Response surface example

Stat 566
5-21-15

①

We wish to maximize the variable called "Yield".

There are 2 factors: "Time" and "Temperature".

Start with a $2^k$ design with 5 center points.

Result: Hyp. test for curvature $p = .814$
Fail to reject.

Gradient vector, found by fitting a $1^{st}$ order regression model, is $\begin{bmatrix} .775 \\ .325 \end{bmatrix}$, or $\begin{bmatrix} 1 \\ .4 \end{bmatrix}$

②

$1^{st}$ step: Let 1 "unit" in the time dir. be 5 secs
So add 5 secs to time, add 2° to temp.

| | | |
|---|---|---|
| Center: | 35 | 155 |
| | 40 | 157 |
| | ⋮ | |

Find peak around 85, 175.

So start over with that center point

At the new center point, we found significant curvature $p = .000$

$$\vec{b} = \begin{bmatrix} .995 \\ .5152 \end{bmatrix}$$

$$B = \begin{bmatrix} -1.376 & .125 \\ .125 & -1.001 \end{bmatrix}$$

$$\hat{x}_m = -\frac{1}{2} B^{-1} \vec{b} = \begin{bmatrix} .389 \\ .306 \end{bmatrix}$$

Time $85 + .389(5) = 86.945$

Temp $175 + .306(5) = 176.53$

$\boxed{\text{HW \#7} \quad 11.1,\ 11.8}$

Most of the quantities on the EVOP calculation sheet follow directly from the analysis of the $2^k$ factorial design. For example, the variance of any effect, such as $\frac{1}{2}(\bar{y}_3 + \bar{y}_5 - \bar{y}_2 - \bar{y}_4)$, is simply $\sigma^2/n$ where $\sigma^2$ is the variance of the observations ($y$). Thus, two standard deviation (corresponding to 95 percent) error limits on any effect would be $\pm 2\sigma/\sqrt{n}$. The variance of the change in mean is

$$V(\text{CIM}) = V\left[\frac{1}{5}(\bar{y}_2 + \bar{y}_3 + \bar{y}_4 + \bar{y}_5 - 4\bar{y}_1)\right]$$
$$= \frac{1}{25}(4\sigma_{\bar{y}}^2 + 16\sigma_{\bar{y}}^2) = \left(\frac{20}{25}\right)\frac{\sigma^2}{n}$$

Thus, two standard deviation error limits on the CIM are $\pm(2\sqrt{20/25})\sigma/\sqrt{n} = \pm 1.78\sigma/\sqrt{n}$.

The standard deviation $\sigma$ is estimated by the range method. Let $y_i(n)$ denote the observation at the $i$th design point in cycle $n$ and $\bar{y}_i(n)$ the corresponding average of $y_i(n)$ after $n$ cycles. The quantities in row (iv) of the EVOP calculation sheet are the differences $y_i(n) - \bar{y}_i(n - 1)$. The variance of these differences is

$$V[y_i(n) - \bar{y}_i(n-1)] \equiv \sigma_D^2 = \sigma^2\left[1 + \frac{1}{(n-1)}\right] = \sigma^2 \frac{n}{(n-1)}$$

The range of the differences, say $R_D$, is related to the estimate of the standard deviation of the differences by $\hat{\sigma}_D = R_D/d_2$. The factor $d_2$ depends on the number of observations used in computing $R_D$. Now $R_D/d_2 = \hat{\sigma}\sqrt{n/(n-1)}$, so

$$\hat{\sigma} = \sqrt{\frac{(n-1)}{n}}\frac{R_D}{d_2} = (f_{k,n})R_D \equiv s$$

can be used to estimate the standard deviation of the observations, where $k$ denotes the number of points used in the design. For a $2^2$ design with one center point, we have $k = 5$ and for a $2^3$ design with one center point, we have $k = 9$. Values of $f_{k,n}$ are given in Table 11.27.

■ **TABLE 11.27** Values of $f_{k,n}$

| $n =$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| $k = 5$ | 0.30 | 0.35 | 0.37 | 0.38 | 0.39 | 0.40 | 0.40 | 0.40 | 0.41 |
| 9 | 0.24 | 0.27 | 0.29 | 0.30 | 0.31 | 0.31 | 0.31 | 0.32 | 0.32 |
| 10 | 0.23 | 0.26 | 0.28 | 0.29 | 0.30 | 0.30 | 0.30 | 0.31 | 0.31 |

## 11.8 Problems

**11.1.** A chemical plant produces oxygen by liquifying air and separating it into its component gases by fractional distillation. The purity of the oxygen is a function of the main condenser temperature and the pressure ratio between the upper and lower columns. Current operating conditions are temperature ($\xi_1$) = −220°C and pressure ratio ($\xi_2$) = 1.2. Using the following data, find the path of steepest ascent:

| Temperature ($\xi_1$) | Pressure Ratio ($\xi_2$) | Purity |
|---|---|---|
| −225 | 1.1 | 82.8 |
| −225 | 1.3 | 83.5 |
| −215 | 1.1 | 84.7 |
| −215 | 1.3 | 85.0 |
| −220 | 1.2 | 84.1 |
| −220 | 1.2 | 84.5 |
| −220 | 1.2 | 83.9 |
| −220 | 1.2 | 84.3 |

**11.2.** An industrial engineer has developed a computer simulation model of a two-item inventory system. The decision variables are the order quantity and the reorder point for each item. The response to be minimized is total inventory cost. The simulation model is used to produce the data shown in Table P11.1. Identify the experimental design. Find the path of steepest descent.

■ **TABLE P11.1**
**The Inventory Experiment, Problem 11.2**

| Item 1 | | |
|---|---|---|
| **Order Quantity ($\xi_1$)** | **Reorder Point ($\xi_2$)** | |
| 100 | 25 | |
| 140 | 45 | |
| 140 | 25 | |
| 140 | 25 | |
| 100 | 45 | |
| 100 | 45 | |
| 100 | 25 | |
| 140 | 45 | |
| 120 | 35 | |
| 120 | 35 | |
| 120 | 35 | |
| **Item 2** | | |
| **Order Quantity ($\xi_3$)** | **Reorder Point ($\xi_4$)** | **Total Cost** |
| 250 | 40 | 625 |
| 250 | 40 | 670 |
| 300 | 40 | 663 |
| 250 | 80 | 654 |
| 300 | 40 | 648 |
| 250 | 80 | 634 |
| 300 | 80 | 692 |
| 300 | 80 | 686 |
| 275 | 60 | 680 |
| 275 | 60 | 674 |
| 275 | 60 | 681 |

**11.3.** Verify that the following design is a simplex. Fit the first-order model and find the path of steepest ascent.

| $x_1$ | $x_2$ | $x_3$ | $y$ |
|---|---|---|---|
| 0 | $\sqrt{2}$ | $-1$ | 18.5 |
| $-\sqrt{2}$ | 0 | 1 | 19.8 |
| 0 | $-\sqrt{2}$ | $-1$ | 17.4 |
| $\sqrt{2}$ | 0 | 1 | 22.5 |

**11.4.** For the first-order model

$$\hat{y} = 60 + 1.5x_1 - 0.8x_2 + 2.0x_3$$

find the path of steepest ascent. The variables are coded as $-1 \leq x_i \leq 1$.

**11.5.** The region of experimentation for three factors are time ($40 \leq T_1 \leq 80$ min), temperature ($200 \leq T_2 \leq 300$°C), and pressure ($20 \leq P \leq 50$ psig). A first-order model in coded variables has been fit to yield data from a $2^3$ design. The model is

$$\hat{y} = 30 + 5x_1 + 2.5x_2 + 3.5x_3$$

Is the point $T_1 = 85$, $T_2 = 325$, $P = 60$ on the path of steepest ascent?

**11.6.** The region of experimentation for two factors are temperature ($100 \leq T \leq 300$°F) and catalyst feed rate ($10 \leq C \leq 30$ lb/in). A first-order model in the usual $\pm 1$ coded variables has been fit to a molecular weight response, yielding the following model:

$$\hat{y} = 2000 + 125x_1 + 40x_2$$

(**a**) Find the path of steepest ascent.

(**b**) It is desired to move to a region where molecular weights are above 2500. On the basis of the information you have from experimentation in this region, about how many steps along the path of steepest ascent might be required to move to the region of interest?

**11.7.** The path of steepest ascent is usually computed assuming that the model is truly first order; that is, there is no interaction. However, even if there is interaction, steepest ascent ignoring the interaction still usually produces good results. To illustrate, suppose that we have fit the model

$$\hat{y} = 20 + 5x_1 - 8x_2 + 3x_1x_2$$

using coded variables ($-1 \leq x_i \leq +1$).

(**a**) Draw the path of steepest ascent that you would obtain if the interaction were ignored.

(**b**) Draw the path of steepest ascent that you would obtain with the interaction included in the model. Compare this with the path found in part (a).

**11.8.** The data shown in the Table P11.2 were collected in an experiment to optimize crystal growth as a function of three variables $x_1$, $x_2$, and $x_3$. Large values of $y$ (yield in grams) are desirable. Fit a second-order model and analyze the fitted surface. Under what set of conditions is maximum growth achieved?

■ **TABLE P11.2**
**The Crystal Growth Experiment, Problem 11.8**

| $x_1$ | $x_2$ | $x_3$ | $y$ |
|---|---|---|---|
| −1 | −1 | −1 | 66 |
| −1 | −1 | 1 | 70 |
| −1 | 1 | −1 | 78 |
| −1 | 1 | 1 | 60 |
| 1 | −1 | −1 | 80 |
| 1 | −1 | 1 | 70 |
| 1 | 1 | −1 | 100 |
| 1 | 1 | 1 | 75 |
| −1.682 | 0 | 0 | 100 |
| 1.682 | 0 | 0 | 80 |
| 0 | −1.682 | 0 | 68 |
| 0 | 1.682 | 0 | 63 |
| 0 | 0 | −1.682 | 65 |
| 0 | 0 | 1.682 | 82 |
| 0 | 0 | 0 | 113 |
| 0 | 0 | 0 | 100 |
| 0 | 0 | 0 | 118 |
| 0 | 0 | 0 | 88 |
| 0 | 0 | 0 | 100 |
| 0 | 0 | 0 | 85 |

**11.9.** The data in Table P11.3 were collected by a chemical engineer. The response $y$ is filtration time, $x_1$ is temperature, and $x_2$ is pressure. Fit a second-order model.

