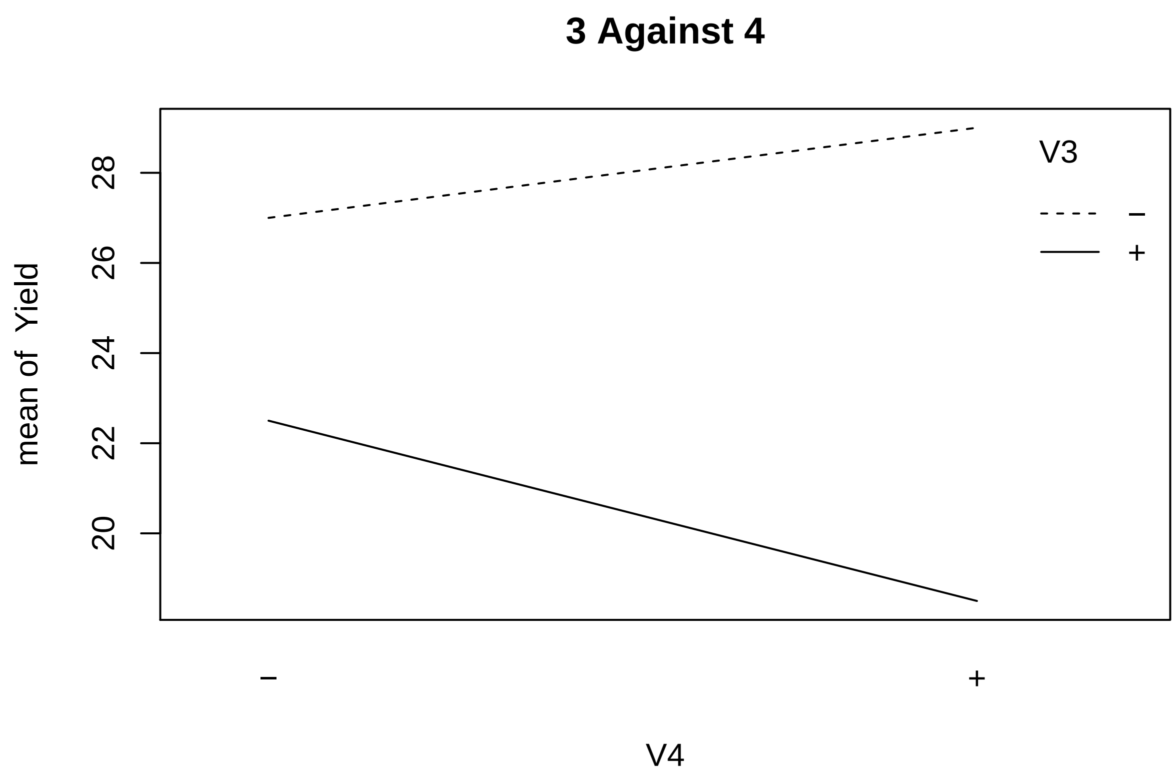
Homework 6 - 6413A

Peter Williams

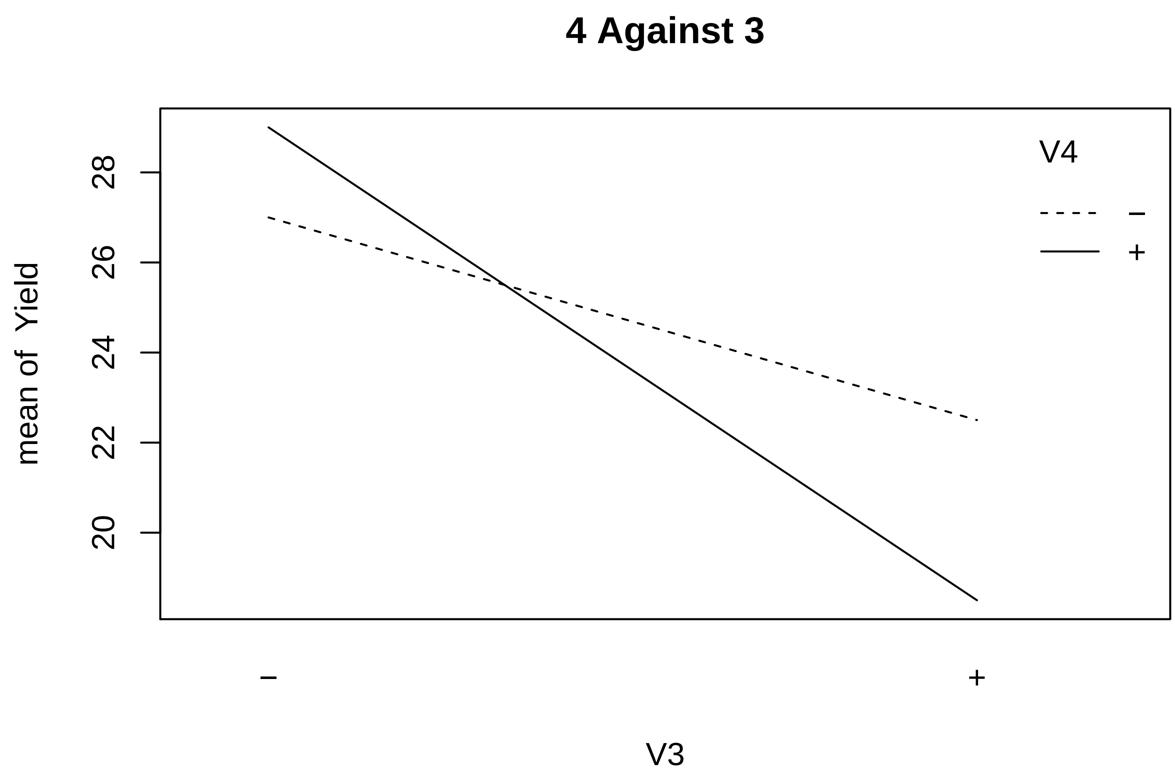
5.7) *An experimenter obtained eight yields for the design given in Table 5.10*

(a) *Make two interaction plot for factors 3 and 4, one for 3 against 4, and the other for 4 against 3.*

```
yieldex <- read.table('http://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/Yield.dat',
                      header=F,stringsAsFactors = F)
colnames(yieldex)[6] <- 'Yield'
with(yieldex, interaction.plot(V4,V3,Yield, main='3 Against 4'))
```




```
with(yieldex, interaction.plot(V3,V4,Yield, main= '4 Against 3'))
```



(b) *Calculate the main effect of 3 and of 4 and the 3 X 4 interaction.*

$ME(3) = \bar{z}(3+) = \bar{z}(3-) = 20.5 - 28.0 =$ -7.5

$ME(4) = \bar{z}(4+) = \bar{z}(4-) = 23.75 - 24.75 =$ -1

$INT(3,4) = \frac{1}{2}[\bar{z}(3+|4+) - \bar{z}(3-|4+)] - \frac{1}{2}[\bar{z}(3+|4-) - \bar{z}(3-|4-)] =$

$\frac{1}{2}(18.5 - 29.0) - \frac{1}{2}(22.5 - 27.0) =$ -3

(c) *If it is know that the standard deviation of each observation (i.e., each run) is 1.5, what is the standard error of the estimates in (b)?*

We can use the formula in (4.14) where $Var(\hat{\theta}) = \frac{4}{N}(1.5^2)$ and $N = 8$, to compute the standard error which is $\sqrt{\frac{4}{N}(1.5^2)} = 1.0606602$.

