



Educated Unemployed

An Analytical Approach Towards Time to Employment

Encephal
Ramanujan College



Presented by Students of
Encephal
(Society of Department of
Statistics)

Contact:
encephal@ramanujan.du.ac.in

Ramanujan College
(University Of Delhi)

Table of Contents

- 1. Introduction**
- 2. Previously Research Done**
- 3. Methodology**
- 4. Analysis**
 - a. Introduction to Data**
 - b. Unemployed**
 - c. Employed**
 - d. General Awareness**
 - e. Distribution Fitting**
- 5. Team Members**

Introduction

India is the country with the 2nd largest population in the world. As per India's census 2011, the youth population (19-24 yrs) which is basically the undergraduate population of India, constitutes one-fifth of the total population of our country. India is expected to have a youth population of 33.5% of the total population of our country till 2020. With this huge sum of people paving way for employment, there must be good quality of education available to train these people for the industrial sector. Mass education of youth for their use in the industrial sector is a very important factor for India's demography.

The motive of earning is not only to work for oneself but for the other dependent people. The social effect which it creates is that, it gives one a sense of dignity, self-respect and social recognition. Lastly the most important impact is on the national income, as by working, an employee not only supports its family but also contributes to the national income of the country.

1) Trends of Employment

The study of employment patterns also helps us analyse the amount of contribution of different sectors to the national income. The trends of employment has changed over the last four decades in various sectors due to several factors.

Following are the three service sectors:

- **Primary Sector-** The percentage of workforce in this sector has decreased from 74.3% (in 1972-73) to 48.9% (in 2011-12). There are various reasons for this pattern, such as presence of automation of various agricultural operations, unequal land distribution and many more.
- **Secondary Sector-** The percentage of people employed in this sector has increased from 10.9% (in 1972-73) to 24.3% (in 2011-12).
- **Tertiary Sector-** The percentage of workforce in this sector has increased from 19.8% (in 1972-73) to 26.8% (in 2011-12). The major reason for this increase is because of globalization and the growth of IT sector in our country.

2) Causes of Unemployment

Educated unemployment in India has become a major problem over the years, with numerous inter-related causes. One of the major reasons is the continuously booming population of the country. With a constant increase in the number of job seeking

people, the demand for jobs is also at a constant rise. But India, with its developing and slow growing economy, is not able to meet the rising demands of employment. According to SoE in figures, 2017 another cause is the lack of skills required for jobs available. Today education lays more stress on the theory rather than on the practical applications and there is no focus on vocational training. Due to this, students though, become educated but lack the required skills to get a job. As the International Commission for Financing Global Education Opportunities reported in 2017, about 40% of employers worldwide find it difficult to recruit people with the skills they need. Moreover, young graduates or post-graduates are unwilling to take on jobs which doesn't meet their expectations. The system of reservation is also a cause for unemployment. Jobs are reserved for certain castes which often times leave the deserving candidates unemployed. In India, a large population is dependent on agriculture as their only source of income, which provides only seasonal employment.

3) Statistics related to Educated Unemployment

The most noteworthy feature of the Indian unemployment rates is that among job seekers (male and female both) the unemployment rate among people with secondary level education is potentially higher as compared with illiterates. This is true for both rural and urban settlements.

According to the Government's periodic labor force survey PLFS data, the following information is retrieved

Unemployment rates	RURAL	URBAN
EDUCATED	10.5%	9.2%
ILLITERATE	1.7%	2.1%

“As per the results of Periodic Labor Force Survey (PLFS) conducted by National Sample Survey Office, Ministry of Statistics and Programme Implementation during 2017-18, the estimated unemployment rate of educated people (secondary & above) on usual basis in the country was 11.4%,” labor minister Santosh Kumar Gangwar said in a recent meeting of Lok Sabha.

This is much higher than the average unemployment rate of the country as a whole which stands at 6.1%. Also, in 2017-18 Unemployment in India was at a 45-year high.

In their report, Kotak institutional equities added, “Unemployment rate is highest among youth with education until diploma (37%), graduate (36%) and post graduate and above (36%), suggesting the lack of opportunities for a higher skilled workforce.” A similar result was reported by the Centre for Monitoring Indian Economy (CMIE), a Mumbai based think-tank.

These figures can be worrying for the overall economic growth of India as well, given that half of India’s 1.3 billion population is below the age of 25.

According to Statistics, in 2019, 43.21 percent of the workforce in India were employed in agriculture, while the other half was almost evenly distributed among the two other sectors, industry and services. While the share of Indians, working in agriculture is declining, it is still the main sector of employment. While most of the Indian workforce is still employed in the agricultural sector, it is the services sector that generates most of the country’s GDP.

Country	Unemployment Rate
Bangladesh	4.29
China	3.62
Singapore	2.30
Nepal	1.29
Sri Lanka	5.1
Malaysia	3.3
Pakistan	5.79

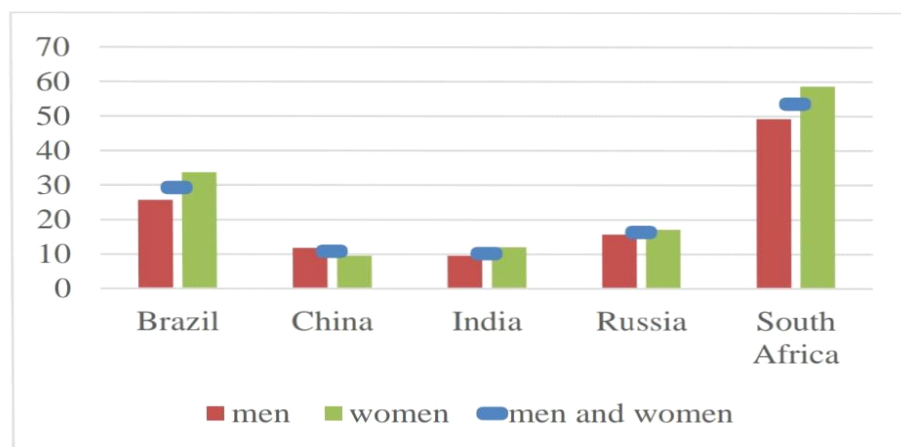
India: Distribution of the workforce across economic sectors from 2009 to 2019



	⬆️ Agriculture ⬆️	Industry ⬆️	Services ⬆️
2019	43.21%	24.89%	31.9%
2018	43.86%	24.69%	31.45%
2017	44.52%	24.47%	31.01%
2016	45.12%	24.29%	30.59%
2015	45.56%	24.34%	30.1%
2014	46.07%	24.38%	29.55%
2013	46.6%	24.36%	29.04%
2012	47%	24.36%	28.64%
2011	48.96%	23.52%	27.52%
2010	51.06%	22.38%	26.57%
2009	52.12%	21.61%	26.27%

Over the past decade or so, there has been a shift in employment from the agricultural sector to the service sector. The data signifies the same.

Figure 2: Youth (15-24) unemployment rates, BRICS countries, latest year (per cent)



Source: China - ILO modelled estimates, May 2018; Brazil - Continuous National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios Contínua); Russian Federation - Population Survey on Employment Problems; South Africa - Quarterly Labour Force Survey; India - National Sample Survey.

This table presents a data showing the recent unemployment figures of the neighboring countries of India. When compared with BRICS nations, we see somewhat good figures for India.

4) Literacy Rate in India

Literacy and level of education are the basic indicators of the level of development achieved by a society. Spread of literacy is generally associated with important traits of modern civilization such as modernization, urbanization, industrialization, communication and commerce. Literacy forms an important input in overall development of individuals enabling them to understand their social, political and cultural environment better and respond to it appropriately. Higher levels of education and literacy lead to much greater awareness and also contributes in improvement of economic and social conditions. It acts as a catalyst for social upliftment enhancing the returns on investment made in almost every aspect of development effort, be it population control, health,

hygiene, environmental degradation control, employment of weaker sections of the society.

According to the Census 2001, as many as 560,687,797 persons in the country are literate. Of these 336,533,716 are males and 224, 154,081 are females. While the overall literacy rate works out to be 64.8 %, the male literacy rate is 75.3% and that for females is 53.7%, showing a gap of 21.6 percentage points between the sexes at the national level. The gap is more in the rural areas. In the urban areas, higher literacy rate has been recorded both for males and females and the difference among the sexes is lower. Kerala, Mizoram, Lakshadweep, Goa and Chandigarh occupy the top five positions in literacy while Dadra & Nagar Haveli, Uttar Pradesh, Jammu & Kashmir, Arunachal Pradesh, Jharkhand, and Bihar, have the least.

Out of 561 million literates in the country, 145 million literates are educated only up to 'Below Primary' level and another 147 million up to 'Primary' level. Number of literates educated up to these two levels account for 52 % of the total literates in the country. 79 million (or 14.1%) of literates have attained Matric/Secondary level. There are 37 million persons in the country who are 'Graduates & above' which also include 12 million females.

Latest data puts India's adult literacy rate at 73.2 percent. While the country has made significant progress in improving literacy over the years, it continues to be home to 313 million illiterate people of which 59 percent are women. On comparing this data to the one obtained in 2001 (mentioned above), we observe that literacy in the nation has increased in many aspects. However, there are much higher milestones to be covered

Data shows that the gender gaps in education, occupation and wages have shrunk sharply between 1983 and 2010 in most indicators; the gaps have narrowed most sharply for the youth. The illiteracy of older adults and the elderly is driving down the overall literacy numbers of India. India is on track to achieve universal literacy for youth by 2030. The need is to focus on education quality, digital literacy, and skilling.

Age	Literacy Rate (%)								
	Rural			Urban			Rural & Urban		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
Age 5 & above	80.3	62.4	71.4	91	80.9	86.1	83.6	68.1	76
Age 7 & above	79.8	61.3	70.8	91.1	80.8	85.9	83.2	67.1	75.4
Age 15 & above	75	53.1	64.1	89.7	77.9	84	79.8	60.8	70.5
All age (age 0 & above)	72.3	56.8	64.7	83.7	74.8	79.5	75.7	62	69.1

Data Source: National Sample Survey Office.

The literacy rate for population in urban areas is greater than that in rural areas. The percentage of people attending educational is 76% in case of Age 5 and above, 75.4% in case of Age 7 and above, 70.5% in case of Age 15 and above and 69.1% in case of All age.

5) Government measures to fight off Unemployment and Educated Unemployment

- National Rural Employment Program:** The NREP was launched in year 1980 to create better employment opportunities in rural areas. The main objectives of this scheme were to generate the employment for underemployed or unemployed people in rural areas and to create the productive community assets for the benefit of poor and backward class people. The preference was given to landless labour in any kind of work under NREP. For employment, SC/ST were given preference amongst landless labour. Also, people working in NREP were paid partly in cash and partly in food grains.
- Jawahar Rozgar Yojana:** The Jawahar Rozgar Yojana was launched in year 1989. It was formed by merging two programs which were National Rural Employment Program and Rural Landless Employment Guarantee Program. It was the biggest employment program of that time. The main objective of this program was to create employment in every rural area under Panchayati Raj Institutions. The budget for this project was divided between central and state government in the ratio 80:20 which helped the project gain pace in a very short interval of time.
- Training of Rural Youth for Self-Employment:** TRYSEM was launched in 1979. It was one of the first major steps taken by the government for the youth at that time. The scheme aimed to train the rural youth for self-employment. The main

objective of this scheme was to provide opportunities to the rural youth from families below poverty line in the form of training and technical skills so that they can be self- employed in different fields and can lift up their families from poverty line. 40 youths were chosen from each block and were trained in technical skills and entrepreneurship. They were not only taught the physical skills but also the psychological skills, change in attitude, enhancement of motivation and skills in human relations etc.

- **Pradhan Mantri Rozgar Yojana:** This scheme was launched in 1993 with the aim to provide self- employment opportunities to the urban youth. It provides the opportunity to the educated youth to start their business if they are wishing to. The PMRY is targeted to the individuals of age group of 18-35 yrs. who have 240000 as their minimum annual family income and 10th pass as their minimum educational qualification. Due to the intense competition in government and private jobs, this scheme provides the opportunity to the people who want to launch a start-up. In the present time, there are around 1 crore 25 lakhs large business enterprises while the number of small businesses are 12 crores. So focusing on small businesses is very essential for the economy of the nation.
- **Swarna Jayanti Shahari Rozgar Yojana (SJSRY):** SJSRY was launched on 1 December 1997. Its main aim was to generate employment for the unemployed or underemployed urban poor by encouraging them to create self- employment opportunities. The cost sharing of this scheme between central and the state government is 75:25. Around 2 lakhs of urban poor were benefited by skill development whereas 50 thousand of urban poor were benefitted under self-employment annually till 2004-05.

Previously Done Research Work

- Challenges Faced By Educated Unemployed Youth in India by DR.V. GOMATHI, Mrs. M. NEELA

January 2016

In India, there has been an alarming phase of mass-unemployment among the educated youth. Post-graduates and graduates are walking pillar to post in search of employment. When there is slowdown in economy, there is surge in the number of uneducated unemployed people. Many educated people are knocking at the gates of offices just for the position of a petty clerk and get disappointed when they read the words 'No Vacancy' on the gate. Numerous cases of suicide among the unemployed young men have been a feature of our time.

- Unemployment among educated youth: implications for India's demographic dividend by Majumder, Rajarshi and Mukherjee, Dipa

9 May 2013

Researchers claim that India is poised for reaping demographic dividend and leapfrog to a higher level of income-employment situation utilizing the relatively larger share of youth or working age persons in total population. However, the outcome depends on the contribution of youth to national product. India at present suffers from remarkably high educated unemployment and questions are also raised about the employability of the youth because of their inadequate education, training, and market ready skill. Huge youth unemployment, especially educated unemployment is the surest way to social tension, unrest, and unlawful activities turning the demographic dividend into a demographic nightmare. In this paper we look at the issue of education, skill formation and unemployment among youth in India, focusing specially on educated unemployment. We find that current skill/training situation of youth in India is inadequate. Surplus and shortage coexists in the labor market indicating serious mismatch between supply and demand. There is an urgent need to relook at human resource development strategies in the country.

- Youth Employment and Unemployment in India by S. Mahendra Dev and M. Venkatanarayana

January 2011

Increase in the share of youth population due to demographic 'dividend' or the 'youth bulge' seems to be one of the sources of future economic growth in India. Although with increase in school and college enrolment rates, the proportion of youth in the labor force has been declining, their high proportions in the labor force indicate that the problem of youth unemployment and underemployment would remain a serious policy issue for many more years to come in India. In this context, this paper examines the employment and unemployment situation of the youth in India during the last two-and-half decades viz., 1983 to 2007-08. It analyses the trends in labor force and workforce participation rates, unemployment, joblessness, working poor, growth and employment elasticities etc. The paper also offers policy recommendations for increasing productive employment and reduction in unemployment for the youth. The poor employability of the workforce would hamper the advantages due to demographic dividend if measures are not taken to improve the educational attainment and skill development of the youth.

- Socio economic determinants of educated unemployment in India by Bairagya, Indraj

Despite the presence of a large volume of sub-standard informal employment in India, the economy is not free from a perennial unemployment problem. As per the National Sample Survey Organization (NSSO) Data, the unemployment rate is more among the educated (secondary and above) persons as compared to those whose education level is lower. This paper, using different rounds of NSSO data, seeks to explore the socio-economic and regional factors responsible for educated unemployment in India. We find that possessing technical education by itself doesn't necessarily guarantee employment which, in fact, questions the rationale behind the present Indian government's initiative to promote technical education on a large scale. Although high-industrialized states account for low educated unemployment rates, these states are also unable to engage all educated people in the production process. Moreover, since the number of unemployed is higher for both the educated and uneducated people among those who are not registered with employment exchanges, search and matching problem may also be an important reason for educated unemployment in India. Besides, differences have been found across social groups, gender, religions, regions and other socio-economic characteristics.