(**a**) What operating conditions would you recommend if the objective is to minimize the filtration time?

(**b**) What operating conditions would you recommend if the objective is to operate the process at a mean filtration rate very close to 46?

■ **TABLE P11.3**
**The Experiment for Problem 11.9**

| $x_1$ | $x_2$ | $y$ |
|---|---|---|
| −1 | −1 | 54 |
| −1 | 1 | 45 |
| 1 | −1 | 32 |
| 1 | 1 | 47 |
| −1.414 | 0 | 50 |
| 1.414 | 0 | 53 |
| 0 | −1.414 | 47 |
| 0 | 1.414 | 51 |
| 0 | 0 | 41 |
| 0 | 0 | 39 |
| 0 | 0 | 44 |
| 0 | 0 | 42 |
| 0 | 0 | 40 |

**11.10.** The hexagon design in Table P11.4 is used in an experiment that has the objective of fitting a second-order model:

■ **TABLE P11.4**
**A Hexagon Design**

| $x_1$ | $x_2$ | $y$ |
|---|---|---|
| 1 | 0 | 68 |
| 0.5 | $\sqrt{0.75}$ | 74 |
| −0.5 | $\sqrt{0.75}$ | 65 |
| −1 | 0 | 60 |
| −0.5 | $-\sqrt{0.75}$ | 63 |
| 0.5 | $-\sqrt{0.75}$ | 70 |
| 0 | 0 | 58 |
| 0 | 0 | 60 |
| 0 | 0 | 57 |
| 0 | 0 | 55 |
| 0 | 0 | 69 |

(**a**) Fit the second-order model.

(**b**) Perform the canonical analysis. What type of surface has been found?

(**c**) What operating conditions on $x_1$ and $x_2$ lead to the stationary point?

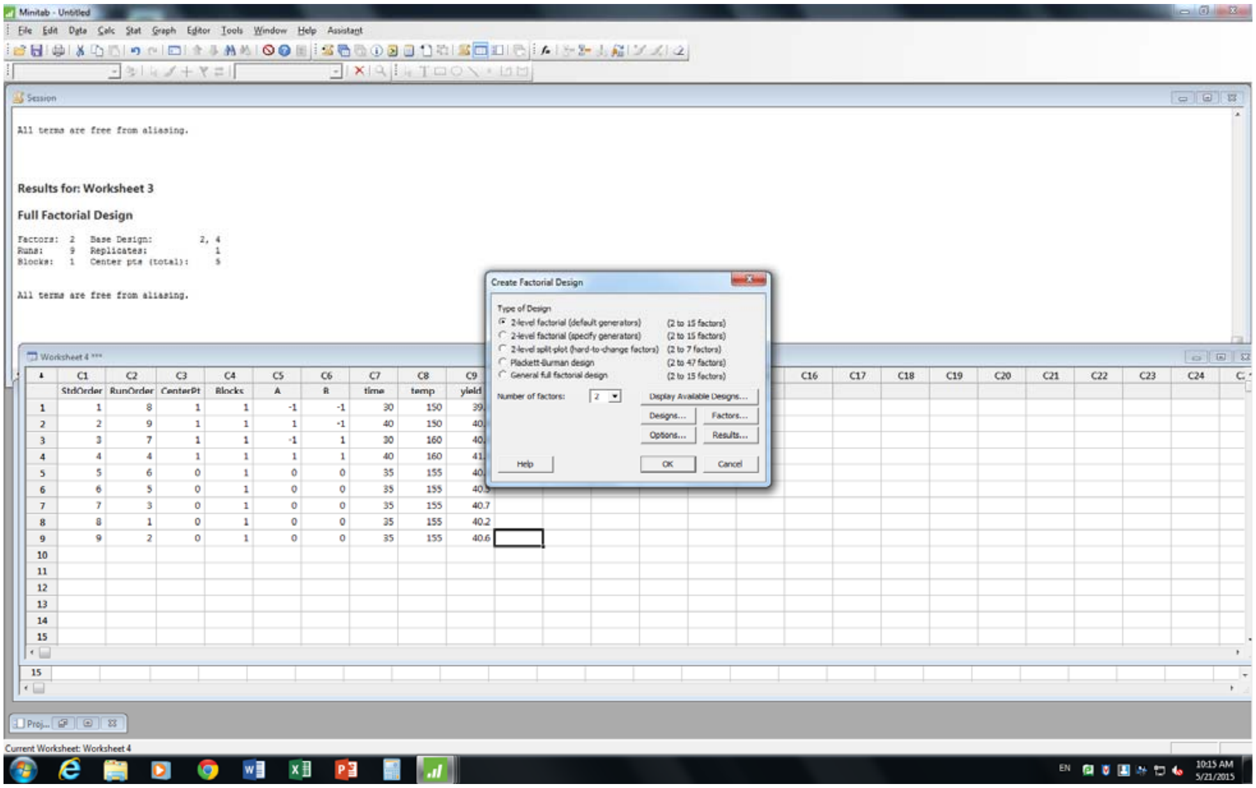
(**d**) Where would you run this process if the objective is to obtain a response that is as close to 65 as possible?

## STEP 1

**$2^2$ design with center points (5 center points)**

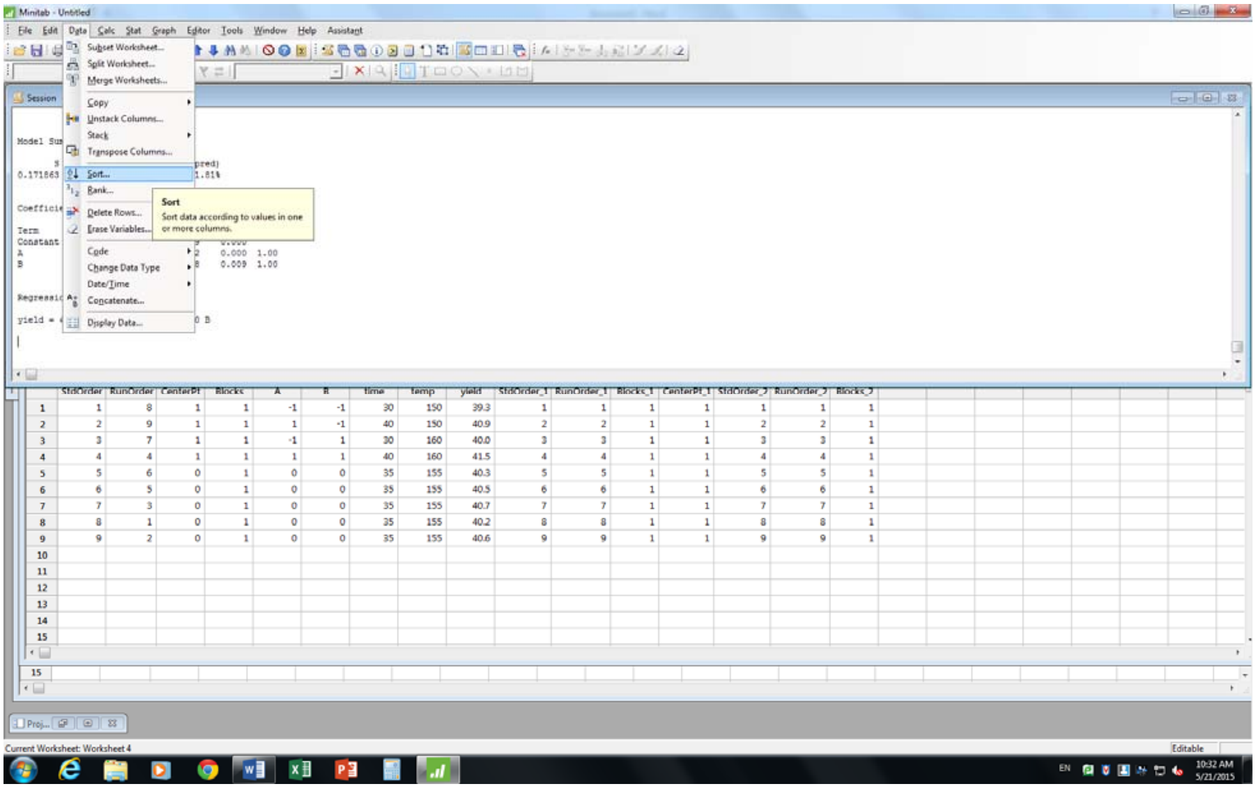
**Stat →DOE→Factorial→create factorial design**

