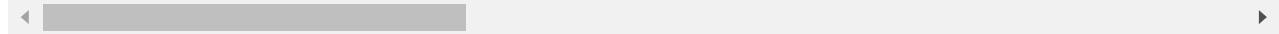
```
Out[6]: Cst_Cnt          int64
         Btl_Cnt          int64
         Sta_ID            object
         Depth_ID          object
         Depthm            int64
                     ...
         TA1              float64
         TA2              float64
         pH2              float64
         pH1              float64
         DIC  Quality  Comment    object
Length: 74, dtype: object
```

## To view statistical summary

In [7]: `data.describe()`

	<b>Cst_Cnt</b>	<b>Btl_Cnt</b>	<b>Depthm</b>	<b>T_degC</b>	<b>Salnty</b>	<b>O2ml_L</b>
<b>count</b>	864863.000000	864863.000000	864863.000000	853900.000000	817509.000000	696201.000000
<b>mean</b>	17138.790958	432432.000000	226.831951	10.799677	33.840350	3.392468
<b>std</b>	10240.949817	249664.587267	316.050259	4.243825	0.461843	2.073256
<b>min</b>	1.000000	1.000000	0.000000	1.440000	28.431000	-0.010000
<b>25%</b>	8269.000000	216216.500000	46.000000	7.680000	33.488000	1.360000
<b>50%</b>	16848.000000	432432.000000	125.000000	10.060000	33.863000	3.440000
<b>75%</b>	26557.000000	648647.500000	300.000000	13.880000	34.196900	5.500000
<b>max</b>	34404.000000	864863.000000	5351.000000	31.140000	37.034000	11.130000

8 rows × 70 columns



## To Print no of elements

In [8]: `data.size`

Out[8]: 63999862

In [9]: `data.ndim`

Out[9]: 2

## To print no of rows and columns

In [10]: `data.shape`

Out[10]: (864863, 74)

## To find missing values

In [11]: `data.isna()`

Out[11]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_PH
0	False	False	False	False	False	False	False	True	False	True	...	
1	False	False	False	False	False	False	False	True	False	True	...	
2	False	False	False	False	False	False	False	True	False	True	...	
3	False	False	False	False	False	False	False	True	False	True	...	
4	False	False	False	False	False	False	False	True	False	True	...	
...	...	...	...	...	...	...	...	...	...	...	...	
864858	False	False	False	False	False	False	False	False	False	False	False	...
864859	False	False	False	False	False	False	False	False	False	False	False	...
864860	False	False	False	False	False	False	False	False	False	False	False	...
864861	False	False	False	False	False	False	False	False	False	False	False	...
864862	False	False	False	False	False	False	False	False	False	False	False	...

864863 rows × 74 columns



## To fill null values with constants

In [12]: `data.fillna(5)`

Out[12]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R
0	1	1	054.0 056.0	HY-060- 0930- 05400560- 0000A-3	19- 4903CR- HY-060- 0930-	0	10.500	33.4400	5.000	25.64900	5.00	...
1	1	2	054.0 056.0	HY-060- 0930-	19- 4903CR- HY-060- 0930-	8	10.460	33.4400	5.000	25.65600	5.00	...

0	1	1	054.0 056.0	HY-060- 0930- 05400560- 0000A-3	19- 4903CR- HY-060- 0930-	0	10.500	33.4400	5.000	25.64900	5.00	...
1	1	2	054.0 056.0	HY-060- 0930-	19- 4903CR- HY-060- 0930-	8	10.460	33.4400	5.000	25.65600	5.00	...