- NBER WORKING PAPERS SERIES EDUCATION AND UNEMPLOYMENT

By Jacob Mincer Sep 1991

A major benefit of education is the lower risk of unemployment at higher educational levels. In PSID (Panel Study of Income Dynamics) data on the male labor force¹ the reduction of the incidence of unemployment is found to be far more important than the reduced duration of unemployment in creating the educational differentials in unemployment rates. In turn, the lesser unemployment incidence of the more educated workers is, in about equal measure, due to their greater attachment to the firms employing them, and to the lesser risk of becoming unemployed when separated from the firm. The lesser frequency of job turnover of more educated workers, which creates fewer episodes of unemployment, is in large part attributable to more on-the-job training. In explaining the lesser conditional unemployment of educated workers and the somewhat shorter duration of their unemployment, indirect evidence is provided that (1) costs of on-the-job search for new employment relative to costs of searching while unemployed are lower for more educated workers; (2) that these workers are also more efficient in acquiring and processing job search information; and (3) that firms and workers search more intensively to fill more skilled vacancies.

- The Impact of Higher Education on Unemployment By Savannah Horner, Athena Zhang, and Madeline Furlong

2018

This paper explores the relationship between higher education and unemployment using regression analysis. Our hypothesis is that the greater the government expenditure on higher education, the lower a state's unemployment will be. Other independent variables such as state GDP per capita, the percentage of the population with bachelor degrees or higher, the cost of college attendance, the share of manufacturing in the state economy, and financial aid as a percentage of state revenue were used in a multi-regression analysis in order to account for bias. Our results found that there is a strong negative relationship between higher education expenditures and unemployment.

Methodology

Descriptive Statistics:

Descriptive statistics are basically brief descriptive coefficients that summarize a given data set, which can either be a representation of entire set or a sample of population.

It is broken into 4 measures-

1. Measure of central tendency
2. Measure of dispersion
3. Measure of skewness
4. Measure of kurtosis

Measure Of Central Tendency:

This measure includes calculation of mean, median, mode of the data which basically gives us the central tendency of the data

Measure Of Dispersion

This helps us to know the dispersion in the given data with the help of variance and standard deviation

Measure Of Skewness

Skewness means lack of symmetry in data. There is a need to study skewness to know about the shape of the curve which is drawn with the given data.

A curve is said to be skewed if

1. Mean, Median and Mode all three falls at different points i.e. they are unequal
2. Quartiles are not equidistant from the Median.

Measure Of Kurtosis

This measure helps us to know about the flatness or the peakedness of the frequency curve.

There are 3 types of curves under this-

1. Leptokurtic curve
2. Normal curve
3. Platykurtic curve

The main purpose for studying descriptive statistics is to provide a brief summary of the samples and the measures done on a particular study.

Graphical Presentation Of The Data

Graphical presentation is basically a way of analyzing the numerical data. It exhibits the relation between data, ideas, information and concepts of the diagrams.

Graphs are very easy to understand and one of the most important learning method. The type of diagram to be constructed depends upon the information given to us.

There are different types of graphs which can be made depending upon the requirement.

1. LINE GRAPHS

Line graphs are used to display the continuous data and it is useful for predicting the future events over time.

2. BAR GRAPHS

Bar graphs is used to display the category of data and it compares the data using solid bars to represent the quantities

3. HISTOGRAM

Histograms are the bars used to represent the frequency of the numerical data which is organized in the form of class intervals.

4.LINE PLOTS

It shows the frequency of the data on a given number line.

5.FREQUENCY TABLE

The table shows the number units of data that fall into a particular class interval

6.FREQUENCY POLYGON

Frequency polygon can be obtained by joining the mid points of all the bars of the histogram in case of the equal class intervals

Inferential Statistics

Generally, Inferential statistics takes data from a sample and makes inferences about the larger population from which the sample was drawn. Because the aim of inferential statistics is to draw conclusions from a sample and generalize them to a population. So, we need to have confidence that our sample accurately reflects the population.

There are two main areas of inferential statistics:

- **ESTIMATING PARAMETERS:**

This means taking a statistic from a sample data (i.e., as sample mean) and use it to say something about a population parameter (i.e., population mean)

- **HYPOTHESIS TESTING:**

This is where you can use sample data to answer research questions.

TESTING OF HYPOTHESIS

Hypothesis testing was introduced by Ronald Fisher, Jerzy Neyman, Karl Pearson and Pearson's son Egon Pearson. Hypothesis testing is an act in statistics whereby an analyst tests an assumption regarding a population parameter.

KEY TERMS:

Null Hypothesis: A null hypothesis is a type of hypothesis used in statistics that proposes no statistical significance exists in a set of given observations. It is denoted by H_0 .

Alternate hypothesis: A hypothesis complementary to null hypothesis. It is denoted by H_1 .

Types Of Error :

Type 1 error: Reject H_0 , when H_0 is true. It is denoted by α .

Type 2 error: Accept H_0 , when H_0 is false. It is denoted by β .

LEVEL OF SIGNIFICANCE: It is size of type 1 error or also called maximum procedure's risk.

ONE TAILED TEST: One-tailed test is a statistical test in which the critical area of the distribution is one sided. And if the sample being tested falls into the one side, the alternative hypothesis will be accepted. It is either right tailed or left tailed.

TWO TAILED TEST: A test of any statistical hypothesis where the alternate hypothesis being two tailed. This is called two-tailed test.

ANOVA (Analysis Of Variance):

Analysis of variance is a statistical method used to test difference between two or more means. According to Prof. R.A. Fisher, Analysis of Variance (ANOVA) is the "Separation of variance ascribable to one group of causes from the variance ascribable to other groups". By this technique the total variation the sample data is expressed as the sum of variation due to the non-negative components where each of these components is a measure of the variation due to some specific independent source or factor.

ONE-WAY ANOVA:

The scheme of classification due to one factor is called one-way classification. Its analysis is called one-way analysis of variance.

One-way ANOVA determines whether there are any statistically significant differences between the two means from two independent (unrelated) groups using the F-distribution.

Statistical Hypothesis:

- The null hypothesis is that the means are all equal, such that,

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

- The alternate hypothesis is that at least one of the means is different.

Mathematical model:

$y_{ij} = \mu_j + \epsilon_{ij}$, where μ_j = mean of the j th population = $\mu + \tau_j$.

$y_{ij} = \mu + \tau_j + \epsilon_{ij}$, where $j = 1, J$ (# groups) and $i = 1, 2, \dots, N_j$

TWO WAY ANOVA:

The scheme of classification due to two factors is called two-way classification. Its analysis is called two-way analysis of variance.

A two-way ANOVA test is a statistical test used to determine the effect of two nominal predictor variables on a continuous outcome variable.

NOTE: The Non-Normality of Error terms Assumption in ANOVA

If we consider a Fixed Effects model (model with known number of treatment levels), the non-normality assumption does not make any significant impact over the estimation process and variance related tests. Estimation will always be unbiased using the OLS approach, but normality assumption is directly related to the testing of means and variances. Therefore, it is expected that variance ratio tests, including the ANOVA F test would be affected with non-normality assumption. But past researches and several simulations have come to the conclusion, that as the sample size increases, the severity of non-normality tends to get reduced (due to Central Limit Theorem). Therefore, ANOVA under non-normality won't show any significant deviations from the standard procedure, and hence, it will remain robust too.

However, under the case of Random Effects Model (model with a few sampled levels of treatments from a large number of treatments), the implications would be severe, because the number of treatments are themselves random variables. Therefore, under this model, it is better to have strong assumptions in the form of population distribution, making normality assumption a desired characteristic.

However, in cases where non normality prevails, it is better to switch to Non-Parametric Methods of testing, as these methods are not generally based on the assumptions of any population distribution.

T-Test

A t-test allows us to compare the differences (measured in means or average values) of the two data sets and determine if they came from the same population. The t-test takes a sample from each of the two data sets and establishes the problem by stating

the null hypothesis that there is no significant difference between the means of the two data sets. Based on the standard formulas, the value of calculated t is calculated and is compared against the standard (tabulated) value and accordingly the decision regarding the acceptance and rejection of null hypothesis is made.

One sample t-test

$$t = (\bar{X} - \mu) / (\hat{\sigma} / \sqrt{n})$$

where,

\bar{X} is the sample mean from a sample X_1, X_2, \dots, X_n of size n ,

$\hat{\sigma}$ is the estimate of the standard deviation of population,

μ is the population mean

F-Test

F-test is also known as F-test to Compare Two Variances. The F distribution is calculated by the ratio of two independent chi square variates divided by their respective degrees of freedom.

$$F = \frac{\frac{df_1 \cdot s_1^2}{\sigma_1^2} / df_1}{\frac{df_2 \cdot s_2^2}{\sigma_2^2} / df_2}$$

All hypothesis testing is done under the assumption that the null hypothesis is assumed to be true.

Assumptions are F values are always positive because the variances can never be negative, the distribution is non symmetric, the samples must be independent, the population from which samples are drawn must be normal and there are two independent degrees of freedom (numerator and denominator).

X² Goodness Of Fit

Chi square goodness of fit test is a non-parametric test that is used to find out how the observed value of a given phenomenon is significantly different from the expected value. When an analyst attempts to fit the statistical model to the observed values, he/she might check how close are the observed values to the expected values under the fitted model. The test is applied when you have one categorical variable from one single population. It is used to determine whether sample data are consistent with a hypothesis distribution.

$$\chi^2 = \left[\frac{(O - E)^2}{E} \right]$$

Where,

O_i is the observed values

E_i is the expected values.

Survival Analysis

Survival analysis is the branch of the statistics for analyzing the expected duration of time until one or more events happen, such as death in biological organisms and failure in mechanical systems.

It is set of statistical approaches for data analysis where the outcome variable of interests is time until an event occurs.

Survival analysis is used in the variety of fields such as:

1. Cancer studies for patient's survival time analysis
2. Sociology for event-history analysis
3. Also used in engineering for failure time analysis

There are two main functions which forms the backbone of this topic-

1. Survival function represented by S(t)
2. Hazard function represented by h(t)

Kaplan Meier Estimation

Kaplan Meier method or product limit method is a widely used and trusted non parametric method when we seek to estimate distribution of the data allowing appropriate adjustments for censored data. Censored data refers to the data corresponding to those observations whose information is partly known. This method has the following assumptions:

1. Censoring present is non informative.
2. All observations follow the same distribution irrespective of the time of entry in the study
3. The hazard of occurrence of events is zero at times where events did not occur in the given data.

The survival probabilities are calculated using the following formula:

$$S(t) = \prod_{t_j \leq t} \frac{\text{number of at risk observations} - \text{number of events}}{\text{number of at risk observations}}$$

Where;

$S(t)$ = survival probability at the t

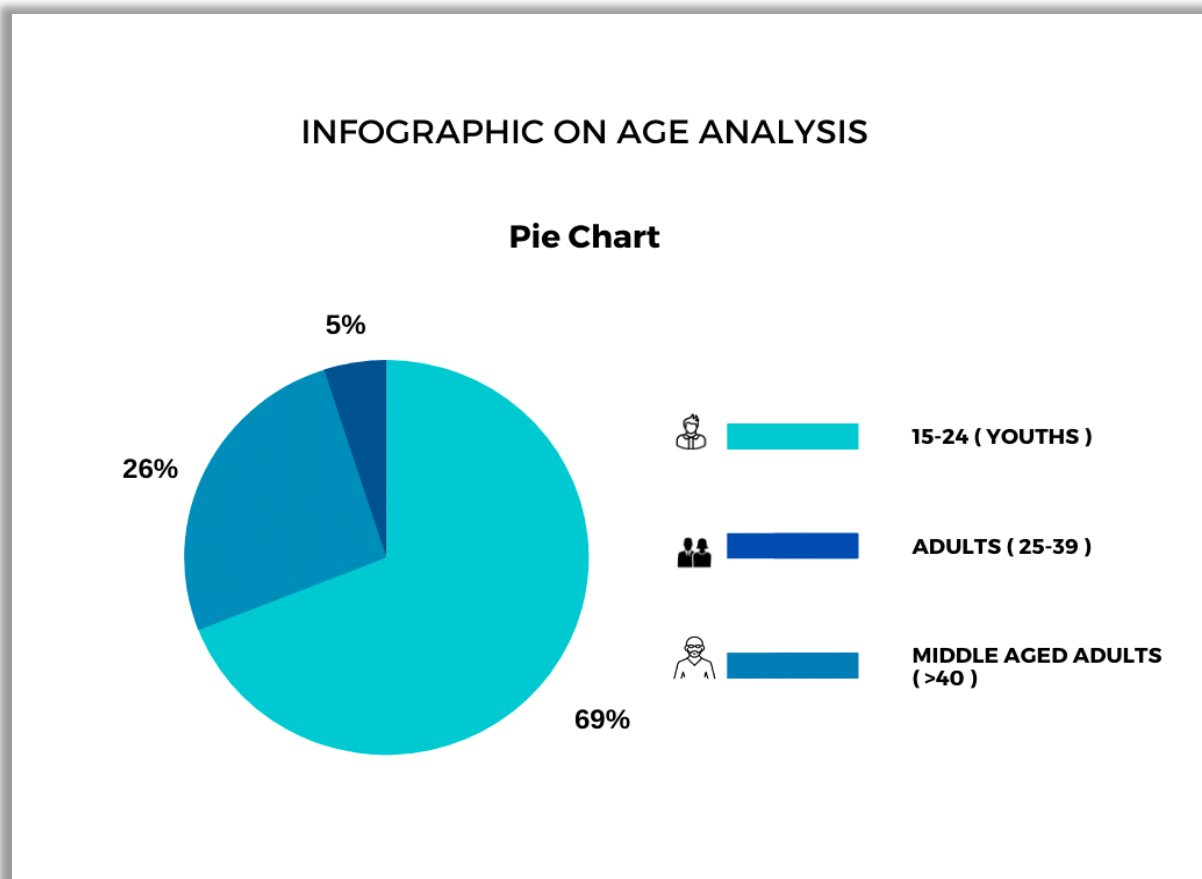
t_j = time of j^{th} occurrence of event

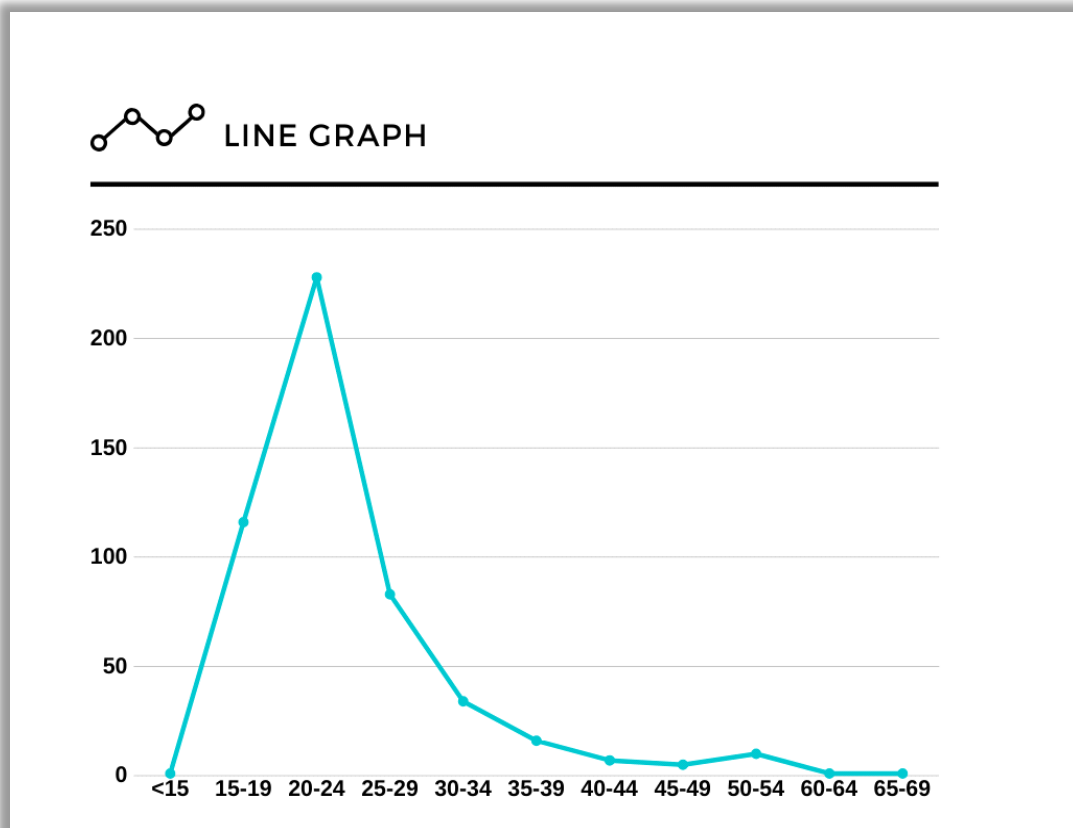
Log Rank Test

Analysis

SECTION – A (INTRODUCTION)