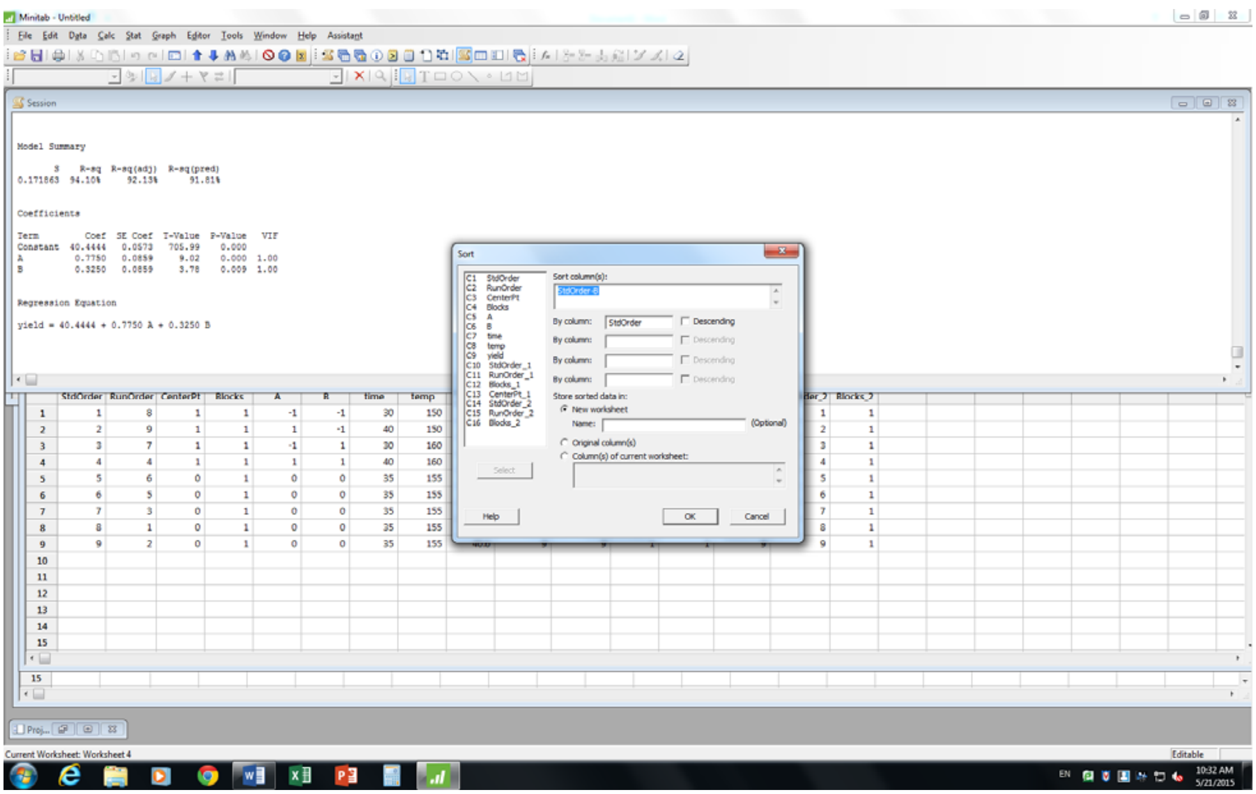
**(choose 2 factors)**



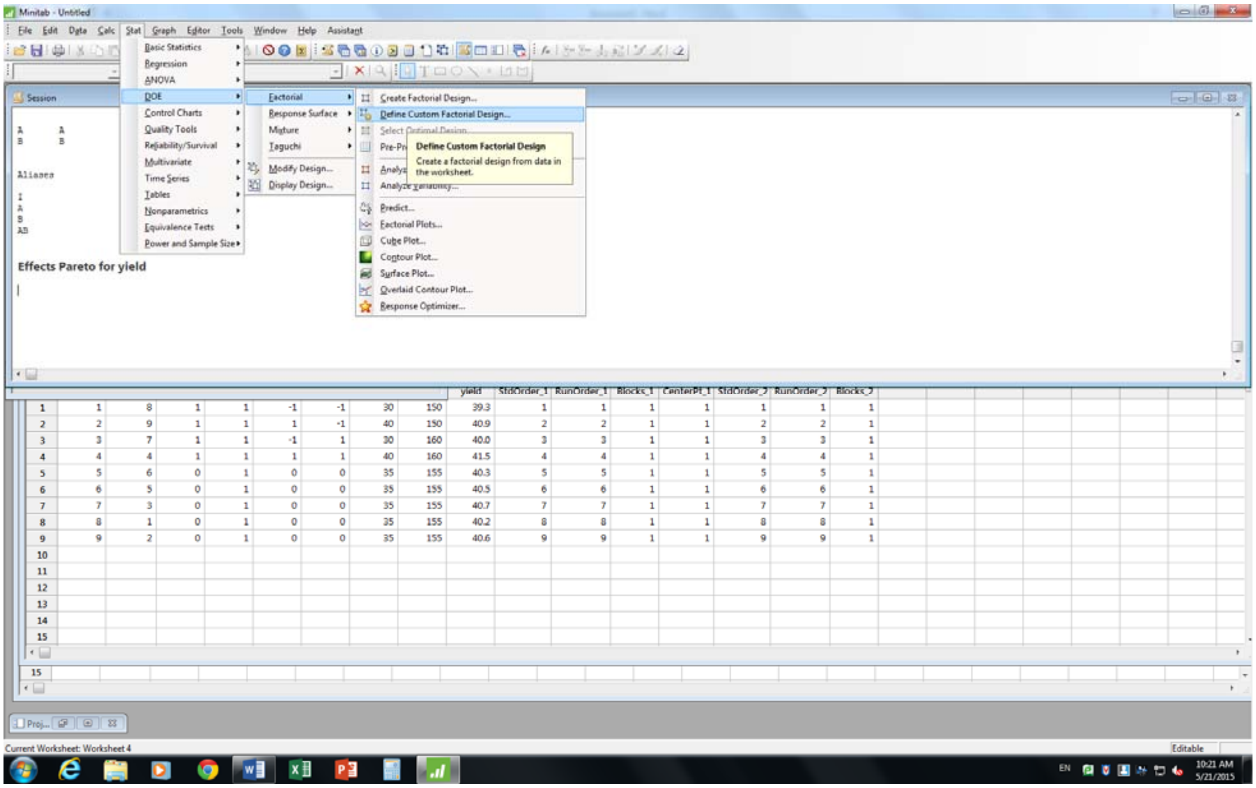
**Sort it by standard order**

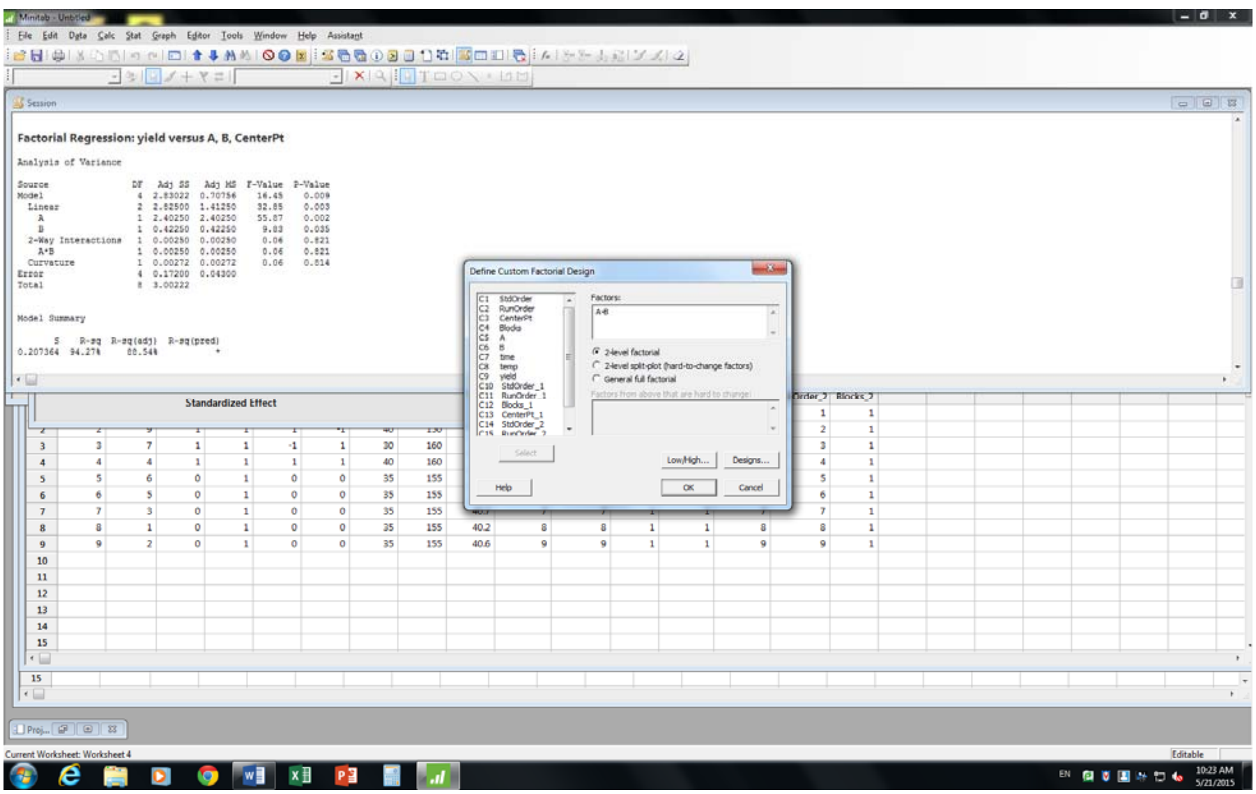
# Response surfaces



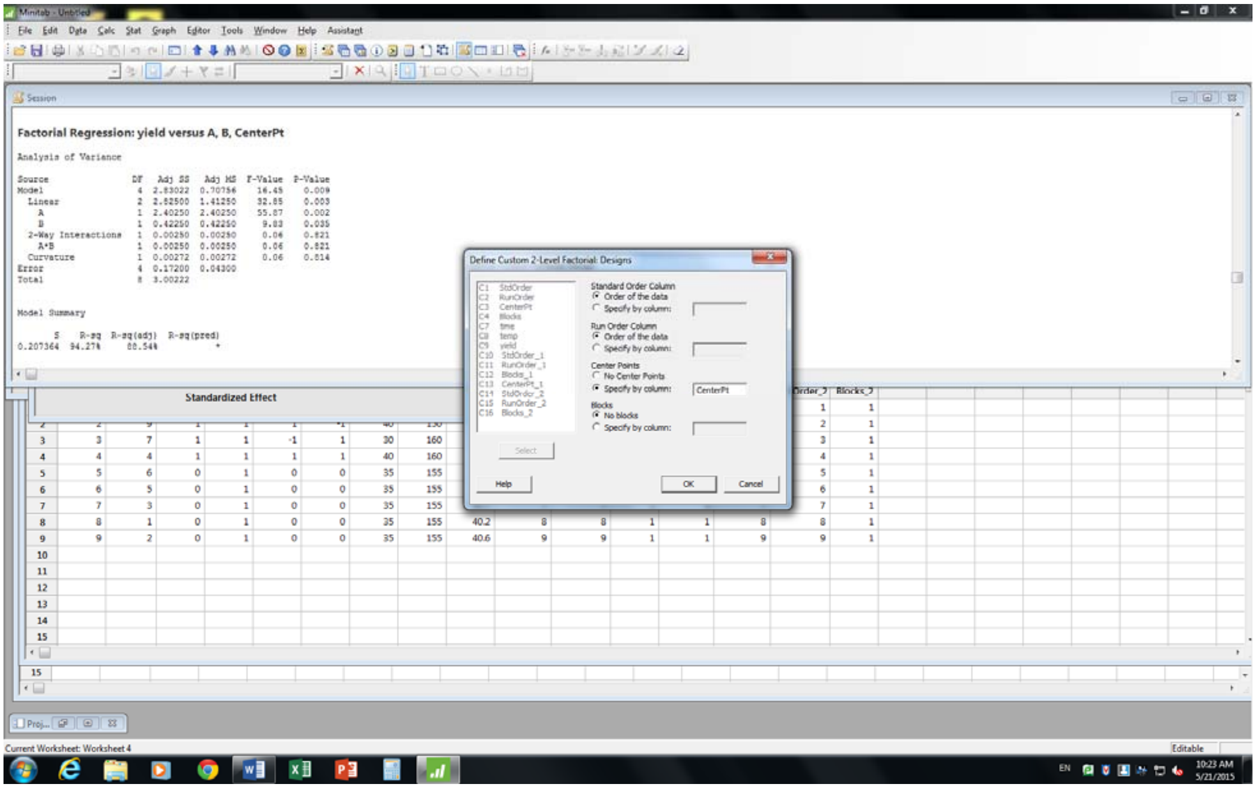


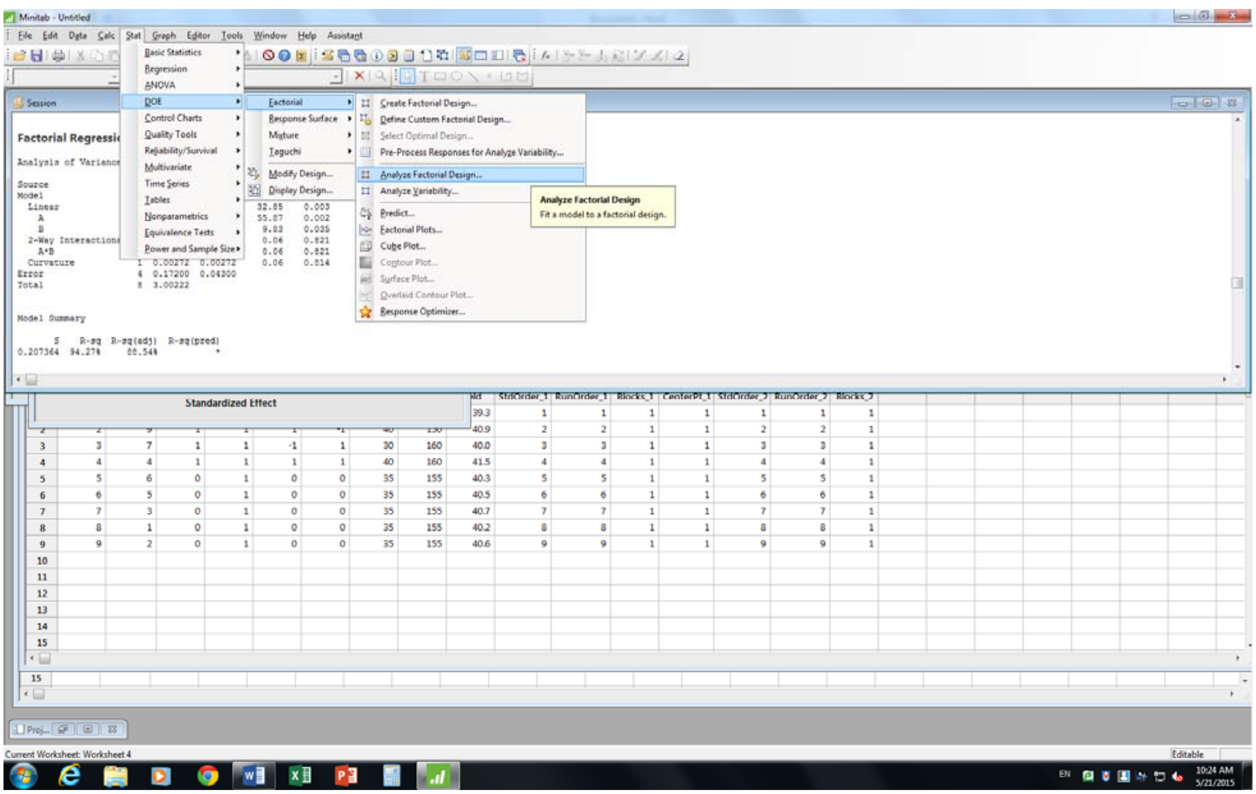
## Now analyze the data



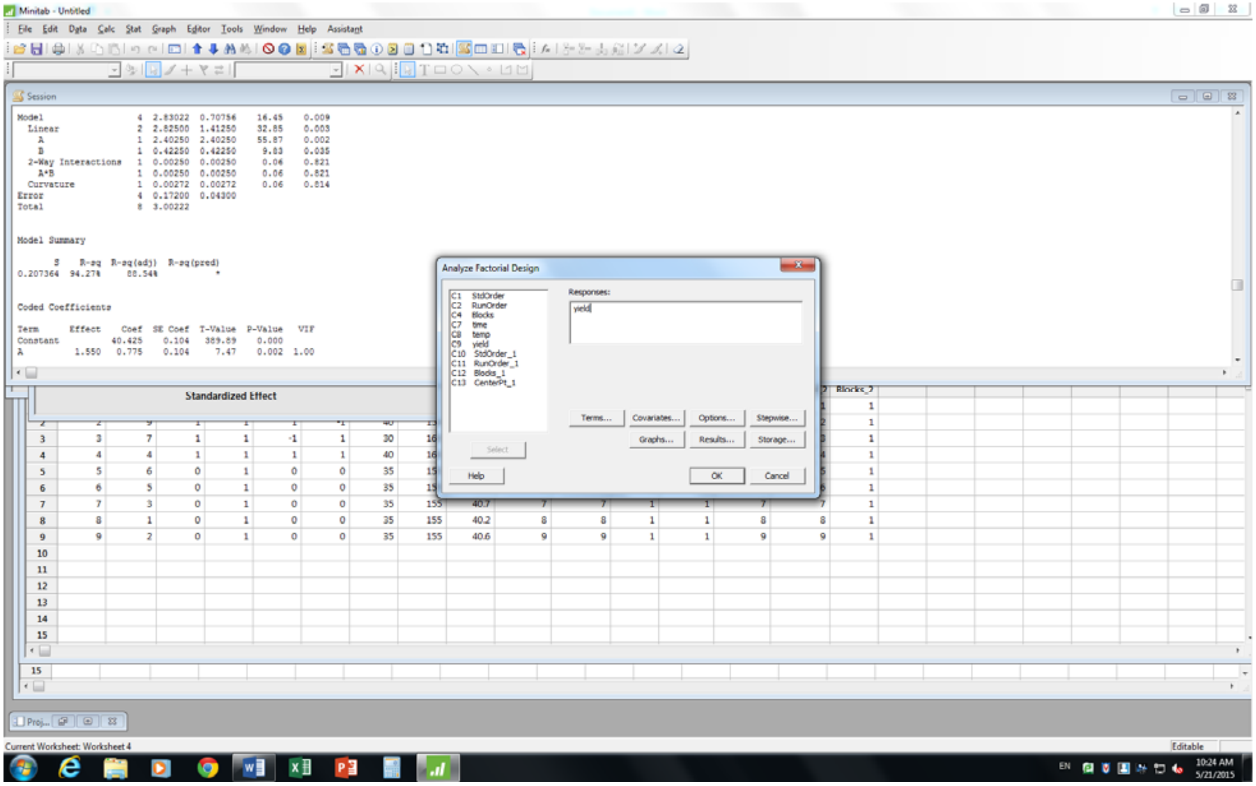


Response surfaces

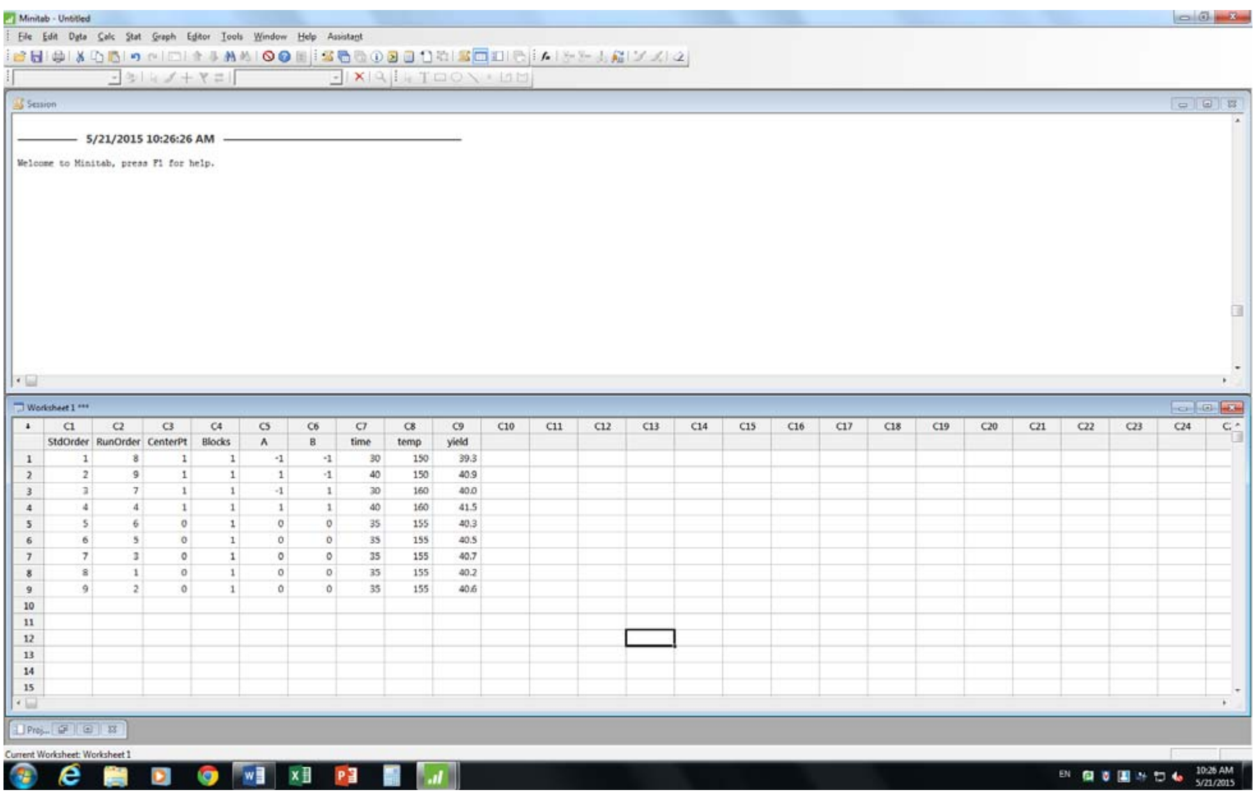




Response surfaces




# ***DATA***



# OUTPUT:

## Factorial Regression: yield versus A, B, CenterPt

Analysis of Variance

| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|---|---|---|---|---|---|
| Model | 4 | 2.83022 | 0.70756 | 16.45 | 0.009 |
