Homework 4

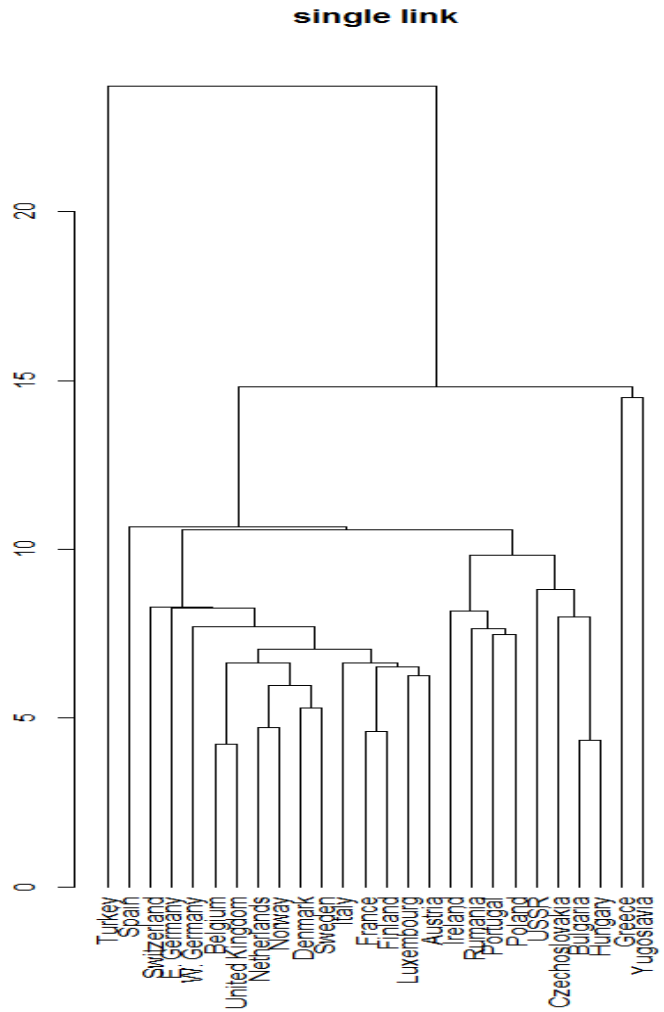
1 PART 1

1.1 PART 1.1

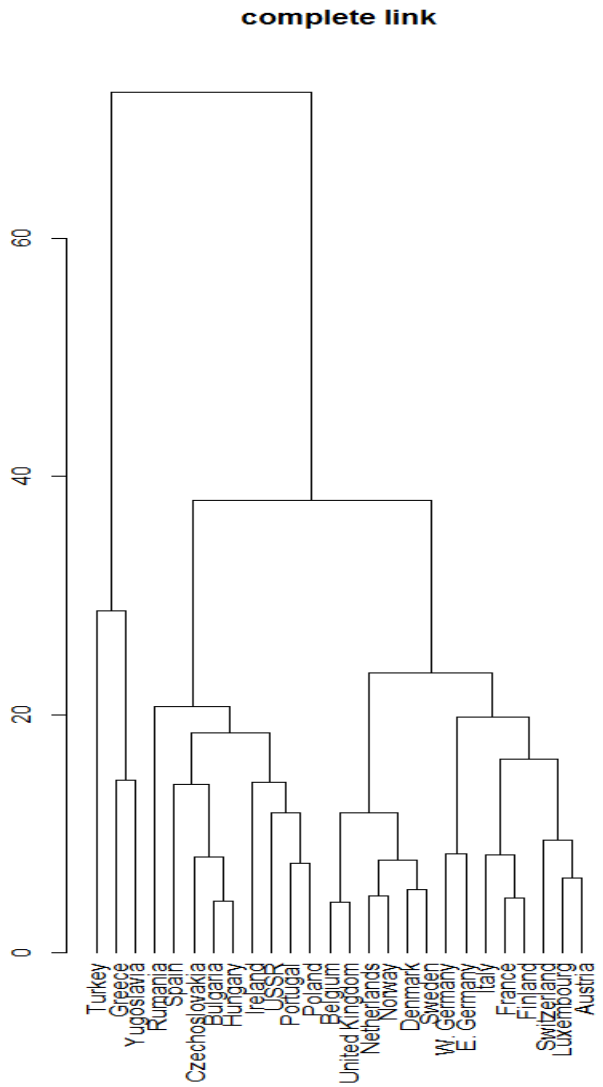
1.1.1 Library used

- Cluster

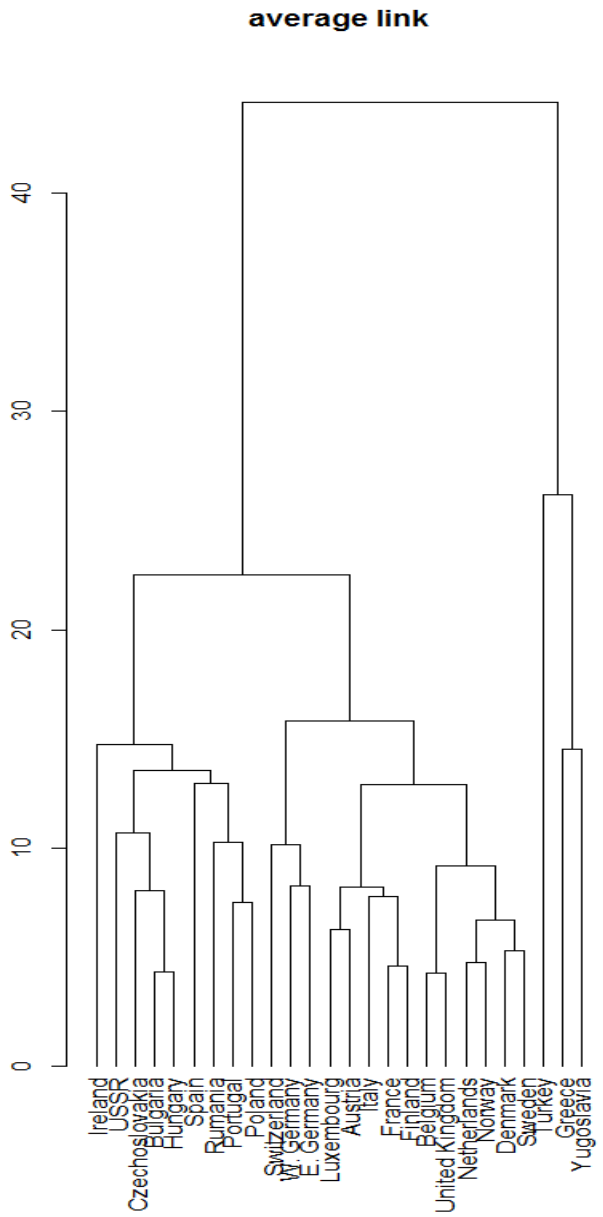
1.1.2 Dendogram: Single Link



1.1.3 Dendogram: Complete Link



1.1.4 Dendogram: Average Link

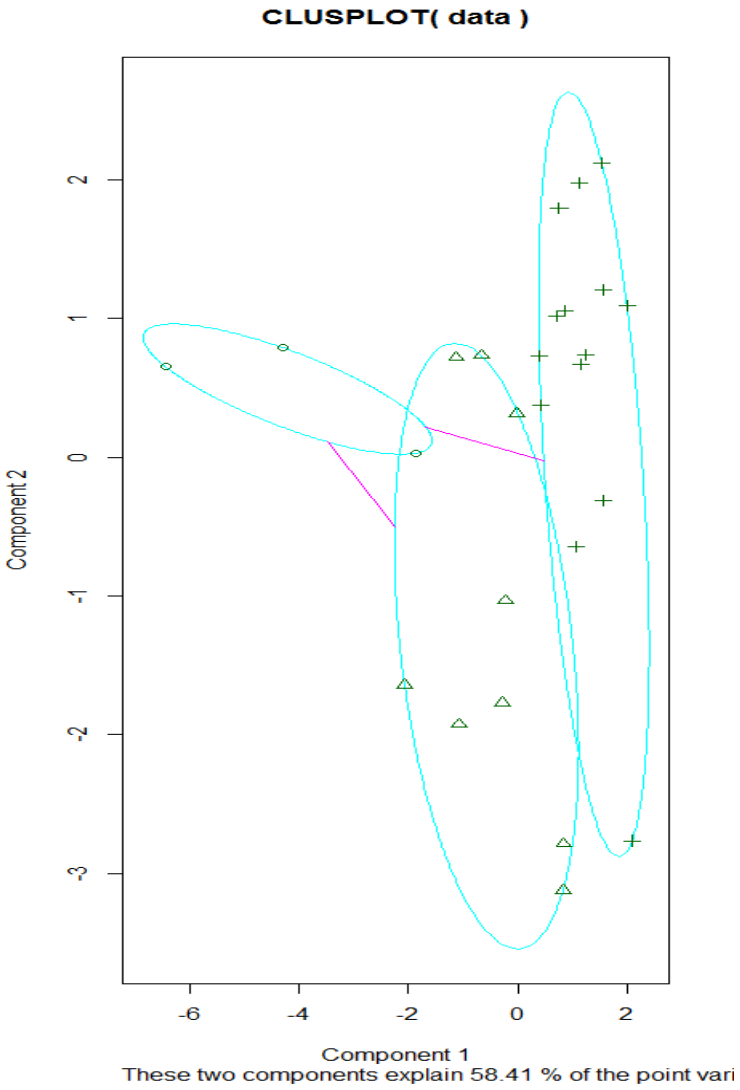


1.2 PART 1.2

1.2.1 What is the good choice of K for this data?

In my opinion, '3' is a good choice for k. It fairly forms clusters without over fitting.

1.2.2 K-means plot



2 PART 2

2.1 ERROR RATE

```
0.4121849
```

2.2 CONFUSION MATRIX

Prediction	Brush_teeth	Climb_stairs	Comb_hair	Descend_stairs