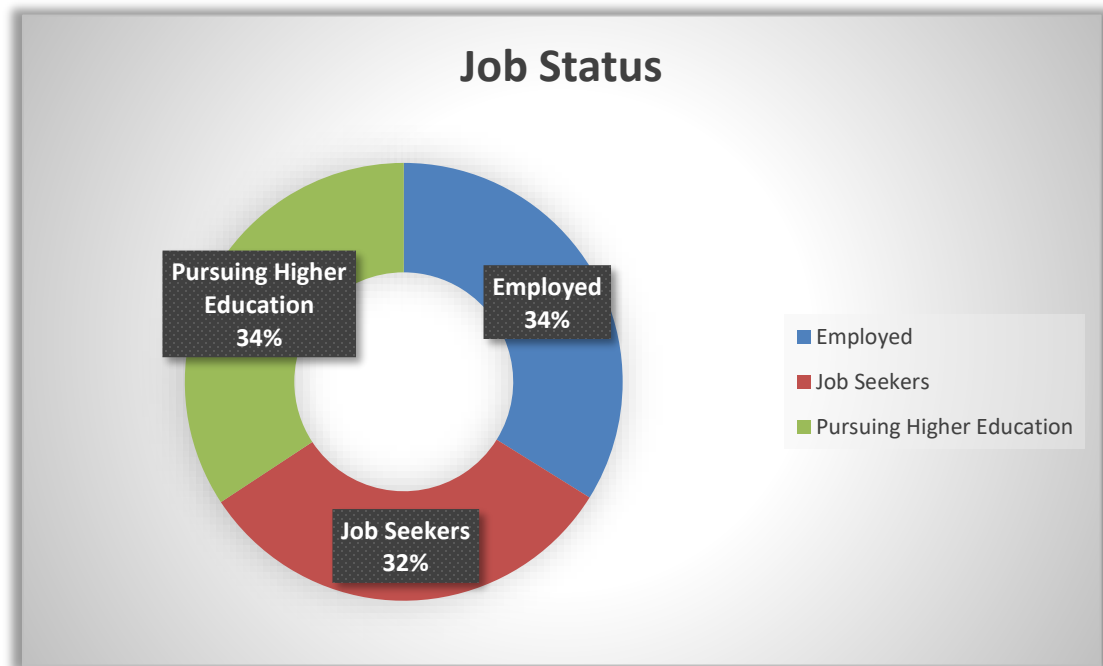
1) Age Analysis:





69 % are the youths (15-24) which had participated in the survey. 26 % are the adults (25-39) which had participated in the survey. 5 % are the middle-aged adults (>40) participated in the survey. Therefore, it is evident that the project covered a fair share of young aspirants who do face problems like educated unemployment after completion of their minimum qualifications.

2) Job Status



In the study conducted, there were a total of 502 responses, out of which 170 respondents were employed. Job seekers counted up to 160 in numbers. Therefore, a combined score of 330 provided a good insight into the job status and related information. With this data, limbo period could be easily calculated and distribution fitting seems easier.

3) Field Of Education

Row Labels	Count of Field of Education	%age of Field of Education2
Arts	111	22%
Commerce	138	27%
Computer Sciences	16	3%
Engineering	49	10%
Law	5	1%
Medical Sciences	13	3%
Science	165	33%
Sports	5	1%
Grand Total	502	100%

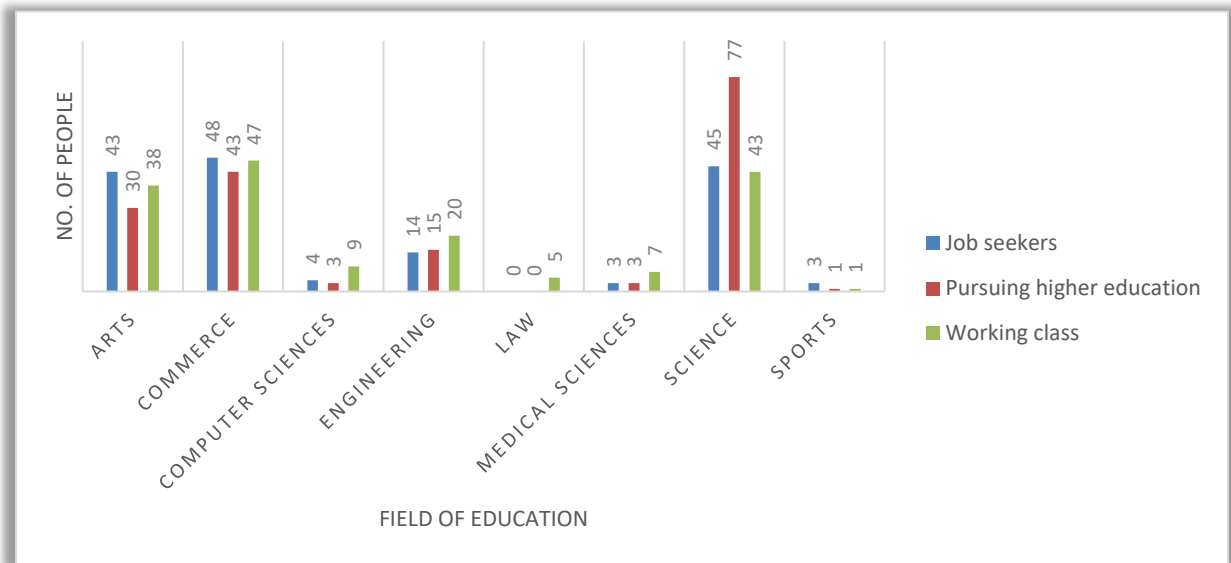
In the study conducted, nearly 1/3rd of the responses came from the educated persons who have Sciences as their Field of Education. Arts and Commerce background students were also included in the study.

4) Qualification Of Respondents



The majority of respondents in the study are Graduates. Senior Secondary pass out students accounted for 1/3rd of the data. Therefore, this way each strata got some representative sample to facilitate consistent and unbiased estimation procedure. Moreover, the targeted section of society was successfully approach in the study.

5) Field Of Education V/S Job Status



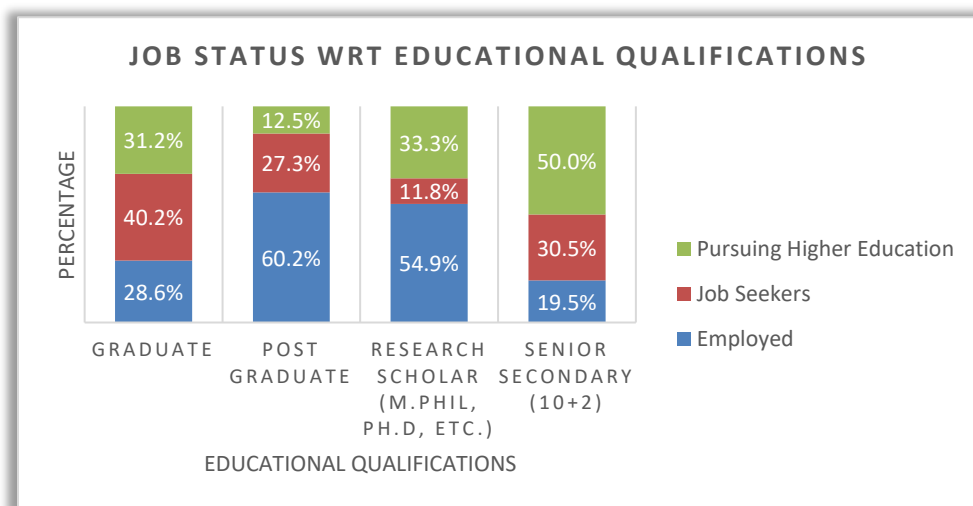
From the graph, the job status in various fields of education can be known. It also shows the scope of being employed in various fields. As the study had a frame of educated youth, we have the following responses:

They are employed

They are seeking jobs

They are pursuing higher education.

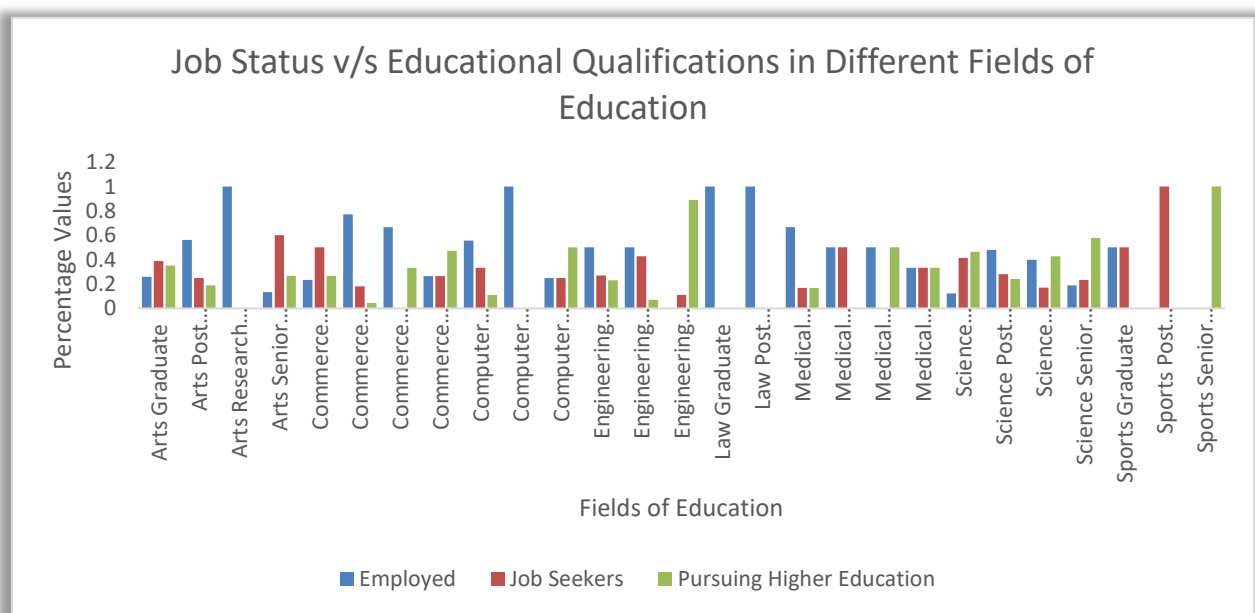
6) Job Status V/S Educational Qualifications



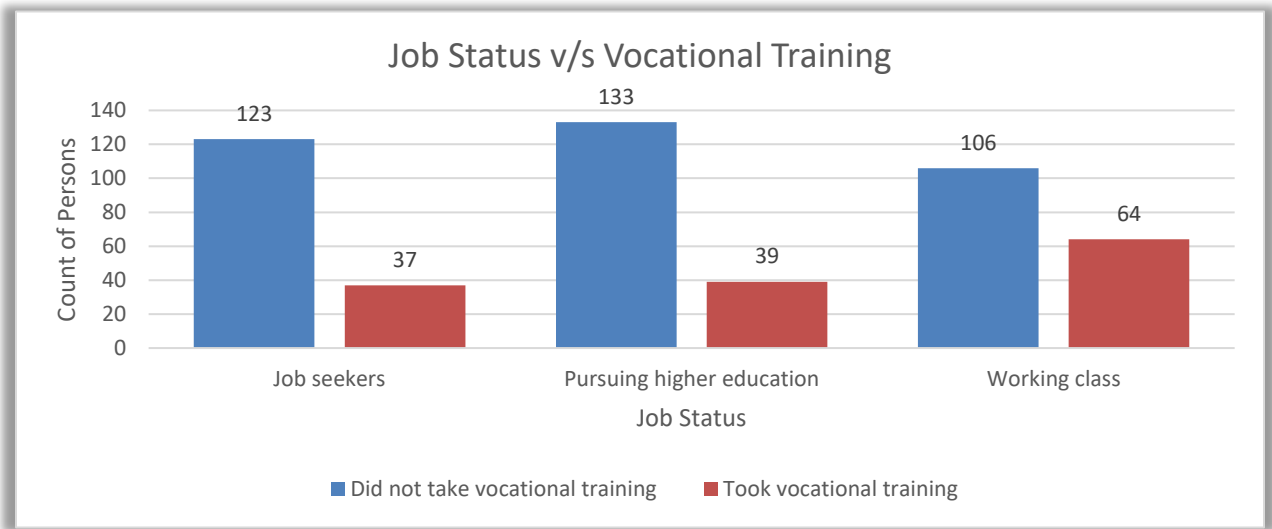
This throws light upon the fact that Post Graduates and Research Scholars do find employment easily, simply because of the fact that their educational qualifications are pretty much decent. The demand for educated youth in employment market is high, and therefore, they tend to receive employment with good remuneration and incentives. While for the Graduates and Senior Secondary pass outs, there is an opposite scenario, as competition among the educated youth leads to rise in educational standards for job seekers. Moreover, 50% of the job seekers do pursue higher education in consideration of the aforementioned fact.

