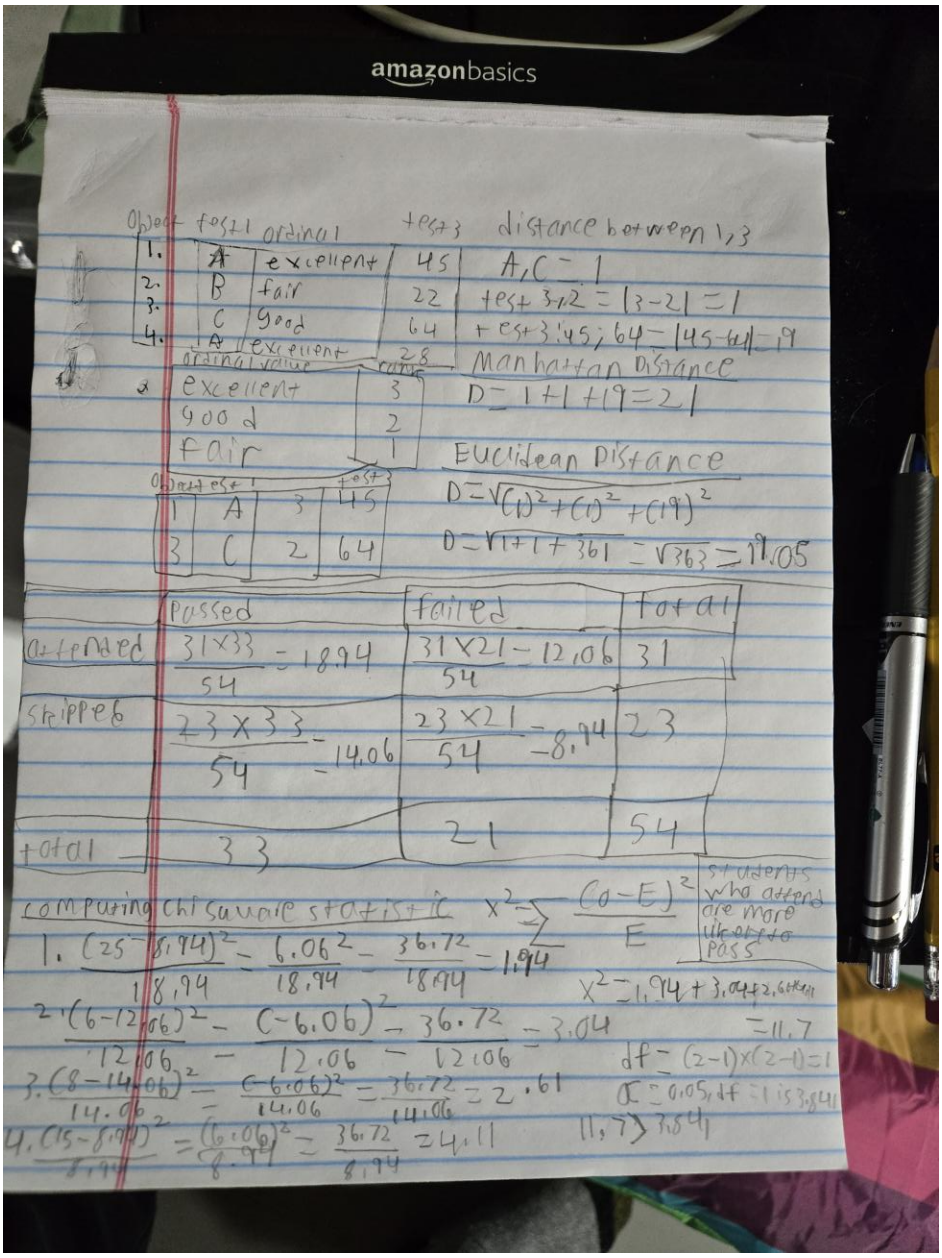

amazonbasics
Object test1 ordinal test3 distance between 1,3
1. A excellent 45 A,C = 1
2. B fair 22 test 3:2 = |3-2| = 1
3. C good 64 test 3:45; 64 - |45-64| = 19
4. A excellent 28
Manhattan distance
D = 1+1+19 = 21
ordinal value rank
excellent 3
good 2
fair 1
Euclidean distance
D = √(1)² + (1)² + (19)²
D = √1+1+361 = √363 = 19.05
passed failed Total
attended 31X33/54 = 18.94 31X21/54 = 12.06 31
skipped 23X33/54 = 14.06 23X21/54 = 8.94 23
total 33 21 54
computing chi square statistic
students who attend are more likely to pass
df = (2-1)x(2-1) = 1
α = 0.05, df = 1 is 3.841
11.7 > 3.841

