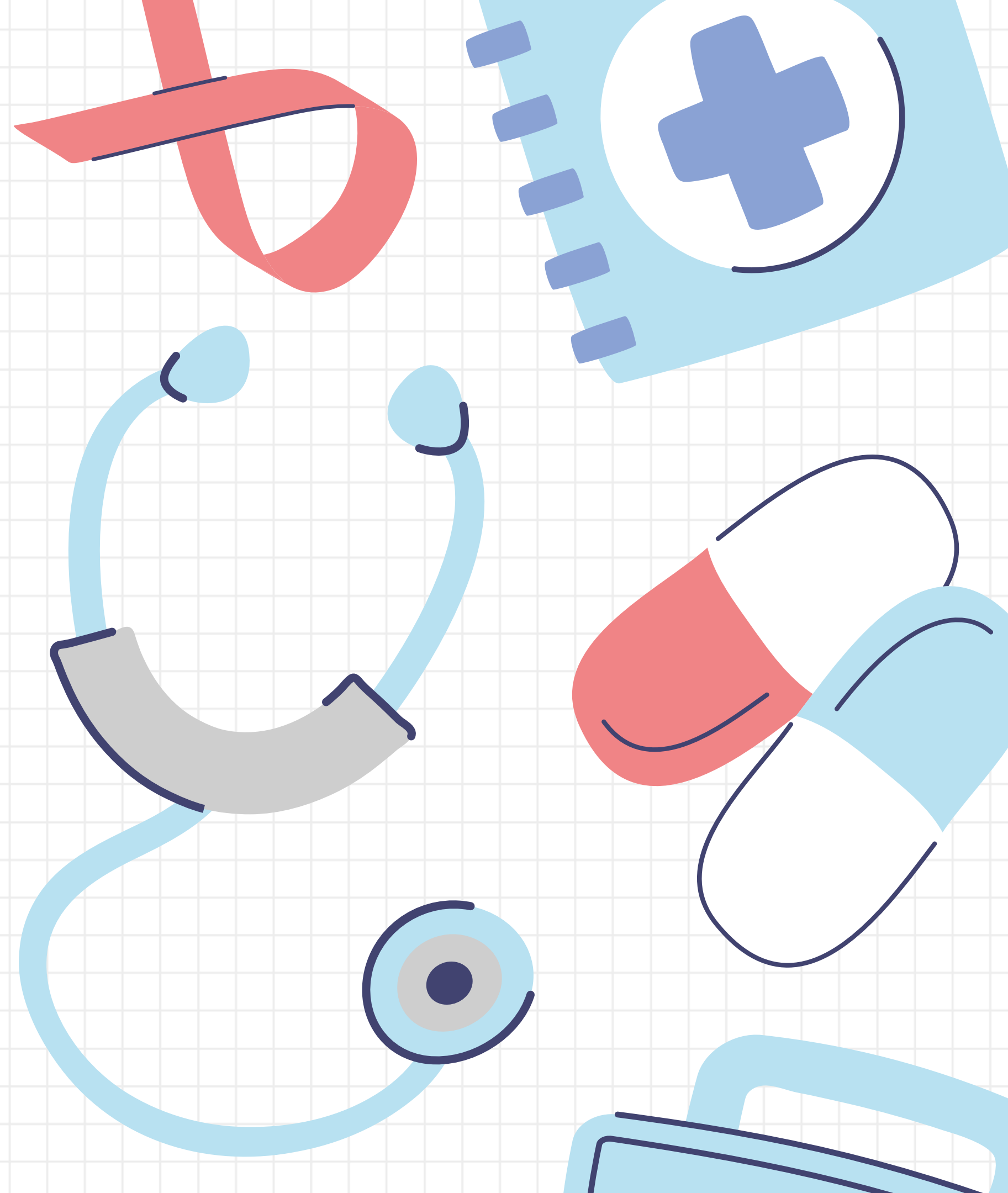
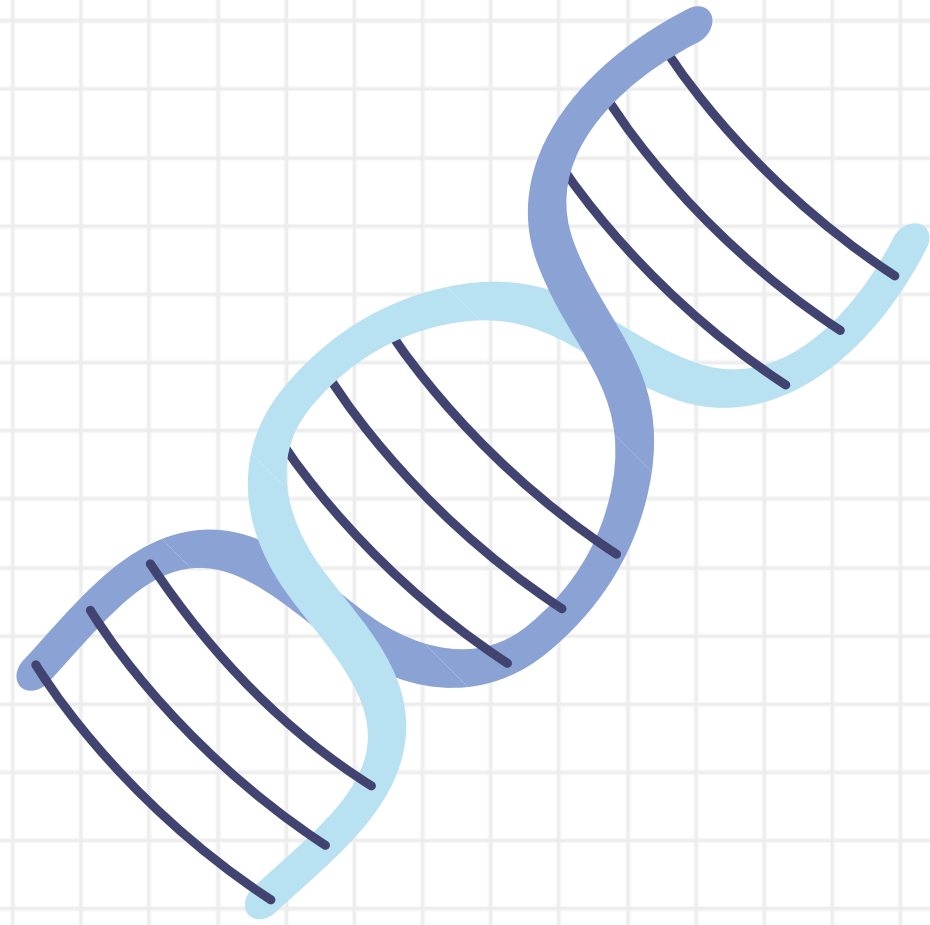