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R
					05400560-0008A-3							
<b>2</b>	1	3	054.0 056.0	HY-060-0930- 05400560-0010A-7	19-4903CR-	10	10.460	33.4370	5.000	25.65400	5.00	...
<b>3</b>	1	4	054.0 056.0	HY-060-0930- 05400560-0019A-3	19-4903CR-	19	10.450	33.4200	5.000	25.64300	5.00	...
<b>4</b>	1	5	054.0 056.0	HY-060-0930- 05400560-0020A-7	19-4903CR-	20	10.450	33.4210	5.000	25.64300	5.00	...
...	...	...	...	...	...	...	...	...	...	...	...	...
<b>864858</b>	34404	864859	093.4 026.4	MX-310-2239- 09340264-0000A-7	20-1611SR-	0	18.744	33.4083	5.805	23.87055	108.74	...
<b>864859</b>	34404	864860	093.4 026.4	MX-310-2239- 09340264-0002A-3	20-1611SR-	2	18.744	33.4083	5.805	23.87072	108.74	...
<b>864860</b>	34404	864861	093.4 026.4	MX-310-2239- 09340264-0005A-3	20-1611SR-	5	18.692	33.4150	5.796	23.88911	108.46	...
<b>864861</b>	34404	864862	093.4 026.4	MX-310-2239- 09340264-0010A-3	20-1611SR-	10	18.161	33.4062	5.816	24.01426	107.74	...
<b>864862</b>	34404	864863	093.4 026.4	MX-310-2239-	20-1611SR-	15	17.533	33.3880	5.774	24.15297	105.66	...

```
Cst_Cnt  Btl_Cnt  Sta_ID  Depth_ID  Depthm  T_degC  Salnty  O2ml_L  STheta  O2Sat ...  R
```

```
09340264-  
0015A-3
```

864863 rows × 74 columns

```
In [13]: data=data[['T_degC','STheta']] [0:10000]
```

```
In [14]: data
```

```
Out[14]:    T_degC  STheta
```

<b>0</b>	10.50	25.649
<b>1</b>	10.46	25.656
<b>2</b>	10.46	25.654
<b>3</b>	10.45	25.643
<b>4</b>	10.45	25.643
...	...	...
<b>9995</b>	15.71	24.769
<b>9996</b>	15.35	24.835
<b>9997</b>	14.64	24.904
<b>9998</b>	14.04	24.991
<b>9999</b>	12.60	25.276

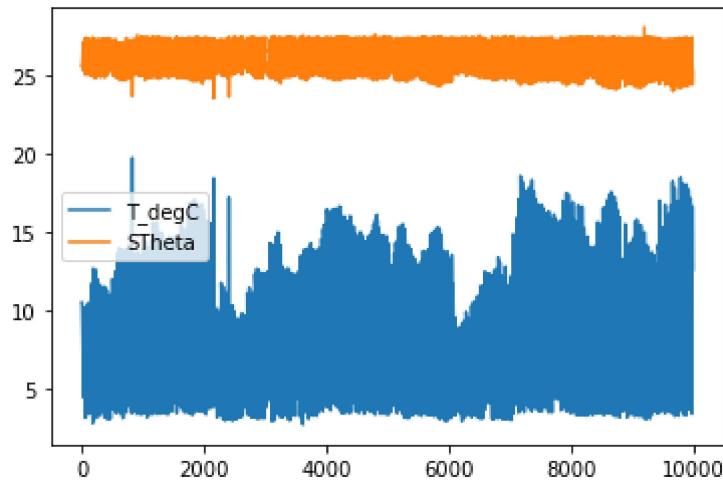
10000 rows × 2 columns

```
In [15]: data=data[['T_degC','STheta']][0:10000]
```

