





**PROJECT**

**On**

PERFORMANCE LAWN EQUIPMENT

**Dr. Arnav Kumar**



**Submitted by:**



**Submitted to:**



**Jhalak**

**Parvesh Thakur**

**Yashi Agarwal**

**PERFORMANCE LAWN EQUIPMENT**



**PLE, headquartered in St. Louis, Missouri, is a privately owned designer and producer of traditional lawn mowers used by homemakers. In the past 10 years, PLE has added another key product, a medium sized diesel power lawn tractor.**

**PLE provides most of the products to the dealerships, which, in turn, sell directly to end users. Also**

**PLE employs 1660 people worldwide.**

**We worked on some datasets provided by PLE for doing business analysis using MS Excel. All the**

**databases used and their analysis is provided on the next page for the reference of a reader.**



**For Reference**

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Sheet Name** | **Attachment** |
| 1 | [Customer Survey](#_SHEET_TITLE:_CUSTOMER) |  |
| 2 | [Blade Weight + Engine Production](#_SHEET_TITLE:_BLADE) |  |
| 3 | [Unit Shipping Cost](#_SHEET_TITLE:_UNIT) |  |
| 4 | [Unit Tractor Transmission Cost](#_SHEET_TITLE:_UNIT_1) |  |
| 5 | [Employee Retention](#_SHEET_TITLE:_EMPLOYEE) |  |
| 6 | [Employee Satisfaction](#_SHEET_TITLE:_EMPLOYEE_1) |  |
| 7 | [Tractor Unit Sales + Industry Tractor Total Sales](#_SHEET_TITLE:_TRACTORS) |  |
| 8 | [Mower Unit Sales + Tractor Unit Sales + Complaints](#_SHEET_TITLE:_MOWER) |  |
| 9 | [Defects After Delivery](#_SHEET_TITLE:_DEFECTS) |  |
| 10 | [Response Time](#_SHEET_TITLE:_RESPONSE) |  |
| 11 | [Time To Pay Suppliers](#_SHEET_TITLE:_TIME) |  |

# SHEET TITLE: CUSTOMER SURVEY



## Reason for choosing this sheet:

This sheet provides us with the ratings of customers of PLE tractors on scale of 1-5 on basis of Quality, Ease of use, Service and Price which may help us to find the relation between Quality, Ease of Use, Price, etc. and region wise comparison on the basis of same variable.

## Business Objective:



We want to test the relation between all attributes with help of correlation matrix and average Price with help of test so that we can take decisions accordingly on which attribute we should start working; and comparison between Quality of North America and South America so that improvement can be done.

## Hypothesis (Null and Alternate):

For Price:-

Null Hypothesis(Ho): Average Price is more than or equal to 3.9

Alternate Hypothesis(Ha): Average Price is less than 3.9

For Quality of tractors of North America(NA) and South America(SA):-

Null Hypothesis(Ho): Average Quality of tractors of NA is less than or equal to Average Quality of tractors of SA.

Alternate Hypothesis(Ha): Average Quality of tractors of NA is more than Average Quality of tractors of SA.

## ABOUT DATA: Descriptive and Visualization



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Quality | | Ease of Use | | Price | | Service | |
| Mean | 4.395 | Mean | 4.165 | Mean | 3.67 | Mean | 4.14 |
| Standard Error | 0.053939059 | Standard Error | 0.055264326 | Standard Error | 0.075391774 | Standard Error | 0.063022689 |
| Median | 5 | Median | 4 | Median | 4 | Median | 4 |
| Mode | 5 | Mode | 4 | Mode | 4 | Mode | 4 |
| Standard Deviation | 0.762813491 | Standard Deviation | 0.781555593 | Standard Deviation | 1.066200694 | Standard Deviation | 0.891275412 |
| Sample Variance | 0.581884422 | Sample Variance | 0.610829146 | Sample Variance | 1.13678392 | Sample Variance | 0.794371859 |
| Kurtosis | 3.19467826 | Kurtosis | 3.240404591 | Kurtosis | 0.017527215 | Kurtosis | 1.421738921 |
| Skewness | -1.492830118 | Skewness | -1.319144614 | Skewness | -0.740647733 | Skewness | -1.139503646 |
| Range | 4 | Range | 4 | Range | 4 | Range | 4 |
| Minimum | 1 | Minimum | 1 | Minimum | 1 | Minimum | 1 |
| Maximum | 5 | Maximum | 5 | Maximum | 5 | Maximum | 5 |
| Sum | 879 | Sum | 833 | Sum | 734 | Sum | 828 |
| Count | 200 | Count | 200 | Count | 200 | Count | 200 |



**About Data**



* We have 200 samples of Customer Satisfaction data of PLE’s Tractors measured on scale of 1-5 on the basis of Quality, Ease

of

Use, Service and Price. And 100 samples of Quality of North America and 50 samples of Quality of South America.

* Quality has average ratings of 4.395 which lies in the range of 4.3411 to 4.4489.
* Similarly, Ease of Use, Price and Service have average ratings of 4.165, 3.67 and 4.14 which lie in the range of 4.1098 to

4.2202; 3.5946 to 3.7454; 4.077 to 4.203 respectively.

* Here, Quality has the highest average whereas Price has least average.
* On an average the deviation of data from its mean is highest in case of Service and least in case of Price as presented

by values of Standard Deviation in Table before.

* As skewness is less than 0 in all(Quantity, Price, etc.) their data is negatively skewed.
* As kurtosis is more than 3 which implies positive kurtosis in Quality; and Ease of Use which means their distribution is

leptokurtic but it is opposite in case of Price and Service their distribution is platykurtic.

* Variables-> Quality; Price; Ease of Use; Service; NA: North America; SA: South America

## Correlation



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Quality | Ease of Use | Service | Price |
| Quality | 1 |  |  |  |
| Ease of Use | 0.134566701 | 1 |  |  |
| Service | 0.162163555 | -0.018900636 | 1 |  |
| Price | 0.099289992 | -0.054937193 | 0.048861637 | 1 |

To know the relationship between all variables with one another we are using Correlation Matrix.

* + There is significantly positive correlation in ratings Ease of use and Quality; and Price and Quality.
  + There is negative correlation in ratings Ease of use and Service; and Price and Ease of use.
  + There is positive correlation between price and quality; and price and service.
  + 1 in the table means the relation between one variable and that variable itself.
  + Blank cells means their correlation is already presented somewhere in the table.

## Statistical Techniques



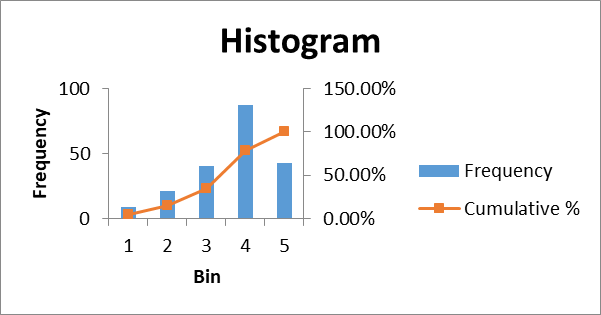
|  |  |
| --- | --- |
| Ho: µ(Price) >= | 3.9 |
| Ha: µ(Price) < | 3.9 |
| Significance Level(α) | 0.05 |
| Sample Mean (x)̅ | 3.67 |
| Sample Std Deviation.(s) | 1.066200694 |
| Sample Size(n) | 200 |
| Standard Error | 0.075391774 |
| tcal | -3.0507307 |
| tcritical(benchmark) | 1.652546746 |
| tcritical for lower tail test | -1.65254675 |
| p value | 0.001296957 |

Hypothesis For Price:- Ho: µ(Price) > = 3.9 Ha: µ(Price) < 3.9

We want to check whether average Customer Satisfaction on the basis of price

is more

than or equal to 3.9 or not.

We performed t-Test: one tailed (lowered tailed) to see the average satisfaction. As p value is less than 0.05, we can say that null hypothesis can be rejected.

Hence, the average Price ratings has reduced to less than 3.9 out of 5

The histogram is of Price is also provide below.

It shows that 35% people rated 3 or less than 3 for Price of Tractors. And 78.50% rated 4 or less than that showing very less percentage of people rating 5.

|  |  |  |
| --- | --- | --- |
| Bin | Frequency | Cumulative % |
| 1 | 9 | 4.50% |
| 2 | 21 | 15.00% |
| 3 | 40 | 35.00% |
| 4 | 87 | 78.50% |
| 5 | 43 | 100.00% |



## Statistical Techniques

Hypothesis For Regional Quality:-

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances | | |
|  | NA | SA |
| Mean | 4.6 | 4.28 |
| Variance | 0.424242424 | 0.613877551 |
| Observations | 100 | 50 |
| Hypothesized Mean  Difference | 0 |  |
| df | 84 |  |
| t Stat | 2.48969004 |  |
| P(T<=t) one-tail | 0.007378558 |  |
| t Critical one-tail | 1.663196679 |  |
| P(T<=t) two-tail | 0.014757116 |  |
| t Critical two-tail | 1.988609667 |  |

Ho: µ(Quality of tractors of NA) - µ(Quality of tractors of SA) <= 0

Ha: µ(Quality of tractors of NA) - µ(Quality of tractors of SA) > 0

We want to check whether average of ratings quality of North America is less than or equal to average ratings of quality of South America.

We performed t-Test: Two-Sample Assuming Unequal Variances to see the average quality. As t stat > t critical and p value is less than 0.05, we have enough evidence to reject null, so null hypothesis can be rejected.

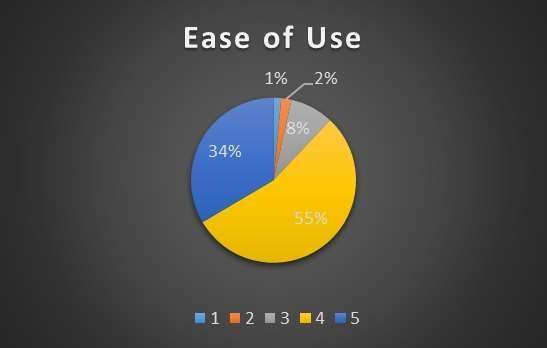
Therefore, North America has a significantly higher average customer

ratings on the basis of Quality in comparison to South America.



## Overall Average Customer Ratings







**Result and Discussion:**



As we can see from test for price that on an average price rating has been reduced to 3.9 or less than that, and also as shown by descriptive statistics price has least average in comparison to other factors. From charts also we can see that 4% customers rated 1 for price whereas only 1% customers rated 1 for others. Similar 10% customers rated 2 for price whereas only 1%, 2% or 4% customers rated 2 for others.

As we can see from test for regional quality that North America has a significantly higher average customer ratings on the basis of Quality in comparison to South America. Which might be because of the difference in sample sizes. As we had 100 samples from North America and just 50 samples from South America. So, this can be the reason.

## Suggestion:

PLE should try to make the product a bit cheaper as we can see from ratings that customers are not satisfied with the price.

PLE should collect more data from customers of South America and should again test the difference in Quality if again it

detects significance difference then PLE should investigate the reason for difference in Quality and take actions accordingly.

As the overall rating for PLE’s Tractors are pretty good as can be seen from the histogram but still there is 18.75% of sample

rating 3 or less for the product, so PLE can try to find the reason and take action accordingly.

PLE should do all these things to get a hold on it’s existing customers in the long run which will help the firm to grow as well

cause customers matter a lot for business.

# SHEET TITLE: BLADE WEIGHT AND ENGINE PROD.



## Reason for choosing this sheet:

### Blade Weight

The blade weight sheet provided us with 350 observations about the weight of the blades of mowers during production.

Here, we wanted to check the consistency of the manufacturing process for the blade weight

### Engine Production Time

The engine production sheet provided us with 50 observations about the production time of the

engines of tractors during production.

Here, we wanted to check the consistency of the time taken manufacturing process for the

engines.

## Business Objective:

We want to test for the consistency of the manufacturing process for the blade weight.

We want to test for the consistency of the time taken for manufacturing process of the engines for tractors.

**Blade Weight**

## Hypothesis (Null and Alternate):



The process of manufacturing blade weight is consistent.

Null Hypothesis: The average weight of the blades is more than or equal to 5.01.

Alternate Hypothesis: The average weight of the blades is less than 5.01.

**Engine Production Time**

The process of manufacturing of engines is consistent.

Null Hypothesis: The average time taken for the production of engine is less than or equal to 49.5

Alternate Hypothesis: The average time taken for the production of engine is more than 49.5

## About Data – Blade Weight

We have the sample size of 305 observations for the blade weight variable. The average blade weight is 4.99, which can be in the range from 4. 98 and 5.00. According to the empirical rule, 99.7% of the weight will lie between 4.66 and

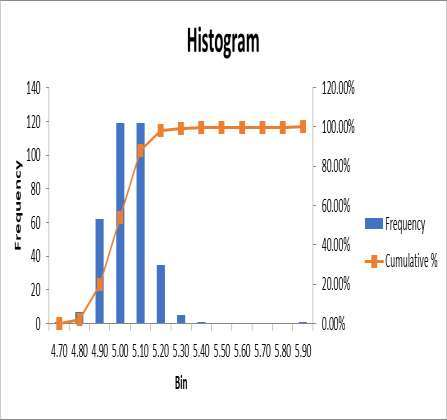
5.32, Which can be seen by histogram as well.