JOB STATUS WITH RESPECT TO EDUCATIONAL QUALIFICATIONS				
Count of Are you Employed ?	Column Labels			
Row Labels	Employed	Job Seekers	Pursuing Higher Education	Grand Total
Arts	34.23%	38.74%	27.03%	100.00%
Graduate	25.93%	38.89%	35.19%	100.00%
Post Graduate	56.25%	25.00%	18.75%	100.00%
Research Scholar (M.Phil, Ph.D, etc.)	100.00%	0.00%	0.00%	100.00%
Senior Secondary (10+2)	13.33%	60.00%	26.67%	100.00%
Commerce	34.06%	34.78%	31.16%	100.00%
Graduate	23.33%	50.00%	26.67%	100.00%
Post Graduate	77.27%	18.18%	4.55%	100.00%
Research Scholar (M.Phil, Ph.D, etc.)	66.67%	0.00%	33.33%	100.00%
Senior Secondary (10+2)	26.42%	26.42%	47.17%	100.00%
Computer Sciences	56.25%	25.00%	18.75%	100.00%
Graduate	55.56%	33.33%	11.11%	100.00%
Post Graduate	100.00%	0.00%	0.00%	100.00%
Senior Secondary (10+2)	25.00%	25.00%	50.00%	100.00%
Engineering	40.82%	28.57%	30.61%	100.00%
Graduate	50.00%	26.92%	23.08%	100.00%
Post Graduate	50.00%	42.86%	7.14%	100.00%
Senior Secondary (10+2)	0.00%	11.11%	88.89%	100.00%
Law	100.00%	0.00%	0.00%	100.00%
Graduate	100.00%	0.00%	0.00%	100.00%
Post Graduate	100.00%	0.00%	0.00%	100.00%
Medical Sciences	53.85%	23.08%	23.08%	100.00%
Graduate	66.67%	16.67%	16.67%	100.00%

Post Graduate	50.00%	50.00%	0.00%	100.00%
Research Scholar (M.Phil, Ph.D, etc.)	50.00%	0.00%	50.00%	100.00%
Senior Secondary (10+2)	33.33%	33.33%	33.33%	100.00%
Science	26.06%	27.27%	46.67%	100.00%
Graduate	12.20%	41.46%	46.34%	100.00%
Post Graduate	48.00%	28.00%	24.00%	100.00%
Research Scholar (M.Phil, Ph.D, etc.)	40.00%	17.14%	42.86%	100.00%
Senior Secondary (10+2)	18.75%	23.44%	57.81%	100.00%
Sports	20.00%	60.00%	20.00%	100.00%
Graduate	50.00%	50.00%	0.00%	100.00%
Post Graduate	0.00%	100.00%	0.00%	100.00%
Senior Secondary (10+2)	0.00%	0.00%	100.00%	100.00%
Grand Total	33.86%	31.87%	34.26%	100.00%



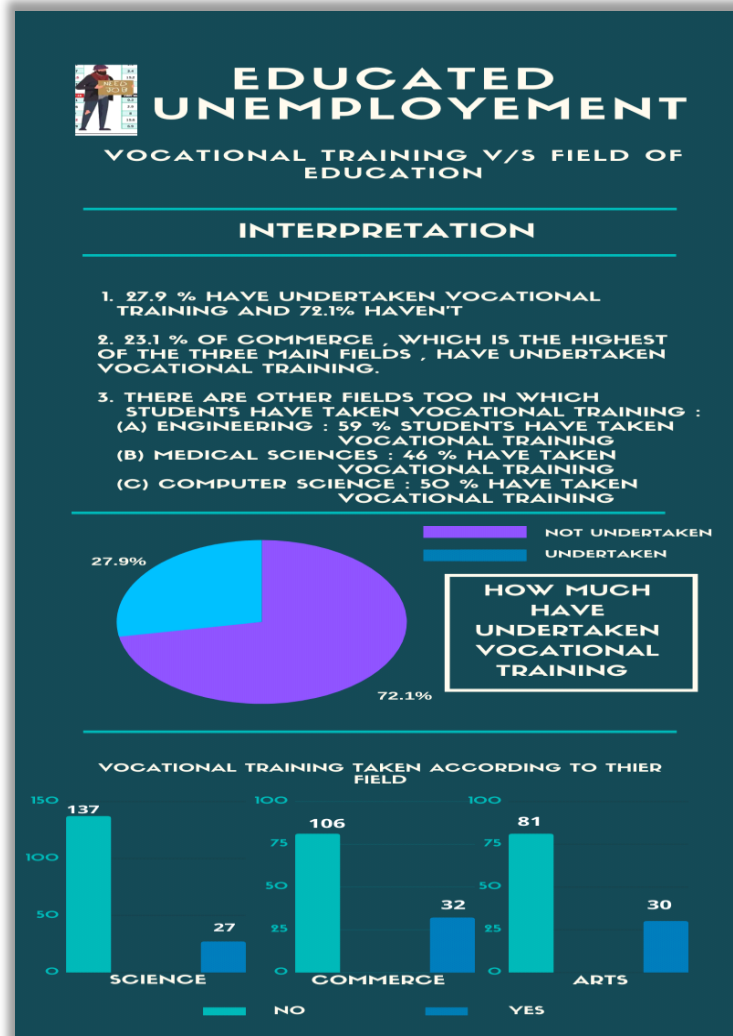
7) Vocational Training V/S Job Status



The above graph shows the relation between vocational training and job status. It represents if undertaking of vocational training is useful in getting a job.

Among the people who have taken vocational training most are in the working class which shows that it is beneficial to some extent. Moreover, those who pursue higher education do not find taking vocational training as a feasible alternative, because they are already involved in skill enhancement. Job seekers seem like not interested in undertaking vocational training, but it could be a determining factor for the cause of good, helping them to seek better employment.

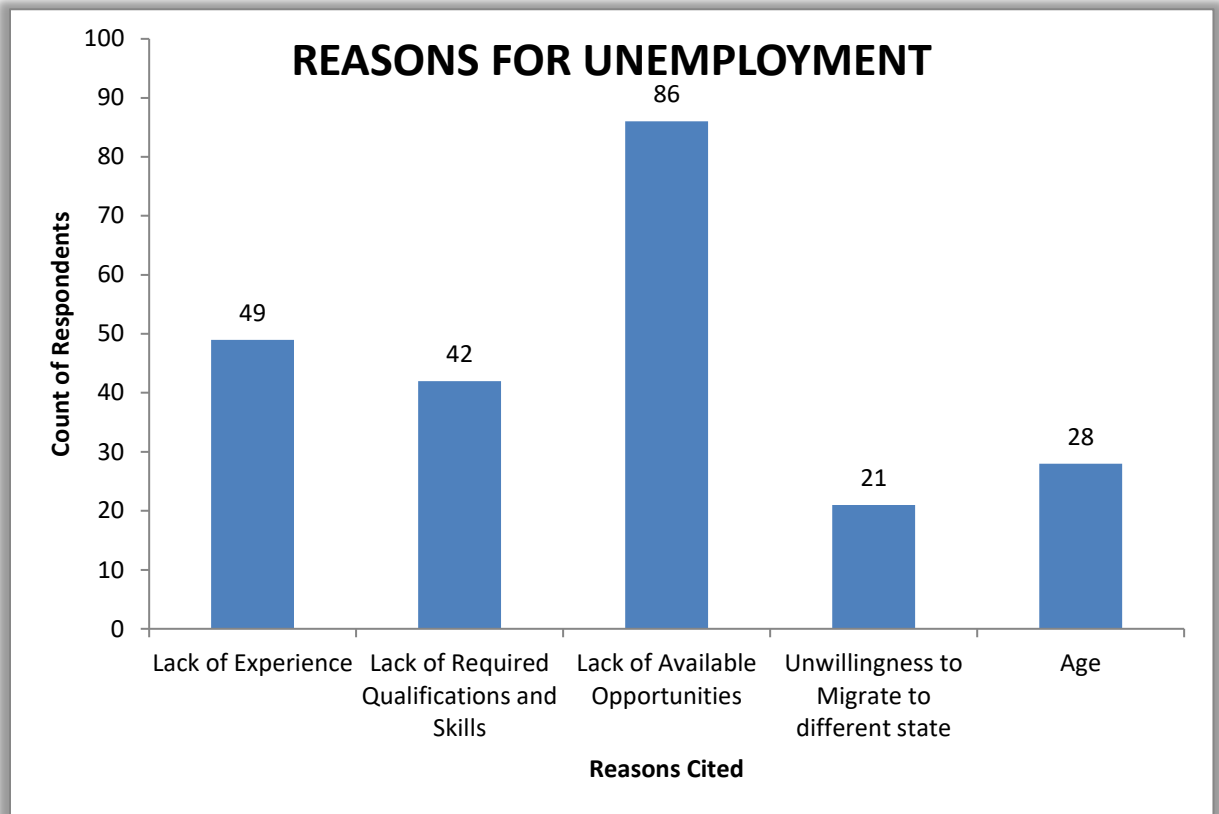
8) Vocational Training Vs Field Of Education



- 27.9 % have undertaken vocational training and 72.1% haven't. Therefore, either respondents are not aware of skill enhancement programmes available to them, or they do not find it really helpful.
- 23.1 % of Commerce students, which is the highest of the three main fields, have undertaken vocational training.
- There are other fields too in which students have taken vocational training:
A) Engineering: 59 % of students have undertaken vocational training.
B) Medical Sciences: 46 % of the students have taken vocational training.
C) Computer Science: 50 % have taken vocational training.

SECTION - B (JOB SEEKERS)

1. Reasons For Unemployment

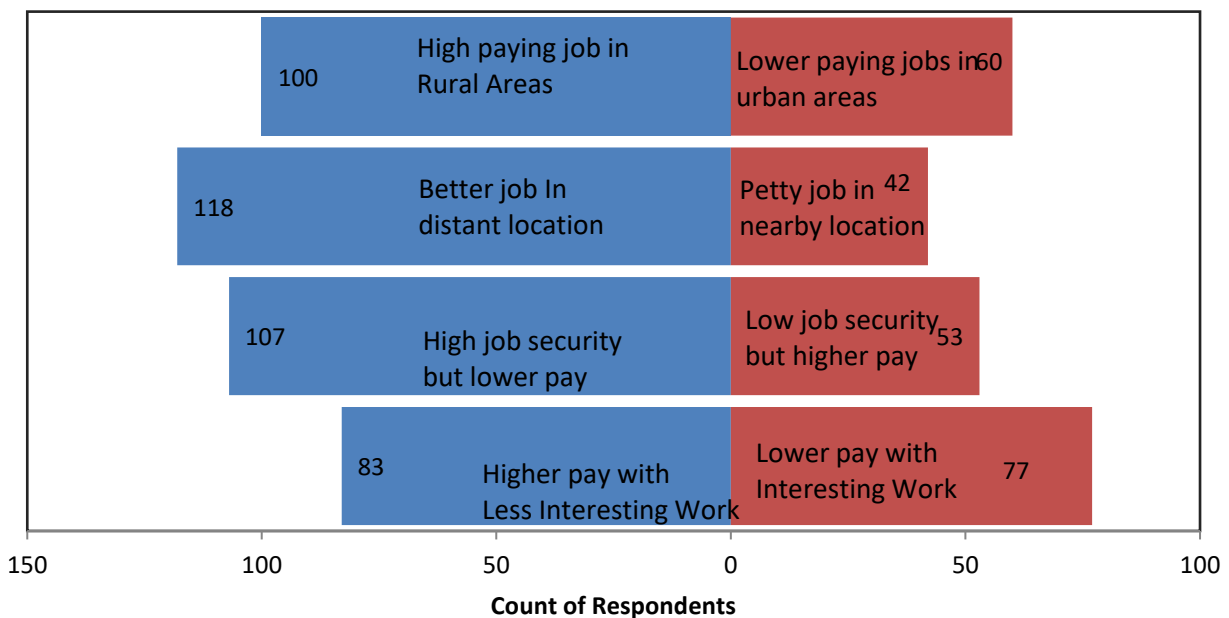


Lack of Available Opportunities and Lack of Experience are the most cited reasons by those respondents who are seeking jobs. Lack of Required Qualifications and Skills is surfaced as another reason, which is in sharp contrast to the fact that job seekers do find undertaking any vocational training or skill enhancement program to help their cause.

2. Preferences Of The Respondents

Higher pay with Less Interesting Work	83
High job security but lower pay	107
Better job In distant location	118
High paying job in Rural Areas	100
Lower pay with Interesting Work	77
Low job security but higher pay	53
Petty job in nearby location	42
Lower paying jobs In Urban Areas	60

Preferred Job Characteristics



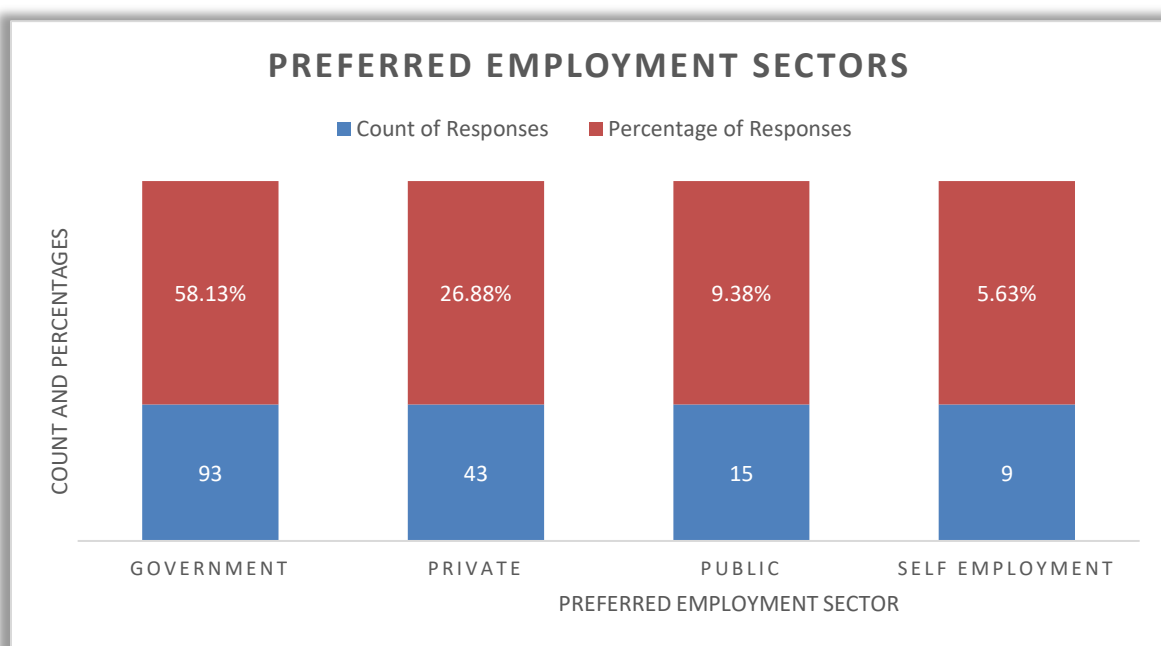
The following are the responses which are recorded more frequently than their complement ones:

1. High paying job in Rural Areas.
2. Better job in distant location.
3. High job security but lower pay.
4. Higher pay with Less Interesting Work.

This represents the tendency of job seekers to assume finding a higher paying employment. Against a better pay, they seem perfectly fine if they can compromise and accommodate their issues in daily life.

3. Preferred Employment Fields

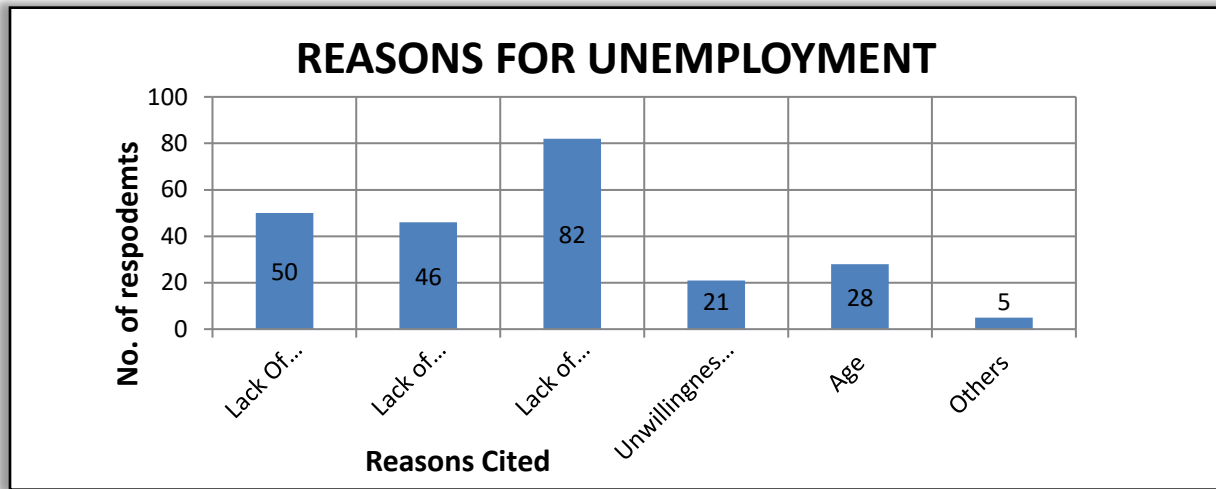
QUALIFICATION OF RESPONDENTS		
Row Labels	Count of Responses	Percentage of Responses
Government	93	58.13%
Private	43	26.88%
Public	15	9.38%
Self employment	9	5.63%
Grand Total	160	100.00%



Job seekers do find the government jobs highly desirable as more than half of them corresponded to selecting Government jobs as the preferred ones, followed by private sector with only 27% responses.