Cesar, Santiago y Jorge

Modelo de Agrupación

# Medi-Alpes

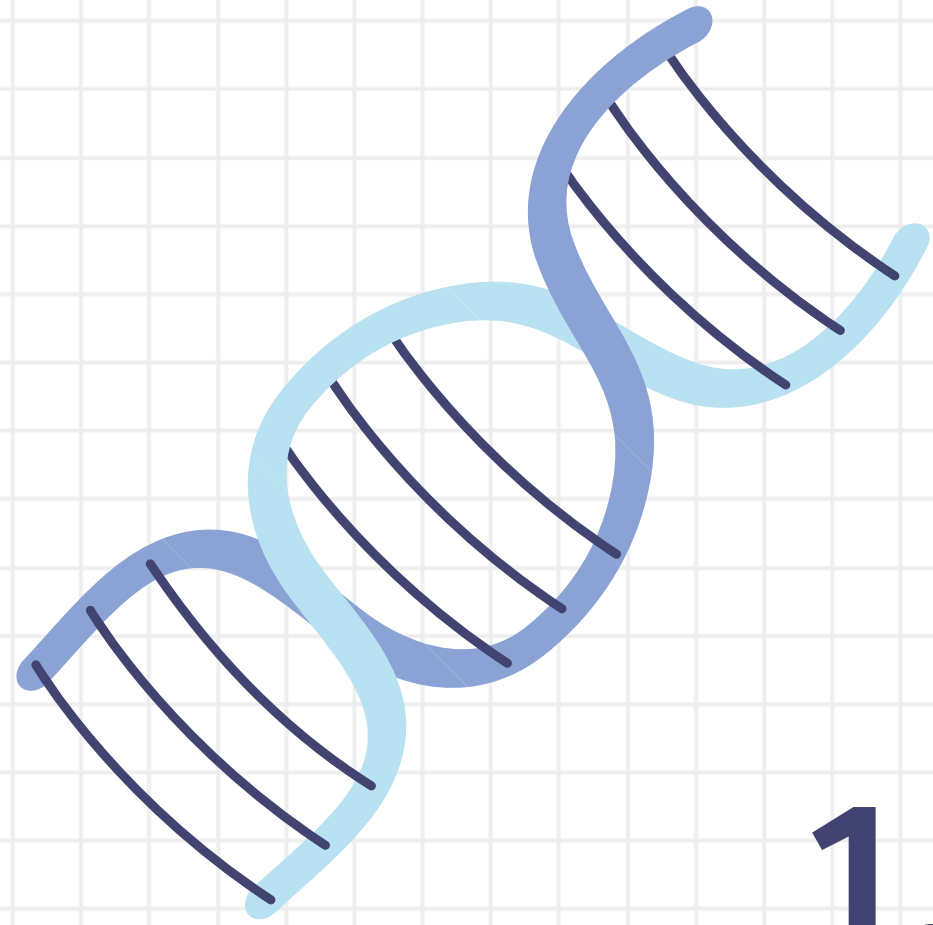




# Contenido

1. Entendimiento de los datos.
2. Análisis de calidad
3. Preparación Datos
4. Modelado e hiperparámetros
5. Evaluación cuantitativa
6. Evaluación cualitativa

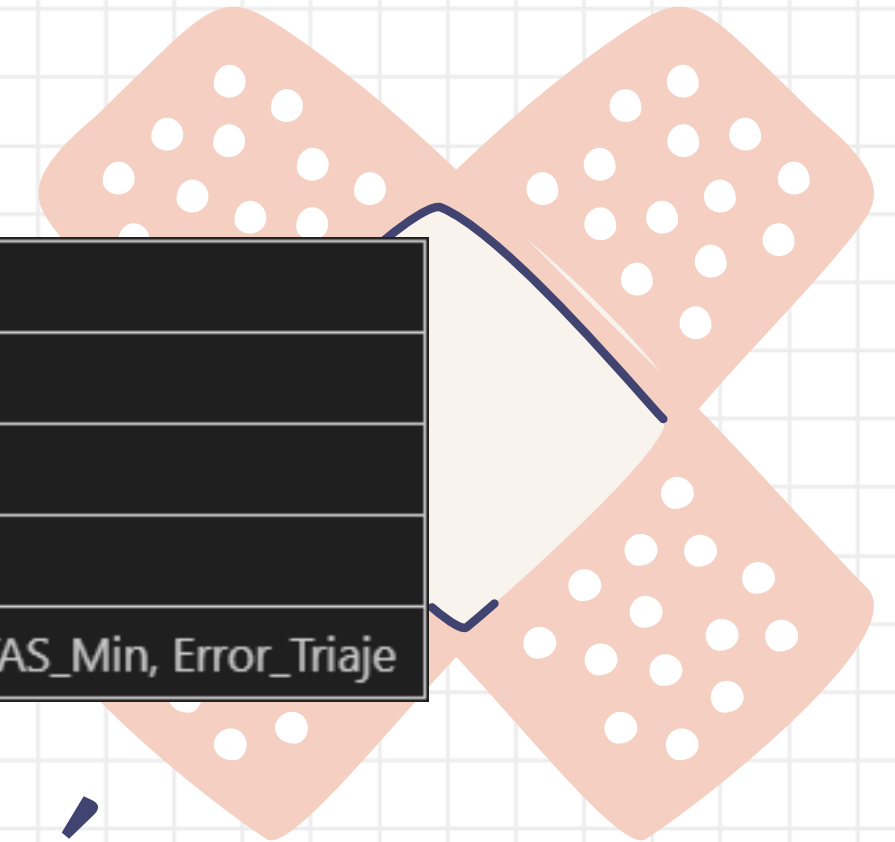




# 1. Entendimiento de los Datos



# Agrupación



Category	Fields
Demografía	Sexo, Edad, Grupo
Accidente	Modo Llegada, Lesión, Queja, Principal
Signos Vitales	Estado Mental, SBP, DBP, HR, RR, BT, Saturación, Dolor
Diagnósticos	dolor NRS, KTAS enfermera, Diagnóstico En Urgencias, Disposición, KTAS experto, Duración_Estancia_Min, Duración_KTAS_Min, Error_Triage