| Linear | 2 | 2.82500 | 1.41250 | 32.85 | 0.003 |
| A | 1 | 2.40250 | 2.40250 | 55.87 | 0.002 |
| B | 1 | 0.42250 | 0.42250 | 9.83 | 0.035 |
| 2-Way Interactions | 1 | 0.00250 | 0.00250 | 0.06 | 0.821 |
| A*B | 1 | 0.00250 | 0.00250 | 0.06 | 0.821 |
| Curvature | 1 | 0.00272 | 0.00272 | 0.06 | 0.814 |
| Error | 4 | 0.17200 | 0.04300 | | |
| Total | 8 | 3.00222 | | | |

Model Summary

| S | R-sq | R-sq(adj) | R-sq(pred) |
|---|---|---|---|
| 0.207364 | 94.27% | 88.54% | * |

Coded Coefficients

| Term | Effect | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|---|
| Constant | | 40.425 | 0.104 | 389.89 | 0.000 | |
| A | 1.550 | 0.775 | 0.104 | 7.47 | 0.002 | 1.00 |
| B | 0.650 | 0.325 | 0.104 | 3.13 | 0.035 | 1.00 |
| A*B | -0.050 | -0.025 | 0.104 | -0.24 | 0.821 | 1.00 |
| Ct Pt | | 0.035 | 0.139 | 0.25 | 0.814 | 1.00 |

Regression Equation in Uncoded Units

yield = 40.425 + 0.775 A + 0.325 B - 0.025 A*B + 0.035 Ct Pt

Alias Structure

| Factor | Name |
|---|---|
| A | A |
| B | B |

Aliases

I
A
B
AB

Fits and Diagnostics for Unusual Observations

| Obs | yield | Fit | Resid | Std Resid | |
|---|---|---|---|---|---|
| 1 | 39.300 | 39.300 | 0.000 | * | X |
| 2 | 40.900 | 40.900 | 0.000 | * | X |

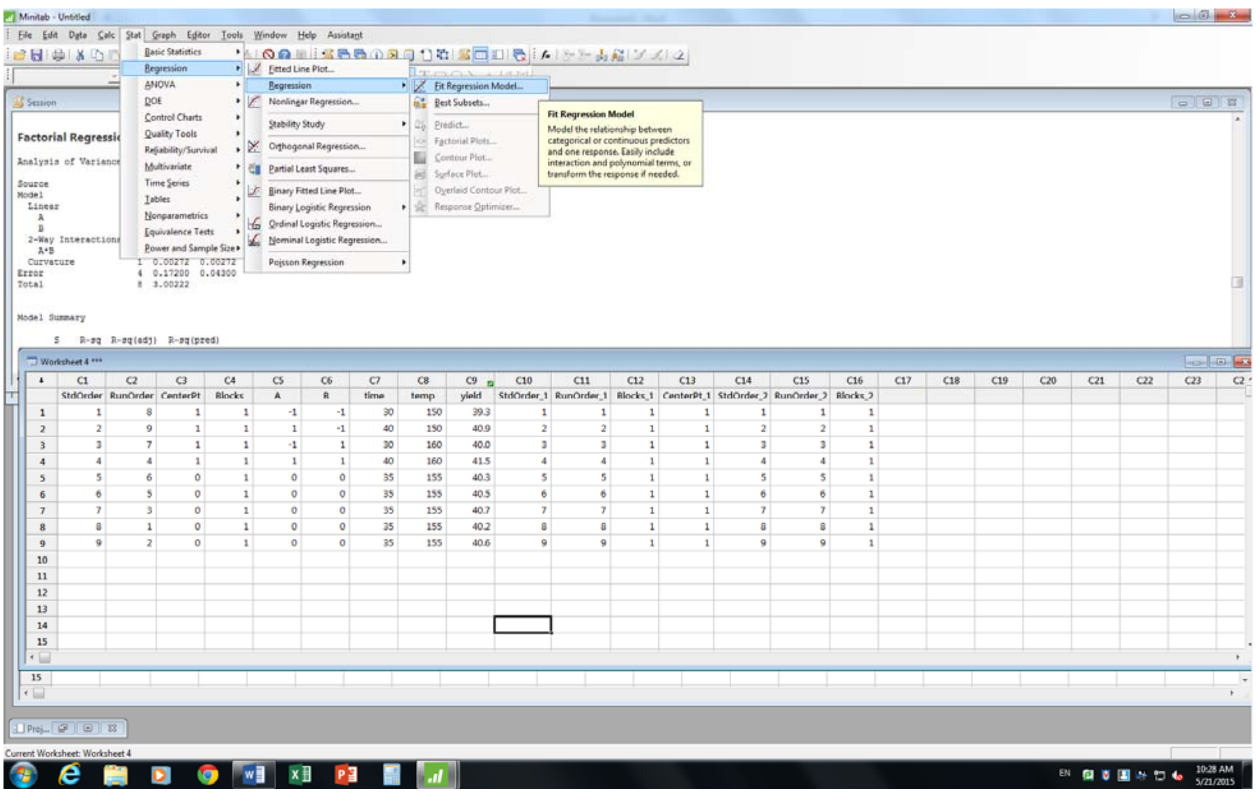
```
  3  40.000  40.000  0.000       *  X
  4  41.500  41.500  0.000       *  X

X  Unusual X
```