3. Senior count: 30+5+3+10+4 = 52

Junior count: 40+40+20+3+4+6 = 113

Total count: 30+40+40+20+5+3+3+10+4+6 = 165

$P_{junior} = \frac{113}{165}$ $H(S) = -\left(\frac{113}{165}\log_2\frac{113}{165} + \frac{52}{165}\log_2\frac{52}{165}\right)$

$P_{senior} = \frac{52}{165}$ $= 0.8990308$

Sales:	Juniors	Seniors	Total
	80	30	110
Systems:	23	8	31
marketing:	4	10	14
secretary:	6	4	10

$H(sales) = -\left(\frac{80}{110}\log_2\frac{80}{110} + \frac{30}{110}\log_2\frac{30}{110}\right)$
$= 0.845359$

$H(systems) = -\left(\frac{23}{31}\log_2\frac{23}{31} + \frac{8}{31}\log_2\frac{8}{31}\right)$
$= 0.823816$

$H(secretary) = -\left(\frac{6}{10}\log_2\frac{6}{10} + \frac{4}{10}\log_2\frac{4}{10}\right) = 0.9709506$

$H(marketing) = -\left(\frac{4}{14}\log_2\frac{4}{14} + \frac{10}{14}\log_2\frac{10}{14}\right) = 0.863129$

$H(department) = \sum \frac{|Sv|}{|S|} H(Sv) = 0.850424$

$IG(department) = H(S) - H(department) =$
$0.899 - 0.850424 = 0.0486079$

	Junior	Senior	total
21-25	20	0	20
26-30	49	0	49
31-35	40	35	75
36-40	0	10	10
41-45	0	3	3
46-50	4	4	8

$H(21\text{-}25) = -\left(\frac{20}{20}\log_2\frac{20}{20} + \frac{0}{20}\log_2\frac{0}{20}\right)$
$= 0$

$H(26\text{-}30) = -\left(\frac{49}{49}\log_2\frac{49}{49} + \frac{0}{49}\log_2\frac{0}{49}\right)$
$= 0$

$H(31\text{-}35) = -\left(\frac{40}{75}\log_2\frac{40}{75} + \frac{35}{75}\log_2\frac{35}{75}\right)$
$= 0.9906175$

$H(36\text{-}40) = -\left(\frac{0}{10}\log_2\frac{0}{10} + \frac{10}{10}\log_2\frac{10}{10}\right)$
$= 0$

$H(41\text{-}45) = -\left(\frac{0}{3}\log_2\frac{0}{3} + \frac{3}{3}\log_2\frac{3}{3}\right)$
$= 0$

$H(46\text{-}50) = -\left(\frac{4}{8}\log_2\frac{4}{8} + \frac{4}{8}\log_2\frac{4}{8}\right)$
$= 0$

$H(Age) = 0.4742957$

$IG(Age = H(S) - H(Age) = 0.8[illegible]0.474295$
$= \boxed{0.4247351}$

salary	junior	senior	total
26k-30k	46	0	46
31k-35k	40	0	40
36k-40k	0	4	4
41k-45k	4	0	4
46k-50k	23	40	63
66k-70k	0	8	8

$H(26-30) = -\left(\frac{46}{46}\log_2\frac{46}{46} + \frac{0}{46}\log_2\frac{0}{46}\right)$
$= 0$

$H(31-35) = -\left(\frac{40}{40}\log_2\frac{40}{40} + \frac{0}{40}\log_2\frac{0}{40}\right)$
$= 0$

$H(36-40) = -\left(\frac{0}{4}\log_2\frac{0}{4} + \frac{4}{4}\log_2\frac{4}{4}\right)$
$= 0$

