

jOB SATISFACTION

DATA ANALYSIS

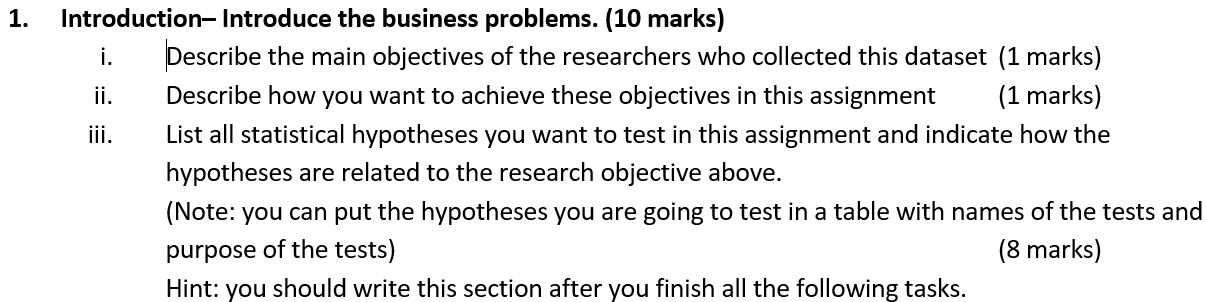


VIJAY KASOTIA

email: vk73742@gmail.com

**Job Satisfaction**

**Tasks:**

****

Answer 1- (i): The objective of the study is that the how job is satisfied for employees in a large company that operates in many different countries/regions (such as east, west, south, north). Also, how life happiness score talks about the job satisfaction, etc.

(ii): After collecting the data set regards their job satisfaction, stylish action plan can make it easier to achieve these objectives.

(iii): List of statistical hypotheses shown below:

1. Null Hypothesis: There is no difference in Job satisfaction after training at any level of promotion.

Alternative Hypothesis: There is a difference in Job satisfaction after training at any level of promotion.

Test: Chi-Square Test

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Observed** | **Expectation** | **O-E** | **(O-E)^2** | **((O-E)^2)/E** |
| 35 | 24.84 | 10.1600 | 103.2256 | 4.1556 |
| 68 | 52.44 | 15.5600 | 242.1136 | 4.6170 |
| 83 | 84.87 | -1.8700 | 3.4969 | 0.0412 |
| 21 | 37.95 | -16.9500 | 287.3025 | 7.5706 |
| 0 | 6.90 | -6.9000 | 47.6100 | 6.9000 |
| 1 | 11.16 | -10.1600 | 103.2256 | 9.2496 |
| 8 | 23.56 | -15.5600 | 242.1136 | 10.2765 |
| 40 | 38.13 | 1.8700 | 3.4969 | 0.0917 |
| 34 | 17.05 | 16.9500 | 287.3025 | 16.8506 |
| 10 | 3.10 | 6.9000 | 47.6100 | 15.3581 |
|  |  |  | **=Chi-Square** | **75.111** |
|  |  |  | **=p-value** | **0.000** |
|  |  |  | |  |  | | --- | --- | | **=Chi-sq (*Χ2*) value** |  | | **9.488** |

The Chi-square (calculated) value is greater than the Chi-square tabulated value, so we can reject the null hypothesis. On the other hand, the p-value is close to zero which indicates that it is less than the level of significance. So, based on the above evidence, we can reject the H0, and support alternative hypothesis. Thus, we can conclude that there is a difference in Job satisfaction after training at any level of promotion.

