Assignment 1

1.

a). n = 1000 P_1=26.2% P_2=39.62%-10%=29.62%

S1: H_0: $P_1 = P_2$

H_a: $P_1 < P_2$ (Left Tail Test)

S2: n_1=1000*26.2%=262 n_2=1000*29.62%=296.2

$SE(\hat{p}_1 - \hat{p}_2) = sqrt(\hat{p}_1\hat{q}_1/n_1 + \hat{p}_2\hat{q}_2/n_2)$

=sqrt(0.262*0.738/262+0.2962*0.7038/296.2)=0.038

$Z_{calc} = Z_{Test} = (\hat{p}_1 - \hat{p}_2)/SE(\hat{p}_1 - \hat{p}_2)$ = (0.262-0.2962)/0.038=-0.9

S3: Level of Significance, LS=0.01, so CI=1-LS=99%, then

$Z_{crit} = Z_{\alpha}$=2.33

S4: Reach an appropriate conclusion, since $\{|Z_{calc}|=0.9\} < \{Z_{crit}=2.33\}$ ➔Do Not Reject H_0

So, there is insufficient evidence to show that the Conservative vote has dropped by more than 10% from their share of the popular vote in 2011.

b). ME=+/-1% 99%CI= 2.576 $\hat{p}$=0.262 $\hat{q}=1-\hat{p}$=0.738

n= $(Z^*/ME)^2 \hat{p}\hat{q}$

= 2.576^2*0.262*0.738/ $(0.01)^2$=12831

c).

p=0.3962 n=17 $\hat{p}$= 2/17 =0.1176

$n\hat{p}$ =17*0.1176=1.9992<10

$n\hat{q}$ =17*0.8824=15.0008>10

S1: H_0: $\hat{p}=p$

H_a: $\hat{p}<p$ (Left Tail Test)

Since $n\hat{p}$ <10, the normal Hypothesis Test cannot be used, we should use the Binomial Distribution

P-Value=1-P(0)-P(1)-P(2)

$=1-[17!/0!(17-0)!*0.3692^0*0.6038^{17}]-[17!/1!(17-1)!*0.3692^1*0.6038^{16}]-[17!/2!(17-2)!*0.3692^2*0.6038^{15}]$=0.9867

LS=α=0.01

Reach an appropriate conclusion, since [P-Value=0.9867]>[α=0.01]➔ Do Not Reject H_0

So, this is insufficient evidence to indicate that the level of support for the Conservatives among U of O students is lower than the 39.62% share of the popular vote in 2011.

2.

Descriptive Statistics: BMImale, BMIfemale

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
BMImale	50	0	26.300	0.531	3.754	21.400	23.750	25.400	28.325
BMIfemale	40	0	24.812	0.855	5.409	17.900	21.300	23.150	27.000

Variable	Maximum
BMImale	38.800
BMIfemale	43.100

According to the data;
Female: n=40 SD=5.409

S1. H_0: $\mu=(\mu_0=26)$
H_a: $\mu<(\mu_0=26)$ (Left Tail Test)

S2. $T_{calc}=(\bar{x}-\mu_0)/SE(\bar{x})$
$SE(\bar{x})=S/sqrt(n)=5.409/sqrt40=0.855$
$T_{calc}=(24.812-26)/0.855=-1.389$

S3. Level of Significance, LS=0.05, so CI=1-LS=95%, then
$T_{calc}=T_{\alpha}(n-1)=T_{0.05}(39)=0.520$

S4. Reach an appropriate conclusion, since $\{|T_{calc}|=1.389\}>\{T_{crit}=0.520\}$→Do Not Reject H_0
So, it is insufficient evidence to show that the average female BMI(in the population) is less than 26.

3.

Descriptive Statistics: BMImale, BMIfemale

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
BMImale	25	25	28.988	0.701	3.507	25.500	26.400	28.300	30.000
BMIfemale	11	28	32.12	1.45	4.81	26.40	28.80	30.80	34.50

Variable	Maximum
BMImale	38.800
BMIfemale	43.10

Male: $\bar{x}_1=25$ Female: $\bar{x}_2=11$ $n_1=50$ $n_2=40$

a).