The minimum and maximum blade weight for mowers is 4.63 and 5.87

respectively.

We can clearly observe that the blade weight seems very consistent. We will

draw further inferences by conducting one sample t-test.

**Weight**

5.90

5.70

5.50

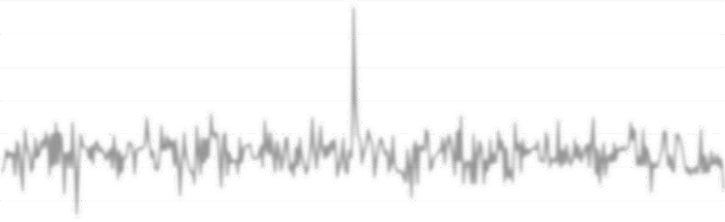
5.30

5.10

4.90

4.70

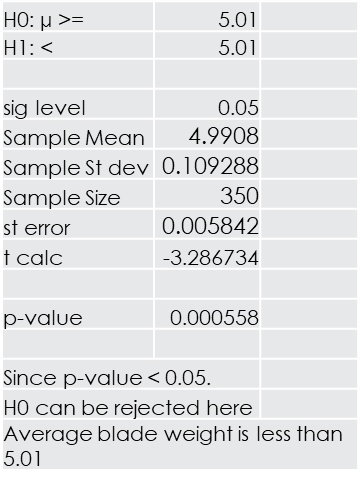
5.87



4.50

## Statistical Techniques – Blade Weight



Hypothesis to check the average blade weight of the mower in production.

H0: µ >= 5.01

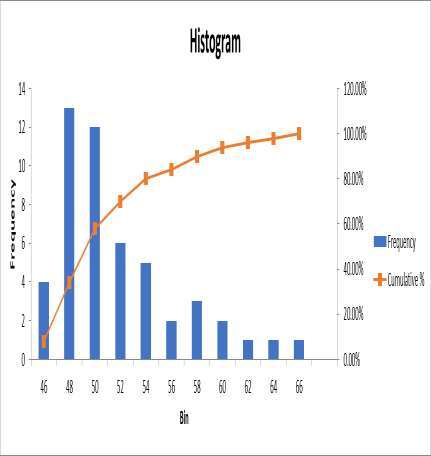
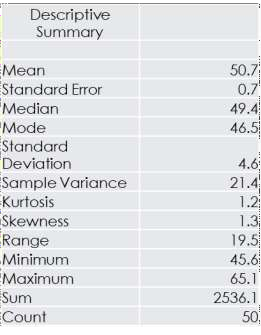
H1: µ < 5.01

We want to check that is the average blade weight for mowers is less 5.01 or not.

We performed one sample left tail t-test to analyze the average blade weight is less than 5.01 or not and as we get the p-value 0.00055 / 00.055%, means the null hypothesis that the average blade weight is more than or equal to 5.01 can be rejected and thus we can conclude that the average blade weight for mowers is less than 5.01.

The average blade weight for mowers is significantly lesser than 5.01.

## About Data – Engine Production



We have the sample size of 50 observations for the blade engine production

time (minutes) variable.

The average production time is 50.7 minutes, which can be in the range from 50. 1

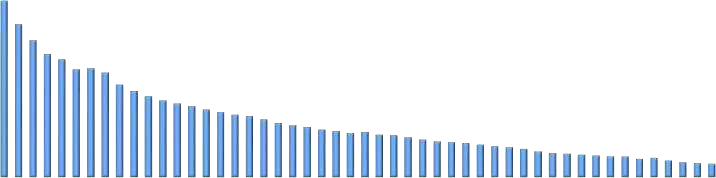
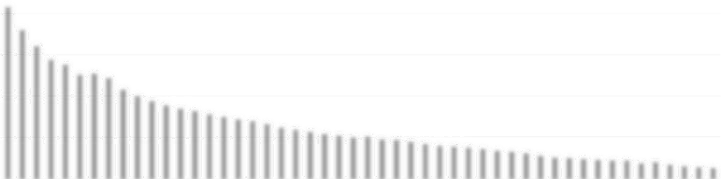
and 51.4 minutes.

According to the empirical rule, 95% of the production time will lie between 41.5

and 60 minutes, which can be seen by histogram as well.

The minimum and maximum production time for engines is 45.6 minutes and 65.1 minutes respectively.

We can clearly observe that the production time for engine have a decreasing trend and there is a high deviation in values as well. We will draw further inferences by conducting one sample t-test.



**Engine-Production Time**

64.0

59.0

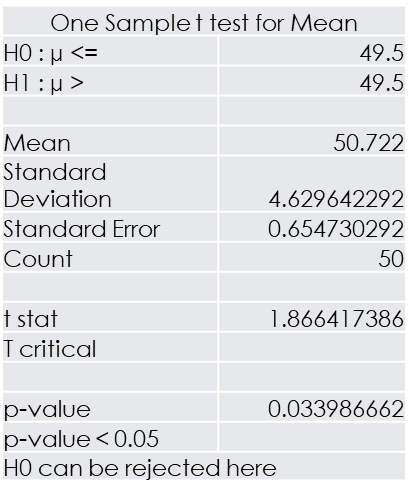
54.0

49.0

44.0



## Statistical Techniques - Engine

Hypothesis to check the average production time for the engine. H0: µ <= 49.05

H1: µ > 49.05

We want to check that is the average production time for engine

is less than 49.05 or not.

We performed one sample right tail t-test to analyze that the average production time is more than 49.05 or not and as we get the p-value 0.0339 / 03.39%, means the null hypothesis that the average production time is less than or equal to 49.05 can be rejected and thus we can conclude that the average production time for engines of tractors is more than 49.05.

The average production time for engine is significantly higher than 49.05.

**Blade Weight**

**Result and Discussion:**



From our inferential test we can conclude that our average blade weight is significantly lesser than 5.01 which means the

production process of blades is very much consistent which very a few outliers.

**Engine Production Time**

From our inferential test we can conclude that our average production time is significantly higher than 49.05 minutes which means the production process of engine is taking significantly higher time than our benchmark and also there is a decreasing trend in production, for the first 22 observations the production time is between 65 and 50 minutes and for the last 28 observation the production time is between 50 and 45 minutes. Also, there is a high variation in production time observations, nearly 4.6 minutes. Even 95% of the production time will lie between 41.5 and 60 minutes, which implies that there is 10 minutes of deviations from average production time.

**Suggestion:**

**Blade Weight**

PLE should continue with the current process, procedure for manufacturing of the blades for mowers.

**Engine Production Time**

There is not a good line of consistency in the production of engine for tractors, high variance, average production time is also higher than the benchmark value. Production is the base of any product created and PLE should consider these problems and try to make the process consistent and efficient as well, which will further help PLE to save a lot of time and money in production of these engines.

# SHEET TITLE: UNIT SHIPPING COST



## Reason for choosing this sheet:

This sheet provided us with 49 observations about per unit shipping cost for both products i.e., mowers and tractors, from 7 Plants to 7 customer locations.

Here, we wanted to compare the per unit shipping cost for both the products across all plants and customer locations.

## Business Objective:



We want to test that is there any difference between the per unit shipping costs for the two

products and also for any other pattern between plant and customer location, if any.

## Hypothesis (Null and Alternate):

There is difference between the unit shipping costs across both products

Null Hypothesis: There is no significant difference between average unit shipping costs for both products.

Alternate Hypothesis: The average unit shipping cost for Tractors is more than average unit shipping cost

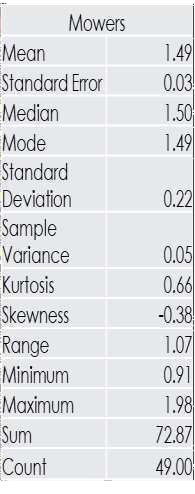
for Mowers.

For difference in variation in the average shipping costs for both the products

Null Hypothesis: There is no significant variation between average unit shipping costs for both products. Alternate Hypothesis: There is difference between average unit shipping costs for both products.

## About Data:



We have the sample size of 49 observations for 7 Plants of PLE and also, we

have 7 customer locations.

The average unit shipping cost for tractors is $1.94, which can be in the range from $1.89 and $1.98 and the average unit shipping cost for mowers is $1.49, which can be in the range from $1.46 and $1.52

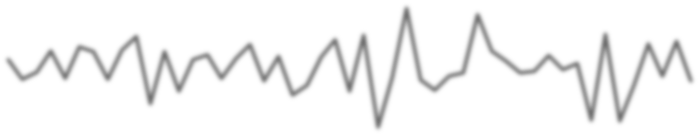
According to the empirical rule, 95% of the costs will lie between $1.31 and

$2.56 for tractors and between $1.05 and $1.92 for mowers.

The minimum and maximum shipping cost for tractors is $1.17 and $2.68

respectively and the same for mowers is $0.91 and $1.98 respectively.

We can clearly observe that the trend for both products is very much alike but the shipping costs for tractor seems higher than that for mowers. We will draw further inferences by conducting t-test and f-test.



**Shipping Costs**

$2.75

$2.25

$1.75

$1.25

$0.75

Mowers

Tra ctors

## About Data - Visualization



**To Toronto Location**

$2.50

$2.00

$1.50

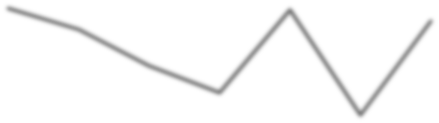
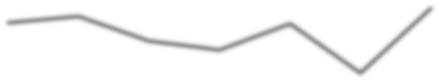
$1.00

$0.50

$0.00

Mowers Tra ctors

Tra ctors



**From Mumbai Plant**

2.30

2.10

1.90

1.70

1.50

1.30

1.10

Atlanta C ara cas London Melbourne Mexic o City Shanghai Toronto

Mower

Tractors



**To Caracas Location**

$2.50

$2.00

$1.50

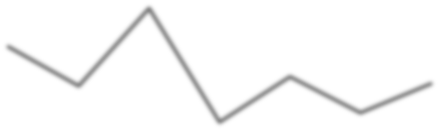
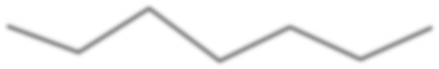
$1.00

$0.50

$0.00

Mowers Tra ctors

Tra ctors



**From Singapore Plant**

2.70

2.50

2.30

2.10

1.90

1.70

1.50

1.30

Atlanta Caracas London Melbourne Mexic o City Shanghai Toronto

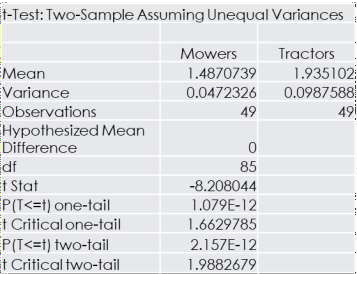
Mower

Tractors

**Statistical Techniques**



Hypothesis for no significant difference between the average shipping costs for tractors and mowers.

H0: µ(Tractors) - µ(Mowers) <=0

H1: µ(Tractors) - µ(Mowers) >0

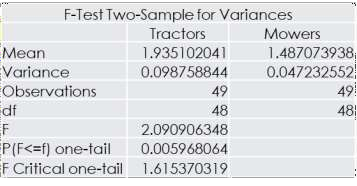
We want to check that is the average shipping cost for tractors is more than that of mowers.

We performed two sample one tail t-test to analyze the difference between average costs of both products and as we get the p-value 0.000 / 00.00%, means the null hypothesis that is there is no significant difference between average shipping cost for tractors and mowers can be rejected and thus we can conclude that the average shipping costs for mowers is less than the average shipping costs for tractors.

Tractors and Mowers have significant difference between their average shipping costs. The average shipping costs for tractors is significantly higher than that of mowers.



## Statistical Techniques

Variance of Tractor is $ 0.9875. Variance of Mower is $ 0.0472. H0: Var.(Tractors) = Var (Mowers) H1: Var.(Tractors) ≠Var (Mowers)

Hypothesis for no significant variance between the average shipping costs for tractors and mowers.

We performed two sample F-test to check whether there is any variance in the average shipping costs across two products for the organization, we get p-value 0.0059/ .59%, which is less than .025 or 2.5% and this means the null hypothesis that is there is no significant variation for average shipping costs between both products can be rejected, which means that there is significant variation between average costs across the two products.

This will confirm the fact that variance in average shipping cost for tractors is significantly higher from that for mowers. ( as visible from sample data)

## Result and Discussion:



We can see from the data visualization that average shipping costs for both the products have very similar trends but costs for tractor is always higher to the costs for mowers. Then, we concluded the same results from the inferential tests, the average shipping costs for tractors is higher than that for mowers, this may be because of the “Volume benefits” as Mowers is the highest selling product for PLE may be that is decreasing its shipping cost.

