

## **PROJECT REPORT**

***"A study on human migration process from 'rural and urban' to urban area in Amravati"***

**YEAR: -2022 - 2023**



**Department of Statistics**

**G.V.I.S.H., Amravati**

**Government Vidarbha Institute of Science and Humanities  
(Autonomous)**

**For the degree of M.Sc. in Statistics**

**Submitted by:**

**Miss. Vaibhavi Sanjay Kalmegh**

**Under the guidance of**

**Dr. (Mrs.) P. B. Lakhani**

**Professor, Department of Statistics,**

**G.V.I.S.H. Amravati**

## **CERTIFICATE**

This is to certify that **Miss. Vaibhavi Sanjay Kalmegh** have completed the project entitled “**A study on human migration process from ‘rural and urban’ to urban area in Amravati**” under the guidance of Dr. (Mrs.) P. B. Lakhani, Professor in Statistics, for the partial fulfillment of degree of M.Sc. II [Statistics].

Place: Amravati

Date:

**Dr. (Mrs.) P. B. Lakhani**

**Professor**

**Department of Statistics**

**G.V.I.S.H. Amravati**

**Dr. S. R. Kolhekar**

**Prof. & Head of the**

**Department of Statistics**

**G.V.I.S.H. Amravati**

**Year: 2022-2023**

## ACKNOWLEDGEMENT

It is an honor to express my deep gratitude to Dr. (Mrs.) P. B. Lakhani for her valuable guidance, enthusiastic encouragement, constant support, and inspiration throughout the project.

My sincere thanks to Dr. S. R. Kolhekar, Head of the Department of Statistics, G.V.I.S.H., Amravati, for her cooperation throughout the project work. I thank her for providing me with a good library with excellent books and a computer laboratory.

I also thank Dr. N.W. Andure for her guidance and well wishes during the project work. To conclude, I express my sincere thanks to the non-teaching staff of the department and to all those names I might have forgotten to add here who contributed directly or indirectly to complete this project.

Miss. Vaibhavi S. Kalmegh.

Place: Amravati

Date:

## **DECLARATION**

I hereby declare that the work incorporated in this report entitled  
**“A STUDY ON HUMAN MIGRATION PROCESS FROM ‘RURAL AND  
URBAN’ TO URBAN AREA IN AMRAVATI”**

in partial fulfillment of the requirements for the degree of

**Master of Science in Statistics**

is the outcome of the original study undertaken by me and it has not been submitted earlier to any other university or institute for the award of any degree or diploma. Also, the data has not been derived from any thesis/publication of any university or scientific organization. The sources of materials used and all assistance received during the investigation have been duly acknowledged.

Miss. Vaibhavi S. Kalmegh .....

Place: Amravati

Date

## INDEX

<b>Sr. no</b>	<b>Content</b>	<b>Page no.</b>
	<b><i>Chapter 1: Introduction</i></b>	<b>1</b>
	1.1 What is Migration?	2
	1.2 Types of Migration.	3 - 8
	1.2.1 International Migration (world migration).	
	1.2.2 External Migration.	
	1.2.3 Internal Migration.	
1.	1.2.4 Subcategories of Migration.	
	1.3 Root causes of Migration.	9 - 10
	1.4 Positive impact of Migration.	11
	1.5 Negative impact of Migration.	12
2.	<b><i>Chapter 2: Review of Literature</i></b>	13 - 16
	<b><i>Chapter 3: Research Methodology and plan work</i></b>	17
	3.1 Objectives.	18
	3.2 Data collection.	19 - 22
	3.3 Research Methodology.	23
3.	3.4 Software used.	24
	3.5 Opinion about the sample collected.	24
	<b><i>Chapter 4: Data Analysis</i></b>	25
	4.1 Sample Profile: Frequency distribution.	26 - 40
4.	4.2 In-depth Analysis.	41 - 69
5.	<b><i>Chapter 5: Conclusions</i></b>	70 - 74
6.	<b><i>Chapter 6: Reference</i></b>	75 - 76
7.	<b><i>Annexure</i></b>	77 - 86

# Chapter No. I

## Introduction



- 1.1 What is Migration?
- 1.2 Types of Migration.
  - 1.2.1 International Migration (world migration).
  - 1.2.2 External Migration.
  - 1.2.3 Internal Migration.
  - 1.2.4 Subcategories of Migration.
- 1.3 Root cause of Migration.
- 1.4 Positive impact of Migration.
- 1.5 Negative impact of Migration.

## **I.I What is Migration?**

### **Definition of migration:**

The basic definition of migration is given as: Migration is defined as the movement of people from one place to another with the intention of settling, permanently or temporarily, at a new location (geographical region).

### **What is migration?**

Migration has become a universal phenomenon in today's modern world. In the view of fact that, due to the expansion of transport and communication, it has become a part of the worldwide process of urbanization and industrialisation.

In almost many countries, it has been observed that industrialisation and economic development has been accompanied by large-scale movements of people from villages to towns, from towns to other towns, and from one country to another country.

From the demographic point of view, migration is one of the three essential components of population growth of any particular area or region, second is fertility and third is mortality. But whereas both fertility and mortality operate within the biological framework, migration does not. Migration influences the size, composition, and distribution of the population. More importantly, migration affects the social, political, and economic life of the people.

In the past years during the pandemic, covid-19 migration has shown an unbroken upward trend, be it of people who have left their homelands because of the economic or other reasons, or of those who has been forced to leave their homes (such as natural disasters or civil disturbance).

### **Who is a migrant?**

The UN Migration Agency (IOM) defines a migrant as any person who is moving or has moved across an international border or within a state away from his / her habitual place of residence, regardless of

- (1) the person's legal status;
- (2) whether the movement is voluntary or involuntary;
- (3) what the causes for the movement are
- (4) what the length of the stay is.

## I.2 Types of Migration.

### I.2.1 International Migration (world migration).

The movement of people across international borders for the purpose of settling is known as international migration. International migrants relocate from one country to another on a regular basis.

Shorter stays in another country should not be classified as permanent international migration, according to the United Nations. When passengers arrive in a country, they are asked whether they plan to stay for less than 3 months, in which case they are classified as visitors; between 3 and 12 months, in which case they are classified as short-term migrants; or for 12 months or more, in which case they are classified as long-term migrants.

India has not always been affected by in-migration, but historical events such as Partition (1947-51), Buddhist Migration (1954-59), Bangladesh Liberation (1971), and Tamil migrants have all contributed to international in-migration in India.

Indian out-migration began during colonial times, primarily after the end of the slave trade (1833), when a labour shortage prompted Indian migration from states such as Uttar Pradesh and Bihar to countries such as the Philippines and Indonesia's Java for plantation agriculture.

### I.2.2 External Migration.

India's external migration can be divided into three categories:

**Emigration:** India's emigration to various parts of the world.

**Immigration:** People from various countries are immigrating to India.

**Refugee Migration:** Involuntary or forced migration to India in the form of refugees has also been a significant trend.



- **Emigration**

India (17.5 million), Mexico (11.8 million), and China are the top three countries of origin for international migrants (10.7 million).

India (\$78.6 billion), China (\$67.4 billion), and Mexico (\$35.7 billion) were the top three remittance recipients.

- **Immigration**

Immigration is the process by which people become permanent residents or citizens of a different country. Immigration has historically benefited states in terms of social, economic, and cultural benefits. According to the UN Department of Economic and Social Affairs (UNDESA) International Migrant Stock 2019 report, India has overtaken China as the world's leading country of origin for immigrants. The UAE is the most popular destination for Indian migrants, followed by the United States and Saudi Arabia.

- **Refugee Migration**

A refugee is defined as "a person who is outside his country of nationality because of a well-founded fear of persecution for reasons of race, religion, nationality, membership of a particular social group, or political opinion." After leaving his mother country, a refugee does not change his nationality. In India, for example, Tibetan refugees.

### **I.2.3 Internal Migration.**

Internal migration is the movement of people within a country from one defined area to another. At the macro level, there are inter-regional and intra-regional disparities, and at the micro-level, there is a fundamental lack of employment opportunities, resulting in low living standards among various socio-economic groups. Internal migration and international migration are the two types of migration. Internal migration is the movement of people within a country from

one location to another. International migration is the movement of people from one country to another in search of work, a place to call home, or a better standard of living.

**In is general further four streams of migration are identified:**

- (a) Rural to Rural (R-R)
- (b) Rural to Urban (R-U)
- (c) Urban to Urban (U-U) and
- (d) Urban to Rural (U-R)

✓ **Rural to Rural migration:**

According to the 2011 Census, this is the country's most dominant migratory movement.

Marriage ceremonies and working as agricultural laborers are the two main reasons for this migration.

For example, agricultural laborers move from overpopulated to underpopulated areas, such as the Awadh-Rohilkhand plains to the Delta plains.

Government decisions can sometimes have a significant impact.

The resettlement of Sikhs in the Tarai region of Uttar Pradesh is an excellent example.

The government also provided for the resettlement of 20,000 families as part of the Dandakaranya project.

✓ **Rural to urban migration:**

It is the second most noticeable migratory pattern.

Since independence, inter-sectoral migration has been the most dominant migratory trend. It is the result of both push and pull factors in rural and urban areas. However, since the 1981 census, the positive impact of rural development programs has resulted in a reduction in push factors, causing this trend to fall to

second place. Megacities serve as a destination, with rural residents primarily from population surplus states like Uttar Pradesh, Bihar, and MP serving as source areas.



Rural-urban migration is primarily responsible for the country's exponential growth in population and the number of megacities.

Migration from one area to another in search of improved livelihood is a key feature of human history. While some regions and sectors fall behind in their capacity to support populations, others move ahead and people migrate to access these emerging opportunities. Industrialization widens the gap between rural and urban areas, including a shift of the workforce towards industrialized areas.

Migration from rural to urban areas expands due to the following three factors: natural growth of population, reclassification of rural areas as urban in course of time and rural to urban migration.

Around 2/5th of the total urban growth in the Developing countries is accounted by the rural to urban migration. More than a half of the urban labour force works in the informal sector of low-skilled, low productivity, often self-employed jobs in petty sales and services. Any social policy that affects rural and urban incomes will be persuaded by migration; this will affect the economics of the state or the income distribution and even the population growth.

Urban cities provide several services, such as marketing of agricultural products and inputs, repair of agricultural equipment and also education and health services to rural areas. Urban areas are becoming important because they are providing services to the rural population. All the statistics will tell us the number of migrations from rural areas to urban areas, there is no specific job-related migration to explain the exact reason for migrating people.

✓ **Urban to Urban migration:**

It primarily involves people moving from small towns to larger cities. This migration takes place in stages, with people moving from rural areas to small towns and then to larger cities. The majority of urban-to-urban migration occurs in search of better opportunities and a higher standard of living. It is dominated by middle-class individuals. In small towns, it creates a void.

✓ **Urban to Rural migration:**

It is a reversal or push-back migration. It occurs at a high level of urbanization when cities are characterized by overcrowding, hazard growth, and high living costs. It is less because it involves the elderly population migrating primarily after their professional commitments are completed. The technical term for this migratory movement is "counter-current migration".

## I.2.4 Subcategories of Migration.

There are further more subcategories of migration which are:

- **Long term Migration:** Long term migration results in the relocation of an individual or household. A person who moves to a geographical region other than that of his or her usual residence for a period of at least a year (12 months), so that the region of destination effectively becomes his or her new residence. Then the person will be a long-term immigrant.
- **Short term migration:** short term migration involves back and forth movement between a source and destination. A person who moves to a geographical region other than that of his or her usual residence for a period of at least three months but less than a year (12 months) except in cases where the movement to that region is for purposes of recreation, holiday, visits to friends or relatives, business, medical treatment or religious pilgrimage. Then the person will be a short-term migrant.
- **Female Migration:** Women account for 70.7 percent of all internal migrants, and marriage is one of the main reasons for female migration in both rural and urban areas.
- **Male Migration:** One of the most common reasons for male migration in both rural and urban areas is job-related migration. Construction, domestic work, textile, brick kilns, transportation, mines, quarries, and agriculture are among the most common employment sectors for migrants. Rates of urbanization have an impact on wage disparities between rural and urban areas, as well as an increase in demand for labour in urban areas, which can drive up urban wages and increase migration.

## **I.3 THE ROOT CAUSES OF RURAL MIGRATION**

Migration is often a deliberate decision and an important component of household livelihood strategies. The root causes of people deciding to move out from rural to urban areas are as follows:

### **❖ Rural poverty, food insecurity and agricultural:**

More than 75% of the world's poor and food insecure live in rural areas, mostly depending on agricultural production for their subsistence. The rural poor, and especially smallholder family farmers, face considerable difficulties in accessing credit, services, technologies and markets that would allow them to improve the productivity of their natural resources and labour. Migration has become an important part of the strategies of rural households for improving their livelihoods. Moreover, it is seen many of the time that whatever amount is invested by the farmer in agriculture is more than the output from the agriculture. Smallholder family farmers, small-scale fishers, forest-dependent communities and pastoralists are hardest hit by weather related disasters, which are increasing in frequency and intensity. Droughts and related food price volatility increase poverty and hunger.

### **❖ Lack of employment and income generating opportunities:**

Most available jobs in agriculture are associated with low and unstable incomes, poor safety and health conditions, gender inequality in pay and opportunities, and limited social protection. Due to restricted access to training, financial and extension services and processing facilities, more attractive prospects may be limited in rural areas.

### **❖ Marriage: 7 out of 10 migrants moved due to marriage**

Marriage was the prime reason behind migration for more than 71% of the migrants. As we discussed earlier, out of the total internal migrants, 7 percent are women (Census of India 2001) and marriage is one of the major reasons for female migration in both the rural and urban areas.

### ❖ **Education:**

Among the various reasons, education ranks high, not always as a primary motivation, but inherent in two leading causes of migration from these regions—economic opportunity and, to a lesser degree, violence. A solid education, culminating in a high school diploma, translates into improved job prospects, more liveable incomes, and greater economic stability. When students don't have the chance to complete their education, the consequences profoundly impact their lives and their children's, in far-reaching and often irreversible ways.

