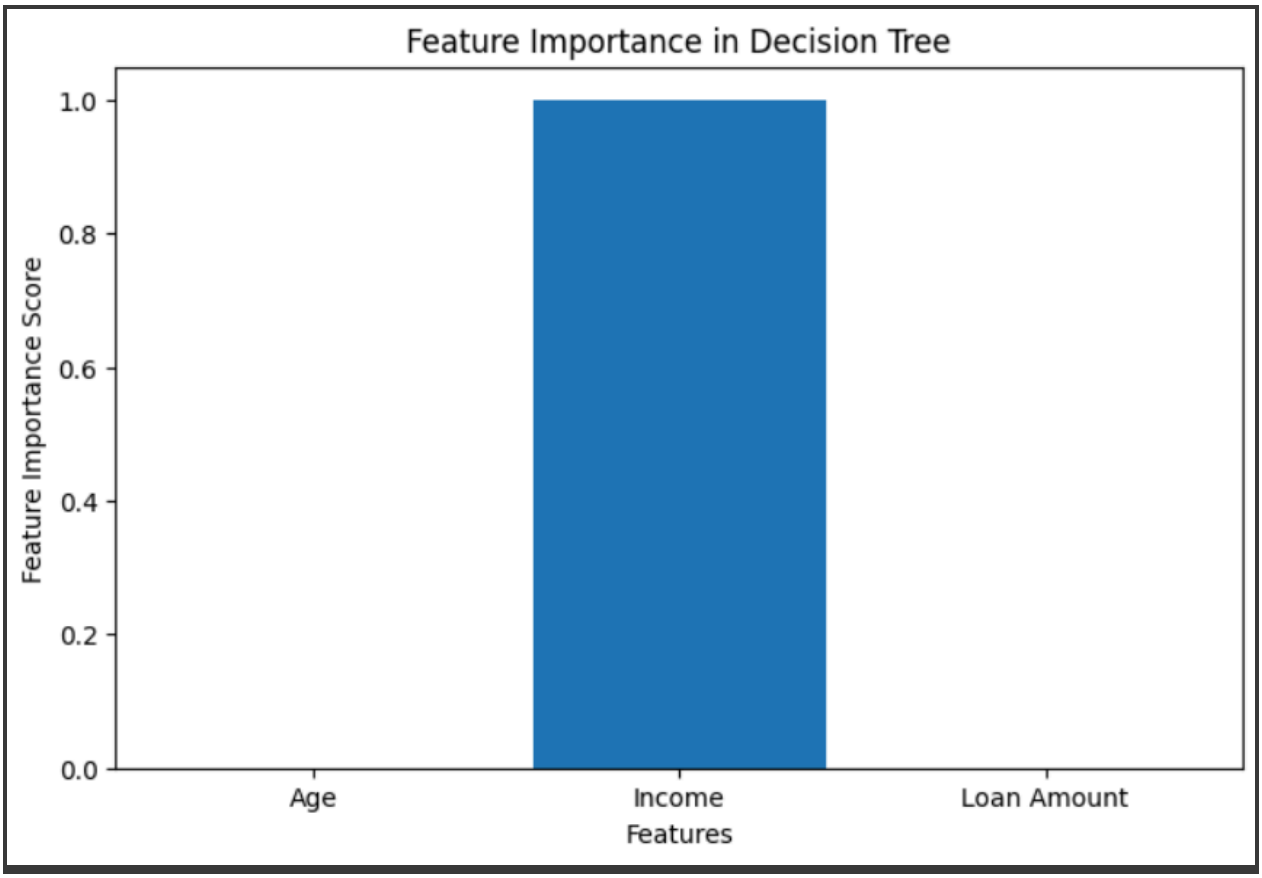
ASSIGNMENT 1

Q1.

```
Method Equal binning:

Rule 0: Age == (44.0, 49.5]
            Age  Income            Loan Amount Credit Risk (Target)
1  (44.0, 49.5]  Medium  (9166.667, 13333.333]             Low Risk

Rule 1: Age == (49.5, 55.0]
            Age  Income         Loan Amount Credit Risk (Target)
3  (49.5, 55.0]  Medium  (4975.0, 9166.667]             Low Risk
6  (49.5, 55.0]  Medium  (4975.0, 9166.667]             Low Risk
Mthod Equal Frequency

Rule 0: Age == (41.667, 49.167]
Selected data points:
                Age  Income            Loan Amount Credit Risk (Target)
1  (41.667, 49.167]  Medium  (8333.333, 11666.667]             Low Risk

Rule 1: Age == (49.167, 55.0]
Selected data points:
              Age  Income           Loan Amount Credit Risk (Target)
3  (49.167, 55.0]  Medium  (4999.999, 8333.333]             Low Risk
6  (49.167, 55.0]  Medium  (4999.999, 8333.333]             Low Risk
```

Q2.

```
Income: {'High': 0, 'Low': 1, 'Medium': 2}
Credit Risk (Target): {'High Risk': 0, 'Low Risk': 1}
```

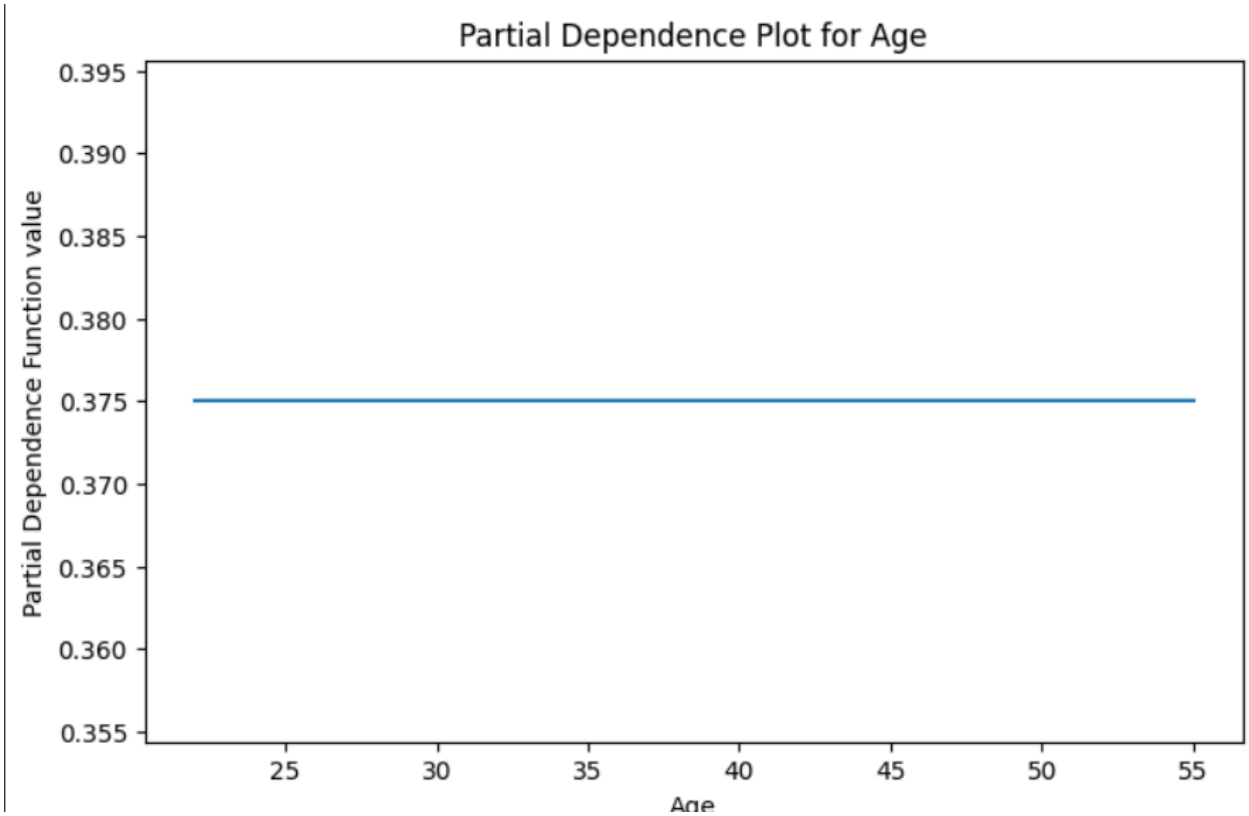


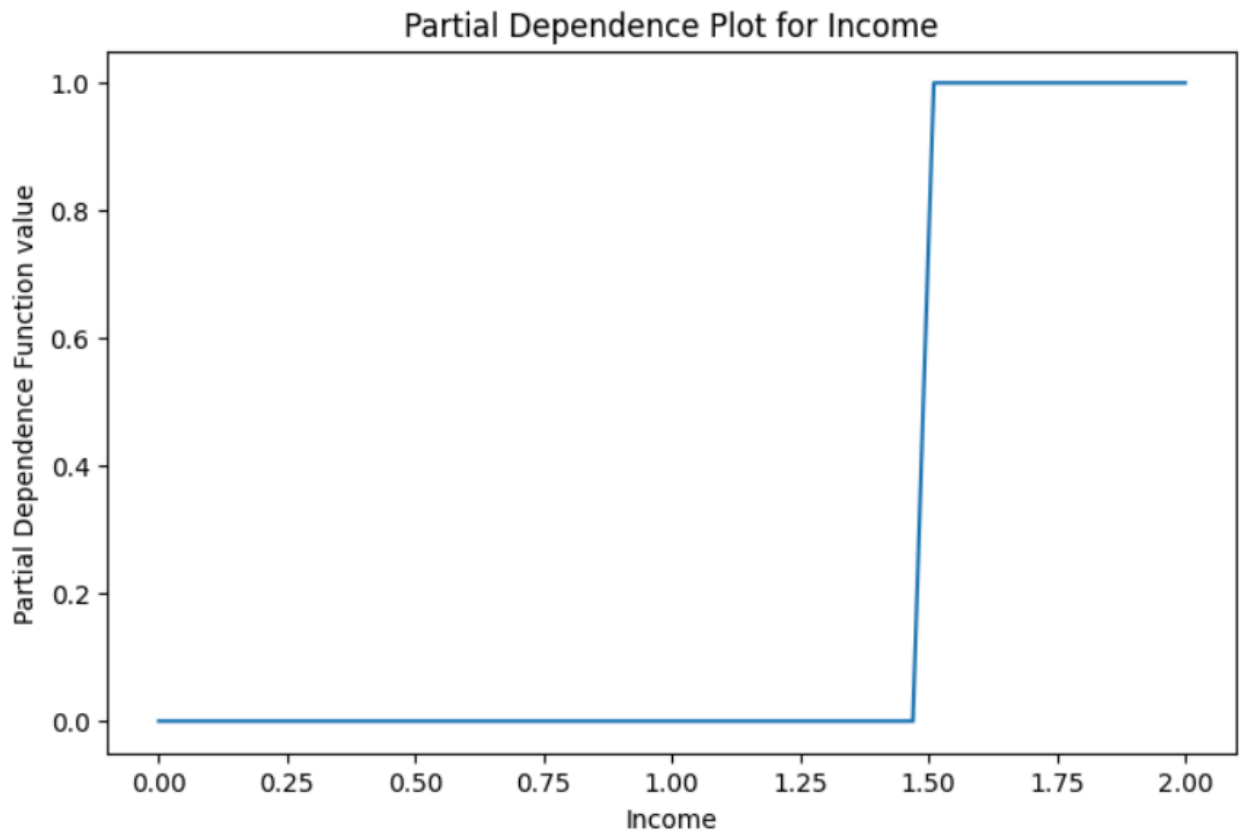

```
Rules are: |--- Income <= 1.50
|   |--- class: 0
|--- Income >  1.50
|   |--- class: 1
```

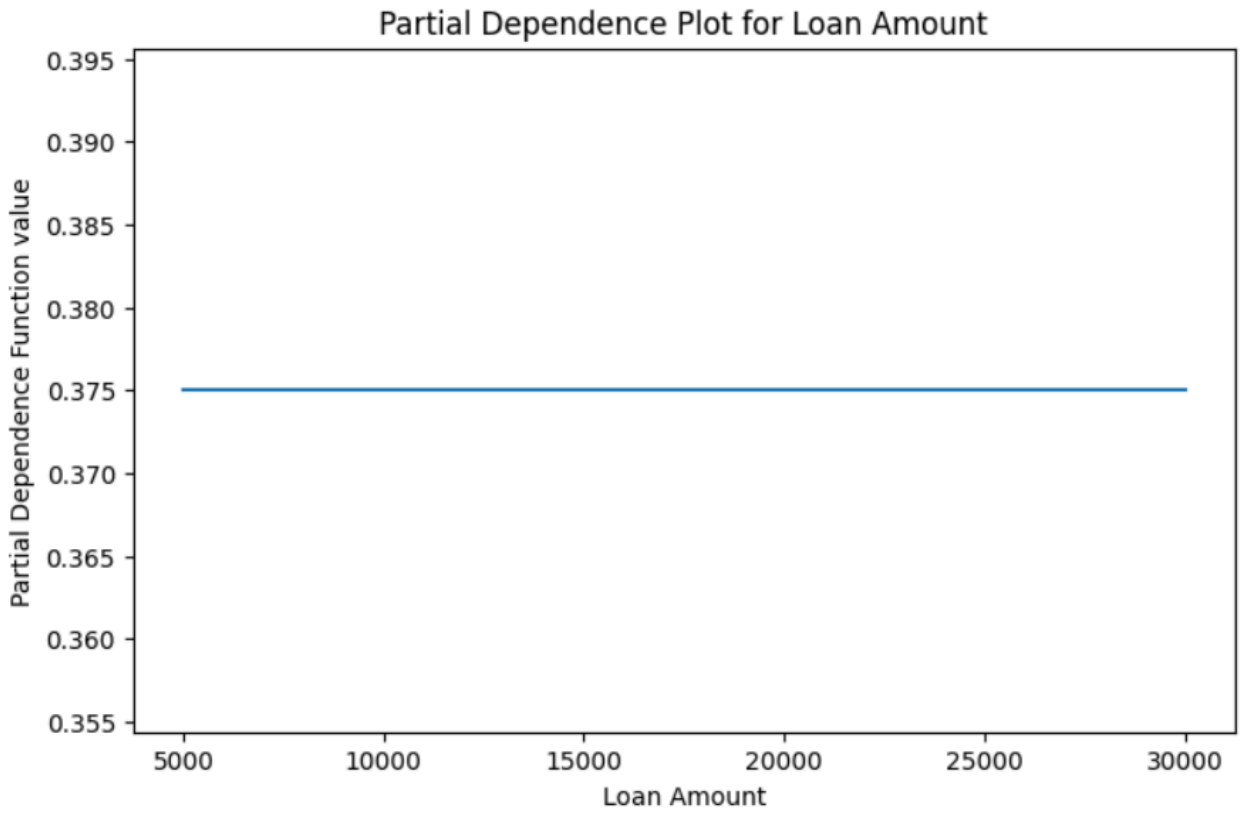
This graph shows that income directly impacts the predictions. Hence it has the most feature importance. According to rule if income is less that 1.50 then it is class 0 else it it class 1.

Q3.

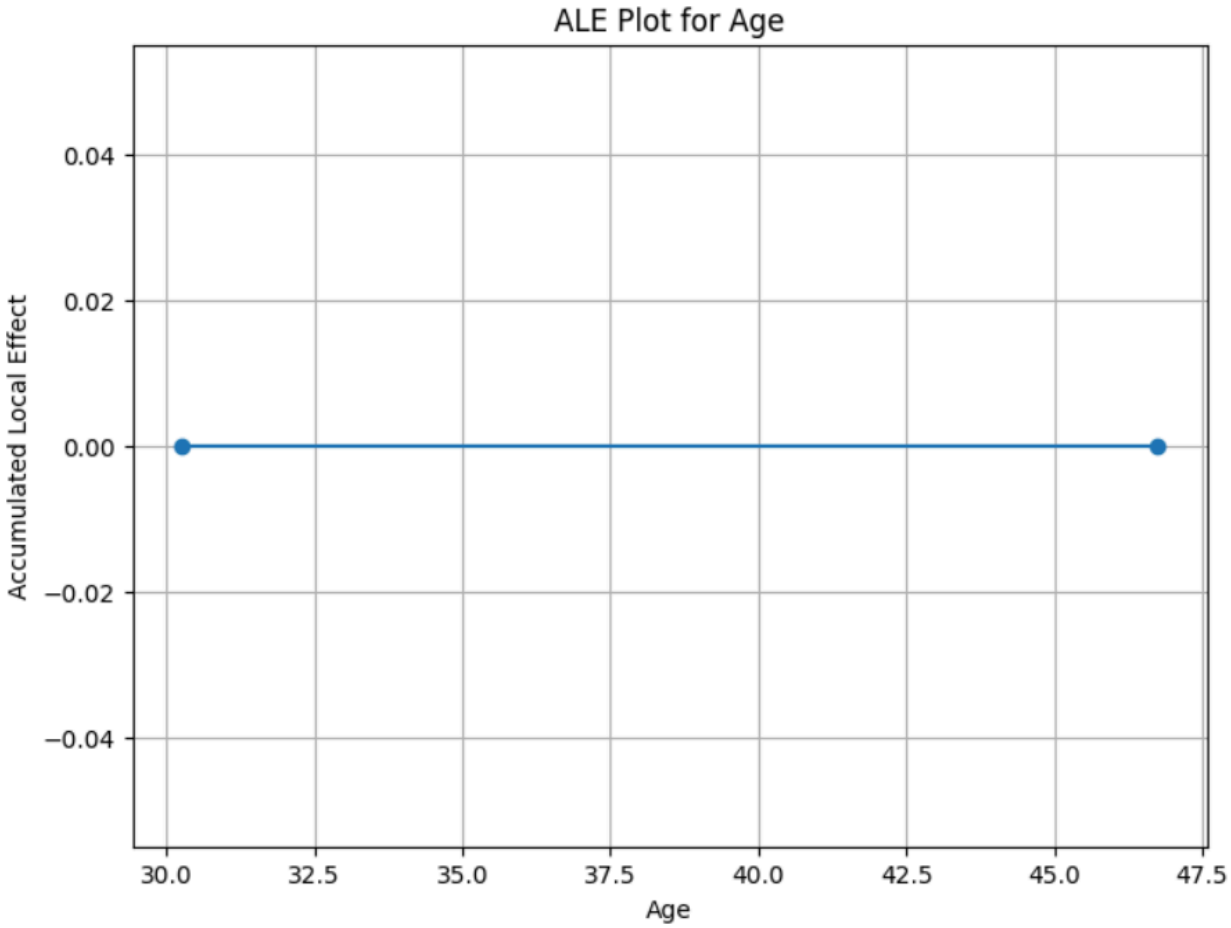
According to the Partial Dependence plots, we see that the age and load amount are completely flat. This means that they are not contributing to the prediction.
The income increases when it is 1, this means that higher income means higher prediction value.

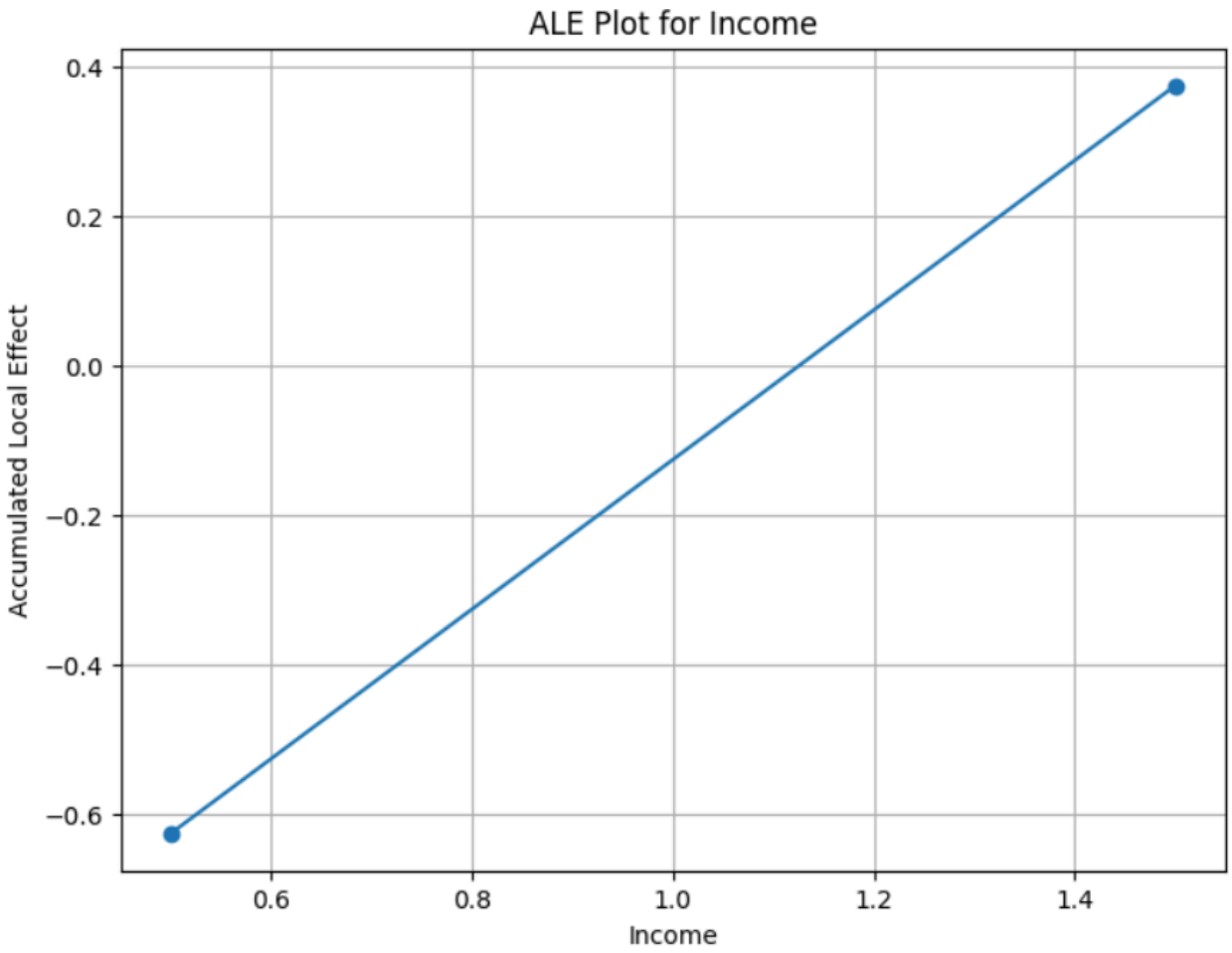

Partial Dependence Plot for Age
Partial Dependence Function value
0.395
0.390
0.385
0.380
0.375
0.370
0.365
0.360
0.355
25
30
35
40
45
50
55
Age

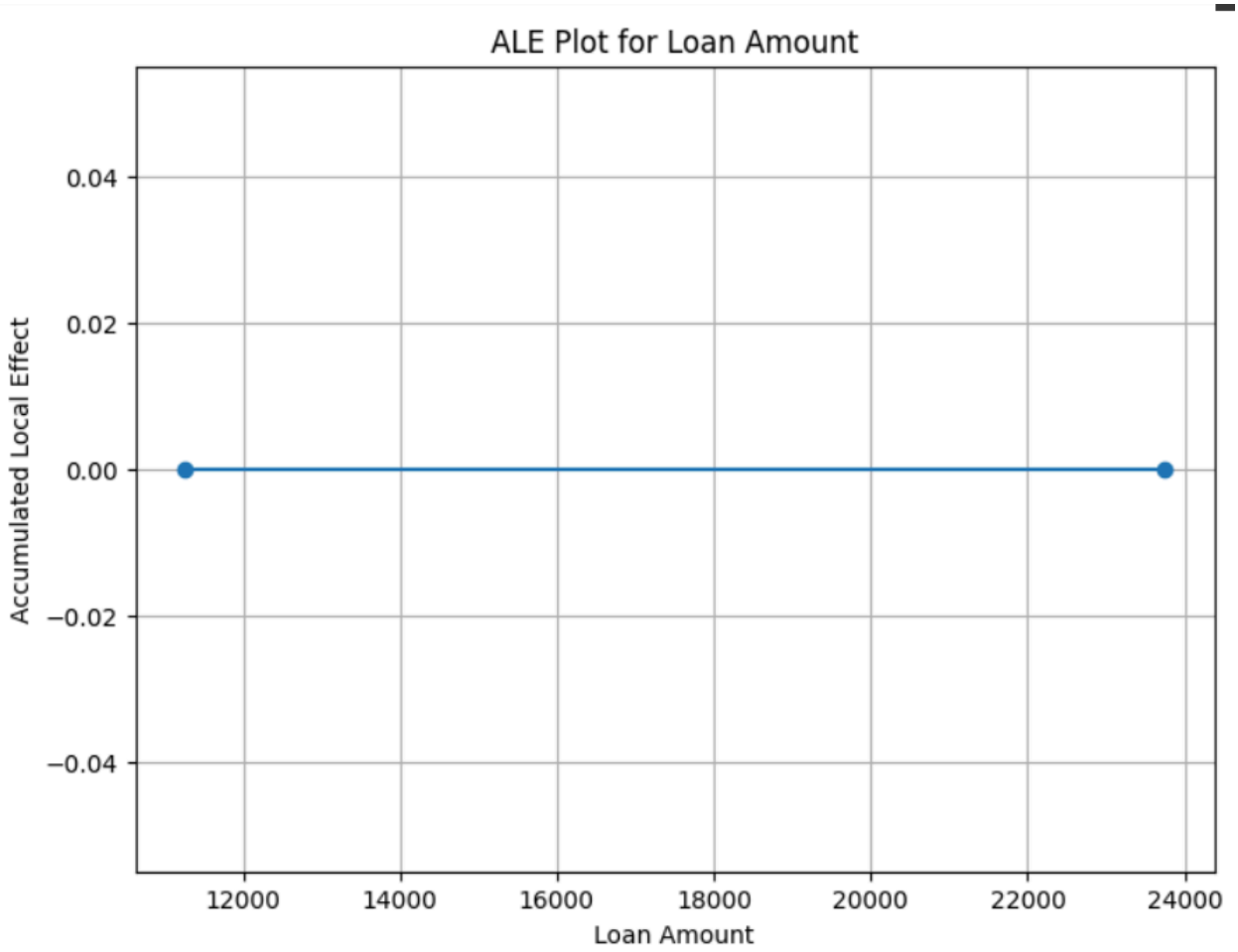

Partial Dependence Plot for Income
Partial Dependence Function value
1.0
0.8
0.6
0.4
0.2
0.0
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
Income

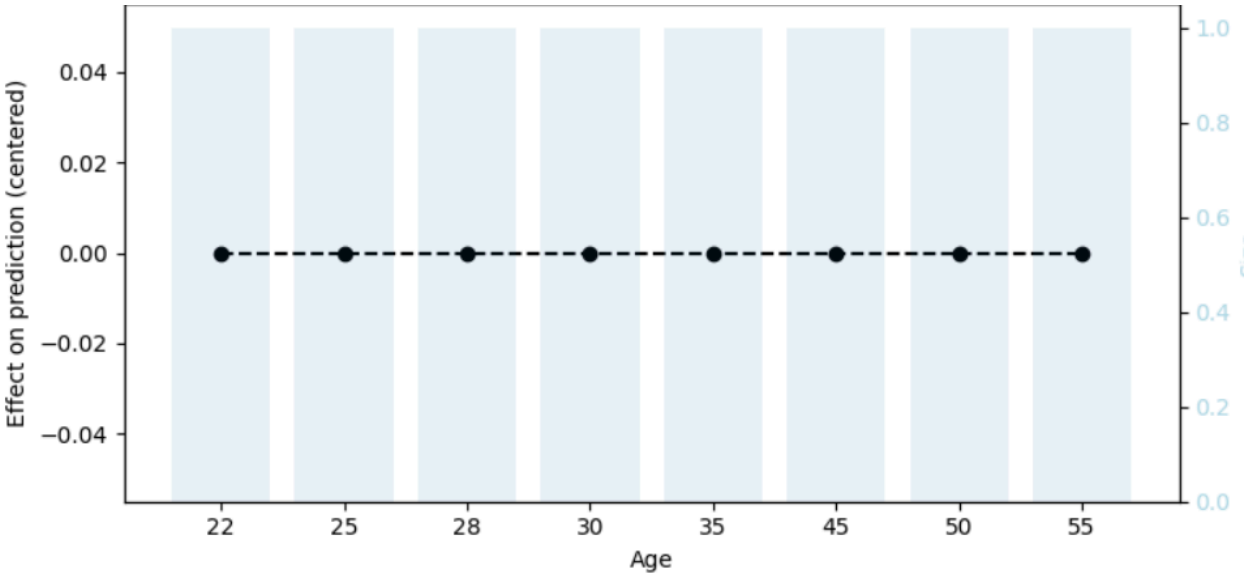


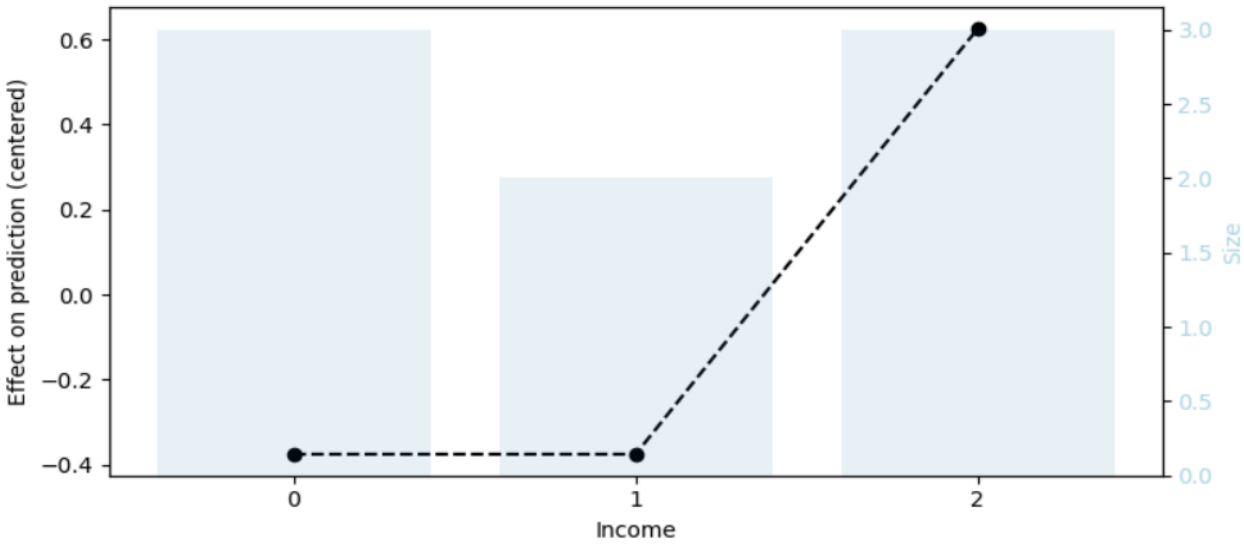
The ALE plots for the feature show that Age and Loan amount do not affect prediction as they are flat. The income increases when it is 1, this means that higher income means higher prediction value which is similar to PDP trend.

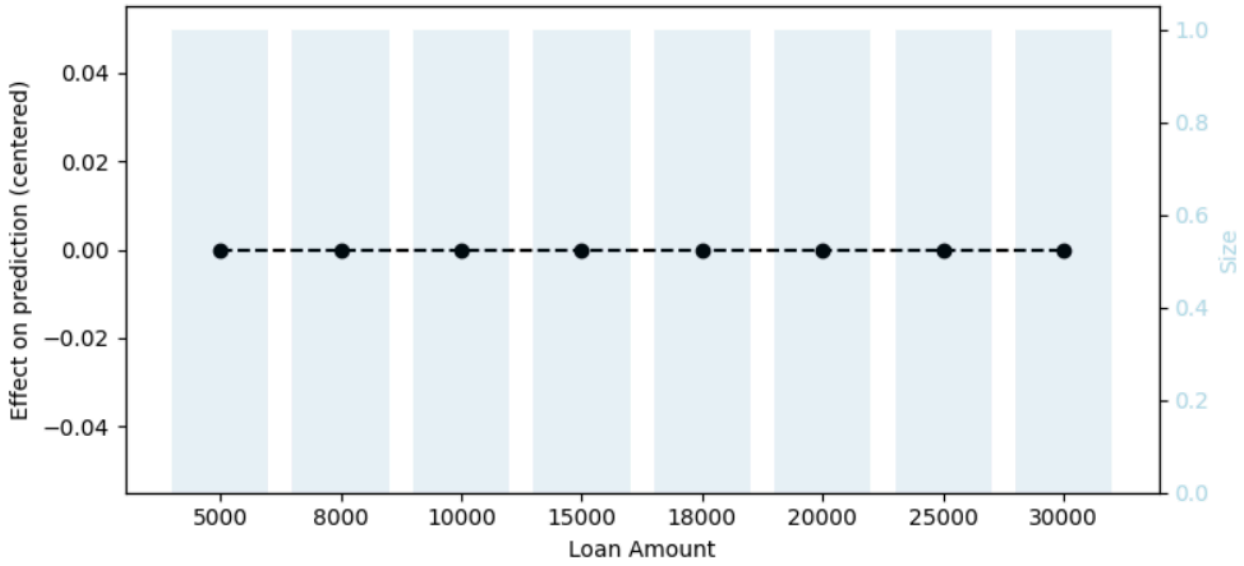

ALE Plot for Age
Accumulated Local Effect
0.04
0.02
0.00
−0.02
−0.04
30.0
32.5
35.0
37.5
40.0
42.5
45.0
47.5
Age


ALE Plot for Income
Accumulated Local Effect
0.4
0.2
0.0
−0.2
−0.4
−0.6
0.6
0.8
1.0
1.2
1.4
Income


ALE Plot for Loan Amount
Accumulated Local Effect
0.04
0.02
0.00
−0.02
−0.04
12000
14000
16000
18000
20000
22000
24000
Loan Amount


Effect on prediction (centered)
0.04
0.02
0.00
−0.02
−0.04
1.0
0.8
0.6
0.4
0.2
0.0
22
25
28
30
35
45
50
55
Age



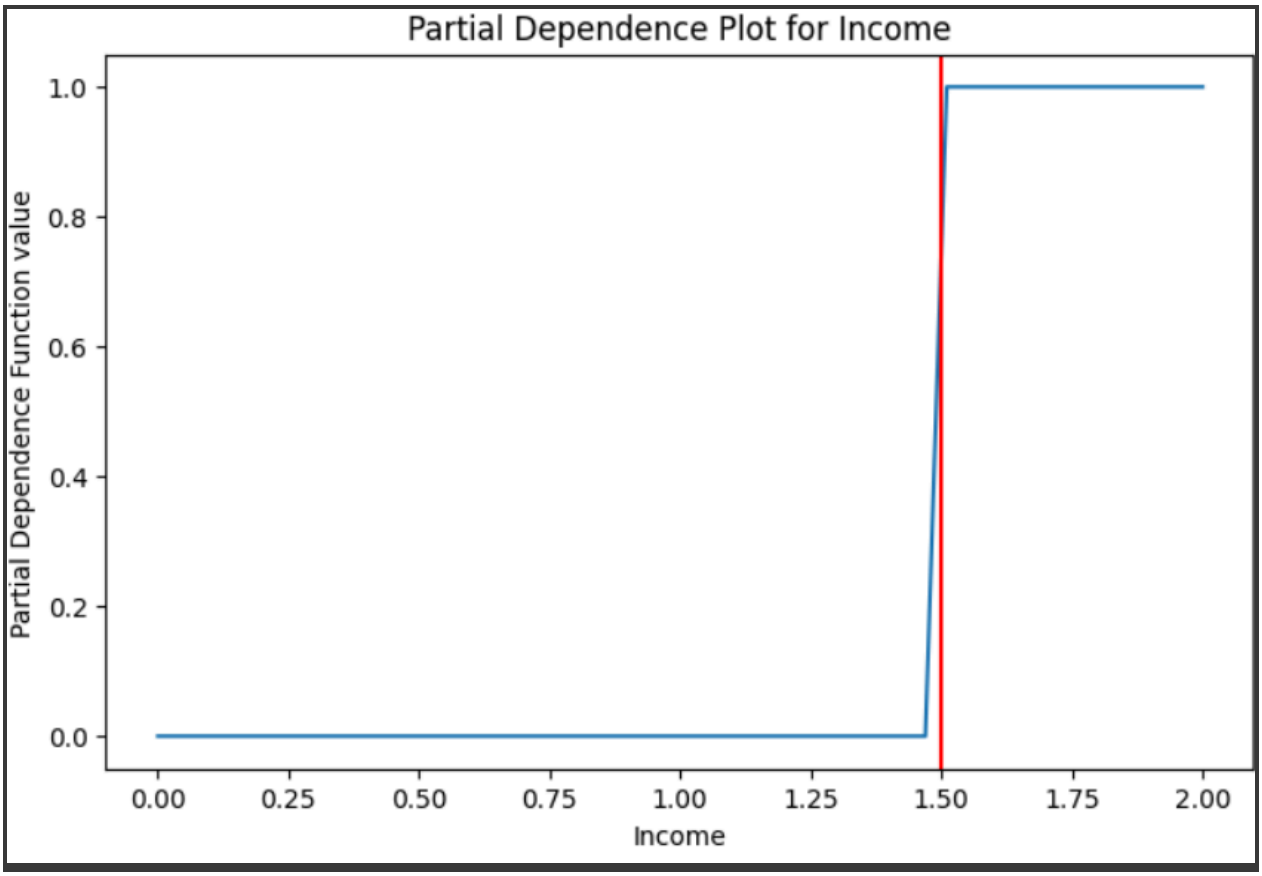


Q4.

The breaking point for income is 1.5 hench values before that are class o and above that are class 1.

```
Rules are: |--- Income <= 1.50
|   |--- class: 0
|--- Income >  1.50
|   |--- class: 1
```

```
Thresholds (breaking points) for feature 'Income': [1.5]
```



Used chatgpt for references for syntax, functions and their usage.