Also there is higher variances in the average shipping costs for tractor as compared to mower which means that the average shipping costs for tractors observe a much higher uncertainty in costs than mowers which again may lead to the problem of uncertain sales in the five regions.

## Suggestion:

PLE should focus on minimizing the unit shipping costs for tractors by taking necessary steps required or at least

try to lower the variance in the unit shipping cost for tractors.

# SHEET TITLE: UNIT TRACTOR



**TRANSMISSION COSTS**

## Reason for choosing this sheet:

This sheet provided us with the results of 30 samples each for the current process used to produce tractor transmission and the two new proposed processes, and therefore we wanted to check if the two new processes can help to decrease the Average cost for transmission.

## Business Objective:

We want to test that is the introduction of the two new processes can decrease the average

cost per transmission or not?

For Process A

## Hypothesis (Null and Alternate):



Null Hypothesis: The Average Transmission Cost does not change.

Alternate Hypothesis: The Average Transmission Cost decreased from previous average transmission cost.

For Process B

Null Hypothesis: The Average Transmission Cost does not change.

Alternate Hypothesis: The Average Transmission Cost decreased from previous average transmission cost.

For Current Process

Null Hypothesis: The average transmission cost is equals to or greater than $305

Alternate Hypothesis: The average transmission cost is less than $305

## About Data:



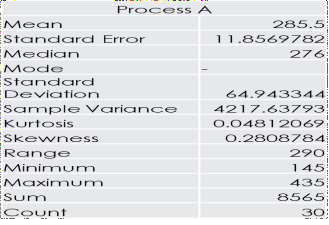
We have the sample size of 30 observations each for 3 processes which gives us unit transmission

costs for tractors. Further details and comparison of these 3 are provided below.

### Average total cost in different processes

In Current Process, Average cost is $289.6 within the range of $281.31 to $297.8

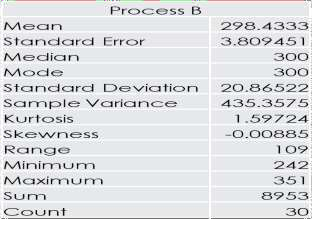
In Process A, Average cost is $285.5 within the range of $273.64 to $297.3 In Process B, Average cost is $298.4 within the range of $294.62 to $302.24



Here, Average Cost of Process B is Maximum with high range for variations in average cost. Average Cost of Process A is less than Current process, but it has a high degree of variation in its average cost.

### Variations in Costs

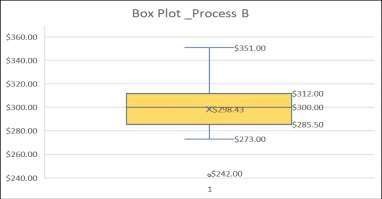
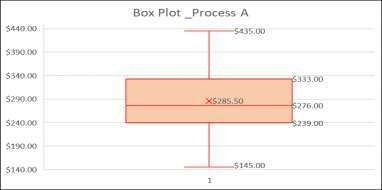
In Current Process, the cost can vary within the range of $244.20 to $334.9



In Process A, the cost can vary within the range of $220.55 to $350.4 In Process B, the cost can vary within the range of $277.56 to $319.29

Here, Process B has the least deviation in the overall Costs observations but have overall high average cost and thus high total cost. Process A has the most deviation in the overall Costs observations but also include the economical option as well as the most expensive one.

## About Data - Visualization

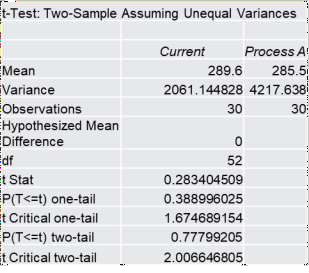


**Statistical Techniques**



Hypothesis for no significant difference between the average

transmission cost for the current process and the proposed process A.



H0: µ (Current Process) <= µ (Process A) H1: µ (Current Process) > µ (Process A)

We want to check that is the average transmission cost for the new

process is less than the current process.

We performed two sample one tail t-test to analyze the difference between average costs of both and as we get the p-value 0.3889 / 38.89%, means the null hypothesis that is there is no significant difference between average transmission cost for the current and the proposed process A can not be rejected and thus the proposed process is no better in comparison to the current one.

Our Average Transmission Cost has no significant difference to the

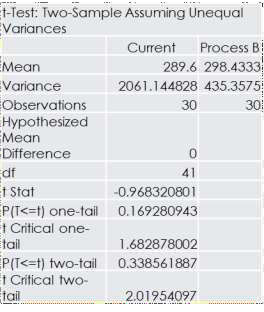
Average transmission cost from Process A.

## Statistical Techniques



Hypothesis for no significant difference between the average transmission

cost for the current process and the proposed process B.



H0: µ (Current Process) <= µ (Process A) H1: µ (Current Process) > µ (Process A)

We want to check that is the average transmission cost for the new process

B is less than the current process.

We performed two sample one tail t-test to analyze the difference between average costs of both and as we get the p-value 0.1692 / 16.92%, means the null hypothesis that there is no significant difference between average transmission cost for the current and the proposed process B can not be rejected and thus the proposed process is no better in comparison to the current one.

Our Average Transmission Cost has no significant difference to the Average

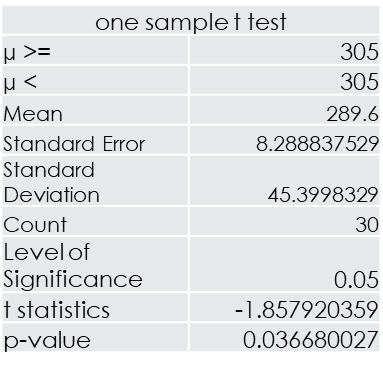
transmission cost from Process B.

## Statistical Techniques



Hypothesis for taking the inference about the average transmission

cost in current process for population data.



H0: the average transmission cost is(µ) >= $305 H1: the average transmission cost is(µ) < $305

We want to check that is the average transmission cost for current

process is less than $305.

We performed one sample left tail t-test to analyze the average cost of current process and as we get the p-value 0.0366 / 03.66%, means the null hypothesis that the average transmission cost for current process is less than $305 can be rejected and thus the current process have average transmission cost of less than $305.

Our Average Transmission Cost for current process is less than $305.

## Result and Discussion:



As we can concluded from the inferential tests, the new proposed processes A and B are

not helping PLE to decrease the cost of transmission for Tractors.

Average Transmission Cost from current process is less than $305, but this amount is higher than the average cost derived from the sample data, i.e., $281.31 to $297.8.

We can collect more samples and do tests again for more accurate and better insights

about the average cost of current process.

## Suggestion:

PLE should continue with the current process for transmissions.

The Current process have a large deviation in average transmission cost, so we recommend for further analysis by taking a larger sample size and if we get significant higher average transmission cost then PLE should try to brought it down by trying to making the process more efficient.

# SHEET TITLE: EMPLOYEE RETENTION



## Reason for choosing this sheet:

This sheet has a plenty of variables with a combination of scale and categorical Variables and we wanted to dig this database for more insights about different- different objectives like presence of younger generation of employees, factors deciding retention years, etc.

## Business Objective:

1. We wanted to perform analysis for the following objectives.
2. Average age of the employees, their average retention time span. (for

population)

1. Relation between Average age across Gender and Local or Not Local.
2. A Regression model for Retention time span.
3. Relation between Average age across Gender.



## Hypothesis (Null and Alternate):

* + Local or not and age have no significant relation with retention years via

Regression Modeling.

* + Average Age of Employees is more than 24.5 years.
  + There is no significant relation between Local or Not Local with Gender.
  + There is no significant difference between Average Age across Gender.

## About Data:

This sheet comprises many type of variables like Gender, Age, Years of Education etc. which are further providing information about Retention of Employees in PLE over last 10 years.

Our sample database has 40 observations for 7 variables, 4 scale variables and 3

categorical variables.

## About Data: Visualization



**Proportions via Local**

Not Local, 17, 43%

Local,23, 57%

Local Not Loc al

Retention Years shows the number of years any employee served PLE in the last 10 years.

Local Variable have 2 categories as Y and N, means employee is



**Proportion for Local/**

**Not Local**

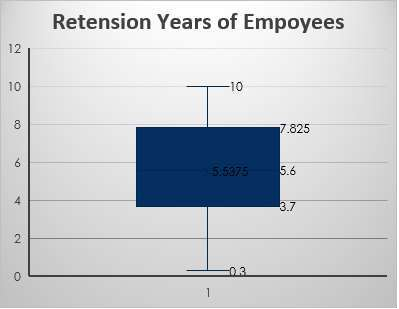
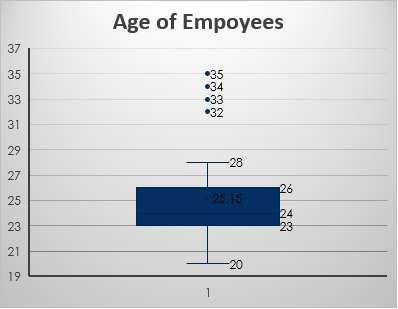
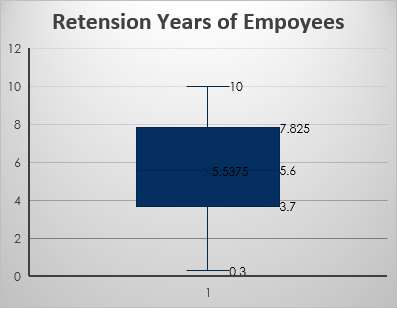
Female, 13, 32%

Male,27,

68%

Female Male

resident of the local region or not.

College Graduation defines whether the employee finished college education or not, whether employee went for higher education or not provided he/she has completed graduation degree.



**Proportions via Education**

Not College

G rad, 13, 33%

College

G rad, 27,

67%

College G rad

Not College G rad



## Statistical Techniques

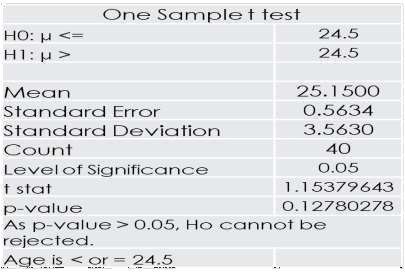
|  |  |
| --- | --- |
| Age | |
| Mean | 25.2 |
| Standard Error | 0.6 |
| Median | 24.0 |
| Mode | 23.0 |
| Standard Deviation | 3.6 |
| Sample Variance | 12.7 |
| Kurtosis | 1.5 |
| Skewness | 1.4 |
| Range | 15.0 |
| Minimum | 20.0 |
| Maximum | 35.0 |
| Sum | 1006.0 |
| Count | 40.0 |

Hypothesis for Average Age of the Employees.

H0: µ (average age) <= 24.5 years

H1: µ (average age) > 24.5 years

We want to check that the average age for employee is more than 24.5 years.

We performed one sample one tail t-test(right tail) to analyze the threshold age of 24.5 years with this sample data and as we get the p-value 0.1278 / 12.78%, means the null hypothesis that is average age is equals to or less than

24.5 years, cannot be rejected.

Average age for employee is less than or equal to 24.5 years.

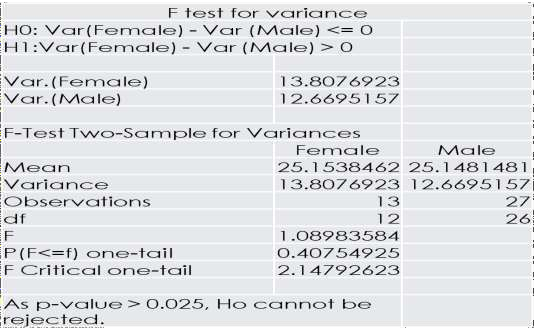


## Statistical Techniques

Hypothesis for Difference between age of the employees across gender.

H0: µ(Female) - µ(Male) = 0 H1: µ(Female) - µ(Male) ≠ 0

We want to check whether is there any difference between the average age of male and female, and also is there any variance between their average ages.

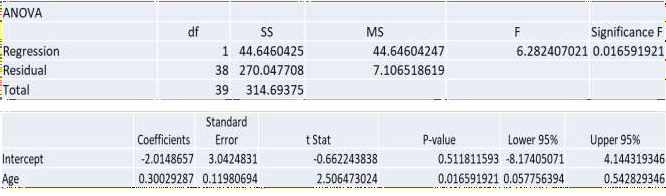
We performed two sample two tail t-test to check the first statement, and we get p-value 0.99 / 99%, means the Null hypothesis that there is no significant difference between average age across gender, can not be rejected. Thus, there is no difference between average age of across gender.

We also performed two sample F-test to check whether there is any variance in the average age across gender for the organization, we get p- value 0.40 / 40%, means the null hypothesis that there is no significant variation for average age between both genders can not be rejected.

This will confirm that there is no disparity for the age group across gender

working in the PLE.