S1. H_0: $\hat{p}_1$- $\hat{p}_2$=Δ_0

H_A: $\hat{p}_1$- $\hat{p}_2$≠Δ_0 (Two Tail Test)

S2. Z_{calc}=($\hat{p}_1$- $\hat{p}_2$)/SE($\hat{p}_1$- $\hat{p}_2$)

p-bar=(x_1+x_2)/(n_1+n_2)=(25+11)/(50+40)=0.4

q-bar=1-0.4=0.6

SE($\hat{p}_1$- $\hat{p}_2$)=sqrt[p-bar*q-bar(1/n_1+1/n_2)]=sqrt[0.4*0.6*(1/50+1/40)]=0.1039

Z_{calc}=(25/50-11/40)/0.1039=2.1655

S3. For two tail test, level of Significance, LS=0.05, so CI=1-LS=95%, then

Z_{crit}=1.960

S4. Reach an appropriate conclusion, since {Z_{calc}=2.165}>{Z_{crit}=1.960}→Reject H_0

So, it is sufficient evidence to show that the proportion of overweight females in the population.

b). P-val=(Z>Z_{calc}=2.165)*2=(1-0.984806)*2=0.0304

c). CI: (25/50-11/40)+1.96(0.1039)=0.428644

(25/50-11/40)-1.96(0.1039)=0.021356

CI: 0.021356 ~ 0.428644

d). From question b, since {p-Val=0.0304}<{α=0.05}→Reject H_0

From question c, $\hat{p}_1$- $\hat{p}_2$=25/50-11/40=0.225, since CI is from0.021356 to 0.428644, 0.225 is in the range of CI, so it is Reject H_0.

Therefore, the results from b and c are all reject H_0, they are as same as the conclusion in a.

4.

a).

```
MTB > desc c1
```

Descriptive Statistics: Liab07

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3
Liab07	479	0	39.33	7.57	165.73	0.12	2.18	6.17	17.67

Variable	Maximum
Liab07	2074.03

b)
One-Sample T: C2, C3, C4, C5, C6, C7, C8, C9, ...

