Cassie Noble
STAT 541
Project Report

For my project I sent out emails to subjects at various times and days, and measured the email open rate as my response variable. The subjects were individuals who had recently created a profile in the online system I work with. A random number generator was used to randomly assign subjects to treatment groups. The design was a two-factor factorial design.

The two factors were the day of the week the email was sent out, and the time of the day the email was sent out. There were three levels of both factors. For days of the week I was interested in measuring responses at the beginning of the week (Monday), mid-week (Wednesday), and the end of week (Friday). For time of day I was interested in measuring responses in the morning (8AM), noon (12PM), and afternoon (4PM). The response was the proportion of individuals who opened the email as measured one week after the email was sent out. There were 108 subjects in this study, with 12 subjects in each of the 9 treatment groups. Each individual response was measured as either opening or not opening the email, and then the proportion of emails opened for each treatment group was compared.

The two main hypotheses for this study were regarding the two factors. The null hypothesis for the day factor was that the email open rate is equal across different days of the week that they receive an email. The null hypothesis for the time factor was that the email open rate is equal across different times of the day that they receive an email.

The model for this design is: $y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$. In this model we have the baseline mean open rate μ. We also have α_i as the effect associated with the ith day of the week they received the email. And we have β_j as the effect associated with the jth time of the day they received the email. The random error of the kth observation from the (i,j)th cell is represented by ε_{ijk}, and is IDD N(0, σ^2). All effects in this model are considered fixed.

The results of the overall F test are shown in Table 1. The overall model was significant with F=8.57 and p=0.031, providing strong evidence that not all groups were equal in terms of open rates. Additional results from this test are shown in Table 2. The R^2 of 0.90 tells us that the effects in the model explain 90% of the variability in the response.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	0.06286667	0.01571667	8.57	0.0305
Error	4	0.00733333	0.00183333		
Corrected Total	8	0.07020000			

Table 1. Overall F test for the model.

R-Square	Coeff Var	Root MSE	open Mean
0.895537	9.445024	0.042817	0.453333

Table 2. Other results for the F test, including an R^2 of 0.90.

Table 3 provides results for the two factors of interest. We have weak evidence to reject the null hypothesis that the day of the week an email is received has no effect on the email open rate (F=5.29, p=0.0752). Table 4 shows the means and standard deviations for the response of open rate by day of the week. Monday has the lowest open rate, and Friday the highest open rate. Wednesday has the greatest variability. These patterns are also shown in Figure 1.

Source	DF	Type III SS	Mean Square	F Value	Pr > F
day	2	0.01940000	0.00970000	5.29	0.0752
time	2	0.04346667	0.02173333	11.85	0.0208

Table 3. F tests for the day and time factors.

Level of day	N	open Mean	open Std Dev
FRI	3	0.50000000	0.08000000
MON	3	0.39000000	0.05196152
WED	3	0.47000000	0.12767145

Table 4. Means and standard deviations of open rate by day.