### ❖ **Moved with family or family reunification:**

Family migration is the term used to categorize the migration of people who migrate due to new or established family ties or other reasons (job transfer) or it sometimes the reunification with a family member who migrated earlier due to the reasons like (job, business, education etc).



### ❖ **Inequality:**

Rural people are drawn to urban areas where they expect to have better employment opportunities and improved access to health, education, and basic services.

### ❖ **Limited access to social protection:**

About 73% of the world population have no adequate access to social protection. The majority live in the rural areas of developing countries, where they face difficulties in managing social, economic and environmental risks.

## I.4 Positive Impact of migration:

- **Labour Demand and Supply:**

Migration fills gaps in demand for and supply of labour, efficiently allocates skilled labour, unskilled labour, and cheap labour.

- **Economic Remittances:**

The economic wellbeing of migrants provides insurance against risks to households in the areas of origin, increases consumer expenditure and investment in health, education and asset formation.

- **Skill Development:**

Migration enhances the knowledge and skills of migrants through exposure and interaction with the outside world.

- **Quality of Life:**

Migration enhances chances of employment and economic prosperity, which in turn improves quality of life. The migrants also send extra income and remittances back home, thereby positively impacting their native place.

- **Social Remittances:**

Migration helps to improve the social life of migrants, as they learn about new cultures, customs, and languages, which help to improve brotherhood among people and ensures greater equality and tolerance.

- **Food and Nutrition Security:**

According to the 2018 State of Food and Agriculture report by Food and Agriculture Organization (FAO), outmigration often leads to improved food and nutrition security for migrants.

- **Demographic Advantage:**

As a result of outmigration, the population density of the place of origin is reduced and the birth rate decreases.

- **Climate Change Adaptive Mechanism:**

Migration has also emerged as a possible adaptive mechanism in the context of climate change and the occurrence of extreme weather events like floods, droughts, and cyclones.



## **I.5 Negative Impact of migration:**

- **Demographic Profile:**

Emigration in large numbers can alter demographic profiles of communities, as most of the young men move out, leaving only the women and elderly to work on the land.

- **Political Exclusion**

Migrant workers are deprived of many opportunities to exercise their political rights, like the right to vote.

- **Population Explosions and the influx of workers** in the place of destinations increase competition.

- **Brain Drain:**

Source state suffers from the loss of human capital.

- **Increased Slum:**

Mass migration results into an increase in slum areas, compromising quality of infrastructure and life at the destination, which further translates into many other problems such as unhygienic conditions, crime, pollution, etc.

# Chapter: 2

## Review Of Literature



*Literature review*

**Srivastava Ravi et al. (2016)**

**“Labour Migration to the Construction Sector in India and its Impact on Rural Poverty”**

This article studies the impact of migration to the construction sector on rural poverty in India based on field survey. The survey has been carried out in two phases, the first phase involving the survey of construction workers in National Capital Region and the second phase comparing the well-being of migrant households with non-migrant households in selected source villages. The study finds that at destinations, workers have poor living and working conditions, lack citizenship rights, entitlements and voice. However, at the cost of hardship, low consumption levels, and possibly a smaller working life span, they manage to save a good portion of their income which they remit or take back home. At origin, migrant households report higher expenditure on consumption, residential housing, and children’s education, as well as some other assets. These differences are a result of higher employment and wage incomes among migrants. The methodology of the study was a quantitative household survey of the construction of workers in the rural areas. 150 workers were interviewed as a sample for this study.

**Abhay Garg and Priyanshu Agarwal (International Journal of Policy Sciences and Law Volume 1, Issue 4) (September 2020)**

**“Analysis of Rural-Urban Migration in India and Impact of COVID-19”.**

This paper examines the Rural-Urban Migration in India and the Impact of the Covid-19 pandemic on the Migrants and the governmental stance and policies on migration during the global pandemic. A major determinant of migration is the higher expected wages offered in the Urban Sector along with better employment opportunities as depicted by the theoretical framework in Harris-Todaro Model. A Case Study of Uttarakhand State is also presented to depict the migrants’ situation in the view of this model. Further, the paper examines

the trends and patterns of migration for the Census years 2001 and 2011 and analyses the reasons behind inter-state migration from an econometric viewpoint. The results correspond to the Harris-Todaro Model depicting an inverse relation between In-Migration Rate and Unemployment rate in contrast to positive relation with poverty rate and State's Net Domestic Product. While there has been a significant jump in all the streams of migration except urban to rural areas from 2001 to 2011, a close analysis of the data reflects that employment is not a major factor responsible for migration in developing countries. Rather sociological factors also influence a substantial flow of migration.

### **Mukhtar Ahmed(2021)**

#### **“Study on Human Migration and its Impact on Education and Society in India”**

This study was especially focused on types of human migration and particularly how migration affects education, positive as well as negative impacts are mentioned.

### **Kyaing kyaing thet**

#### **“Pull and push factors of migration. A case study in the urban area of monywa township, Myanmar”.**

This report emphasizes on what are the factors responsible for migration are explained and what is the socio-economic profile of migrants.

### **Naveen p. singh and Nisha Varghese. (2010)**

#### **“Labour migration and its implications on rural economy of Indo-Gangetic plains of India”.**

In this, they explain that the massive out migration of male population from rural to urban areas has increased at an enormous rate in recent years in India, especially in the Indo Gangetic regions of the country. This indicates that most

of the developed regions of India are utilizing labour force of backward areas. A detailed investigation on labour migration was done in the states of Uttar Pradesh and Bihar to study the determinants of migration and their impacts on farm economies. This study reveals that the farm efficiency of the migrant households has not been significantly affected by this phenomenon, indicating that rural-urban migration is in the form of surplus unskilled labour. The determinants analysis shows that number of members in the family and their education status had a positive impact on migration.

### **Kallio Enni (2016)**

#### **“Human Migration: Implications and Opportunities for Conservation”**

This study was especially focused on the theory of human migration, impact on rural and urban areas, the impact of human migration on conservation and implications for biodiversity.

### **Rajesh Singh and Florentin Smarandache. (July 2022)**

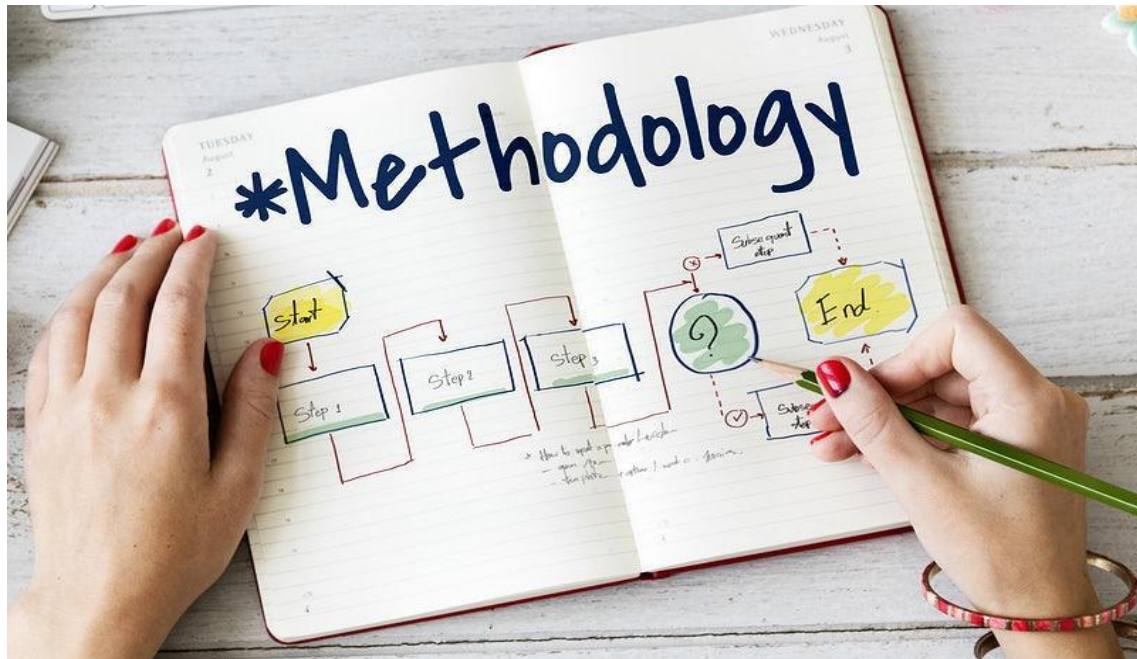
#### **“Structural Dynamics of Various Causes of Migration in Jaipur”.**

This paper gives a brief about the composition of people migrating for various reasons and displays a meticulous trend. The dominance of people moving due to marriages is increasing, whereas Jaipur city is losing its lustre in attracting people for education and business. short duration migration from Jaipur district to urban areas has gone down to a very low level. The flow of migrants from rural areas to Jaipur outpaced the migrants from urban areas and its composition from various in terms long and short distance migration has substantially changed over two consecutive decades. Movements of males and females differed on many criteria, as males move faster than females for employment & education and females move faster than males for marriages and moving along family was found evident in short, medium and long-distance migration. A gender gap of 45 people migration for different reasons was observed and a gender specific trend was seen.



# Chapter No. 3

## Research Methodology & Plan Work



3.1 Objectives

3.2 Data Collection

3.3 Research Methodology

3.4 Software used

3.5 Opinion about the sample collected.

### 3.1 Objectives.

#### *The objective of the study:*

1. To examine the proportions for males and females who are migrating 1<sup>st</sup> time.
2. To check whether there is an association between gender and reasons for migration.
3. To study whether there is an association between age and reasons for migration.
4. To inspect whether there is an association between gender and duration.
5. To suggest the suitable model for the prediction of future migration.



### **3.2 Data collection:**

For the collection of data, I have personally interviewed the residents of Amravati from different localities to control the response bias and to increase the reliability of the data. I collected primary data by directly interviewing the respondents from different localities of Amravati. A structured pattern of questionnaire as well as Google forms were used during the sample survey. The questionnaire I have constructed consist of 16 questions. The questionnaire consists of different variable like gender, age, qualification, annual income etc. All the questions in the questionnaire have the appropriate options as per the requirements of my project. Data was collected in such a manner that it included respondents of age above 16 years.

I have collected a sample of 400 respondents by the method of Snowball Sampling. Statistical data provides the raw material for investigation and gives support for decision-making in that particular manner.

#### **3.2.1 Questionnaire:**

Dear respondents, I Ms. Vaibhavi Sanjay Kalmegh student of Statistics department of Government Vidarbha Institute of Science and Humanities (Autonomous) pursuing Master's degree (M.Sc. II Sem IV). As a part of my curriculum, I require to do a statistical project on "A study on human migration process from 'rural and urban' to urban area in Amravati". This questionnaire is intended to collect data about the migration all responses given by you will be strictly kept confidential and will be used for academic purpose only.

1) Name / E-mail ID: \_\_\_\_\_

2) Gender

- ☐ Male
- ☐ Female



3) Age

- ☐ 16 – 25
- ☐ 26 – 35
- ☐ 36 – 45
- ☐ 46 – 55
- ☐ 56 above

4) Are you currently attending school/college?

- ☐ Yes
- ☐ No

If yes then attempt Q5 otherwise attempt Q6

5) Which level of schooling are you currently attending?

- ☐ Primary
- ☐ Secondary
- ☐ Diploma
- ☐ Graduate
- ☐ Post graduate
- ☐ Doctorate

6) Do you have any occupation?

- ☐ Yes
- ☐ No

If yes attempt Q7 and Q8 otherwise attempt Q9

7) What is your occupation?

- ☐ Business
- ☐ Government service
- ☐ Private service
- ☐ Others (daily wages / labor/part time/housewife)

8) What is your total annual income?

- ☐ 50,000
- ☐ 1 lakh
- ☐ 2 lakhs
- ☐ 4 lakhs
- ☐ 5 lakh and above

9) What is your current marital status

- ☐ Single
- ☐ Married
- ☐ Divorcee
- ☐ Widow

10) Family structure

- ☐ Single
- ☐ Joint family
- ☐ Nuclear family

11) What is the reason for your migration?

- ☐ Employment
- ☐ Marriage
- ☐ Education
- ☐ Family reunification
- ☐ Others - \_\_\_\_\_

12) Are you migrating 1<sup>st</sup> time?

- ☐ Yes
- ☐ No

13) From where did you come?

- ☐ Urban – Urban
- ☐ Rural – Urban

14) At what age did you leave your place of residence?

- ☐ Below 15
- ☐ 16 – 25
- ☐ 26 – 35
- ☐ 36 - 45
- ☐ 46 – 55
- ☐ 56 and above

15) How long you have been living away from your place of residence?

- ☐ 6 months (short term migration)
- ☐ 1 year and more (long term migration)

16) Are there any chances for future migration?

- ☐ Yes
- ☐ No
- ☐ Maybe

Thank you for sparing your precious time to fill this questionnaire.

Have a good day and best luck for your future.

### **3.3 Research methodology**

As our research is mainly focused on the rural and urban to urban migration, which are the subtypes of internal migration, hence the data was collected from urban areas of Amaravati district. The sample will not be a random one.

Completely different is the approach used by Snowball Sampling, which relies on the network structure of a target population. It is assumed that all persons of the target population are linked through social contacts and, therefore, every member of the target population can be reached through referral chains of other members of the target group. Selection through referral chains entails selecting first a number of individuals of the target population, and subsequently selecting additional members of the target population who are identified through references made by the initial respondents. At every subsequent stage, every respondent is asked to refer to another person of the target population. Door to door data collection was done, whereas Google forms were also used for the locations where it is difficult to reach. A specially structured interview was developed for the people and used for primary data collection. Hence, the data consists of information from 400 respondents.

#### **Research diagram:**

Sampling design: as stated earlier, data is collected from urban areas of Amaravati district. This data is collected by using a proper questionnaire from the different age groups of respondents. Information from all age groups (from 16 to 56 and above) was collected. The sample size is 400 respondents.

### **3.4 Software used.**

1. Microsoft Excel.
2. R-software.
3. SPSS (Statistical Package for Social Sciences).
4. Python.

The data thus collected has been tabulated and analysed. For data entry and data compilation I used MS-EXCEL. Also, I have used MS-EXCEL for analysis and diagrammatic and graphical representation.

