DATA ANALYSIS OF JOB SATISFACTIONS ON YEARS OF EXPERIENCE

Aung Cho[1], Khin Khin Lay[2]
[1]Application Department, [2]Faculty of Computer Science,
[1,2]University of Computer Studies, Maubin, Myanmar
Email: am6244052@gmail.com

ABSTRACT

SPSS (Statistical Package for the Social Sciences) is powerful to analyze business and marketing data. This paper intends to support business and marketing leaders the benefits of data analyzing with applied SPSS. It showed the data analysis of job satisfactions on years of experience. As SPSS's background algorithms, it showed the cross tabulation algorithm for cross tabulation table and Pearson chi-square algorithm for data significant. And then Sample data 'demo.sav' was free downloaded from Google and was analyzed and viewed. It used IBM SPSS statistics version 23 and PYTHON version 3.7.

KEYWORDS: *Statistical Package for the Social*

Sciences

I. INTRODUCTION

Nowadays, businesses are competing with others not to lose their market places in. local and external regions. To avoid the loss of market places they should use data science technology. This paper used SPSS integrated with Python software. It showed the data analysis of job satisfactions on years of experience that includes four tables, the result of each table, two chi-square test for data significant and data analytical view.

A. SPSS standing for Statistical Package for the Social Sciences, is a powerful, user- friendly software package for the manipulation and statistical analysis of data.

The package is particularly useful for students and researchers in psychology, sociology, psychiatry, and other behavioral sciences, containing as it does an extensive range of both univariate and multivariate procedures. Data analysis generally begins with the calculation of a number of summary statistics such as the mean, median, standard deviation, etc., and by creating informative graphical displays of the data such as histograms, box plots, and stem- and-leaf plots.[2]

B. SPSS Crosstabs Cross tabulation tables, or contingency tables are frequently employed to examine the relationship between two variables (usually nominal or ordinal) that have a small number of categories.[4]

C. Chi-square test The Chi-square test allows us to determine whether or not there is a statistically significant association between two variables. If the variables are not associated they are said to be statistically independent (hence Chi-square is often referred to as the 'Chi-square test of independence').[4]

D. SPSS Graph SPSS can display your data in a bar chart, a line graph, an area graph, a pie chart, a scatterplot, a histogram, a collection of high-low indicators, a box plot, or a dual-axis graph. Adding to the flexibility, each of these basic forms can have

multiple appearances. For example, a bar chart can have a two- or three-dimensional appearance, represent data in different colors, or contain simple lines or I-beams for bars. The choice of layouts is almost endless.[3]

II. ALGORITHM

A. Notation(cross tabulation algorithm) [1]

Notation	Description
X_i	Distinct values of row variable arranged in ascending order: $X_1 < X_2 < ... X_R$
Y_j	Distinct values of row variable arranged in ascending order: $Y_1 < Y_1 < ... Y_C$
f_{ij}	Sum of cell weights for cases in cell(i,j)
c_j	$\sum_{i=1}^{R} fij$ the jth column subtotal
r_i	$\sum_{j=1}^{C} fij$ the ith row subtotal
w	$\sum_{j=1}^{C} fij = \sum_{i=1}^{R} fij$ the grand total

B. Marginal and Cell Statistics [1]

count $\qquad = f_{ij}$

Expected Count $\quad E_{ij} = \dfrac{r_i c_j}{W}$

row percent $\quad = 100 \times (f_{ij} / r_i)$

column percent $\quad = 100 \times (f_{ij} / rC_j)$

total percent $\quad = 100 \times (f_{ij} / W)$

C. cross tabulation algorithm [1]

Residual $\qquad R_{ij} = f_{ij} - E_{ij}$

Standardized Residual $\quad SR_{ij} = \dfrac{R_{ij}}{\sqrt{E_{ij}}}$

Adjusted Residual $\quad AR_{ij} = \dfrac{R_{ij}}{\sqrt{E_{ij}\left(1 - \dfrac{r_i}{W}\right)\left(1 - \dfrac{C_j}{W}\right)}}$

D. Pearson's Chi-square(cross tabulation algorithm) [1]

$$\chi_p^2 = \sum_{ij} \dfrac{(f_{ij} - E_{ij})^2}{E_{ij}}$$

The degrees of freedom are $(R - 1)(C - 1)$.