1. Null Hypothesis: Salary is related to life happiness is truly equal to zero

Alternative Hypothesis: Salary is related to life happiness is truly different from zero

Test: One-Way ANOVA

We use ANOVA to test if there is a statistically significant difference in salary with respect to life happiness. Salary will serve as the dependent variable, and life happiness will act as the independent variable.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **One factor ANOVA** | | | | | |
|  | *Mean* | *n* | *Std. Dev* |  |  |
|  | 52.4 | 300 | 9.13 | Salary(000) |  |
|  | 5.6 | 300 | 1.82 | Life Happiness Score (1-10) |  |
|  | 29.0 | 600 | 24.33 | Total |  |
|  |  |  |  |  |  |
| ANOVA table |  |  |  |  |  |
| *Source* | *SS* | *df* | *MS* | *F* | *p-value* |
| Treatment | 3,28,723.23 | 1 | 3,28,723.227 | 7593.75 | 0.0000 |
| Error | 25,886.61 | 598 | 43.289 |  |  |
| Total | 3,54,609.83 | 599 |  |  |  |

We conclude that the mean salary is statistically significantly different for at least one of the life happiness scores (F1, 598= 7593.75, p < 0.05).

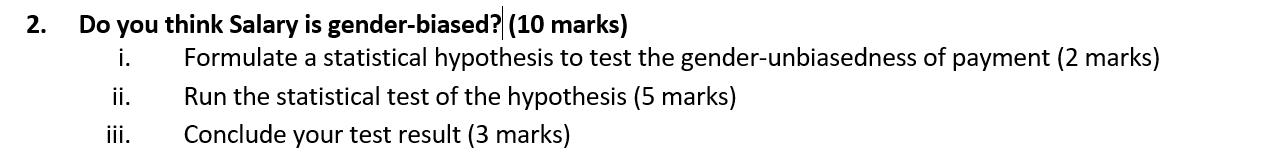
1. H0: There is no difference in between the job satisfaction before training and job satisfaction after training.

H1: There is a difference in between the job satisfaction before training and job satisfaction after training.

Test: Chi-Square Test



The Chi-square (calculated) value is less than the Chi-square tabulated value, so we cannot reject the null hypothesis. On the other hand, the p-value is greater which indicates that it is more than the level of significance. So, based on the above evidence, we cannot reject the H0, and does not support alternative hypothesis. Thus, we can conclude that there is no difference between the Job satisfaction before training and after training.



Answer 2: (i)- Set up hypothesis:

Null Hypothesis: Salary is gender biased i.e., truly equal to zero.

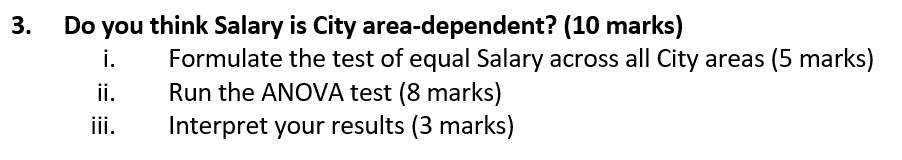
Alternative Hypothesis: Salary is gender un-biased i.e., truly not equal to zero.

(ii): Test: Independent Groups (t-test)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Hypothesis Test: Independent Groups (t-test, pooled variance) | | | | | | |
|  |  |  |  |  |  |  |
|  | Gender | Salary(000) |  |  |  |  |
|  | 1.62 | 52.42 | mean |  |  |  |
|  | 0.49 | 9.13 | std. dev. |  |  |  |
|  | 300 | 300 | n |  |  |  |
|  |  |  |  |  |  |  |
|  |  | 598 | df |  |  |  |
|  |  | -50.803 | difference (Gender - Salary(000)) | | | |
|  |  | 41.759 | pooled variance | |  |  |
|  |  | 6.462 | pooled std. dev. | |  |  |
|  |  | 0.528 | standard error of difference | | |  |
|  |  | 0 | hypothesized difference | | |  |
|  |  |  |  |  |  |  |
|  |  | -96.286 | t |  |  |  |
|  |  | 0.0000 | p-value (two-tailed) | |  |  |

(iii): We can conclude the result based on the p-value. Thus, the p-value is close to zero which indicates that it is less than the level of significance (i.e., 0.05 > 0.00). Therefore, we can reject the null hypothesis and support the alternative hypothesis.