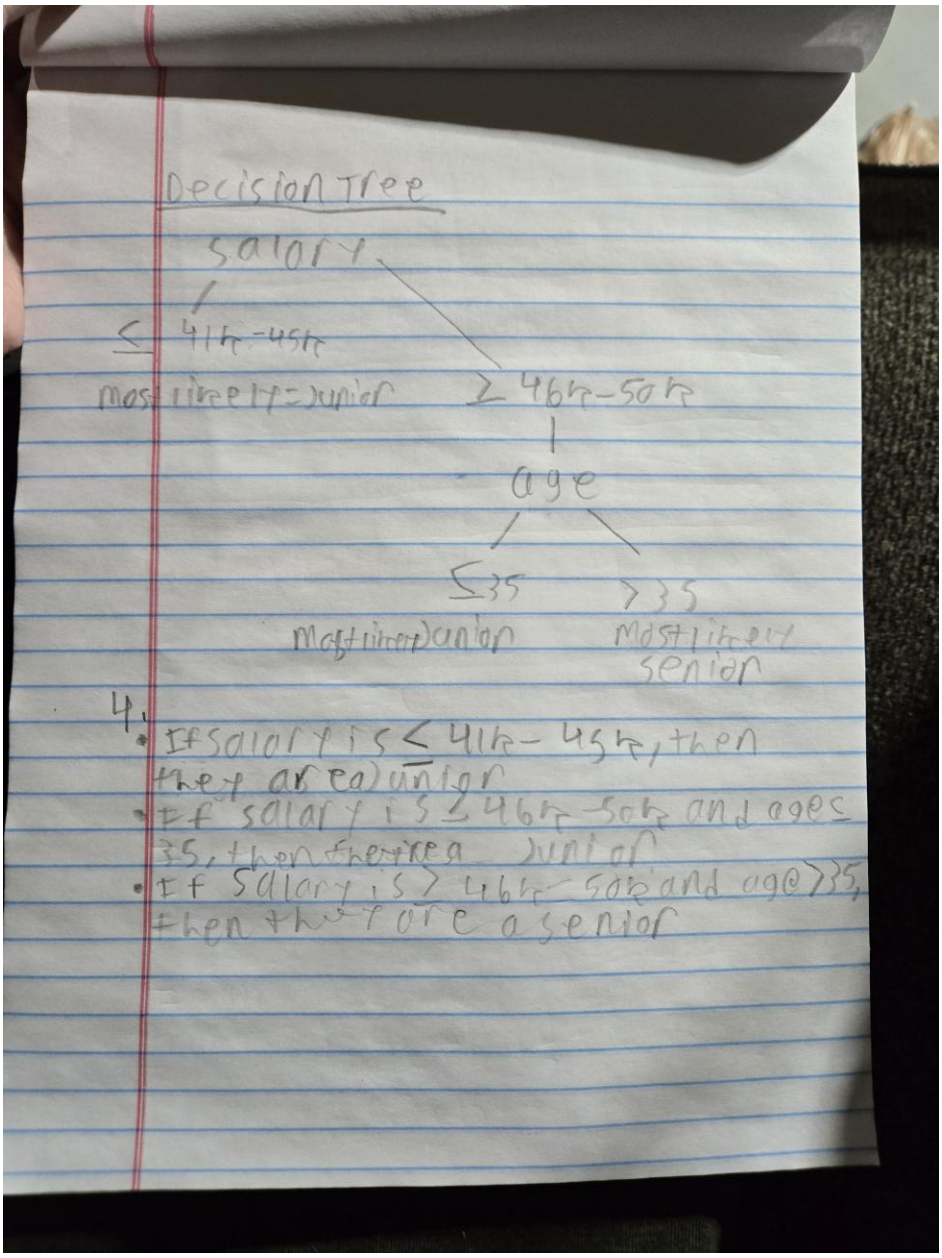
$H(41-45) = -\left(\frac{4}{4}\log_2\frac{4}{4} + \frac{0}{4}\log_2\frac{0}{4}\right)$
$= 0$

$H(46-50) = -\left(\frac{23}{63}\log_2\frac{23}{63} + \frac{40}{63}\log_2\frac{40}{63}\right)$
$= 0.946818$

$H(66-70) = -\left(\frac{0}{8}\log_2\frac{0}{8} + \frac{8}{8}\log_2\frac{8}{8}\right)$
$= 0$

$H(\text{salary}) = 0.361526$

$IG(\text{salary}) = H(S) - H(\text{salary}) = 0.$ [illegible] $= 0.537518$



4.

- If salary is ≤ 41k-45k, then they are a junior
- If salary is ≥ 46k-50k and age ≤ 35, then they're a junior
- If salary is ≥ 46k-50k and age > 35, then they're a senior