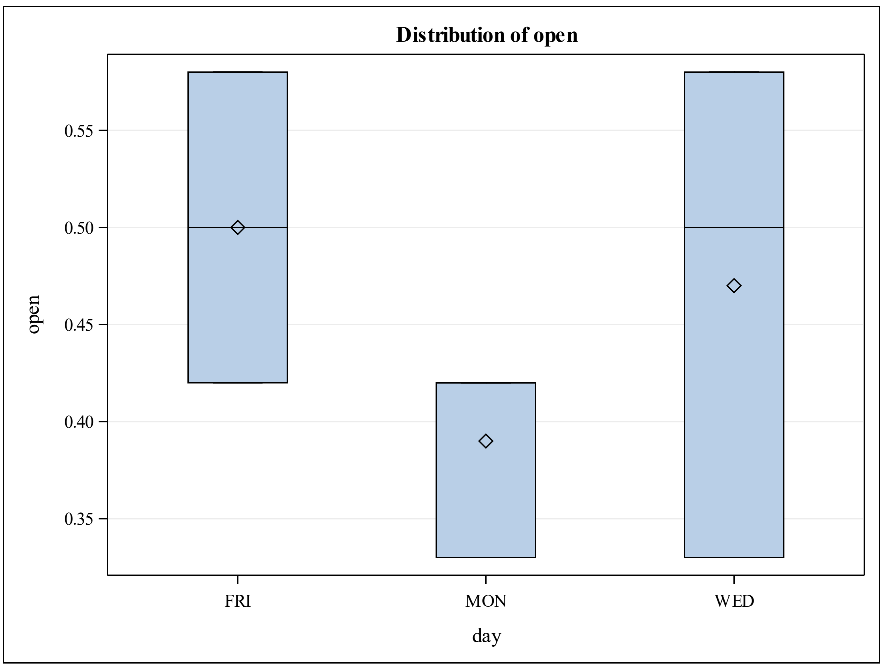

Figure 1. Open rate by day of the week.

Back to Table 3, we also have moderate evidence to reject the null hypothesis that the time of the day an email is received has no effect on the email open rate (F=11.85, p=0.0208). Looking at Table 5, we can see the afternoon had the highest open rate, and mid-day had the lowest open rate. The afternoon had the largest variability among these groups. This pattern can also been seen in Figure 2.

Level of time	N	open Mean	open Std Dev
AM	3	0.47333333	0.04618802
NOON	3	0.36000000	0.05196152
PM	3	0.52666667	0.09237604

Table 5. Means and standard deviations of open rate by time.

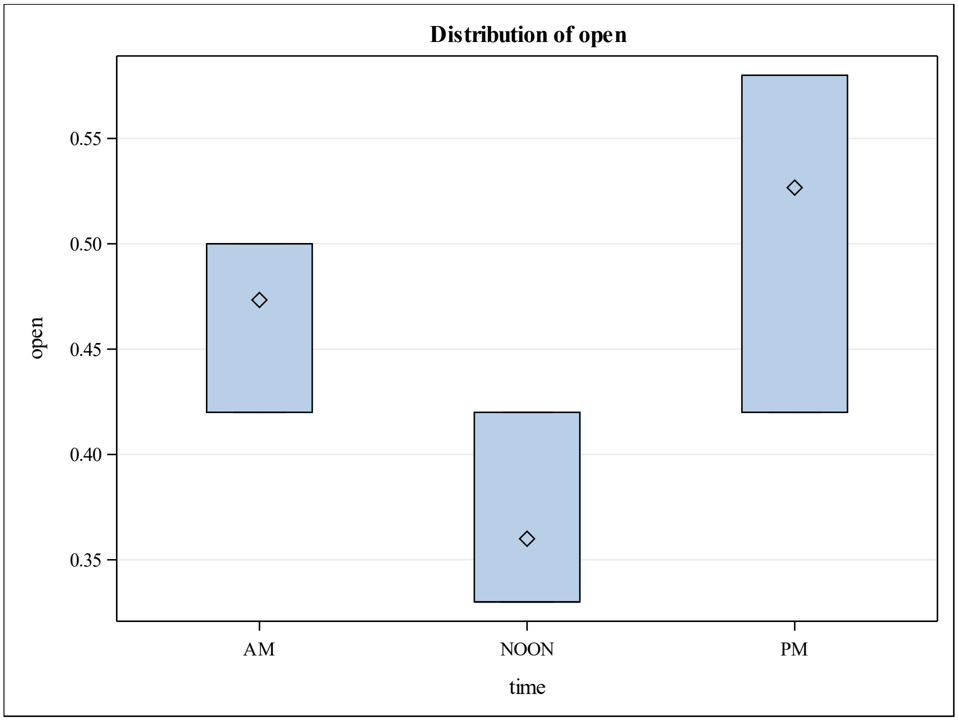

Figure 2. Open rate by time of the day.