I also used R programming and SPSS for other analytical purposes.

### **3.5 Opinion about the sample collected.**

While collecting the sample, I faced many problems, such as respondent's behaviour playing an important role in data collection, because respondents refuse to answer the questions asked about their income. Respondents do not reveal their actual income and profession. Most of the time, people have ignored the questionnaires sent through google forms via WhatsApp, Instagram, Facebook. Many people hesitate to fill their e-mail-Id on Google form.

Many respondents were not ready to cooperate as I was unknown to them and continuously kept asking what kind of information they had. What is the intention/ motive of mine behind this?

I interviewed respondents with various professions, such as businessmen, government servants, employees, labour, housewife etc. When the respondents did not mention their income, I considered their income on the basis of their profession so as to avoid non-response errors.

# Chapter 4

## *Data analysis*



4.1. Sample profile: frequency distribution.

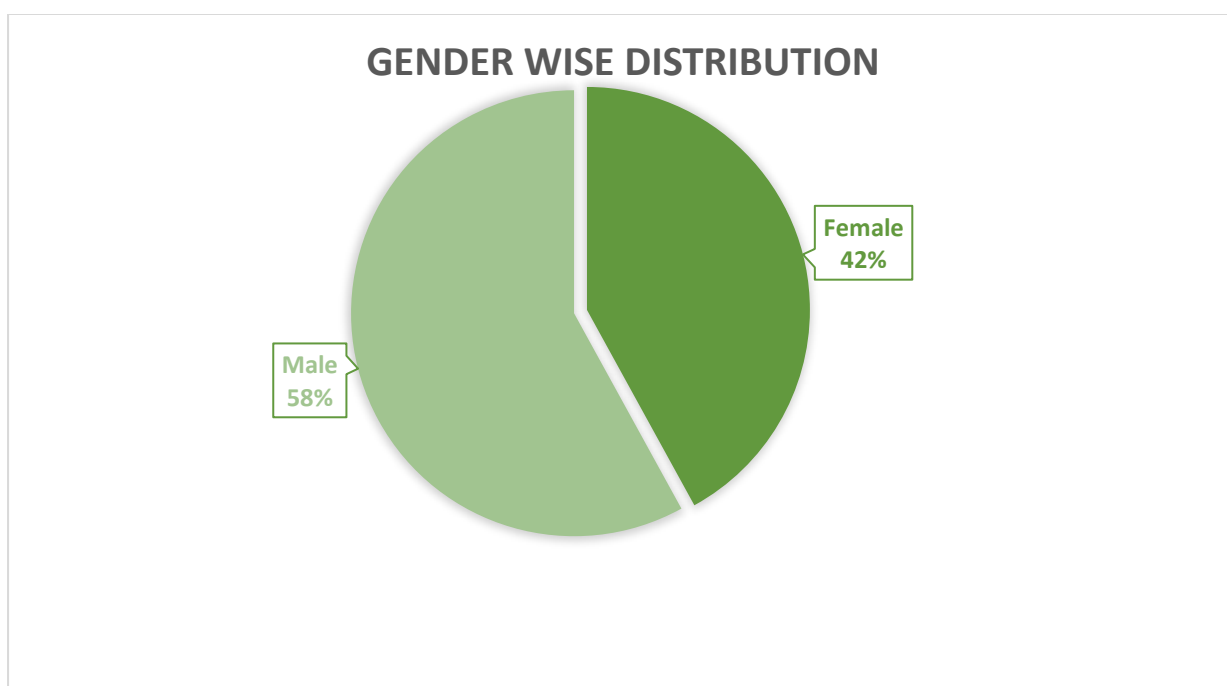
4.2. In-depth analysis.

#### 4.1. Sample profile: frequency distribution.

##### Q.2 Gender wise distribution

Gender	No of response	Percentage
Female	168	42%
Male	232	58%
Total	400	100%

Graphical representation of Gender wise Distribution.

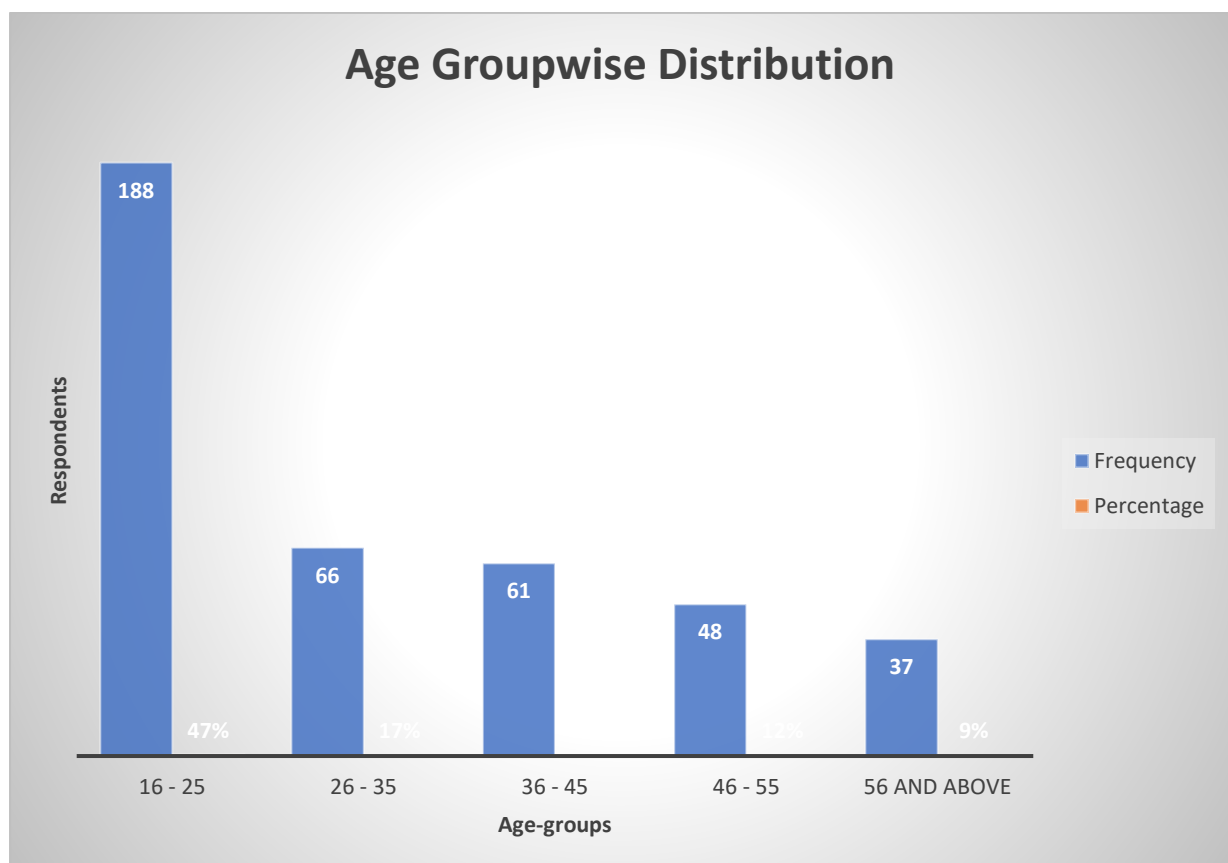


**Interpretation:** From the above figure, we conclude that 58% are males, and 42% are females.

### Q.3 Age groupwise distribution

Age	No of response	Percentage
16 - 25	188	47%
26 - 35	66	17%
36 - 45	61	15%
46 - 55	48	12%
56 and above	37	9%
Total	400	100%

Graphical representation of Age Group-wise Distribution.



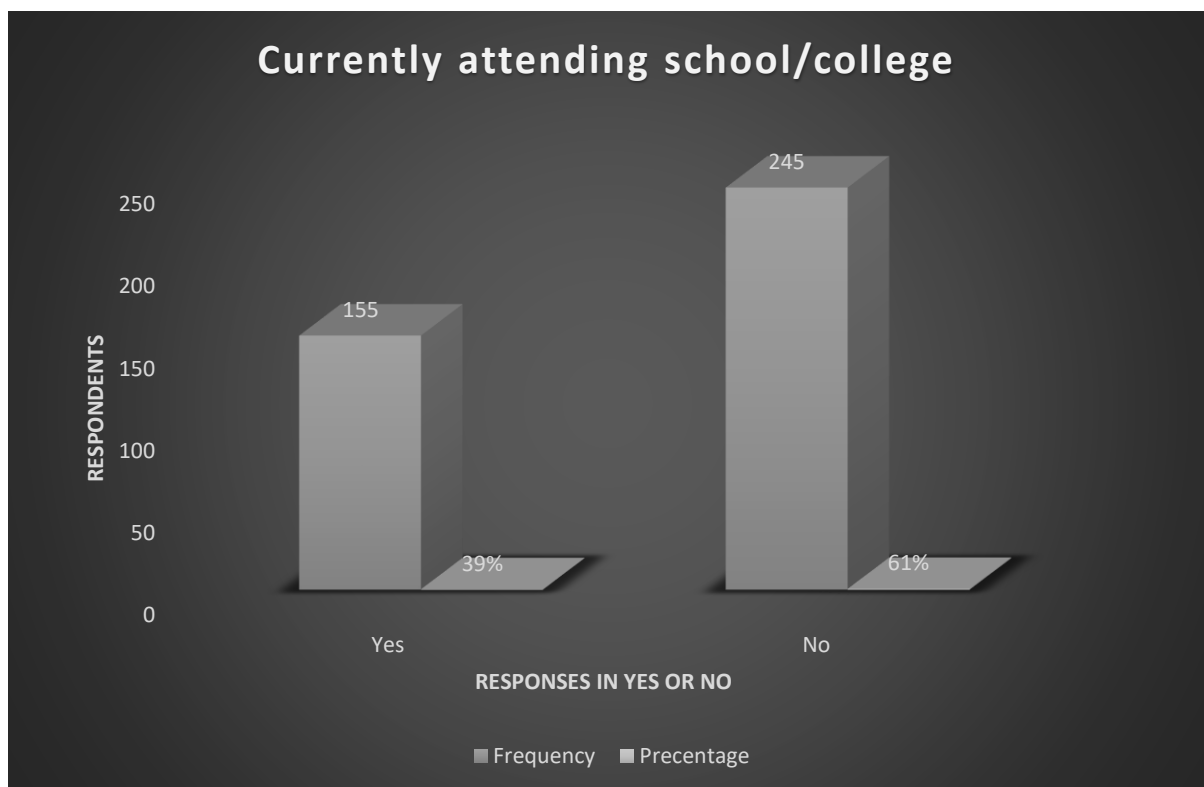
**Interpretation:** From the above figure, we conclude that there are 47% of respondents from the age group 16-25, 17% of respondents are from the age group 26-35, 15% of respondents are from the age group 36-45, 12% of respondents are from the age group of 46-55 and 9% of respondents are from the age group of 56 and above.



#### Q.4 Are you currently attending school/college?

Currently attending	No of response	Percentage
Yes	155	39%
No	245	61%
Total	400	100%

Graphical representation of how many is attending school/college



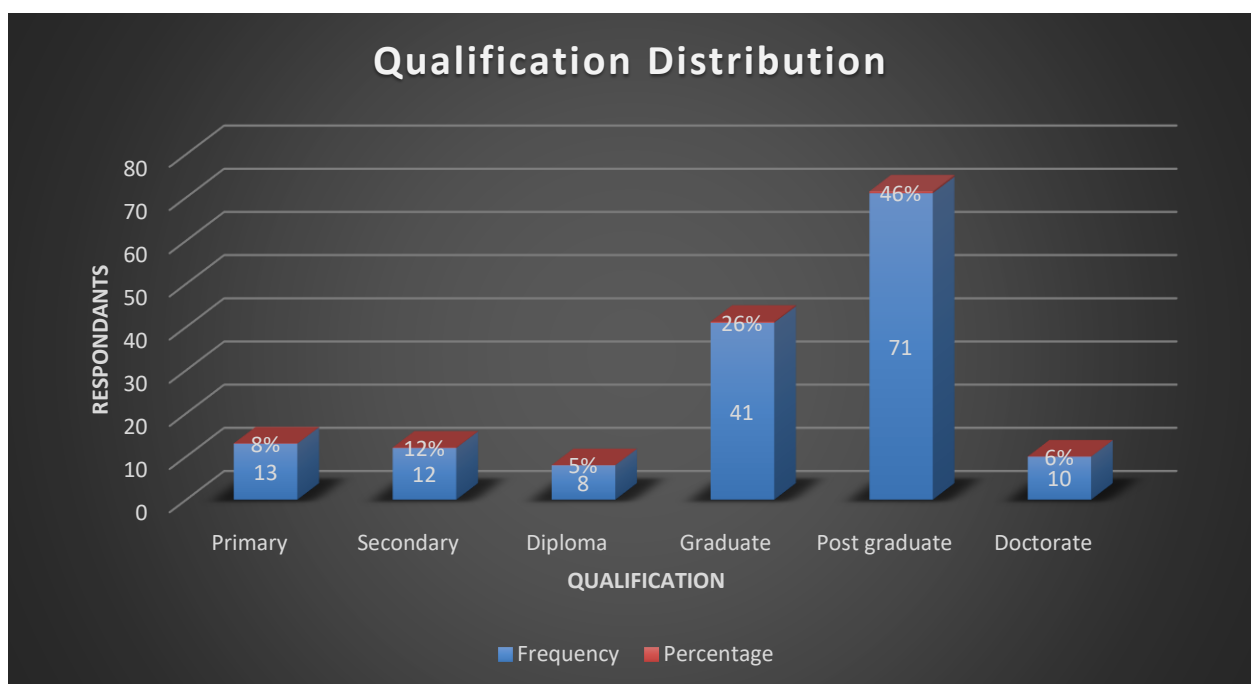
**Interpretation:** From the above figure, we conclude that 39% of respondents are attending school/college and 61% of respondents are not attending school/college.

## If Yes

Q.5 which level of schooling are you currently attending?