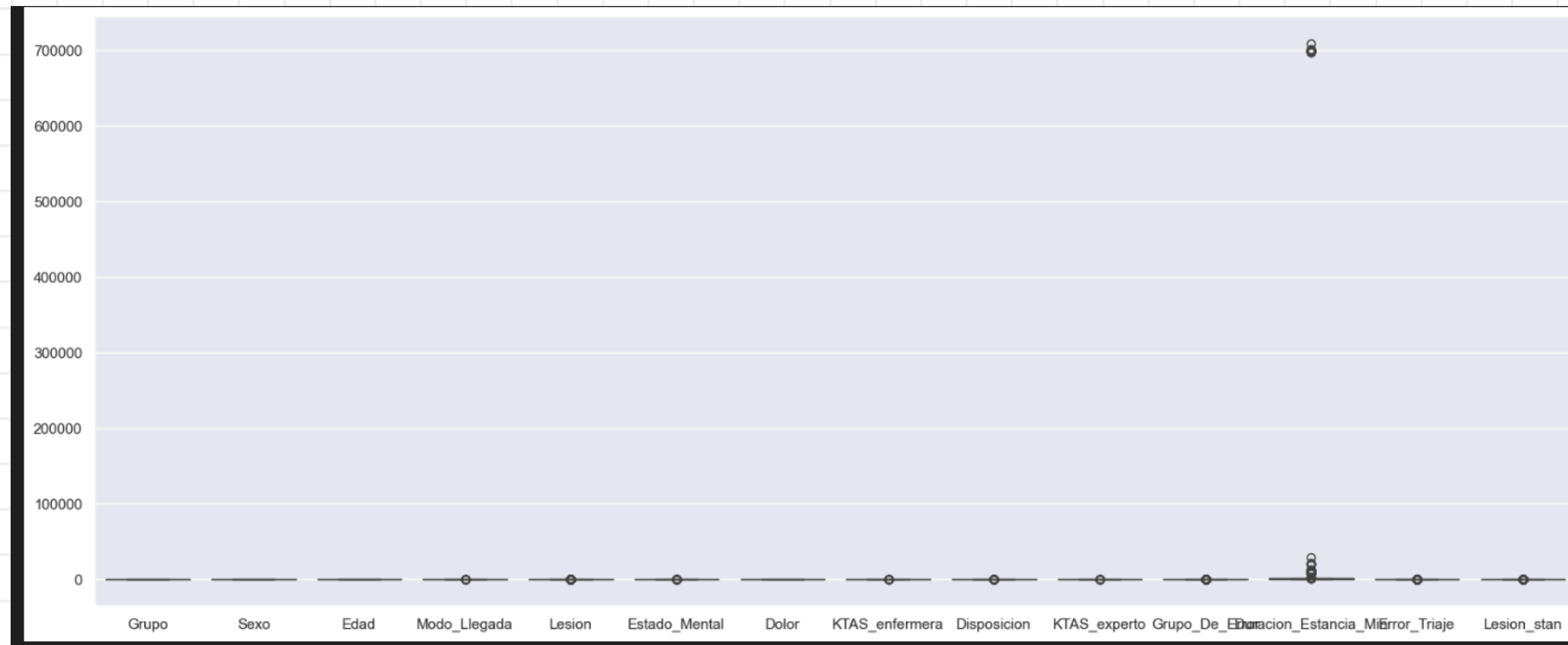
## Características más Importantes



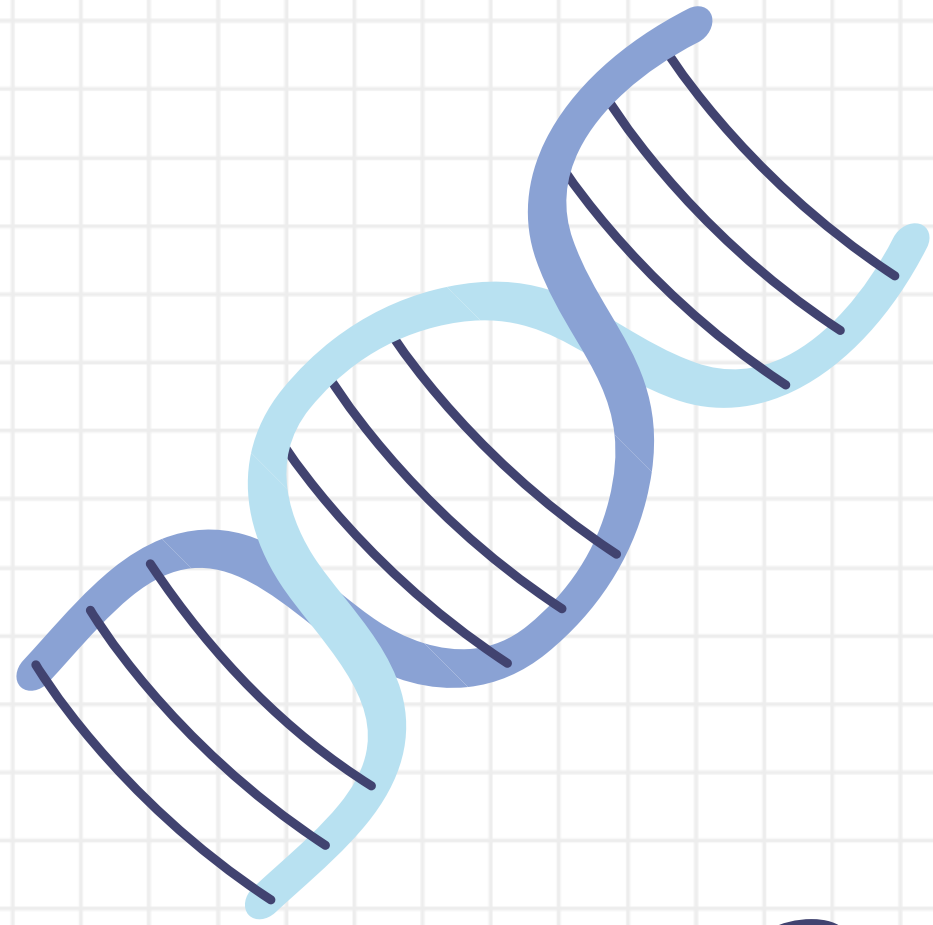
- Los signos vitales:
  - Estado mental
  - presión Arterial Sistólica
  - Presión Arterial Diastólica
  - Frecuencia cardíaca
  - Frecuencia Respiratoria
  - Temperatura Corporal
  - Saturación de Oxígeno.

- La edad de los pacientes.
- El Triage realizado por la enfermera y los expertos.
- El modo en el que llevan al hospital.

# Graficos



**En general los datos no tienen comportamientos muy atípicos (solo duración estancia)**



## 2. Análisis de calidad




# Compleitud

De manera general los datos no tienen problemas de completitud. Los que sí que lo tienen son los signos vitales, pero en este caso creemos que es por la gravedad del estado del paciente.

Grupo	0.000
Sexo	0.000
Edad	0.000
Modo_Llegada	0.000
Lesion	0.000
Queja_Principal	0.000
Estado_Mental	0.000
Dolor	0.000
dolor_NRS	0.000
SBP	0.017
DBP	0.021
HR	0.012
RR	0.015
BT	0.011
Saturacion	0.556
KTAS_enfermera	0.000
Diagnostico_En_Urgencias	0.001
Disposicion	0.000
KTAS_experto	0.000
Grupo_De_Error	0.000
Duracion_Estancia_Min	0.000
Duracion_KTAS_Min	0.000
Error_Triaje	0.000
Lesion_stan	0.000

# Unicidad

**Aunque nuestro análisis nos diga que hay datos duplicados, no podemos eliminar los datos sin más porque nada nos asegura que sean duplicados, puede ser que hayan dos pacientes que tengan los mismos problemas al momento de llegar al hospital.**



