

José Daniel Molina Galindo

Carné: 1007420

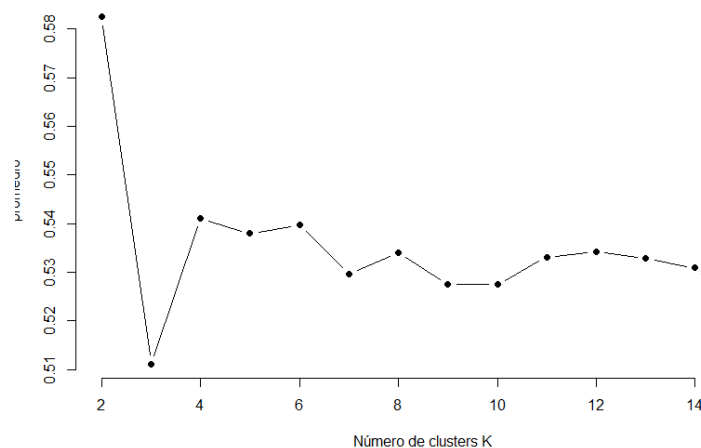
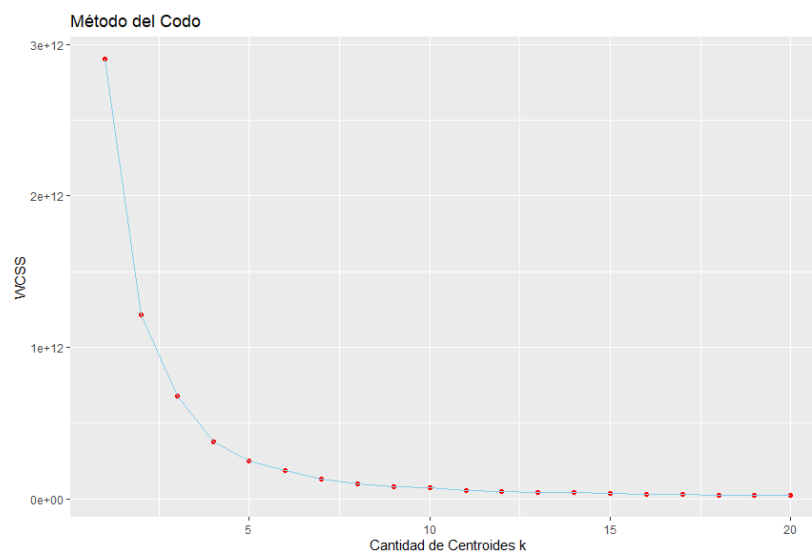
Análisis de Datos



## LABORATORIO CLUSTERING

### Parte I – clustering y la ley del codo

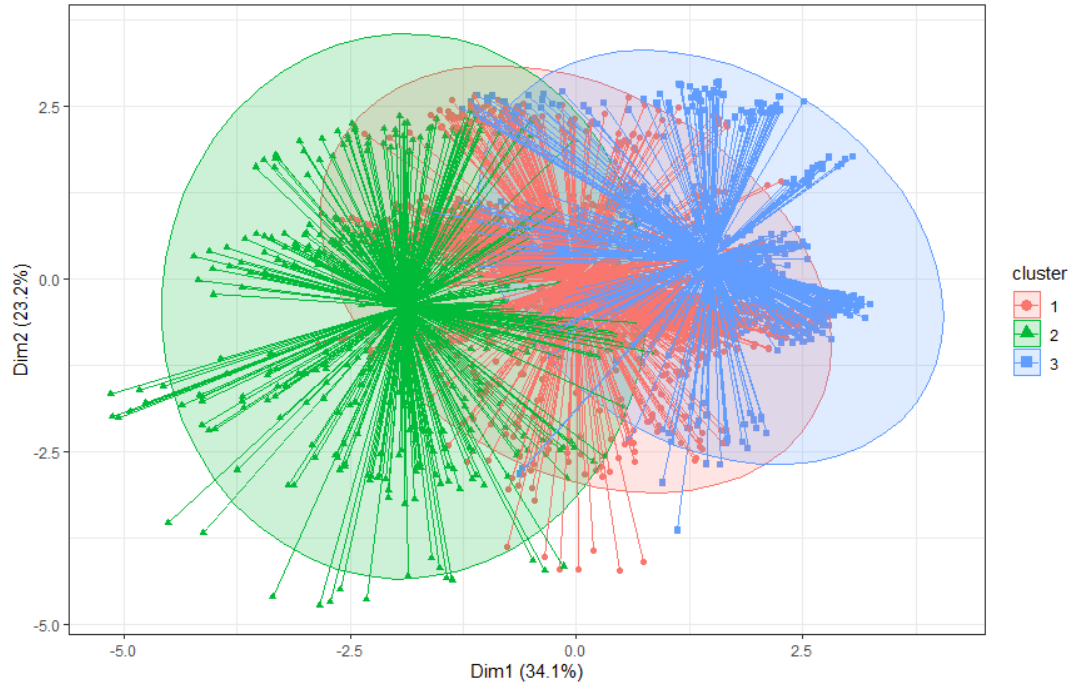
1. Determinar el numero ideal de clusters, de acuerdo a la ley codo, justificando de manera grafica su decision.