Qualification	No of response	Percentage
Primary	13	8%
Secondary	12	12%
Diploma	8	5%
Graduate	41	26%
Post graduate	71	46%
Doctorate	10	6%
Total	155	100%

Graphical representation of Qualification distribution

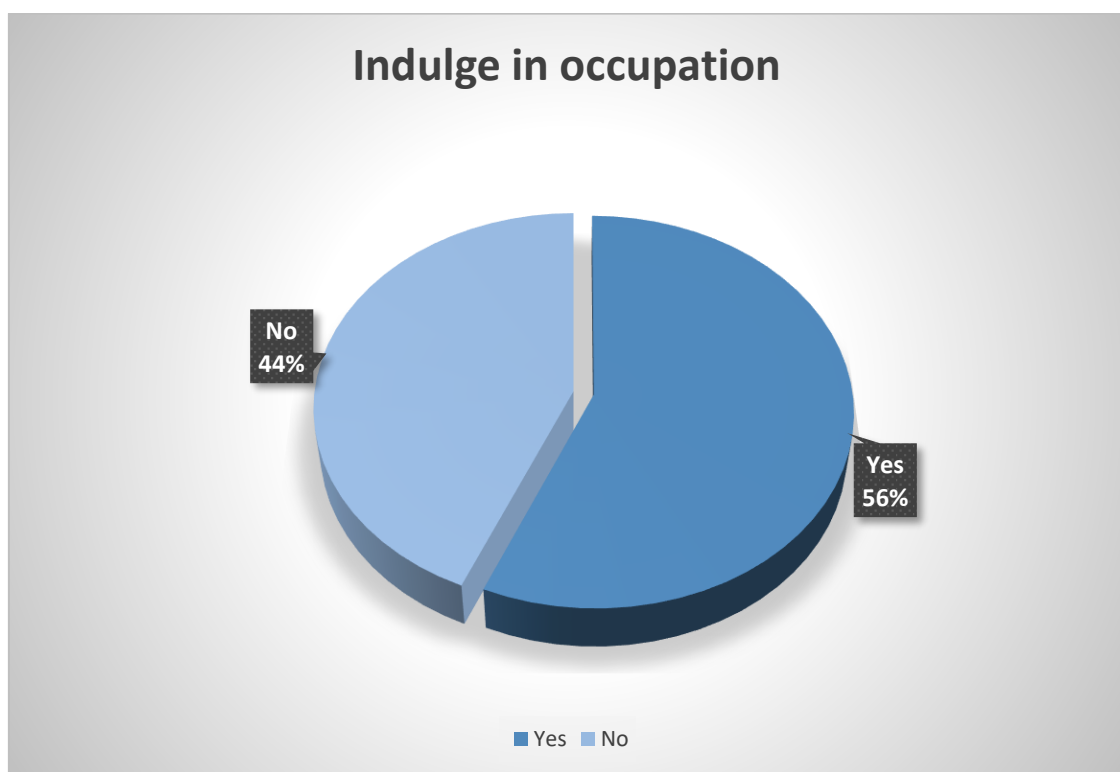


**Interpretation:** From the above figure, we conclude that there 8% of students pursuing primary education, 12% of students are pursuing secondary education, 5% of students are pursuing diploma, 26% of students are pursuing graduate degree, 46% of students are pursuing post graduate degree and 6% of students are pursuing doctorate.

Q.6 do you have any occupation?

Have any occupation	No of response	Percentage
Yes	226	57%
No	174	44%
Total	400	100%

Graphical representation of respondents having occupation or not



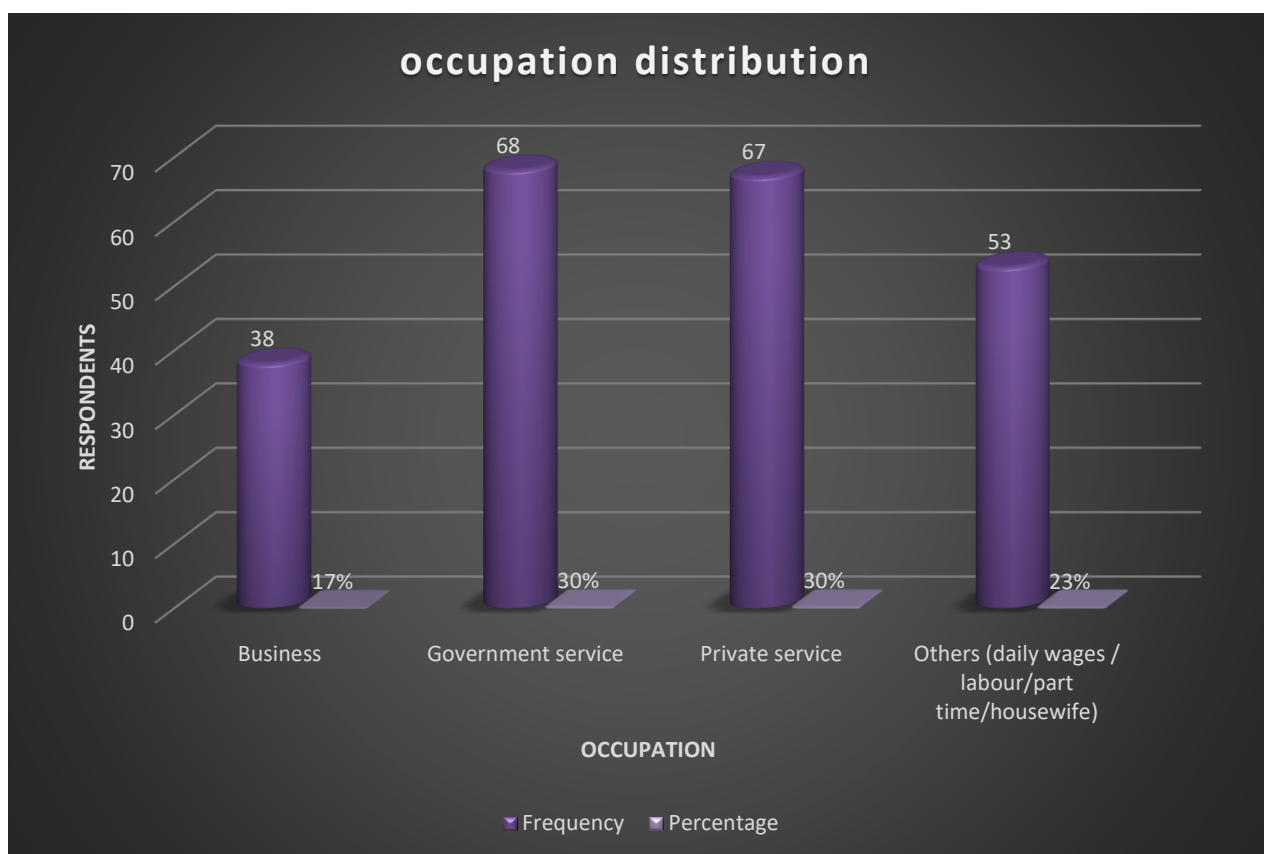
**Interpretation:** From the above figure, we conclude that 56% of people have indulged themselves in occupation and 44% of people are not having any occupation.

**If yes**

**Q.7 What is your occupation?**

Occupation	No of response	Percentage
Business	38	17%
Government service	68	30%
Private service	67	30%
Others (daily wages / labour/part time/housewife)	53	23%
Total	226	100%

Graphical representation of occupation distribution

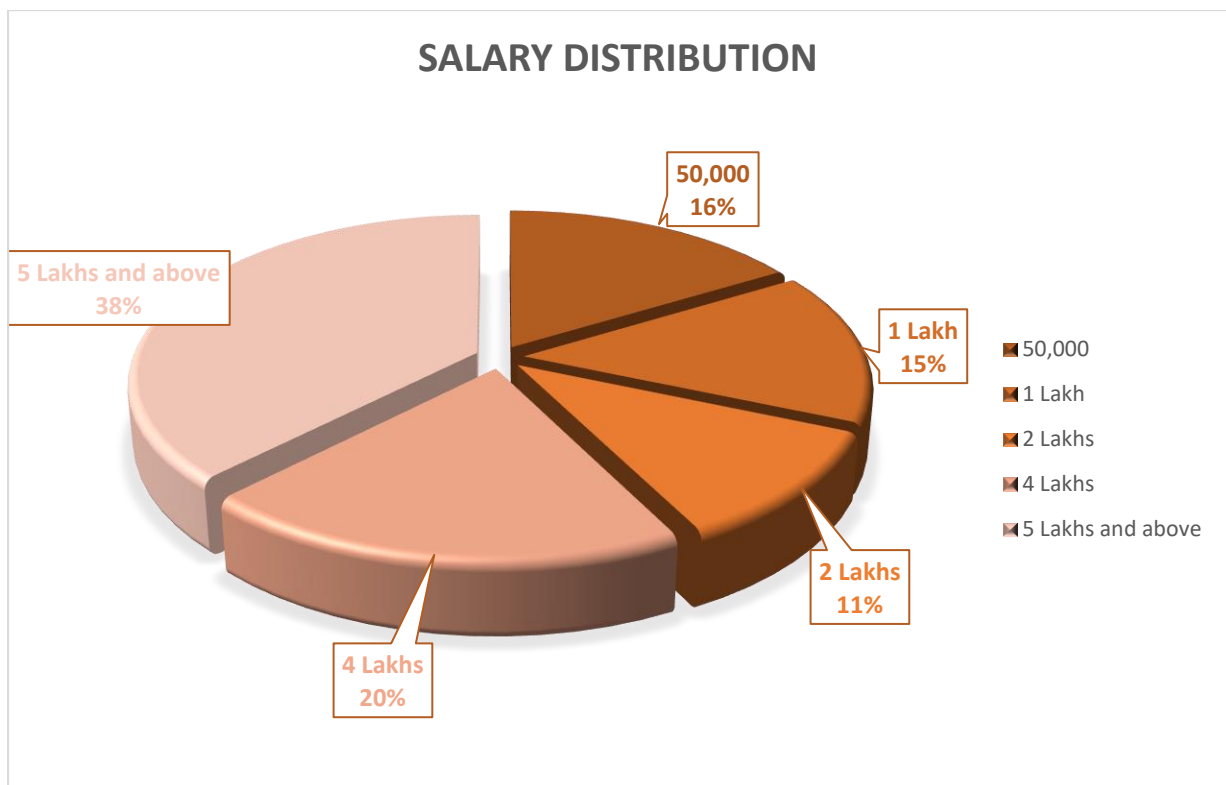


**Interpretation:** From the above figure, we conclude that 17% of people are in the business sector, 30% of people are in government sector, 30% of people are in the private sector and 23% of people have selected others, which includes daily wages, labour, part time jobs and housewife).

Q.8 What is your total annual income?

Salary Distribution	No of response	Percentage
50,000	37	16%
1 Lakh	34	15%
2 Lakhs	25	11%
4 Lakhs	45	20%
5 Lakhs and above	85	38%
Total	226	100%

Graphical representation of salary distribution



**Interpretation:** From the above figure, we conclude that there are 16% of respondents have 50,000 salary, 15% of respondents have 1 lakh salary, 11% of respondents have 2 lakh salary, 20% of respondents have a 4 lakh salary and 38% of respondents have a salary in the range of 5 lakhs and above.

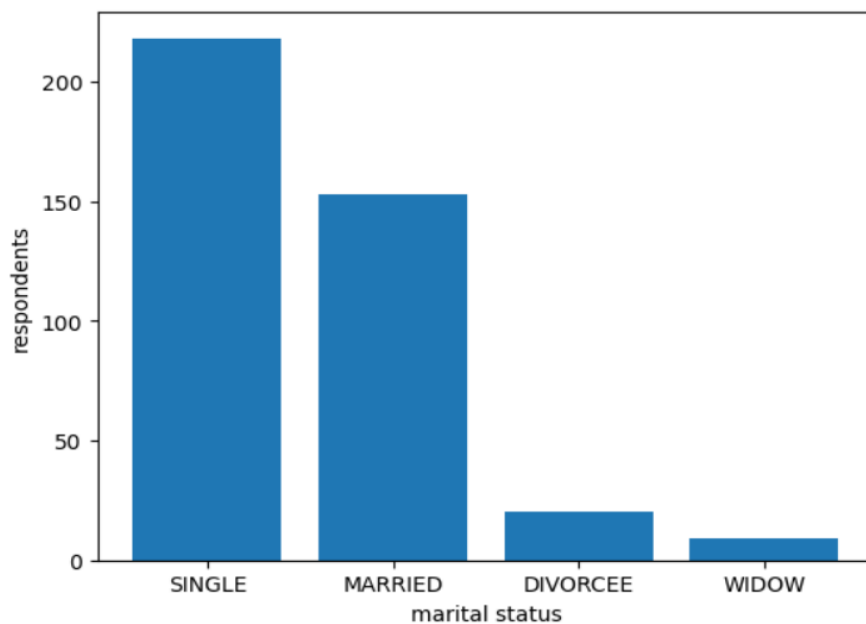
### Q.9 What is your current marital status

Marital status	No of response	Percentage
Single	218	55%
Married	153	38%
Divorcee	20	5%
Widow	9	2%
Total	400	100%

#code for graphical representation in Python

#### Graphical representation of Marital status

```
In [5]: import matplotlib.pyplot as pyplot
labels=('SINGLE','MARRIED','DIVORCEE','WIDOW')
index=(1,2,3,4)
sizes=(218,153,20,9)
pyplot.bar(index,sizes, tick_label=labels)
pyplot.ylabel('respondents')
pyplot.xlabel('marital status')
pyplot.show()
```