## Statistical Techniques

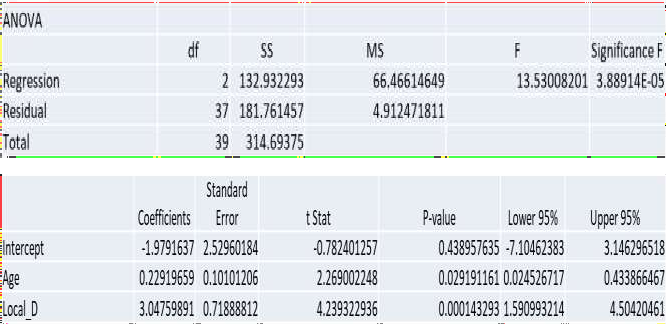
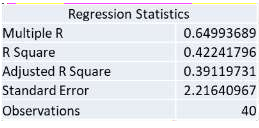


We perform regression analysis for retention years (Dependent Variable) by taking age and local or not as the independent variables. When we did the regression with only age variable, we get R2 = 0.14 and p-value is also < 0.05, implying there is relation between age and retention years.

We also want to check the significance of Local or Not local on retention years. Then we add this another variable which gives us the best fit model, here we get R2

= 0.4224 and Adjusted R2 = 0.3911 and also p value is <

0.05



In our best fit model, we get better Adjusted R2 and also

the p-value improved significantly.

Regression Model - **1.979164 + 0.229197\*Age +**

**3.047599\*Local\_D**.

Moreover, we can also make this regression model more dynamic by considering the lower 95% and upper 95% values for intercept and coefficient, which will tell us that the predicted value will lie within these two calculated values. (with 95% confidence level)

**Note: Here we are considering absolute model for**

**predictions.**



## Statistical Techniques

Further we also check for any correlation between gender and local or not with assumption that may be there are more number of females from local.

We run chi- square test for this, by considering the null hypothesis that there is no relation between both these variables.

H0: No relation between Local / Not Local and Gender

H1: There is significant relation between Local / Not Local

and Gender

After running chi-square we get p-value to be 0.7199 / 71.9% and now we can draw the inference that there is no significant relation between gender and local or not local, and the null hypothesis can not be rejected.

This will tell us that there is no case of gender inequality basis from where they belongs to, meaning they are local or not local.

## Result and Discussion:



* Average Age of Employees is 25.15 years, and for population it can be 24.5 years or less, means there is youth- generation employed in PLE. There are some outliers in the Age database which shows the availability of the experienced staff in the PLE.
* But Median is 24 years, means 50% employees are of this age.
* A significant number of employees (22.5%) are less than the average age of employees, derived from the Mode( age =23

years).

* The Range of Age of employees is from 21.6 years to 28.6 years (from Standard deviation).
* Age and Local / Not Local has a significant relation with Retention Years in the PLE.
* There is no significant relation between Gender and Local or Not Local.
* There is no significant difference in age across Gender, also there is no significant differences in variations in average age

for both genders.

* Regression Model for Years Retention is = -1.979164 + 0.229197\*Age + 3.047599\*Local\_D.

## Suggestions

There is good gender equality among employees and also a good share of youth is working in PLE. As we can see in our regression model, Local or Not Local has a larger coefficient value (3.0476) which means PLE can take initiatives to have onboard more people among locals which will increase and improve the retention span of employees.

# SHEET TITLE: EMPLOYEE SATISFACTION



## Reason for choosing this sheet:

This sheet provides us with the quarterly data of past 4 years of internal surveys of employees to determine their overall satisfaction in jobs. It provides the average ratings of three employee categories namely Design and Production, Manager, and Sales and Administration. By which we were able to find in which category employees are more satisfied and in which one they are less satisfied.

## Business Objective:



We want to test the variation in employee’s satisfaction across all categories.

## Hypothesis (Null and Alternate):

For All categories:-

Null Hypothesis(Ho): Average Satisfaction is same across all three employee categories.

Alternate Hypothesis(Ha): Average Satisfaction is different across all three employee categories.

For Sales & Administration and Design & Production:-

Null Hypothesis(Ho): Average satisfaction of employees of Sales and Administration is less than or equal to average

satisfaction of employees of Design and Production.

Alternate Hypothesis(Ha): Average satisfaction of employees of Sales and Administration is more than

average satisfaction of employees of Design and Production.

## ABOUT DATA: Descriptive and Visualization

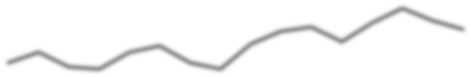


|  |  |
| --- | --- |
| Design & Production | |
| Mean | 2.95 |
| Standard Error | 0.02 |
| Median | 2.95 |
| Mode | 2.86 |
| Standard Deviation | 0.09 |
| Sample Variance | 0.01 |
| Kurtosis | -1.09 |
| Skewness | 0.24 |
| Range | 0.29 |
| Minimum | 2.83 |
| Maximum | 3.12 |
| Sum | 47.18 |
| Count | 16 |

|  |  |
| --- | --- |
| Sales & Administration | |
| Mean | 3.52 |
| Standard Error | 0.03 |
| Median | 3.52 |
| Mode | 3.37 |
| Standard Deviation | 0.10 |
| Sample Variance | 0.01 |
| Kurtosis | -0.62 |
| Skewness | 0.23 |
| Range | 0.34 |
| Minimum | 3.37 |
| Maximum | 3.71 |
| Sum | 56.26 |
| Count | 16 |

|  |  |
| --- | --- |
| Manager | |
| Mean | 3.82 |
| Standard Error | 0.04 |
| Median | 3.85 |
| Mode | 3.92 |
| Standard Deviation | 0.14 |
| Sample Variance | 0.02 |
| Kurtosis | 0.96 |
| Skewness | -1.06 |
| Range | 0.53 |
| Minimum | 3.48 |
| Maximum | 4.01 |
| Sum | 61.19 |
| Count | 16 |

|  |  |
| --- | --- |
| Total | |
| Mean | 3.18 |
| Standard Error | 0.02 |
| Median | 3.19 |
| Mode | 3.04 |
| Standard Deviation | 0.09 |
| Sample Variance | 0.01 |
| Kurtosis | -1.77 |
| Skewness | -0.19 |
| Range | 0.25 |
| Minimum | 3.04 |
| Maximum | 3.29 |
| Sum | 50.83 |
| Count | 16 |



**Design & Production**

3.15

3.05

2.95

2.85

2.75

1st Q- 2nd 3rd 4th 1st Q- 2nd 3rd 4th 1st Q- 2nd 3rd 4th 1st Q- 2nd 3rd 4th 11 Q-11 Q-11 Q-11 12 Q-12 Q-12 Q-12 13 Q-13 Q-13 Q-13 14 Q-14 Q-14 Q-14

**Manager**

4.10 4.00

3.90

3.80

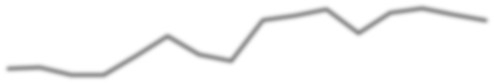
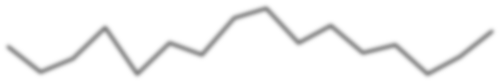
3.70

3.60

3.50

3.40

1st Q- 2nd 3rd 4th 1st Q- 2nd 3rd 4th 1st Q- 2nd 3rd 4th 1st Q- 2nd 3rd 4th 11 Q-11 Q-11 Q-11 12 Q-12 Q-12 Q-12 13 Q-13 Q-13 Q-13 14 Q-14 Q-14 Q-14



**Sales & Administration**

3.8 3.7

3.6

3.5

3.4

3.3

1st Q- 2nd 3rd Q- 4th Q- 1st Q- 2nd 3rd Q- 4th Q- 1st Q- 2nd 3rd Q- 4th Q- 1st Q- 2nd 3rd Q- 4th Q-

11 Q-11 11 11 12 Q-12 12 12 13 Q-13 13 13 14 Q-14 14 14

**Total**

3.40

3.30 3.20

3.10

3.00 1st Q- 2nd 3rd Q-4th Q- 1st Q- 2nd 3rd Q-4th Q- 1st Q- 2nd 3rd Q-4th Q- 1st Q- 2nd 3rd Q-4th Q- 11 Q-11 11 11 12 Q-12 12 12 13 Q-13 13 13 14 Q-14 14 14

## About Data:



We have quarterly average satisfaction of employee data of four year. (2011 – 2014)

* Employees belonging to Design and Production category has average satisfaction of 2.95 which lies in the range of 2.927 to 2.9709.
* Employees belonging to Sales and Administration category has average satisfaction of 3.52 which lies in the range of

3.4906 to 3.5415.

* Employees belonging to Manager category has average satisfaction of 3.82 which lies in the range of 3.48 to

4.01

* **Here, Average satisfaction level of employees belonging to managerial position is higher as compared to other**

**categories. But overall average satisfaction is 3.18 as presented by mean of total.**

* On an average the deviation of data from its mean is very small in cases of all the categories and total (0.09) as

well. But in comparison, deviation is a bit more in Manager (0.14) followed by Sales and Administration (0.10).

* As skewness is more than 0 in Design and Production, and Sales and Administration, their data is positively skewed.

But data of Manager and Total are negatively skewed.

* As kurtosis is less than 3 which implies negative kurtosis in all categories both individually and overall, our data is

platykurtic.

* Variables: Design and Production; Sales and Administration; and Manager.

## Statistical Techniques



Hypothesis For All categories: -

Ho: Average Satisfaction is same across all three categories

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anova: Single  Factor |  |  |  |  |  |  |
| SUMMARY |  |  |  |  |  |  |
| Groups | Count | Sum | Average | Varianc e |  |  |
| Design & Production | 16 | 47.18 | 2.94875 | 0.008425 |  |  |
| Manager | 16 | 61.19 | 3.824375 | 0.02089  3 |  |  |
| Sales & Administratio n | 16 | 56.26 | 3.51625 | 0.010398 |  |  |
| ANOVA |  |  |  |  |  |  |
| Source of Variation | SS | df | MS | F | P-value | F crit |
| Between Groups | 6.313154167 | 2 | 3.156577 | 238.4347 | 1.13E-  24 | 3.204317 |
| Within Groups | 0.59574375 | 45 | 0.013239 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 6.908897917 | 47 |  |  |  |  |

Ha: Average Satisfaction is different across all three categories

We want to check that whether employees of all categories are satisfied equally or not.

We performed ANOVA: Single factor test to analyze the difference between average satisfaction of all categories and as F > F Critical and p value is less than 0.05, we can say the null hypothesis can be rejected.

There is indeed a significant difference in Satisfaction level of employees across their Categories.

As Average of Manager is highest we can say employees of belonging to Managerial position have highest satisfaction level as compared to other two.

## Statistical Techniques



For Sales & Administration and Design & Production

|  |  |  |
| --- | --- | --- |
| t-Test:Two- Sample Assuming Unequal  Variances |  |  |
|  | Sales & Administration | Design & Production |
| Mean | 3.51625 | 2.94875 |
| Variance | 0.010398333 | 0.008425 |
| Observations | 16 | 16 |
| Hypothesized Mean  Difference | 0 |  |
| df | 30 |  |
| t Stat | 16.54541218 |  |
| P(T<=t) one-tail | 6.29356E-17 |  |
| t Critical one-tail | 1.697260887 |  |
| P(T<=t) two-tail | 1.25871E-16 |  |
| t Critical two-tail | 2.042272456 |  |

Ho: µ(Sales & Administration)-µ(Design & Production) <= 0

Ha: µ(Sales & Administration)-µ(Design & Production) > 0

We want to check whether average satisfaction in Design & Production is

less than or equal to Sales & Administration or not.

We performed t-Test: Two-Sample Assuming Unequal Variances to analyze the objective and as we get the p-value < 0.05 and t stat > t Critical, the null hypothesis can be rejected .

Hence, on an average employees belonging to Sales and Administration are more satisfied as compared to employees belonging to Design and Production.

## Analysis



The average satisfaction of every year is presented here, more the bar is filled with color, more averagely it is rated during the particular year by particular category of employees.



3

9

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Design & Production** | | **Sales & Admin.** | **Manager** | | **Total** | |
| 2011 | 2.8 | 6 | 3.49 | 3.7 |  | 3.0 | 6 |
| 2012 | 2.8 | 9 | 3.51 | 3.7 |  | 3.1 | 3 |
| 2013 | 2.9 | 9 | 3.59 | 3.90 | | 3.25 | |
| 2014 | 3.06 | | 3.49 | 3.89 | | 3.27 | |
| Sparkline Trends |  | |  |  | |  | |

Spark lines denote that on an average satisfaction of Design and Production has been increasing constantly, satisfaction of Sales and Administration has increased and then there is a sharp decline in 2014, satisfaction of Managers have been increasing only and satisfaction in total has also been increasing only since four years.

