Statistics

Q1 i) Modal Class = class that appears most often ⇒ $\boxed{6-8}$

ii) Continuous as the times are in a particular range and are not set values e.g 2mths, 3mths, 4mths

iii) Calculator work – see handout in class for method:

Mean = 6 Std Dev = 2·9

By hand:

x	f		(x - x̄)		
Time in months	People	xf	d	d^2	d^2f
(1) 0-2	14	14	-5	25	350
(3) 2-4	17	51	-3	9	153
(5) 4-6	24	120	-1	1	24
(7) 6-8	36	252	1	1	36
(9) 8-10	18	162	3	9	162
(11) 10-12	11	121	5	25	275
	120	720			1000

$$\text{Mean} = \frac{\Sigma xf}{\Sigma f} = \frac{720}{120} = \boxed{6}$$

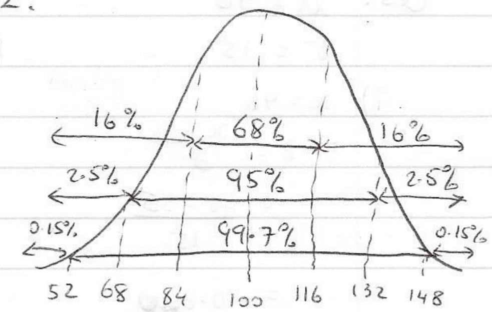
$$\text{Std Dev} = \sqrt{\frac{1000}{120}} = 2{\cdot}88 = \boxed{2{\cdot}9}$$

iv) No of people between (6 - 2·9) and (6 + 2·9) which is between 3·1 and 8·9 – As 3 and 9 are mid interval values we can take ½ the no. of people in those intervals as an estimate

$$\Rightarrow \frac{17}{2} + 24 + 36 + \frac{18}{2} \approx \boxed{78}$$

Q2.

i)



ii) a) 50%

b) 95% + 2·5% = 97·5%

c) 16%

d) 68% + 16% = 84%

e) 95% + 2·5% = 99·7%

Q3.

i) $\hat{p} = \frac{37}{100} = 0{\cdot}37$

If Rejected ⇒ Margin of error < 0.03

$$\Rightarrow \frac{1}{\sqrt{n}} < 0{\cdot}03$$

$$\Rightarrow \frac{1}{n} < 0{\cdot}0009$$

$$\Rightarrow n > \frac{1}{0{\cdot}0009}$$

$$\Rightarrow n > 1111{\cdot}11$$

$$\Rightarrow \boxed{n = 1112}$$

Q4. $\hat{p} = \frac{34}{100} = 0{\cdot}34$

$$E = 1{\cdot}96\sqrt{\frac{p(1-p)}{n}}$$

$$= 1{\cdot}96\sqrt{\frac{0{\cdot}34(1-0{\cdot}34)}{100}} = 0{\cdot}093$$

$$\Rightarrow 95\%\ CI = 0{\cdot}34 - 0{\cdot}09 < p < 0{\cdot}34 + 0{\cdot}09$$

$$\Rightarrow \boxed{25\% < p < 43\%}$$

Q5. $\mu = 56$

$\sigma = 15$

$x = 46$

$\Rightarrow z = \frac{x-\mu}{\sigma}$

$= \frac{46-56}{15}$

$= -0.6$

$\Rightarrow$ Using Tables

$P(z < -0.6)$

$= 1 - P(z < 0.6)$

$= 1 - 0.7257$

$= 0.2743$

$\Rightarrow$ 27% of the popn is below his score

$\Rightarrow \boxed{27^{th} \text{ percentile}}$

Q6. i) No as this was only the mean of the sample.

ii) Not told the standard deviation so can only use the simplified version of the margin of error formula i.e.

