

04_Hierarchical_clustering

March 27, 2020

1 Hierarchical agglomerative clustering of cases

Computational notebook 04 for *NAME* Dal Cin, Fleischmann.. - ADD Reference.

Contact: martin@martinflfleischmann.net

Date: 27/03/2020

This notebooks standardize contextual characters and generates hierarchical clustering using Ward's method.

The only input required is `summative_data.csv` generated by `03_Calculate_contextual_characters.ipynb`.

```
[3]: import matplotlib
import matplotlib.pyplot as plt
import pandas as pd
import sklearn
from sklearn import preprocessing
import scipy as sp
from scipy.cluster import hierarchy
```

```
[19]: sklearn.__version__, sp.__version__, pd.__version__, matplotlib.__version__
```

```
[19]: ('0.22.2.post1', '1.4.1', '1.0.3', '3.2.1')
```

```
[5]: path = 'data/summative_data.csv'
data = pd.read_csv(path, index_col=0)
```

```
[14]: data
```

```
[14]:
```

	sdbAre_meanIQ	sdbAre_rangeIQ	sdbAre_TheilID	sdbPer_meanIQ	\
foz	0.145669	0.074895	0.193100	0.525362	
aguda	-0.690154	-0.615406	-0.330468	-0.622020	
esposende	1.115875	0.587245	-0.074543	1.378664	
espinho	-0.488655	-0.837593	-1.292406	-0.390449	
povoa	0.236331	-0.456906	-1.022082	0.668069	
vila_do_conde	-0.641281	-0.900349	-0.805594	-0.418497	

vila_praia	-0.236344	-0.309269	-0.563741	-0.073743
costa_nova	-0.308560	-0.464835	-0.476125	-0.101839
praia_mira	-0.156279	-0.156343	-0.156052	-0.096126
palheiros1	0.901566	-0.200327	-1.192465	0.846955
palheiros2	-0.593671	-0.946392	-1.562552	-0.718485
palheiros3	-0.598359	-0.786099	-0.937243	-0.772217
figueira_foz1	-1.098891	-0.959705	-0.561010	-1.332771
figueira_foz2	-0.619773	-0.573249	0.031223	-0.646134
figueira_foz3	0.493904	0.614052	0.684350	0.564935
pedrogao	0.473856	0.596048	0.288538	0.477814
sao_martinho	1.139194	0.844062	0.246320	1.140482
ericeira	-0.646403	-0.776185	-0.811816	-0.660839
azenhas	-0.875571	-0.925823	-1.115090	-1.022967
quaaios	0.133678	0.468040	0.606887	0.268789
vieira1	-0.901778	-0.933488	-0.923896	-1.109101
vieira2	0.476182	0.413256	0.350032	0.545655
foradouro	-0.164652	-0.269247	-0.527675	0.071673
nazare1	-0.872834	-1.048490	-1.377594	-1.123900
nazare2	2.932040	3.700397	1.741189	2.587010
cascais	-0.750294	-0.182028	1.516532	-0.824684
costa_caparica	-0.592074	0.072117	1.793385	-0.648654
praia_corvoeiro	-1.041269	-0.529299	0.742258	-1.237639
zambujeira	-0.600126	-0.485447	-0.090612	-0.740082
sesimbra	-0.867133	-0.897043	-0.682418	-1.035329
armacao	0.304595	1.462163	2.373260	0.332502
olhao	-0.003125	0.869389	1.797273	0.096579
monte_gordo	2.723822	1.277364	-0.296739	2.442285
albufeira	-0.638526	-0.269896	1.089003	-0.702064
faro	-0.104699	0.226760	0.994317	0.122445
quarteira	2.413738	2.317628	0.352456	2.208322

	sdbPer_rangeIQ	sdbPer_TheilID	ssbCCo_meanIQ \
foz	0.350970	0.224825	-1.582175
aguda	-0.285571	0.065931	-1.052284
esposende	0.714241	0.637960	-0.652857
espinho	-0.837013	-1.101580	-0.588273
povoa	-0.988813	-1.366325	-2.686536
vila_do_conde	-0.874376	-0.637180	-2.273327
vila_praia	-0.406707	-0.725052	-0.558345
costa_nova	-0.228026	-0.358866	-1.124209
praia_mira	0.027534	-0.099592	0.255449
palheiros1	-0.757842	-1.256083	1.682114
palheiros2	-1.226491	-1.438577	1.202416
palheiros3	-1.024424	-1.023904	1.762862
figueira_foz1	-0.970175	-0.523551	-0.291106
figueira_foz2	-0.384762	-0.045132	-0.091578
figueira_foz3	0.532740	0.613300	0.016107

