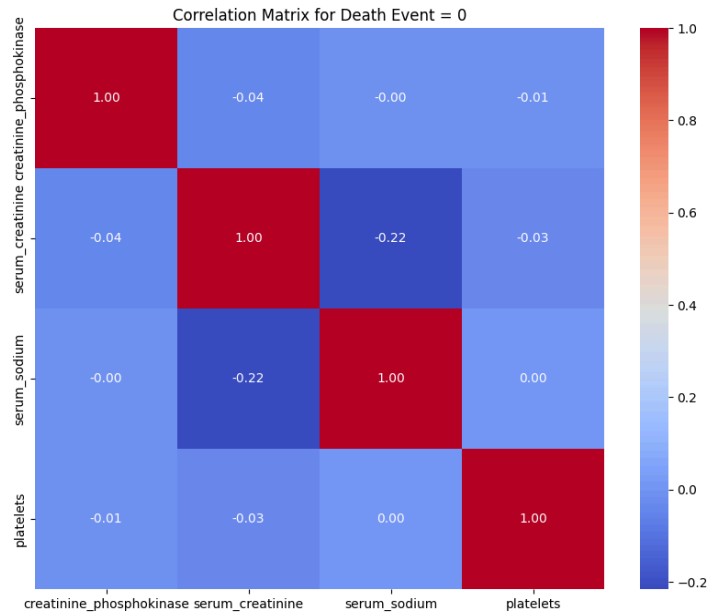


**CS677 HW 3**  
**By: Ahnaf Tajwar**

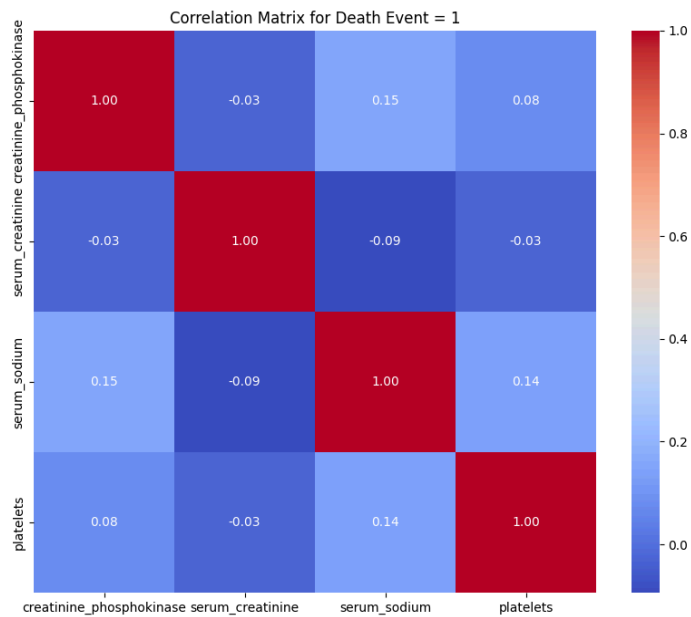
**Question 1**

1. See code
2. Correlation Matrices
  - a. Death Event = 0



i.

- b. Death Event = 1



i.

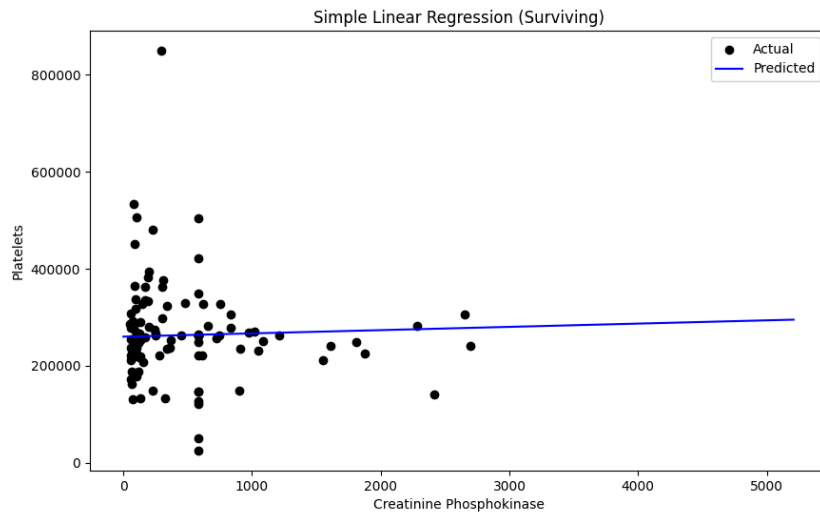
### 3. Correlations

- a. For surviving patients, serum\_sodium and serum\_creatinine have the highest correlation of -0.22.
- b. For surviving patients, serum\_sodium and creatinine\_posphokinase, as well as serum\_sodium and platelets have the lowest correlation of 0.0.
- c. For deceased patients, serum\_sodium and creatinine\_posphokinase have the highest correlation of 0.15.
- d. For deceased patients, serum\_creatinine and creatinine\_posphokinase, as well as serum\_creatinine and platelets have the lowest correlation of -0.03.
- e. The results are not the same for both cases. In fact, the lowest feature correlations of surviving patients are the highest correlations for the same features for deceased patients.

