

```
import warnings
warnings.filterwarnings("ignore")
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

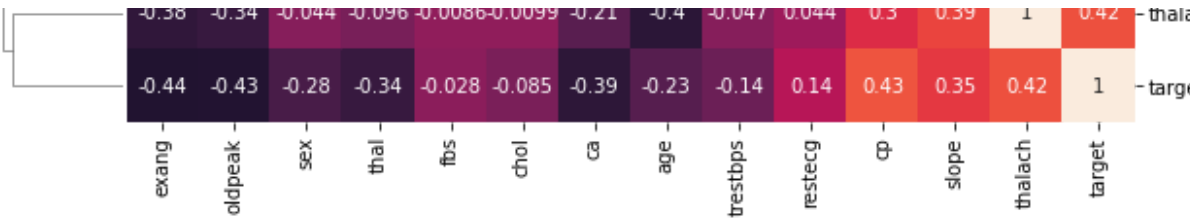
```
dt = pd.read_csv("heart.csv")
desired_width = 520
pd.set_option('display.width', desired_width)

np.set_printoptions(linewidth=desired_width)

pd.set_option('display.max_columns',25)
```

```
print(dt.head(10))
print((dt.info()))
print(dt.describe(include='all'))
print(dt.columns)
print(dt.isnull().sum())
sns.clustermap(dt.corr(),annot = True)
plt.show()
```





```
dt.columns = ['age', 'sex', 'chest_pain_type', 'resting_blood_pressure', 'cholesterol', 'fasting_blood_sugar', 'max_heart_rate', 'exercise_induced_angina', 'major_vessels', 'resting_heart_rate', 'slope', 'thalach', 'target']
```

```
dt['sex'][dt['sex'] == 0] = 'female'
dt['sex'][dt['sex'] == 1] = 'male'
#print(len(dt.columns))
dt['chest_pain_type'][dt['chest_pain_type'] == 0] = 'typical angina'
dt['chest_pain_type'][dt['chest_pain_type'] == 1] = 'atypical angina'
dt['chest_pain_type'][dt['chest_pain_type'] == 2] = 'non-anginal pain'
dt['chest_pain_type'][dt['chest_pain_type'] == 3] = 'asymptomatic'

dt['fasting_blood_sugar'][dt['fasting_blood_sugar'] == 0] = 'lower than 120mg/ml'
dt['fasting_blood_sugar'][dt['fasting_blood_sugar'] == 1] = 'greater than 120mg/ml'

dt['rest_ecg'][dt['rest_ecg'] == 0] = 'normal'
dt['rest_ecg'][dt['rest_ecg'] == 1] = 'ST-T wave abnormality'
dt['rest_ecg'][dt['rest_ecg'] == 2] = 'left ventricular hypertrophy'

dt['exercise_induced_angina'][dt['exercise_induced_angina'] == 0] = 'no'
dt['exercise_induced_angina'][dt['exercise_induced_angina'] == 1] = 'yes'

dt['st_slope'][dt['st_slope'] == 0] = 'upsloping'
dt['st_slope'][dt['st_slope'] == 1] = 'flat'
dt['st_slope'][dt['st_slope'] == 2] = 'downsloping'

dt['thalassemia'][dt['thalassemia'] == 1] = 'normal'
dt['thalassemia'][dt['thalassemia'] == 2] = 'fixed defect'
dt['thalassemia'][dt['thalassemia'] == 3] = 'reversable defect'

dt['sex'] = dt['sex'].astype('object')
dt['chest_pain_type'] = dt['chest_pain_type'].astype('object')
dt['fasting_blood_sugar'] = dt['fasting_blood_sugar'].astype('object')
dt['rest_ecg'] = dt['rest_ecg'].astype('object')
dt['exercise_induced_angina'] = dt['exercise_induced_angina'].astype('object')
dt['st_slope'] = dt['st_slope'].astype('object')
dt['thalassemia'] = dt['thalassemia'].astype('object')

print(dt.head(30))
print(dt.dtypes)
dt = pd.get_dummies(dt,prefix=['st_slope'],columns=['st_slope'])
dt = pd.get_dummies(dt, drop_first=True)
print(dt.columns)
print(dt.head(30))
print(len(dt.columns))
sns.countplot(x='target',data=dt,hue = 'sex_male')
print(dt['target'].value_counts())
print(dt['sex_male'].value_counts())
plt.show()
sns.clustermap(dt.corr().annot = True)
```

```
plt.show()
```



```

1    165
0    138
Name: target, dtype: int64
1    207
0     96
Name: sex_male, dtype: int64

```