Figure 3 shows the interaction plot for day by time combinations. A test for an interaction could not be performed because there were no replicates in the way this experiment was designed. However the clear parallelism in this plot indicates that we likely do not have an interaction effect.

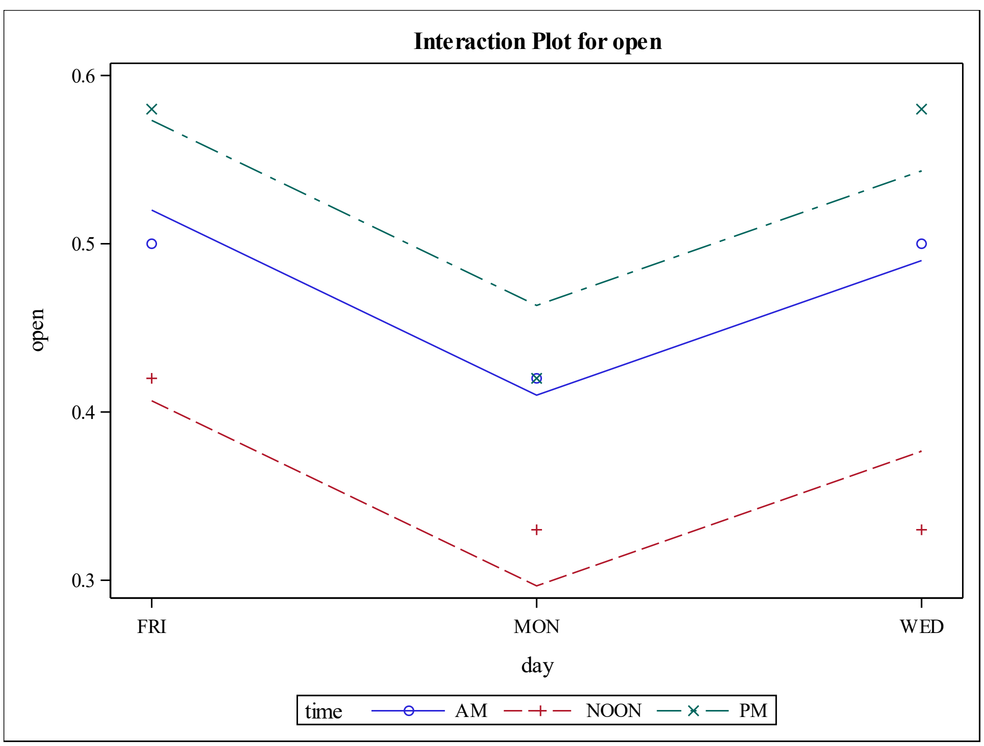

Figure 3. Interaction plot indicating no evidence of an interaction effect.

Table 6 shows the, parameter estimates for the model. This includes the baseline mean $\hat{\mu} = 0.4533$, the day effects, $\hat{\alpha}_{MON} = 0.0467, \hat{\alpha}_{WED} = -0.0633, \hat{\alpha}_{FRI} = 0.0167$, and the time effects $\hat{\beta}_{AM} = 0.0200,\ \hat{\beta}_{NOON} = -0.0933$, and $\hat{\beta}_{PM} = 0.0733$.

Parameter	Estimate	Standard Error	t Value	Pr > |t|
mu	0.45333333	0.01427248	31.76	<.0001
day=MON	0.04666667	0.02018434	2.31	0.0819
day=WED	-0.06333333	0.02018434	-3.14	0.0349
day=FRI	0.01666667	0.02018434	0.83	0.4554
time=AM	0.02000000	0.02018434	0.99	0.3778
time=NOON	-0.09333333	0.02018434	-4.62	0.0099
time=PM	0.07333333	0.02018434	3.63	0.0221

Table 6. Parameter estimates for the model.