### Question 2

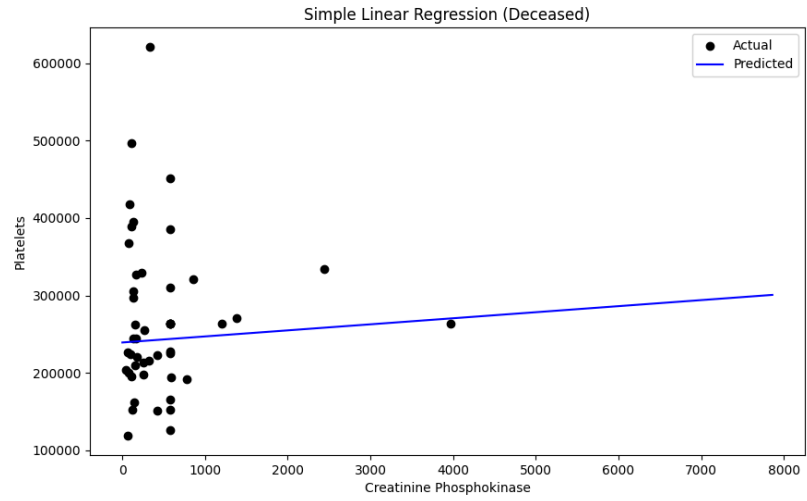
#### 1. See code

- a. Linear Regression
  - i. Surviving



1.

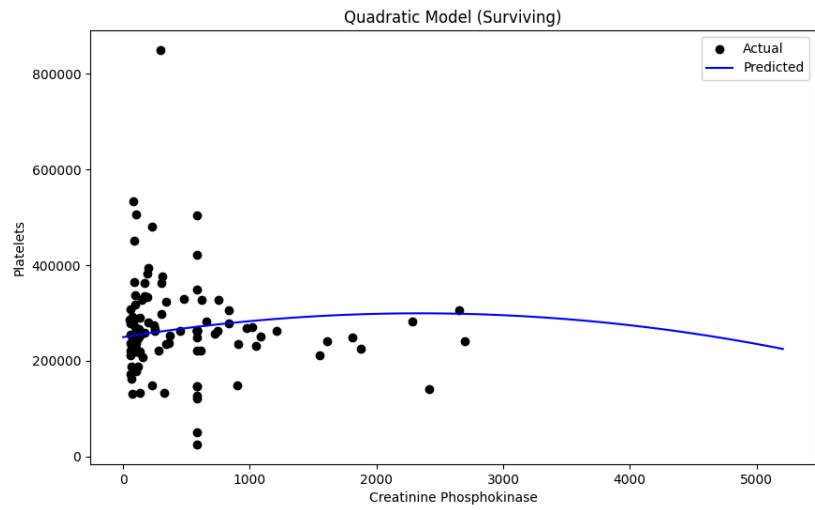
- ii. Deceased



1.