(d) *Based on the results in (a)-(c), which of the three effect in (b) are signficant? (You can use the effect estimate of ± 2 standard error as the 95% confidence interval to answer this question.) Comment on synergistic and antagonistic interactions.*

We can form a confidence interval using the formula $effect \pm t_{\frac{\alpha}{2},df} S.E.(effect)$

```
SE.effect = sqrt(0.5*(1.5^2))
tvalue = qt(0.975,7)
paste('Confint for Effect 3 = [',round(-7.5+tvalue*SE.effect,digits=3),',',
      round(-7.5-tvalue*SE.effect,digits=3) ,']')
```

```
## [1] "Confint for Effect 3 = [ -4.992 , -10.008 ]"
```

```r
paste('Confint for Effect 4 = [',round(-1+tvalue*SE.effect,digits=3),',',
      round(-1-tvalue*SE.effect,digits=3) ,']')
```

```
## [1] "Confint for Effect 4 = [ 1.508 , -3.508 ]"
```

```r
paste('Confint for Effect 3x4 = [',round(-3+tvalue*SE.effect,digits=3),',',
      round(-1-tvalue*SE.effect,digits=3) ,']')
```

```
## [1] "Confint for Effect 3x4 = [ -0.492 , -3.508 ]"
```

Based on the confidence intervals, the main effect for 3 and the interaction INT(3,4) effect is found to be significant at $\alpha = 0.05$. Since 0 falls in the the range of the confidence interval for effect 4, the effect is not estimated to be significant under the chose α level.

Comments on plots:

The 4 against 3 plot is synergistic, that is Yield decreases at the higher (+) 3 setting, for both settings of variable 4 (+, -). In contrast, the 3 against 4 plot is antagonistic. Yield increases from (-) to (+) settings for variable 4 when variable 3 is at the (-), however, Yield decreases from (-) to (+) when variable 3 is at the (+) setting.

5.21) *(a) By comparing the two graphs in Figures 5.7 and 5.8, and would note that the former has one more line than the latter. Identify this line.* The two factor interaction 56 (BE) line is not present in figure 5.8.

(b) By using the assignment of factors A, B, C, D, E to columns 2,5,3,4,6 as in Section 5.5, identify which additional interaction among the factors can be estimated clearly With $I = ABCDE = 25346, I = 23456 = 125 = 1346$, we can look at the aliasing relationships for two-order fi's that are clear:
$23 = 456 = 135 = 1246$,
$24 = 356 = 145 = 1236$,
$26 = 345 = 156 = 1234$,
$35 = 246 = 123 = 1456$,
$45 = 236 = 124 = 1356$,
$56 = 234 = 126 = 1345$

Figure 5.8 visualizes interactions for BD (45), BC (35), AE (26), AD (24), AC (23), but the interaction 56 (DE), can also be estimated clearly.

5.26) *An experimenter used the following design for studying five variables in eight runs in four blocks of size 2.*

(a) By reordering the runs, write down the design matrix in four blocks of size 2. This design matrix should contain five columns and eight rows, Indicate which two run numbers occur in each block. With block I referring to $(B_2 = (-), B_1 = (-))$, block II $(B_2 = (+), B_1 = (-))$, block III $(B_2 = (-), B_1 = (+))$, and block IV $(B_2 = (+), B_1 = (+))$. The design can be reordered by block as so (retaining original run #s):

Block	Run	V1	V2	V3	V4	V5	B1	B2
I	2	+	-	-	-	+	-	-
I	7	-	+	+	-	-	-	-
II	3	-	+	-	+	+	-	+
II	6	+	-	+	+	-	-	+
III	4	+	+	-	+	-	+	-
III	5	-	-	+	+	+	+	-
IV	1	-	-	-	-	-	+	+
IV	8	+	+	+	-	+	+	+

(b) Explain why and under what conditions the main effect 3 is (or is not) confounded with a block effect. $B_1 = 12$ and $B2 = 13$, further $B_1B_2 = 23$ so with this design the main effect 3 is not confounded with a block effect.