pedrogao	0.542392	0.219599	0.697154
sao_martinho	0.848636	0.452623	0.025071
ericeira	-0.846296	-0.686899	-0.001176
azenhas	-1.045405	-1.040783	0.018076
quaiaos	0.763578	0.709374	0.884984
vieira1	-0.990776	-0.844770	1.284124
vieira2	0.719786	0.450407	1.173539
foradouro	-0.462550	-0.801038	-0.862021
nazare1	-1.456273	-1.420101	1.414961
nazare2	3.142604	1.970070	-0.530249
cascais	0.193065	1.181609	-0.243466
costa_caparica	0.595914	1.564967	0.438035
praia_corvoeiro	-0.408792	0.393931	-0.095620
zambujeira	-0.428670	-0.328774	1.030106
sesimbra	-1.004221	-0.681124	0.306456
armacao	1.654776	2.141651	0.624147
olhao	1.520997	2.022318	-0.109315
monte_gordo	0.272042	-0.690248	0.203878
albufeira	0.079461	1.097082	-0.070199
faro	0.804126	1.145030	-0.776799
quarteira	1.864321	0.178903	0.570057

	ssbCCo_rangeIQ	ssbCCo_TheilID	ssbCor_meanIQ	...	\
foz	1.289722	1.625614	1.418570	...	
aguda	0.651961	0.757708	0.178081	...	
esposende	1.004727	2.057803	3.159203	...	
espinho	0.393074	0.295110	-0.233439	...	
povoa	1.250063	1.949153	-0.219334	...	
vila_do_conde	1.609617	1.781211	-1.563553	...	
vila_praia	-0.296443	-0.394528	0.348525	...	
costa_nova	1.882528	2.021040	-0.371011	...	
praia_mira	0.101699	-0.231061	-0.226495	...	
palheiros1	-2.462428	-1.518385	-0.612414	...	
palheiros2	-1.735955	-1.567289	-1.563553	...	
palheiros3	-2.239058	-1.561990	-1.563553	...	
figueira_foz1	-0.024027	-0.210634	0.157251	...	
figueira_foz2	0.318859	0.021771	-0.354524	...	
figueira_foz3	0.507669	0.293725	0.259799	...	
pedrogao	-0.798909	-0.824334	-0.063065	...	
sao_martinho	-0.003693	-0.331330	-0.097306	...	
ericeira	-0.077809	-0.542039	0.887174	...	
azenhas	0.165658	-0.068862	-0.197032	...	
quaiaos	-1.133161	-0.614224	-0.533963	...	
vieira1	-1.117329	-1.099133	-0.233782	...	
vieira2	-0.491948	-0.465432	-0.050995	...	
foradouro	1.539276	1.431744	-0.532598	...	
nazare1	-1.763870	-1.508549	-1.563553	...	

nazare2	0.589902	-0.004860	-0.738789	...
cascais	0.274451	-0.000360	0.823353	...
costa_caparica	0.427505	0.150223	-0.373019	...
praia_corvoeiro	0.182786	-0.422794	0.126041	...
zambujeira	-0.495271	-0.561681	-1.563553	...
sesimbra	-0.358866	-0.746685	-0.346469	...
armacao	-0.089550	-0.478580	1.316718	...
olhao	0.131751	0.379221	0.207481	...
monte_gordo	-0.118474	-0.320397	0.986175	...
albufeira	-0.051202	-0.400535	0.123464	...
faro	0.668719	0.956021	1.998350	...
quarteira	0.268025	0.153340	1.011813	...