66



# Consistencia

Como podemos observar, dolor\_NRS tiene un problema. Hay datos con el valor #BOP!, los cuales representan una gran cantida del total.

```
Estado_Mental
1      936
2       33
3       23
4        8
Name: count, dtype: int64
```

```
dolor_NRS
#BOP!    440
3        209
4        111
5        109
6         59
2         31
7         27
8          8
10         3
1          2
9          1
Name: count, dtype: int64
```

```
KTAS_experto
3      391
4      369
2      171
5       50
1       19
Name: count, dtype: int64
```

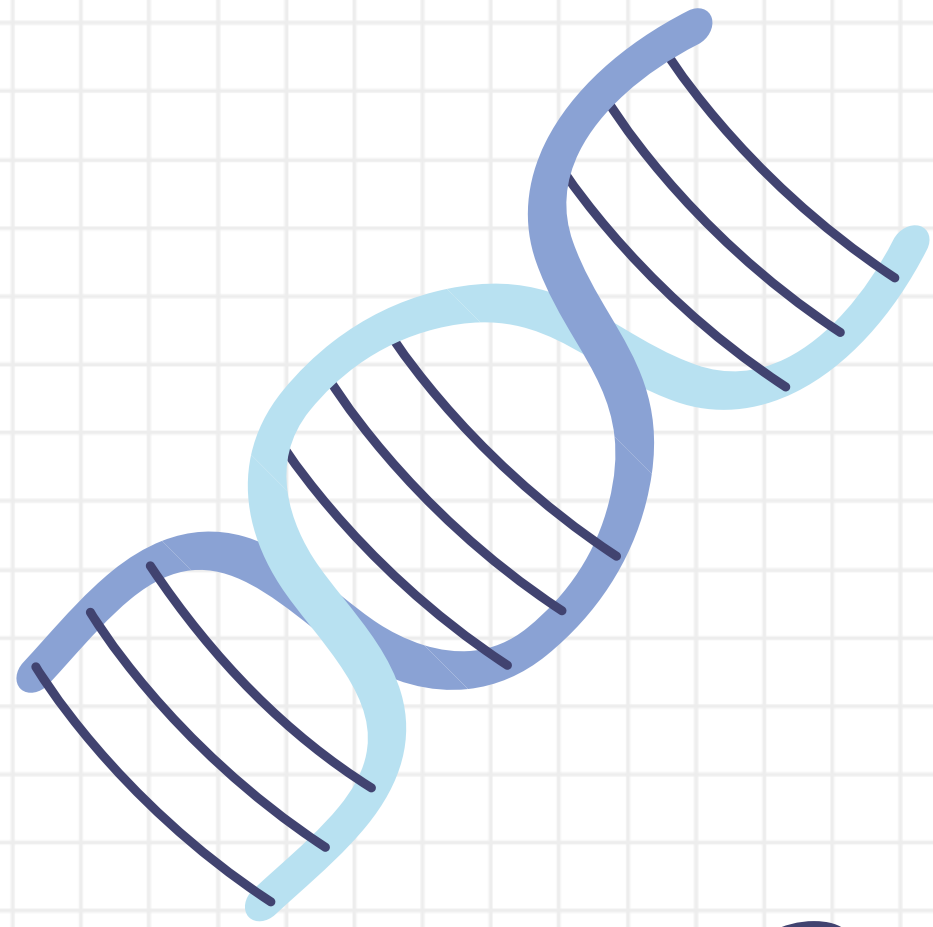
# Validez

Como podemos observar, la validez de estos datos es correcta, el único problema como mencionamos anteriormente está en dolor\_NRS.

```
array([2, 3, 4, 1, 5], dtype=int64)
```

```
array(['2', '4', '3', '#B0b!', '6', '5', '7', '8', '1', '9', '10'],  
      dtype=object)
```

```
array([1, 4, 2, 3], dtype=int64)
```



# 3. Preparación Datos



# Calidad dolor\_NRS

Primero el dolor reportado por la enfermera, el cual tiene un 44% de sus entradas en null.

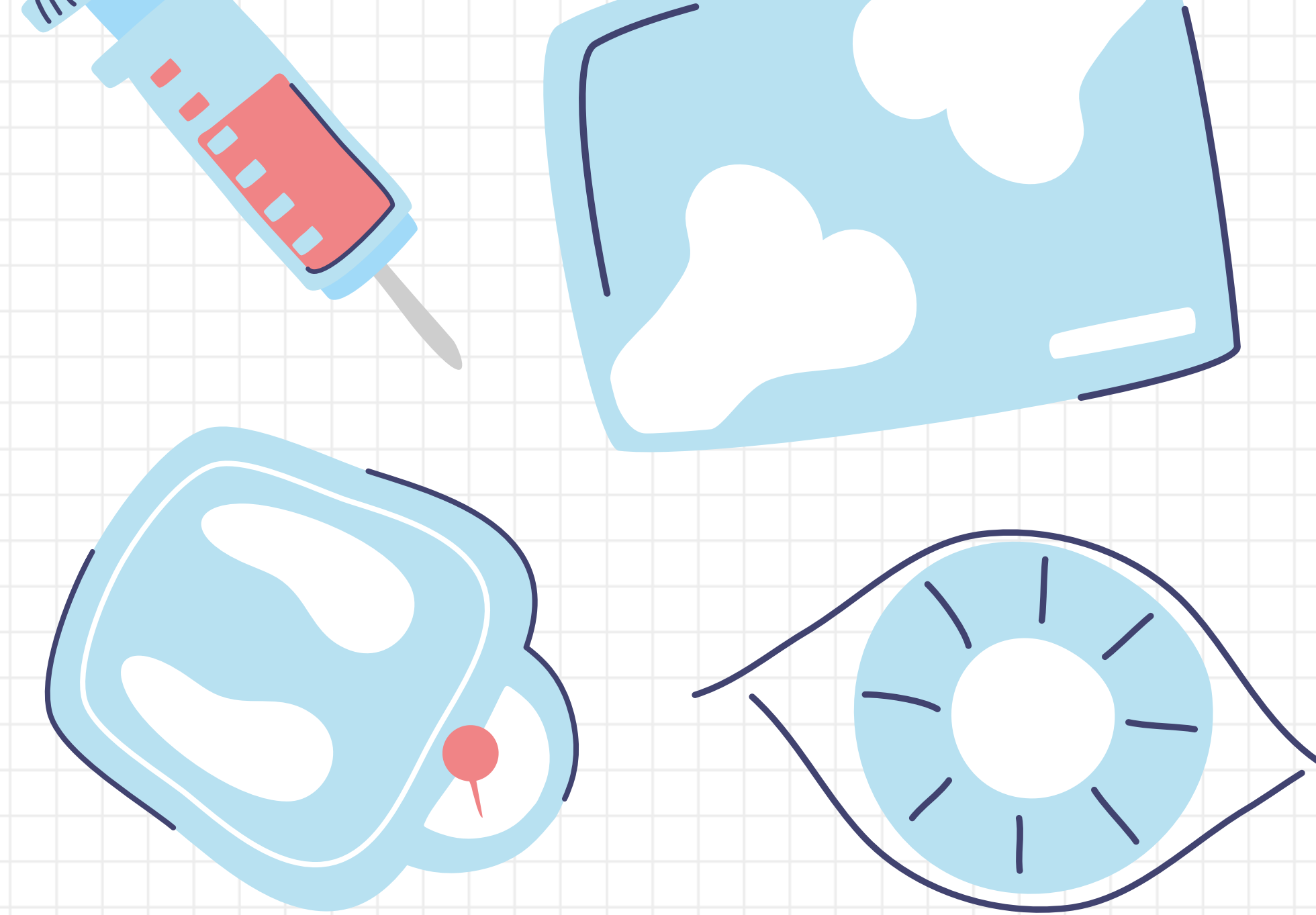
Nos dimos cuenta de que estas se correspondían cuando el paciente no tenía dolor, por lo que las asignamos a 0.

