**Data Analysis Report** 

**Techno Pulse Electronics** 

Name: - Atharva Santosh Mavale

Summary: -

This report analyzes data from Techno Pulse Electronics (TPE), mainly focused on warehouses with the sales and customer ratings. Furthermore, we summarized key variables using descriptive statistics. We did find outliers in the sales data and proportion of customer ratings. We also analyzed the relationship between Manager Gender and Customer Rating. Moreover, annual sales by males manager in the warehouse were calculated. Finally, the sample size needed analyzing the

proportion of warehouse have high customer revies with 95% confident level.

Introduction: -

The main goal is to provide meaningful insights of Techno Pulse Electronics (TPE), as I have used many tools to explore and summarize the variables like sales, customer ratings etc. These insights will Techno Pulse Electronics (TPE) for future predictions and as I have used the data of 150 warehouses and used them to answer specific business questions that will help and guide company decisions.

Q1. Summaries of key variables of interest

a) "Sales" - Annual Sales in \$'000.

Report

Based on the sample of sales, approximately 50% of the Sales are below \$240.83 and the average sales are \$253.47. However, most of the sales are between \$214.1 and \$282.08. As there is a small group with very high sales amounts approximately \$450 or more. To conclude, as most sales (44%) of its frequency occur around \$250.

b) "Cus\_Rating" - The typical Customer review rating

# Report

- Based on the sample size there are 150 customers ratings in total.
- Distribution of Ratings explain the frequency counts of Cus\_Rating ("Medium," "High,"
   "Low")
- As most of the customers gave a **Medium Rating**, making up 54% of all ratings. This explains most people gave medium rating to warehouses.
- Approximately 31% of the total proportion of High ratings as this explain there are some High Reviews but not as common as Medium Ratings.
- 23 ratings (15%) are the least ratings given by the customer; this seems some customers are unhappy with warehouse services.

# Q2. Exploring relationships between two variables

a) I would like to know if there is a link between the typical customer review rating ("Cus\_Rating") and Managers gender ("Mgr\_Gender"). I suspect that the customer review rating may differ based on the location managers gender, but I'd like to know if this is actually the case. Therefore, I'd like you to establish from your sample data if there is any relationship between these two variables.

# Report

- As data collected from 150 Warehouses, we observed that is there any relationship between Customer reviews and Manager gender. However, customer reviews are being categorized in "High", "Medium", "Low".
- As Chi-square test was conducted to find the relationship between two categorial variable and test produced **Chi-Square statistic of 0.1377**.
- P value was 0.14 as this was greater than 0.05, we reject null hypothesis. Hence this
  indicates that there was no significant association between Manager gender and
  customer ratings.
- As the data show no significant association between customer and Manager genders is Male or Female.
- However, 83% of manager are female is fair to understand that rating distribution between two genders is similar.
- b) I 'm also interested to establish if there is a relationship between the number of items sent from warehouses ("Items\_Sent") and the number of team members operating at that warehouse location ("T\_Members").

# Report

- This Dataset includes 150 samples, and we observed the relationship between ("Items\_Sent") and ("T\_Members") using scatter plot as plot show no clear relationship between number of items sent and team member.
- Then I tried to find **correlations** between the items 0.013 which is close to **zero** as this also suggests that there was no relationship between the two variables.
- Then I applied linear regression to check if their possibility this variable related, but from the analysis we can conclude that there was no significant relationship between ("Items\_Sent") and ("T\_Members") as proof the very low correlation and R-squared value explains that team size does not affect number of items shipped.
- c) Further, it would be helpful if we knew if the annual sales of the warehouse ("Sales") has any relationship with the length of time this store has been operating with a web presence ("Length").

### Report

- As we analyzed the relationship between length of time and annual sales of warehouses with help correlation, I tried to find the relationship between two variables is approximately -0.159.
- This suggests that very weak relationships, as there would be likely warehouse length would affect warehouse sales.
- With the help of the linear regression R square value (0.0256) which is around 2.5% as this also indicates lengths as no effect on sales performance.
- Conclusion, this analysis gives insights that other factors are more important for warehouse sales then determining the length of the warehouse.

# Q3. Estimating warehouse measures

a) I would like you to estimate the average annual sales of all warehouses managed by Males.

#### Report

- The goal was to find the average annual sales of all warehouses managed by males (25 males in total of all warehouses)
- Used Average function to find Average sales in \$'000 by males is \$242.7
- b) I 'm also interested to know if you can estimate the proportion of all our stores that have a customer review rating of "High".

# Report

- To estimate the proportion of all stores with high ratings as Number of stores is 150, number of stores with "High Rating" is 46 and proportion of stores with high ratings is 31%.
- $Proportion = (Number\ of\ stores\ with\ High\ ratings\ \div\ Total\ Number\ of\ stores)X100$
- Out of the 150 warehouses, 46 received 'High Ratings' as this express it's only one third
  of all stores

# Q4. Claims about warehouses

a) I read somewhere that the industry average wage for warehouse staff is \$65,000. I think that the average wage is higher than this figure for TPE warehouse staff. Is there any evidence to suggest that this is the case?

### Report

- We are determining that the average wage for staff is higher than the wages of other industries average wages which is \$65,000.
- By using the Hypothesis Test for finding average wages null (H0) = 65,000 we are checking here that wage is higher than the given value.
- T-statics is 2.7613 which is greater than critical value (1.6551) and p value (0.003) is less than 0.05 so we reject the null hypothesis.
- Hence, on this hypothesis test there is calculated evidence which supports the claim that average TPE of warehouse staff is greater than \$65000 and P-value (0.003) proof the result.
- b) Another claim based on industry observations was that fewer than 22% of warehouse managers were male. Can you also check this claim against your survey data?

# Report

- Our objective is here to claim 22% of warehouse manager are male there are total of 150 managers.
- There are 25 male warehouse managers out of total 150 warehouse managers. This gives sample of proportion of 0.17 or 17%.
- As industry observation claim to 22% of manager are male and we tested this using Hypothesis Test. As the standard error was calculated 0.03 and z-score for the sample proportion was -1.56 as this indicates that proportion of manager is below the hypothesized proportion (0.22).
- Our p-value is 0.057, as this mean there is 5.7 probability that sample proportions is lower than the true proportion of male managers is 22%. We fail to reject the null hypothesis.
- Hence, we don't have enough evidence to claim that fewer than 22% of warehouse managers are male.

### Q5. Appropriate sample size

Finally, I am concerned that the sample of 150 is too small to provide accurate results as this seems hardly enough data. If we ever decide to repeat the analysis, I would like to be able to

calculate approximately the true proportion of warehouses that have a high customer review rating ("Cus\_Rating") to within 5%. Therefore, what sample size would we need to include in the next analysis to satisfy this requirement?

### Report: -

- Based on the data 150 warehouse, approximately 31% of warehouses have high review ratings.
- To calculate the true proportion of warehouse of the high review rating with margin error of 5% at a 95% confidence level, our Z score is 1.959.
- However, 31% of proportion of warehouses have high ratings as this shows a significant portion of warehouses have high customer satisfaction.
- To ensure the margin error of 5% the sample size of the warehouse needs to be 327.
- Conclusion, with the current 150 warehouse data we will not be able to find or capture variation accurately. Hence, we need larger sample size

#### Conclusion: -

Finally, this report will provide useful information about TechnoPulse Electronics (TPE) to improve their performance in future. This information analyzes the help company to make decisions about staff, warehouse operation and guides them to create growth strategies. As I have analyzed the data and provided the suggestions that will help TechnoPulse Electronics (TPE) to improve their experience with customers.