	lskElo_TheilID	likGra_meanIQ	likGra_rangeIQ	\
foz	-0.500558	-0.697226	-0.574718	
aguda	-0.646004	0.214105	0.671245	
esposende	-0.925142	-1.415258	-1.361038	
espinho	0.459406	0.392503	-0.281639	
povoa	0.831341	-0.326971	-0.628027	
vila_do_conde	-0.070472	-0.198830	-0.096624	
vila_praia	0.939187	-0.260624	-0.311059	
costa_nova	-0.538544	-0.190119	-0.892778	
praia_mira	0.443046	-0.254146	0.846564	
palheiros1	-1.134958	-1.130611	-1.233345	
palheiros2	-1.111674	0.451276	-1.219248	
palheiros3	0.420562	0.327512	1.565384	
figueira_foz1	-0.362172	2.800785	1.680444	
figueira_foz2	-0.867828	0.451542	0.806984	
figueira_foz3	-0.223815	-0.165446	0.422766	
pedrogao	0.410632	-0.797134	0.363970	
sao_martinho	1.690686	-1.079796	-0.891570	
ericeira	0.834854	0.976562	-0.733056	
azenhas	-1.307642	0.061480	-0.616883	
quaios	0.386376	-1.153721	-1.112917	
vieira1	0.404325	1.028746	0.578978	
vieira2	0.820905	-1.090640	-0.400422	
foradouro	-1.055496	-0.255042	-1.179590	
nazare1	3.521481	1.801182	1.051530	
nazare2	-0.431192	-1.463759	-1.199687	
cascais	-0.369707	1.307384	1.440742	
costa_caparica	-0.296671	-0.377394	-0.668649	
praia_corvoeiro	-0.140697	0.616143	-0.136710	
zambujeira	-0.593274	0.407838	1.463611	
sesimbra	-0.329867	2.168130	0.999789	
armacao	-0.526223	-0.455090	0.459516	
olhao	0.128509	0.893664	2.465878	
monte_gordo	2.307504	-1.259007	-0.643967	

albufeira	-0.114712	0.174255	0.945155
faro	-0.994964	-0.162346	-0.552422
quarteira	-1.057200	-1.339946	-1.028204

	likGra_TheilID	meshedness_meanIQ	meshedness_rangeIQ \
foz	-0.106314	0.653276	1.034210
aguda	0.384864	3.105235	-0.506306
esposende	-0.732511	-0.493201	0.752514
espinho	-0.869100	1.332040	1.199922
povoa	-0.925107	0.984730	-1.043610
vila_do_conde	-0.492363	0.275276	-1.016189
vila_praia	-0.872127	0.670659	0.543844
costa_nova	-0.955551	1.154867	-0.383700
praia_mira	0.523876	0.986359	0.098466
palheiros1	-1.014853	-1.174709	-0.805875
palheiros2	-1.325050	-1.045194	-0.638967
palheiros3	1.780842	-1.007964	-1.228614
figueira_foz1	-0.874055	0.008264	-0.295389
figueira_foz2	0.260599	-0.855275	0.288392
figueira_foz3	1.284578	0.516899	-0.656976
pedrogao	1.059512	0.285523	1.497316
sao_martinho	0.740992	0.696023	-0.853814
ericeira	-1.168394	0.016498	2.084130
azenhas	-1.119369	-1.739709	-1.548973
quaios	-0.715410	-0.571928	-1.107820
vieira1	-0.304447	0.989722	-0.151493
vieira2	0.448504	0.965242	-1.680004
foradouro	-1.181150	0.517357	0.288544
nazare1	-0.777983	-1.115321	1.923833
nazare2	1.299911	-0.703578	-0.804982
cascais	0.624743	-0.072762	-0.500818
costa_caparica	0.622603	0.632274	-0.042493
praia_corvoeiro	-0.975604	-1.285390	-0.220612
zambujeira	0.314223	0.424661	-0.591777
sesimbra	-0.270193	0.492678	0.532565
armacao	2.548534	-0.542030	-0.443133
olhao	1.032281	-0.999335	2.411589
monte_gordo	0.478139	0.108114	-0.023688
albufeira	2.063143	-0.475239	0.570500
faro	-0.494007	-1.505744	1.189653
quarteira	-0.293754	-1.228314	0.129756

	meshedness_TheilID	sdsLen_meanIQ	sdsLen_rangeIQ \
foz	-0.241022	0.461052	0.724088
aguda	-0.406828	-0.615518	-0.818046
esposende	-0.121131	1.705268	0.855035
espinho	-0.237977	-0.760362	-0.396820