Esto tenía una excepción en 2 entradas, las cuales registraban que el paciente tenía dolor, pero no especificaba cuál, por lo que le asignamos la media.

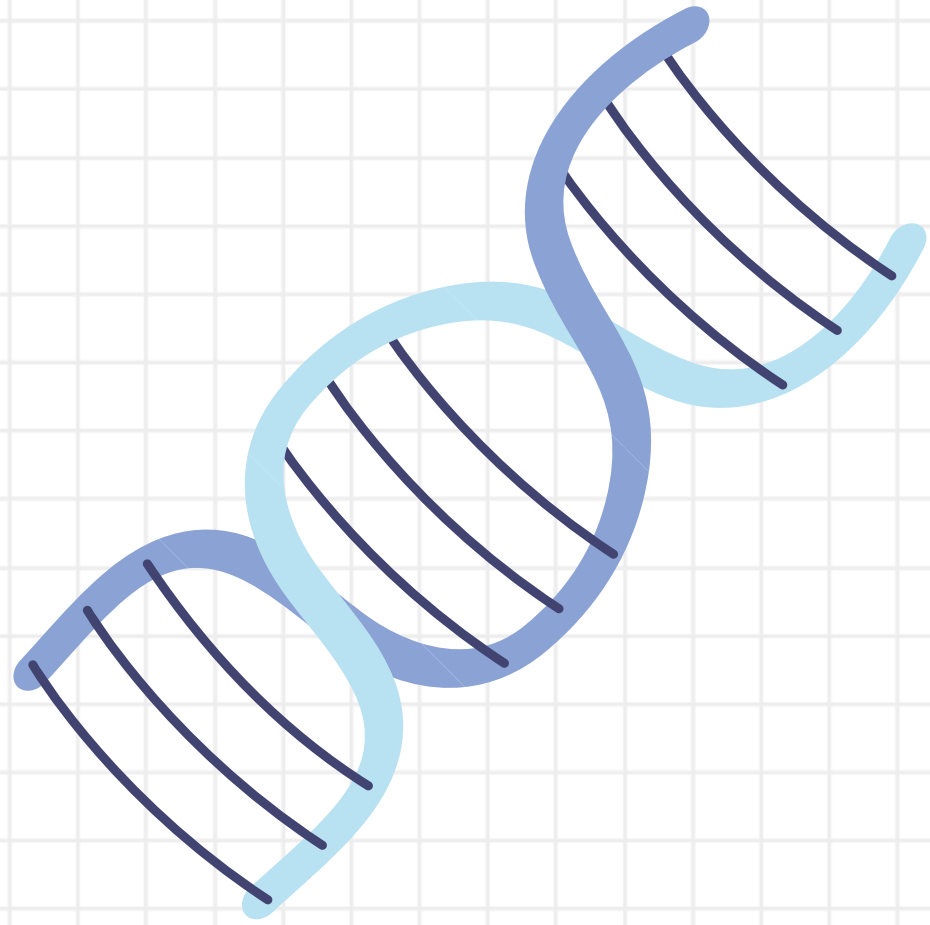
dolor_NRS	
0	438
3	209
4	111
5	109
6	59
2	33
7	27
8	8
10	3
1	2
9	1

**Hicimos uso del Early Warning Score (EWS) es una herramienta clínica utilizada para identificar a los pacientes que están en riesgo de deterioro.**

**El cálculo del EWS generalmente se basa en una serie de parámetros fisiológicos como la frecuencia cardíaca, la presión arterial, la temperatura, la frecuencia respiratoria, y el nivel de conciencia.**