The average satisfaction on Quarterly basis is presented and with the help of conditional formatting we can see in which quarter of which year average satisfaction was highest. Darker the shade of green the less is the average satisfaction level and the more it tends towards white the more is the satisfaction level of employees.

|  |  |  |  |
| --- | --- | --- | --- |
| **Quarter** | **Design & Production** | **Sales & Administration** | **Manager** |
| **1st Q-11** | 2.86 | 3.51 | 3.81 |
| **2nd Q-11** | 2.91 | 3.38 | 3.76 |
| **3rd Q-11** | 2.84 | 3.45 | 3.86 |
| **4th Q-11** | 2.83 | 3.61 | 3.48 |
| **1st Q-12** | 2.91 | 3.37 | 3.75 |
| **2nd Q-12** | 2.94 | 3.53 | 3.92 |
| **3rd Q-12** | 2.86 | 3.47 | 3.89 |
| **4th Q-12** | 2.83 | 3.66 | 3.58 |
| **1st Q-13** | 2.95 | 3.71 | 3.82 |
| **2nd Q-13** | 3.01 | 3.53 | 4.01 |
| **3rd Q-13** | 3.03 | 3.62 | 3.92 |
| **4th Q-13** | 2.96 | 3.48 | 3.84 |
| **1st Q-14** | 3.05 | 3.52 | 3.92 |
| **2nd Q-14** | 3.12 | 3.37 | 4.00 |
| **3rd Q-14** | 3.06 | 3.46 | 3.93 |
| **4th Q-14** | 3.02 | 3.59 | 3.70 |

We can see that Average satisfaction of Design and Production is lowest in Quarter 4 of 2011 and 2012 which average satisfaction level of 2.83 each. Average satisfaction is highest in Quarter 2 of 2013 Manager with average satisfaction level of 4.01.

## Result and Discussion:



As we can see from tests that Managers on an average have highest satisfaction level as compared to others and from sparklines also we can see its average satisfaction has been rising only but in 2014 it has been rising on diminishing rate so we can work on that.

As we can see from test that Sales and Administration on an average have higher satisfaction level as compared to Design and Production but we can see the sharp fall in average satisfaction level of Sales and Administration in 2014 from sparklines and continuous increase in the satisfaction level of Design and Production says that we have been doing something for betterment employees of Design and Production category since years but in 2014 we lacked somewhere for betterment of employees of Sales and Administration.

## Suggestion:

PLE should continue the betterment of employees of Design and Production and also should work on knowing the needs of that department and if it is not fulfilling some basics then it should try to do that as well to cover up the gap.

PLE should see why Manager’s satisfaction is rising on diminishing rate in 2014, it can be because of lack of motivation or maybe something else, so PLE should find the reason of that and should start working on that.

PLE should see the reason of sharp fall in satisfaction level of employees of Sales and Administration and should

try to convert that fall in rise in increasing rate in 2015.

# SHEET TITLE: TRACTORS UNIT SALES + INDUSTRY



**TRACTORS TOTAL SALES**

## Reason for choosing this sheet:

Sales figures are the backbone for any company as it is the source of the revenue generation

which is the ultimate goal of any organization.

Here we wanted to predict and analyze the market condition of the last launched product that

is tractor. (Launched 10 years ago).

We also wanted to compare sales region-wise and year-wise to look for any specific trends etc.

## Business Objective:

We want to test for the market condition of Tractors across 5 regions in the 5 years. (2010-2014)

We want to analyze the market share of Tractors for PLE with Industry’s sales figures.

**Europe**

## Hypothesis (Null and Alternate):



The average sales in Eur for tractors has been declining.

Null Hypothesis: The average sales from Eur is more than or equal to 675 units.

Alternate Hypothesis: The average sales from Eur is less than 675 units.

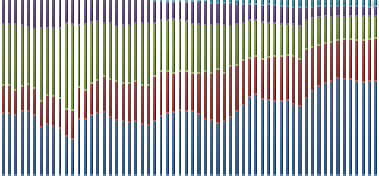
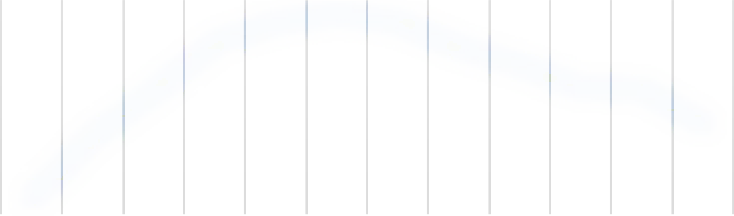
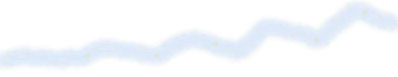
**North America**

The average sales in NA for tractors has been increasing .

Null Hypothesis: The average sales from Eur is less than or equal to 940 units.

Alternate Hypothesis: The average sales from Eur is more than 940 units.

## About Data – Tractor Sales



**Tractor Sale - Region**

100%

90%

80%

70%

60%

50%

40%

30%

20%

10%

0%

NA SA Eur Pac China

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

3000

2800

2600

2400

2200

2000

**Average Sales\_Month-wise**

**PLE Sales**

PLE sold the tractors in 5 regions, North America (NA), South America (SA), European Market (Eur), Pacific region (Pac) and China. We have dataset for sales in these region for the 5 years from 2010 to 2014 and that is why we have the sample size of 60 observations for each region.

NA has the highest average sales of 1075 (35% - 40%) units, then Eur have average sales of 648 units, followed by SA, Pac and China by average sales of 598, 272 and 47 units respectively. Average Sales for PLE across all regions is 2640 units, with minimum sale of 1592 units (January 2010) and maximum of 4476 units (June 2014).

Jan-10

Apr-10 Jul-10 Oct-10 Jan-11 Apr-11 Jul-11 Oct-11 Jan-12 Apr-12 Jul-12 Oct-12 Jan-13 Apr-13 Jul-13 Oct-13 Jan-14 Apr-14 Jul-14

Oct-14

According to the empirical rule, 95% of the sales for PLE will lie between 1013 and

4267 units.

Overall PLE has a increasing trend for sales from all regions as average sales in

2010 was 1742, in 2011 was 2080, in 2012 was 2466, in 2013 was 3062 and in

Jan-10

Apr-10 Jul-10 Oct-10 Jan-11 Apr-11 Jul-11 Oct-11 Jan-12 Apr-12 Jul-12 Oct-12 Jan-13 Apr-13 Jul-13 Oct-13 Jan-14 Apr-14 Jul-14

Oct-14

2014 was 3849.

But the sale is not uniform across the months. There is trend of increase in first two quarters then peak in May – June and then decrease in sales in the last two quarters, which is referring to the presence of seasonal scenario.

**Note: The average sales figures we have , are on monthly basis.**

## About Data



**Sales - Region**

100%

80%

60%

40%

20%

0%

2037

908

1280

844

675

538

NA

547

617

700

699

677

Eur 2012

229

316

286

277

254

Pac

2014

129

768

612

430

274

SA

2011

80

204

China

2010

2013

When we have a look at the sales figures from each region’s point of view, we observe a significant increase in the sales from NA, SA and China, also a constant increase in sales for Pac region too. But In Europe the trend is not promising as in the year 2010 and 2011. In those two years it was the lead region for sales, then in 2012 it slips to second position and then in 2013 and 2014 it become the third region in respect of sales. Also observe that PLE has been selling a significant share of 10.90% of industry’s total sales in 2010 in Eur region as compared to 09.31% in 2014. Also the overall average share of PLE in Eur went to 10.06%



**NA**

2350

1850

1350

850

350

However NA has emerged as the lead sales generating region for PLE in these 5 years. It started as the number second region for highest sales in 2010, 538 units and then become lead sales region in 2012 with 844 units and in 2014 it managed to sell 2037 units of tractors in the region. From Industry’s perspective, PLE has a share of 11.26% of industry’s total sales in 2010 in NA, which reaches to 16.37% in 2014. Now the overall average share of PLE in NA went to 13.91%



**Eur**

870

770

670

570

470

Jan-10

Apr-10 Jul-10 Oct-10 Jan-11 Apr-11 Jul-11 Oct-11 Jan-12 Apr-12 Jul-12 Oct-12 Jan-13 Apr-13 Jul-13 Oct-13 Jan-14 Apr-14 Jul-14 Oct-14

Hence we are focusing on these two markets of PLE, Europe and North America

for drawing further inferences about their performances.

Jan-10

Apr-10 Jul-10 Oct-10 Jan-11 Apr-11 Jul-11 Oct-11 Jan-12 Apr-12 Jul-12 Oct-12 Jan-13 Apr-13 Jul-13 Oct-13 Jan-14 Apr-14 Jul-14 Oct-14



## Statistical Techniques – North America

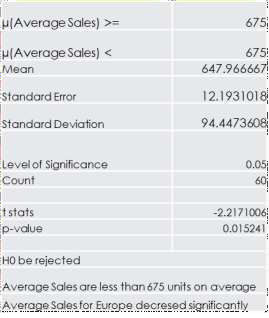
Hypothesis to check the average sales of the Europe region.

µ(Average Sales) >= 675

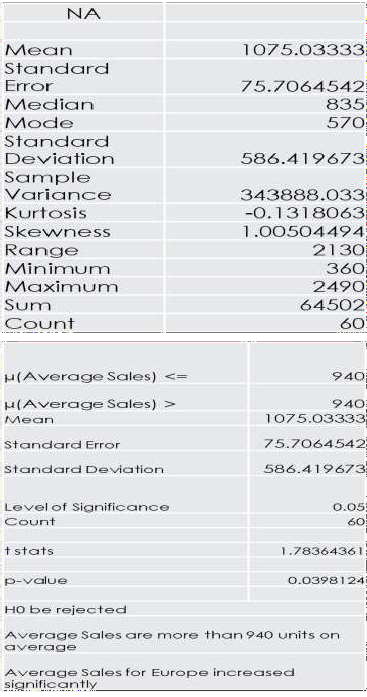
µ(Average Sales) < 675

We want to check that is the average sales of the European market for

tractors is less 675 units or not.

We performed one sample left tail t-test to analyze the average sales is less than 675 or not and as we get the p-value 0.01524 / 01.52%, means the null hypothesis that the average sales is more than or equal to 675 can be rejected and thus we can conclude that the average sales for tractors is significantly less than 675.

The average sales in Europe for tractors is significantly lesser than 675 units.





## Statistical Techniques – North America

Hypothesis to check the average sales of the NA region. µ(Average Sales) <= 940

µ(Average Sales) > 940

We want to check that is the average sales of the NA market for tractors is

more than 940 units or not.

We performed one sample right tail t-test to analyze the average sales is more than 940 or not and as we get the p-value 0.03981 / 03.98%, means the null hypothesis that the average sales is less than or equal to 940 can be rejected and thus we can conclude that the average sales for tractors is significantly greater than 940 units.

The average sales in NA for tractors is significantly greater than 940 units.

**Europe**

## Result and Discussion:



From our inferential test we can conclude that the average sales have been decreasing significantly for PLE.

**North America**

From our inferential test we can conclude that our average sales have increased significantly in these last 5 years.

## Suggestion:

#### Europe

PLE should analyze the condition in this region and also conduct any further analysis on industry data to confirm

that is there any problem going on with overall tractor industry or it is happening with only with PLE.

#### North America

NA has been the lead sales generating region for PLE, so PLE should take care of this region like that. PLE can conduct further analysis for customer satisfaction, dealer satisfaction, industry data to more accurately judge its place in NA.

As NA is the sales generating regions, performance of this region must have been the top priority for PLE. Good performance in NA can help PLE to consider more aggressive decision and policies in other regions of concern for sales maximization and to increase its share in tractor industry for those regions.

# SHEET TITLE: MOWER UNIT SALES + TRACTOR



**UNIT SALES + COMPLAINTS**

## Reason for choosing this sheet:

These sheets provide us with the five years data of monthly unit sales of mower and tractor; and complaints registered by customers. This can show the comparison between complaints of different regions and also complaints can be compared with total sales.

## Business Objective:

We want to test the relation between Total Sales and Total Complaints; and comparison between

complaints of different regions.

For Total Sales and Total Complaints:-

## Hypothesis (Null and Alternate):



Null Hypothesis (Ho): There is no relation between Total Unit Sales and Complaints. Alternate Hypothesis (Ha): There is relation between Total Unit Sales and Complaints.

For Regional Complaints:-

Null Hypothesis(Ho): Average Complaints of Europe is less than or equal to Average Complaints of South

America.

Alternate Hypothesis(Ha): Average Complaints of Europe is more than Average Complaints of South America.

For Total Sales:-

Null Hypothesis(Ho): Average Total Sales is less than or equal to 11000. Alternate Hypothesis(Ha): Average Total Sales is more than 11000.