Likelihood Ratio

$$\chi_{LR}^2 = 2 \sum_{ij} f_{ij} \ln (f_{ij}/E_{ij})$$

The degrees of freedom are $(R - 1)(C - 1)$.

Note: when $f_{ij} = 0$, the entire term $f_{ij} \ln (f_{ij}/E_{ij})$ is treated as 0, because $\lim_{n \to 0} n \log (n) = 0$, and thus has no effect on the sum.

Fisher's Exact Test

If the table is a 2×2 table, not resulting from a larger table with missing cells, then the Fisher exact test is calculated. For more information, see the topic "Significance Levels for Fisher's Exact Test" on p. 1037.

Yates Continuity Corrected for 2 x 2 Tables

$$\chi_c^2 = \begin{cases} \dfrac{W(|f_{11}f_{22} - f_{12}f_{21}| - 0.5W)^2}{r_1 r_2 c_1 c_2} & \text{if} |f_{11}f_{22} - f_{12}f_{21}| > 0.5W \\ 0 & \text{otherwise} \end{cases}$$

The degrees of freedom are 1.

III. ANALYSIS

A. Job Satisfactions on Years of Experience

SPSS Code

CROSSTABS /TABLES=empcat BY jobsat /FORMAT=AVALUE TABLES /STATISTICS=CHISQ /CELLS=COLUMN BPROP /COUNT ROUND CELL /BARCHART.
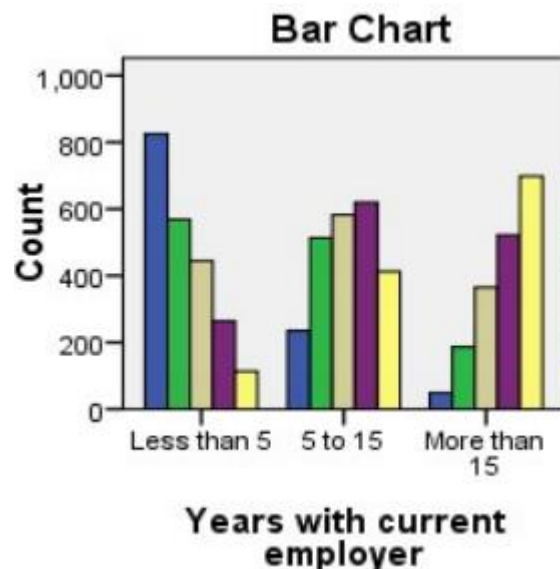
TABLE.I YEARS WITH CURRENT EMPLOYER* JOB SATISFACTION CROSS TABULATION

Years of Experience	Highly satisfied
Less than 5	9.2%$_e$
5 to 15	33.8%$_c$
More than 15	57.0%$_e$
Total	100.0%

TABLE.II CHI-SQUARE TESTS

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1689.561[a]	8	.000
Likelihood Ratio	1747.380	8	.000
Linear-by-Linear Association	1525.767	1	.000
N of Valid Cases	6400		

Bar Chart-1: Years with current employer*job satisfaction

B. Data analytical view:

As a result of table-1 and bar chart-1, more years of experience, higher job satisfaction. Years of experience and job satisfaction are direct correlation and real association.

As a result of table-2, p-value is less than 0.05, so variables are significantly variance. In the business, employees with more experience should share the knowledge and happiness to employees with lower experience. Business leader and owners should award promotion and bonus to all tried employees in the end of the year to get more job satisfactions.