Further, in the sample data, we will use two variables: gender and salary. The variable gender has values of either “1” (male) or "2" (female). It will function as the independent variable in this T test. The variable salary is a numeric variable, and it will function as the dependent variable. The negative t value tells than the mean of female is less than the male. Hence, we conclude that the variance in salary of males is statistically significantly different than that of females.



Answer (i): Null Hypothesis: There is no difference between Salary across all City areas i.e., is truly equal to zero.

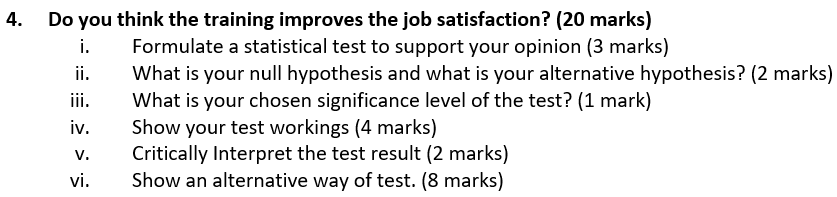
Alternative Hypothesis: There is a difference between Salary across all City areas i.e., is truly equal to zero.

(ii) ANOVA Test:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| One factor ANOVA |  |  |  |  |  |
|  |  |  |  |  |  |
|  | *Mean* | *n* | *Std. Dev* |  |  |
|  | 2.7 | 300 | 1.40 | City Area |  |
|  | 52.4 | 300 | 9.13 | Salary(000) |  |
|  | 27.6 | 600 | 25.71 | Total |  |
|  |  |  |  |  |  |
| ANOVA table |  |  |  |  |  |
| *Source* | *SS* | *df* | *MS* | *F* | *p-value* |
| Treatment | 3,70,463.80 | 1 | 3,70,463.802 | 8692.23 | 0.0000 |
| Error | 25,486.82 | 598 | 42.620 |  |  |
| Total | 3,95,950.63 | 599 |  |  |  |

(iii) We use ANOVA to test if there is a statistically significant difference in salary with respect to city areas (from 1 to 5). Salary will serve as the dependent variable, and city area will act as the independent variable.

The p-value of the test is 0.000 which is less than the level of the significance i.e., 0.00 < 0.05. Therefore, based on the above evidence, we can reject the null hypothesis and support the alternative hypothesis. Thus. we conclude that the mean salary is statistically significantly different for at least one of the city areas (F1, 598= 8692.23, p < 0.05).



Answer (i): Without any statistical test, we cannot tell that the training improves the job satisfaction or not. However, the given statement may true, because training enhance the level of skill which result job satisfaction level is higher. The statistical test will **One sample T-test for Proportion** use.

(ii). Null Hypothesis: there is no difference in job satisfaction after training.

Alternative Hypothesis: there is a difference in job satisfaction after training.

(iii) The chosen significance level of the test or alpha is 5% or 0.05.

(iv): Test Working:

|  |  |
| --- | --- |
| Hypothesis Test: Mean vs. Hypothesized Value | |
| 0.000 | hypothesized value |
| 3.243 | mean Job Satisfaction Score after training(1-5) |
| 0.997 | std. dev. |
| 0.058 | std. error |
| 300 | n |
| 299 | df |
|  |  |
| 56.343 | t |
| 0.0000 | p-value (two-tailed) |

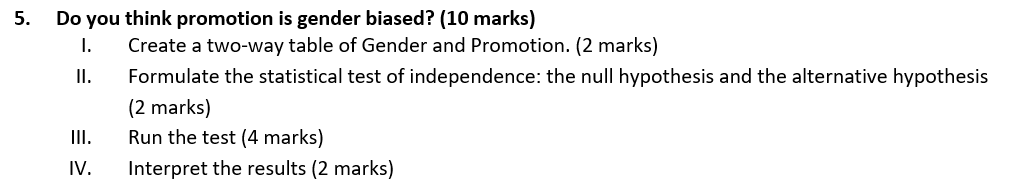
(v) The p-value of the test is 0.000 which is less than the level of the significance i.e., 0.00 < 0.05. Therefore, based on the above evidence, we can reject the null hypothesis and support the alternative hypothesis. Thus. we conclude that the training improves the job satisfaction. Additionally, in part (i), the judgmental statement or opinion is true.