## ABOUT DATA: Descriptive and Visualization



**Complaints/Total Sales**

3.00%

2.50%

2.00%

1.50%

1.00%

0.50%

0.00%

Jul-09

Nov-10

Apr-12

Aug-13

Dec-14

May-16

**Months**

**Months**

400

350

300

250

200

150

100

50

0

**Complaints**

Complaints

**Percentage**

**No. of Complaints**

Jan-10

Apr-10 Jul-10 Oct-10 Jan-11 Apr-11 Jul-11 Oct-11 Jan-12 Apr-12 Jul-12 Oct-12 Jan-13 Apr-13 Jul-13 Oct-13 Jan-14 Apr-14 Jul-14

Oct-14

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mower Unit Sales | | Tractor Unit Sales | | Total Unit Sales | | Complaints | |
| Mean | 9148.05 | Mean | 2640.183333 | Mean | 11788.23333 | Mean | 248.9166667 |
| Standard Error | 267.296501 | Standard Error | 105.013514 | Standard Error | 309.3331879 | Standard Error | 6.013651685 |
| Median | 9390 | Median | 2408 | Median | 11785.5 | Median | 244 |
| Mode | 7020 | Mode | 2324 | Mode | #N/A | Mode | 232 |
| Standard Deviation | 2070.469794 | Standard Deviation | 813.431182 | Standard Deviation | 2396.08457 | Standard Deviation | 46.58154565 |
| Sample Variance | 4286845.167 | Sample Variance | 661670.2879 | Sample Variance | 5741221.267 | Sample Variance | 2169.840395 |
| Kurtosis | -1.189149598 | Kurtosis | -0.652116779 | Kurtosis | -0.926674391 | Kurtosis | -0.569504798 |
| Skewness | -0.188903124 | Skewness | 0.661643538 | Skewness | -0.058304941 | Skewness | 0.078336549 |
| Range | 6930 | Range | 2884 | Range | 9338 | Range | 196 |
| Minimum | 5350 | Minimum | 1592 | Minimum | 7220 | Minimum | 154 |
| Maximum | 12280 | Maximum | 4476 | Maximum | 16558 | Maximum | 350 |
| Sum | 548883 | Sum | 158411 | Sum | 707294 | Sum | 14935 |
| Count | 60 | Count | 60 | Count | 60 | Count | 60 |



**About Data:**



We have quarterly average satisfaction of employee data of four year (2011 – 2014).

* Mower Unit Sales has average sales of 9148.05 units which lies in the range of 8880.75 to 9415.35 units.
* Tractor Unit Sales has average sales of 2640.18 units which lies in the range of 2535.17 to 2745.19 units.
* Here, Average Sales of Mower are higher in comparison to Tractor.
* Average complaints are 248.92 which lies in the range of 154 to 350.
* On an average the deviation of data from its mean is highest in case of Mower in comparison to Tractor. In case of complaints

the deviation of data from its mean is 46.58

* As skewness is more than 0 in Tractor Unit Sales, and Complaints, their data is positively skewed. But data of Mower Unit

Sales and Total Unit Sales are negatively skewed.

* As kurtosis is less than 3 which implies negative kurtosis in both Mower Unit Sales; Tractor Unit Sales; and Total Unit Sales;

and Complaints as well, our data is platykurtic.

* Variables-> Total: Total Unit Sales; Mower: Mower Unit Sales; Tractor: Tractor Unit Sales; Complaints: Total region-wise

complaints; COS: Complaints/Total Sales.

## Statistical Techniques



Hypothesis For Total Sales and Total Complaints:- Ho: Total Unit Sales and Complaints has no relation Ha: Total Unit Sales and Complaints are related

|  |  |  |
| --- | --- | --- |
| t-Test: Paired Two Sample for Means | | |
|  | Total Unit  Sales | Complaints |
| Mean | 11788.23333 | 248.916666  7 |
| Variance | 5741221.267 | 2169.84039  5 |
| Observations | 60 | 60 |
| Pearson Correlation | 0.875556458 |  |
| Hypothesized Mean  Difference | 0 |  |
| df | 59 |  |
| t Stat | 37.94807052 |  |
| P(T<=t) one-tail | 1.88616E-43 |  |
| t Critical one-tail | 1.671093032 |  |
| P(T<=t) two-tail | 3.77232E-43 |  |
| t Critical two-tail | 2.000995378 |  |

We want to check the relation between Total Sales and Complaints. We performed t-Test: Paired Two Sample for Means and as t stat > t Critical and p value is less than 0.05, we can say that null hypothesis can be rejected.

Hence, Total Unit Sales and Complaints are related.

Correlation value shows that these variables have high positive relation this might be because of the reason that more the products of PLE will be sold in market more will be Complaints that can be received.

|  |  |
| --- | --- |
| Year | Average Complaints |
| 2010 | 1.87% |
| 2011 | 2.15% |
| 2012 | 2.21% |
| 2013 | 2.22% |
| 2014 | 2.19% |

Average Complaints have risen from 2010 to 2011, but we hold the average

percentage in another years.

## Statistical Techniques



|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal  Variances |  |  |
|  | Sales &  Administration | Design &  Production |
| Mean | 3.51625 | 2.94875 |
| Variance | 0.010398333 | 0.008425 |
| Observations | 16 | 16 |
| Hypothesized Mean Difference | 0 |  |
| df | 30 |  |
| t Stat | 16.54541218 |  |
| P(T<=t) one-tail | 6.29356E-17 |  |
| t Critical one-tail | 1.697260887 |  |
| P(T<=t) two-tail | 1.25871E-16 |  |
| t Critical two-tail | 2.042272456 |  |

For Complaints of Eur and SA:-

Ho: µ(Eur) - µ(SA) <= 0

Ha: µ(Eur) - µ(SA) > 0

We want to check whether average complaints of Eur is less than or equal to SA or not.

We performed t-Test: Two-Sample Assuming Unequal Variances to analyze the objective and as we get the p-value < 0.05 and t stat > t Critical, the null hypothesis can be rejected.

We can see from graph as well, the trend line shows difference between

complaints of Eur and SA.

Hence, on an average Eur has higher complaints than SA.

## Statistical Techniques



Hypothesis For Total Sales :- Ho: µ <= 11000

|  |  |
| --- | --- |
| Ho: µ <= | 11000 |
| Ha: µ > | 11000 |
| Significance Level(α) | 0.05 |
| Sample Mean (x)̅ | 11788.23333 |
| Sample Std Deviation.(s) | 2396.08457 |
| Sample Size(n) | 60 |
| Standard Error | 309.3331879 |
| t-cal. | 2.548169302 |
| t-critical (benchmark) | 1.671093032 |
| t critical for upper tail test | 1.671093032 |
| p value | 0.006726409 |

Ha: µ > 11000

We want to check whether average Total Sales are less than or equal to

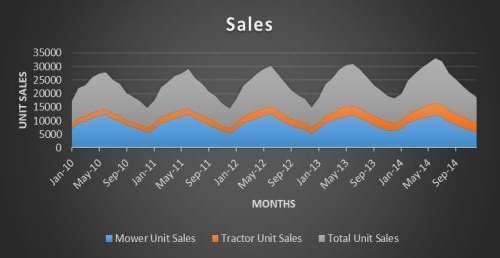
11000 or not.

We performed t-Test: One tailed (upper tailed) and as t stat > t Critical

and

p value is less than 0.05, we can say that null hypothesis can be

rejected.

Hence, on an average Total Sales of PLE is more than 11000.

We can see from the data and graph that in the 1st six months of the year i.e. from January to June the sales of both products keep on increasing and reaches its peak in June and then starts falling and falls till it reaches its bottom in December and this cycle continues for all years. This shows that the product sales are seasonal and our products are made to be best used in summers.

## Result and Discussion:



As we can see from test of Total Sales and Complaints that total sales and complaints are related and have high positive correlation as well which means complaints keep on rising with increase in Total Sales. But as shown by table PLE holds the average complaints per Total Sales in 2011 – 2014.

PLE is doing great as the average total sales are more than 11000.

As we can see from test and graph that Europe has higher complaints than South America, this can be because Europe has higher sales than South America as Total Sales and Complaints are highly related.

## Suggestion:

PLE should make an objective and try to minimize the percentage of complaints.

PLE should also try to increase the sales a bit more in South America and other regions as well.



# SHEET TITLE: DEFECTS AFTER DELIVERY

## Reason for choosing this sheet:

This sheet provides us with the number of Defects after delivery in supplier-provided material found in all

shipments received from suppliers in different months for 5 years (2010-2014).

By using this data we can perform defect tracking using different analytical techniques. The analysis can provide us useful insights about:

* + Whether there is any improvement in quality in supplier-provided material or not.
  + Whether there is increasing trend in improvement or not.
  + We can also forecast the no of defects for the upcoming year according to current trend. This could help PLE in analyzing whether there is any improvement in the forecasted year or not.

## Business Objective:



The mined information can help PLE in determining whether there is any positive impact of the measures taken for defect reduction or not. Moreover, if there would be any change in the current measures for improvement in supplier-provided material, in the forecasted year, then on comparing the actual value with the forecasted value PLE could determine whether the change in measures were effective or not.

## Hypothesis (Null and Alternate):

For ANOVA: Single Factor test:

Null Hypothesis (Ho): Average defects were same over the past 5 years Alternate Hypothesis (Ha): Average defects were not same over the past 5 years For t-Test: Two-Sample Assuming Unequal Variances:

Null Hypothesis (Ho): Average number of defects would be same in 2010 and 2014.

Alternate Hypothesis (Ha): Average number of defects would not be same in 2010 and 2014.

## About Data: Descriptive and Visualization

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2010 | | 2011 | | 2012 | | 2013 | | 2014 | |
| Mean | 826.3333 | Mean | 837.4166667 | Mean | 785.9166667 | Mean | 669.0833 | Mean | 496.25 |
| Standard Error | 3.35824 | Standard Error | 3.182476483 | Standard Error | 15.13748024 | Standard Error | 8.941272 | Standard Error | 15.65254 |
| Median | 826.5 | Median | 839 | Median | 805 | Median | 681.5 | Median | 484 |
| Mode | #N/A | Mode | 828 | Mode | 804 | Mode | #N/A | Mode | #N/A |
| Standard Deviation | 11.63329 | Standard Deviation | 11.02442193 | Standard Deviation | 52.43776976 | Standard Deviation | 30.97347 | Standard Deviation | 54.22198 |
| Sample Variance | 135.3333 | Sample Variance | 121.5378788 | Sample Variance | 2749.719697 | Sample Variance | 959.3561 | Sample Variance | 2940.023 |
| Kurtosis | -0.58025 | Kurtosis | 0.253354213 | Kurtosis | -0.210010309 | Kurtosis | 0.807776 | Kurtosis | -1.74896 |
| Skewness | 0.218469 | Skewness | -0.142075754 | Skewness | -1.212020884 | Skewness | -1.39278 | Skewness | 0.271461 |
| Range | 38 | Range | 41 | Range | 150 | Range | 93 | Range | 139 |
| Minimum | 810 | Minimum | 816 | Minimum | 686 | Minimum | 603 | Minimum | 436 |
| Maximum | 848 | Maximum | 857 | Maximum | 836 | Maximum | 696 | Maximum | 575 |
| Sum | 9916 | Sum | 10049 | Sum | 9431 | Sum | 8029 | Sum | 5955 |
| Count | 12 | Count | 12 | Count | 12 | Count | 12 | Count | 12 |
| Confidence Level(95.0%) | 7.391437 | Confidence Level(95.0%) | 7.004583512 | Confidence Level(95.0%) | 33.31736937 | Confidence Level(95.0%) | 19.67961 | Confidence Level(95.0%) | 34.451 |

900

**AVERAGE NUMBER**

**OF DEFECTS**

800

700

600

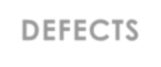
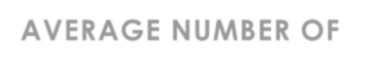
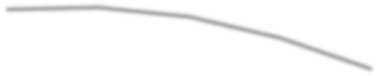
500

400

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 2010 | 2011 | 2012 | 2013 | 2014 |
| Series1 | 826.333333 | 837.416667 | 785.916667 | 669.083333 | 496.25 |

**AVERAGE NUMBER OF**

**DEFECTS**



826.33 837.42 785.92

669.08

496.25



**YEAR**

## About Data:



* From the descriptive analysis is clearly visible that:
* 2010 has average defects of 826.333 which lies in the range of 810 to 848.
* Similarly, 2011, 2012,2013 and 2014 have average defects of 837.417, 785.917, 669.083 and 496.25 which lie in the range of 816 to 857; 686 to 836; 603 to 696; and 436 to 575 respectively.
* Here, 2010 has the highest average defects whereas 2014 has least average defects.
* On an average the deviation of data from its mean is highest in case of 2014 and least in case of 2011 as presented by values

of Standard Deviation in Table already provided.

* As skewness is less than 0 in 2011, 2012, 2013 their data is negatively skewed whereas 2010 and 2014 has positively skewed

data as their skewness is more than 0.

* Kurtosis is less than 3 which implies positive kurtosis in all years (2010 – 2014) which means their distribution is platykurtic.
* Variables-> Number of defects after delivery.

## Statistical Techniques



Hypothesis For Average Defects Trend:-

Ho: Average defects were same over the past 5 years

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anova: Single Factor | |  |  |  |  |  |
| SUMMARY |  |  |  |  |  |  |
| Groups | Count | Sum | Average | Variance |  |  |
| 2010 | 12 | 9916 | 826.3333333 | 135.3333 |  |  |
| 2011 | 12 | 10049 | 837.4166667 | 121.5379 |  |  |
| 2012 | 12 | 9431 | 785.9166667 | 2749.72 |  |  |
| 2013 | 12 | 8029 | 669.0833333 | 959.3561 |  |  |
| 2014 | 12 | 5955 | 496.25 | 2940.023 |  |  |
| ANOVA |  |  |  |  |  |  |
| Source of  Variation | SS | df | MS | F | P-value | F crit |
| Between Groups | 984600.3333 | 4 | 246150.0833 | 178.2154 | 8.67824E-  31 | 2.53968863 |
| Within Groups | 75965.66667 | 55 | 1381.193939 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 1060566 | 59 |  |  |  |  |

Ha: Average defects were not same over the past 5 years.

