ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 5, Issue 3, September 2015

Reduction of Attrition Rate in a BPO Company Using Design of Experiments (DOE)

Dr. P Beaulah Soundarabai
Associate professor, Computer Science Department, Christ University, Bangalore – 29
Aneesh Das
Student – BSc (CMS), Computer Science Department, Christ University, Bangalore – 29

Abstract:-Today's corporate world is driven by the thought that more they contract ability, more will be their increment in income & billings from customers. Be that as it may, the real issue that the human asset branch of an organization countenances is the whittling down for different reasons. Keeping in mind the end goal to address that issue we have done an investigation utilizing the configuration of tests ,anova& populace research for a specific sort of organization in respect to how we can contract an impeccable applicant, who will be better ability as well as, will likewise stay with the firm for a more drawn out period accordingly, lessening the attrition rate.

Keywords: **Design, ANOVA, Data– Observational, Experimental, Treatments, Blocks, Yield, Error.**

I. INTRODUCTION

Design of experiment was given by Ronald A Fisher and is concerned with the following:

- Planning of the experiment
- Obtaining relevant information from it regarding statistical hypothesis under study.
- Making the statistical analysis of the data.

Thus, design of experiment can be defined as the logical construction of the experiment in which inference is drawn maybe well defined.

II. OBSERVATIONAL DATA

Observational data is represented by observations on the elementary units in a population or of a sample and are not changed or modified by the investigator. It is difficult to assign cause and affect relationship using observational data. E.g. the data collected from 2014 recruitment of students, viz. the number of students applied, the number of students who passed the interview, the number of people who agreed to join, the number of people who actually joined by passing the training, etc.

III. EXPERIMENTAL DATA

Experimental data is obtained from experimentation and here certain factors can be concluded or modified by varying certain factors in order to determine if they have any effect on the data. It helps in understanding cause and effect relationship. Experimental data result from logically designed experiment and can be used to prove or disprove certain theories. E.g. after experimentation we will come to know if there is any significant variance to neglect a certain course and will try to analyze the reason for that with population research/study.

IV. POPULATION STUDY

In the various fields of healthcare, census a population study is a study of a group of individuals taken from the general population who share a common characteristic, such as age, sex, or health condition. This group may be studied for different reasons, such as their response to a drug or risk of getting a disease. When planning an experiment, you will likely use groups of participants. This lesson explores the types of groups an experimenter can collect data from and the reason why there are different groups. Now, let's change the fact that it's a healthcare project. What if we design an experiment so that we can study different classes of population of how they will react to a certain offer/contract when it is placed in front of them by using ANOVA & population research.

Important Terms & Definitions:

(i) **Experiment**: It is a device or means of getting an answer to the problem under consideration. It can be divided into 2 categories: absolute & comparative. Absolute experiment consists of finding the absolute value of some characteristics whereas comparative experiment are designed to compare the effect of 2 or more objects on the same population characteristic.

(ii) **Treatments**: Various objects of comparison in a comparative experiment are called treatments. In many experiment we try to establish the effect of one or more independent variables on a response (dependent variable). The independent variables are called treatments or factors which are often qualitative in nature. The value of response will reflect the effect of different treatments.

(iii) **Experimental Unit**: The smallest division of the experimental material to which we apply the treatments & on which we make observations on the variable under study is called the experimental unit.

(iv) **Blocks**: In population research, most of the times we divide the whole experimental unit i.e. the population, into relatively homogeneous subgroups orstrata. These are called blocks.

(v) **Yield**: The measurement of variable under study on different experimental units are called yields. E.g. the measurement of samples in different population researches is called yield.

(vi) **Experimental Error**: Consider a large homogeneous population which is divided into

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 5, Issue 3, September 2015

samples of equal shape & size & different treatments are applied to these plots. If the yield from some of these treatments are more than the others then there is a doubt whether this difference is due to treatment effects or due to chance.

Such variations which are due to random or chance or non-assignable factors beyond human control is called experimental error. It arises due to 3 factors:

- The inherent variability in the experiment.
- Material to which treatments are applied.
- Lack of uniformity in the methodology of conducting the experiment.

(vii) Lack of representativeness of the sample to the population under study.

Analysis Of Variance (ANOVA)

ANOVA is a powerful statistical tool which is used when we have 3 or more samples at the same time. The basic purpose of ANOVA is to test the homogeneity of different means. ANOVA was introduced by Professor R.A. Fisher in the field of agriculture.