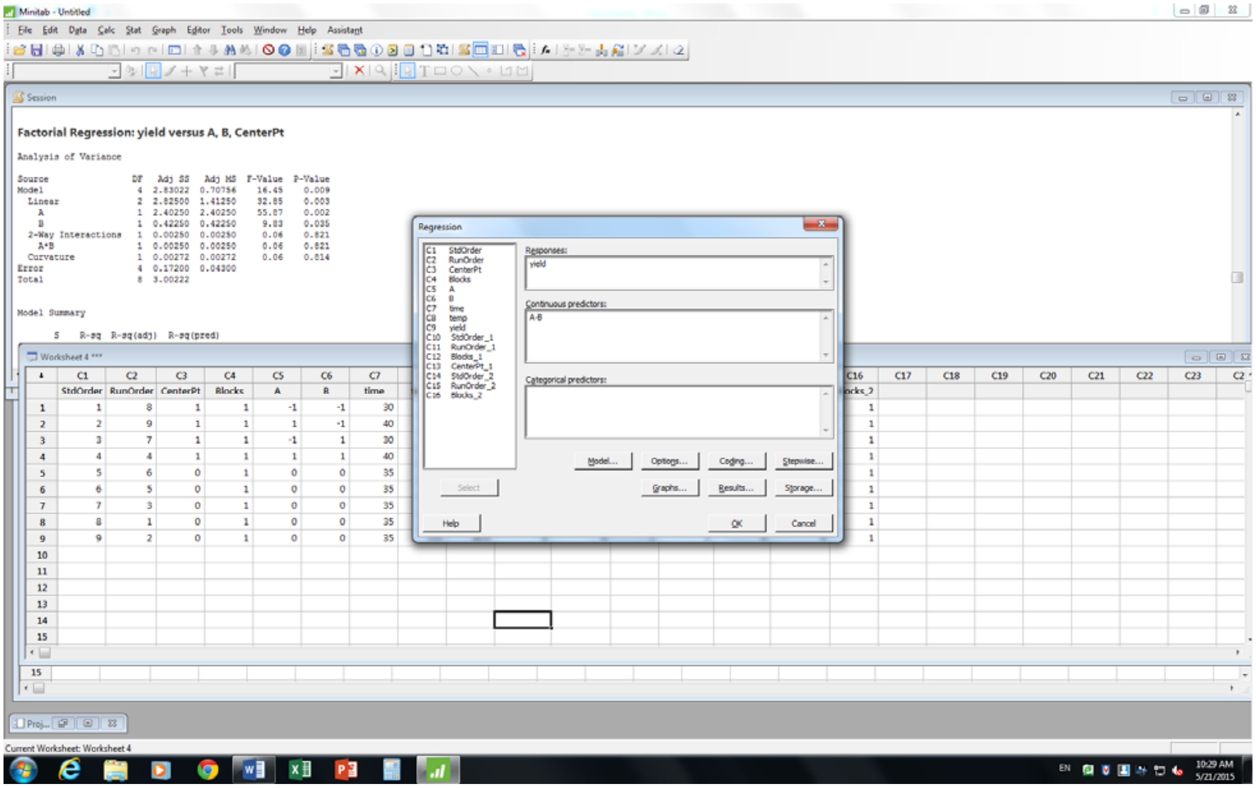
$$H_0: no\ curvature$$

$$H_1: there\ is\ curvature$$

## STEP 2:

**Fit a regression line**






# OUTPUT

## Regression Analysis: yield versus A, B

Analysis of Variance

| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|---|---|---|---|---|---|
| Regression | 2 | 2.82500 | 1.41250 | 47.82 | 0.000 |
| A | 1 | 2.40250 | 2.40250 | 81.34 | 0.000 |
| B | 1 | 0.42250 | 0.42250 | 14.30 | 0.009 |
| Error | 6 | 0.17722 | 0.02954 | | |
| Lack-of-Fit | 2 | 0.00522 | 0.00261 | 0.06 | 0.942 |
| Pure Error | 4 | 0.17200 | 0.04300 | | |
| Total | 8 | 3.00222 | | | |

Model Summary

| S | R-sq | R-sq(adj) | R-sq(pred) |
|---|---|---|---|
| 0.171863 | 94.10% | 92.13% | 91.81% |

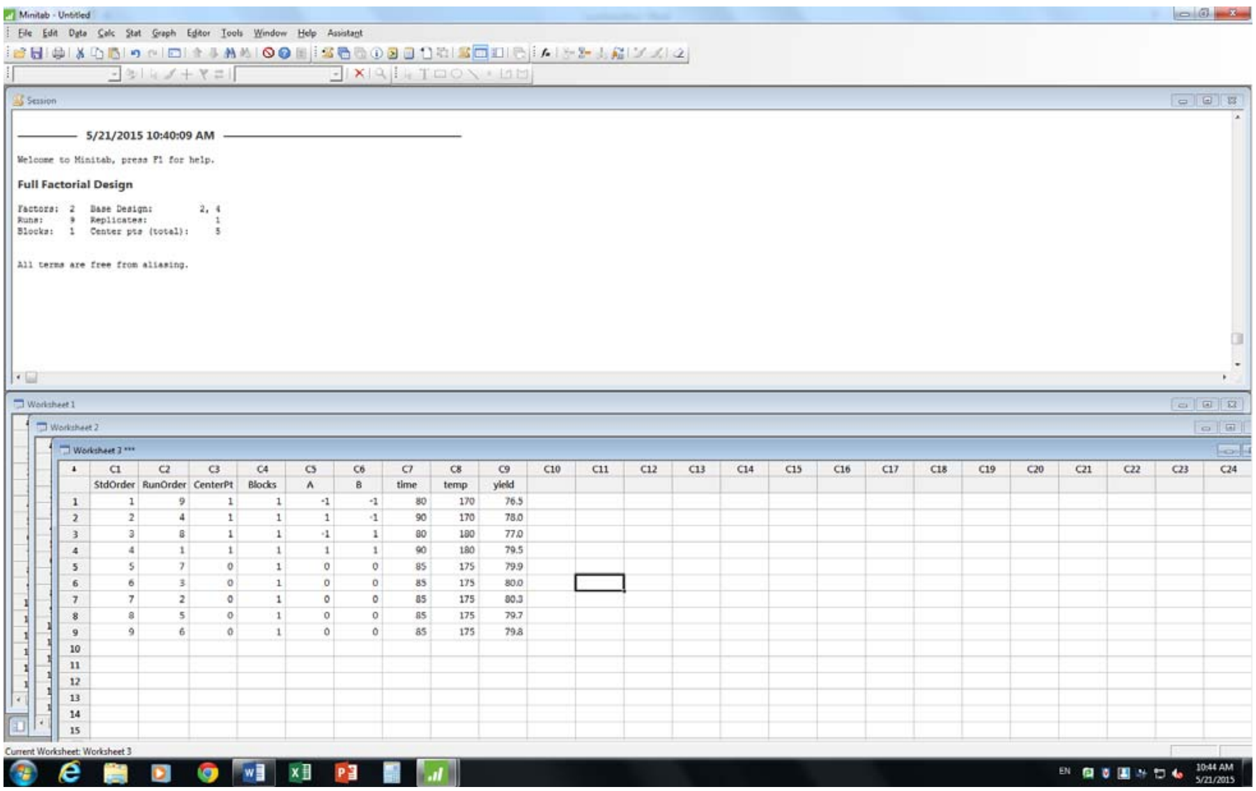
Coefficients

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|
| Constant | 40.4444 | 0.0573 | 705.99 | 0.000 | |
| A | 0.7750 | 0.0859 | 9.02 | 0.000 | 1.00 |

```
B           0.3250   0.0859     3.78    0.009  1.00
```

```
Regression Equation

yield = 40.4444 + 0.7750 A + 0.3250 B
```

## **NEW DATA**



# Step 1 output for new data

## Factorial Regression: yield versus A, B, CenterPt

Analysis of Variance

| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|---|---|---|---|---|---|
| Model | 4 | 15.9080 | 3.9770 | 75.04 | 0.001 |
| Linear | 2 | 5.0000 | 2.5000 | 47.17 | 0.002 |
| A | 1 | 4.0000 | 4.0000 | 75.47 | 0.001 |
| B | 1 | 1.0000 | 1.0000 | 18.87 | 0.012 |
| 2-Way Interactions | 1 | 0.2500 | 0.2500 | 4.72 | 0.096 |
| A*B | 1 | 0.2500 | 0.2500 | 4.72 | 0.096 |
| Curvature | 1 | 10.6580 | 10.6580 | 201.09 | 0.000 |
| Error | 4 | 0.2120 | 0.0530 | | |
| Total | 8 | 16.1200 | | | |

Model Summary

| S | R-sq | R-sq(adj) | R-sq(pred) |
|---|---|---|---|
| 0.230217 | 98.68% | 97.37% | * |

Coded Coefficients

| Term | Effect | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|---|
| Constant | | 77.750 | 0.115 | 675.45 | 0.000 | |
| A | 2.000 | 1.000 | 0.115 | 8.69 | 0.001 | 1.00 |
| B | 1.000 | 0.500 | 0.115 | 4.34 | 0.012 | 1.00 |
| A*B | 0.500 | 0.250 | 0.115 | 2.17 | 0.096 | 1.00 |
| Ct Pt | | 2.190 | 0.154 | 14.18 | 0.000 | 1.00 |

