Clustering of countries based on measures

For calculating the epidemic age, we considered t0 the day when the number of confirmed cases reaches 10.
Implemented measures are separated in three groups based on time of implementation: "anticipatory measures": implemented before t0 (day when 10 cases were reported); "early measures": implemented between t0 and day when 200 cases were reported; and "late measures": implemented after day when 200 cases were reported.

Each country is described with three variables:

1. **Number of Anticipatory measures**
2. **Number of Early measures**
3. **Number of Late measures**
4.

KMeans

Optimal number of clusters is chosen based on total wittiness. I removed all "Risk communication" measures.
I chose 40 most frequent measures (at least 15 countries)

[1] "Activate case notification"
[2] "Airport health check"
[3] "Border health check"
[4] "Enhance detection system"
[5] "Isolation of cases"
[6] "Quarantine"
[7] "Surveillance"
[8] "Tracing and tracking"
[9] "Environmental cleaning and disinfection"
[10] "Adapt procedures for patient management"
[11] "Enhance laboratory testing capacity"
[12] "Increase availability of PPE"
[13] "Increase healthcare workforce"
[14] "Increase in medical supplies and equipment"
[15] "Increase isolation and quarantine facilities"
[16] "Increase patient capacity"
[17] "Personal protective measures"
[18] "Research"
[19] "Activate or establish emergency response"
[20] "Crisis management plans"
[21] "Measures to ensure security of supply"

[22] "Police and army interventions"
[23] "The government provide assistance to vulnerable populations"
[24] "Actively communicate with healthcare professionals."
[25] "Actively communicate with managers."
[26] "Educate and actively communicate with the public."
[27] "Travel alert and warning"
[28] "Closure of educational institutions"
[29] "Mass gathering cancellation"
[30] "Measures for special populations"
[31] "Return operation of nationals"
[32] "Small gathering cancellation"
[33] "Special measures for certain establishments"
[34] "Airport restriction"
[35] "Border restriction"
[36] "Cordon sanitaire"
[37] "Individual movement restrictions"
[38] "National lockdown"
[39] "Port and ship restriction"
[40] "Public transport restriction"

RESULT:
K-means clustering with 8 clusters of sizes 9, 7, 6, 6, 10, 1, 6, 11

Cluster means:

	Anticipatory_measures	Early_measures	Late_measures
1	4.222222	8.888889	4.111111
2	7.000000	14.857143	3.428571
3	4.666667	2.000000	21.666667
4	5.000000	9.000000	16.000000
5	12.900000	2.400000	1.900000
6	26.000000	1.000000	0.000000
7	14.833333	7.833333	7.166667
8	7.545455	3.000000	12.909091

Clustering vector:
[1] 2 8 8 1 3 4 3 5 8 8 2 5 1 8 8 4 7 5 5 7 1 4 1 8 2 5 6 1 5 2 1
[32] 5 8 5 3 1 7 3 7 4 4 7 2 1 7 5 1 4 3 8 8 5 2 2 8 3

```
Within cluster sum of squares by cluster:
[1] 163.33333  96.57143 134.66667 108.00000 162.20000   0.00000
[7] 122.50000 169.63636
 (between_SS / total_SS =  82.8 %)

> countries_with_measures[which(cluster.results$cluster==1)]
[1] "Bosnia and Herzegovina" "Estonia"
[3] "Iceland"                "Indonesia"
[5] "Kuwait"                 "Malaysia"
[7] "New Zealand"            "Serbia"
[9] "Slovenia"
>  countries_with_measures[which(cluster.results$cluster==2)]
[1] "Albania"   "Ecuador"   "Japan"     "Lithuania" "Senegal"
[6] "Taiwan*"   "Thailand"
>  countries_with_measures[which(cluster.results$cluster==3)]
[1] "Brazil"      "China"       "Netherlands" "Norway"
[5] "Spain"       "US"
>  countries_with_measures[which(cluster.results$cluster==4)]
[1] "Canada"       "Germany"      "India"        "Portugal"
[5] "Ireland"      "Korea, South"
>  countries_with_measures[which(cluster.results$cluster==5)]
 [1] "Croatia"       "El Salvador"   "Greece"
 [4] "Honduras"      "Kazakhstan"    "Liechtenstein"
 [7] "Mauritius"     "Montenegro"    "Slovakia"
[10] "Syria"
>  countries_with_measures[which(cluster.results$cluster==6)]
[1] "Kosovo"
>  countries_with_measures[which(cluster.results$cluster==7)]
[1] "Ghana"           "Hungary"         "North Macedonia"
[4] "Poland"          "Romania"         "Singapore"
>  countries_with_measures[which(cluster.results$cluster==8)]
 [1] "Austria"        "Belgium"        "Czechia"
 [4] "Denmark"        "Finland"        "France"
 [7] "Italy"          "Mexico"         "Sweden"
[10] "Switzerland"    "United Kingdom”
```

Happy to get feedback and try new things.