Variation is inherent in nature & can be divided into 2 types:

(i) Variation due to assignable causes: These can be detected and measured.

(ii) Variation due to chance causes: These cannot be detected & measured.

ANOVA can be defined as estimation of the amount of variation due to each of the independent factors (causes) separately & then comparing these estimates due to assignable factors with the estimate due to chance causes.

The estimates due to chance causes will be called as experimental error.

Assumptions for ANOVA

(i) The observations are independent.

(ii) Parent population from which observations are taken is normal.

(iii) Various treatments & environment effects are additive in nature.

Uses of ANOVA

ANOVA helps us to compare several population means simultaneously and thus helps in saving time & money as compared to experiments of comparing two means at a time.

Note 1:

The origin of ANOVA was in agricultural experiments but it is now applied to design of experiment (DOE) in various fields.

Note 2:

It consists of classifying & cross classifying statistical results.

One Way Classification

Consider 'N' observations yij, i=1,2,…,k; j=1,2,…,niof a random variable Y. These are grouped into 'k' classes of sizes n1,n2,…,nk respectively such that N=Σni (Figure 1). Then the one way classified table of data can be written in the following manner:

Class	Sample Obs	Total	Mean
1	y11,y12,y13,…,y1n1	T1.	Ŷ1.
2	y21,y22,y23,…,y2n2	T2.	Ŷ2.
…	…		
K	yk1,yk2,yk3,…,yknk	Tk.	Ŷ3.
		y..	Ŷ..

(Fig: 1)

The total variation in the observation yij canbe split into the following two components:

(i) The variation between the classes or the variation due to different basis of classification, commonly known as treatments.

(ii) The variation within the classes.

The First type of variation is due to assignable causes which can be detected & controlled and the second type of variation is due to chance causes which are beyond human control. The main objective of ANOVA is to examine if there is a significant difference between the means of different classes.

The mathematical model will consist of:

- Effect of the treatment.
- The error E which is produced by chance.

ANOVA						
Source of Variation	SS	Df	MS	F	P-value	F crit
Between Groups	26790.96	7	3827.28	7.61269	0.000405	2.657197
Within Groups	8044	16	502.75			
Total	34834.96	23				

(Fig. 2)

(Fig. 3)

Paired Sample	BA	Bcom	BSc Others	BSc Comp Sc	BBA	BBM	BCA	Others	TOTAL
Tier 1	2	89	12	10	10	15	30	0	168
Tier 2	3	114	51	33	41	15	50	7	314
Tier 3	6	153	49	74	20	71	46	12	431
TOTAL	11	356	112	117	71	101	126	19	913

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 5, Issue 3, September 2015

Null Hypothesis:
There is no difference in the means of the students recruited from different courses.

Alternative Hypothesis:
There is a significant difference in the means of the students recruited from different courses. Or atleast one of them is different.

Now, for doing analysis of variance, we find the computed data related to the table:

- Correction factor
- Raw sum of squares
- Total sum of squares
- Sum of squares due to treatments
- Error sum of squares
- Mean sum of squares due to treatments
- Mean sum of squares due to errors
- F Calculated

Decision Theory:
If Fcal< or = Fcrit then, accept Null Hypothesis.
Testing:
Fcal = 7.61269
Fcrit = 2.657197
Fcal>FcritSo, we reject null hypothesis (Figure 3).

So, now we know for the fact that there is a significant difference in the recruitment of different course groups or treatments, which needs to solved using population research & study of the mindset of people.

V. MEMBERS

When you direct a trial or overview you gather data from a gathering of individuals. Presently, while 'gathering of individuals' may appear like a sufficient depiction, it is, truth be told, not. We require more a particular term in light of the fact that the measurements we utilize are diverse relying upon gathering we utilize. Be that as it may, don't stress, there's no convoluted procedure to recognizing the gathering of individuals you utilize. The main gathering of individuals is a populace, which is characterized as the complete accumulation to be concentrated on. The second gathering is a specimen, which is characterized as an area of the populace. We take a gander at a few illustrations to help make this a bit clearer.

VI. POPULACE

When you are endeavoring to study a populace, you need to gather data from everybody in that gathering. This makes it to a great degree hard to study populaces. For instance, suppose you are going to ponder:

- All individuals with schizophrenia in the United States (pretty nearly 3.1 million people).
- Californians' perspective on raisins (38 million).
- Immigrants convictions about the U.S's. outside arrangement (no one truly knows what number of on the off chance that you incorporate illicit and legitimate).