Diagnostics plots were also analyzed to determine if model assumptions were adequately met. These plots are shown in Figure 4. The residual vs. predicted plot does not show any clear funneling, indicating no violation of the constant variance assumption. The residual vs. normal plot shows the residuals follow the 1:1 line pretty closely, indicating no violation of the normality of residuals assumption.

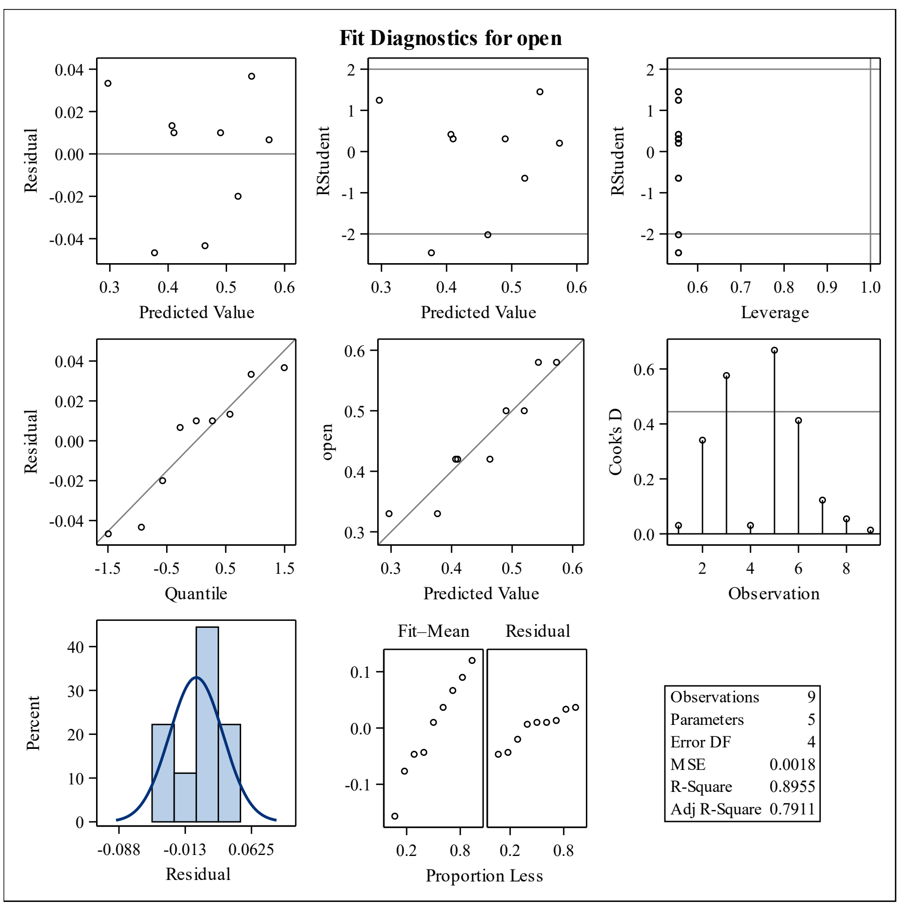

Figure 4. Diagnostics plots for the model.

A multiple comparison test was conducted to test for differences between levels of days and times. Table 7 shows the estimated group means for each day used in the comparison. Table 8 shows the p-values for each comparison. We have very weak evidence to reject the null hypothesis that there is no difference in open rate between the Monday and Friday groups (p=0.1039). We do not have evidence to reject the null hypothesis of no difference between Monday and Wednesday groups (p=0.2520), or Wednesday and Friday groups (p=1).

day	open LSMEAN	LSMEAN Number
FRI	0.50000000	1
MON	0.39000000	2
WED	0.47000000	3

Table 7. Estimated means for each day of the week used in the multiple comparison test.

Least Squares Means for effect day Pr > |t| for H0: LSMean(i)=LSMean(j) Dependent Variable: open			
i/j	1	2	3
1		0.1039	1.0000
2	0.1039		0.2520
3	1.0000	0.2520	

Table 8. Pairwise comparisons for the day effect.

Table 9 shows the estimated group means for each time used in the multiple comparison test. Table 10 shows the p-values for each of these comparisons. We have moderate evidence to reject the null hypothesis of no difference in open rate between the Noon and PM groups (p=0.0266). We have weak evidence to reject the null hypothesis of no difference in open rate between the Noon and AM groups (p=0.0949). We do not have evidence to reject the null hypothesis of no difference between AM and PM groups (p=0.6055).

time	open LSMEAN	LSMEAN Number
AM	0.47333333	1
NOON	0.36000000	2
PM	0.52666667	3

Table 9. Estimated means for each time of the day used in the multiple comparison test.

Least Squares Means for effect time Pr > |t| for H0: LSMean(i)=LSMean(j) Dependent Variable: open			
i/j	1	2	3
1		0.0949	0.6055
2	0.0949		0.0266
3	0.6055	0.0266	

Table 10. Pairwise comparisons for the time effect.

A post-hoc power analysis was conducted to determine the power of the tests conducted. The parameters use for the test, including the estimated error standard deviation, are shown in Table 11. The results of the power analysis are shown in Table 12. For the day factor, we had a power of 0.491. For the time factor, we had a power of 0.820.

Fixed Scenario Elements	
Dependent Variable	open
Alpha	0.05
Error Standard Deviation	0.042817
Total Sample Size	9
Error Degrees of Freedom	4

Table 11. Parameters for the power analysis.

Computed Power			
Index	**Source**	**Test DF**	**Power**
1	day	2	0.491
2	time	2	0.820

Table 12. Estimated power of the tests performed for to test for effects of day and time.

In summary, we have some evidence to suggest that both the day and time that an email is received has an effect on the email open rate. If I were to make a suggestion to my workplace on how to get the highest open rate for emails, I would suggest sending them out Friday afternoons. If that were not possible, I would at least suggest avoiding the middle of the day.

I would also suggest conducting follow up experiments designed in such a way to include replicates. This would greatly increase our power of the test, and make it easier to detect differences between the different days and times. Just looking at the pattern of results, we did not have evidence of an interaction effect, but replication would also allow us to test for that.

Appendix A – SAS code

```
ods rtf file="541experiment.rtf"; /* MS-Word format */

DATA in;
      DO day = 'MON', 'WED', 'FRI';
      DO time = 'AM  ', 'NOON', 'PM  ';
            INPUT open @@; OUTPUT;
      END; END;
      LINES;
    0.42 0.33 0.42
    0.50 0.33 0.58
    0.50 0.42 0.58
    ;

PROC GLM DATA=in PLOTS=(ALL);
    CLASS day time;
    MODEL open = day time / SS3;
    MEANS day time;
      LSMEANS day time / ADJUST=BON;

      *mu*;
      ESTIMATE 'mu' INTERCEPT 1;

      *day effect*;
      ESTIMATE 'day=MON' day 2 -1 -1 / divisor = 3;
      ESTIMATE 'day=WED' day -1 2 -1 / divisor = 3;
      ESTIMATE 'day=FRI' day -1 -1 2 / divisor = 3;

      *time effect*;
      ESTIMATE 'time=AM  ' time 2 -1 -1 / divisor = 3;
      ESTIMATE 'time=NOON' time -1 2 -1 / divisor = 3;
      ESTIMATE 'time=PM  ' time -1 -1 2 / divisor = 3;

TITLE 'TWO-FACTOR FACTORIAL DESIGN';

PROC GLMPOWER DATA=IN;
      CLASS day time;
      MODEL open = day time;
      POWER
            STDDEV = 0.042817
            ALPHA = 0.05
            NTOTAL = 9
            POWER = .
      ;

TITLE 'POWER ANALYSIS';

RUN;

ods rtf close;
```