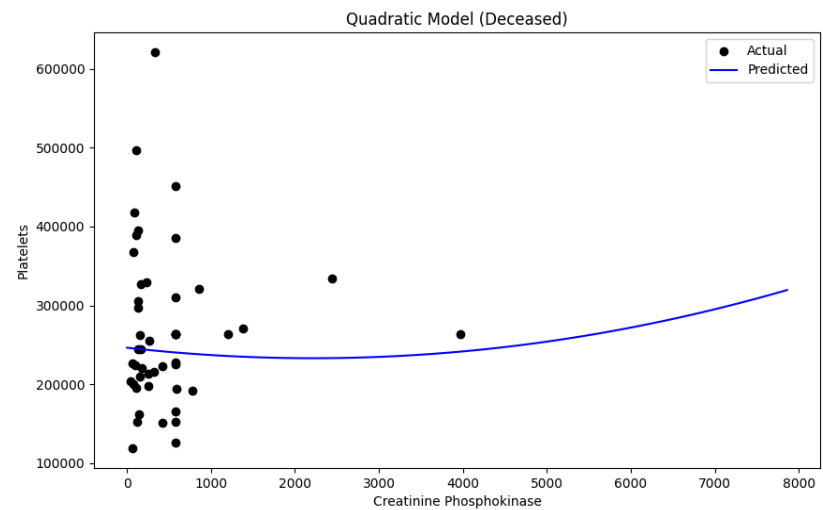
b. Quadratic

i. Surviving



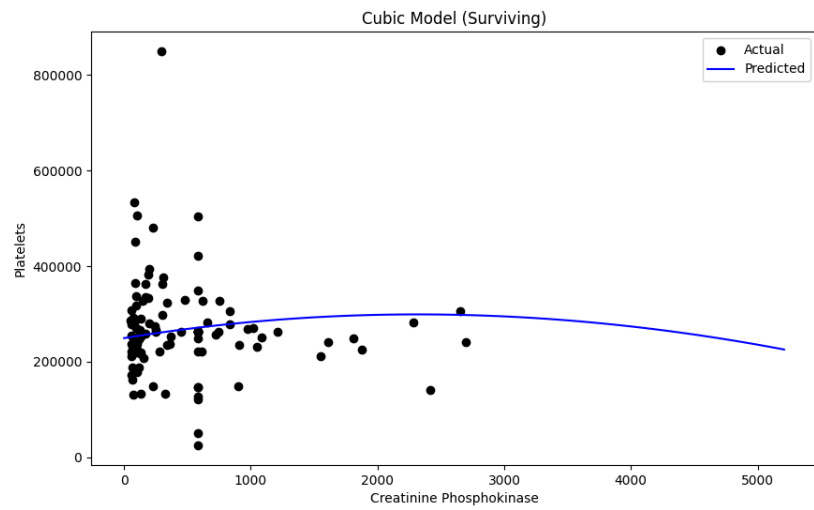
1.

ii. Deceased

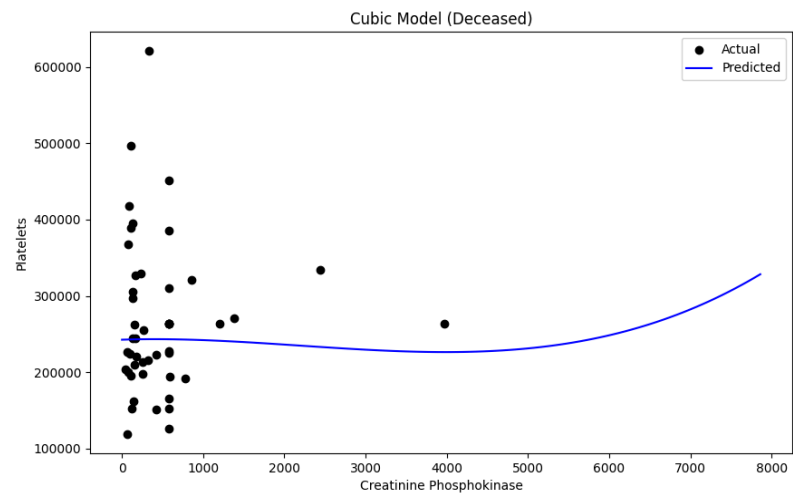


1.