## Line Chart

```
In [16]: data.plot.line()
```

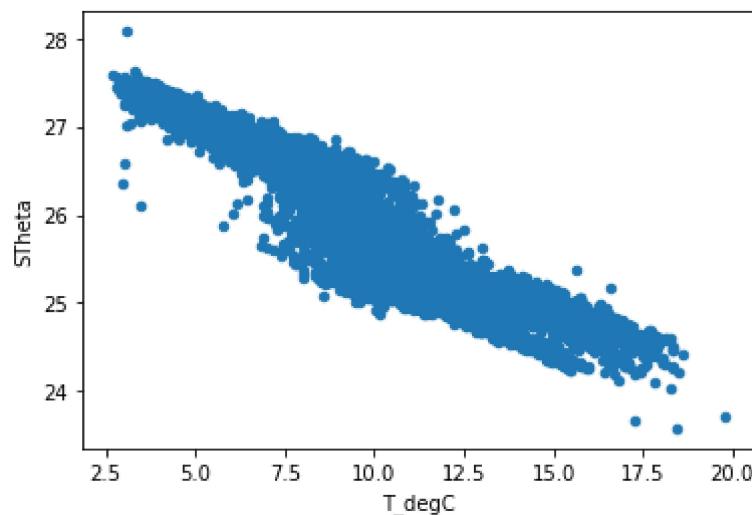
```
Out[16]: <AxesSubplot:>
```



## Scatter Plot

```
In [17]: data.plot.scatter(x='T_degC', y='STheta')
```

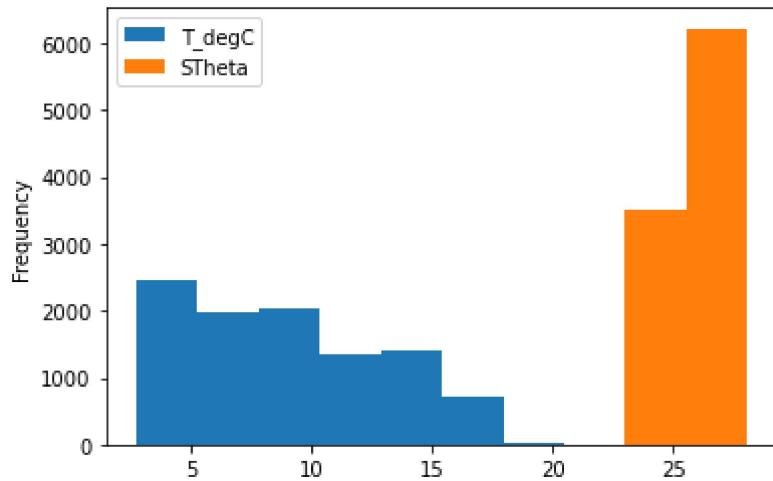
```
Out[17]: <AxesSubplot:xlabel='T_degC', ylabel='STheta'>
```



## Histogram

```
In [18]: data.plot.hist()
```

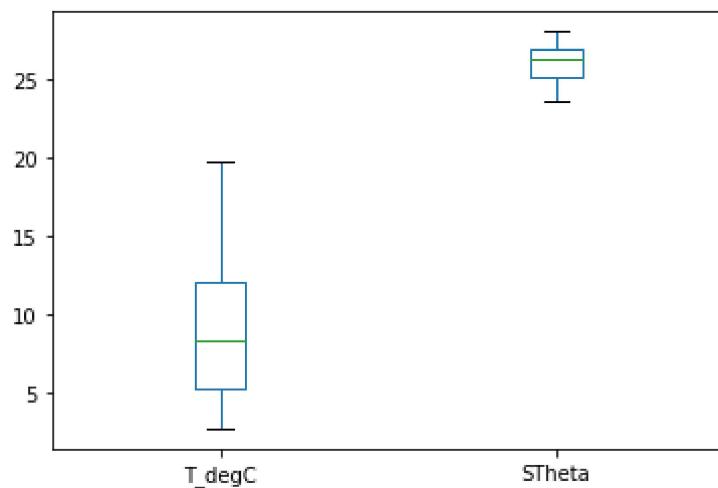
```
Out[18]: <AxesSubplot:ylabel='Frequency'>
```



## Box Plot

```
In [19]: data.plot.box()
```

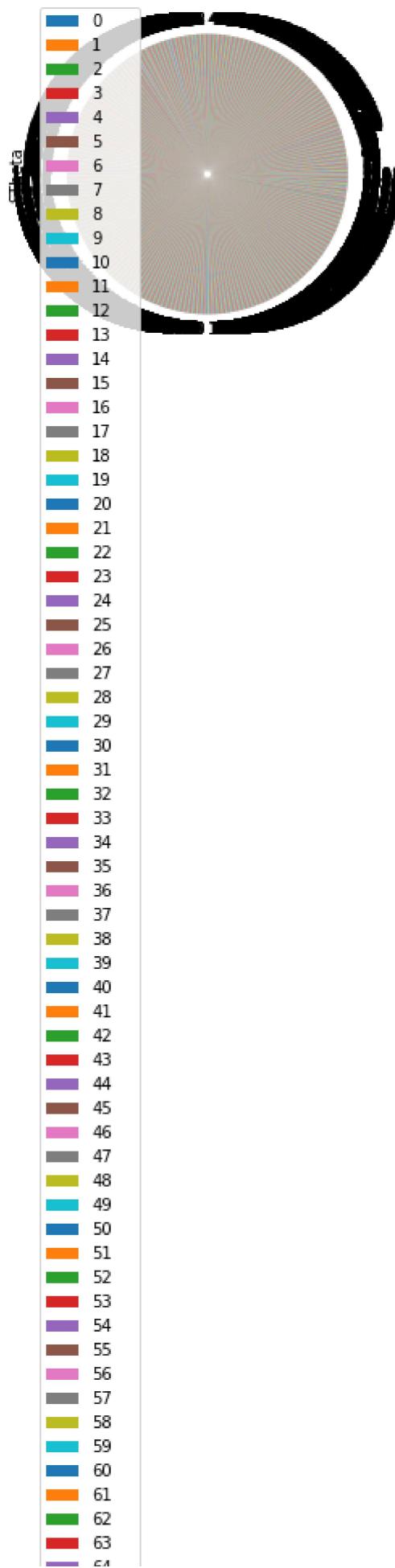
```
Out[19]: <AxesSubplot:>
```



## Pie Chart

```
In [20]: data1=data[['T_degC','STheta']] [0:2000]
data1.plot.pie(y="STheta")
```

```
Out[20]: <AxesSubplot:ylabel='STheta'>
```

