



Carrera:

Ing. Tic's

Materia:

Ingeniería del Conocimiento

Semestre:

8º

Docente:

Flores Gallegos Eduardo.

Alumno:

Reyes Ibarra Sergio Antonio.



Práctica 1:

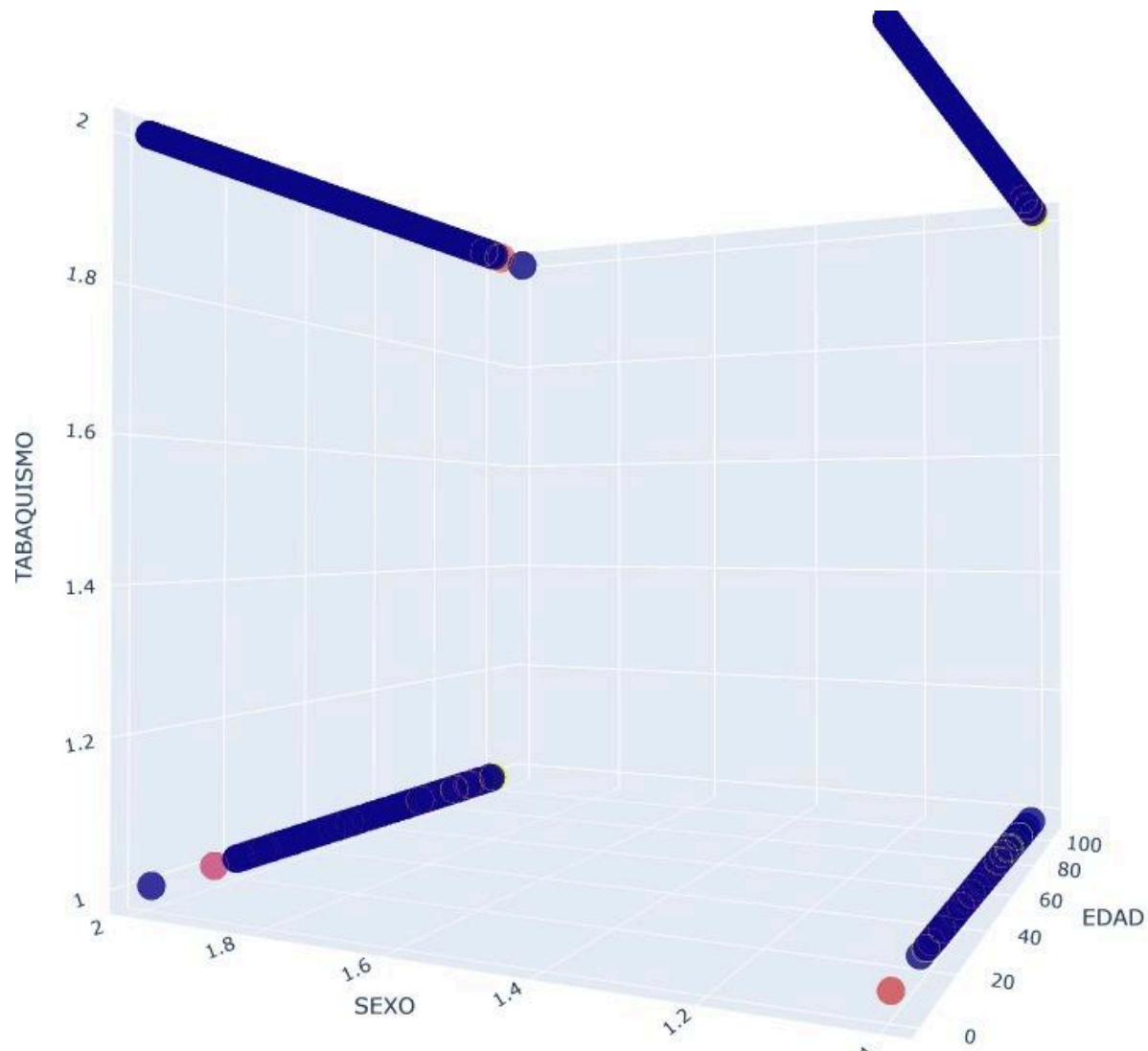
Segmentación de persona según edad, sexo y tabaquismo usando Agrupación Jerárquica. (Agrupación clustering)