(vi) Alternative way: One sample Test

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **One-Sample Statistics** | | | | |
|  | N | Mean | Std. Deviation | Std. Error Mean |
| Job Satisfaction Score after training (1-5) | 300 | 3.24 | .997 | .058 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **One-Sample Test** | | | | | | |
|  | Test Value = 0 | | | | | |
| t | df | Sig. (2-tailed) | Mean Difference | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| Job Satisfaction Score after training (1-5) | 56.343 | 299 | .000 | 3.243 | 3.13 | 3.36 |

The p-value of the test is 0.000 which is less than the level of the significance i.e., 0.00 < 0.05. Therefore, based on the above evidence, we can reject the null hypothesis and support the alternative hypothesis.



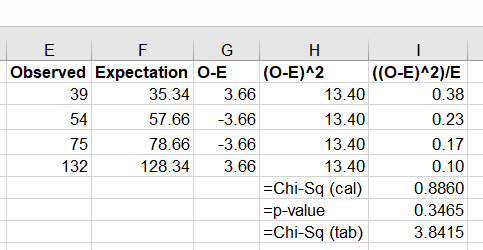
Answer 5: (I): Two-way Table of Gender and Promotion

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gender \* Promoted Crosstabulation** | | | | | |
|  | | | Promoted | | Total |
| No | Yes |
| Gender | 1 | Count | 39 | 75 | 114 |
|  |  |  |  |
| 2 | Count | 54 | 132 | 186 |
|  |  |  |  |
| Total | | Count | 93 | 207 | 300 |
|  |  |  |  |

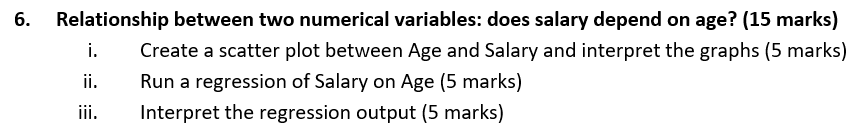
II. Null Hypothesis: There is no difference in between Gender and Promotion

Alternative Hypothesis: There is a difference in between Gender and Promotion

III. Test: Chi-Square Test (Two Categorical Variables)



IV. The P-Value is 0.3465. The chosen of the level of significance is 0.05. Thus, the p-value is higher than the 0.05. Therefore, the result is not statistically significant at p < 0.05. Based on the above evidence, we cannot reject the null hypothesis and in favour of alternative hypothesis. On the other hand, the Chi-square calculated value (= 0.8860) is less than the Chi-square tabulated value (=3.8415), so we cannot reject the null hypothesis. Hence, there is no difference between the gender and promotion.



Answer (i): Simple scatter plot Age by Salary

ii. Regression Test Output:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |
| Multiple R | 0.7551 |  |  |  |  |  |
| R Square | 0.5702 |  |  |  |  |  |
| Adjusted R Square | 0.5688 |  |  |  |  |  |
| Standard Error | 5.9928 |  |  |  |  |  |
| Observations | 300 |  |  |  |  |  |
|  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 1 | 14198.8812 | 14198.8812 | 395.3584 | 0.0000 |  |
| Residual | 298 | 10702.3555 | 35.9139 |  |  |  |
| Total | 299 | 24901.2367 |  |  |  |  |
|  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | 24.0706 | 1.4673 | 16.4045 | 0.0000 | 21.1830 | 26.9582 |
| Age | 0.6770 | 0.0340 | 19.8836 | 0.0000 | 0.6100 | 0.7440 |

(iii). Regression Equation:

Interpretation of the regression equation:

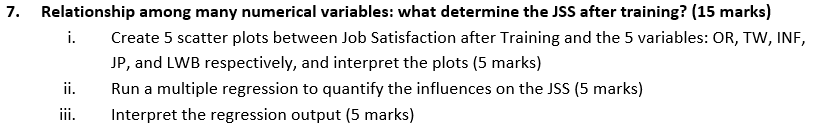
Intercept, , the salary intercept implies that when the value of age is zero, then the mean value of salary is expected to increase by 24.07.

