**Regression analysis of total planned product and total manpower needed**

Dataset preparation

By applying count function onto each of the individual parts in the “process parameter” worksheet, we are able to get the number of planned production and amount of manpower for each part in a particular week. After that we are going to using sum function to sum up the planned production for all the 10 parts to come out with a total planned production. Similar process is applied to manpower to get the total manpower needed as shown below.

|  |  |  |
| --- | --- | --- |
| **Week** | **Total Planned Production (x-axis)** | **Total Manpower Needed (y-axis)** |
| 1 | 39 | 154 |
| 2 | 29 | 110 |
| 3 | 24 | 89 |
| 4 | 22 | 83 |
| 5 | 22 | 83 |
| 6 | 21 | 75 |
| 7 | 18 | 64 |
| 8 | 16 | 58 |
| 9 | 18 | 74 |
| 10 | 17 | 64 |
| 11 | 16 | 59 |
| 12 | 21 | 72 |
| 13 | 24 | 83 |
| 14 | 22 | 77 |
| 15 | 22 | 76 |
| 16 | 23 | 77 |
| 17 | 23 | 80 |
| 18 | 22 | 75 |
| 19 | 22 | 75 |
| 20 | 20 | 67 |
| 21 | 19 | 70 |
| 22 | 21 | 79 |
| 23 | 22 | 82 |
| 24 | 23 | 84 |
| 25 | 23 | 83 |
| 26 | 34 | 113 |
| 27 | 38 | 141 |
| 28 | 30 | 98 |
| 29 | 31 | 99 |
| 30 | 37 | 120 |
| 31 | 36 | 121 |
| 32 | 35 | 120 |
| 33 | 36 | 122 |
| 34 | 36 | 122 |
| 35 | 31 | 97 |
| 36 | 30 | 86 |
| 37 | 29 | 84 |
| 38 | 38 | 130 |
| 39 | 47 | 161 |
| 40 | 54 | 204 |
| 41 | 55 | 206 |
| 42 | 51 | 194 |
| 43 | 50 | 186 |
| 44 | 46 | 180 |
| 45 | 52 | 187 |
| 46 | 57 | 212 |
| 47 | 51 | 187 |
| 48 | 49 | 184 |
| 49 | 40 | 147 |
| 50 | 31 | 120 |
| 51 | 39 | 141 |
| 52 | 37 | 143 |

After we managed to compile the raw data into a more organized data as shown in the table above, we are going to perform simple linear regression analysis onto the table above by using total planned production as independent variables (x-axis) and using total manpower needed as dependent variable (y-axis) to uncover relationship between total planned production and total manpower needed. In other words, we are going to determine how is the effect of total planned production onto the total manpower needed.

Interpreting regression analysis

A picture containing text, screenshot, line, font

Description automatically generated

By using “analytics toolpak” tools in excel, we are able to get to the summary shown above. From the result, it provides lots of valuable information for us to make informed decisions. Multiple R with a value of 0.9851 with is very close to 1 indicate us that this model has a strong linear relationship. R square of 0.9704 indicates that 97.04% of variability in total manpower needed can be explained by the total planned production. As we are performing simple linear over here, adjusted R square would not play much role in our case. Standard error tells us that the average distance that the data points fall from regression line is 7.8478. As the manpower needed data is close to around 100 something, this small deviation would not affect much on the accuracy of data.

After that, we can conduct hypothesis testing to check whether the total planned production will affecting the total manpower needed. Hypothesis to be used in the experiment are shown below:

*H0: The total planned production does not affect the total manpower needed.*

*H1: The total planned production* *affects the total manpower needed*

The p-value of value 6.88e-40 in the highlighted cell gives us strong evidence to reject the null hypothesis, which means total planned production will affects the total manpower needed. At the same time, it also tells us that this is a signification model as the p value is less than 0.05. The confidence interval of 95% in this model leads us to have a range of [3.58, 3.95]. Since the coefficient of total planned production which is 3.77 fall under this range, it further prove the significance of this model.

Building model

By referring to the coefficient column in the regression analysis, we could form a regression equation as shown below:

Y = 3.77(x) - 5.99, where x = total planned production.

Alternatively, we can use scatter plot to get the similar equation.

Residual plot



By observing the standard residuals data above, we notice that all of the data above are fall within the range of [-3,3], thus we can conclude that there are no potential outliers exist in this model. The data in residual plot appears to be random and it shows no serious difference in the spread of the data for different X values. Therefore, it is managed to fulfil the assumption of linear regression.

Conclusion

This experiment shows that as total planned production increases, the total manpower needed increase as well. From the result, it shows that when total planned production increase by 1 unit, the total manpower will increase by approximately about 3.77, which round up to be 4 unit as manpower data cannot keep in continuous form. The significant of this model tells us that all the computed result are reliable, thus we can safely make informed decision based on this model.