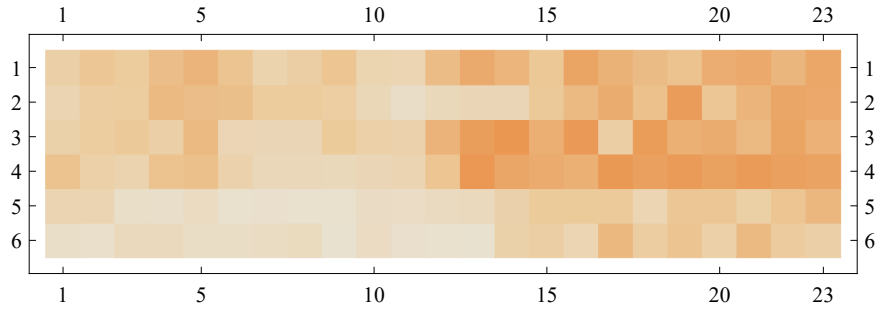
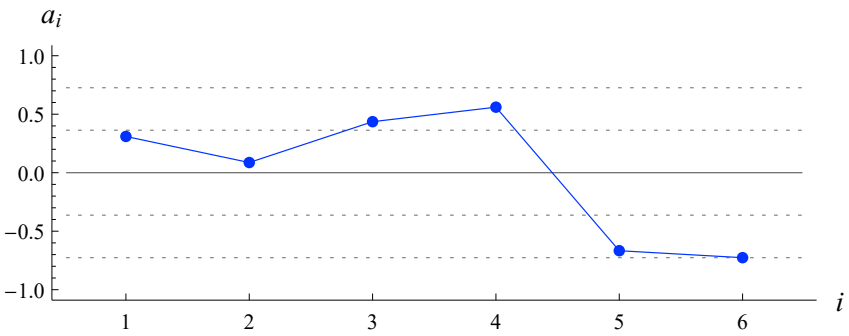
Name: Rufus Petrie

HW8 Problem 5: Oyster Disease Study



Bridgeport Harbor CT	Mystic River CT	Thames River CT	Westport Harbor CT	Oyster Bay NY	Cotuit MA		Overall Mean	Pooled SD
1.69	1.47	1.82	1.94	0.71	0.65		1.38	0.51

AnovaBlocked[prevalences]



	DF	SumOfSq	MeanSq	FRatio	PValue
Group	5	36.3381	7.26763	28.0033	0.
Block	22	47.5151	2.15978	8.32196	7.11×10^{-15}
Error	110	28.548	0.259527		
Total	137	112.401			

```
α = .05;
BonferroniBlocked[prevalences, α]
```

Number of Comparisons:	Rejection Region:	Sampling Distribution:
15	$\|T\| \geq 3.00066$	StudentT (df=110)

Significant mean differences:

$\hat{\mu}_1 - \hat{\mu}_5 = 0.977$ $\quad \hat{\mu}_1 - \hat{\mu}_6 = 1.04$ $\quad \hat{\mu}_2 - \hat{\mu}_4 = -0.473$
$\hat{\mu}_2 - \hat{\mu}_5 = 0.754$ $\quad \hat{\mu}_2 - \hat{\mu}_6 = 0.813$ $\quad \hat{\mu}_3 - \hat{\mu}_5 = 1.1$
$\hat{\mu}_3 - \hat{\mu}_6 = 1.16$ $\quad \hat{\mu}_4 - \hat{\mu}_5 = 1.23$ $\quad \hat{\mu}_4 - \hat{\mu}_6 = 1.29$

Because the group differences has a P-value of approximately zero, we can reject the null hypothesis that the proliferation of the Dermo disease does not vary by site. From the Bonferroni blocked analysis, we see that the site pairings of (1,5), (1,6), (2,5), (2,6), (3,5), (3,6), (4,5), and (4,6) all had statistically significant differences.

```
FriedmanTest[prevalences]
```

Statistic:	PValue:	Distribution:
72.6866	2.83×10^{-14}	ChiSquare(df=5)