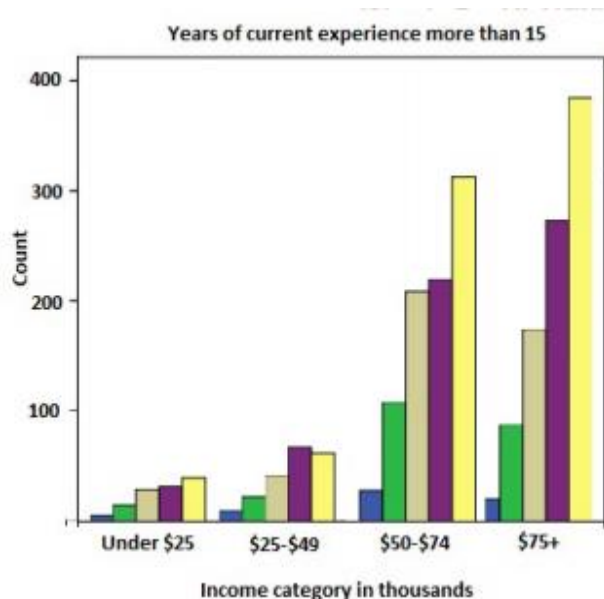### C. Job Satisfactions on Most Years of Experience and income level

If the control variable 'income level' was added on test 3.1 condition, the result was found as the following.

TABLE.III MOST YEARS WITH CURRENT EMPLOYER* JOB SATISFACTION *INCOME LEVEL

CROSS TABULATION More than 15 Years

Income Level	Highly satisfied
Under $25	5.6%
$25-$49	8.7%
$50-$74	16.9%
$75+	68.8%
Total	100.0%

Bar Chart-2: Years with current employer*job satisfaction*income level

Years of current experience more than 15

Job satisfaction
■ Highly dissatisfied
■ Somewhat dissatisfied
□ Neutral
■ Somewhat satisfied
□ Highly satisfied

TABLE.IV CHI-SQUARE TESTS

Chi-Square Tests

Income category in thousands		Value	df	Asymptotic Significance (2-sided)
Under $25	Pearson Chi-Square	243.267[b]	8	.000
	Likelihood Ratio	235.182	8	.000
	Linear-by-Linear Association	213.329	1	.000
	N of Valid Cases	1174		
$25 - $49	Pearson Chi-Square	379.849[c]	8	.000
	Likelihood Ratio	386.484	8	.000
	Linear-by-Linear Association	350.891	1	.000
	N of Valid Cases	2388		
$50 - $74	Pearson Chi-Square	178.911[d]	8	.000
	Likelihood Ratio	169.772	8	.000
	Linear-by-Linear Association	144.894	1	.000
	N of Valid Cases	1120		
$75+	Pearson Chi-Square	321.073[e]	8	.000
	Likelihood Ratio	216.629	8	.000
	Linear-by-Linear Association	190.037	1	.000
	N of Valid Cases	1718		
Total	Pearson Chi-Square	1689.561[a]	8	.000
	Likelihood Ratio	1747.380	8	.000
	Linear-by-Linear Association	1525.767	1	.000
	N of Valid Cases	6400		

D. Data analytical view:

As a result of table-3 and bar chart-2, higher income, higher job satisfaction. Income level and job satisfaction are direct correlation and real association.

As test-3.1 and test-3.2, if there is more job satisfactions, employers have more experience and more income.

As a result of table-4, p-value is less than 0.05, so variables are significantly variance.

IV. CONCLUSION

SPSS data analysis tools are valuable in social science, business and marketing fields. It is very good for presentation report by graphical design. Business leaders can get their goal with good result and can avoid the loss of market places in local and global regions by using SPSS software.

REFERENCES
[1] IBM SPSS Statistics 24 Algorithms pdf book [book style]
[2] A handbook of statistical analyses using SPSS / Sabine, Landau, Brian S. Everitt, ISBN 1-58488-369-3 [book style]
[3] SPSS For Dummies®, 2nd Edition, ISBN: 978-0-470- 48764-8 [book style]
[4] SPSS for Social Scientists Robert L. Miller, Ciaran Acton, Deirdre A. Fullerton and John Maltby, ISBN 0–333– 92286–7 [book style]