Self-Employment recording only for 6% of responses is highly significant to show the Job Creators are missing from among the Educated Youth. Youth desires to find jobs, instead of creating new jobs for themselves and others.

4. Reasons For Unemployment



The above graph represents the perspective of the educated unemployed masses as to what they attribute their unemployment to.

And the prime reason for educated unemployment turns out to be **Lack of Available Opportunities**.

5. Difference Between The Qualification Of The Respondent And The Minimum Qualification Required For The Current Job

When categorized According to the qualifications, it is seen that most of the research scholars (42%) are employed in a job that demands less than a PhD. Those who have completed Post Graduation have about 29% people working in a job for which they are overqualified. About 17% of the working graduates are overqualified and none of the employed respondent, qualified till senior secondary level reported being overqualified. This highlights the fact that there are less job opportunities which require being a research scholar, therefore, about 40% of the research scholars have a job for which they are overqualified.

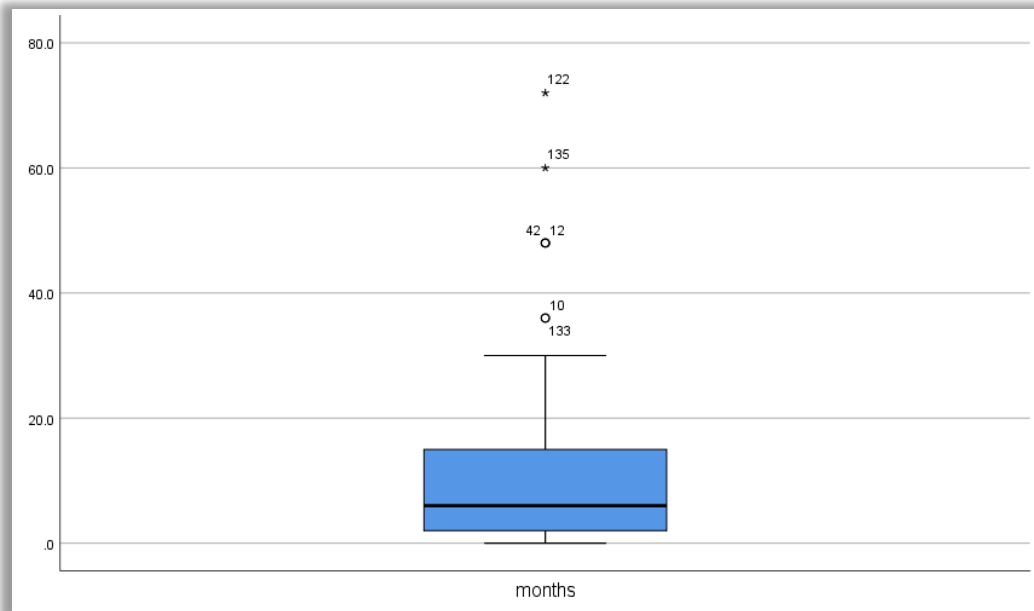
3. SECTION-C (EMPLOYED)

1. Months Taken To Find Employment Since Completion Of Minimum Educational Qualification Requirement For Job

Here are the summary statistics of the variable which records the months taken by an employed respondent to find his job since the completion of his minimum qualifications for the job:

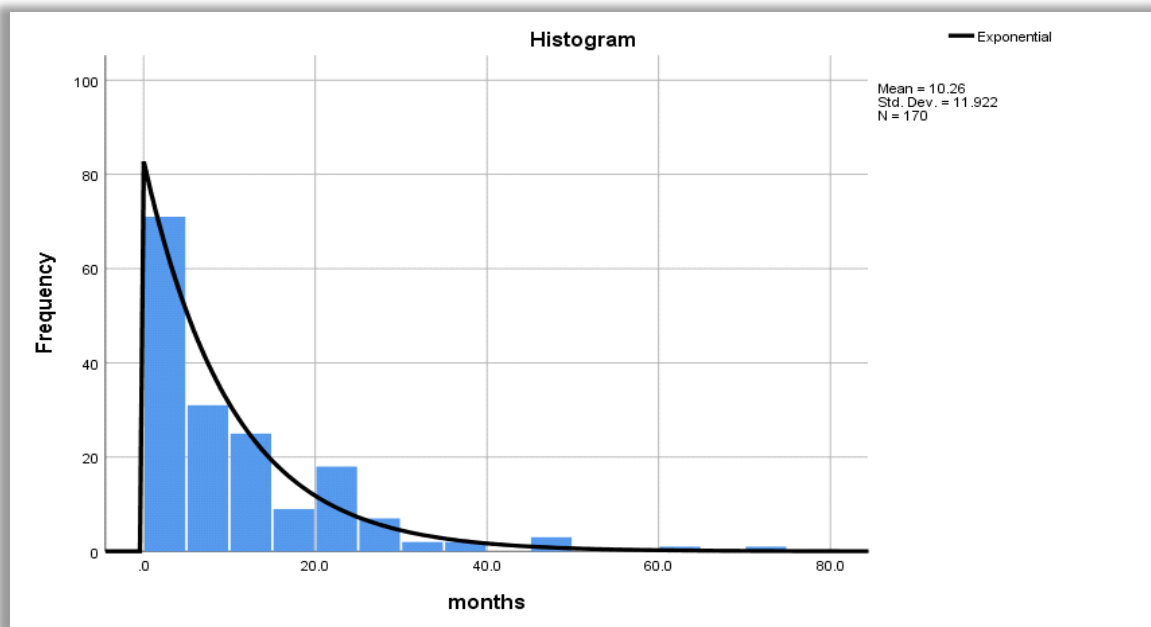
Statistics		
Months taken to find Employment		
N	Valid	170
Mean		10.256
Median		6.000
Mode		1.0
Std. Deviation		11.9219
Skewness		2.080
Kurtosis		5.884
Minimum		.0
Maximum		72.0
Percentiles	25	2.000
	50	6.000
	75	15.000

Below is the boxplot of the variable that represents its positional distribution, and marks



out the outliers as well.

The histogram has a superimposed estimated curve of an exponential distribution, which suggests (after comparison from several other distributions like Gamma, Weibull, Rayleigh) that an exponential distribution could be fitted into the model.



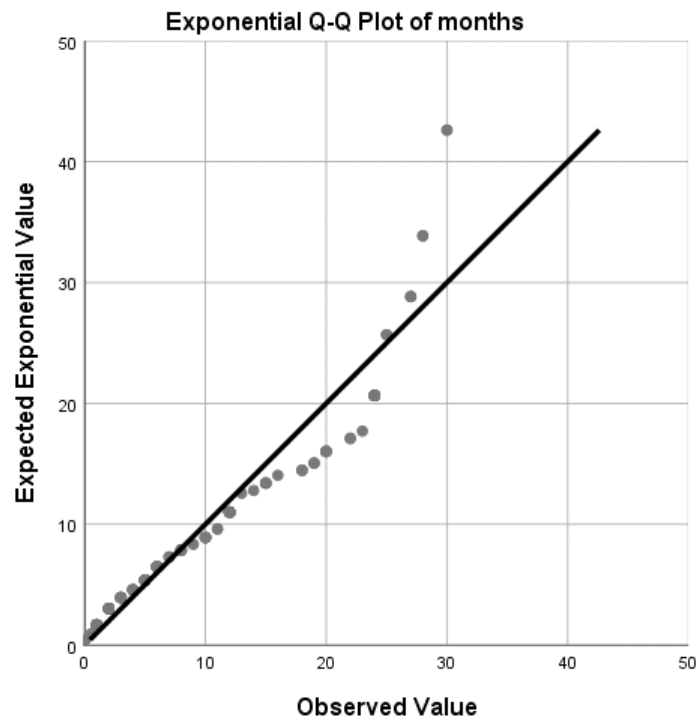
To obtain a better fit, we remove outliers (values greater than or equal to 36) and obtain the estimates as below using a Q-Q Plot as follows:
(we name the variable monthsexp to differentiate from months)

• Q-Q Plot		
Estimated Distribution Parameters		
		Months
Exponential Distribution	Scale	.117
The cases are unweighted.		

The rate parameter equals 0.117 which essentially means that Mean Time taken to find employment is 8.54 months.

Hence, the density function is

$f(x) = (0.117) \cdot \exp(-0.117x)$; for all $x \geq 0$



- **T-Test for Single Sample**

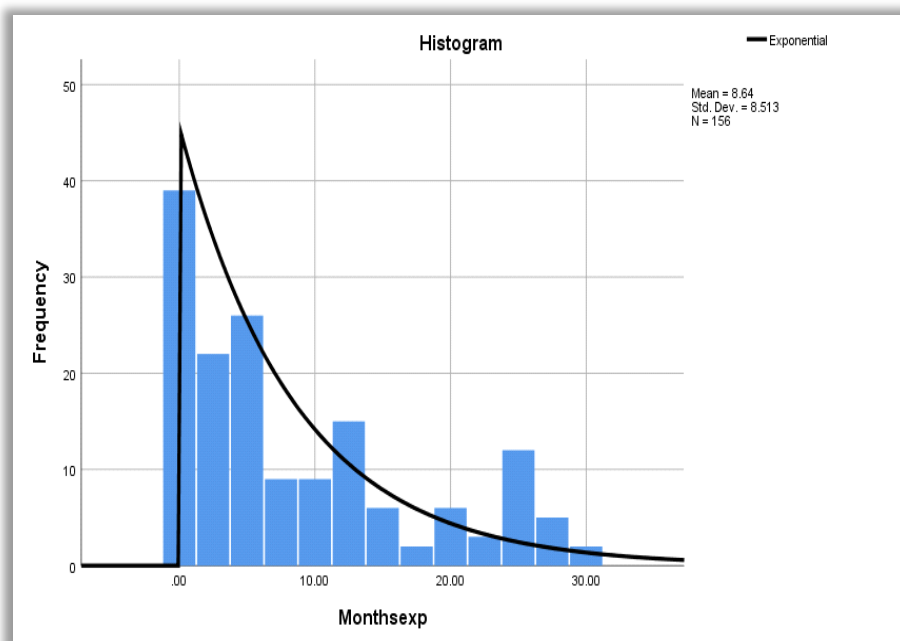
We now conduct a t test for means of a single sample to test whether the sample comes from a population which has a mean equaling to 8.547.

It is known that for large samples, by Central Limit theorem, any distribution will be approximated by a Normal Distribution, and therefore, we apply the t-test for single sample.

One-Sample Test						
Test Value = 8.547009						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
monthsexp	.133	155	.894	.09081	-1.2555	1.4371

Based on the two tailed significance value of 0.894, we compare it with level of significance, i.e., 0.05. It is to be concluded that based on the given data, not enough evidence is found that would support rejection of Null Hypothesis.

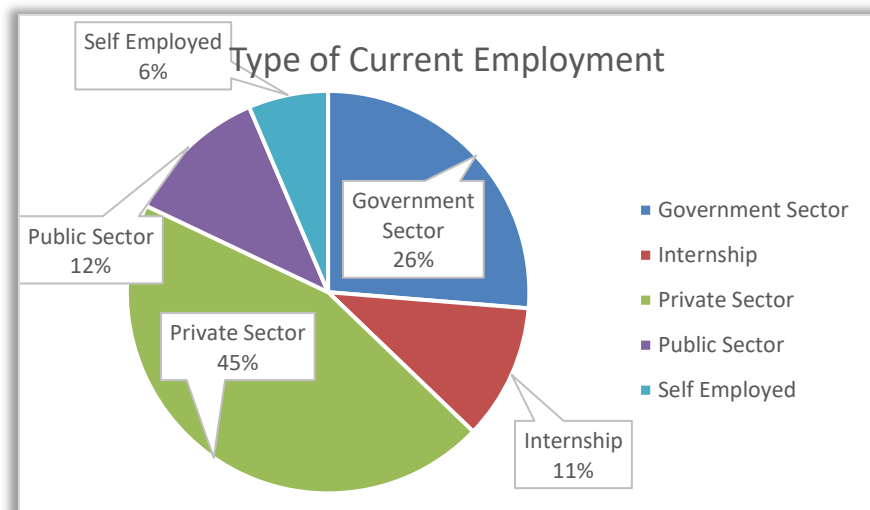
Therefore, the sample comes from a population that has a mean 8.547009, which implies that the fit is good enough.



Most of the respondents, which is roughly 73.5% of the total respondents who were involved in the study were able to find their job within 1 year. Although, time taken to find employment is not so long, it cannot be deduced whether a person seeks satisfaction with his current state of employment, as it might be possible that a relatively better job opportunity could be found out. Therefore, it directly rules out the possibility of having a strong evidence in favour of or against educated unemployment.

2. Type Of Current Employment

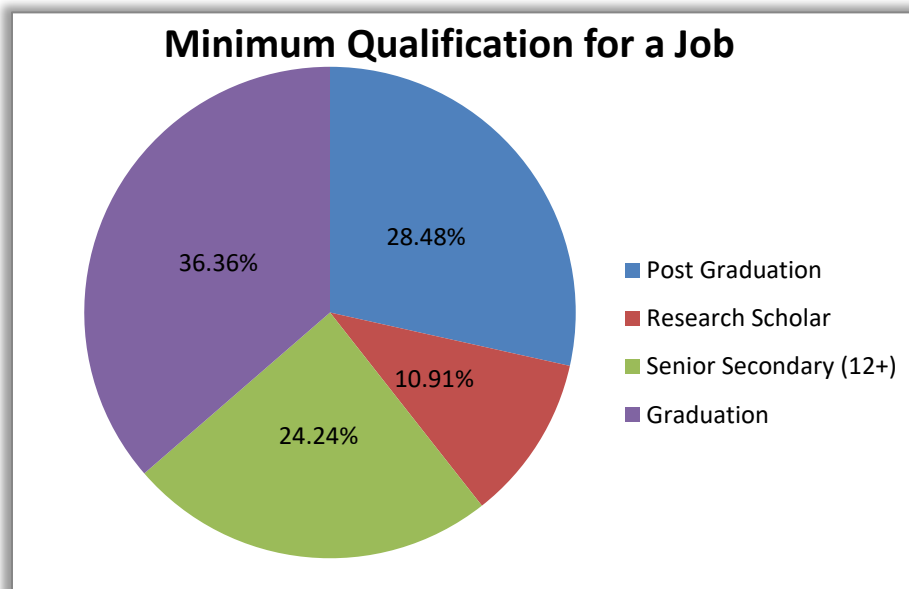
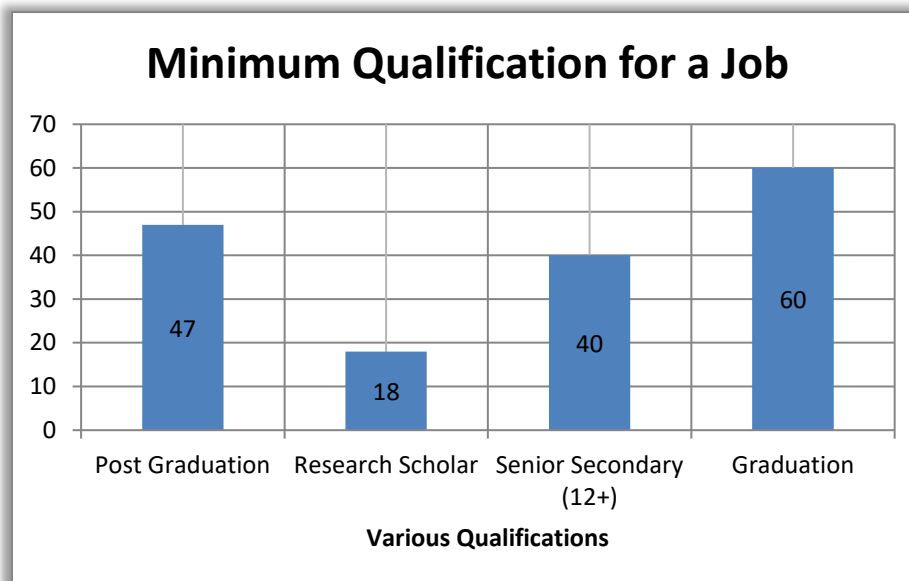
Type of Current Employment	Count Type of Current Employment
Government Sector	43
Internship	17
Private Sector	78
Public Sector	22
Self Employed	10
Grand Total	170



Less than half (45% of the respondents) number of people are employed in the private sector, followed by the government sector that soaks in around 26%.