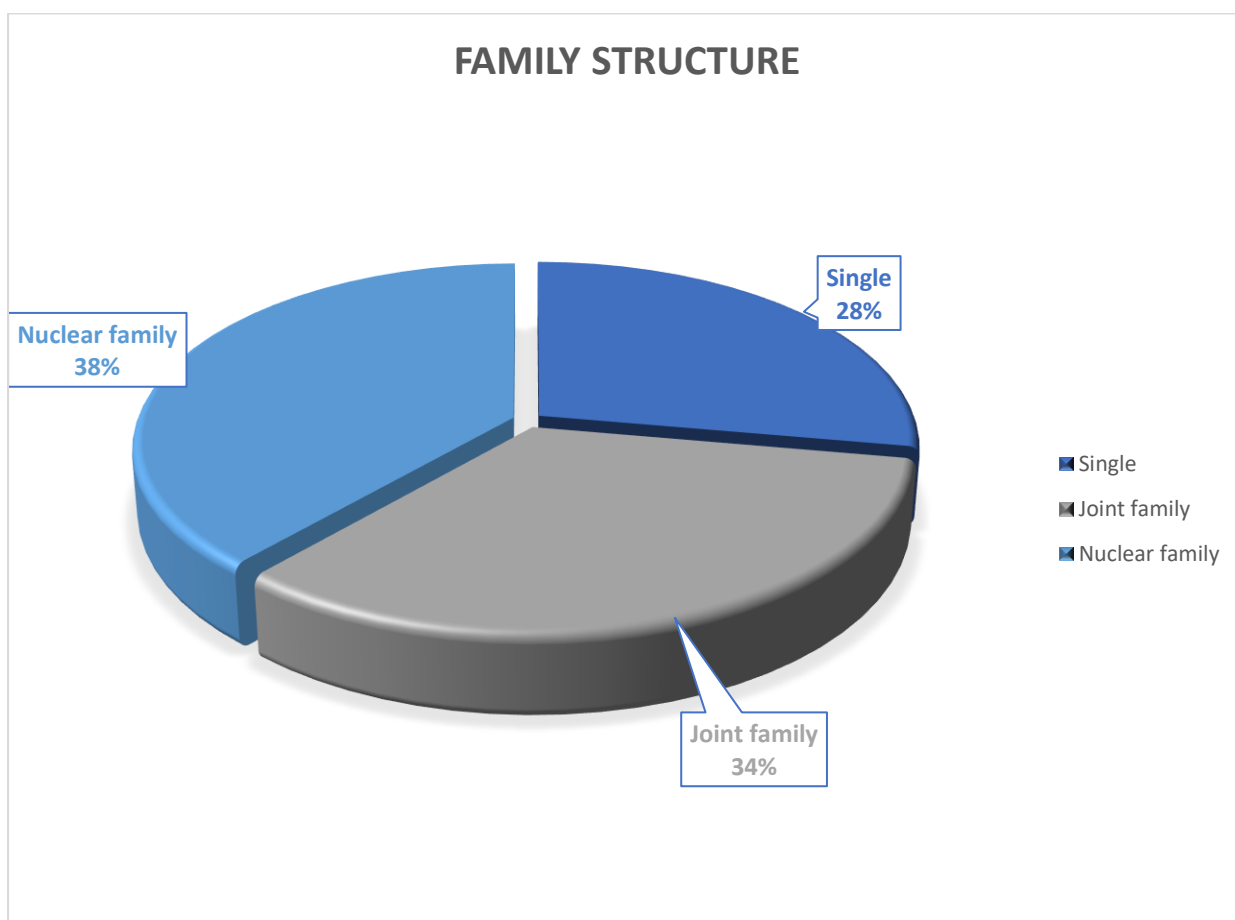


**Interpretation:** From the above figure, we conclude that 55% of respondents are single, 38% of respondents are married, 5% of respondents are divorcee and 2% of respondents are widows.

### Q.10 Family structure

Family Structure	No of response	Percentage
Single	111	28%
Joint family	135	34%
Nuclear family	154	39%
Total	400	100%

Graphical representation of Family structure

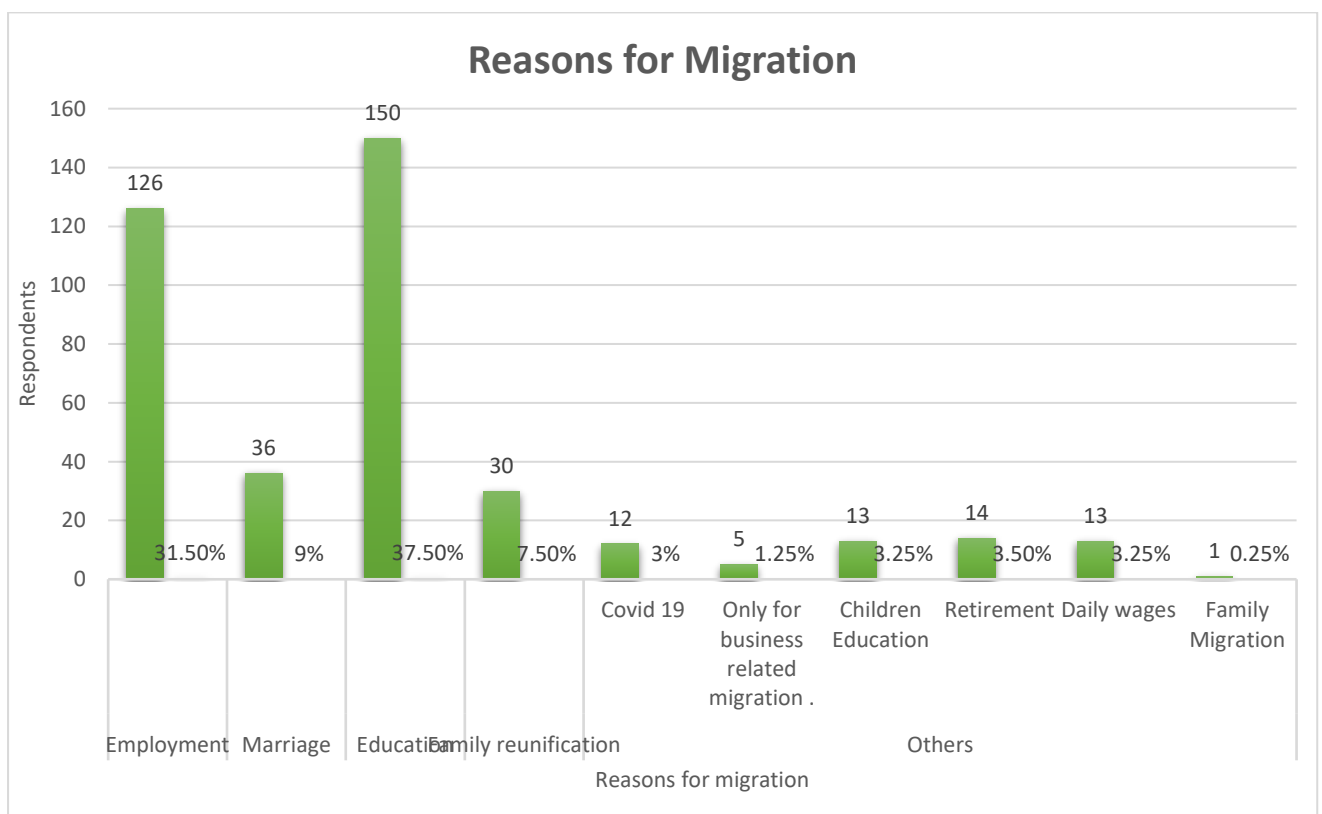


**Interpretation:** From the above figure, we conclude that 28% of respondents are living single, 34% of respondents are living in a joint family and 38% of respondents are living in a nuclear family.

### Q.11 what is the reason for your migration?

Reasons	Other's sub-reasons	No of response	Percentage
Employment		126	31.50%
Marriage		36	9%
Education		150	37.50%
Family reunification		30	7.50%
Others	Covid 19	12	3%
	Only for business related migration	5	1.25%
	Children Education	13	3.25%
	Retirement	14	3.50%
	Daily wages	13	3.25%
	Family Migration	1	0.25%
Total		400	100%

Graphical representation of Reasons for Migration



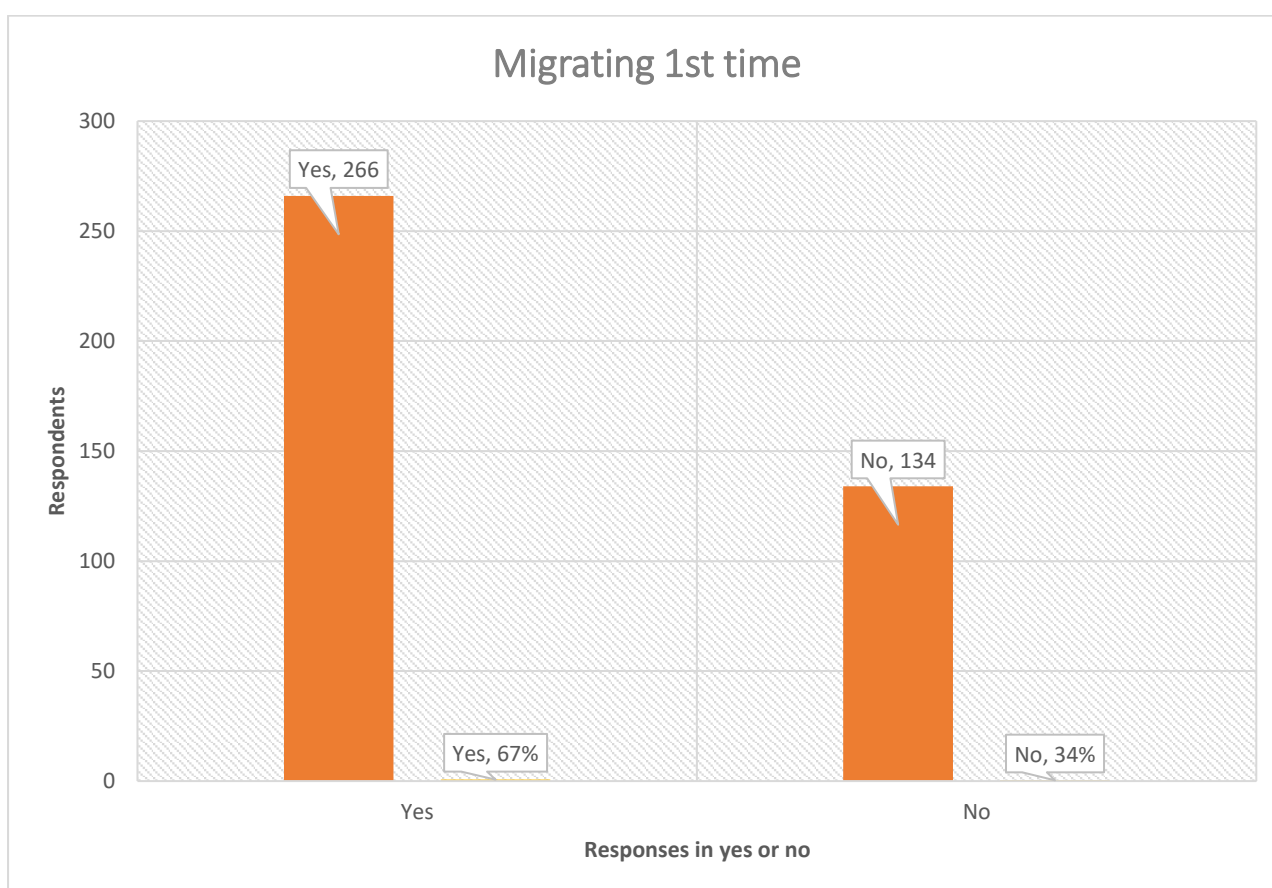
**Interpretation:** From the above figure, we conclude that 31.50% of respondents have migrated for employment, 9% of respondents have migrated because of marriage, 37.50% of respondents have migrated for education purposes, 7.50% of respondents have migrated because of family reunification, while 14.5% of respondents have others reasons where others include the reasons like Covid19(3%) , Only for business related migration(1.25%), Children’s education(3.25%), Retirement(3.50%), Daily wages(3.25%), Family Migration(0.25%).



Q. 12 Are you migrating 1<sup>st</sup> time

Migrating 1st time	No of response	Percentage
Yes	266	67%
No	134	34%
Total	400	100%

Graphical representation of migration whether it is 1<sup>st</sup> time or not.

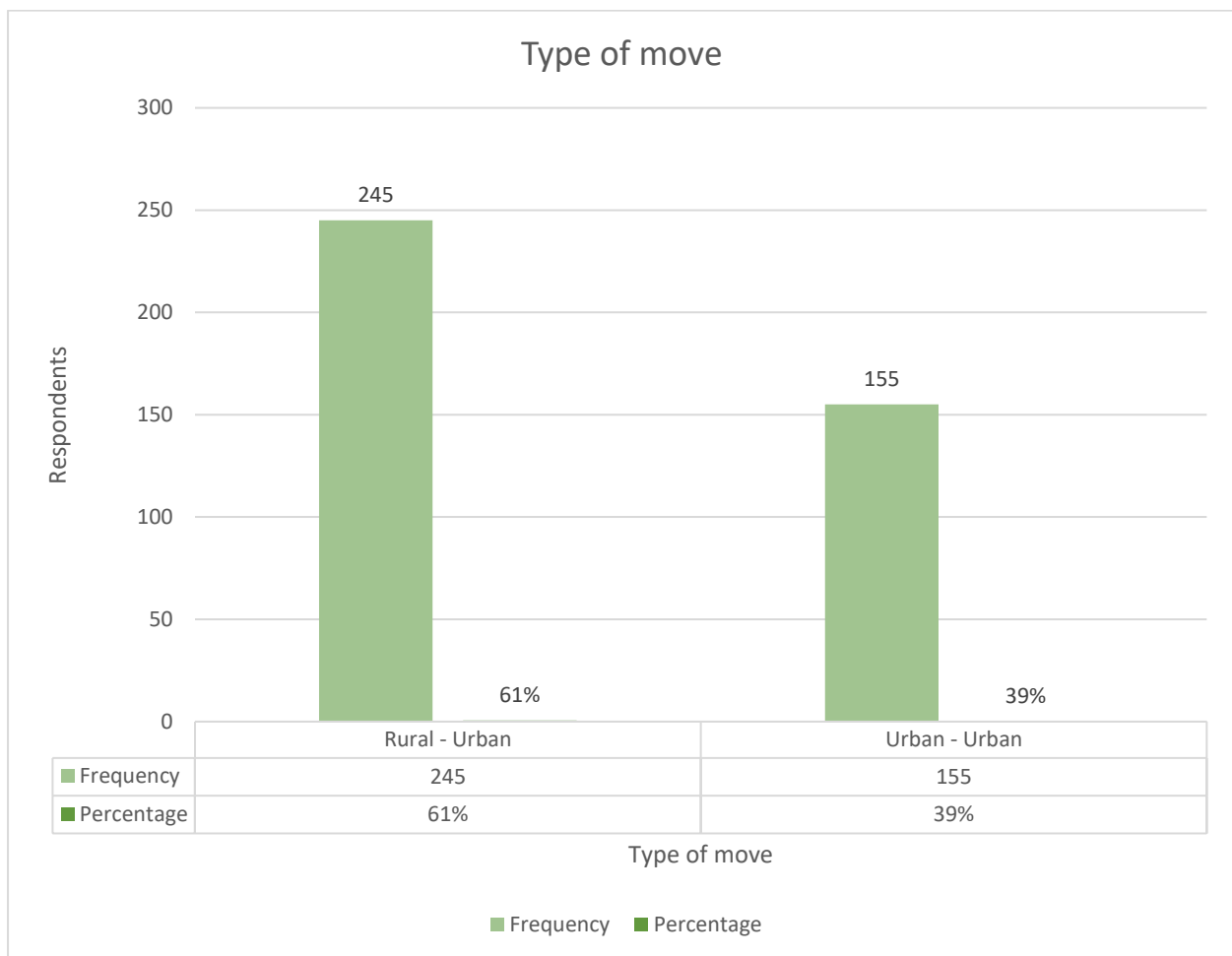


**Interpretation:** From the above figure, we conclude that 67% of people are migrating 1<sup>st</sup> time while 34% of people are not migrating 1<sup>st</sup> time.