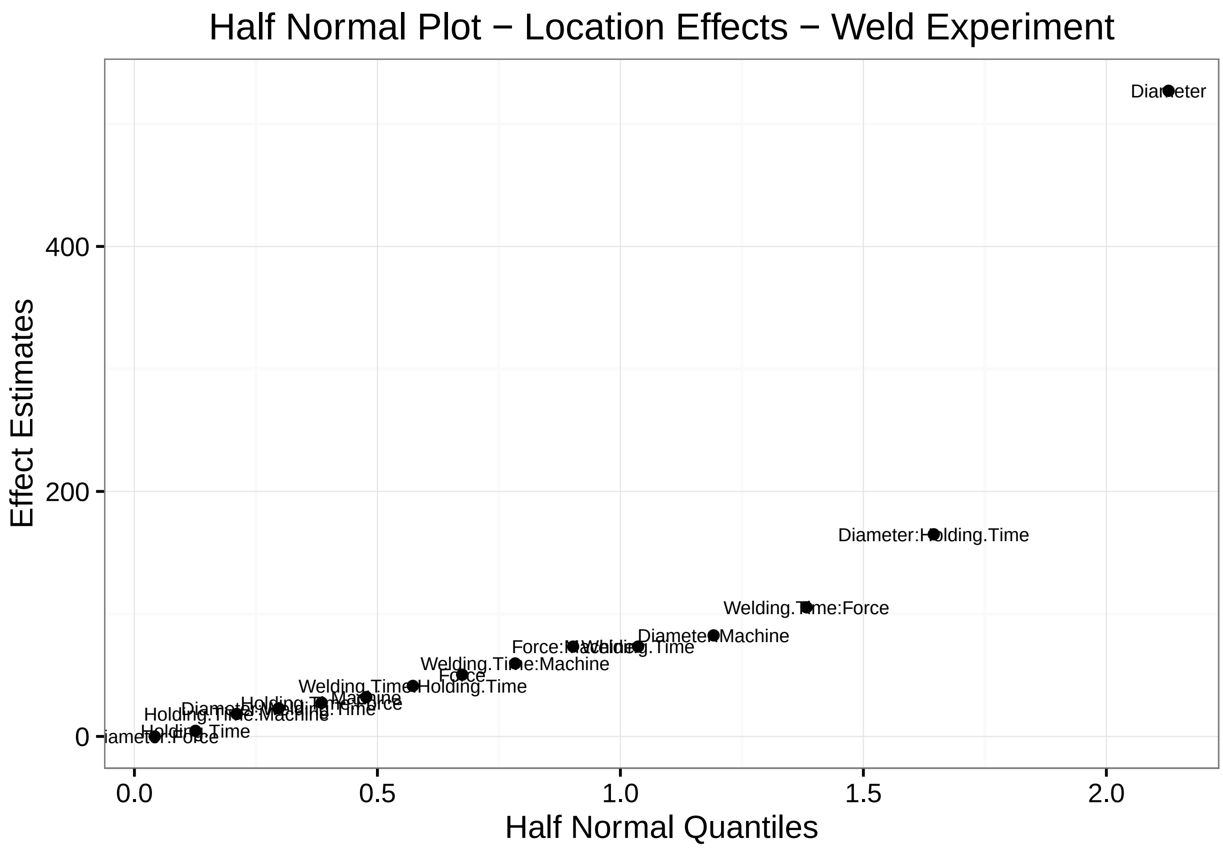
(c) Explain why and under what conditions the main effect 4 is (or is not) confounded with a block effect. In this design $4 = -B_1B_2$. So the main effect 4 is confounded with the (-) block effect B_1B_2. So in this design either B_1B_2 or 4 would be unestimable.

5.28) *In a resistance spot welding experiment, five factors were chosen to study their effect on the tensile strength, which is the maximum load a weld can sustain in a tensile test. The five factor are: button diameter (A), welding time (B), holding time (C), electrode force (D), and machine type (E), each at two levels. The last factor is qualitative, while the others are quantitative. A 2_V^{5-1} design with I = -ABCDE was used for the experiment. Each run has three replicates. The data are given in Table 5.11.*

(a) What effects are clear and strongly clear in this experiment? All the main effects are clear, and all 10 2fi's are clear, but no effects are strongly clear.

(b) Analyze the location and disperion effect separately, including the fitted models for each. Regression is used to fit the model below, and the factorial effects are found by doubling the regression coefficients. The half normal plots below highlight visually the significance of the main effect for diameter, and the interaction between diameter and holding time as significant effects:

```
library(ggplot2)
options(scipen=15)
lmodel <- lm(Strength~Diameter*Welding.Time*Holding.Time*Force*Machine,
             data=weldData)
locCoef <- data.frame(coef(lmodel));
locCoef <- transform(locCoef, effect = row.names(locCoef));
locCoef <- locCoef[2:nrow(locCoef),]; row.names(locCoef) <- NULL;
#using regression coefficients for factorial effects - must double
locCoef$facEffect <- 2*abs(locCoef$coef.lmodel.);
locCoef <- locCoef[!is.na(locCoef$facEffect),2:3]
#get rid of unestimable effects
locCoef$qNorm <- qnorm(0.5+0.5*(rank(locCoef$facEffect)-0.5)/nrow(locCoef))
ggplot(locCoef, aes(x=qNorm,y=facEffect,label=effect)) + geom_point() +
  geom_text(size=2.5) + xlab('Half Normal Quantiles') + ylab('Effect Estimates') +
  ggtitle('Half Normal Plot - Location Effects - Weld Experiment') + theme_bw()
```



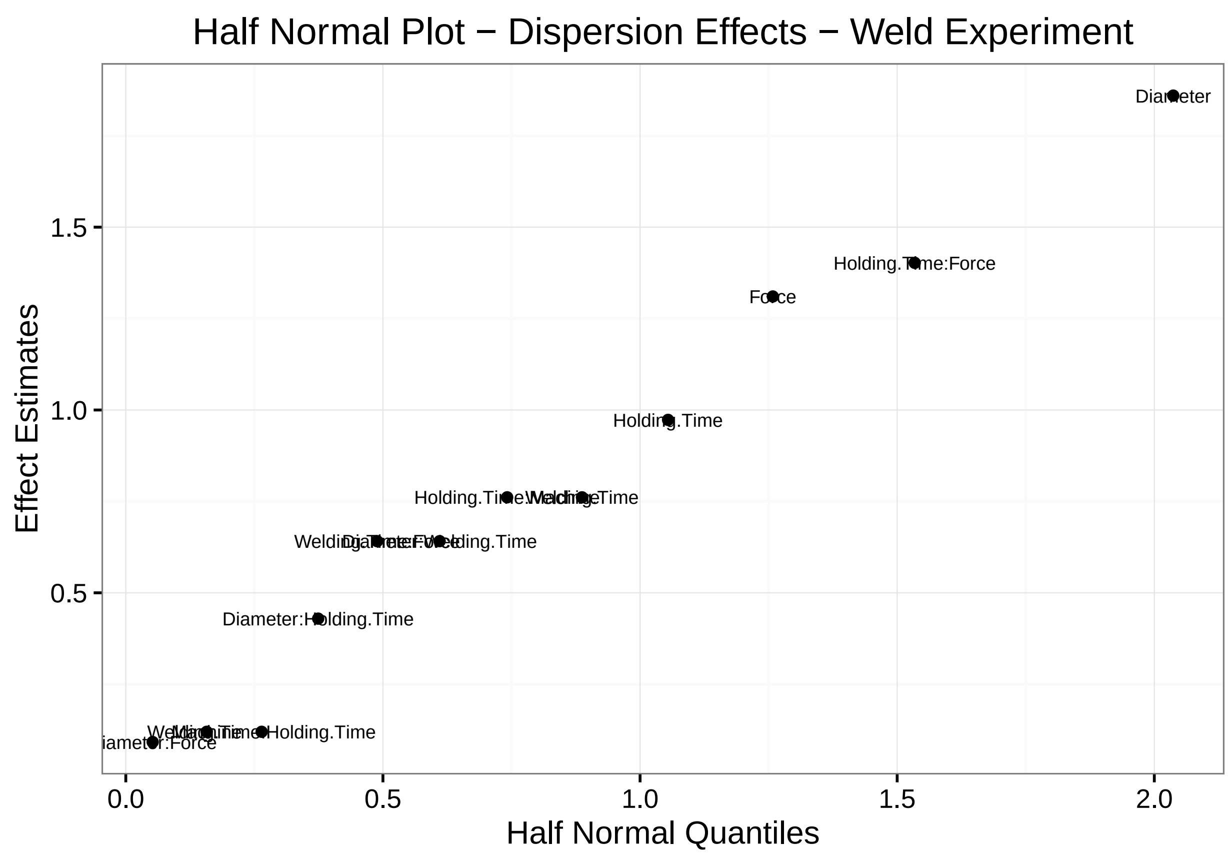

```r
#updated model with variables of interest from half normal plot
lmodel <- lm(Strength~Diameter + Diameter:Holding.Time,data=weldData)
summary(lmodel)
```

```
##
## Call:
## lm(formula = Strength ~ Diameter + Diameter:Holding.Time, data = weldData)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -258.96  -93.96   16.04  101.98  236.04
##
## Coefficients:
##                        Estimate Std. Error t value             Pr(>|t|)
## (Intercept)             1605.00      18.93  84.779 < 0.0000000000000002 ***
## Diameter                 263.54      18.93  13.921 < 0.0000000000000002 ***
## Diameter:Holding.Time     82.50      18.93   4.358            0.0000753 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 131.2 on 45 degrees of freedom
## Multiple R-squared:  0.8254, Adjusted R-squared:  0.8177
## F-statistic: 106.4 on 2 and 45 DF,  p-value: < 0.00000000000000022
```

This yields a regression equation of the following form $\hat{y} = 1605 + 263.54X_D + 82.5X_DX_H$, with D denoting Diameter and H, Holding Time.

For the dispersion analysis, we compute the log sample variance, dropping cases where there is no dispersion among replicates (i.e. $s^2 = 0$).

```
library(plyr)
dispData <- ddply(weldData,.(Diameter,Welding.Time,Holding.Time,Force,Machine),
                  summarise, lns2 = log(var(Strength)),nrep=length(Strength))
#drop where s2 = 0
dispData <- dispData[dispData$lns2 != -Inf,]
dmodel <- lm(lns2~Diameter*Welding.Time*Holding.Time*Force*Machine,data=dispData)
dispCoef <- data.frame(coef(dmodel));
dispCoef <- transform(dispCoef, effect = row.names(dispCoef));
dispCoef <- dispCoef[2:nrow(dispCoef),]; row.names(dispCoef) <- NULL;
#using regression coefficients for factorial effects - must double
dispCoef$facEffect <- 2*abs(dispCoef$coef.dmodel.);
dispCoef <- dispCoef[!is.na(dispCoef$facEffect),2:3]
#get rid of unestimable effects
dispCoef$qNorm <- qnorm(0.5+0.5*(rank(dispCoef$facEffect)-0.5)/nrow(dispCoef))
ggplot(dispCoef, aes(x=qNorm,y=facEffect,label=effect)) + geom_point() +
  geom_text(size=2.5) + xlab('Half Normal Quantiles') + ylab('Effect Estimates') +
  ggtitle('Half Normal Plot - Dispersion Effects - Weld Experiment') + theme_bw()
```




```
#update dispersion model with factors deemed interesting from half normal plot
dmodel <- lm(lns2~Diameter+Force + Holding.Time:Force,data=dispData)
summary(dmodel)
```

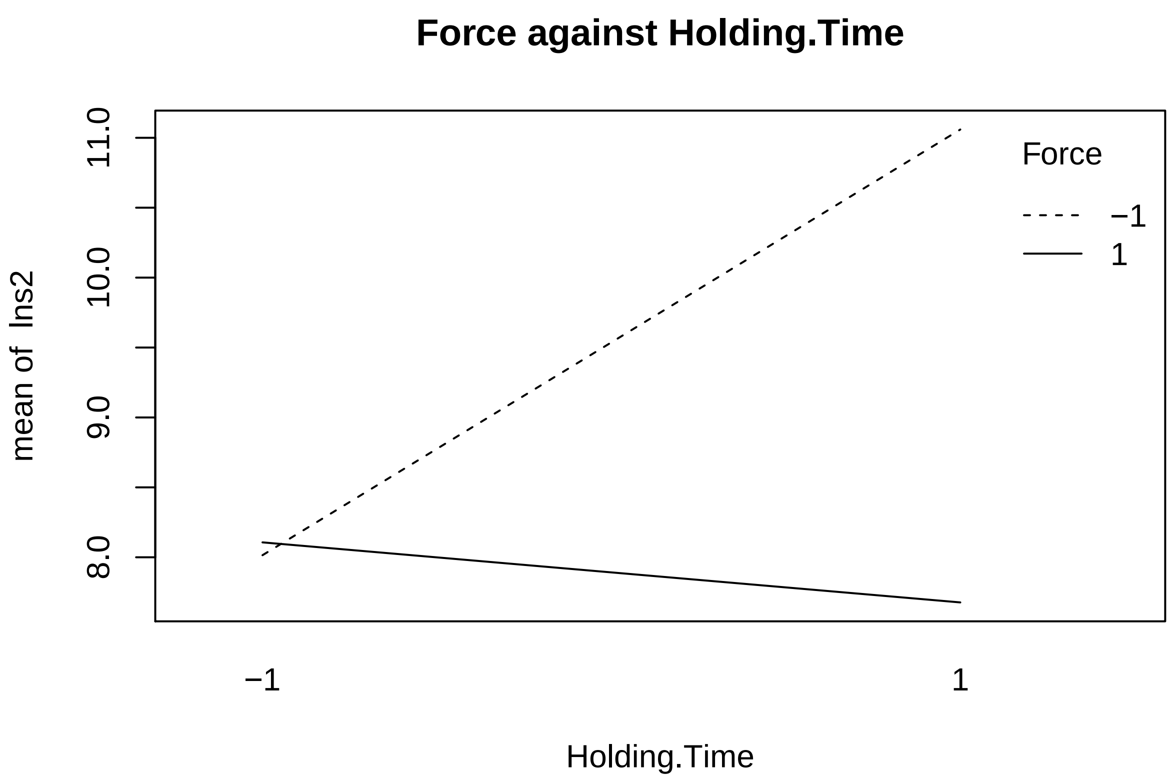
```
##
```

```
## Call:
## lm(formula = lns2 ~ Diameter + Force + Holding.Time:Force, data = dispData)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6326 -0.3191 -0.2088  0.5462  1.4119
##
## Coefficients:
##                    Estimate Std. Error t value       Pr(>|t|)
## (Intercept)          8.2786     0.3134  26.412 0.000000000773 ***
## Diameter            -1.0894     0.2932  -3.716         0.0048 **
## Force               -0.3867     0.3134  -1.234         0.2486
## Force:Holding.Time  -0.4327     0.3134  -1.380         0.2008
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.04 on 9 degrees of freedom
## Multiple R-squared:  0.6776, Adjusted R-squared:  0.5701
## F-statistic: 6.306 on 3 and 9 DF,  p-value: 0.01361
```

This yields a regression equation of the following form $\hat{z} = 8.28 + -1.09X_D + -0.39X_F + -0.43X_F X_H$, with D denoting Diameter and H, Holding Time, and F Force.