So you can see the trouble with mulling over populaces. Littler populaces, similar to junior colleges, are simpler to gather populace information. Notwithstanding, the issue here is whether you have one individual who neglects to contribute, then you don't have a populace.

VII. TEST

Most social specialists understand that acquiring data from each individual in a populace is beside incomprehensible. So as opposed to attempting to gather everybody's data, they gather an example of the populace. In any case, dissimilar to a populace, which is everybody, there are diverse ways you can gather a specimen of a populace. The diverse methods for taking an example are kind of like how there are distinctive approaches to cut a cake. Here is a rundown of the diverse examining systems:

• Random test: every person in the populace has an equivalent shot of being chosen.
• Stratified specimen: a specialist isolates the populace into gatherings in light of attributes, and after that the analyst arbitrarily chooses from every gathering in view of its size.
• Quota test: a scientist intentionally sets a prerequisite to guarantee a specific gathering is spoken to.
• Purposive specimen: a specialist intentionally concentrates on a specific subset of a populace.
• Convenience test: choice of the specimen is taking into account simplicity of availability.

Since now we have officially demonstrated that the decision of courses by the hopefuls are critical to discovering the ideal possibility for the organization we are currently going to manage the mentality of individuals in respect to how a fresher settles on his choices with respect to his career using various surveys and population research.

VIII. CONCLUSIONS

In Tier 1 cities like Bangalore & Chennai (Fig. 4) we focus on students who expect a lot from companies. So, the perfect candidate from Tier 1 cities preferably will be BCom, BCA students who look for quick jobs & hands on training as these courses are believed to be professional in nature. The candidates opting for these two courses prefer to gain work experience before they go for specialization or go for higher education or prefer to climb up the corporate ladder by working at a firm.

Looking at the Tier 2 cities like Mysore, Coimbatore, etc. (Figure 4) we find that it’s usually better to hire BCom professionals, BCA candidates & also BSc candidates preferably BSc Mathematics who are always looking for new projects to work for whereas other combination students go for masters or go for Govt jobs like banking, etc.

Last but not the least comes the Tier 3 towns (Figure 5). The major drive of a BPO organization comes from this pool of students always looking for better opportunities which can be proved by small paired sample sign test. It’s a non-parametric test which

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 5, Issue 3, September 2015

statistically proves that Tier 3 is the best pool to hire from with a lowest affordable possibility of attrition rate, which is also a major area of concern. Even in tier 3 we should focus our campus drives to BCom pass outs, BSc Computer Science for TIG groups or process executive jobs. Looking at the 2014 campus hires data BBM is an add in but, according to the research BBM students tend to move away from the companies to complete their MBA thus increasing our attrition rate.

Paired Sample	BA	BCom	BSc Others	BSc Comp Sc	BBA	BBM	BCA	Others	TOTAL
Tier 1	2	89	12	10	10	15	30	0	168
Tier 2	3	114	51	33	41	15	**50**	7	314
Tier 3	6	153	49	74	20	71	46	12	431
TOTAL	11	356	112	117	71	101	126	19	913

(Fig. 4)

Paired Sample	BA	BCom	BSc Others	BSc Comp Sc	BBA	BBM	BCA	Others	TOTAL
Tier 1	2	89	12	10	10	15	30	0	168
Tier 2	3	114	51	33	41	15	**50**	7	314
Tier 3	6	153	49	74	20	71	46	12	431
TOTAL	11	356	112	117	71	101	126	19	913

(Fig. 5)

IX. FUTURE ENHANCEMENTS

In future, this research work can be completed by the HR groups of the organizations along these lines, that they can employ a flawless hopeful. These information controls should be possible utilizing SPSS or MS Excel. It can likewise be developed upto factorial anova if required to check the information at distinctive levels.

REFERENCES

[1] study.com (Internet Source)

[2] Housila P. Singh, "The Utilization of Kurtosis in the Estimation of the Parameters of the One- Way Random Effects Model", Biometrical Journal, vol. 44, issue 8, 2002.

[3] en.wikipedia.org (Internet Source)

[4] www.iasri.res.in (Internet Source)

[5] www.studyblue.com (Internet Source)

[6] Radwan, M H S, "Studies on the diffusion and mixing of solid particles in a tube mixer-V. Practical studies on a binary system with two radio activated components", The International Journal of Applied Radiation and Isotopes, 1973.