### Q.13 From where did you come?

Type of move	No of response	Percentage
Rural - Urban	245	61%
Urban - Urban	155	39%
Total	400	100%

#### Graphical representation of Type of move among the migrants



**Interpretation:** From the above figure, we conclude that about 61% of migration is done from Rural – Urban and 39% of migration is done from Urban – Urban.

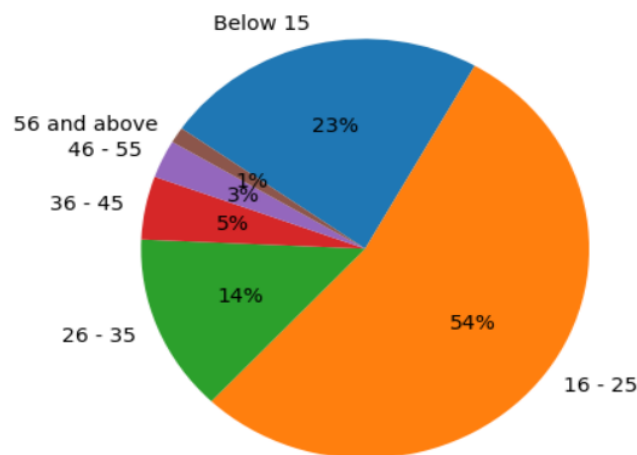
Q.14 At which age did you leave your place of residence?

Age	No of response	Percentage
Below 15	96	23%
16 - 25	211	54%
26 - 35	56	14%
36 - 45	20	5%
46 - 55	12	3%
56 and above	5	1%
Total	400	100%

#code for graphical representation in Python

Graphical representation of Age in which migrant has 1<sup>st</sup> leave their place of residence

```
In [2]: import matplotlib.pyplot as pyplot
labels = ('Below 15', '16 - 25', '26 - 35', '36 - 45', '46 - 55', '56 and above')
sizes = (96, 211, 56, 20, 12, 5)
pyplot.pie(sizes,
            labels=labels,
            autopct='%1.1f%%',
            counterclock=False,
            startangle=145)
pyplot.show()
```

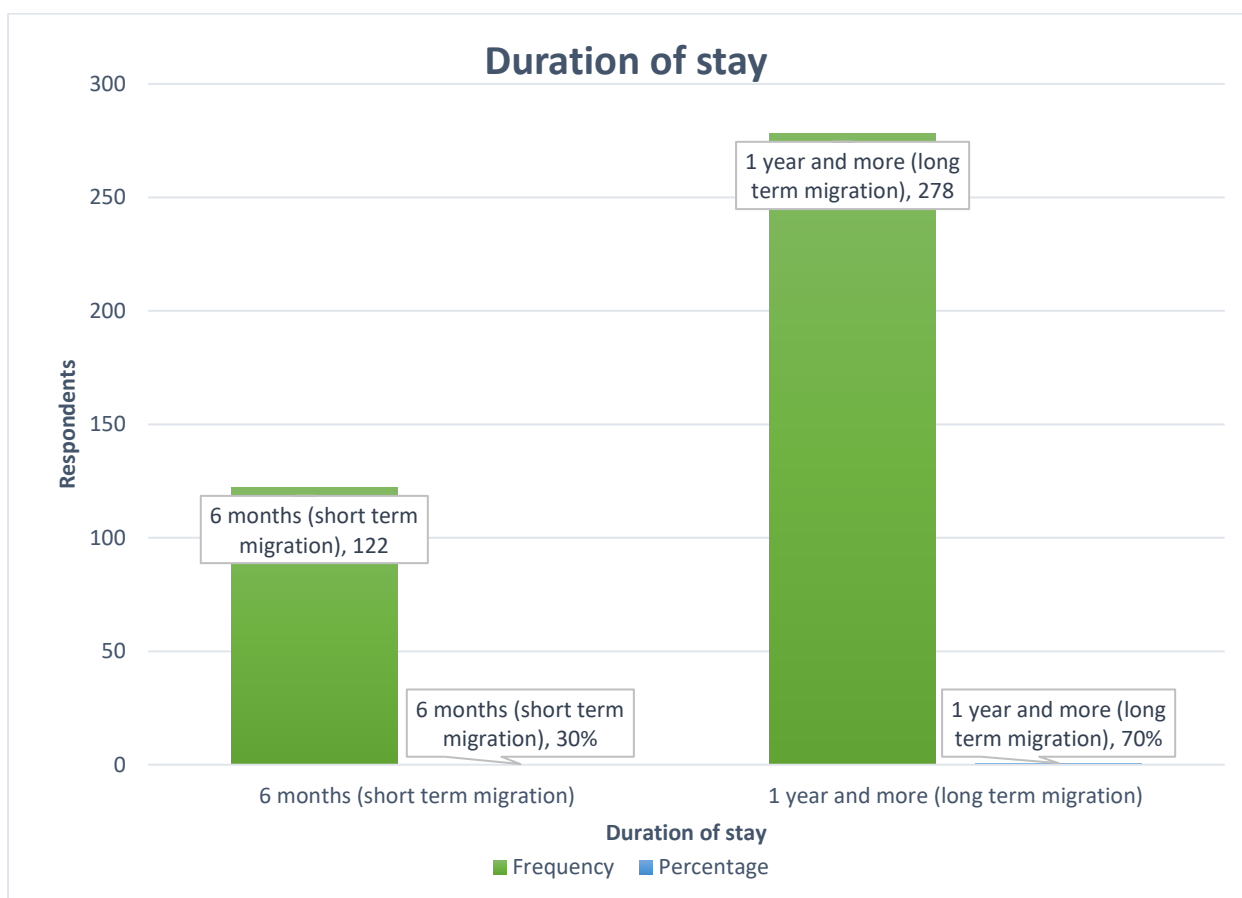


**Interpretation:** From the above figure, we conclude that 24% of the respondents had migrated at the age of less than 15, 53% of the respondents had migrated at the age of 16–25, 14% of the respondents had migrated at the age of 26–35, 5% of the respondents had migrated at the age of 36–45, 3% of the respondents had migrated at the age of 46–55 and 1% respondents had migrated at the age of 56 and above.

Q.15 How long you have been living away from your place of residence?

Duration of stay	No of response	Percentage
6 months (short term migration)	122	30%
1 year and more (long term migration)	278	70%
Total	400	100%

Graphical representation of Duration of stay of migrants

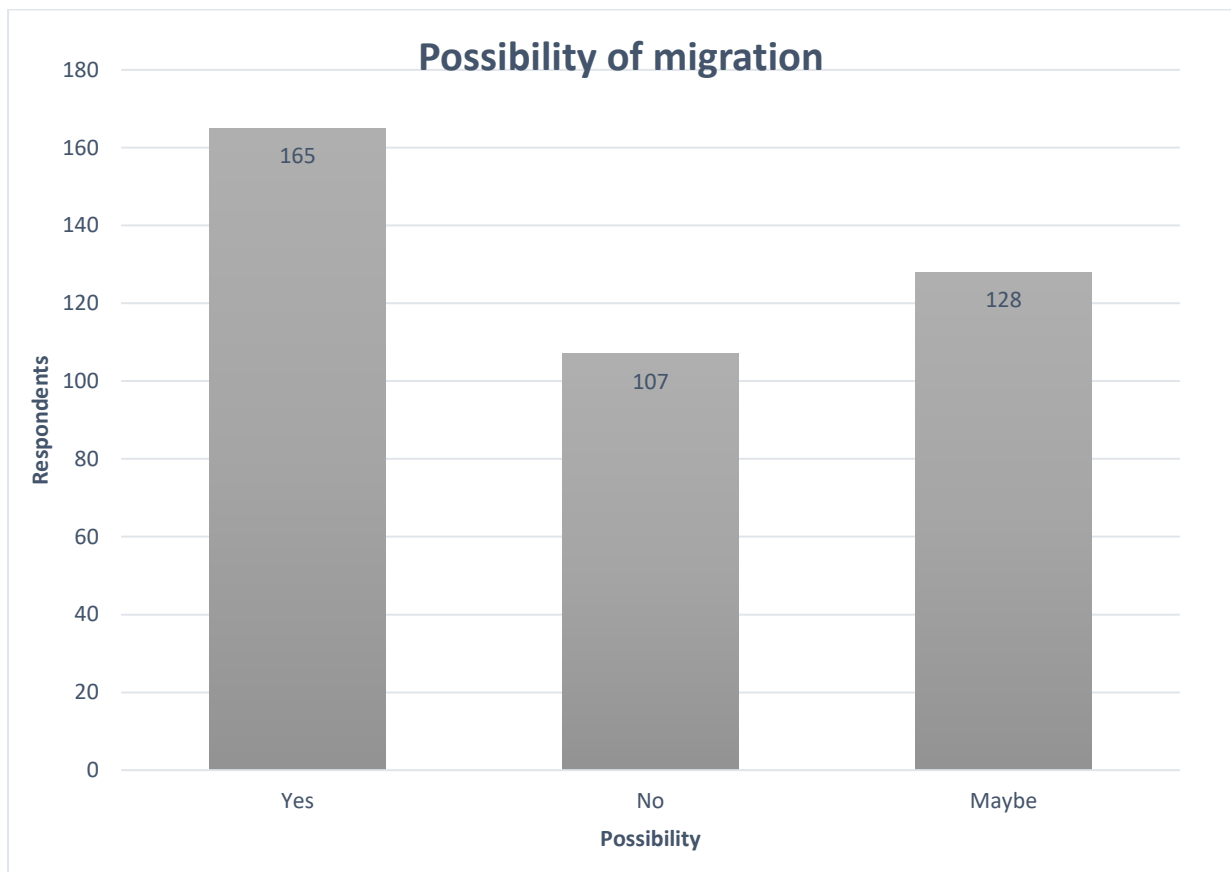


**Interpretation:** From the above figure, we conclude that 30% of the migrants have migrated for 6 months, Hence their migration is short term migration, whereas 70% of migrants have migrated for 1 year and more, hence their migration is long term migration.

Q.16 Is there any chances for future migration?

Chance of migration	No of response	Percentage
Yes	165	41%
No	107	27%
Maybe	128	32%
Total	400	100%

Graphical representation of possibility of migration in future.



**Interpretation:** From the above figure, we conclude that 41% of the observed respondents have the possibility of future migration, 27% of the observed respondents didn't have the possibility of future migration, whereas 32% of the observed respondents are not sure about the possibility of future migration.

1. To examine the proportions for male and female who are migrating 1<sup>st</sup> time.

Test for difference of proportion for large sample test:

Hypothesis:

$H_0$  : Proportions for males and females who are migrating 1<sup>st</sup> time are same i.e.,  
 $P_1 = P_2$

$H_1$  : Proportions for males and females who are migrating 1<sup>st</sup> time are not same  
i.e.  $P_1 \neq P_2$

No of females =  $n_1 = 168$ .

No of males =  $n_2 = 232$ .

No of females who are migrating 1<sup>st</sup> time  $x_1 = 126$

No of females who are migrating 1<sup>st</sup> time  $x_2 = 140$

Now,

$$P_1 = \frac{x_1}{n_1} \\ = \frac{126}{168} = 0.75$$

$$P_2 = \frac{x_2}{n_2} \\ = \frac{140}{232} = 0.6034$$

Since sample are large, the test statistics under the null hypothesis

$$H_0 \text{ is, } |Z| = \left| \frac{P_1 - P_2}{\sqrt{\hat{P} * \hat{Q} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \right|$$

$$\hat{P} = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2} \\ = \frac{168 * 0.75 + 232 * 0.6034}{168 + 232} = 0.6649$$

$$\hat{Q} = (1 - \hat{P})$$

$$= (1 - 0.6649) = 0.3351$$

$$|Z| = \left| \frac{P_1 + P_2}{\sqrt{\hat{P} * \hat{Q} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \right| = \left| \frac{0.75 + 0.6034}{\sqrt{0.6649 * 0.3351 \left( \frac{1}{168} + \frac{1}{232} \right)}} \right| = |3.0745|$$

$$|Z| = 3.0745$$

Level of significance  $\alpha = 0.05$  or 5%

Critical value of  $Z_\alpha = 1.96$  at 5% level of significance for two tailed test

Decision:  $|Z| > Z_\alpha$

i.e.,  $3.0745 > 1.96$ ,  $H_0$  is rejected.