We want to test whether average defects were same over past 5 years or not. We performed ANOVA: Single Factor test to see the same.

As p value is less than 0.05, we can say that null hypothesis can

be rejected.

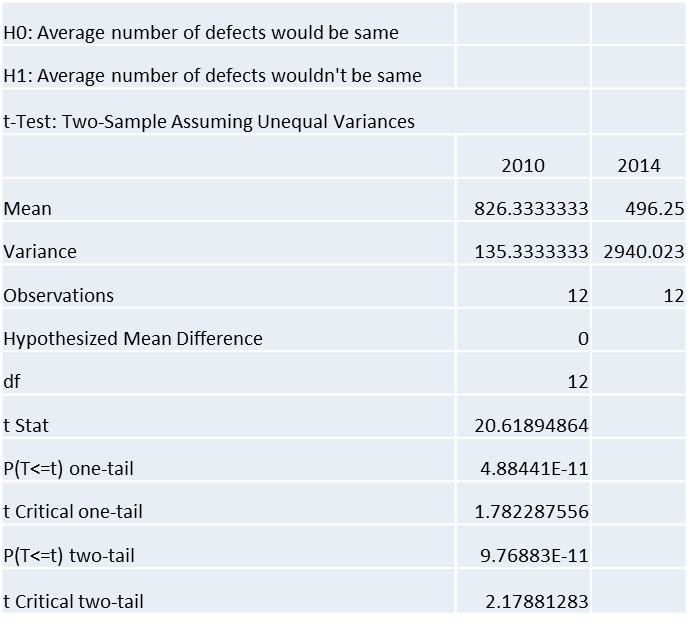
Hence, there is a significant difference in average number of

defects over past 5 years.



## Statistical Techniques

Hypothesis For Average defects: -

Ho: Average number of defects would be same in 2010 and 2014. Ha: Average number of defects would not be same in 2010 and 2014. Ho: µ(Defects in 2010) = µ(Defects in 2014)

Ha: µ(Defects in 2010) ≠ µ(Defects in 2014)

We want to test whether the average no of defects were same or not in 2010 and 2014. We performed t-Test: Two- Sample Assuming Unequal Variances to see the same.

As t stat > t critical and p value is less than 0.05, we have enough

evidence to reject null, so null hypothesis can be rejected.

Therefore, there is significant difference in average number

Of defects in 2010 and 2014.

## Result and Discussion:



From the analysis we found:

There was a significant rise in defects in 2010 to 2011 but we can see a sharp fall

after 2011 and defects were continuously declining till 2014.

## Suggestion:

The continuous decline in defects suggest the current measures undertaken to decrease the defects after delivery are effective. Hence, by proceeding with same measures or by performing any further improvement in existing measures, PLE could reach the goal of minimum Defects after Delivery.

# SHEET TITLE: RESPONSE TIME



## Reason for choosing this sheet:

This sheet provides us the quarterly data on Response times to customer service calls in year 2011-2012.

Since in today’s Consumer dominated economy, customer service is a crucial factor which affect the overall

growth of a company, hence, improvement in Response times to customer service calls is always desired. This data can help PLE in finding the trend Response times to customer service calls in year 2011-2012.

## Business Objective:



By finding out the trend in Response times to customer service calls, PLE can retrospect to find out the reasons of delay or reason of fluctuations, if any, in the response time and can take measures to improve average response time.

## Hypothesis (Null and Alternate):

Hypothesis For Average Time Q4 2011-

|  |  |
| --- | --- |
| Hypothesis For Average Time 2011-  Ho: Average response time is greater than/ equal to | 4.25 |
| Ha: Average response time is lesser than 3.1 |  |
| Null Hypothesis(Ho): µ >=4.25;  Alternate Hypothesis(Ha): µ < 4.25 |  |
| Hypothesis For Average Time 2012-  Ho: Average response time is greater than/ equal to | 3.28 |
| Ha: Average response time is lesser than 3.28 |  |
| Null Hypothesis(Ho): µ >=3.28;  Alternate Hypothesis(Ha): µ < 3.28 |  |

Ho: Average response time is greater than / equal to 5.1

Ha: Average response time is lesser than 5.1

Null Hypothesis(Ho): µ >= 5.1;

Alternate Hypothesis(Ha): µ < 5.1

Hypothesis For Average Time Q4 2012-

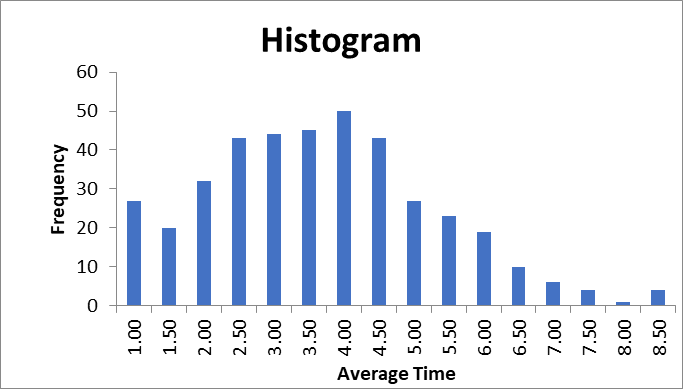
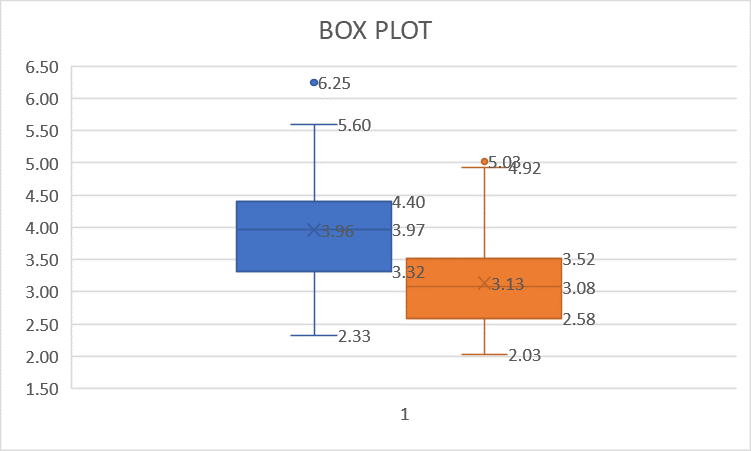
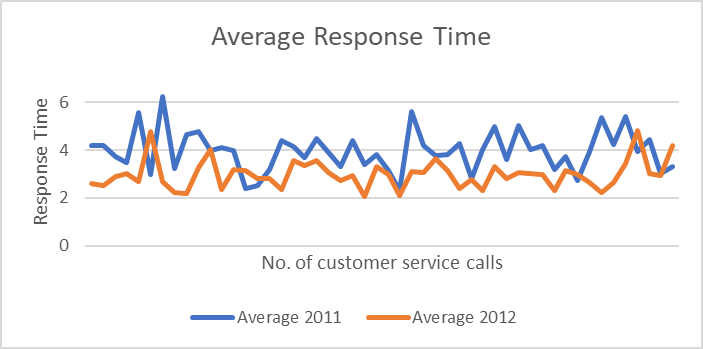
Ho: Average response time is greater than/ equal to 3.1

Ha: Average response time is lesser than 3.1

Null Hypothesis(Ho): µ >= 3.1;

Alternate Hypothesis(Ha): µ < 3.1

## ABOUT DATA: Descriptive and Visualization



**About Data: Descriptive and Visualization**



* We have data of response time of all quarters of 2011 and 2012.
* Average response time of 2011 is 3.96 which lies in the range of 2.32 to 6.25.
* Average response time of 2012 is 2.98 which lies in the range of 2.32 to 6.25
* On an average the deviation of data from its mean is higher in the year 2011 as compared to 2012 as presented by the values

of Standard Deviation in Table.

* As skewness is more than 0 the data is positively skewed.
* Kurtosis is less than 3 which implies negative kurtosis which means the distribution is platykurtic.
* Variables-> Response Time

|  |  |  |  |
| --- | --- | --- | --- |
| Average 2011 | | Average 2012 | |
| Mean | 3.960327443 | Mean | 2.98315732 |
| Standard Error | 0.120300907 | Standard Error | 0.084591952 |
| Median | 3.966630702 | Median | 2.966632305 |
| Mode | #N/A | Mode | #N/A |
| Standard Deviation | 0.85065587 | Standard Deviation | 0.598155427 |
| Sample Variance | 0.72361541 | Sample Variance | 0.357789915 |
| Kurtosis | 0.264210218 | Kurtosis | 2.023819123 |
| Skewness | 0.37553403 | Skewness | 1.114742048 |
| Range | 3.921225527 | Range | 2.770737015 |
| Minimum | 2.3260059 | Minimum | 2.044915244 |
| Maximum | 6.247231427 | Maximum | 4.815652259 |
| Sum | 198.0163721 | Sum | 149.157866 |
| Count | 50 | Count | 50 |



## Statistical Techniques

|  |  |
| --- | --- |
| Ho: µ >= | 5.1 |
| Ha: µ < | 5.1 |
| Significance Level(α) | 0.05 |
| Sample Mean (x̅) | 4.45 |
| Sample Std Deviation (s) | 2.118545827 |
| Sample Size(n) | 50 |
| Standard Error | 0.299607624 |
| t-cal. | -2.15972288 |
| t-critical. (benchmark) | 1.676550893 |
| t-critical for lower tail test | -1.676550893 |
| p value | 0.017857468 |

Hypothesis For Average Time Q4 2011-

Ho: µ >= 5.1 Ha: µ < 5.1

We want to test whether average response time is more than or equal to 5.1 or not.

We performed t test: One tailed(lower tailed) to test the same.

As p value is less than 0.05, we can say that null hypothesis can be rejected.

Hence, the average response time is less than 4.25.

|  |  |
| --- | --- |
| Ho: µ >= | 3.1 |
| Ha: µ < | 3.1 |
| Significance Level(α) | 0.05 |
| Sample Mean (x̅) | 2.53 |
| Sample Std Deviation (s) | 1.130837642 |
| Sample Size(n) | 50 |
| Standard Error | 0.159924593 |
| t-cal. | -3.57775128 |
| t-critical (benchmark) | 1.676550893 |
| t. critical for lower tail test | -1.67655089 |
| p value | 0.000396166 |

Hypothesis For Average Time Q4 2012- Ho: µ >= 3.1

Ha: µ < 3.1

We want to test whether average response time is more than or equal to 3.1 or not.

We performed t test: One tailed(lower tailed) to test.

As p value is less than 0.05, we can say that null hypothesis can be rejected. Hence, the average response is less than 3.28.



## Statistical Techniques

|  |  |
| --- | --- |
| Ho: µ >= | 4.25 |
| Ha: µ < | 4.25 |
| Significance Level(α) | 0.05 |
| Sample Mean (x)̅ | 3.960327443 |
| Sample Std Deviation (s) | 0.85065587 |
| Sample Size(n) | 50 |
| Standard Error | 0.120300907 |
| t-cal. | -2.407900031 |
| t-critical(benchmark) | 1.676550893 |
| t-critical for lower tail test | -1.676550893 |
| p value | 0.009926847 |

Hypothesis For Average Time 2011-

Ho: µ >= 4.25

Ha: µ < 4.25

We want to test whether average response time is more than or equal to 4.25 or not.

We performed t test: One tailed (lower tailed) to test.

As p value is less than 0.05, we can say that null hypothesis can be rejected. Hence, the

average response time is less than 4.25.

Hypothesis For Average Time 2012-

|  |  |
| --- | --- |
| Ho: µ >= | 3.28 |
| Ha: µ < | 3.28 |
| Significance Level(α) | 0.05 |
| Sample Mean | 2.98315732 |
| Sample Std Deviation (s) | 0.598155427 |
| Sample Size(n) | 50 |
| Standard Error | 0.084591952 |
| t-cal. | -3.50911255 |
| t-critical (benchmark) | 1.676550893 |
| t-critical for lower tail test | -1.67655089 |
| p value | 0.000487176 |

Ho: µ >= 3.28 Ha: µ < 3.28

We want to test whether average response time is more than or equal to 3.28 or not. We performed t test: One tailed (lower tailed) to test the same.

As p value is less than 0.05, we can say that null hypothesis can be rejected. Hence, the

average response time is less than 3.28.

## Result and Discussion:



As we can see from test, the average response time has decreased in 2012 compared to 2011 and similar trend is visible in Q4 of 2012 compared to Q4 of 2011. Moreover, the data indicates that the standard deviations i.e., fluctuations in response time are significantly reduced in the year 2012 compared to 2011.