povoa	-0.412517	0.349350	0.067833
vila_do_conde	-0.397444	0.581370	1.509812
vila_praia	-0.272933	0.270958	1.681556
costa_nova	-0.307469	-0.042163	-0.185388
praia_mira	-0.339244	-0.583534	0.030992
palheiros1	-0.347904	-0.333674	-0.206665
palheiros2	-0.333396	-0.661535	-0.931967
palheiros3	-0.361537	-0.760099	-0.170349
figueira_foz1	-0.345735	-0.660135	-0.237054
figueira_foz2	-0.175278	1.479505	-1.007637
figueira_foz3	-0.378479	0.721040	1.869536
pedrogao	-0.074206	0.054505	2.748589
sao_martinho	-0.385081	1.032168	-0.413419
ericeira	-0.060756	-0.907482	-0.441773
azenhas	-0.343724	-1.011384	-0.260812
quaios	-0.390973	-0.038813	-0.753850
vieira1	-0.368978	-0.896935	-0.915641
vieira2	-0.418167	3.837869	-1.007637
foradouro	-0.323453	-1.043737	-0.953351
nazare1	0.938362	-1.149868	-1.007637
nazare2	3.889488	0.601193	1.778987
cascais	-0.376870	0.007725	0.209764
costa_caparica	-0.161275	1.101844	-1.007637
praia_corvoeiro	-0.242353	0.728317	-0.250074
zambujeira	-0.378906	-0.837457	-0.577571
sesimbra	-0.311015	-0.815989	-0.803208
armacao	-0.340535	-0.694528	0.263037
olhao	0.701410	-0.810422	-1.007637
monte_gordo	-0.329474	-0.031411	0.217406
albufeira	-0.185788	-0.605530	0.207218
faro	4.040527	0.827500	1.969977
quarteira	-0.199338	-0.499090	-0.779655

	sdsLen_TheilID
foz	0.028703
aguda	-0.171664
esposende	-0.215043
espinho	-0.561585
povoa	-0.482178
vila_do_conde	0.174819
vila_praia	1.100864
costa_nova	0.377693
praia_mira	-0.248484
palheiros1	-0.623030
palheiros2	-1.028346
palheiros3	-0.007962
figueira_foz1	-0.073076

figueira_foz2	-1.034646
figueira_foz3	0.457077
pedrogao	3.612789
sao_martinho	-0.288137
ericeira	-0.349702
azenhas	0.249238
quaaios	-0.639641
vieira1	-0.869967
vieira2	-1.034646
foradouro	-0.953459
nazare1	-1.034646
nazare2	0.745924
cascais	0.442356
costa_caparica	-0.834119
praia_corvoeiro	-0.927245
zambujeira	-0.604882
sesimbra	1.262271
armacao	0.507301
olhao	-1.034646
monte_gordo	1.628755
albufeira	1.283601
faro	1.820061
quarteira	-0.674349

[36 rows x 96 columns]

1.1 Standardize

```
[15]: x = data.values
      scaler = preprocessing.StandardScaler()
      cols = list(data.columns)
      data[cols] = scaler.fit_transform(data[cols])
      data
```

[15]:	sdbAre_meanIQ	sdbAre_rangeIQ	sdbAre_TheilID	sdbPer_meanIQ	\
foz	0.145669	0.074895	0.193100	0.525362	
aguda	-0.690154	-0.615406	-0.330468	-0.622020	
esposende	1.115875	0.587245	-0.074543	1.378664	
espinho	-0.488655	-0.837593	-1.292406	-0.390449	
povoa	0.236331	-0.456906	-1.022082	0.668069	
vila_do_conde	-0.641281	-0.900349	-0.805594	-0.418497	
vila_praia	-0.236344	-0.309269	-0.563741	-0.073743	
costa_nova	-0.308560	-0.464835	-0.476125	-0.101839	
praia_mira	-0.156279	-0.156343	-0.156052	-0.096126	
palheiros1	0.901566	-0.200327	-1.192465	0.846955	
palheiros2	-0.593671	-0.946392	-1.562552	-0.718485	
palheiros3	-0.598359	-0.786099	-0.937243	-0.772217	

figueira_foz1	-1.098891	-0.959705	-0.561010	-1.332771
figueira_foz2	-0.619773	-0.573249	0.031223	-0.646134
figueira_foz3	0.493904	0.614052	0.684350	0.564935
pedrogao	0.473856	0.596048	0.288538	0.477814
sao_martinho	1.139194	0.844062	0.246320	1.140482
ericeira	-0.646403	-0.776185	-0.811816	-0.660839
azenhas	-0.875571	-0.925823	-1.115090	-1.022967
quaios	0.133678	0.468040	0.606887	0.268789
vieira1	-0.901778	-0.933488	-0.923896	-1.109101
vieira2	0.476182	0.413256	0.350032	0.545655
foradouro	-0.164652	-0.269247	-0.527675	0.071673
nazare1	-0.872834	-1.048490	-1.377594	-1.123900
nazare2	2.932040	3.700397	1.741189	2.587010
cascais	-0.750294	-0.182028	1.516532	-0.824684
costa_caparica	-0.592074	0.072117	1.793385	-0.648654
praia_corvoeiro	-1.041269	-0.529299	0.742258	-1.237639
zambujeira	-0.600126	-0.485447	-0.090612	-0.740082
sesimbra	-0.867133	-0.897043	-0.682418	-1.035329
armacao	0.304595	1.462163	2.373260	0.332502
olhao	-0.003125	0.869389	1.797273	0.096579
monte_gordo	2.723822	1.277364	-0.296739	2.442285
albufeira	-0.638526	-0.269896	1.089003	-0.702064
faro	-0.104699	0.226760	0.994317	0.122445
quarteira	2.413738	2.317628	0.352456	2.208322