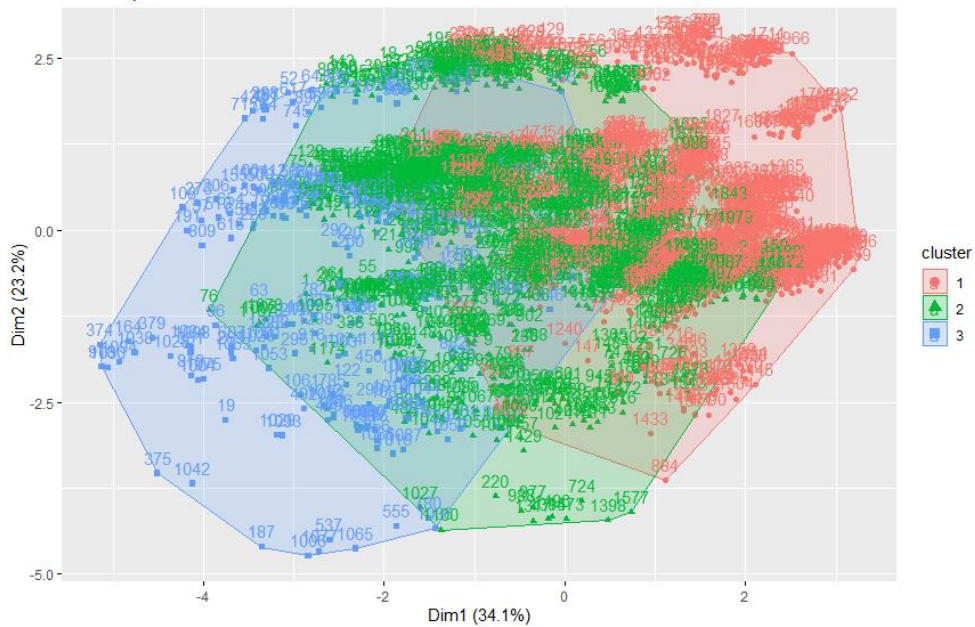
Por el método del codo y con el siguiente gráfico podemos determinar que el número ideal de clustering es el 3 porque en el método del codo se puede ver donde hace el cambio en los datos de la tabla.

2. Efectuar el clustering con el numero ideal de clusters determinado por su persona y graficar (con las herramientas vistas en clase) como quedarían los datos agrupados.