Brush_teeth	1030	7	14	2
Climb_stairs	0	356	4	43
Comb_hair	0	5	276	2
Descend_stairs	2	27	7	168
Drink_glass	0	25	31	10
Eat_meat	0	9	18	5
Eat_soup	0	0	0	0
Getup_bed	64	78	22	16
Liedown_bed	0	0	0	0
Pour_water	70	12	48	1
Sitdown_chair	8	26	4	15
Standup_chair	2	1	0	0
Use_telephone	0	1	7	0
Walk	16	258	39	45

Prediction	Drink_glass	Eat_meat	Eat_soup	Getup_bed	Liedown_bed
Brush_teeth	15	0	0	13	9
Climb_stairs	0	0	0	13	6
Comb_hair	16	5	0	18	6
Descend_stairs	0	0	0	5	3
Drink_glass	368	11	2	35	5
Eat_meat	201	571	19	105	30
Eat_soup	0	0	0	0	0
Getup_bed	28	0	0	419	79
Liedown_bed	0	0	0	1	0
Pour_water	205	37	112	120	46
Sitdown_chair	15	0	0	27	10
Standup_chair	0	0	0	7	0
Use_telephone	5	0	0	7	1
Walk	2	0	0	146	33

Prediction	Pour_water	Sitdown_chair	Standup_chair	Use_telephone	Walk
Brush_teeth	11	15	3	1	25
Climb_stairs	0	8	19	0	154
Comb_hair	5	0	0	17	0
Descend_stairs	0	0	1	0	20
Drink_glass	13	4	3	68	10
Eat_meat	211	105	62	27	25

Eat_soup	0	0	0	0	0
Getup_bed	31	112	168	9	127
Liedown_bed	0	1	0	0	0
Pour_water	553	69	57	30	17
Sitdown_chair	9	72	37	11	25
Standup_chair	0	1	9	0	2
Use_telephone	0	0	0	141	0
Walk	0	113	149	0	1440

2.3 IMPACT OF VALUE OF K

I found value of k from 5-8 optimum.

2.4 MODIFYING SAMPLE LENGTH

I used length values between 30 to 50 and accuracy improves up to some fractions.