En relación al EDA (Exploratory data analysis) se hallaron diversas cuestiones en relación al dataset implementado.

En las primeras y últimas filas podemos observar una distribución uniforme en cuando a los valores de cada columna, no hay valores nulos, por lo tanto, no es necesaria la eliminación ni alteración de ninguna de sus filas. Luego, mediante gráficos de Seaborn y Matplotlib podemos ver los picos más altos en la distribución de edades que va desde mediados de 50 años de edad hasta los mediados de 60 años, luego, en distribución de edad en pacientes positivos de cardiopatía vemos una tendencia bastante regular en pacientes de mediana edad, no así, en pacientes más jóvenes o longevos.

El diagrama de caja de edades según variable objetivo nos arroja un solo outlier dentro de los casos negativos. Después, en las relaciones entre variables del dataset nos fijamos en tendencias claras para ir eligiendo las variables importantes para nuestro modelo de predicción y terminamos de corroborar esas tendencias con dos gráficos Pairplot y un gráfico de correlaciones.

Por último, se realizó el modelo de predicción mediante la utilización de Maquinas de soportes vectoriales (SVM) y se obtuvo un f1-score de 1.0 y opté por no modificar el modelo ya que al ser un dataset de pequeñas cantidades de datos es más viable deducir que se deba al modelo y no significativamente a un sobreajuste del modelo en sí.