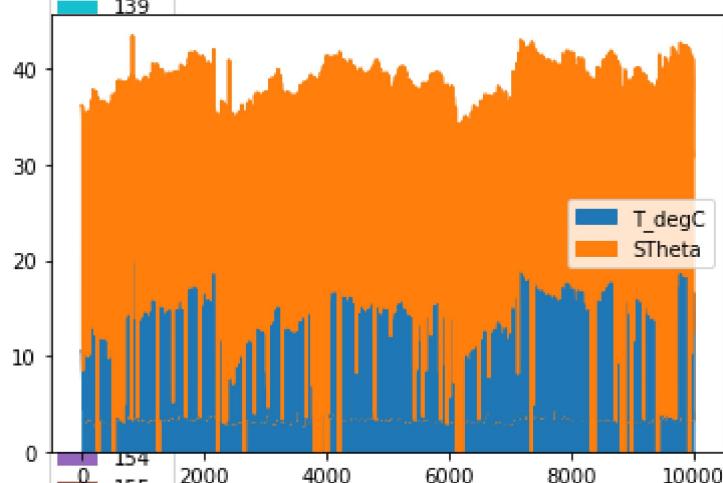


64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129



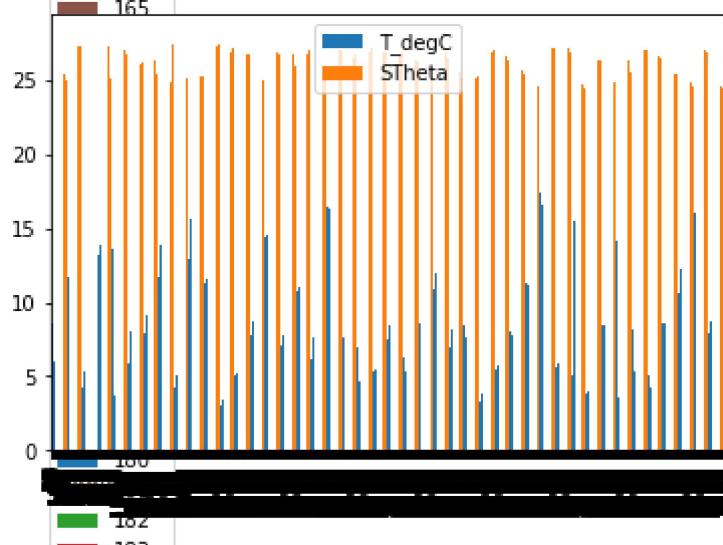
In [21]: `data.plot.area()`

Out[21]: <AxesSubplot: >



In [22]: `data.plot.bar()`

Out[22]: <AxesSubplot: >

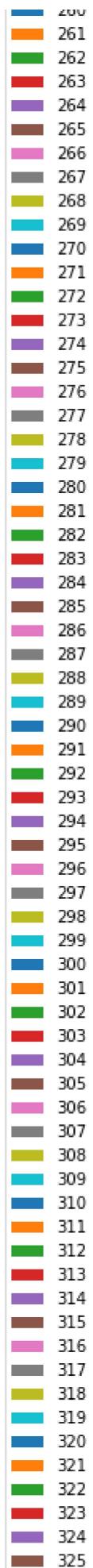


In [ ]:

In [ ]:

In [ ]:

195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260

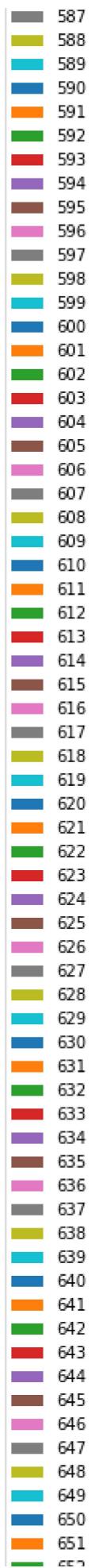


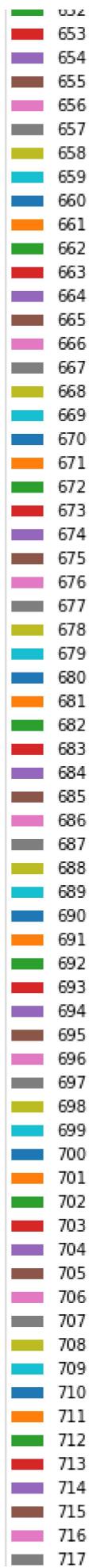
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390

391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456

450
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521

522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586





718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782

783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848

840
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913

914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978

979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044

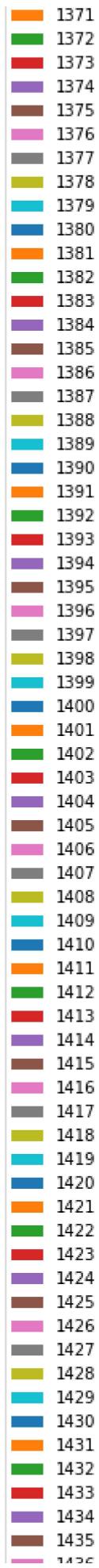
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109

1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174

■	1175
■	1176
■	1177
■	1178
■	1179
■	1180
■	1181
■	1182
■	1183
■	1184
■	1185
■	1186
■	1187
■	1188
■	1189
■	1190
■	1191
■	1192
■	1193
■	1194
■	1195
■	1196
■	1197
■	1198
■	1199
■	1200
■	1201
■	1202
■	1203
■	1204
■	1205
■	1206
■	1207
■	1208
■	1209
■	1210
■	1211
■	1212
■	1213
■	1214
■	1215
■	1216
■	1217
■	1218
■	1219
■	1220
■	1221
■	1222
■	1223
■	1224
■	1225
■	1226
■	1227
■	1228
■	1229
■	1230
■	1231
■	1232
■	1233
■	1234
■	1235
■	1236
■	1237
■	1238
■	1239
■	1240

1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305

■	1306
■	1307
■	1308
■	1309
■	1310
■	1311
■	1312
■	1313
■	1314
■	1315
■	1316
■	1317
■	1318
■	1319
■	1320
■	1321
■	1322
■	1323
■	1324
■	1325
■	1326
■	1327
■	1328
■	1329
■	1330
■	1331
■	1332
■	1333
■	1334
■	1335
■	1336
■	1337
■	1338
■	1339
■	1340
■	1341
■	1342
■	1343
■	1344
■	1345
■	1346
■	1347
■	1348
■	1349
■	1350
■	1351
■	1352
■	1353
■	1354
■	1355
■	1356
■	1357
■	1358
■	1359
■	1360
■	1361
■	1362
■	1363
■	1364
■	1365
■	1366
■	1367
■	1368
■	1369
■	1370



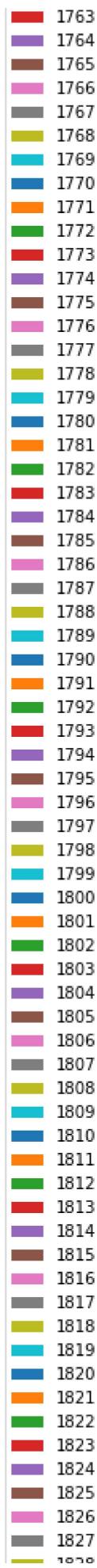
■	1430
■	1437
■	1438
■	1439
■	1440
■	1441
■	1442
■	1443
■	1444
■	1445
■	1446
■	1447
■	1448
■	1449
■	1450
■	1451
■	1452
■	1453
■	1454
■	1455
■	1456
■	1457
■	1458
■	1459
■	1460
■	1461
■	1462
■	1463
■	1464
■	1465
■	1466
■	1467
■	1468
■	1469
■	1470
■	1471
■	1472
■	1473
■	1474
■	1475
■	1476
■	1477
■	1478
■	1479
■	1480
■	1481
■	1482
■	1483
■	1484
■	1485
■	1486
■	1487
■	1488
■	1489
■	1490
■	1491
■	1492
■	1493
■	1494
■	1495
■	1496
■	1497
■	1498
■	1499
■	1500
■	1501

1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566

1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632

1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697

1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762



1820
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893