Regression Equation in Uncoded Units

yield = 77.750 + 1.000 A + 0.500 B + 0.250 A*B + 2.190 Ct Pt

Alias Structure

| Factor | Name |
|---|---|
| A | A |
| B | B |

Aliases

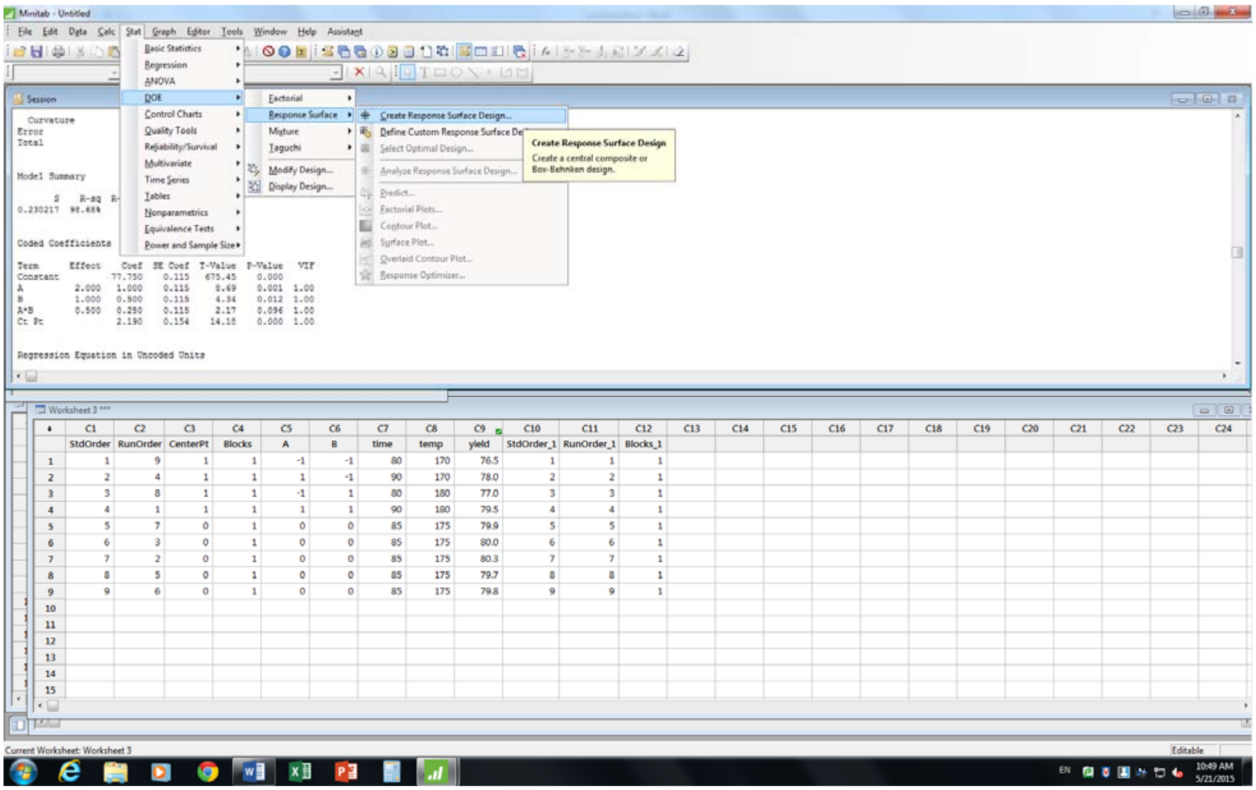
I
A
B
AB

Fits and Diagnostics for Unusual Observations

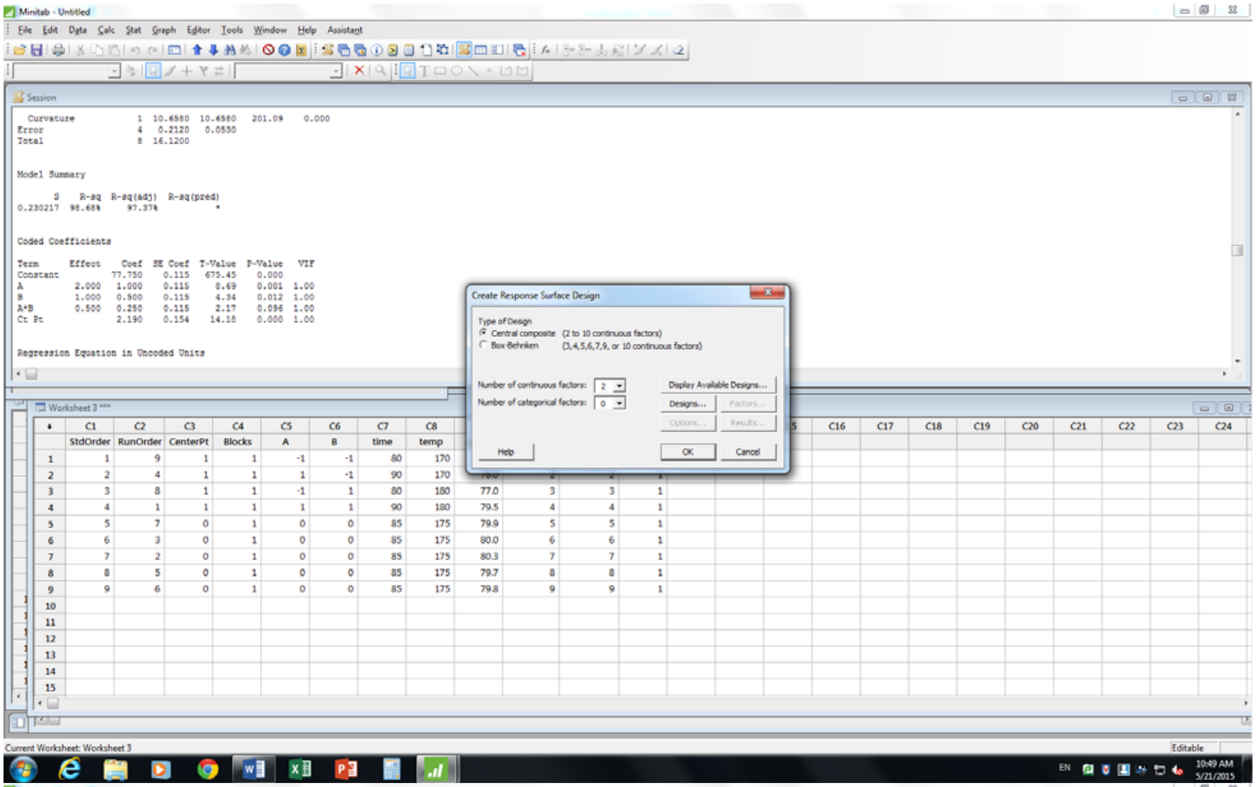
| Obs | yield | Fit | Resid | Std Resid | |
|---|---|---|---|---|---|
| 1 | 76.500 | 76.500 | 0.000 | * | X |
| 2 | 78.000 | 78.000 | 0.000 | * | X |

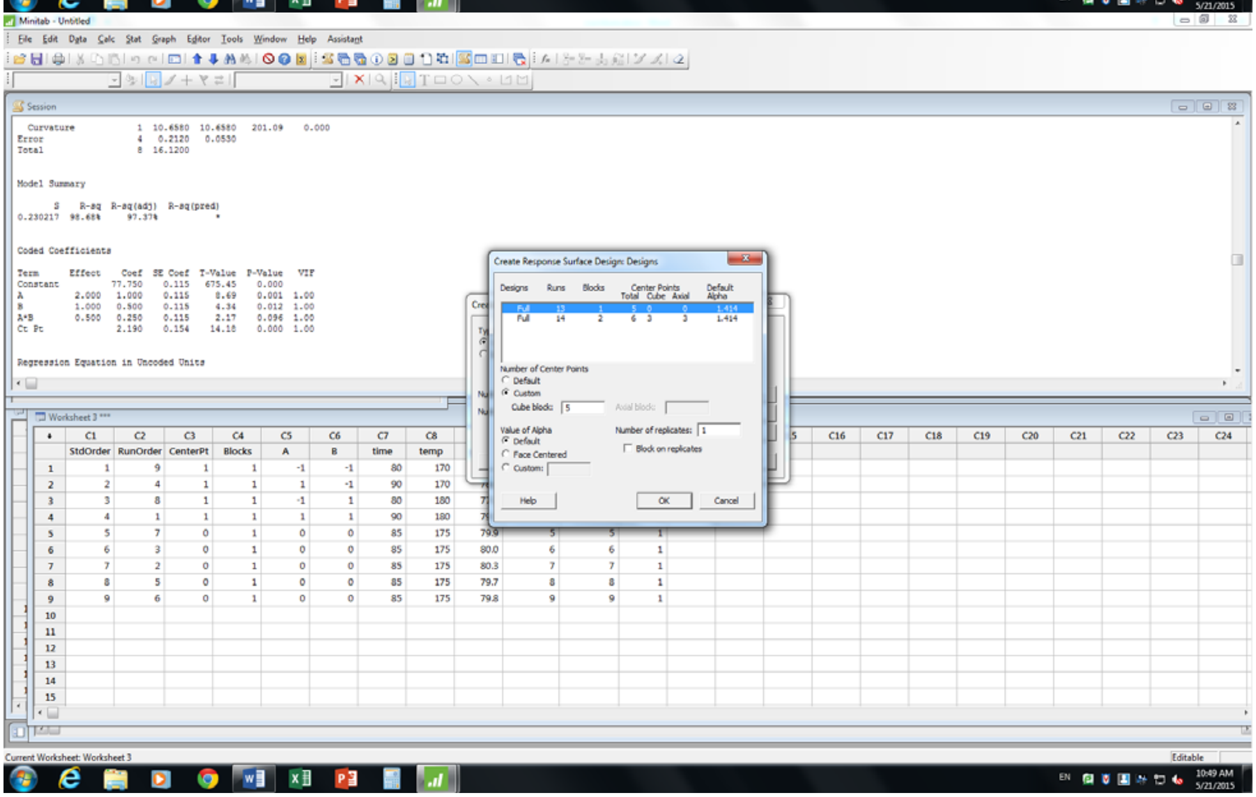
```
  3  77.000  77.000  0.000       *  X
  4  79.500  79.500  0.000       *  X

X  Unusual X
```