	sdbPer_rangeIQ	sdbPer_TheilID	ssbCCo_meanIQ	\
foz	0.350970	0.224825	-1.582175	
aguda	-0.285571	0.065931	-1.052284	
esposende	0.714241	0.637960	-0.652857	
espinho	-0.837013	-1.101580	-0.588273	
povoa	-0.988813	-1.366325	-2.686536	
vila_do_conde	-0.874376	-0.637180	-2.273327	
vila_praia	-0.406707	-0.725052	-0.558345	
costa_nova	-0.228026	-0.358866	-1.124209	
praia_mira	0.027534	-0.099592	0.255449	
palheiros1	-0.757842	-1.256083	1.682114	
palheiros2	-1.226491	-1.438577	1.202416	
palheiros3	-1.024424	-1.023904	1.762862	
figueira_foz1	-0.970175	-0.523551	-0.291106	
figueira_foz2	-0.384762	-0.045132	-0.091578	
figueira_foz3	0.532740	0.613300	0.016107	
pedrogao	0.542392	0.219599	0.697154	
sao_martinho	0.848636	0.452623	0.025071	
ericeira	-0.846296	-0.686899	-0.001176	
azenhas	-1.045405	-1.040783	0.018076	
quaios	0.763578	0.709374	0.884984	
vieira1	-0.990776	-0.844770	1.284124	

vieira2	0.719786	0.450407	1.173539
foradouro	-0.462550	-0.801038	-0.862021
nazare1	-1.456273	-1.420101	1.414961
nazare2	3.142604	1.970070	-0.530249
cascais	0.193065	1.181609	-0.243466
costa_caparica	0.595914	1.564967	0.438035
praia_corvoeiro	-0.408792	0.393931	-0.095620
zambujeira	-0.428670	-0.328774	1.030106
sesimbra	-1.004221	-0.681124	0.306456
armacao	1.654776	2.141651	0.624147
olhao	1.520997	2.022318	-0.109315
monte_gordo	0.272042	-0.690248	0.203878
albufeira	0.079461	1.097082	-0.070199
faro	0.804126	1.145030	-0.776799
quarteira	1.864321	0.178903	0.570057

	ssbCCo_rangeIQ	ssbCCo_TheilID	ssbCor_meanIQ	...	\
foz	1.289722	1.625614	1.418570	...	
aguda	0.651961	0.757708	0.178081	...	
esposende	1.004727	2.057803	3.159203	...	
espinho	0.393074	0.295110	-0.233439	...	
povoa	1.250063	1.949153	-0.219334	...	
vila_do_conde	1.609617	1.781211	-1.563553	...	
vila_praia	-0.296443	-0.394528	0.348525	...	
costa_nova	1.882528	2.021040	-0.371011	...	
praia_mira	0.101699	-0.231061	-0.226495	...	
palheiros1	-2.462428	-1.518385	-0.612414	...	
palheiros2	-1.735955	-1.567289	-1.563553	...	
palheiros3	-2.239058	-1.561990	-1.563553	...	
figueira_foz1	-0.024027	-0.210634	0.157251	...	
figueira_foz2	0.318859	0.021771	-0.354524	...	
figueira_foz3	0.507669	0.293725	0.259799	...	
pedrogao	-0.798909	-0.824334	-0.063065	...	
sao_martinho	-0.003693	-0.331330	-0.097306	...	
ericeira	-0.077809	-0.542039	0.887174	...	
azenhas	0.165658	-0.068862	-0.197032	...	
quaios	-1.133161	-0.614224	-0.533963	...	
vieira1	-1.117329	-1.099133	-0.233782	...	
vieira2	-0.491948	-0.465432	-0.050995	...	
foradouro	1.539276	1.431744	-0.532598	...	
nazare1	-1.763870	-1.508549	-1.563553	...	
nazare2	0.589902	-0.004860	-0.738789	...	
cascais	0.274451	-0.000360	0.823353	...	
costa_caparica	0.427505	0.150223	-0.373019	...	
praia_corvoeiro	0.182786	-0.422794	0.126041	...	
zambujeira	-0.495271	-0.561681	-1.563553	...	
sesimbra	-0.358866	-0.746685	-0.346469	...	

