AGYEMANG ERIC
MAT 450 HW3

3) a) $\bar{y}_u = \frac{t}{N}$ where $t = 66 + 59 + 70 + 83 + 82 + 71 = 431$

$N = 6$

$\therefore \bar{y}_u = \frac{431}{6} = \boxed{71.833}$

$$s^2 = \frac{\sum_{i=1}^{N}(y_i - \bar{y}_u)^2}{N-1} = \left[(66-71.83)^2 + (59-71.83)^2 + (70-71.83)^2 + (83-71.8\right.$$
$$+ [80-71.83]^2 + [71-71.83]^2\Big) \Big/ (6-1)$$
$$= \boxed{86.17}$$

b) For the number of possible SRS of size 4 $= \binom{N}{n} = \binom{6}{4} = \boxed{15}$

c)

Sample	units in sample	y-values in sample	sample means
S_1	{1, 2, 3, 4}	66, 59, 70, 83	$\frac{66+59+70+83}{4} = 69.50$
S_2	{1, 2, 3, 5}	66, 59, 70, 82	$\frac{66+59+70+82}{4} = 69.25$
S_3	{1, 2, 3, 6}	66, 59, 70, 71	$\frac{66+59+70+71}{4} = 66.50$
S_4	{1, 2, 4, 5}	66, 59, 83, 82	$\frac{66+59+83+82}{4} = 72.50$
S_5	{1, 2, 4, 6}	66, 59, 83, 71	$\frac{66+59+83+71}{4} = 69.75$
S_6	{1, 2, 5, 6}	66, 59, 82, 71	$\frac{66+59+82+71}{4} = 69.50$
S_7	{1, 3, 4, 5}	66, 70, 83, 82	$\frac{66+70+83+82}{4} = 75.25$
S_8	{1, 3, 4, 6}	66, 70, 83, 71	$\frac{66+70+83+71}{4} = 72.5$
S_9	{1, 3, 5, 6}	66, 70, 82, 71	$\frac{66+70+82+71}{4} = 72.25$
S_{10}	{1, 4, 5, 6}	66, 83, 82, 71	$\frac{66+83+82+71}{4} = 75.5$
S_{11}	{2, 3, 4, 5}	59, 70, 83, 82	$\frac{59+70+83+82}{4} = 73.50$
S_{12}	{2, 3, 4, 6}	59, 70, 83, 71	$\frac{59+70+83+71}{4} = 70.75$
S_{13}	{2, 3, 5, 6}	59, 70, 82, 71	$\frac{59+70+82+71}{4} = 70.50$
S_{14}	{2, 4, 5, 6}	59, 83, 82, 71	$(59+83+82+71)/4 = 73.75$
S_{15}	{3, 4, 5, 6}	70, 83, 82, 71	$(70+83+82+71)/4 = 76.50$

Now using $v(\bar{y}) = \left(1-\frac{n}{N}\right)\frac{s^2}{n}$; $s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y}_u)^2$

defined in ⓐ above, and $n = 4$ $N = 6$

$\Rightarrow v(\bar{y}) = \left(1-\frac{4}{6}\right)\left(\frac{86.17}{4}\right) = \boxed{7.18083}$

ⓓ The number of stratified random samples of size 4 that are possible with 2 students selected from each stratum is

$$\binom{3}{2}\binom{3}{2} = \frac{3!}{2!(3-2)!} \times \frac{3!}{2!(3-2)!} = \boxed{9}$$

ⓔ

unit in S_1		Unit in S_2		y-values in sample S_1		S_2	
1	2	4	5	65	59	83	82
1	2	4	6	66	59	83	71
1	2	5	6	66	59	82	71
1	3	4	5	66	70	83	82
1	3	4	6	66	70	83	71
1	3	5	6	66	70	82	71
2	3	4	5	59	70	83	82
2	3	4	6	59	70	83	71
2	3	5	6	59	70	82	71

There was no samples from part (c) which contains 3 units from one of the strata.
This eliminates the first 3 samples which contains {1, 2, 3} and the three sample containing students {4, 5, 6}.

