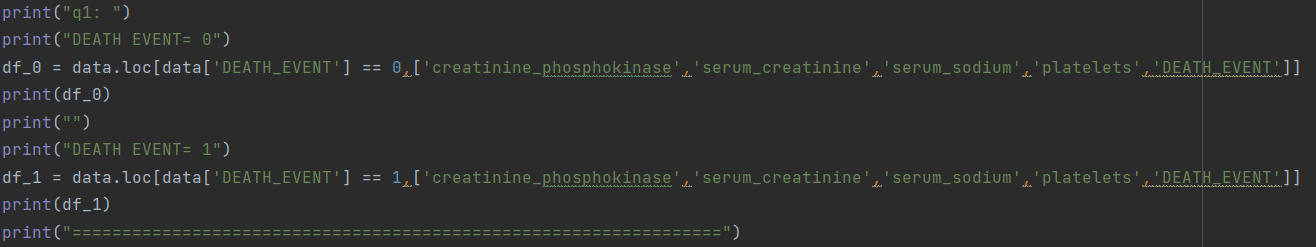
**Assignment 8 Heart**

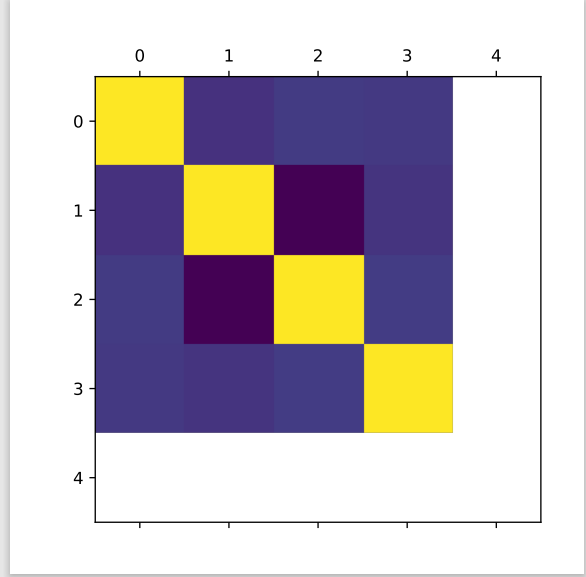
# ***Question 1:***

## **1. load the data into Pandas dataframe. Extract two dataframes with the above 4 features: df 0 for surviving patients (DEATH EVENT = 0) and df 1 for deceased patients (DEATH EVENT = 1)**

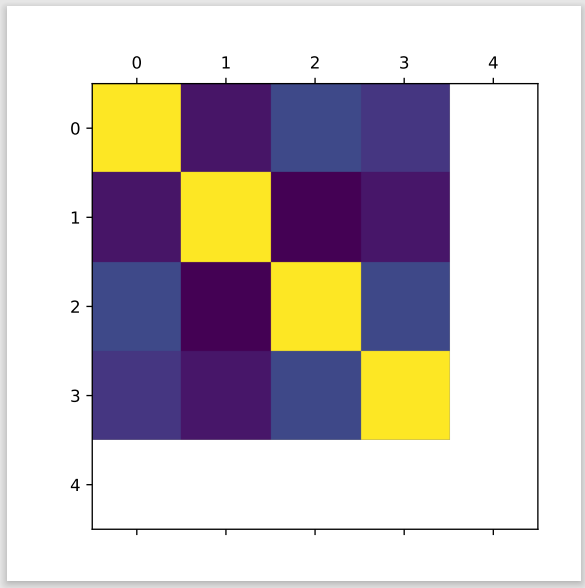


## **2. for each dataset, construct the visual representations of correponding correlation matrices M0 (from df 0) and M1 (from df 1) and save the plots into two separate files**

Plot of Surviving



Plot of Deceased



## **3. examine your correlation matrix plots visually and answer the following:**

### **(a) which features have the highest correlation for surviving patients?**

Serum\_sodium has the highest correlation for surviving patients

### **(b) which features have the lowest correlation for surviving patients?**

Serum\_creatinine has the highest correlation for surviving patients

### **(c) which features have the highest correlation for deceased patients?**

Serum\_sodium has the highest correlation for deceased patients

### **(d) which features have the lowest correlation for deceased patients?**

Serum\_creatinin has the highest correlation for deceased patients

### **(e) are results the same for both cases?**

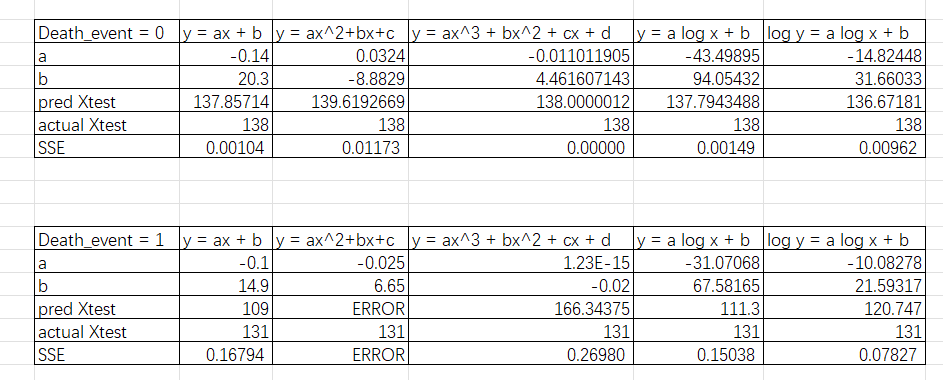
Yes

# ***Question 2:***

The last digit of my BUid is 9, so my group is Group3.

x: serum sodium

y: serum creatinine



# ***Question 3:***

| Model | SSE (death event=0) | (death event=1) |
| --- | --- | --- |
|  | 0.00104 | 0.16794 |
|  | 0.01173 | Error(The result of pred Xtest brings i, so it is the error) |
|  | 0 | 0.26980 |
|  | 0.00149 | 0.15038 |
|  | 0.00962 | 0.07827 |

## **1. which model was the best (smallest SSE) for surviving patients? for deceased patients?**

is the best for surviving patients

is the best for deceased patients

## **2. which model was the worst (largest SSE) for surving patients? for deceased patients?**

is the worst for surviving patients

is the worst for deceased patients