armacao	-0.089550	-0.478580	1.316718	...
olhao	0.131751	0.379221	0.207481	...
monte_gordo	-0.118474	-0.320397	0.986175	...
albufeira	-0.051202	-0.400535	0.123464	...
faro	0.668719	0.956021	1.998350	...
quarteira	0.268025	0.153340	1.011813	...

	lskElo_TheilID	likGra_meanIQ	likGra_rangeIQ	\
foz	-0.500558	-0.697226	-0.574718	
aguda	-0.646004	0.214105	0.671245	
esposende	-0.925142	-1.415258	-1.361038	
espinho	0.459406	0.392503	-0.281639	
povoa	0.831341	-0.326971	-0.628027	
vila_do_conde	-0.070472	-0.198830	-0.096624	
vila_praia	0.939187	-0.260624	-0.311059	
costa_nova	-0.538544	-0.190119	-0.892778	
praia_mira	0.443046	-0.254146	0.846564	
palheiros1	-1.134958	-1.130611	-1.233345	
palheiros2	-1.111674	0.451276	-1.219248	
palheiros3	0.420562	0.327512	1.565384	
figueira_foz1	-0.362172	2.800785	1.680444	
figueira_foz2	-0.867828	0.451542	0.806984	
figueira_foz3	-0.223815	-0.165446	0.422766	
pedrogao	0.410632	-0.797134	0.363970	
sao_martinho	1.690686	-1.079796	-0.891570	
ericeira	0.834854	0.976562	-0.733056	
azenhas	-1.307642	0.061480	-0.616883	
quaios	0.386376	-1.153721	-1.112917	
vieira1	0.404325	1.028746	0.578978	
vieira2	0.820905	-1.090640	-0.400422	
foradouro	-1.055496	-0.255042	-1.179590	
nazare1	3.521481	1.801182	1.051530	
nazare2	-0.431192	-1.463759	-1.199687	
cascais	-0.369707	1.307384	1.440742	
costa_caparica	-0.296671	-0.377394	-0.668649	
praia_corvoeiro	-0.140697	0.616143	-0.136710	
zambujeira	-0.593274	0.407838	1.463611	
sesimbra	-0.329867	2.168130	0.999789	
armacao	-0.526223	-0.455090	0.459516	
olhao	0.128509	0.893664	2.465878	
monte_gordo	2.307504	-1.259007	-0.643967	
albufeira	-0.114712	0.174255	0.945155	
faro	-0.994964	-0.162346	-0.552422	
quarteira	-1.057200	-1.339946	-1.028204	

	likGra_TheilID	meshedness_meanIQ	meshedness_rangeIQ	\
foz	-0.106314	0.653276	1.034210	

aguda	0.384864	3.105235	-0.506306
esposende	-0.732511	-0.493201	0.752514
espinho	-0.869100	1.332040	1.199922
povoa	-0.925107	0.984730	-1.043610
vila_do_conde	-0.492363	0.275276	-1.016189
vila_praia	-0.872127	0.670659	0.543844
costa_nova	-0.955551	1.154867	-0.383700
praia_mira	0.523876	0.986359	0.098466
palheiros1	-1.014853	-1.174709	-0.805875
palheiros2	-1.325050	-1.045194	-0.638967
palheiros3	1.780842	-1.007964	-1.228614
figueira_foz1	-0.874055	0.008264	-0.295389
figueira_foz2	0.260599	-0.855275	0.288392
figueira_foz3	1.284578	0.516899	-0.656976
pedrogao	1.059512	0.285523	1.497316
sao_martinho	0.740992	0.696023	-0.853814
ericeira	-1.168394	0.016498	2.084130
azenhas	-1.119369	-1.739709	-1.548973
quaaios	-0.715410	-0.571928	-1.107820
vieira1	-0.304447	0.989722	-0.151493
vieira2	0.448504	0.965242	-1.680004
foradouro	-1.181150	0.517357	0.288544
nazare1	-0.777983	-1.115321	1.923833
nazare2	1.299911	-0.703578	-0.804982
cascais	0.624743	-0.072762	-0.500818
costa_caparica	0.622603	0.632274	-0.042493
praia_corvoeiro	-0.975604	-1.285390	-0.220612
zambujeira	0.314223	0.424661	-0.591777
sesimbra	-0.270193	0.492678	0.532565
armacao	2.548534	-0.542030	-0.443133
olhao	1.032281	-0.999335	2.411589
monte_gordo	0.478139	0.108114	-0.023688
albufeira	2.063143	-0.475239	0.570500
faro	-0.494007	-1.505744	1.189653
quarteira	-0.293754	-1.228314	0.129756