③ ⑧ from ⓔ above we can find $\bar{y}_{str}$ for each stratified random sample as follows.

units in S_1		units in S_2		y-values in sample				$\bar{y}_{str}$
1	2	4	5	66	59	83	82	72.50
1	2	4	6	66	59	83	71	69.75
1	2	5	6	66	59	82	71	69.50
1	3	4	5	66	70	83	82	75.25
1	3	4	6	66	70	83	71	72.50
1	3	5	6	66	70	82	71	72.25
2	3	4	5	59	70	83	82	73.50
2	3	4	6	59	70	83	71	70.75
2	3	5	6	59	70	82	71	70.50

we calculate $S_1^2 = \frac{1}{N_1-1}\sum_{j=1}^{2}(y_{1j}-\bar{y}_{1U})^2 = 31$

$$S_2^2 = \frac{1}{N_2-1}\sum_{j=1}^{2}(y_{2j}-\bar{y}_{2U})^2 = 44.33333$$

Then $V(\bar{y}_{str}) = \sum_{h=1}^{H}\left(1-\frac{n_h}{N_h}\right)\left(\frac{N_h}{N}\right)^2\frac{S_h^2}{n_h} = \left(1-\frac{2}{3}\right)\left(\frac{3}{6}\right)^2\left(\frac{31}{2}\right)+\left(1-\frac{2}{3}\right)\left(\frac{3}{6}\right)^2\left(\frac{4.333}{2}\right)$

$= \boxed{3.14}$

The variance is smaller due to the fact that the extreme samples from (C) above are excluded by the stratified design. The variances, $S_1^2 = 31$, $S_2^2 = 44.3333$ are much smaller than the population variance S^2.

⑦ (a) $\hat{t}_{str} = \sum_{h=1}^{H} N_h \bar{y}_h$ where $\bar{y}_h = \dfrac{\sum_{j=1}^{n} f y_{hj}}{\sum f}$

For Biology: $\bar{y}_b = \dfrac{(1\times0)+(2\times1)+(0\times2)+(1\times3)+(0\times4)+(2\times5)+(1\times\ldots}{1+2+0+1+0+2+0+1+1+0}$

$= \underline{3.142}$

for Physics: $\bar{y}_p = \dfrac{(10\times0)+(2\times1)+(0\times2)+(1\times3)+(2\times4)+(5\times5)+(1\times6)+\ldots}{10+2+0+1+2+1+1+0+2}$

$= \underline{2.105}$

For Social $\bar{y}_s = \dfrac{(9\times0)+(0\times1)+(1\times2)+(0\times3)+(2\times4)+(0\times5)+(1\times6)}{9+0+1+0+2+0+1+0+0}$

$= \underline{1.23077}$

For Humanities: $\bar{y}_H = \dfrac{(8\times0)+(2\times1)+(3\times1)+(0\times4)}{8+2+1} = \underline{0.4545}$

Then $\hat{t}_{str} = (102\times3.142)+(310\times2.105)+(217\times1.23077)+(178\times0.4545) = \boxed{1{,}321.012}$

$SE(\hat{t}_{str}) = \sqrt{\hat{Var}(\hat{t}_{str})} = \left[\sum N_h^2\left(1-\frac{n_h}{N_h}\right)\left(\frac{N_h}{N}\right)^2 \frac{s_h^2}{n_h}\right]^{0.5}$ and

$s_h^2 = \dfrac{\sum_{j\in S_h} f(y_{hj}-\bar{y}_h)^2}{n_h-1}$. The table below gives the calculated results.

STRATUM	N_h	n_h	$\bar{y}_h = \sum_{j\in S} y_{hj}/n_h$	$\hat{t}_h = N_h\bar{y}_h$	s_h^2	$V(\hat{t}_h) = (1-\frac{n_h}{N_h})N_h^2\cdot\frac{s_h^2}{n_h}$
Biological	102	7	3.142	102×3.142 $\boxed{320.484}$	6.8095	9426.294
Physical	310	19	2.105	310×2.105 $\boxed{652.55}$	8.21052	38982.68
Social	217	13	1.23077	217×1.23077 $\boxed{267.077}$	4.3590	14843.40
Humanities	178	11	0.4545	178×0.4545 $\boxed{80.901}$	0.8727	2358.35
	807	50		1321.01		65610.727

So $SE(\hat{t}_{str}) = \sqrt{\ldots}$

7 (b) from ex... by faculty ... + 5×2+ ... and ... The

7 (a) $\hat{t}_{str} = \sum_{h=1}^{H} N_h \bar{y}_h$ where $\bar{y}_h = \frac{\sum_{j\in S_h} f y_{hj}}{\sum f}$

For Biology: $\bar{y}_b = \frac{(1\times 0)+(2\times 1)+(0\times 2)+(1\times 3)+(0\times 4)+(2\times 5)+(1\times \ldots}{1+2+0+1+0+2+0+1+0}$

$= \underline{3.142}$

for Physics: $\bar{y}_p = \frac{(10\times 0)+(2\times 1)+(0\times 2)+(1\times 3)+(2\times 4)+(1\times 5)+(1\times 6)+(\ldots}{10+2+0+1+2+1+1+0+2}$

$= \underline{2.105}$

for Social $\bar{y}_s = \frac{(9\times 0)+(0\times 1)+(1\times 2)+(0\times 3)+(2\times 4)+(0\times 5)+(1\times 6)}{9+0+1+0+2+0+1+0+0}$

$= \underline{1.23077}$

For Humanities: $\bar{y}_H = \frac{(8\times 0)+(2\times 1)+(3\times 1)+(0\times 4)}{8+2+1} = \underline{0.4545}$

Then $\hat{t}_{str} = (102\times 3.142)+(310\times 2.105)+(217\times 1.23077)+(178\times 0.4545) = \boxed{1,321.012}$

$SE(\hat{t}_{str}) = \sqrt{\hat{Var}(\hat{t}_{str})} = \left[\sum N_h^2\left(1-\frac{n_h}{N_h}\right)\left(\frac{N_h}{N}\right)^2 \frac{s_h^2}{n_h}\right]^{0.5}$ and

$s_h^2 = \frac{\sum_{j\in S_h} f (y_{hj}-\bar{y}_h)^2}{n_h - 1}$. The table below gives the calculated results.

STRATUM	N_h	n_h	$\bar{y}_h = \sum_{j\in S} y_{hj}/n_h$	$\hat{t}_h = N_h\bar{y}_h$	s_h^2	$V(\hat{t}_h) = (1-\frac{n_h}{N_h}) N_h^2 \cdot \frac{s_h^2}{n_h}$
Biological	102	7	3.142	102×3.142 $\boxed{320.484}$	6.8095	9426.294
Physical	310	19	2.105	310×2.105 $\boxed{652.55}$	8.21052	38982.68
Social	217	13	1.23077	217×1.23077 $\boxed{267.077}$	4.3590	14843.40
Humanity	178	11	0.4545	178×0.4545 $\boxed{80.901}$	0.8727	2358.35
	807	50		1321.01		65610.727

$(R\times1)+(3\times2)+$

so $SE(\hat{t}_{str}) = \sqrt{65610.727} = \boxed{256.15}$

7 (b) from exercise 6, the total number of publication by faculty members is $\hat{t}_s = N\bar{y}_s = 807\left[4\times1+2\times3+3\times4+\ldots + 5\times2+6\times1+8\times2+9\times1+10\times1\right]/50 = \boxed{1436.46}$

and $SE(\hat{t}_s) = \sqrt{N^2\left(1-\frac{n}{N}\right)\frac{s^2}{n}}$; $s^2 = \frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{(n-1)}$

Then $s^2 = \frac{1}{49}\left[28(0-1.78)^2 + 4(1-1.78)^2 + \cdots + 1(10-1.78)^2\right]$

$= 7.1931$

So $SE(\hat{t}_s) = \sqrt{807^2\left(1-\frac{50}{807}\right)\frac{7.193}{50}} = \boxed{296.455}$

Comparing we note that $\hat{t}_{str} = 1,321.012$ is less than $\hat{t}_s = 1436.46$ and also $SE(\hat{t}_{str}) = 256.15$ is also less than the $SE(\hat{t}_s) = 296.455$

(c) * This is given by $\hat{p}_{str} = \sum_{h=1}^{H}\frac{N_h}{N}\hat{p}_h = \left(\frac{102}{807}\times\frac{1}{7}\right) + \left(\frac{310}{807}\times\frac{10}{19}\right) + \left(\frac{217}{807}\times\frac{9}{13}\right) + \left(\frac{178}{807}\times\frac{8}{11}\right)$ Reason that $\hat{p}_h = \left(\frac{y_h}{n_h}\right)$

$\Rightarrow \hat{p}_{str} = \boxed{0.5668}$

* $SE(\hat{p}) = \sqrt{\sum_{h=1}^{H}\left(1-\frac{n_h}{N_h}\right)\left(\frac{N_h}{N}\right)^2\frac{\hat{p}_h(1-\hat{p}_h)}{n_h-1}} = \left[\left(1-\frac{7}{102}\right)\left(\frac{102}{807}\right)^2\left(\frac{\frac{1}{7}\left(1-\frac{1}{7}\right)}{(7-1)}\right) + \left(1-\frac{19}{310}\right)\left(\frac{310}{807}\right)^2\left(\frac{\frac{10}{19}\left(1-\frac{10}{19}\right)}{(19-1)}\right) + \left(1-\frac{13}{217}\right)\left(\frac{217}{807}\right)^2\left(\frac{\frac{9}{13}\left(1-\frac{9}{13}\right)}{(13-1)}\right) + \left(1-\frac{11}{178}\right)\left(\frac{178}{807}\right)^2\left(\frac{\frac{8}{11}\left(1-\frac{8}{11}\right)}{(11-1)}\right)\right]^{0.5} = \boxed{0.071}$

7(d) Yes stratification has increase precision.
from part (a) and (b) above we realized that the standard error of the estimate of the population total is smaller when the stratified random sample is considered and so we conclude that stratification has improved the precision in this example

(25) (a) want to prove that $\hat{V}_{Neyman}(\hat{t}_{str}) = \frac{1}{n}\left(\sum_{h=1}^{H} N_h S_h\right)^2 - \sum_{h=1}^{H} N_h S_h^2$

Consider $\hat{V}(\hat{t}_{str}) = \sum_{h=1}^{H}\left(1-\frac{n_h}{N_h}\right)\frac{N_h^2 S_h^2}{n_h}$.

from Neyman allocation, the sample size of stratum is:

$n_h = n\left[\frac{N_h S_h}{\sum_{j=1}^{H} N_j S_j}\right]$ we will substitute into the variance expression above and obtain

$$\hat{V}_{Neyman}(\hat{t}_{str}) = \sum_{h=1}^{H}\left[1 - \frac{n N_h S_h}{N_h \sum_{j=1}^{H} N_j S_j}\right] \cdot \frac{N_h^2 S_h^2}{n\left[\frac{N_h S_h}{\sum_{j=1}^{H} N_j S_j}\right]}$$

$$= \sum_{h=1}^{H}\left[\left(1 - \frac{n S_h}{\sum_{j=1}^{H} N_j S_j}\right)\left(\frac{N_h S_h \sum_{j=1}^{H} N_j S_j}{n}\right)\right]$$

$$= \frac{1}{n}\sum_{h=1}^{H}\left[\left(\frac{\sum_{j=1}^{H} N_j S_j - n S_h}{\cancel{\sum_{j=1}^{H} N_j S_j}}\right)\left(N_h S_h \cancel{\sum_{j=1}^{H} N_j S_j}\right)\right]$$

$$= \frac{1}{n}\sum_{h=1}^{H}\left(N_h S_h \sum_{j=1}^{H} N_j S_j\right) - \frac{1}{n}\sum_{h=1}^{H}\left(N_h S_h \times n S_h\right)$$

$$= \frac{1}{n}\left(\sum_{h=1}^{H} N_h S_h\right)^2 - \sum_{h=1}^{H} N_h S_h^2 \quad \text{as required}$$

(b) want to show for the $V_{prop}(\hat{t}_{str}) = \frac{N}{n}\sum_{h=1}^{H} N_h S_h^2 - \sum_{h=1}^{H} N_h S_h^2$,

$$V_{prop}(\hat{t}_{str}) - \hat{V}_{Neyman}(\hat{t}_{str}) = \frac{N^2}{n}\sum_{h=1}^{H}\frac{N_h}{N}\left(S_h - \sum_{h=1}^{H}\frac{N_h}{N}S_h\right)^2$$