It is highly depicting that the onset of industrialization and globalization in India has offered much lucrative job opportunities to people in the country. Government sector on the other hand remains a constant job provider to the working class.

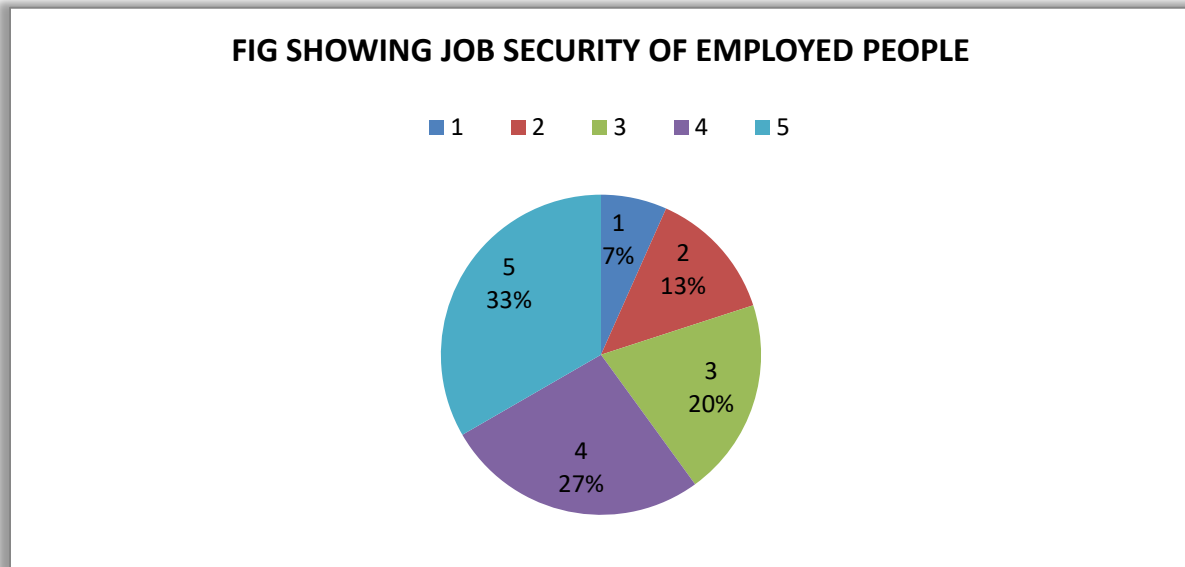
3. Minimum Qualification Required For A Job



Referring to the above charts we infer the following

About 36% of the employed respondents' job requires being at least a **Graduate**. 28% jobs have **Post Graduation** as the minimum qualification. 24% jobs have **Senior Secondary (12+)** as minimum qualification and about 11% jobs require being a **Research Scholar**. So, **Graduation** is the most demanded educational qualification.

4. Scale Of Job Security



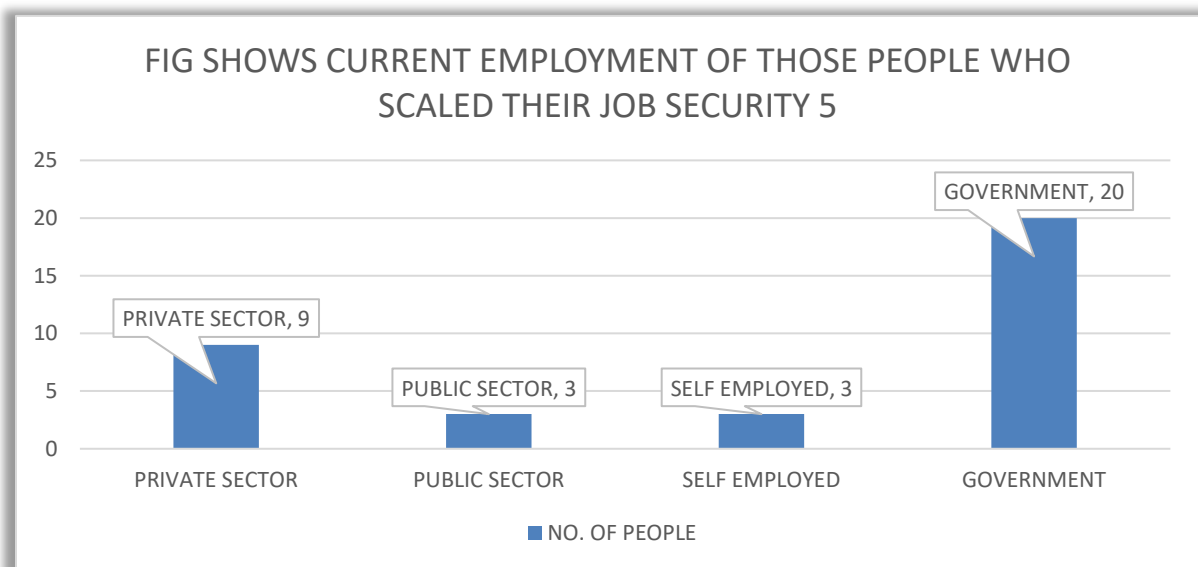
INDEX FOR JOB SECURITY

- 1-LEAST SECURE
- 2-UNSECURE
- 3-SECURE
- 4-MORE SECURE
- 5-MOST SECURE

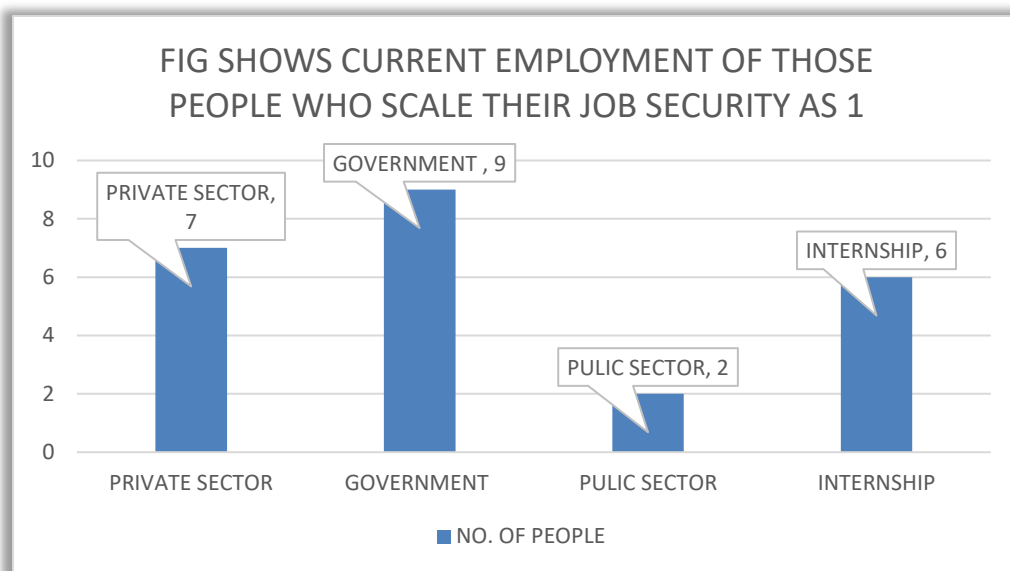
As far as job security is concerned, majority of the respondents corresponded to their job being highly secure. For those who considered giving their job security a grade of 4, their share is significant too. A combined proportion of about 60% is hence, quite representative of the fact that those people would like to get a secured job.

5. Job Security V/S Employment Sector

COUNT OF TYPE OF CURRENT EMPLOYMENT	SCALE FROM MINIMUM TO MAXIMUM JOB SECURITY					
	1: Least Secure	2: Unsecure	3: Secure	4: More Secure	5: Most Secure	Grand Total
Government Sector	9	2	2	10	20	43
Internship	6	3	5	3		17
Private Sector	7	9	20	33	9	78
Public Sector	2	1	7	9	3	22
Self Employed		1	2	4	3	10
Grand Total	24	16	36	59	35	170



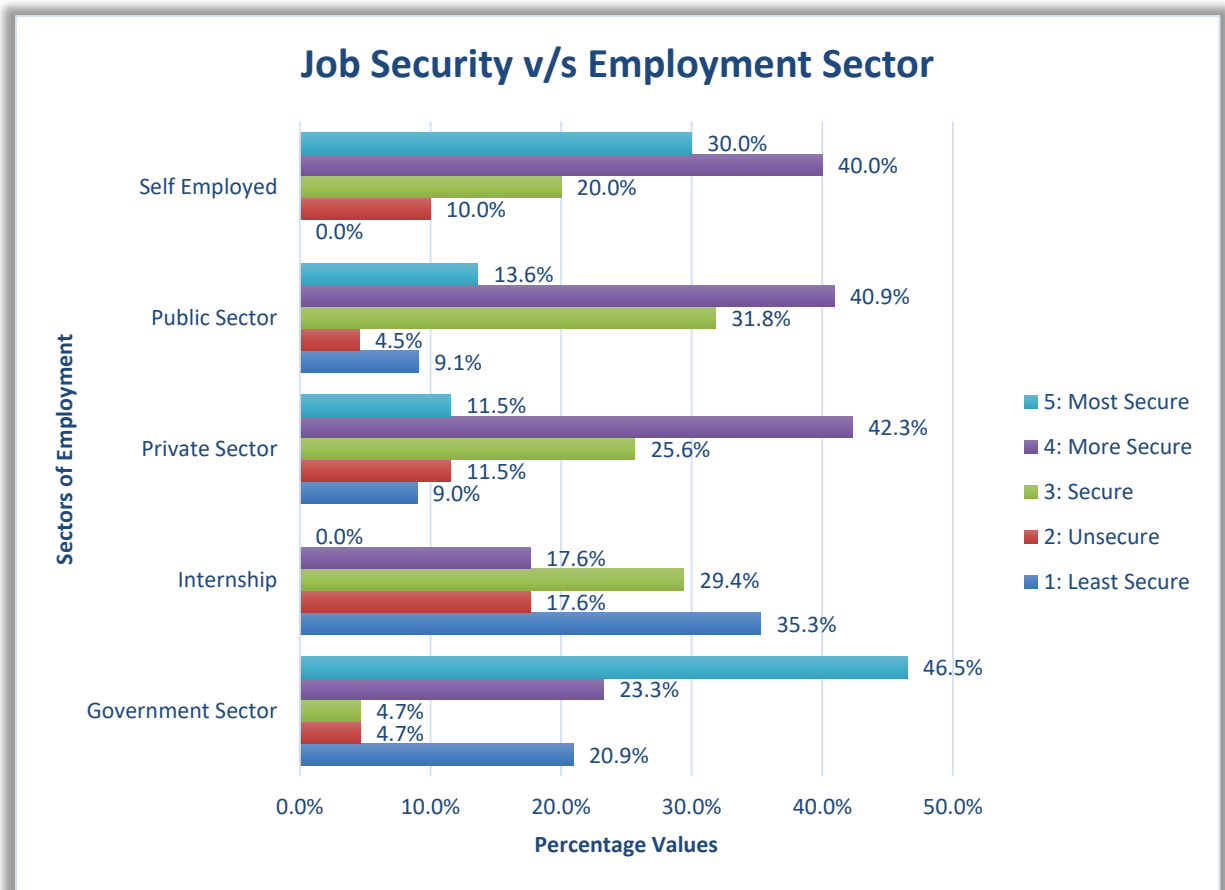
Most of the people who scaled their job security as 5 are working in government sector, followed by private sector. As it is evident that 48.78% of the respondents are engaged in government sector do scale their job security with 5 index. Thus, government jobs are quite secure for the employees. While out of 70 respondents who work in corporate, only 9 feel their jobs are most secured, which corresponds to only 12.8%. Therefore, private sector is quite insecure for people who need a safe job.



When it comes to employed people giving job security scale 1 to their jobs, there can be seen no distinction among different sectors. It can be deduced that these jobs could be quick handy jobs, or jobs which are put on a contractual basis. Therefore, job security as a grade of 1 is quite distributed among different sectors.

PERCENTAGE OF WORKING POPULATION V/S JOB SECURITY

TYPE OF CURRENT EMPLOYMENT	SCALE OF JOB SECURITY (% TABLE)					Grand Total
	1: Least Secure	2: Unsecure	3: Secure	4: More Secure	5: Most Secure	
Government sector	20.93	4.65	4.651	23.26	46.51	100
Internship	35.29	17.6	29.41	17.65	0	100
Private Sector	8.974	11.5	25.64	42.31	11.54	100
Public Sector	9.091	4.55	31.82	40.91	13.64	100
Self Employed	0	10	20	40	30	100
Grand Total	14.12	9.41	21.18	34.71	20.59	100



1. In government sector maximum percentage of people have most secure job.
2. In internship maximum percentage of people have least secure job.
3. In private sector maximum percentage of people have more secure job.
4. In public sector maximum percentage of people have more secure job.
5. In self-employed section maximum percentage of people have more secure job.

From the above data we can infer that government sector provides maximum job security.

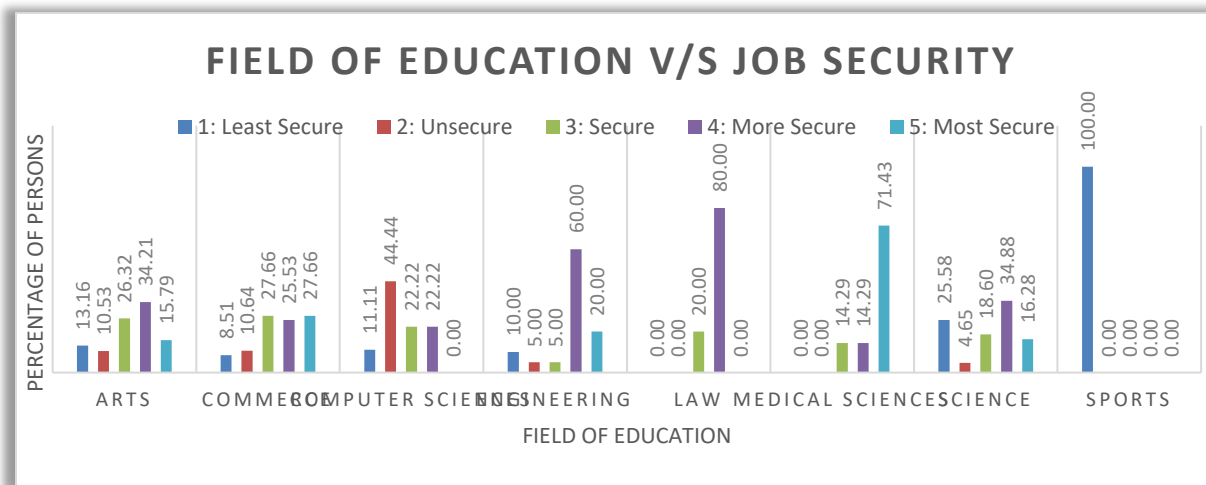
6. Job Security V/S Field Of Education

FIELD OF EDUCATION Row Labels	SCALE OF JOB SECURITY					Grand Total
	1: Least Secure	2: Unsecure	3: Secure	4: More Secure	5: Most Secure	
Arts	5	4	10	13	6	38
Commerce	4	5	13	12	13	47
Computer Sciences	1	4	2	2		9
Engineering	2	1	1	12	4	20
Law			1	4		5
Medical Sciences			1	1	5	7
Science	11	2	8	15	7	43
Sports	1					1
Grand Total	24	16	36	59	35	170

FIELD OF EDUCATION	SCALE OF JOB SECURITY (In terms of %age)					
Row Labels	1: Least Secure	2: Unsecure	3: Secure	4: More Secure	5: Most Secure	Grand Total
Arts	13.16	10.53	26.32	34.21	15.79	100.00
Commerce	8.51	10.64	27.66	25.53	27.66	100.00
Computer Sciences	11.11	44.44	22.22	22.22	0.00	100.00
Engineering	10.00	5.00	5.00	60.00	20.00	100.00
Law	0.00	0.00	20.00	80.00	0.00	100.00
Medical Sciences	0.00	0.00	14.29	14.29	71.43	100.00
Science	25.58	4.65	18.60	34.88	16.28	100.00
Sports	100.00	0.00	0.00	0.00	0.00	100.00

People from medical sciences background have scaled their jobs as secured, as about 85% of them give 4 & 5 ratings to their jobs.

