

```
#Ejercicio 1
lista1=[1,2,3,4,5]
lista2=['a','b','c','d','e']
```

```
#Ejercicio 2
d_test=pd.DataFrame({'columna_A':lista1, 'columna_B':lista2})
d_test
```

```
#Ejercicio 3
url='https://raw.githubusercontent.com/fpineyro/homework-0/master/heart_disease.csv'
d_hd=pd.read_csv(url)
d_hd
```

```
#Ejercicio 3.1) Imprimir la variable `age` y `has_heart_disease` al mismo tiempo.
print(d_hd[["age", "has_heart_disease"]])
```

```
#Ejercicio 3.2) Explorar los datos con DataTable de google colab, usar filtros y ordenar las columnas haciendo click en ellas
from google.colab import data_table
```

```
# Activar el DataTable
data_table.enable_dataframe_formatter()
```

```
# Mostrar la tabla interactiva
d_hd
```

```
#Ejercicio 3.3) Imprimir los primeros 3 registros con head, y los últimos 3 con tail
```

```
# head
d_hd.head(3)
```

```
#tail
d_hd.tail(3)
```

```
#Ejercicio 3.4) Acceder con loc a la columna has_heart_disease, y mostrar todos los registros
```

```
# Acceder a la columna 'has_heart_disease' con loc
print(d_hd.loc[:, "has_heart_disease"])
```

```
#Ejercicio 3.5) Idem 3.4) peor mostrando los primeros 3 registros. Usar
slice, no head.
# Mostrar solo los primeros 3 registros de la columna
'has_heart_disease' usando slice
print(d_hd.loc[:, "has_heart_disease"][:3])

#Ejercicio 3.6) Usar iloc para obtener las primeras 3 columnas, y 5
registros
# Usar iloc para obtener las primeras 3 columnas y 5 registros
print(d_hd.iloc[0:5,0:3])

#Ejercicio 4.3) ¿Qué variables tienen nulos? (usar status)
# Ver cuántos valores nulos tiene cada variable
status = d_hd.isnull().sum()
print(status)

#Ejercicio 4.4) Devolver un data frame con todos los registros que
tienen NaN en la variable "thal".
# Filtrar los registros donde 'thal' es NaN
thal_nulls = d_hd[d_hd["thal"].isnull()]

print(thal_nulls)

#Ejercicio 4.5) Reemplazar los nulos que aparecen en la variable "thal"
por 999. *No* guardarlo en el dataframe original. Usar fillna.
# Reemplazar los NaN en 'thal' por 999, sin modificar el DataFrame
original
thal_filled = d_hd["thal"].fillna(999)

print(thal_filled)

#Ejercicio 4.6) Reemplazar los nulos que aparecen en la variable "thal"
por 999 (como antes),
# y también reemplazar los nulos de "num_vessels_flour" por el promedio
de la variable.
# Guardar el resultado en un dataframe nuevo llamado "d_hd2".
# Crear una copia del dataframe original
d_hd2 = d_hd.copy()

# Reemplazar NaN en 'thal' por 999
d_hd2["thal"] = d_hd2["thal"].fillna(999)

# Reemplazar NaN en 'num_vessels_flour' por el promedio de la columna
```

```
mean_vessels = d_hd2["num_vessels_flour"].mean()
d_hd2["num_vessels_flour"] =
d_hd2["num_vessels_flour"].fillna(mean_vessels)

# Mostrar las primeras filas del nuevo dataframe
print(d_hd2.head())

#Ejercicio 4.7) Chequear resultado con status
# Revisar si quedan nulos en el nuevo dataframe d_hd2
status = d_hd2.isnull().sum()
print(status)

#Ejercicio 4.8)
# Eliminar todas los registros con nulos usando `dropna` de pandas
# Crear un nuevo DataFrame eliminando todas las filas con nulos
d_hd_no_na = d_hd.dropna()

# Mostrar las primeras filas del resultado
print(d_hd_no_na.head())

# Chequear si quedaron nulos
print(d_hd_no_na.isnull().sum())
```