### # code for z proportion test in Python

```
In [1]: # importing all the libraries
from statsmodels.stats.proportion import proportions_ztest
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
data = pd.read_csv("C:/Users/vaibhavi/Desktop/zproportion.csv")
data
```

Out[1]:

	Gender	Are you migrating 1st time?
0	Female	Yes
1	Male	No
2	Female	No
3	Male	No
4	Female	No
...	...	...
395	Male	Yes
396	Male	Yes
397	Male	Yes
398	Male	No
399	Male	Yes

400 rows × 2 columns

```
In [2]: # checking what is in the 'Are you migrating 1st time?' column
# and the data type
data['Are you migrating 1st time?'].unique()
```

```
Out[2]: array(['Yes', 'No'], dtype=object)
```

```
In [3]: # cheching whether there is null value and data type
data.isnull().sum()
```

```
Out[3]: Gender                0
Are you migrating 1st time?  0
dtype: int64
```

```
In [4]: # Replacing Yes by 0, and No by 1
data['Are you migrating 1st time?'].map({'Yes':0, 'No':1})
```

```
Out[4]: 0      0
1      1
2      1
3      1
4      1
..
395    0
396    0
397    0
398    1
399    0
Name: Are you migrating 1st time?, Length: 400, dtype: int64
```

```
In [5]: #Create the Data Frame
data['Are you migrating 1st time?']=data['Are you migrating 1st time?'].map({'Yes':0, 'No':1})
data.head()
```

```
Out[5]:
```

	Gender	Are you migrating 1st time?
0	Female	0
1	Male	1
2	Female	1
3	Male	1
4	Female	1

```
In [8]: # code for Z proportion
number_of_successes = np.array([126, 140])
total_sample_sizes = np.array([168, 232])
(test_stat, p_value) = proportions_ztest(number_of_successes, total_sample_sizes, alternative='two-sided')
print("the computed z test statistic is:", test_stat)
print("the p-value is:", p_value)
```

```
the computed z test statistic is: 3.0649708160798514
the p-value is: 0.002176912380379675
```

Since,  $p\text{-value} > 0.05$

i.e.,  $0.0021 > 0.05$ ,  $H_0$  is rejected at a 5% level of significance.

We accept the  $H_1$  at 5 % level of significance i.e., Proportions for males and females who are migrating 1<sup>st</sup> time are not same i.e.,  $P_1 \neq P_2$

**Result:** As p-value is 0.0021, which is less than 0.05 we may conclude that Proportions for males and females who are migrating 1<sup>st</sup> time are not the same.



2. To check whether there is an association between gender and reasons for migration.

### Chi-square test for independence using python

In learning, it is common to encounter a scenario where the target variable is categorical and the predictors can be either continuous or categorical. When both the target variable and predictors are categorical, a Chi-square test can be employed to gauge the strength of their relationship.

The Chi-square test finds the probability of a Null hypothesis( $H_0$ ).

the p-value came higher than 0.05. Hence,  $H_0$  will be accepted. This means the variables are not associated with each other. This means that if two variables are associated, then the p-value will come very close to zero.

Hypothesis:

$H_0$  = There is no association between gender and reasons for migration.

$H_1$  = There is an association between gender and reasons for migration.

### #Code for chi-square in Python

```
In [2]: # importing the libraries
import pandas as pd
import numpy as np
import scipy.stats as ss
from scipy.stats import chi2_contingency
import seaborn as sns
import matplotlib.pyplot as plt
data=pd.read_excel("C:/Users/vaibhavi/Desktop/CORRELATION.xlsx")
data
```

```
Out[2]:
```

	Gender	Reason
0	Female	Employment
1	Male	Education
2	Female	Family reunification
3	Male	Education
4	Female	Education
...	...	...
395	Male	Others
396	Male	Family reunification
397	Male	Family reunification
398	Male	Family reunification
399	Male	Family reunification

400 rows × 2 columns

```
In [3]: # checking whether there is null value in the data and the data type
data.isnull().sum()
```

```
Out[3]: Gender      0
Reason      0
dtype: int64
```

```
In [10]: # cross tabulation
cross_tab=pd.crosstab(index=data['Reason'],columns=data['Gender'])
cross_tab
```

```
Out[10]:
```

	Gender	Female	Male
Reason			
Education		81	69
Employment		42	84
Family reunification		7	23
Marriage		33	3
Others		5	53

```
In [13]: # Performing Chi-sq test
chi_sq_result=chi2_contingency(cross_tab,)
p, x = chi_sq_result[1], "reject" if chi_sq_result[1] < 0.05 else "accept"
```

```
In [19]: # P-Value is the Probability of H0 being True
# If P-Value>0.05 then only we Accept the assumption(H0)
print(f"The p_value is {chi_sq_result[1]} and hence we {x} the null Hypothesis with {chi_sq_result[2]} degrees of freedom")

The p_value is 1.7197553800899911e-16 and hence we reject the null Hypothesis with 4 degrees of freedom
```

Since  $p\text{-value} < 0.05$ ,

i.e.,  $1.719\text{e-}16 < 0.05$ ,  $H_0$  is rejected at a 5% level of significance.

Hence, we accept the  $H_1$  i.e., There is an association between gender and reasons for migration.

**Result:** As  $p\text{-value} = 1.719\text{e-}16$  which is almost zero and is smaller than 0.05, we may conclude that there is an association between gender and reasons for migration.

### 3. To study whether there is an association between age and reasons for migration.

Analysis of variance (ANOVA) is a statistical technique used to check if the means of two or more groups are significantly different from each other.

ANOVA checks the impact of one or more factors by comparing the means of different samples.

Hypothesis:

$H_{0A}$  = There is no significant difference in migration between the (population) means of the age groups (factor 1).

$H_{1A}$  = There is a significant difference in migration between the (population) means of the age groups (factor 1).

$H_{0B}$  = There is no significant difference in migration between the (population) means of the reasons for migration (factor 2).

$H_{1B}$  = There is a significant difference in migration between the (population) means of the reasons for migration (factor 2).

		Reasons for migration				
		Education	Employment	Family reunification	Marriage	Others
Age	Below 15	44	16	12	4	20
	16 – 25	103	69	11	17	11
	26 - 35	2	27	3	15	9
	36 – 45	1	9	2	0	8
	46 -55	0	3	1	0	8
	56 and above	0	2	1	0	2

# Anova: Two-Factor Without Replication

<i>SUMMARY</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
Below 15	5	96	19.2	227.2
16 - 25	5	211	42.2	1749.2
26 - 35	5	56	11.2	105.2
36 - 45	5	20	4	17.5
46 - 55	5	12	2.4	11.3
56 and above	5	5	1	1
Education	6	150	25	1760
Employment	6	126	21	638.8
Family reunification	6	30	5	26
Marriage	6	36	6	62.8
Others	6	58	9.666667	34.66667

## ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Rows	6155.067	5	1231.013	3.813391	0.013743	2.71089
Columns	1989.333	4	497.3333	1.540622	0.228698	2.866081
Error	6456.267	20	322.8133			
Total	14600.67	29				

Since for the rows  $p\text{-value} < 0.05$

i.e.,  $0.0137 < 0.05$

(or  $F=3.8133 > F_{crit}=2.7108$ ),  $H_{0A}$  is rejected at a 5% level of significance with (5, 20) degree of freedom.

Hence, we accept the  $H_{1A}$  i.e., there is a significant difference in migration between the (population) means of the age groups (factor 1)

**Result 1:** As p-value 0.0137 which is smaller than 0.05, we may conclude that there is a significant difference in migration between the (population) means of the age groups (factor 1).

Also, for the columns p-value  $> 0.05$

i.e.,  $0.2286 > 0.05$

(or  $F=1.5406 > F_{crit}=2.8660$ ),  $H_{0B}$  is accepted at a 5% level of significance with (4, 20) degree of freedom.

**Result 2:** As p-value 0.2286 which is greater than 0.05, we may conclude that there is no significant difference in migration between the (population) means of the reasons for migration (factor 2).

4. To determine whether there is an association between gender and duration.

### What is the chi-square test of independence?

A chi-square ( $X^2$ ) test of independence is a type of Pearson's chi-square test. Pearson's chi-square tests are nonparametric tests for categorical variables. They're used to determine whether your data is significantly different from what you expected.

You can use a chi-square test of independence, also known as a chi-square test of association, to determine whether two categorical variables are related. If two variables are related, the probability of one variable having a certain value is dependent on the value of the other variable.

The chi-square test of independence is an inferential statistical test, meaning that it allows you to draw conclusions about a **population** based on a **sample**. Specifically, it allows you to conclude whether two variables are related in the population

In statistics, Yates's correction for continuity (or Yates's chi-squared test) is used in certain situations when testing for independence in a contingency table.

By ruling out independence of the two variables, the chi-square can be used to assess whether two variables are, in fact, dependent or not. More generally, we say that one variable is "not correlated with" or "independent of" the other if an increase in one variable is not associated with an increase in the another.

Hypothesis:

$H_0$  = There is no association between the duration of stay and Gender.

$H_1$  = There is an association between the duration of stay and Gender.

Gender	Duration of stay		
	Short term	Long term	Total
Female	a 122	b 46	a+b 168
Male	c 156	d 76	c+d 232
Total	a+c 278	b+d 122	400

Observed frequency	Expected frequency	(O – E) <sup>2</sup> / E
122	116.76	0.2351
46	51.24	0.5378
156	161.24	0.1702
76	70.76	0.3880

$$\chi^2_{\text{cal}} = \sum (O - E)^2 / E$$

$$= 0.2351 + 0.5378 + 0.1702 + 0.3880$$

$$\chi^2_{\text{cal}} = 1.3311$$

$$\text{Degree of freedom} = (2-1)(2-1) = 1$$

$$\alpha = 0.05 \text{ (5\%)}$$

$$\chi^2_{\text{tab}} \text{ (critical value)} = 3.84$$

$$\chi^2_{\text{cal}} < \chi^2_{\text{tab}}$$

Hence, we accept  $H_0$ .

#code for chi-square in R

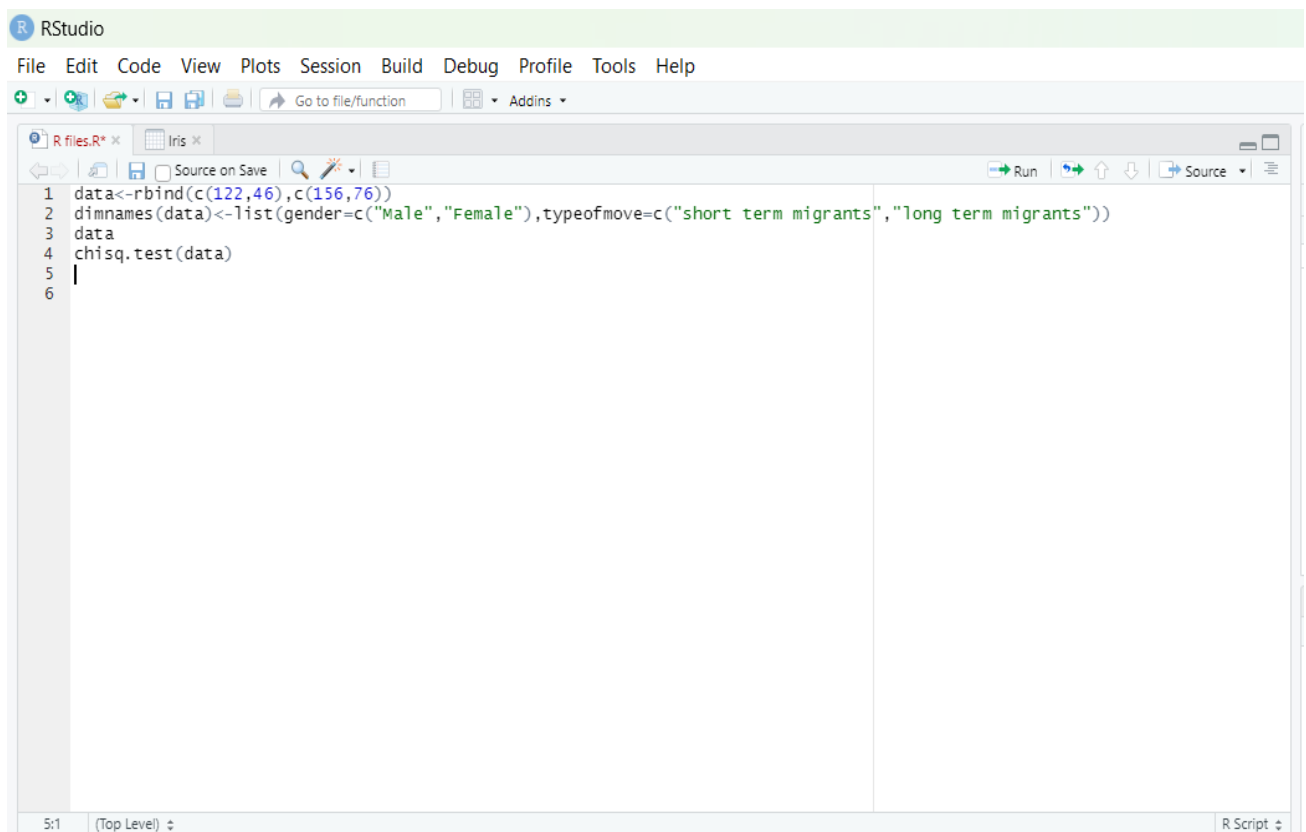
```
data<-rbind(c(122,46),c(156,76))
```

```
dimnames(data)<-list(gender=c("Male","Female"),typeofmove=c("short term  
migrants","long term migrants"))
```

```
data
```

```
chisq.test(data)
```