	meshedness_TheilID	sdsLen_meanIQ	sdsLen_rangeIQ	\
foz	-0.241022	0.461052	0.724088	
aguda	-0.406828	-0.615518	-0.818046	
esposende	-0.121131	1.705268	0.855035	
espinho	-0.237977	-0.760362	-0.396820	
povoa	-0.412517	0.349350	0.067833	
vila_do_conde	-0.397444	0.581370	1.509812	
vila_praia	-0.272933	0.270958	1.681556	
costa_nova	-0.307469	-0.042163	-0.185388	
praia_mira	-0.339244	-0.583534	0.030992	
palheiros1	-0.347904	-0.333674	-0.206665	

palheiros2	-0.333396	-0.661535	-0.931967
palheiros3	-0.361537	-0.760099	-0.170349
figueira_foz1	-0.345735	-0.660135	-0.237054
figueira_foz2	-0.175278	1.479505	-1.007637
figueira_foz3	-0.378479	0.721040	1.869536
pedrogao	-0.074206	0.054505	2.748589
sao_martinho	-0.385081	1.032168	-0.413419
ericeira	-0.060756	-0.907482	-0.441773
azenhas	-0.343724	-1.011384	-0.260812
quaios	-0.390973	-0.038813	-0.753850
vieira1	-0.368978	-0.896935	-0.915641
vieira2	-0.418167	3.837869	-1.007637
foradouro	-0.323453	-1.043737	-0.953351
nazare1	0.938362	-1.149868	-1.007637
nazare2	3.889488	0.601193	1.778987
cascais	-0.376870	0.007725	0.209764
costa_caparica	-0.161275	1.101844	-1.007637
praia_corvoeiro	-0.242353	0.728317	-0.250074
zambujeira	-0.378906	-0.837457	-0.577571
sesimbra	-0.311015	-0.815989	-0.803208
armacao	-0.340535	-0.694528	0.263037
olhao	0.701410	-0.810422	-1.007637
monte_gordo	-0.329474	-0.031411	0.217406
albufeira	-0.185788	-0.605530	0.207218
faro	4.040527	0.827500	1.969977
quarteira	-0.199338	-0.499090	-0.779655

	sdsLen_TheilID
foz	0.028703
aguda	-0.171664
esposende	-0.215043
espinho	-0.561585
povoa	-0.482178
vila_do_conde	0.174819
vila_praia	1.100864
costa_nova	0.377693
praia_mira	-0.248484
palheiros1	-0.623030
palheiros2	-1.028346
palheiros3	-0.007962
figueira_foz1	-0.073076
figueira_foz2	-1.034646
figueira_foz3	0.457077
pedrogao	3.612789
sao_martinho	-0.288137
ericeira	-0.349702
azenhas	0.249238

quaiaos	-0.639641
vieira1	-0.869967
vieira2	-1.034646
foradouro	-0.953459
nazare1	-1.034646
nazare2	0.745924
cascais	0.442356
costa_caparica	-0.834119
praia_corvoeiro	-0.927245
zambujeira	-0.604882
sesimbra	1.262271
armacao	0.507301
olhao	-1.034646
monte_gordo	1.628755
albufeira	1.283601
faro	1.820061
quarteira	-0.674349