**After this, since p-value<0.05 we will do the following:**

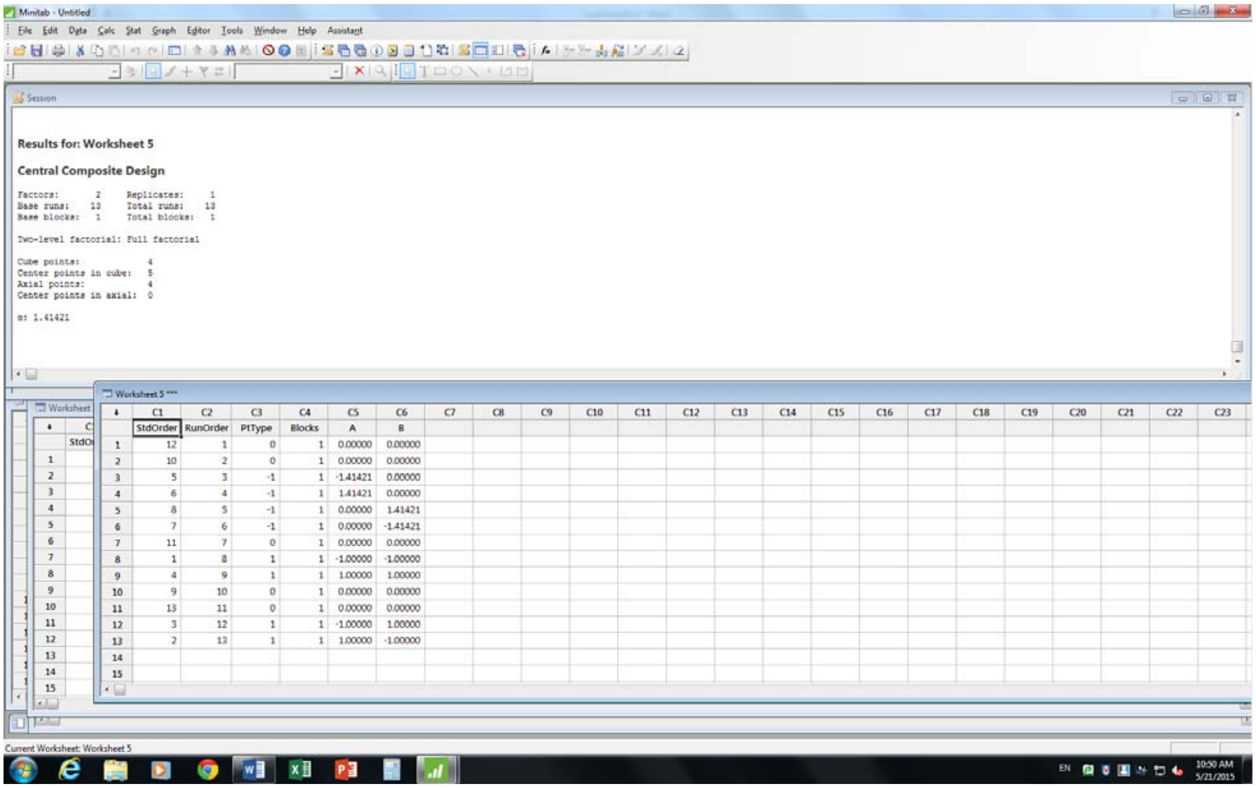


Response surfaces



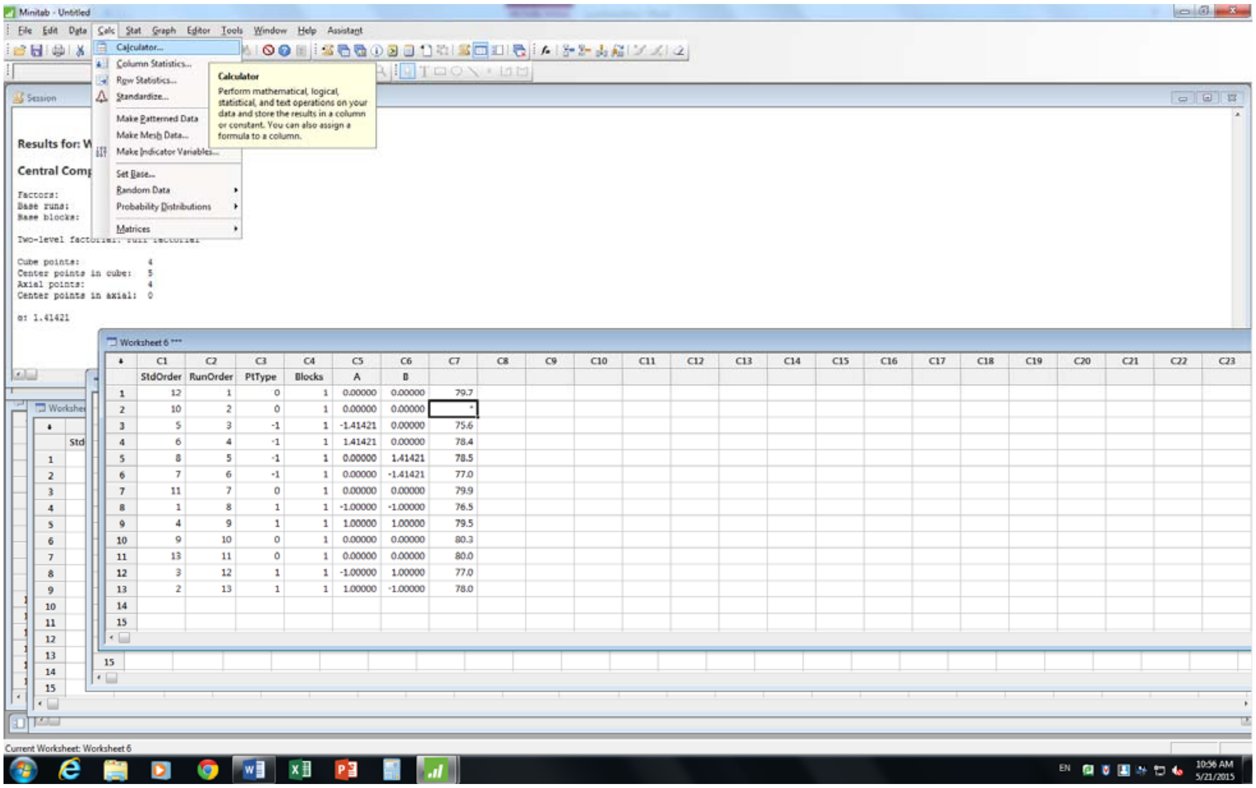



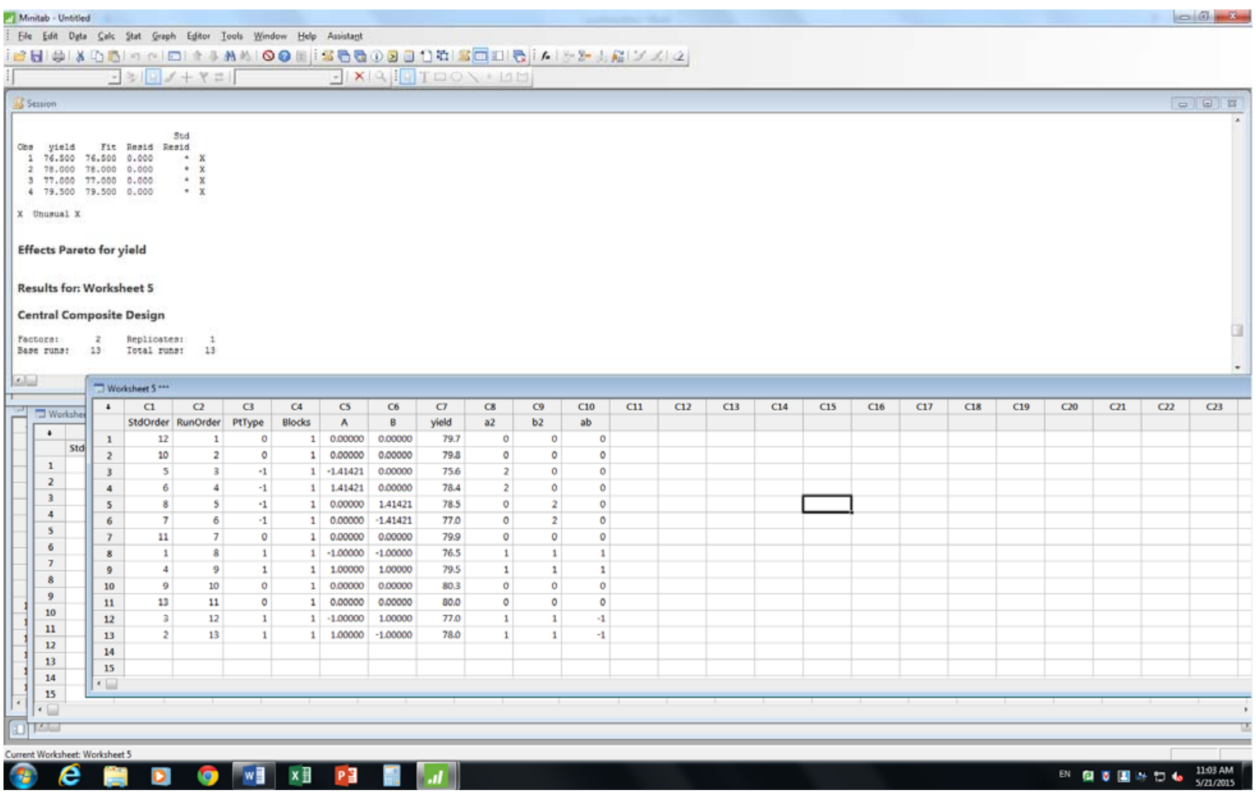
## It gives us




## Find A^2, B^2 and A*B

Response surfaces



**Now run regression with all the quadratic forms that was created in last step.**

## Regression Analysis: yield versus A, B, a2, b2, ab

Analysis of Variance

| Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|---|---|---|---|---|---|
| Regression | 5 | 28.2478 | 5.6496 | 79.85 | 0.000 |
| A | 1 | 7.9198 | 7.9198 | 111.93 | 0.000 |
| B | 1 | 2.1232 | 2.1232 | 30.01 | 0.001 |
| a2 | 1 | 13.1761 | 13.1761 | 186.22 | 0.000 |
| b2 | 1 | 6.9739 | 6.9739 | 98.56 | 0.000 |
| ab | 1 | 0.2500 | 0.2500 | 3.53 | 0.102 |
| Error | 7 | 0.4953 | 0.0708 | | |
| Lack-of-Fit | 3 | 0.2833 | 0.0944 | 1.78 | 0.290 |
| Pure Error | 4 | 0.2120 | 0.0530 | | |
| Total | 12 | 28.7431 | | | |

Model Summary

| S | R-sq | R-sq(adj) | R-sq(pred) |
|---|---|---|---|
| 0.266000 | 98.28% | 97.05% | 91.84% |

Coefficients

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|---|---|---|---|---|---|
| Constant | 79.940 | 0.119 | 672.00 | 0.000 | |
| A | 0.9950 | 0.0940 | 10.58 | 0.000 | 1.00 |
| B | 0.5152 | 0.0940 | 5.48 | 0.001 | 1.00 |
| a2 | -1.376 | 0.101 | -13.65 | 0.000 | 1.02 |
| b2 | -1.001 | 0.101 | -9.93 | 0.000 | 1.02 |
| ab | 0.250 | 0.133 | 1.88 | 0.102 | 1.00 |

Regression Equation

yield = 79.940 + 0.9950 A + 0.5152 B - 1.376 a2 - 1.001 b2 + 0.250 ab

## NEW STEP