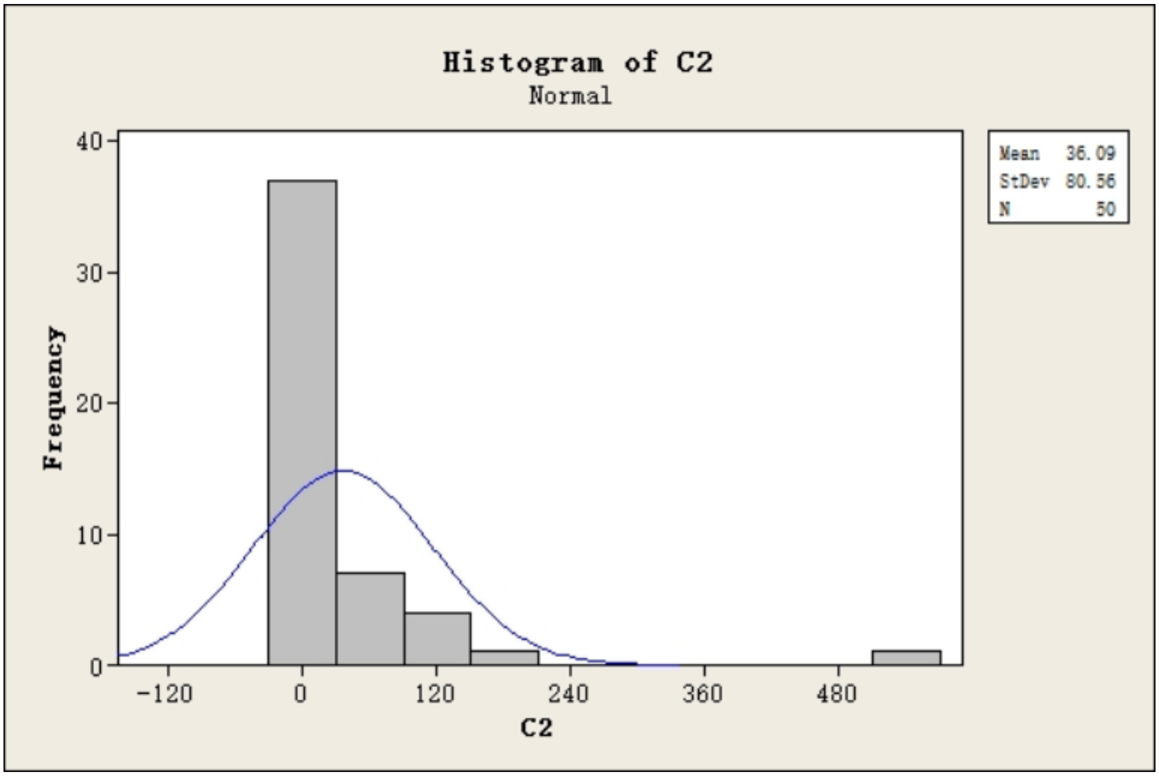
Variable	N	Mean	StDev	SE Mean	95% CI
C2	50	36.1	80.6	11.4	(13.2, 59.0)
C3	50	26.52	44.63	6.31	(13.83, 39.20)
C4	50	31.7	101.5	14.4	(2.8, 60.5)
C5	50	41.7	221.0	31.3	(-21.1, 104.5)
C6	50	16.79	31.46	4.45	(7.85, 25.73)
C7	50	33.9	132.9	18.8	(-3.9, 71.7)
C8	50	85.9	324.1	45.8	(-6.2, 178.1)
C9	50	48.6	203.3	28.8	(-9.2, 106.4)
C10	50	24.93	69.40	9.81	(5.21, 44.66)
C11	50	40.8	157.6	22.3	(-4.0, 85.6)
C12	50	58.6	157.4	22.3	(13.8, 103.3)
C13	50	24.51	49.11	6.94	(10.55, 38.46)
C14	50	26.3	76.6	10.8	(4.5, 48.0)
C15	50	58.8	231.4	32.7	(-7.0, 124.5)
C16	50	31.18	51.43	7.27	(16.56, 45.79)
C17	50	28.9	133.2	18.8	(-9.0, 66.7)
C18	50	70.6	325.9	46.1	(-22.0, 163.2)
C19	50	33.4	100.7	14.2	(4.8, 62.0)
C20	50	13.93	27.29	3.86	(6.17, 21.68)
C21	50	43.5	221.1	31.3	(-19.3, 106.3)