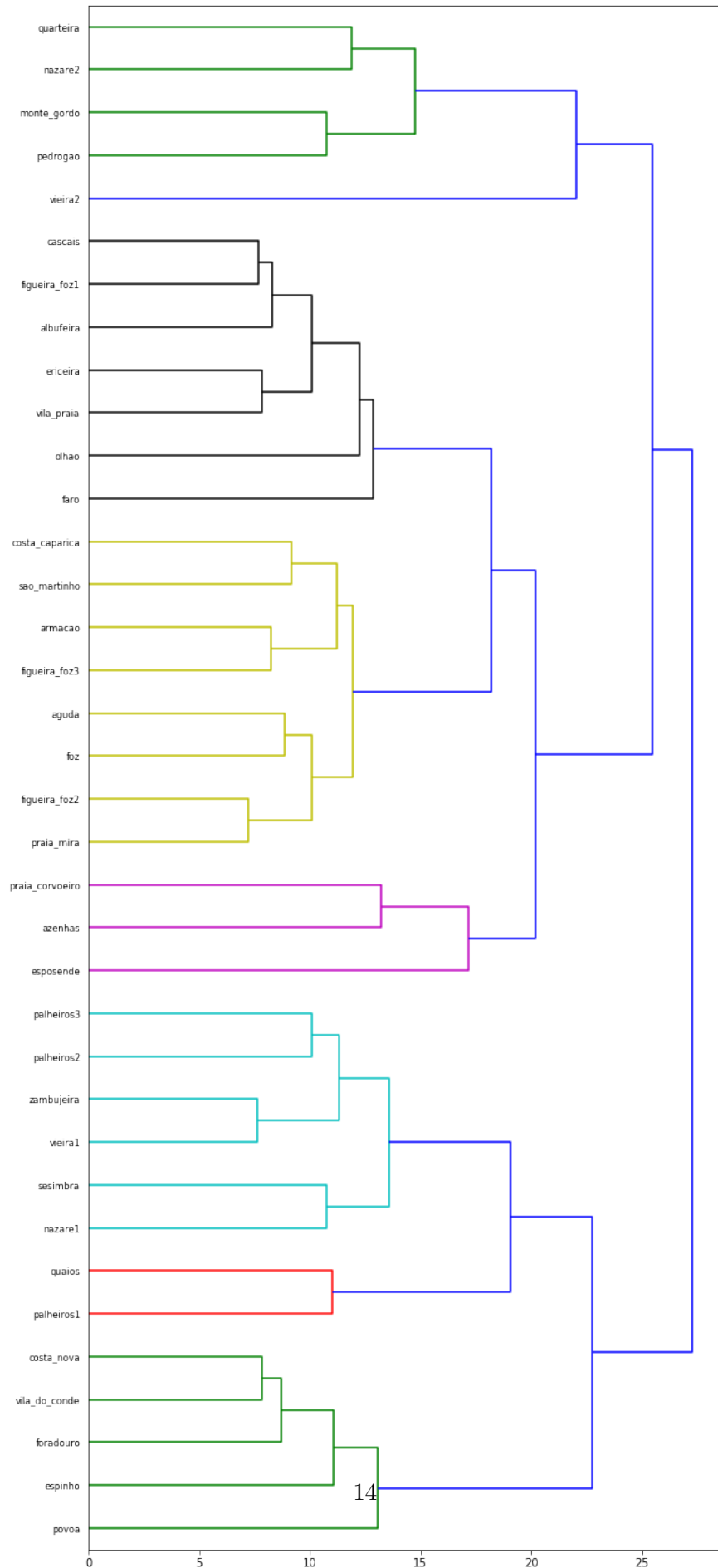
[36 rows x 96 columns]

```
[16]: data.to_csv('data/summative_data_norm.csv')
```

1.2 Clustering

```
[17]: Z = hierarchy.linkage(data, 'ward')
```

```
[18]: plt.figure(figsize=(10, 25))
dn = hierarchy.dendrogram(Z, labels=data.index, orientation='right',
                           color_threshold=18)
#plt.savefig('dendrogram_right.svg')
```



1.3 Assing to places

```
[ ]: folder = 'data/'
parts = ['atlantic', 'preatl', 'premed', 'med']

c_parts = []
geoms = []
for part in parts:
    path = folder + part + '.gpkg'
    layers = [x[:-4] for x in fiona.listlayers(path) if 'blg' in x]
    for l in layers:
        buildings = gpd.read_file(path, layer=l + '_blg')

        if 'part' in buildings.columns:
            for bpart in set(buildings.part):
                subset = buildings.loc[buildings.part == bpart]
                geoms.append(subset.geometry.unary_union.centroid)
                c_parts.append(part)
        else:
            geoms.append(buildings.geometry.unary_union.centroid)
            c_parts.append(part)
```

```
[ ]: gdf = gpd.GeoDataFrame(data, geometry=geoms)
gdf['part'] = c_parts
```

```
[ ]: gdf['cl'] = hierarchy.fcluster(Z, 18, criterion='distance')
gdf['cl'] = gdf.cl.replace(8, 7)
gdf['cl'] = gdf.cl.replace(2, 3)
```

```
[ ]: gdf.reset_index().to_file('data/points.gpkg', driver='GPKG', layer='ward')
```