1) a) $\hat{\beta} = (X'X)^{-1}X'Y$

$2\hat{\beta}_i = (i^{th} \text{ factor effect})$

With the given X and Y, $\hat{\beta} = (40.917, 0.417, 5.75, -3.5, -0.917, 1.3, 4.5, 1.167)$

$\Rightarrow A = 0.83$

$B = 11.5$

$C = -7$

$AB = -1.83$

$BC = 2.67$

$AC = 9$

$ABC = 2.3$

$\therefore$ The effects of B, C, and AC are relatively higher than the other combinations of treatments.

-0.1 full regression model

b)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
A	1	4.2	4.2	0.129	0.724134	
B	1	793.5	793.5	24.573	0.000143	***
C	1	294.0	294.0	9.105	0.008178	**
A:B	1	20.2	20.2	0.625	0.440930	
B:C	1	42.7	42.7	1.321	0.267254	
A:C	1	486.0	486.0	15.050	0.001330	**
A:B:C	1	32.7	32.7	1.012	0.329486	
Residuals	16	516.7	32.3			

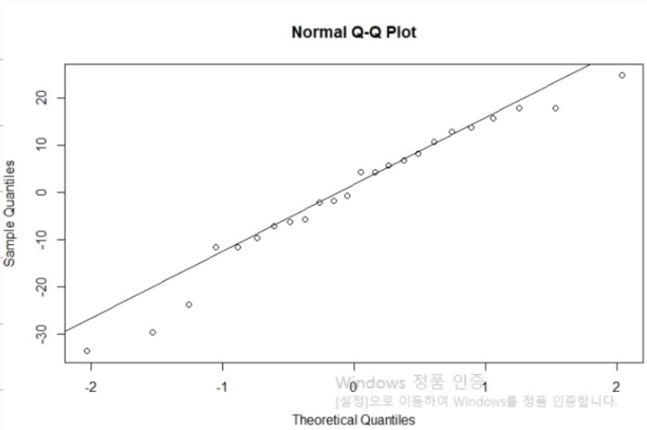
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

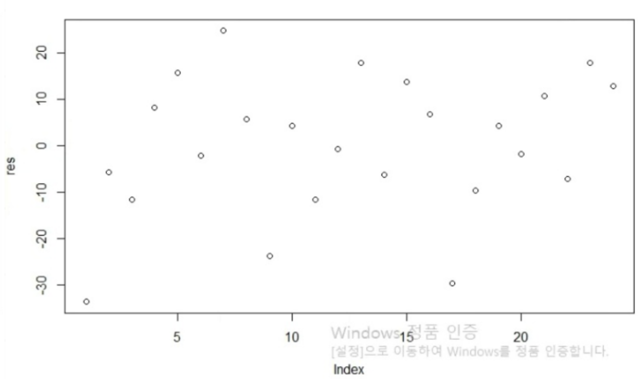
$\therefore$ The table suggests that the conclusion made previously was accurate because the only significant p-values are those of B, C, and AC.

-3.5 4.5

c) $\hat{y} = 40.917 + 5.75X_2 - 3.5X_3 + 4.5X_1X_3$

d)

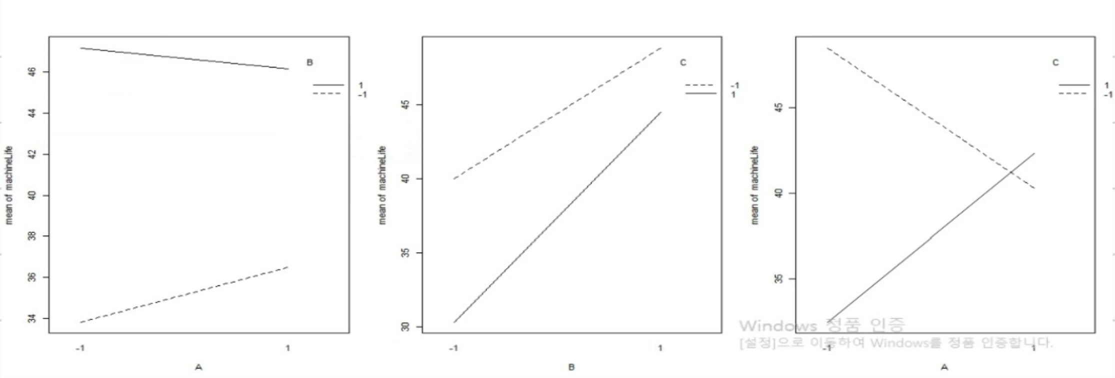




$\therefore$ The QQ-plot and residual plot suggest there is a polynomial trend in residuals, which is against the assumption that $\hat{\varepsilon} \sim N(0, \sigma^2)$. There is an indication suggesting the model used is inappropriate, and I believe the assumption fails in the presence of the interaction term of A and C because A individually does not contribute to the life of machines, yet it is added to the model.

-0.1

e)



∴ As the conclusions made previously suggest, the interaction between A and C clearly exists, and the argument is strengthened noticing the last interaction plot. However, considering the fact that A alone barely affects the results thus eliminated from the model, that the residuals from the model indicates a polynomial trend, and that the crossing point of the interaction of A and C is almost at the tip, I would recommend eliminating the interaction term from the model, which would result in having a model with 2 levels.

2)

-0.5 derive an accurate model and clearly specify range of subscripts

-0.5 : check normality

-0.3 : ANOVA Table

Block 1	Block 2
(1)	a
ab	b
ac	c
bc	abc

A	B	C	Treatment Combination	Re I
-	-	-	(1)	21
+	-	-	a	33
-	+	-	b	34
+	+	-	ab	56
-	-	+	c	43
+	-	+	ac	41
-	+	+	bc	61
+	+	+	abc	40

$\Rightarrow$ Block Effect $= \bar{y}_{b1} - \bar{y}_{b2}$

$$= \frac{(21+56+41+61)-(33+34+43+40)}{4}$$

$$= 7.25$$

$$SS_b = \frac{(179)^2+(150)^2}{4} - \frac{(179+150)^2}{8} = 105.125$$

$\Rightarrow$

	Df	Sum Sq	Mean Sq
B.new	1	351.1	351.1
C.new	1	210.1	210.1
A.new:C.new	2	421.3	210.6
A.new:B.new:C.new	3	160.4	53.5

∴ Because it is unreplicated, the degrees of freedom was not given for the error term so that it is combined within the ABC term. Therefore, I, instead, assumed that the $SS_E = SS_{ABC} - SS_b = 160.4 - 150.125 = 10.275$, and this suggests that all of the terms tested significantly affect the variability.

3)

-1.4 model, ANOVA table, check residual and pairwise comparison

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
B	1	793.5	793.5	25.831	6.62e-05	***
C	1	294.0	294.0	9.571	0.005979	**
A:C	1	486.0	486.0	15.821	0.000807	***
A:B:C	1	32.7	32.7	1.063	0.315392	
Residuals	19	583.7	30.7			

∴ The table suggests that the term B, C, and AC all affect the variability, which agrees to the conclusion made in Problem 1