$E = \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{100}} = \frac{1}{10}$

$\Rightarrow \boxed{E = 10\%}$

iii) $E = 5\%$

$\Rightarrow 0.05 = \frac{1}{\sqrt{n}}$

$\Rightarrow \frac{1}{400} = \frac{1}{n}$

$\Rightarrow \boxed{n = 400}$

Q7. $\hat{p} = \frac{513}{950} = 0.54$

$E = 1.96\sqrt{\frac{p(1-p)}{n}}$

$= 1.96\sqrt{\frac{(0.54)(1-0.54)}{950}}$

$= 0.0317$

$\Rightarrow 0.54 - 0.0317 < p < 0.54 + 0.0317$

$0.508 < p < 0.5717$

$\Rightarrow \boxed{50.8\% < p < 57.2\%}$

As 0.54 is within the confidence interval, there is enough evidence (with 95% confidence)

Q9. $\mu = 66\%$

$\sigma = 12\%$

Top 15% $\Rightarrow$ 85% below

$\Rightarrow$ 0.85 from Tables = z score of 1.04

$\Rightarrow z = \frac{x-\mu}{\sigma}$

$\Rightarrow 1.04 = \frac{x - 0.66}{0.12}$

$\Rightarrow x - 0.66 = 1.04(0.12)$

$\Rightarrow x - 0.66 = 0.1248$

$x = 0.78 = \boxed{78\%}$

Q10. i) Rank data first i.e.

Q_1

202, 213, 321 | 332, 523, 543, | 611, 615, 653, | 715, 755, 1187

Median

Q_3

$$\text{Median} = \frac{12+1}{2} = 6.5^{th} \text{ value}$$
$$= \frac{543+611}{2} = \boxed{577}$$

$$\text{Median of lower 6} = \frac{6+1}{2} = 3.5^{th}$$
$$= \frac{321+332}{2} = \boxed{326.5}$$

$$\text{Median of upper 6} = 3.5^{th} \text{ value}$$
$$= \frac{653+715}{2} = \boxed{684}$$

$$\Rightarrow LQ = 326.5$$
$$UQ = 684$$

ii) IQ Range = UQ − LQ

$$= 684 - 326.5$$
$$= \boxed{357.5}$$

iii) $P_{15} = 15\% \text{ of } (12+1)$

$$= 1.95$$

⇒ average of 1st + 2nd

$$= \frac{202+213}{2}$$
$$= \boxed{207.5}$$

$P_{80} = 80\% \text{ of } (13)$

$$= 10.4$$

⇒ Average of 10th + 11th

$$= \frac{715+755}{2}$$
$$= \boxed{735}$$

Q11. a) Positively Skewed

b) Normal Distribution

c) Negatively Skewed

Q12. i) $E = \frac{1}{\sqrt{n}}$

$$0.01 = \frac{1}{\sqrt{n}}$$
$$\frac{1}{10000} = \frac{1}{n}$$
$$\Rightarrow \boxed{n = 10{,}000}$$

ii) Costs & effort required might be too large to collect a sample of that size.

Q13. $\hat{p} = \frac{92}{200} = 0.46$

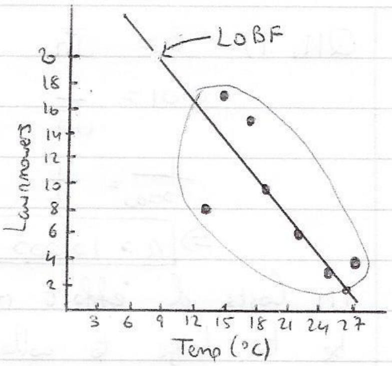
i)

$$E = 1.96\sqrt{\frac{0.46(1-0.46)}{200}}$$
$$= 0.0691$$
$$\Rightarrow 0.46 - 0.0691 < p < 0.46 + 0.0691$$
$$\boxed{0.391 < p < 0.529}$$

ii) No as the CI only refers to the support level on the day of the survey. Opinions may have changed before the election itself

iii) 45% lies in our CI so we can accept with 95% certainty.