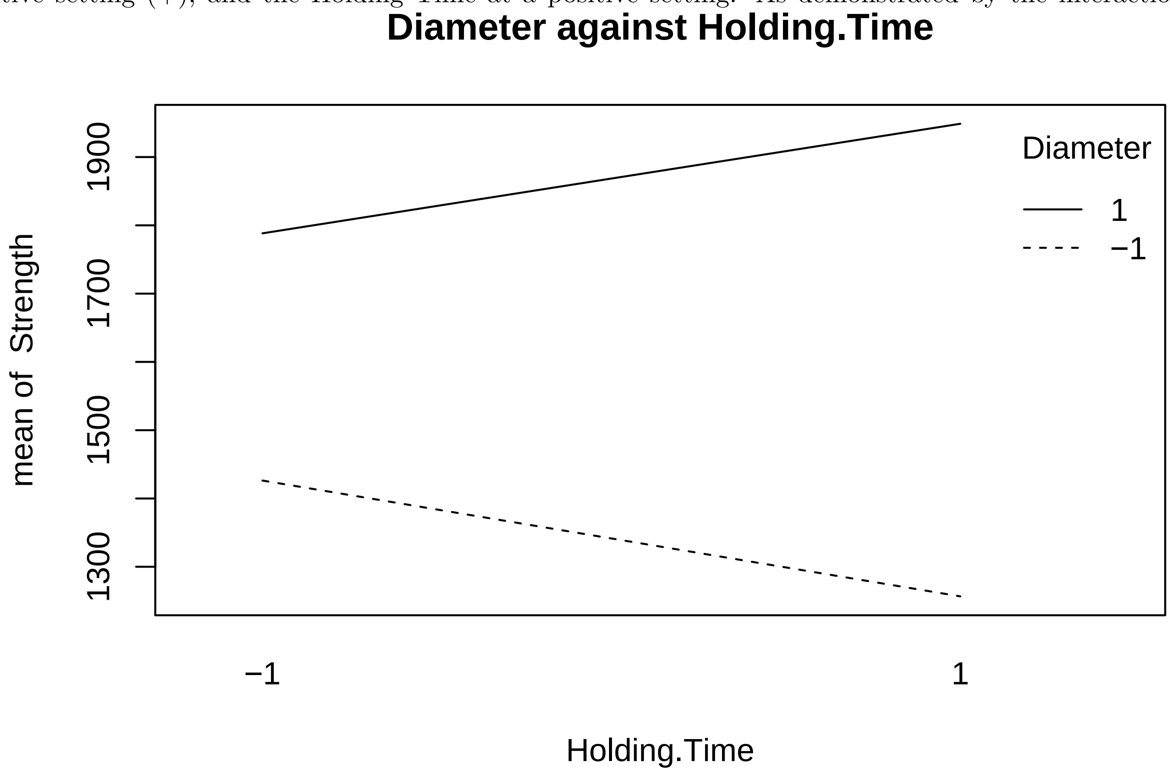
(c) For the dispersion effect model in (b), interpret the meaning of the significant effects. Use this model or an interaction effects plot to select optimal factor settings.
Based on the coefficients from the updated regression model above, variance is reduced at a (+) setting of Diameter, and can be minimized at the (+) setting for Force, and Holding.Time which is further demonstrated by the interaction plot below:

```
with(dispData, interaction.plot(Holding.Time,Force,lns2, main='Force against Holding.Time'))
```



(d) Based on the location effect model in (b), select the optimal factor settings in terms of maximizing the tensile strength.

Based on the updated, reduced model, tensile strength is maximized at the Diameter, D at a positive setting (+), and the Holding Time at a positive setting. As demonstrated by the interaction plot:



(w) Is there any adjustment factor? In the case there is, identify this factor.

Since there is not a target value for tensile strength in this problem, only the objective to maximize the value, there isn't an adjustment factor. So just the objective of maximizing the tensile strength while minimizing its variance is considered.