Objetivo de la práctica:

El objetivo de la practica es la segmentación de personas que contagiaron covid19, esto gracias a los datos proporcionados por el gobierno federal de mexico, utilizando Agrupación Jerárquica es identificar grupos homogéneos de personas basados en su edad, sexo y si sufren de tabaquismo.

Técnica utilizada:

La técnica utilizada en la agrupación jerárquica es la agrupación aglomerativa, que es un enfoque iterativo que comienza tratando cada punto de datos como un clúster individual y luego fusiona gradualmente los clústeres más cercanos entre sí hasta que todos los puntos de datos estén en un solo clúster o en el número deseado de clústeres en este ejemplo son 3 clusters.

Grafica 3D:



Datos previamente filtrados en Rstudio:

Environment History Connections Tutorial

Import Dataset 122 MB

Data

- Aguascalientes 20665 obs. of 40 variables
- COVID19MEXICO 1296717 obs. of 40 variables
- datos_aguas 20665 obs. of 3 variables

Files Plots Packages Help Viewer Presentation

New Folder New Blank File Delete Rename More

C:\Users\seri\OneDrive\Escritorio\Taller_IA\Taller_IA\Hierarchical Cluste

Name

- COVID19MEXICO.csv
- Mail_Customers.csv
- segmented_customers.csv
- aguascalientes.csv
- RESULT.csv
- ags.csv

Console Terminal Background Jobs

```
R 4.4.0 ~ /> # EPOC <db1>, ASMA <db1>, INMUSUPR <db1>, HIPERTENSION <db1>, OTRA_COM <db1>, CARDIOVASCULAR <db1>, OBESIDAD <db1>,  
# RENAL_CRONICA <db1>, TABAQUISMO <db1>, OTRO_CASO <db1>, TOMA_MUESTRA_LAB <db1>, RESULTADO_LAB <db1>,  
# TOMA_MUESTRA_ANTIGENO <db1>, RESULTADO_ANTIGENO <db1>, CLASIFICACION_FINAL <db1>, MIGRANTE <db1>,  
# PAIS_NACIONALIDAD <chr>, PAIS_ORIGEN <chr>, UCI <db1>  
> View(COVID19MEXICO)  
>  
>  
> # Crear un subconjunto de los datos para el estado de Aguascalientes  
> # suponiendo que el archivo csv tiene una columna que identifica el estado  
> # y que se llama "ENTIDAD_RES" (ajustar según el nombre real de la columna)  
> datos_aguas <- subset(Aguascalientes, select = c(EDAD, SEXO, TABAQUISMO))  
> View(datos_aguas)  
> write.csv(datos_aguas, "ags.csv")  
> View(datos_aguas)  
> View(datos_aguas)  
> View(Aguascalientes)  
> aguascalientes2 <- data_frame2[datos_aguas$TABAQUISMO < 3, ]  
Error: objeto 'data_frame2' no encontrado  
> View(datos_aguas)  
> View(Aguascalientes)  
> aguascalientes2 <- data_frame2[datos_aguas$TABAQUISMO < 3, ]
```

Segundo filtro en Rstudio:

COVID19MEXICO x Aguascalientes x Untitled1* x datos_aguas x

Filter

	EDAD	SEXO	TABAQUISMO
1	43	1	2
2	44	2	2
3	30	1	2
4	80	1	2
5	1	2	2
6	54	1	2
7	22	2	2
8	39	2	2
9	23	2	2
10	43	1	2
11	44	1	2
12	37	1	2
13	63	2	2
14	64	1	2
15	35	1	2
16	28	2	2
17	22	2	2
18	35	2	2
19	37	2	2

Showing 1 to 19 of 20,665 entries, 3 total columns