While the people coming from arts, commerce and sciences section have a fairly distributed job security index for their jobs. Computer sciences jobs do not offer any significant job security to employees. While sports jobs could be seen a least secured ones from the given figure, but it is highly absurd to draw any conclusion for there was a single response in the category of sports as minimum qualification.



7) Months Taken To Find Employment Since Completion Of Minimum Educational Qualifications Requirement For The Job V/S Vocational Training

METHOD # 01: ONE WAY ANOVA

We conduct a Single Factor ANOVA to test the significance of vocational training's effect over the months taken to find employment since completion of minimum qualifications for the post.

0: Respondent has taken Vocational Training

1: Respondent hasn't taken Vocational Training

While conducting a Single Factor ANOVA, we have hypothesize as follows:

H₀: All the levels (0 and 1) are homogenous in nature

H₁: There is significant difference between the two levels.

Between-Subjects Factors			
		Value Label	N
VocTra	0	No	106
	1	Yes	64

ANOVA					
Months					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.925	1	5.925	.041	.839
Within Groups	24014.194	168	142.942		
Total	24020.119	169			

Since, the p value of 0.839 > Level of Significance (taken to be 0.05), we conclude that on the basis of the given data, we do not find enough evidence to reject the null hypothesis.

Therefore, all the levels are homogenous in nature, and it follows that taking vocational training does not affect at all the response variable value, which are months taken to find employment.

However, one may question the validity of F test, if the population values do come from an exponential distribution. Therefore, we consider conducting non-parametric tests for the same.

METHOD # 02: MANN WHITNEY U TEST

We conduct a Mann Whitney U Test to determine

Ho: The samples originating from the four categories belong to the same population,

H1: At least one of those samples belongs to different population.

→ NPar Tests

Mann-Whitney Test

		Ranks		
	VocTra	N	Mean Rank	Sum of Ranks
Months	No	106	85.66	9079.50
	Yes	64	85.24	5455.50
	Total	170		

Test Statistics^a

Months	
Mann-Whitney U	3375.500
Wilcoxon W	5455.500
Z	-.053
Asymp. Sig. (2-tailed)	.958

a. Grouping Variable: VocTra

Based on the sample data, we do not find enough evidence to support rejection of Null Hypothesis, and therefore, again, being parallel to ANOVA 1 Way test, it is concluded that Vocational Training does not affect the Months Taken to find Employment.

This effectivly throws light upon the fact that educated unemployment is an independent factor that affects the time to employment since completion of minimum qualifications for job.

8) Months Taken To Find Employment Since Completion Of Minimum Educational Qualifications Requirement For The Job V/S Minimum Educational Qualification

It is to be tested that level of minimum qualification has some impact over the months taken to find employment. As our independent variable is categorical in nature, we have four levels to classify minimum qualifications of the respondents who have found a job:

- 0:Senior Secondary (10+2)
- 1: Graduate
- 2:Post Graduate
- 3:Research Scholars

(We have coded the levels for ease of applying regression analysis over the data)

METHOD #01: ONE WAY ANOVA & GENERAL LINEAR MODEL APPROACH USING INDICATOR VARIABLES

We establish a Univariate General Linear Model between the Months taken to find Employment and the Minimum Educational Qualifications (Using Indicator Variables) of the respondents.

The model assumed is as follows:

$$Y = \mu + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 +$$

where,

Y corresponds to the Months value submitted by the respondent.

X_1 is indicator variable representing the label values (0:absence & 1: presence) for Graduate Students

X_2 is indicator variable representing the label values (0 & 1) for Post Graduate Students

X_3 is indicator variable representing the label values (0 & 1) for Research Students

β_k are the regression coefficients in the model

is the error term associated with a response value

We hypothesize as follows:

H₀: There is no valid regression between the variables, ie, all regression coefficients equal to 0.

H₁: At least one regression coefficient does not equal to 0, and hence contributes to the model.

The Model Summary and ANOVA table is attached below:

Between-Subjects Factors		
		N
MinEdQu	0	32
	1	57
	2	53
	3	28

Tests of Between-Subjects Effects					
Dependent Variable: Months					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1157.361 ^a	3	385.787	2.801	.042
Intercept	14306.906	1	14306.906	103.878	.000
MinEdQu	1157.361	3	385.787	2.801	.042
Error	22862.759	166	137.727		
Total	41901.250	170			
Corrected Total	24020.119	169			
a. R Squared = .048 (Adjusted R Squared = .031)					

On observing the p value of 0.042, it is to be concluded that the given data gives enough evidence to reject null hypothesis. It is therefore understood that some (or even a single

one) regression coefficients are not equal to 0, and do contribute to the model. But a major determiner would be the value of coefficient of determination, R Squared which equals 0.048, and is extremely low, thereby suggesting to discard the model.

Hence, our regression approach fails.

But, as far as ANOVA is concerned, we can see that there exists some significant difference between the levels so classified.

Hence, we apply the Least Significant Difference method to determine which of the pairs differ significantly from each other:

We hypothesize as follows at 0.05 level of significance:

H₀: The two levels are homogenous, and do not show significant difference between their means.

H₁: The two class means are not equal to 0, and they do differ significantly from each other.

Post Hoc Tests

MinEdQu

Multiple Comparisons						
Dependent Variable: Months						
LSD						
(I) MinEdQu	(J) MinEdQu	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
0	1	-5.80 [*]	2.592	.027	-10.92	-.68
	2	-7.25 [*]	2.627	.006	-12.44	-2.06
	3	-3.53	3.037	.247	-9.53	2.46
1	0	5.80 [*]	2.592	.027	.68	10.92
	2	-1.45	2.239	.520	-5.87	2.98
	3	2.27	2.708	.403	-3.08	7.62
2	0	7.25 [*]	2.627	.006	2.06	12.44
	1	1.45	2.239	.520	-2.98	5.87
	3	3.72	2.742	.177	-1.70	9.13
3	0	3.53	3.037	.247	-2.46	9.53
	1	-2.27	2.708	.403	-7.62	3.08
	2	-3.72	2.742	.177	-9.13	1.70

Based on observed means.
The error term is Mean Square(Error) = 137.727.
*. The mean difference is significant at the 0.05 level.

Since the post hoc test including LSD is necessarily two sided, we compare it against Level of Significance = 0.05, and now on comparison with the p values, we find the following pairs of levels which differ significantly from each other:

- Code 1 and Code 0 differ significantly from each other, that shows values Months taken to find a job for Senior Secondary and Post Graduate students differ significantly from each other.
- Code 2 and Code 0 differ significantly from each other, that shows values Months taken to find a job for Senior Secondary and Graduate students differ significantly from each other.

However again, one may question the validity of t and F tests, if the population values do come from an exponential distribution. Therefore, we consider conducting non-parametric tests for the same.

METHOD # 02: KRUSKAL WALLIS RANK TEST FOR ANOVA

We conduct a Kruskal Wallis Rank Test to determine

Ho: The samples originating from the four categories belong to the same population,

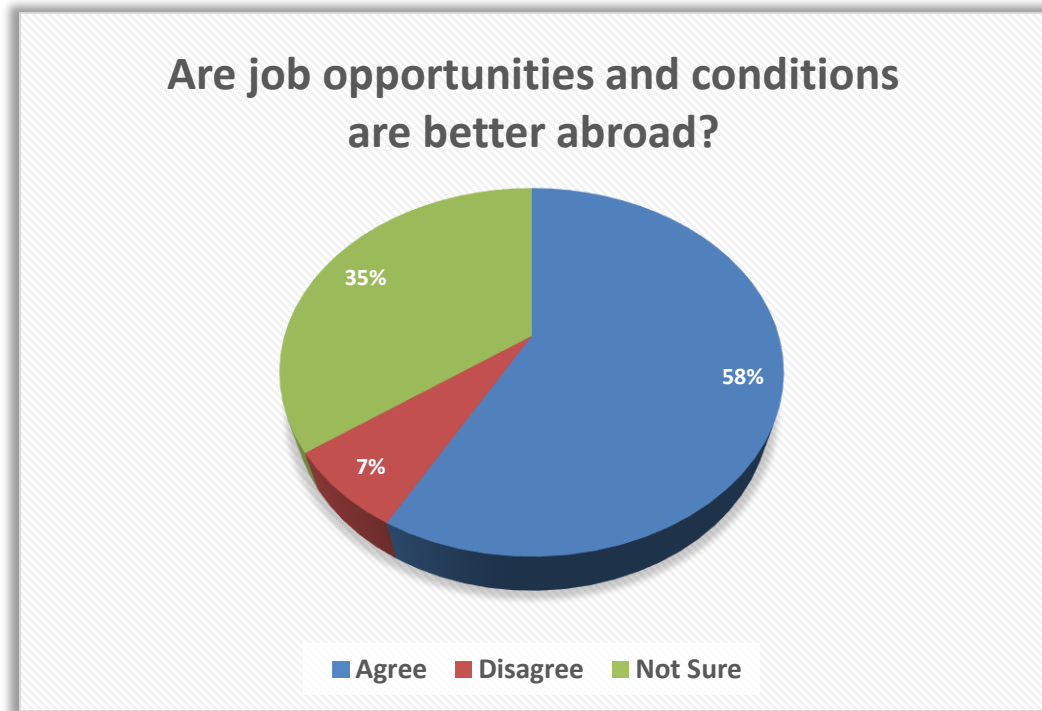
H1: At least one of those samples belongs to different population.

→ Kruskal-Wallis Test			
Ranks			
	MinEdQu	N	Mean Rank
Months	0	32	70.28
	1	57	84.37
	2	53	96.21
	3	28	84.93
	Total	170	
Test Statistics ^{a,b}			
Months			
Kruskal-Wallis H		5.635	
df		3	
Asymp. Sig.		.131	
a. Kruskal Wallis Test			
b. Grouping Variable: MinEdQu			

Since the asymptotic significance value of 0.131 > Level of Significance (0.05), we need to accept the null hypothesis, and conclude that all the samples belong to the same population. **Therefore, with Non Parametric Tests, the conclusion is to assume that level of educational qualifications does not affect the Number of Months taken to find Employment.**

4. SECTION-D (GENERAL AWARENESS)

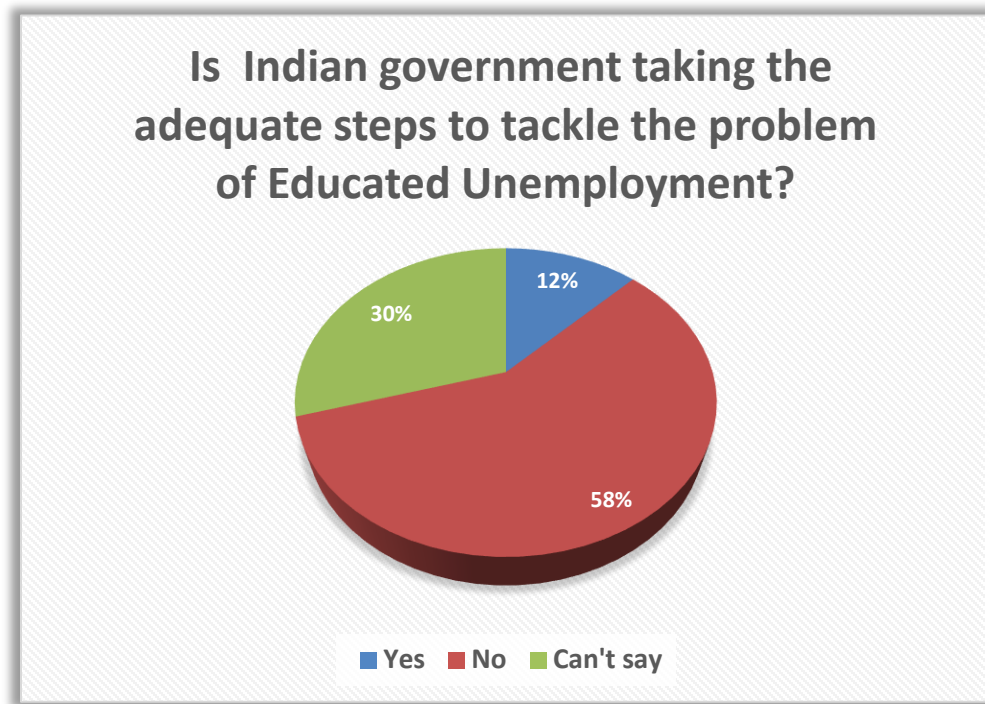
1. Better Opportunities Abroad??



More than half of the respondents, which account to 58% agreed that job opportunities are better abroad. 35% of the people disagree with the above statement.

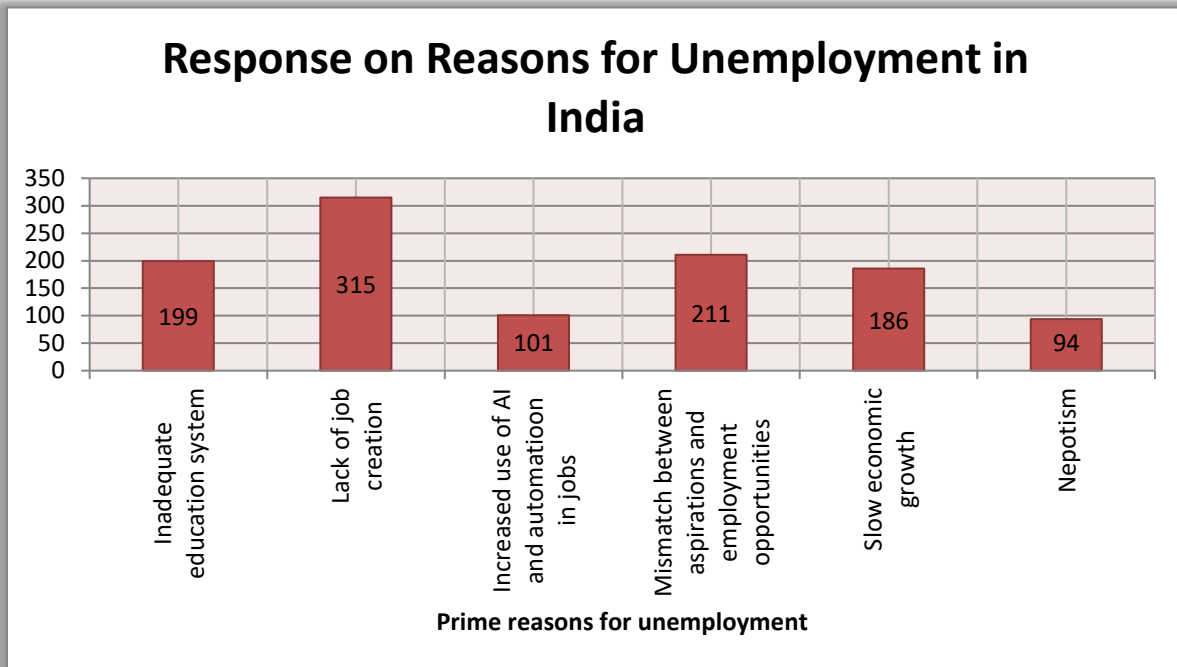
This might be putting light to the fact that Brain Drain is a common statement when it comes to describing the quality of educated youth left in the country itself. Since, young educated class of the society thinks that they could have better jobs in foreign countries, they tend to move out where they can find remuneration and opportunities best suited to their educational qualifications. Therefore, the country is left with no sufficient qualitative educated youth to count upon.