Q14. i)



ii) $\boxed{r = -0.74}$

iii) $(x_1, y_1) = (8, 20)$

$(x_2, y_2) = (26, 1)$

$\Rightarrow \text{Slope} = \frac{y_2 - y_1}{x_2 - x_1} = \boxed{-0.76}$

For every 1°C rise in temperature the no. of lawnmowers decreases by 0.76.

iv) Outlier $= (13, 8)$

Q15. $\mu = 10 \quad \sigma = 0.05$

i) $z = \frac{x - \mu}{\sigma} = \frac{10.08 - 10}{0.05}$

$= 1.6$

$\Rightarrow P(z > 1.6) = 1 - P(z < 1.6)$

$= 1 - 0.9452$

$= 0.0548$

ii) $z_1 = \frac{9.9 - 10}{0.05} = -2$

$z_2 = \frac{10.12 - 10}{0.05} = 2.4$

$\Rightarrow P(z < -2) \text{ OR } P(z > 2.4)$

$= [1 - P(z < 2)] + [1 - P(z < 2.4)]$

$= (1 - 0.9772) + (1 - 0.9918)$

$= \boxed{0.031}$

iii) Out of 1000 discs number that would be rejected

$= 0.031 \times 1000 = 31$

$\Rightarrow$ Should make $\boxed{1031}$

Q16.

i) $\hat{p} = \frac{107}{200} = 0.535$

$E = 1.96 \sqrt{\frac{0.535(1 - 0.535)}{200}}$

$= 0.07$

ii) CI:

$0.535 - 0.07 < p < 0.535 + 0.07$

$0.465 < p < 0.605$

As 62.5% is not in the interval we can say with 95% confidence that support has changed.

Q17. $0.28 < p < 0.36$

$\Rightarrow E = \frac{0.36 - 0.28}{2} = 0.04$

$\Rightarrow 0.04 = \frac{1}{\sqrt{n}}$

$\Rightarrow \frac{1}{625} = \frac{1}{n}$

$\Rightarrow \boxed{n = 625}$

Q18. $H_0 : \mu = 1600$

$H_1 : \mu \neq 1600$

$\bar{x} = 1571$

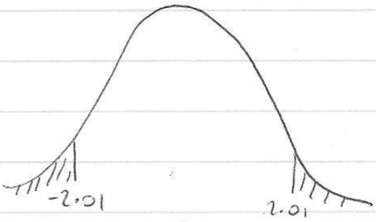
$s = 250$

$n = 300$

i) $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{1571 - 1600}{\frac{250}{\sqrt{300}}}$

$= -2.01$

ii)



$p\text{-value} = P(z > 2.01) \times 2$

$= [1 - P(z < 2.01)] \times 2$

$= (1 - 0.9778) \times 2$

$= \boxed{0.0444}$

iii) @ 5% ⇒ $\alpha = 0.05$

As the p-value < 0.05

⇒ the result is significant

Q19. i) $\mu = 23.42$

$n = 800$

$\bar{x} =$ €22.92

$s =$ €8.56

$z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{22.92 - 23.42}{\frac{8.56}{\sqrt{800}}}$

$= \boxed{-1.65}$

ii) $p\text{-Value} = P(z < -1.65) \times 2$

$= [1 - P(z < 1.65)] \times 2$

$= (1 - 0.9505) \times 2$

$= 0.0495 \times 2$

$= \boxed{0.099}$

iii) as $\alpha = 0.05$ and the p-value > 0.099, the result is not significant @ the 5% level of significance.

Q20. $H_0 : \mu = 67.5$

$H_1 : \mu \neq 67.5$

Method 1: 95% CI

$E = 1.96 \dfrac{\sigma}{\sqrt{n}} = 1.96 . \dfrac{10}{\sqrt{100}} = 1.96$

$\Rightarrow 69 - 1.96 < \mu < 69 + 1.96$

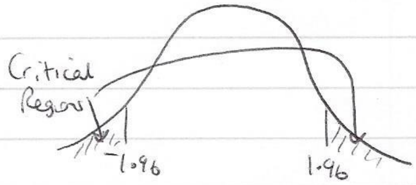
$67.04 < \mu < 70.96$

As 67.5% is in our interval we fail to reject H_0.

⇒ Not enough evidence the students have improved.

Method 2: z-Scores

$z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{69 - 67.5}{\frac{10}{\sqrt{100}}} = 1.5$



As 1.5 is outside the critical region we fail to reject H_0.