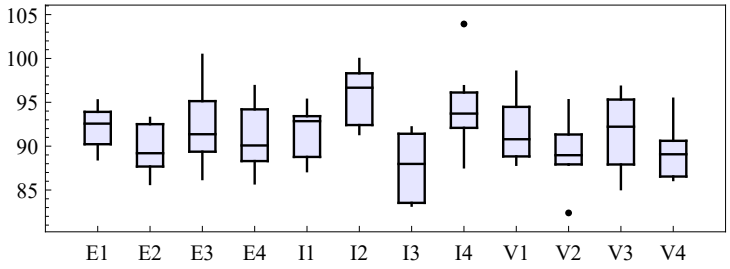
Signed Ranks P Values	*Site 2*	*Site 3*	*Site 4*	*Site 5*	*Site 6*
Site 1	0.1242	0.5449	0.1491	2.3842×10^{-7}	2.3842×10^{-7}
Site 2		0.1974	0.0282	1.001×10^{-5}	1.6689×10^{-6}
Site 3			0.6380	7.1526×10^{-7}	6.6757×10^{-6}
Site 4				2.3842×10^{-7}	2.3842×10^{-7}
Site 5					0.6168

From the Friedman test, we can reject the null hypothesis that each group has the same incidence of Dermo. From the signed rank tests, we see that the site pairings of (1,5), (1,6), (2,5), (2,6), (3,5), (3,6), (4,5), and (4,6) all have statistically significant incidences of Dermo at the (0.05/15)=0.0033 level of significance.

Results: From the results of this study, it appears that sites 5 and 6 (Oyster Bay, NY, and Cotuit, MA) have statistically significant differences in their proliferations of the Dermo disease. We can see this from as our graph as Cotuit and Oyster bay have mostly light colors their rows, while the other locations have a mixture of light and dark cells (lighter cells indicate a lower incidence of Dermo). Furthermore, our blocked ANOVA analysis indicated that there's a statistically significant difference between groups with a P-value of roughly zero. We also reached this conclusion from non-parametric analysis, as the Friedman test also has a P-value close to zero, which suggests a statistically significant difference between the groups. Furthermore, both the Bonferroni blocked analysis and the signed rank tests suggested that the site pairings of (1,5), (1,6), (2,5), (2,6), (3,5), (3,6), (4,5), and (4,6) all had statistically significant differences, which means that we have once again reached the same conclusions through both para-

metric and non-parametric analysis. These results also suggest that sites 5 and 6 have a different incidence of Dermo with a high degree of statistical significance. From this, we can conclude with a high degree of confidence that both Oyster Bay, NY and Cotuit, MA have a lower incidence of Dermo with a high degree of robustness. Judging from the map of the sample site, it appears that the Connecticut coastline may have some idiosyncratic factor that makes it less friendly to oysters.

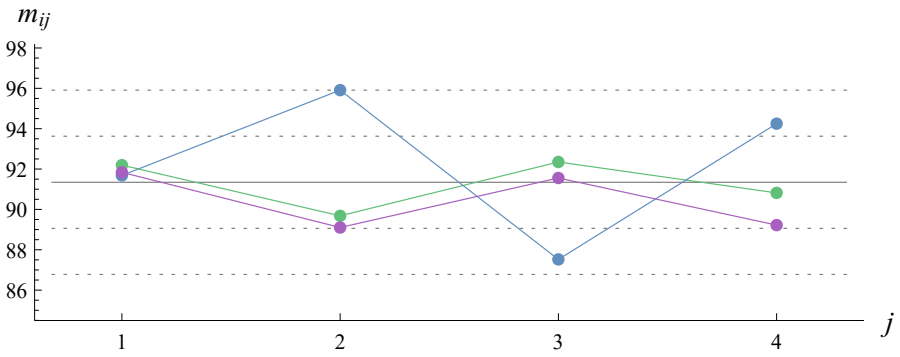
HW8 Problem 6: Industrial Experiment



	Zone 1	*Zone 2*	*Zone 3*	*Zone 4*
Externally Recycled	*92.194* *2.211*	*89.681* *2.646*	*92.34* *4.396*	*90.818* *3.643*
Internally Recycled	*91.684* *2.806*	*95.91* *3.124*	*87.521* *3.712*	*94.249* *4.699*
Virgin (i.e. unused)	*91.838* *3.558*	*89.103* *3.663*	*91.556* *4.033*	*89.221* *3.001*

The ratio of the maximum to the minimum standard deviation equals (4.699/2.211)=2.125. This indicates that it may be more appropriate to use non-parametric statistical tests for this data.