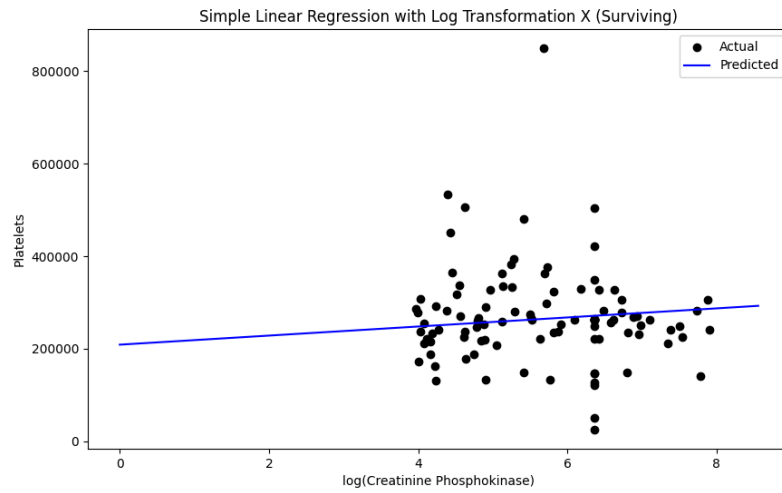
- c. Cubic
  - i. Surviving



- 1.
- ii. Deceased

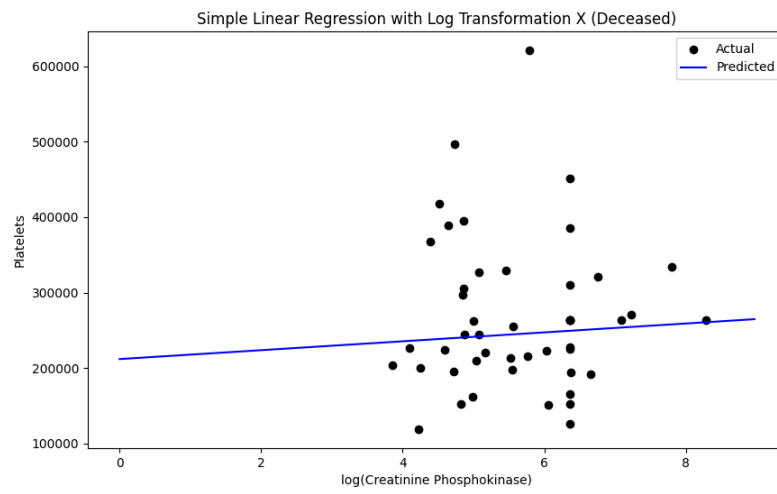


- 1.
- d.  $y = a \log x + b$ 
  - i. Surviving



1.

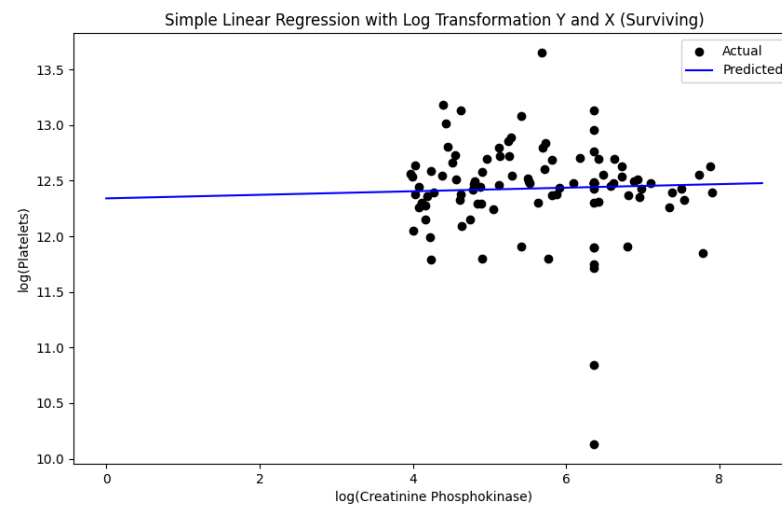
ii. Deceased



1.

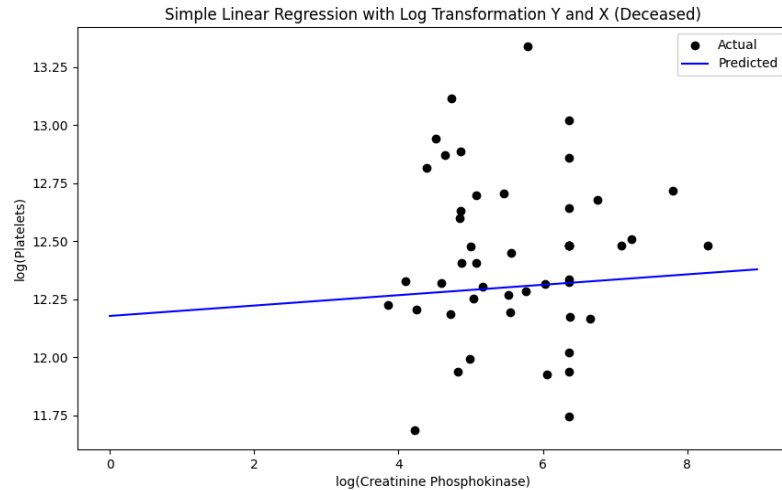
e.  $\log y = a \log x + b$

i. Surviving



1.

ii. Deceased



1.

### Question 3

Model	SSE (death event=0)	SSE (death event=1)
$y = ax + b$	1114142677733	489272557690
$y = ax^2 + bx + c$	1156251304018	490804335614
$y = ax^3 + bx^2 + cx + d$	1156401636003	491966262897
$y = a \log x + b$	1133210760949	485925037113
$\log y = a \log x + b$	1144612941344	564539865998

1. Best Model

a. Surviving Patients

i.  $y = ax + b$ , lowest SSE

b. Deceased Patients

i.  $y = a \log x + b$ , lowest SSE

2. Worst Model

a. Surviving Patients

i.  $y = ax^3 + bx^2 + cx + d$ , highest SSE

b. Deceased Patients

i.  $\log y = a \log x + b$ , highest SSE