## Suggestion:

The decrease in average response time clearly indicates that the current measures to improve the average

response time are effective and PLE can continue with these for further improvement in future.

# SHEET TITLE: TIME TO PAY SUPPLIERS



**AND ON TIME DELIVERY**

## Reason for choosing this sheet:

These sheets provide us with the data on time to pay suppliers and percentage of number of deliveries on time for five years (2010-2014), on monthly basis. The data can help in giving significant insights upon:

* + What is the average time to pay suppliers.
  + Whether the average time to pay suppliers is same for all

years.

* + Is there any improvement in 2014 compared to 2010.
  + By using data from the sheet "on-time delivery ", we can analyze whether their is a relation between the working days in which payment is made and percentage of on time deliveries.

## Business Objective:

The mined information can help PLE to restructure its financial policy in future to further improve its time to pay suppliers. This would further help PLE in generating confidence in the market and thus may improve its statistics of on-time delivery from suppliers.

## Hypothesis (Null and Alternate):

Hypothesis For Average Time To Pay



Ho: Average time to pay is greater than / equal to 8 days.

Ha: Average time to pay is lesser than 8 days.

Null Hypothesis(Ho): µ >= 8;

Alternate Hypothesis(Ha): µ < 8

Hypothesis For Average Time To Pay 2010 and 2014

Ho: No difference in time to pay between 2010 and 2014.

Ha: Significant difference in time to pay between 2010 and 2014.

Null Hypothesis(Ho): µ (2010) - µ (2014) = 0

Alternate Hypothesis(Ha): µ (2010) - µ (2014) ≠ 0

Hypothesis For Average Time To Pay

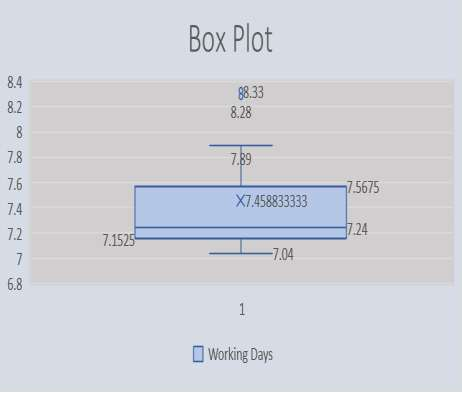
Ho: No difference in time to pay in 2012, 2013 and 2014

Ha: Significant difference in time to pay in 2012, 2013 and 2014

Null Hypothesis(Ho): µ (2012) - µ (2013) - µ (2014) = 0

Alternate Hypothesis(Ha): µ (2012) - µ (2013) - µ (2014) ≠ 0

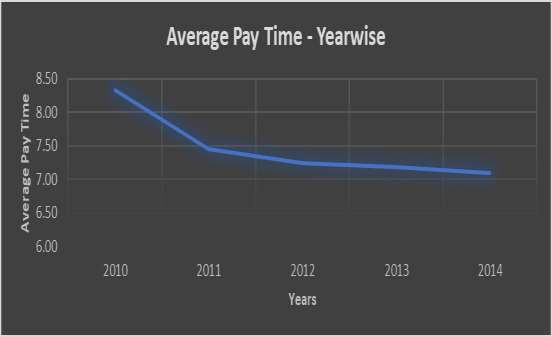
## ABOUT DATA: DESCRIPTIVE AND VISUALIZATION



|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2010 |  | 2011 |  | 2012 |  | 2013 |  | 2014 |  |
| Mean | 8.323333333 | Mean | 7.4525 | Mean | 7.240833333 | Mean | 7.184166667 | Mean | 7.093333333 |
| Standard Error | 0.007719842 | Standard Error | 0.054189972 | Standard Error | 0.007533595 | Standard Error | 0.009882517 | Standard Error | 0.007212503 |
| Median | 8.32 | Median | 7.405 | Median | 7.235 | Median | 7.18 | Median | 7.1 |
| Mode | 8.32 | Mode | #N/A | Mode | 7.22 | Mode | 7.19 | Mode | 7.11 |
| Standard Deviation | 0.026742317 | Standard Deviation | 0.187719568 | Standard Deviation | 0.026097138 | Standard Deviation | 0.034234043 | Standard Deviation | 0.024984844 |
| Sample Variance | 0.000715152 | Sample Variance | 0.035238636 | Sample Variance | 0.000681061 | Sample Variance | 0.00117197 | Sample Variance | 0.000624242 |
| Kurtosis | -0.892071244 | Kurtosis | 1.332154017 | Kurtosis | -0.446630479 | Kurtosis | -1.339553378 | Kurtosis | 0.297059101 |
| Skewness | -0.188870974 | Skewness | 1.185402182 | Skewness | 0.664497794 | Skewness | 0.372469335 | Skewness | -0.959028007 |
| Range | 0.08 | Range | 0.64 | Range | 0.08 | Range | 0.1 | Range | 0.08 |
| Minimum | 8.28 | Minimum | 7.25 | Minimum | 7.21 | Minimum | 7.14 | Minimum | 7.04 |
| Maximum | 8.36 | Maximum | 7.89 | Maximum | 7.29 | Maximum | 7.24 | Maximum | 7.12 |
| Sum | 99.88 | Sum | 89.43 | Sum | 86.89 | Sum | 86.21 | Sum | 85.12 |
| Count | 12 | Count | 12 | Count | 12 | Count | 12 | Count | 12 |



**About Data: Descriptive and Visualization**



* We have data of past five years i.e., 2010-2014.
* Average working days in which suppliers have been paid is 7.459 days.
* Data is not much deviated from mean as Standard Deviation is just 0.46.
* As skewness is more than 0 the data is positively skewed.
* Kurtosis is less than 3 which implies negative kurtosis which means the distribution is platykurtic.
* Variables-> time to pay suppliers

|  |  |
| --- | --- |
| Working Days | |
| Mean | 7.458833333 |
| Standard Error | 0.059352005 |
| Median | 7.24 |
| Mode | 8.32 |
| Standard Deviation | 0.459738657 |
| Sample Variance | 0.211359633 |
| Kurtosis | -0.176019774 |
| Skewness | 1.237138403 |
| Range | 1.32 |
| Minimum | 7.04 |
| Maximum | 8.36 |
| Sum | 447.53 |
| Count | 60 |



## Statistical Techniques

Hypothesis For Average Time to Pay Suppliers- Ho: µ >= 8

|  |  |
| --- | --- |
| Ho: µ >= | 8 |
| Ha: µ < | 8 |
| Significance Level(α) | 0.05 |
| Sample Mean (x̅) | 7.458833333 |
| Sample Std Deviation.(s) | 0.059352005 |
| Sample Size(n) | 60 |
| Standard Error | 0.059352005 |
| t cal | -9.117917127 |
| t critical | 1.671093032 |
| tcritical for lower tail test | -1.671093032 |
| P-Value | 3.63504E-13 |

Ha: µ < 8

We want to check whether average time to pay suppliers (in all four

years, 2010-2014) is less than or equal to 8 days or not. We

performed t test: One tailed (lower tailed) to test the same. As p value

is less than 0.05, we can say that null hypothesis can be rejected. Hence, the average time to pay suppliers is less than 8 days.



## Statistical Techniques

Hypothesis is For Average Time to Pay Suppliers-

|  |  |  |
| --- | --- | --- |
| t-Test: Two-Sample Assuming Unequal Variances | | |
|  | 2010 | 2014 |
| Mean | 8.32333333 | 7.093333333 |
| Variance | 0.00071515 | 0.000624242 |
| Observations | 12 | 12 |
| Hypothesized Mean Difference | 0 |  |
| df | 22 |  |
| t Stat | 116.423746 |  |
| P(T<=t) one-tail | 1.7025E-32 |  |
| t Critical one-tail | 1.71714437 |  |
| P(T<=t) two-tail | 3.4049E-32 |  |
| t Critical two-tail | 2.07387307 |  |

Ho: No difference in time to pay suppliers between 2010 and 2014.

Ha: Significant difference in time to pay suppliers between 2010 and 2014.

Ho: µ(average time to pay in 2010) - µ(average time to pay in 2014) = 0

Ho: µ(average time to pay in 2010) - µ(average time to pay in 2014) ≠ 0

We want to test whether the time to pay suppliers is same between the years 2010

and 2014.

We performed t-Test: Two- Sample Assuming Unequal Variances for the same.

As t stat > t critical and p value is less than 0.05, we have enough evidence to reject

null, so null hypothesis can be rejected.

Therefore, the average time to pay suppliers is not same in years 2010 and 2014

## Statistical Techniques



Hypothesis For Average Time to Pay Suppliers-

Ho: No difference in time to pay across 2012, 2013 and 2014

Ha: Significant difference in time to pay across 2012, 2013 and 2014

We want to test whether the time to pay suppliers is same in the

years 2012, 2013 and 2014

We performed ANOVA: Single Factor for the same. As F > F critical and p value is less than 0.05, we have

enough evidence to reject null, so null hypothesis can be rejected.

Therefore, the average time to pay suppliers is not same in years 2012, 2013 and 2014

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anova: Single Factor |  |  |  |  |  |  |
| SUMMARY |  |  |  |  |  |  |
| Groups | Count | Sum | Average | Variance |  |  |
| 2012 | 12 | 86.89 | 7.240833333 | 0.000681061 |  |  |
| 2013 | 12 | 86.21 | 7.184166667 | 0.00117197 |  |  |
| 2014 | 12 | 85.12 | 7.093333333 | 0.000624242 |  |  |
| ANOVA |  |  |  |  |  |  |
| Source of Variation | SS | df | MS | F | P-value | F crit |
| Between Groups | 0.132872222 | 2 | 0.066436111 | 80.45474006 | 2.04219E-13 | 3.2849176  51 |
| Within Groups | 0.02725 | 33 | 0.000825758 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 0.160122222 | 35 |  |  |  |  |

## Statistical Techniques



**Relation Between Time to pay Suppliers and On Time Delivery**

We perform regression analysis for On Time Delivery i.e., Percentage of Number of Deliveries on Time (Dependent Variable(Y)) and Time to Pay Suppliers i.e., Working Days in which payment is made to suppliers (Independent Variable(X)).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |
| Regression Statistics | |  |  |  |  |  |
| Multiple R | 0.533734082 |  |  |  |  |  |
| R Square | 0.28487207 |  |  |  |  |  |
| Adjusted R Square | 0.272542278 |  |  |  |  |  |
| Standard Error | 0.002766809 |  |  |  |  |  |
| Observations | 60 |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
|  | df | SS | MS | F | Significance F |  |
| Regression | 1 | 0.000176869 | 0.000176869 | 23.1043697 | 1.12599E-05 |  |
| Residual | 58 | 0.000444003 | 7.65523E-06 |  |  |  |
| Total | 59 | 0.000620873 |  |  |  |  |
|  | Coefficients | Standard Error | t Stat | P-value | Lower 95% | Upper 95% |
| Intercept | 1.015868251 | 0.005854944 | 173.5060689 | 1.7934E-80 | 1.004148308 | 1.02758819  4 |
| Working Days | -0.003766077 | 0.000783506 | - 4.806700501 | 1.12599E-  05 | -0.005334434 | -0.00219772 |

In the model, we get R² = 0.28487 and Adjusted R² = 0.27254 **Regression Model = 1.015868251 + -0.003766077\*Working Days**

From this Regression model we expect that the ontime deliveries should increase as the time to pay suppliers would decrease.

This is also visible from the regression analysis carried out, as the coefficient of working days is negative, this implies as the working days (in which payment is made) would increase on time delivery would decrease, and vice versa

**Moreover, with this data we can predict that If the time to**

|  |  |  |
| --- | --- | --- |
|  | Working Days | Percent |
| Working Days | 1 |  |
| Percent | -0.533734082 | 1 |

**pay suppliers would decrease from the current average of**

**7.45 to 6 working days, the percentage of on-time**

**deliveries could increase to 99.33%.**

From the analysis it is found that:

## Result and Discussion:



* + The time to pay suppliers is not same in 2010-2014 and also not same for years 2012, 2013 and 2014.
  + The standard deviation in time to pay suppliers in different years is also not much significant. And it is lowest in 2014.
  + From the line chart (provided earlier), it is visible that there is a significant increase in time to pay suppliers in between the year 2010-2011 but then the time gradually decreased over the years (2011- 2014)
  + The analysis of both regression and correlation indicates that time to pay suppliers is inversely related with on-time delivery (as the sign of correlation as well as coefficient of working days in which payment is made is negative). Hence it is probable that, more will be the time taken by PLE to make payment to suppliers, more would be the delay in delivery.
  + Moreover, with this data we can predict that If the time to pay suppliers would decrease from the current average of 7.45 to 6 working days, the percentage of on-time deliveries could increase to 99.33%.

## Suggestion:

The average time to pay suppliers is less than 8 days. Moreover, the standard deviation is not much high in the data, i.e. there is no significant fluctuation in time to pay suppliers. This can generate a further confidence among suppliers for PLE. Alongside, PLE can further improve its time to pay suppliers by restructuring its financial policy.



THANK YOU