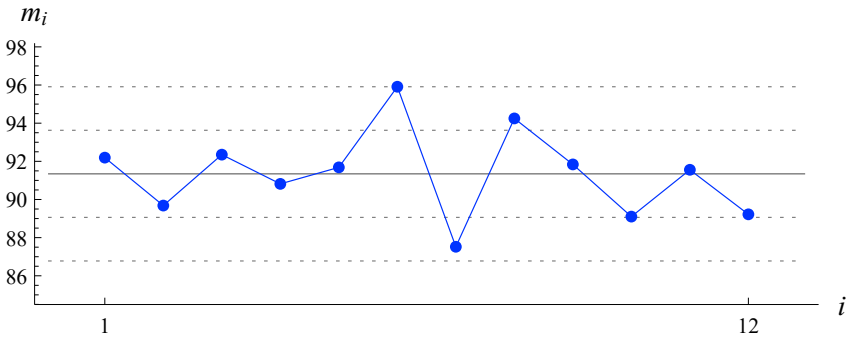
```
AnovaTwoWay[thickness]
```



	DF	SumOfSq	MeanSq	FRatio	PValue
Row	2	58.796	29.398	2.36205	0.1
Col	3	26.9843	8.99475	0.722703	0.541
Row x Col	6	383.098	63.8497	5.13014	0.000156
Error	84	1045.46	12.446		
Total	95	1514.34			

From the two way ANOVA table, we see that there aren't any statistically significant row and column effects at the 5% level. However, we do observe statistically significant interaction effects at the 5% level. In particular, it appears that there are significant row effects for the second, third, and fourth zones, but not for any other particular interactions.

```
samples = Flatten[thickness, 1];
AnovaOneWay[samples]
```



	DF	SumOfSq	MeanSq	FRatio	PValue
Group	11	468.878	42.6253	3.42482	0.000584
Error	84	1045.46	12.446		
Total	95	1514.34			

```
α = .05;
BonferroniBlocked[samples, α]
```

Number of Comparisons:	Rejection Region:	Sampling Distribution:
66	$\|T\| \geq 3.50784$	StudentT (df=77)

Significant mean differences:

$\hat{\mu}_2 - \hat{\mu}_6 = -6.23$ $\hat{\mu}_6 - \hat{\mu}_7 = 8.39$ $\hat{\mu}_6 - \hat{\mu}_{10} = 6.81$

$\hat{\mu}_6 - \hat{\mu}_{12} = 6.69$ $\hat{\mu}_7 - \hat{\mu}_8 = -6.73$

From our one-way analysis of the samples list, we can reject the null hypothesis that there is no difference between each sample at the 5% level, From our Bonferroni analysis, it appears that the sample pairings of (2,6), (6,7), (6,10), (6,12), and (7,8) have statistically significant differences in their means.

Results: From our statistical analyses, we can conclude that sixth, seventh, and eighth samples had some distinguishing characteristic from the other sample (these samples correspond to I2, I3, and I4 in the graph above and the second, third, and fourth columns in the second row of the table). From the results of the two-way ANOVA table, we see that while there were no statistically significant row or column effects, there were statistically significant interactions effects. Furthermore, the interaction effects were significant to the extent that we saw group effects when we performed a one way ANOVA on the samples. In addition, the Bonferroni analysis concluded statistically significant differences in means between the pairings of (2,6), (6,7), (6,10), (6,12), and (7,8). These results suggest that internally recycled wafers had significantly different thicknesses in zones 2, 3, and 4. This implies that silicon wafers recycled from in-house technicians had statistically significant differences from the target thickness when placed away from the oxygen inlet in the new furnace. This suggests that the technicians may either want to make sure that they heat the silicon wafers towards the oxygen inlet in the furnace or focus on using virgin and externally recycled wafers if wish to ensure that their wafers stay closer to 90 angstroms in thickness.