c)
Data Display

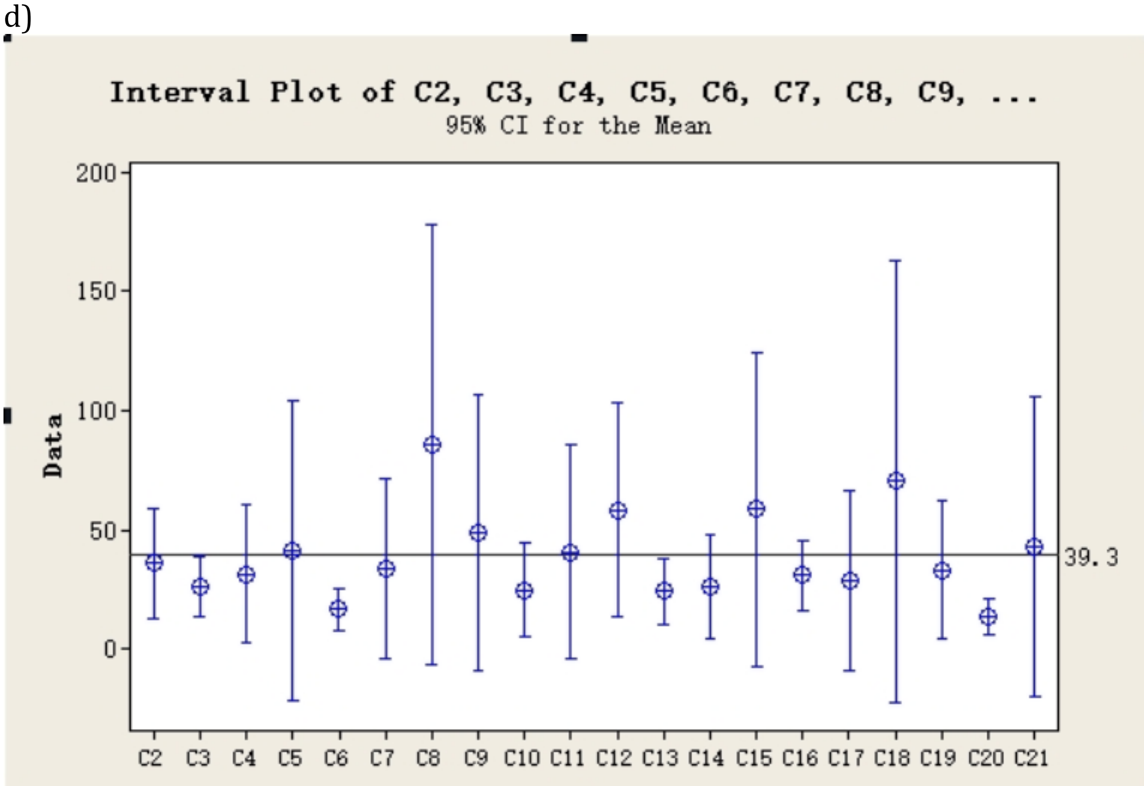
C2

27.233	30.179	6.085	0.755	30.170	57.214	4.351
11.344	1.485	7.863	2.959	0.925	0.174	5.150
3.710	1.073	7.826	13.257	3.164	527.814	24.215
29.388	161.521	0.769	18.930	2.182	1.407	2.498
12.934	31.777	1.978	21.716	0.767	10.717	134.557
121.219	57.010	54.017	18.604	27.759	4.853	9.574
17.124	131.244	47.812	0.401	90.482	1.373	10.013
14.938						

95%CI=1.96
SE($\bar{x}$)=S/sqrt(n)=80.6/sqrt(50)=11.399
CI: 36.1+1.96*11.399=58.442
CI: 36.1-1.96*11.399=13.758
CI:(13.758,58.442)
So, CI95% is (13.758, 58.442)


Histogram of C2
Normal
Mean 36.09
StDev 80.56
N 50
Frequency
C2
0
10
20
30
40
-120
0
120
240
360
480

d)


Interval Plot of C2, C3, C4, C5, C6, C7, C8, C9, ...
95% CI for the Mean
Data
0
50
100
150
200
39.3
C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13 C14 C15 C16 C17 C18 C19 C20 C21

According the graph above, C6, C13, C20 are not reach the mean line. Therefore, there are only 17 out of 20 intervals that contain the true value of the population mean from part a.

e). From the question, the samples that we use to calculate are 50 that mean the sample size is 50. It is not large enough to get the accurate results. There are some errors will appear when we count the data. From question d we only get 17, so it is not necessarily equal to the "expected value" to 19.

Personal Ethics Statement
Individual Assignment:
By signing this Statement, I am attesting to the fact that I have reviewed the entirety of my attached work and that I have applied all the appropriate rules of quotation and referencing in use at the Telfer School of Management at the University of Ottawa, as well as adhered to the fraud policies outlined in the Academic Regulations in the University's Undergraduate Studies Calendar. I further attest that I have knowledge of and have respected the "Beware of Plagiarism" brochure found on the Telfer School of Management web site.

__Xi Liu___	January 31, 2014______
Signature	Date
__Liu Xi________________	6967614_________
Last Name (print), First Name (print)	Student Number