Slope, = 0.68, the slope coefficient implies that for each increase of 1 year in age, then the value of salary is estimated to increase by 0.68. Additionally, the age variable is statistically significant at 5% level of significance.

Model Diagnose:

The R-square of the linear model is 0.57 (or 57%), it indicates that the 57% of the variation in the salary variable can be explained by the variation in the age variable. Thus, the model is moderately good fit with the data.

The p-value of the linear model is close to zero, which implies that it is less than the chosen level of significance (5%). Thus, the model is statistically significant.



Answer (i) A) Scatter plots JSS vs OR

The scatter dots moving upward direction, indicates that there is a positive but weak correlation (i.e., 0.39) between the job satisfaction after training and organization relation and employee satisfaction.

B) Scatter plots JSS vs TW

The scatter dots moving upward direction, indicates that there is a positive but weak correlation (i.e., 0.40) between the job satisfaction after training & the teamwork and employee satisfaction.

C) Scatter plots JSS vs INF

The scatter dots moving upward direction, indicates that there is a positive but moderately strong correlation (i.e., 0.54) between the job satisfaction after training & the information and employee satisfaction.

D) Scatter plots JSS vs JP

The scatter dots moving upward direction, indicates that there is a positive but moderately strong correlation (i.e., 0.50) between the job satisfaction after training & the Job passion and self-evaluation employee satisfaction.

E) Scatter plots JSS vs WLB

The scatter dots moving upward direction, indicates that there is a positive but moderately strong correlation (i.e., 0.502) between the job satisfaction after training & the work/Life balance and employee satisfaction.

(iii). Multiple regression Output:



Regression Equation:

Interpretation of the multiple regression output:

* Slope, = 0.07, the slope coefficient implies that for each increase of 1 good organization relation, then the value of JSS is estimated to increase by 0.07. Additionally, the OR variable is not significant at 5% level of significance.
* Slope, = 0.33, the slope coefficient implies that for each increase of 1 good teamwork, then the value of JSS is estimated to increase by 0.33. Additionally, the TW variable is statistically significant at 5% level of significance.
* Similarly, we can interpret the rest of the beta coefficients.

Model Diagnose:

* The R-square of the linear model is 0.60 (or 57%), it indicates that the 60% of the variation in the JSS (dependent) variable can be explained by the variation all independent variable. Thus, the model is moderately good fit with the data.
* The p-value of the linear model is close to zero, which implies that it is less than the chosen level of significance (5%). Thus, the model is statistically significant.

# **Reference**

Heiberger, R.M. and Neuwirth, E., 2009. One-way anova. In R through excel (pp. 165-191). Springer, New York, NY. <https://link.springer.com/chapter/10.1007/978-1-4419-0052-4_7>

Kim, H.Y., 2014. Analysis of variance (ANOVA) comparing means of more than two groups. Restorative dentistry & endodontics, 39(1), pp.74-77. <https://synapse.koreamed.org/articles/1090099>

McHugh, M.L., 2011. Multiple comparison analysis testing in ANOVA. Biochemia medica, 21(3), pp.203-209. <https://hrcak.srce.hr/clanak/108599>

Seber, G.A. and Lee, A.J., 2012. Linear regression analysis. John Wiley & Sons. <https://books.google.com/books?hl=en&lr=&id=X2Y6OkXl8ysC&oi=fnd&pg=PR5&dq=linear+regression&ots=semQC1oQmq&sig=LAodlmAANdyOwHRR1toKXoHYUHk>

Tranmer, M. and Elliot, M., 2008. Multiple linear regression. The Cathie Marsh Centre for Census and Survey Research (CCSR), 5(5), pp.1-5. <http://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/multiple-linear-regression.pdf>