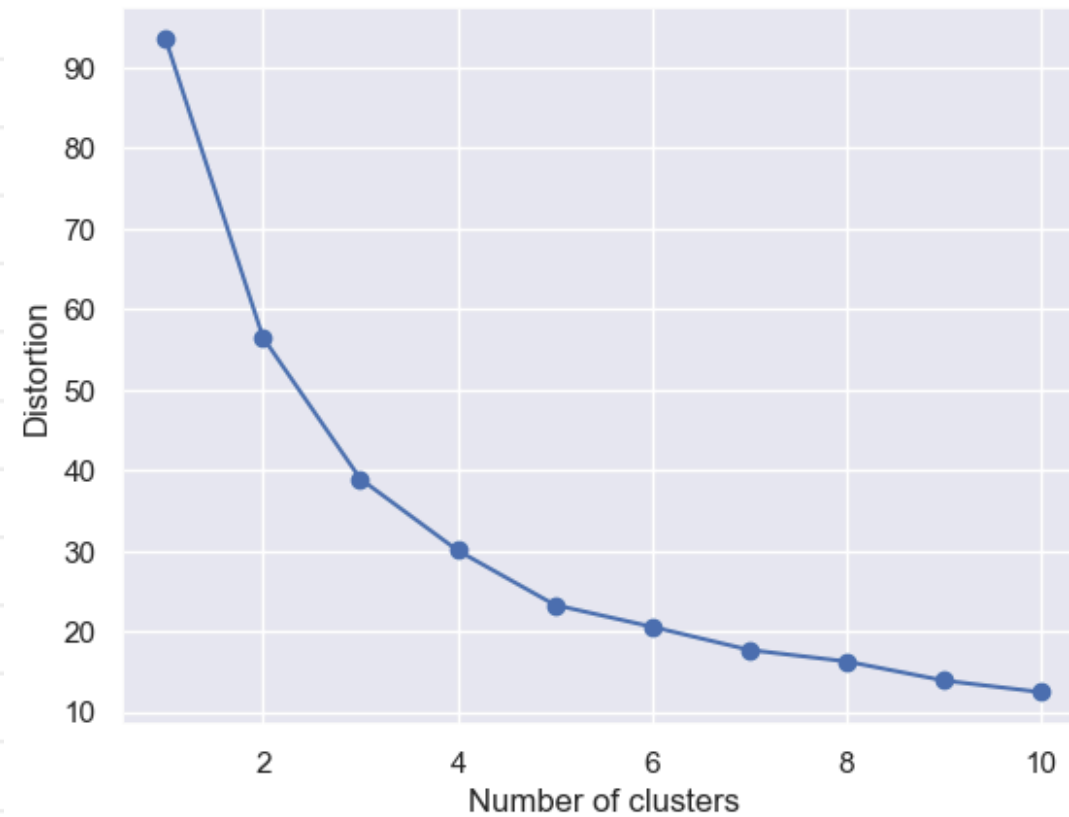
**EWS**



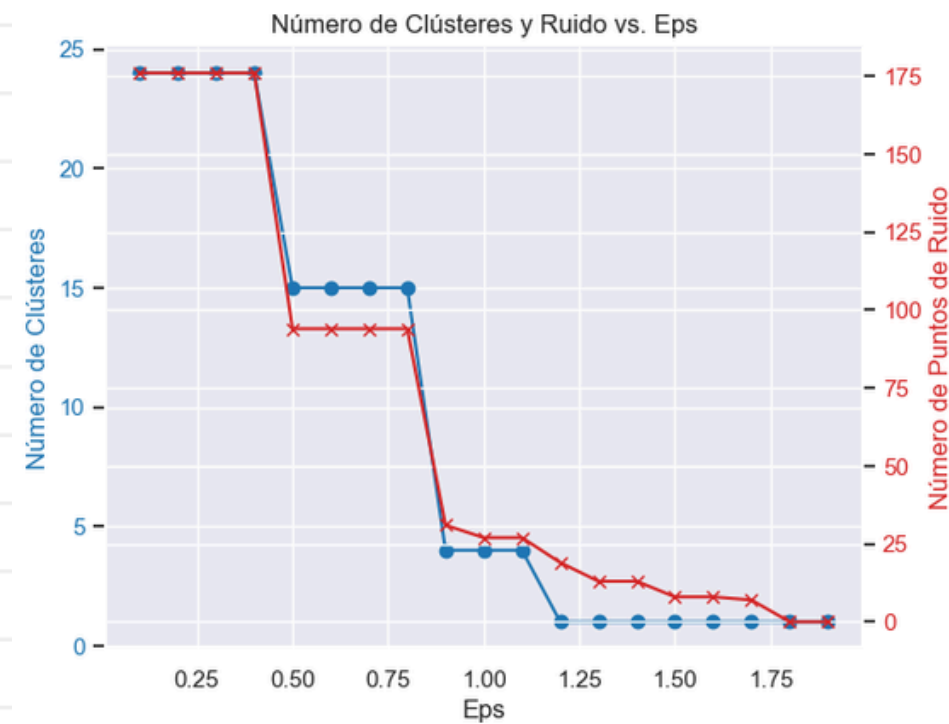
## 4. Modelado



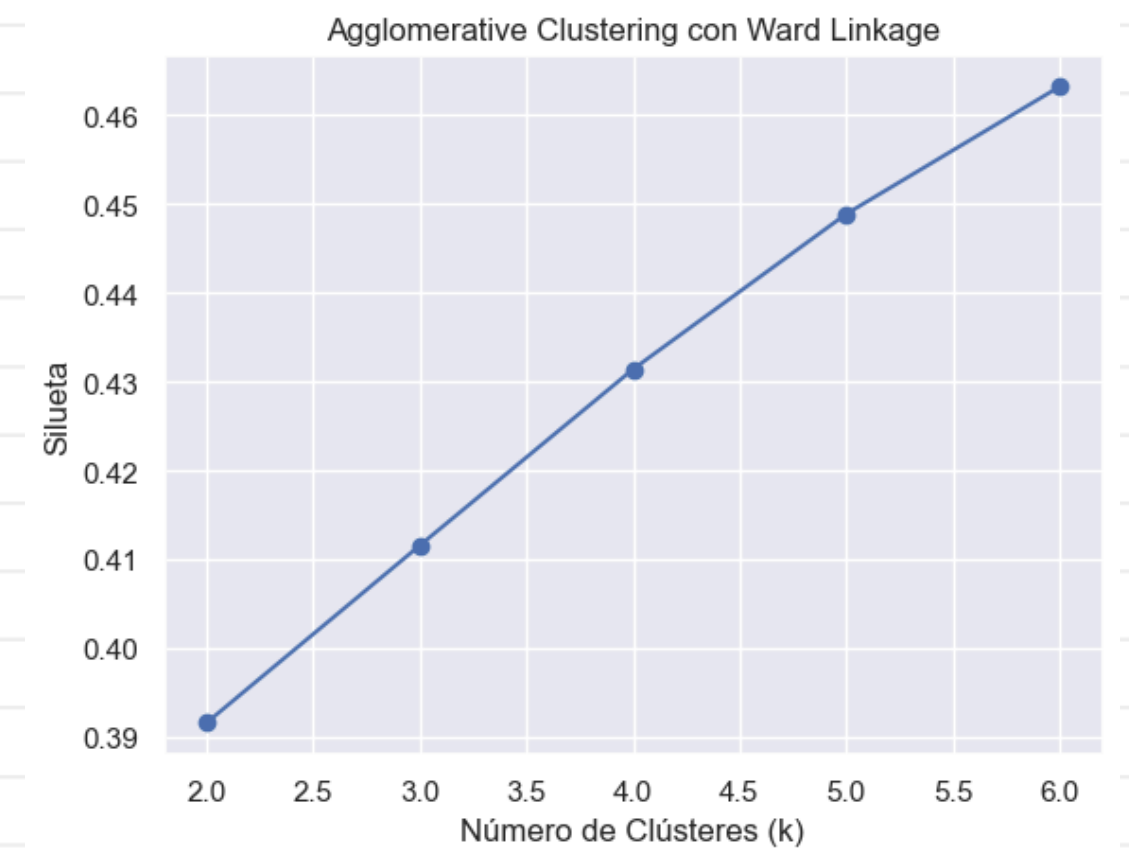
# Número de clusters para cada modelo



Kmeans

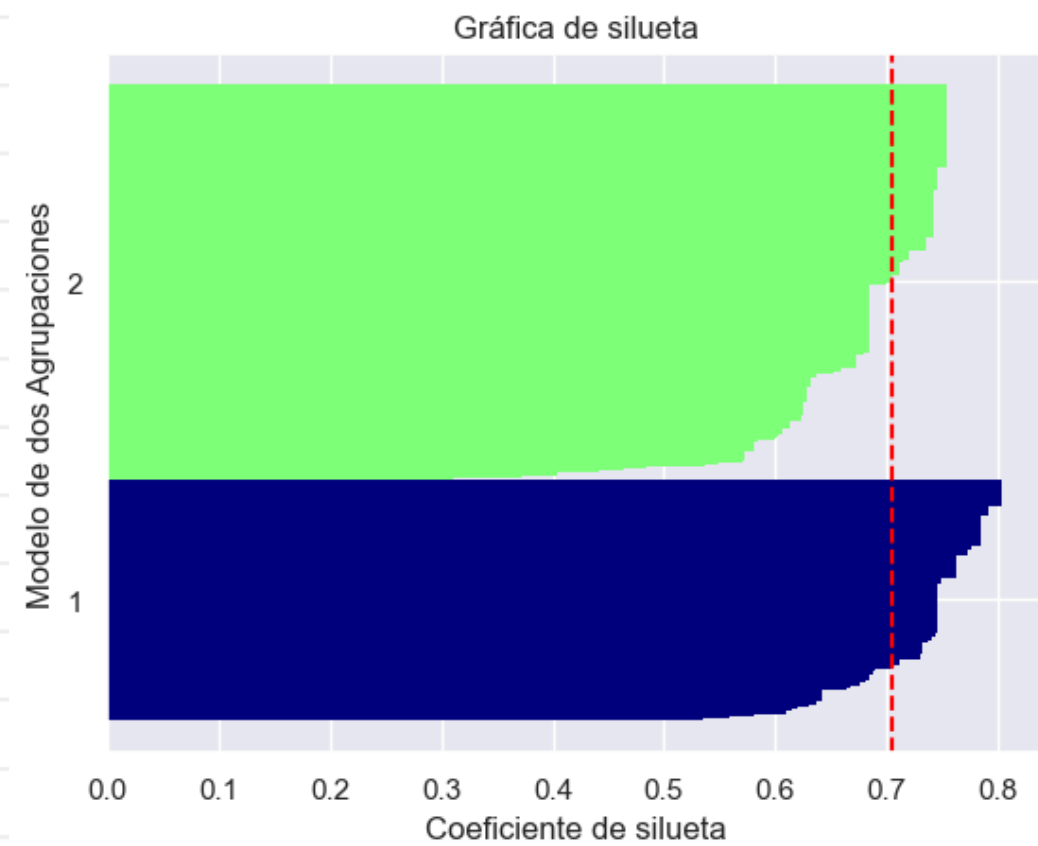


DBSCAN

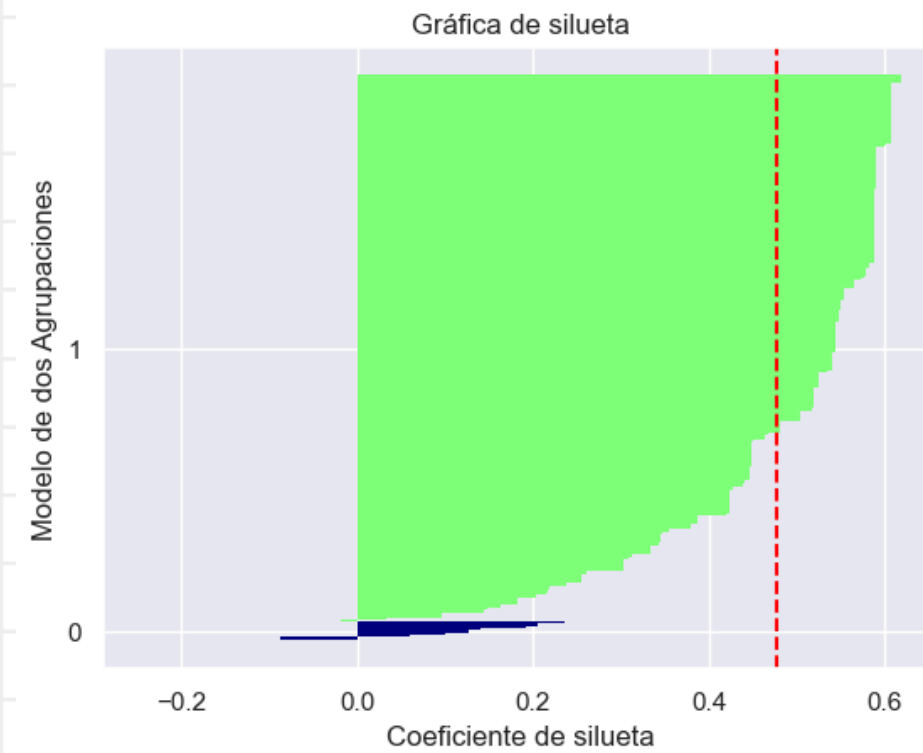


AgglomerativeClustering

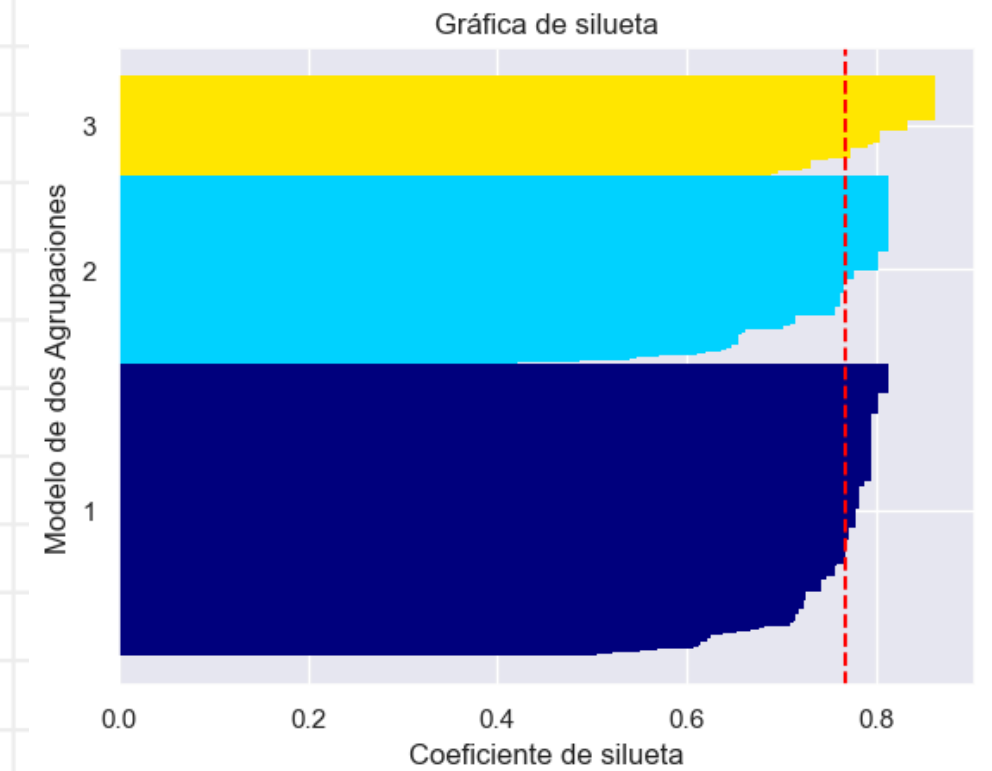
# Gráfica de silueta para cada modelo



Kmeans

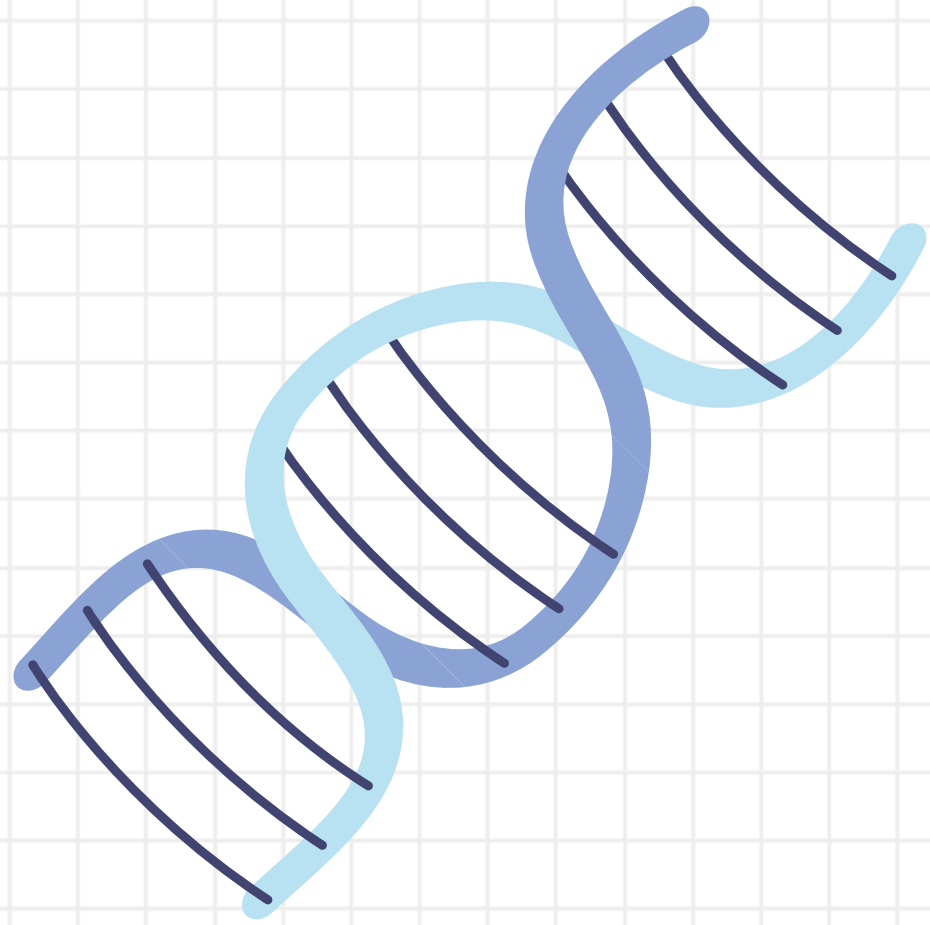


DBSCAN

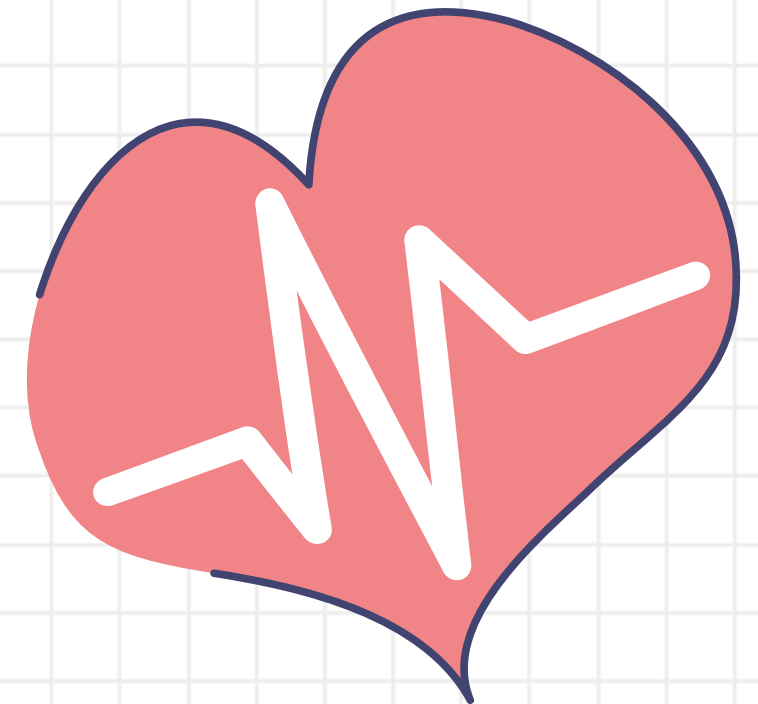


AgglomerativeClustering





## 5. Evaluación Cualitativa



# Análisis y conclusiones

A nivel cuantitativo, en el último modelo, AgglomerativeClustering, se observa un coeficiente de silueta cercano a 0.8, lo que indica una buena cohesión en general. Sin embargo, en el grupo 1 (azul oscuro) y el grupo 2 (celeste) hay registros con coeficientes por debajo de 0.6, lo que sugiere posibles valores atípicos o una mala asignación. El segundo modelo, DBSCAN, muestra un coeficiente de silueta moderado, cercano a 0.4 para