Resultados clustering K-means



Cluster plot



## Parte 2 – Clustering y feature engineering

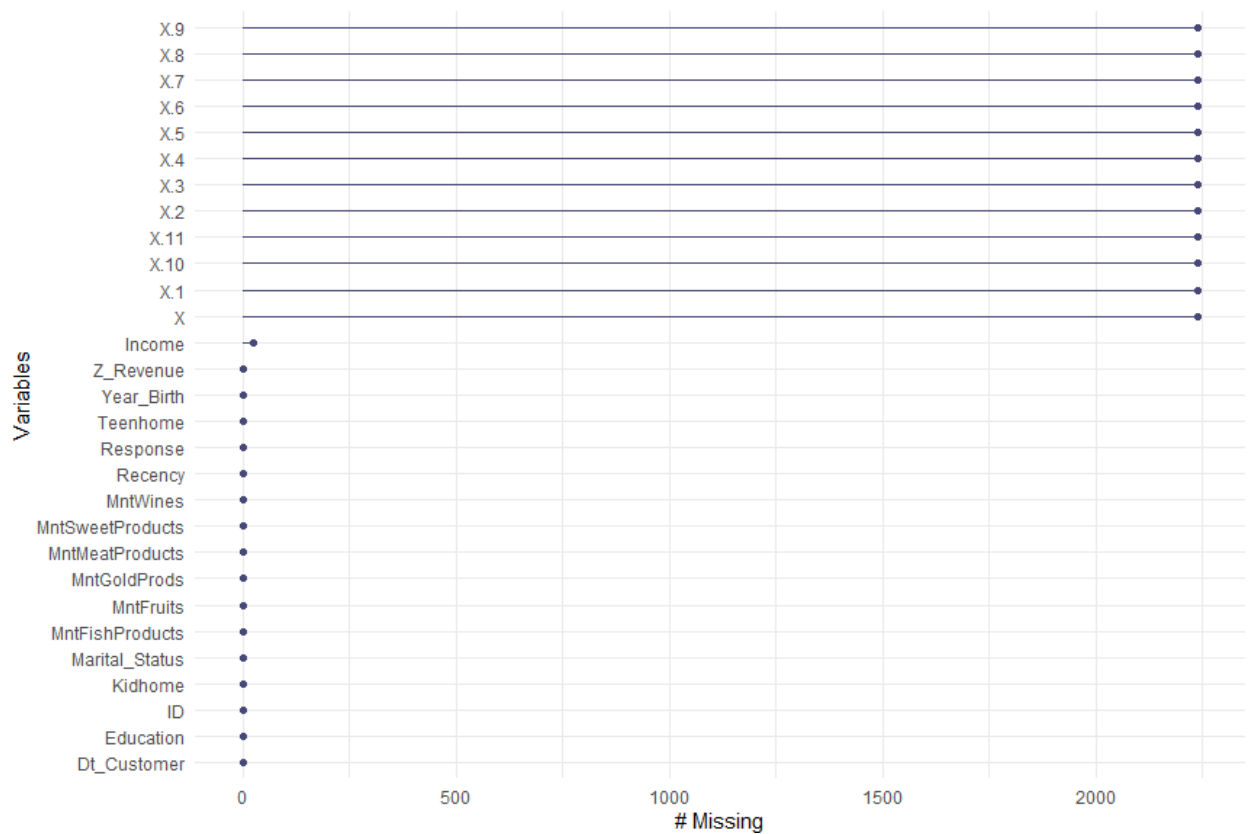
Una vez efectuado este paso previo:

1. Determinar el numero ideal de clusters, de acuerdo a la ley codo, justificando de manera grafica su decision.

### Cantidad de valores NA

```
[1] 26904
```

Con esta tabla podemos saber los números de los datos NA que hay en cada columna



1. Determinar el número ideal de clusters, de acuerdo a la ley codo, justificando de manera grafica su decision.

ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	MntWines	MntFruits
5524	1957	Graduation	Single	58138	0	0	04-09-2012	58	635	88
2174	1954	Graduation	Single	46344	1	1	08-03-2014	38	11	1
4141	1965	Graduation	Together	71613	0	0	21-08-2013	26	426	49
6182	1984	Graduation	Together	26646	1	0	10-02-2014	26	11	4
5324	1981	PhD	Married	58293	1	0	19-01-2014	94	173	43
7446	1967	Master	Together	62513	0	1	09-09-2013	16	520	42
965	1971	Graduation	Divorced	55635	0	1	13-11-2012	34	235	65
6177	1985	PhD	Married	33454	1	0	08-05-2013	32	76	10
4855	1974	PhD	Together	30351	1	0	06-06-2013	19	14	0
5899	1950	PhD	Together	5648	1	1	13-03-2014	68	28	0
387	1976	Basic	Married	7500	0	0	13-11-2012	59	6	16
2125	1959	Graduation	Divorced	63033	0	0	15-11-2013	82	194	61
8180	1952	Master	Divorced	59354	1	1	15-11-2013	53	233	2
2569	1987	Graduation	Married	17323	0	0	10-10-2012	38	3	14
2114	1946	PhD	Single	82800	0	0	24-11-2012	23	1006	22
9736	1980	Graduation	Married	41850	1	1	24-12-2012	51	53	5
4939	1946	Graduation	Together	37760	0	0	31-08-2012	20	84	5
6565	1949	Master	Married	76995	0	1	28-03-2013	91	1012	80
2278	1985	2n Cycle	Single	33812	1	0	03-11-2012	86	4	17
9360	1982	Graduation	Married	37040	0	0	08-08-2012	41	86	2
5376	1979	Graduation	Married	2447	1	0	06-01-2013	42	1	1
1993	1949	PhD	Married	58607	0	1	23-12-2012	63	867	0
4047	1954	PhD	Married	65324	0	1	11-01-2014	0	384	0
1409	1951	Graduation	Together	40608	0	1	10-03-2013	68	270	2

```

Error in do_one(nmeth) : NA/NaN/Inf in foreign function call (arg 1)
In addition: Warning message:
In storage.mode(x) <- "double" : NAs introduced by coercion
>

```

Ya estan los datos NA eliminados pero aun asi genera el aviso y evita que genere ley de codo y clusters