2. Steps Taken by Government



Only 1/3rd of population agrees that Indian government is taking adequate steps to tackle the problem of educated unemployment. While more than half of the people of the people disagree with the above statement. This shows that either government is not taking steps to curb the issue, or if it is taking the steps, then people are unaware of the government policies.

3. Reasons for Educated Unemployment in India



Most of the people (65%) reckon that the prime reason for educated unemployment is **Lack of Job Creation** which highlights the fact that the pace of job creation remains far below than what is needed to absorb people added to workforce every year. **Other reasons like Mismatch between aspirations and employment opportunities, Inadequate education system, Slow economic growth, Increased use of AI and automation in jobs and Nepotism** follow its lead in that very chronology.

5. SECTION-D (DISTRIBUTION FITTING)

The objective of this section is to find out a distribution for the duration of unemployment after completion of desired level of education. The first step is to find out the variables that significantly effect the duration of unemployment.

The possible candidates are:

1) Qualification

Here it is to be tested whether the level of an individual's highest qualification have a significant effect on the months taken to find employment since completion of their studies. Qualification of an individual is divided into four levels as follows:

1. Senior Secondary (10+2)
2. Graduate
3. Post Graduate
4. Research Scholar

The hypothesis for the test is set up as:

H₀: The levels are homogenous and do not show a significant effect on the time taken to find employment

H₁: The levels of qualification show a significant effect on time taken to find employment

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	16.684	3	.001

Test of equality of survival distributions for the different levels of Qualifications.

The log test conducted above provides a p-value equals to $0.001 < 0.05$. Hence at 5% level of significance we have enough evidence to reject the null hypothesis, therefore concluding that the level of an individual's qualifications has significant effect on the time taken to find employment.

2) Field of education

Here it is to be tested whether an individual's field of education have a significant effect on the months taken to find employment after completion of their studies. The variable of field of education is divided into the following levels:

1. Commerce
2. Science
3. Arts
4. Computer Science
5. Engineering
6. Medical Science
7. Sports
8. Law

The hypothesis for the test is set up as:

H₀: Field of education does not have a significant effect on the time taken to find employment.

H₁: Field of education has a significant effect on time taken to find employment

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	10.730	7	.151

Test of equality of survival distributions for the different levels of Field of Education.

The log test conducted above provides a p-value equals to $0.151 > 0.05$. Hence at 5% level of significance we do not have enough evidence to reject the null hypothesis,

therefore concluding that the level of an individual's field of education does not have a significant effect on the time taken to find employment.

3) Vocational training

Here it is to be tested whether engaging in vocational training have a significant effect on the months taken to find employment since completion of their studies.

The hypothesis for the test is set up as:

H0: Undertaking vocational training does not have a significant effect on time taken to find employment.

H1: Undertaking vocational training has a significant effect on time taken to find employment.

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	2.562	1	.109

Test of equality of survival distributions for the different levels of Have you undertaken any Vocational Training ?

The log test conducted above provides a p-value equals to $0.109 > 0.05$. Hence at 5% level of significance we do not have enough evidence to reject the null hypothesis, therefore concluding that undertaking vocational training does not have a significant effect on the time taken to find employment. This contradicts the usual belief that vocational training helps in finding employment.

4) Kaplan Meier

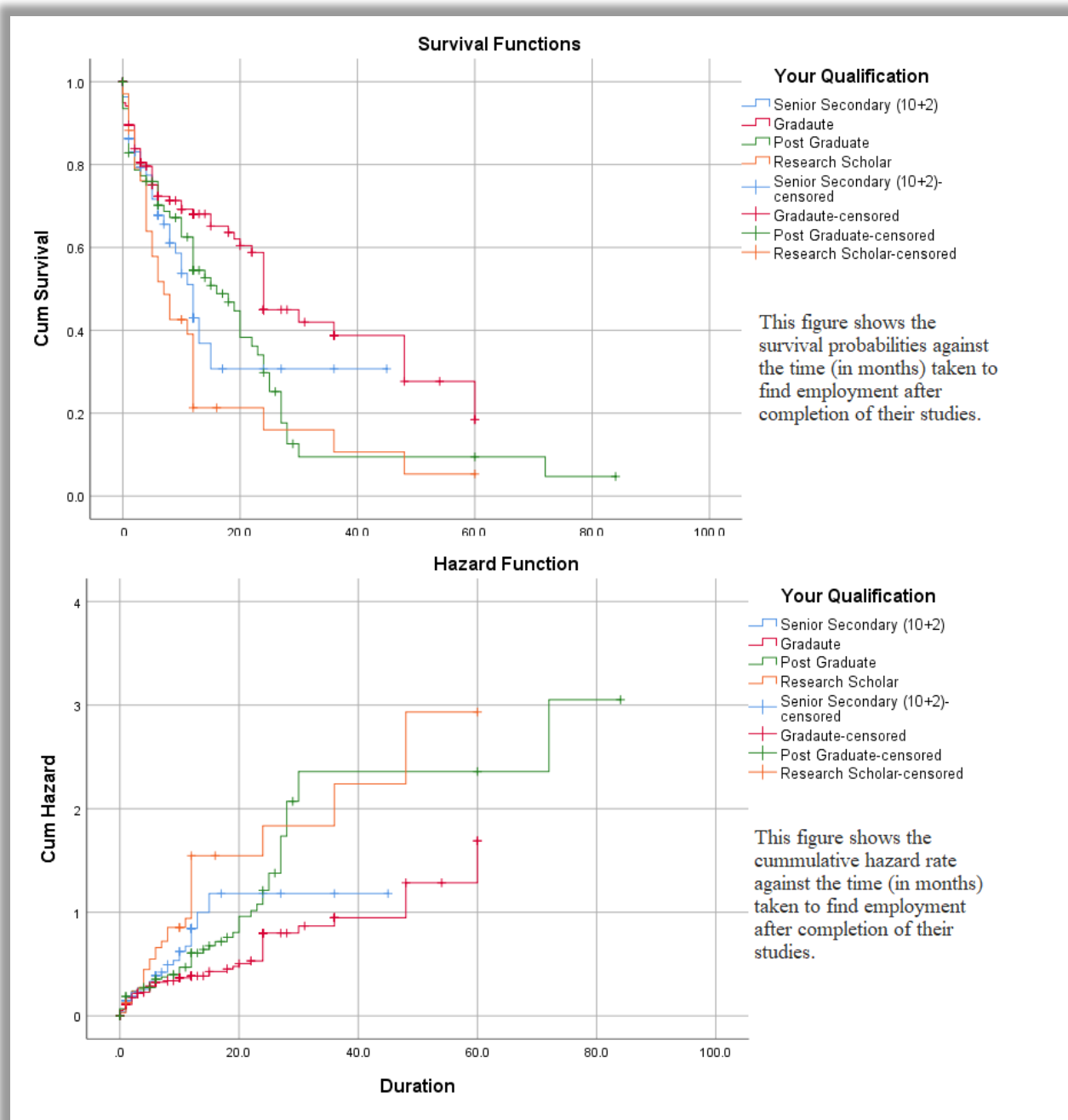
From the tests above it can be concluded that only an individual's level of qualification has a significant effect on the months taken to find employment after completion of their studies, hence Kaplan Meier is applied while factoring in the effect of different levels of qualifications.

Finding employment is treated as the concerned event and the data values from unemployed section of the sample are treated as censored.

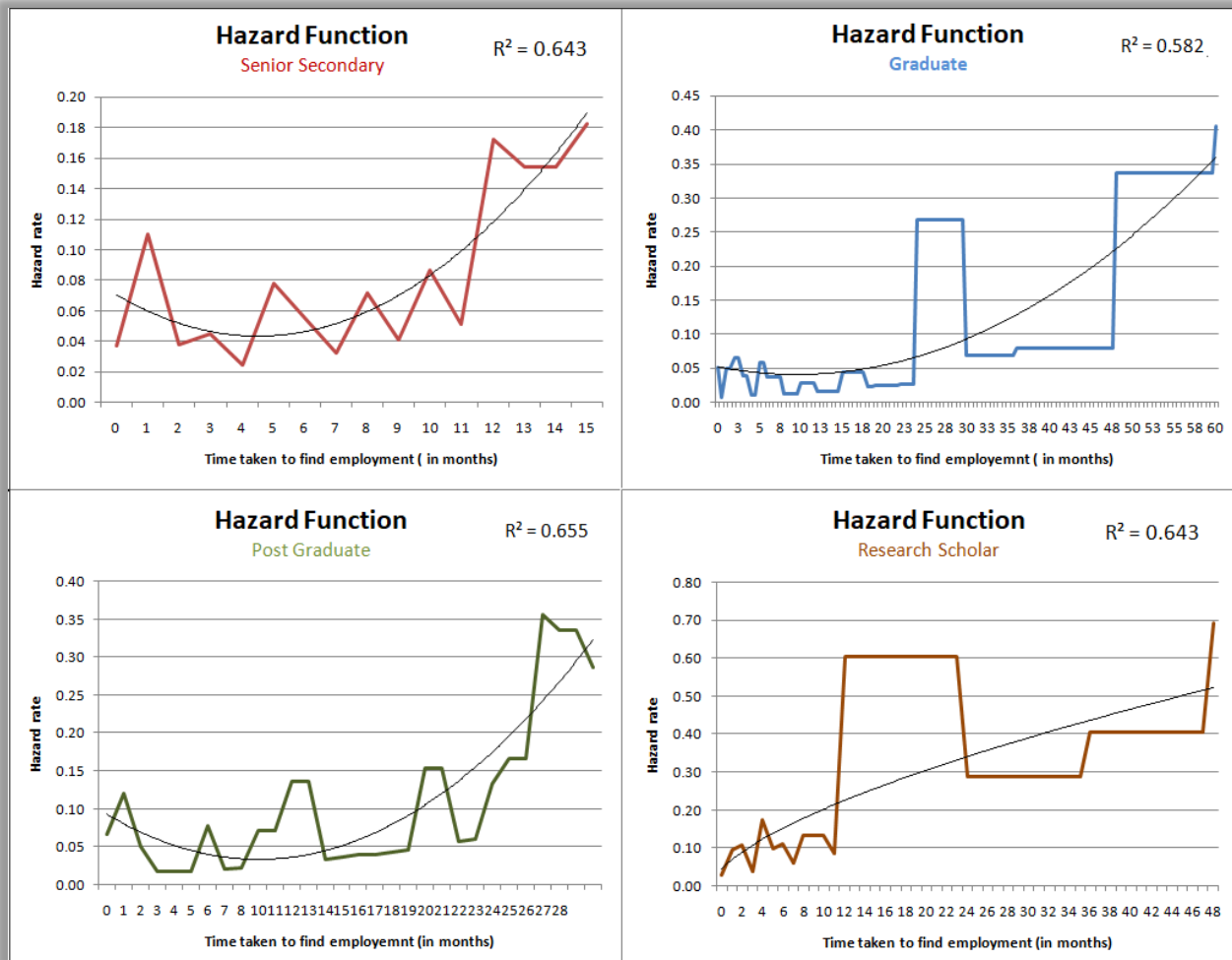
In the following estimation it is assumed that if censoring and event occurs at the same time then censoring occurs right after the occurrence of event.

The survival table obtained through Kaplan Meier is given below:

Qualification	Time	Survival Probability	N of Cumulative Events	N of Remaining Cases	Qualification	Time	Survival Probability	N of Cumulative Events	N of Remaining Cases
Senior Secondary (10+2)	.000	.963	3	79	Post Graduate	.000	.935	5	72
	1.000	.863	10	60		1.000	.828	13	62
	2.000	.831	12	52		2.000	.787	16	57
	3.000	.794	14	43		3.000	.773	17	56
	4.000	.775	15	40		4.000	.759	18	55
	5.000	.716	18	37		6.000	.702	22	49
	6.000	.678	20	35		7.000	.687	23	46
	7.000	.656	21	30		8.000	.672	24	45
	8.000	.611	23	27		10.000	.625	27	40
	9.000	.586	24	24		12.000	.545	32	34
	10.000	.537	26	22		14.000	.527	33	29
	11.000	.510	27	19		15.000	.508	34	27
	12.000	.430	30	16		16.000	.488	35	25
	13.000	.368	31	6		18.000	.468	36	23
	15.000	.307	32	5		19.000	.447	37	21
Graduate	.000	.949	7	130		20.000	.383	40	18
	.500	.941	8	124		22.000	.362	41	17
	1.000	.896	14	118		23.000	.340	42	16
	2.000	.838	21	102		24.000	.298	44	14
	3.000	.805	25	97		25.000	.252	46	11
	4.000	.796	26	90		27.000	.176	49	7
	5.000	.751	31	83		28.000	.126	51	5
	6.000	.723	34	78		30.000	.095	52	3
	8.000	.713	35	72		72.000	.047	53	1
	10.000	.692	37	66	Research Scholar	.000	.971	1	33
	12.000	.681	38	60		1.000	.882	4	30
	15.000	.651	40	44		2.000	.791	7	26
	18.000	.636	41	42		3.000	.761	8	25
	19.000	.620	42	39		4.000	.639	12	21
	20.000	.604	43	38		5.000	.578	14	19
	22.000	.588	44	36		6.000	.517	16	17
	24.000	.450	52	26		7.000	.487	17	16
	30.000	.420	53	14		8.000	.426	19	14
	36.000	.387	54	12		11.000	.390	20	11
	48.000	.277	56	5		12.000	.213	25	6
	60.000	.184	57	2		24.000	.160	26	3
						36.000	.106	27	2
						48.000	.053	28	1



Hazard rates obtained through the above method are then used to estimate the distribution followed by the data factored by the levels of qualifications. Crude fitting of second degree polynomials to the above obtained hazard rates revealed the following findings.



The hazard functions corresponding to the three levels viz. Senior Secondary, Graduation, Post Graduate show a bath-tub shaped hazard function while the hazard function corresponding to level Research Scholar show a monotonously increasing hazard rate. Hence we can conclude that the time taken to find employment follow distributions with the hazard functions as described above.

A three parameter distribution with increasing, decreasing, bathtub and upside-down bathtub shaped hazard rates is given below as:

Probability Distribution Function:

$$f(t) = \alpha \beta \lambda^{\beta-1} (1 + \lambda t^{\beta})^{\alpha-1} \exp \{1 - (1 + \lambda t^{\beta})^{\alpha}\}$$

Survival Function:

$$S(t) = \exp \{1 - (1 + \lambda t^{\beta})^{\alpha}\}$$

Hazard Function:

$$h(t) = \alpha \beta \lambda^{\beta-1} (1 + \lambda t^{\beta})^{\alpha-1}$$

Where; α, β are shape parameters

λ is the scale parameter

Through the method of maximum likelihood estimation method, maximizing the following likelihood function provides the required estimates:

$$L = \prod_{\text{censored obs.}} S(t_i) * \prod_{\text{uncensored obs.}} f(t_i)$$

Qualification	Parameters		
	α	β	λ
Senior Secondary	60.0000	0.8499987	0.0009716114
Graduate	30.0000	1.099998	0.0004922494
Post Graduate	92.8924	0.8999999	0.0005181526
Research Scholar	11.99905711	0.56584717	0.01382976

The fitted survival functions are given below as:

