1 A)

Risk scores using model
3.32
9.6
4
10.84
7.12

1 B)

RMSE : 1.92

1 C)

The most important attribute is the Height

1 D)

Blood pressure Mean : 129
Blood pressure Std : 20.432

Height Mean: 5.76
Height Std: .4615

1 E)

Risk = 0.4752*Z_bp + 1.91126*Z_h + 4.5885

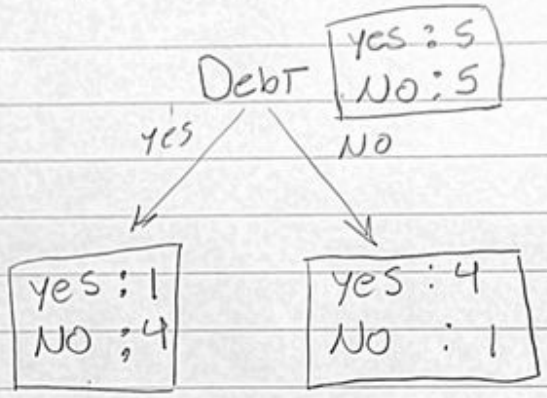
1 F)

The most important attribute is Height

1 G)

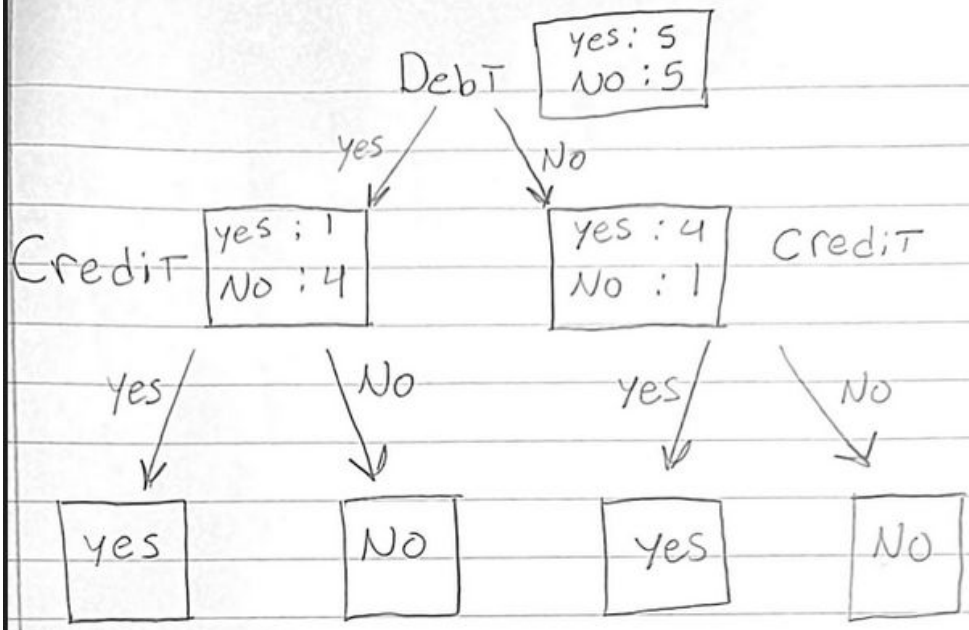
My answer for F is consistent with my answer for C

		class Y	class N	E_{child}	E_{TOTAL}	
Debt	Y	1	4	.7219	.7219	Best Split
	N	4	1	.7219		
UNEmp	Y	0	2	0	.7635	
	N	5	3	.9544		
Credit	Y	2	1	.9183	.9317	
	N	3	4	.9852		



Debt = yes		class Y	class N	E_{child}	E_{TOTAL}	
UNEmp.	Y	0	1	0	.6490	
	N	1	3	.81127		
Credit	Y	1	0	0	0	Best Split
	N	0	4	0		

Debt = NO		class Y	class N	E_{child}	E_{TOTAL}	
UNemp	Y	0	1	0	.6490	
	N	3	1	.81127		
Credit	Y	0	2	0	0	Best Split
	N	3	0	0		



Accuracy : Depth 2 (1 + 1) / 10 = 2/10

Error = 1 - 2/10 = 4/5 80 %

3 A)

gini = 1 - (1/20)^2 - (19/20)^2
gini = .095

3 B)

Same Calculation for both left and right handedness
gini = 1 - (1/10)^2 - (9/10)^2
gini = .18

3 C)

Unique IDs will be the best split chose

3 D)

I think the answer in C is reasonable. If you split the training set by a unique ID you are looking at how well a baseball player will bat against each individual pitcher. This makes sense because you can have a really good left handed pitcher and a really bad right handed pitcher, and what hand they throw with would not be as important as the pitcher's overall skill in the game.