el grupo 1 (verde), lo que indica una agrupación razonable, aunque con registros cercanos a 0.0, sugiriendo una mala asignación, especialmente en el grupo 2 (azul oscuro), donde se observan coeficientes negativos, lo que es preocupante y sugiere que estos registros están mal agrupados. Finalmente, en el primer modelo, KMeans, la mayoría de los registros tiene coeficientes cercanos a 0.7, lo que indica una buena agrupación; sin embargo, en el grupo 1 (verde) y el grupo 2 (azul oscuro) hay registros con coeficientes inferiores a 0.6, lo que también indica posibles valores atípicos o mala asignación de ciertos puntos.

# Análisis y conclusiones

A nivel cualitativo, aunque seleccionamos las variables EWS, DOLOR\_NRS y KTAS\_Experto porque inicialmente parecían mostrar correlaciones útiles para agrupar a los pacientes de manera coherente, los resultados de los modelos de agrupación sugieren lo contrario. Si bien era razonable esperar que un peor EWS y un mayor nivel de dolor se relacionaran con un peor registro en KTAS\_Experto, los resultados muestran que esta relación no es consistente en los datos. El análisis mediante AgglomerativeClustering, DBSCAN y KMeans confirma que los datos son muy dispersos y no presentan clústeres naturales claramente diferenciados. En AgglomerativeClustering, aunque algunos grupos tienen un coeficiente de silueta alto, existen registros con mala cohesión. DBSCAN destaca la presencia de puntos atípicos y agrupaciones inconsistentes, con coeficientes negativos y cercanos a cero, lo que indica que el modelo no es capaz de captar patrones claros. Por último, KMeans presenta grupos más coherentes, pero aún con registros mal asignados o atípicos.

En conclusión, después de probar múltiples modelos, los algoritmos de agrupación no parecen ofrecer un valor significativo para la toma de decisiones rápidas en el hospital, ya que no logran identificar patrones claros y consistentes en los datos que puedan ser útiles para la clasificación de pacientes y la optimización de decisiones en el mismo.