From the prove in part (a) above, $\hat{V}_{Neyman}(\hat{t}_{str}) = \frac{1}{n}\left(\sum_{h=1}^{H} N_h S_h\right)^2 - \sum_{h=1}^{H} N_h S_h^2$

Consider $V_{prop}(\hat{t}_{str}) - \hat{V}_{Neyman}(\hat{t}_{str}) = \frac{N}{n}\sum_{h=1}^{H} N_h S_h^2 - \sum_{h}^{H} N_h S_h^2 - \left[\frac{1}{n}\left(\sum_{h=1}^{H} N_h S_h\right)^2 - \sum_{h=1}^{H} N_h S_h^2\right]$

$$= \frac{N}{n}\sum_{h=1}^{H} N_h S_h^2 - \frac{1}{n}\left(\sum_{h=1}^{H} N_h S_h\right)^2$$

$$= \frac{N^2}{n}\left[\sum_{h=1}^{H}\frac{N_h}{N}S_h^2 - \left(\sum_{h=1}^{H}\frac{N_h}{N}S_h\right)^2\right]$$

$$= \frac{N^2}{n}\sum\frac{N_h}{N}\left[S_h^2 - S_h\sum_{h=1}^{H}\frac{N_h}{N}S_h\right]$$

$$= \frac{N^2}{n}\sum_{h=1}^{H}\frac{N_h}{N}\left[S_h^2 - S_h\sum_{h=1}^{H}\frac{N_h}{N}S_h\right]$$ as prove of first part

Next, consider $\hat{V}_{prop}(\hat{t}_{str}) - \hat{V}_{Neyman}(\hat{t}_{str}) = \frac{N^2}{n}\sum_{h=1}^{H}\frac{N_h}{N}\left[S_h^2 - S_h\sum_{h=1}^{H}\frac{N_h}{N}S_h\right]$

$$= \frac{N^2}{n}\left[\sum_{h=1}^{H}\frac{N_h}{N}S_h^2 - 2S_h\sum_{h=1}^{H}\frac{N_h}{N}S_h + \left(\sum_{h=1}^{H}\frac{N_h}{N}S_h\right)^2\right]$$

$$= \frac{N^2}{n}\left[\sum_{h=1}^{H}\frac{N_h}{N}\left(S_h - \sum_{h=1}^{H}\frac{N_h}{N}S_h\right)^2\right]$$ as the second part of our prove. Hence proven.

25 ⓒ When $H=2$, we want $\hat{V}_{prop}(\hat{t}_{str}) - V_{Neyman}(\hat{t}_{str}) = \frac{N_1 N_2}{n}(S_1-S_2)^2$.

From part ⓑ above, $V_{prop}(\hat{t}_{str}) - \hat{V}_{Neyman}(\hat{t}_{str}) = \frac{N^2}{n}\left[\sum_{h=1}^{H}\frac{N_h}{N}S_h^2 - \left(\sum \frac{N}{N}S_h\right)^2\right]$. Put $H=2$ into this

$$\Rightarrow V_{prop}(\hat{t}_{str}) - \hat{V}_{Neyman}(\hat{t}_{str}) = \frac{N^2}{n}\left[\frac{N_h}{N}\left(S_h - \sum_{h=1}^{H}\frac{N_h}{N}S_h\right)^2\right]$$

$$= \frac{N^2}{n}\left[\frac{N_1}{N}\left(S_1 - \frac{N_1S_1+N_2S_2}{(N_1+N_2)}\right)^2 + \frac{N_2}{N}\left(S_2 - \frac{N_1S_1+N_2S_2}{(N_1+N_2)}\right)^2\right]$$

$$= \frac{N^2}{n}\left[\frac{N_1}{N}\left(\frac{N_2}{N}S_1 - \frac{N_2}{N}S_2\right)^2 + \frac{N_2}{N}\left(\frac{N_1}{N}S_1 - \frac{N_1}{N}S_2\right)\right]$$

$$= \frac{N^2}{n}\cdot\frac{N_1}{N}\frac{N_2}{N}\left(\frac{N_1}{N}+\frac{N_2}{N}\right)(S_1-S_2)^2 = \underline{\underline{\frac{N_1N_2}{n}(S_1-S_2)^2}}$$

proven.