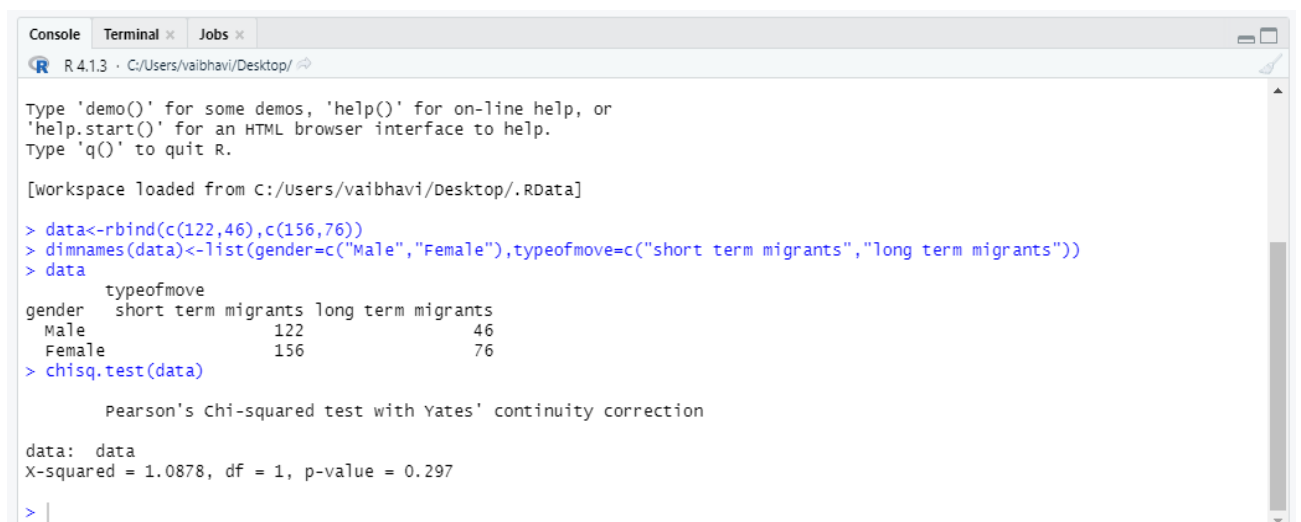
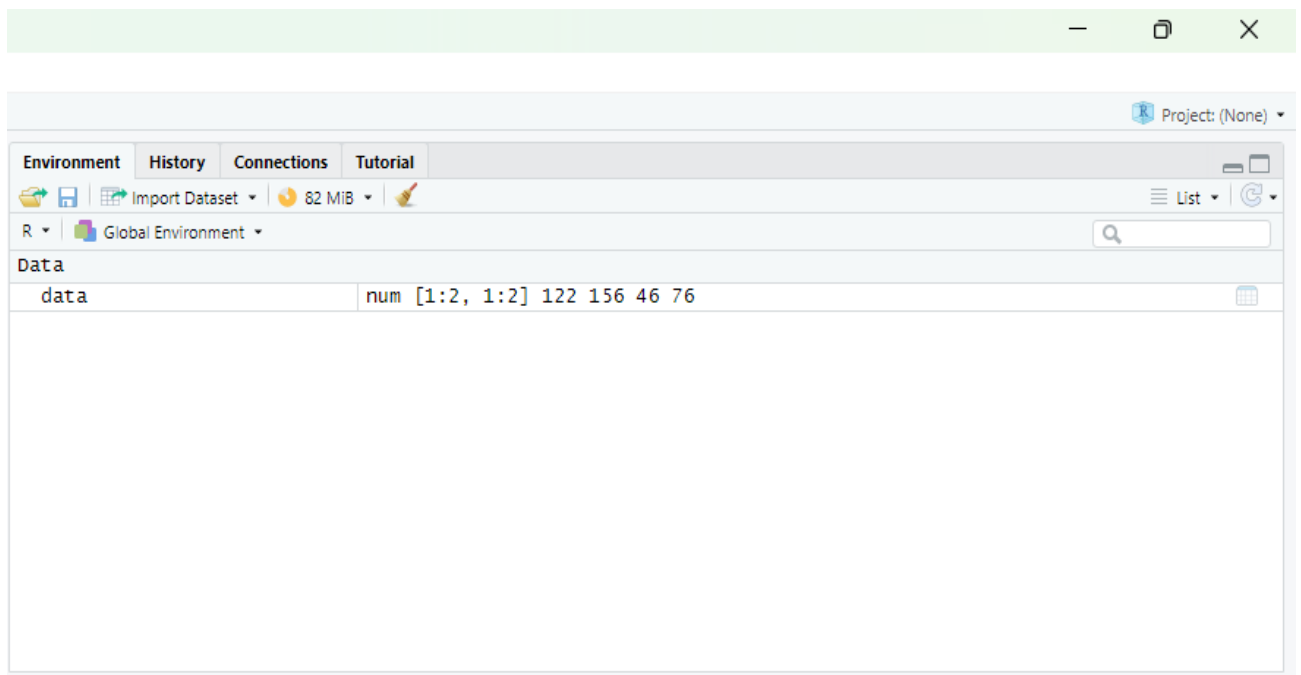
Output is as follows:

A screenshot of the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for saving, running, and other functions. The main editor window shows an R script with the following code:

```
1 data<-rbind(c(122,46),c(156,76))
2 dimnames(data)<-list(gender=c("Male","Female"),typeofmove=c("short term migrants","long term migrants"))
3 data
4 chisq.test(data)
5 |
6
```

The status bar at the bottom indicates the current position is 5:1 (Top Level) and the file type is R Script.





Since  $p\text{-value} > 0.05$ ,

i.e.,  $0.297 > 0.05$ ,  $H_0$  is accepted at 5% level of significance with 1 degree of freedom.

**Result:** As  $p\text{-value} = 0.297$  which is greater than 0.05, we may conclude that there is no association between the duration of stay and Gender.

## 5. To suggest the suitable model for the prediction of future migration.

### **Logistic Regression Overview**

Logistic regression is a fundamental classification technique. It belongs to the group of linear classifiers and is somewhat similar to polynomial and linear regression. Logistic regression is fast and relatively uncomplicated, and it's convenient for you to interpret the results. Although it's essentially a method for binary classification, it can also be applied to multiclass problems.

### **Problem Formulation**

When you're implementing the logistic regression of some dependent variable  $y$  on the set of independent variables  $\mathbf{x} = (x_1, \dots, x_r)$ , where  $r$  is the number of predictors (or inputs), you start with the known values of the predictors  $\mathbf{x}_i$  and the corresponding actual response (or output)  $y_i$  for each observation  $i = 1, \dots, n$ .

Your goal is to find the logistic regression function  $p(\mathbf{x})$  such that the predicted responses  $p(\mathbf{x}_i)$  are as close as possible to the actual response  $y_i$  for each observation  $i = 1, \dots, n$ . Remember that the actual response can be only 0 or 1 in binary classification problems!

Once you have the logistic regression function  $p(\mathbf{x})$ , you can use it to predict the outputs for new and unseen inputs, assuming that the underlying mathematical dependence is unchanged.

### **Classification Performance**

Binary classification has four possible types of results:

1. True negatives: correctly predicted negatives (zeros)
2. True positives: correctly predicted positives (ones)
3. False negatives: incorrectly predicted negatives (zeros)
4. False positives: incorrectly predicted positives (ones)

You usually evaluate the performance of your classifier by comparing the actual and predicted outputs and counting the correct and incorrect predictions.

```
In [1]: import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn import metrics
from sklearn import linear_model
import statsmodels.api as sm
from sklearn.metrics import mean_squared_error
import seaborn as sns
import matplotlib.pyplot as plt
data=pd.read_csv("C:/Users/vaibhavi/Desktop/logistic model.csv")
data
```

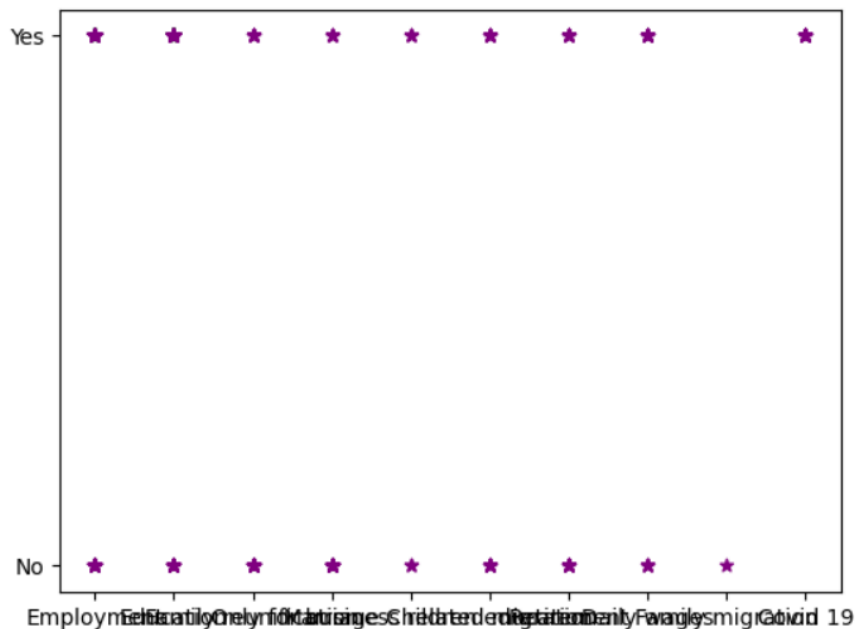
```
Out[1]:
```

	attending college/school	reason	chances
0	No	Employment	No
1	Yes	Education	Yes
2	No	Family reunification	No
3	Yes	Education	Yes
4	Yes	Education	Yes
...	...	...	...
395	No	Daily wages	Yes
396	No	Family reunification	No
397	No	Family reunification	No
398	No	Family reunification	No
399	No	Family reunification	No

400 rows × 3 columns

```
In [2]: plt.scatter(data.reason,data.chances,marker='*',color='purple')
```

```
Out[2]: <matplotlib.collections.PathCollection at 0x26bb0dc1f70>
```



```
In [3]: data['attending college/school'].unique()
```

```
Out[3]: array(['No', 'Yes'], dtype=object)
```

```
In [4]: data['reason'].unique()

Out[4]: array(['Employment', 'Education', 'Family reunification', 'Marriage',
              'Only for business related migration .', 'Children education',
              'Retirement', 'Daily wages', 'Family migration', 'Covid 19'],
              dtype=object)

In [5]: data['reason'].map({'Employment':4,'Education':5,'Family reunification':1,'Marriage':2,'Only for business related migration .':3,
                             'Children education':4,'Retirement':1,'Daily wages':1,'Family migration':1,'Covid 19':1})

Out[5]: 0      4
        1      5
        2      1
        3      5
        4      5
        ..
        395    3
        396    1
        397    1
        398    1
        399    1
        Name: reason, Length: 400, dtype: int64
```

```
In [6]: data['attending college/school'].map({'Yes':0,'No':1})

Out[6]: 0      1
        1      0
        2      1
        3      0
        4      0
        ..
        395    1
        396    1
        397    1
        398    1
        399    1
        Name: attending college/school, Length: 400, dtype: int64
```

```
In [7]: data['reason']=data['reason'].map({'Employment':4,'Education':5,'Family reunification':1,'Marriage':2,'Only for business related
data['attending college/school']=data['attending college/school'].map({'Yes':0,'No':1})
data.head()
```

```
Out[7]:
```

	attending college/school	reason	chances
0	1	4	No
1	0	5	Yes
2	1	1	No
3	0	5	Yes
4	0	5	Yes

```
In [8]: #seperating train and test data
independent_variables = ['attending college/school','reason']
x=data[independent_variables]
y=data['chances']

x_train, x_test, y_train, y_test = train_test_split(x, y,test_size=0.05)
print('train:', len(x_train),'test:', len(x_test))

train: 380 test: 20
```

```
In [9]: # instantiate a logistic regression model, and fit with x and y
model = LogisticRegression(solver="lbfgs")
model = model.fit(x_train, y_train)

#check the accuracy on the training set
model.score(x_train , y_train)
```

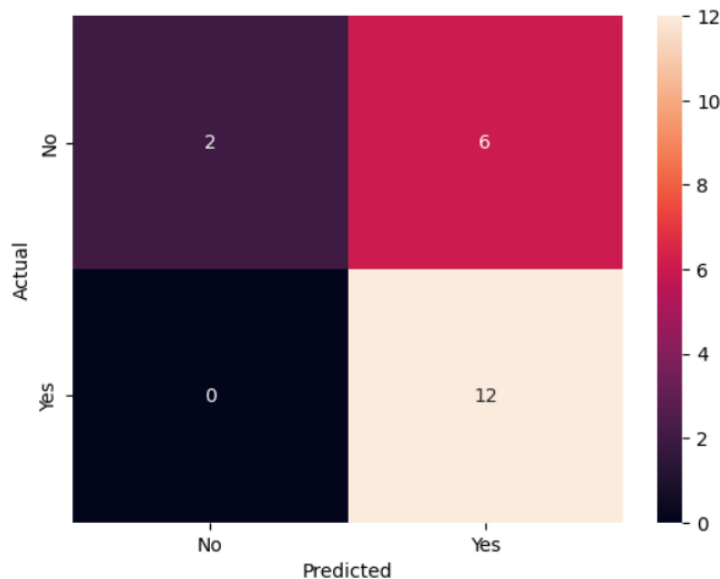
```
Out[9]: 0.7894736842105263
```

```
In [10]: from sklearn import metrics
print("Accuracy ON TEST SET :", metrics.accuracy_score(y_test,model.predict(x_test)))

Accuracy ON TEST SET : 0.7
```

```
In [11]: confusion_matrix = pd.crosstab(y_test, model.predict(x_test), rownames=['Actual'], colnames=['Predicted'])
sns.heatmap(confusion_matrix, annot=True)
```

```
Out[11]: <AxesSubplot:xlabel='Predicted', ylabel='Actual'>
```



```
In [12]: print(model.predict(x_test))
```

```
['Yes' 'No' 'Yes' 'No' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes'
 'Yes' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes' 'Yes']
```

As can be observed from the matrix:

TP = True Positives = 12

TN = True Negatives = 2

FP = False Positives = 6

FN = False Negatives = 0

You can then also get the Accuracy using:  $\text{Accuracy} = (\text{TP} + \text{TN}) / \text{Total} = (12 + 2) / 20 = 0.7$

The accuracy is therefore 70% for the test set

Recall that our original dataset (from step 1) had 400 observations. Since we set the test size to 0.05, then the confusion matrix displayed the results for 20 records (=400\*0.05). These are the 20 test records:

In the actual dataset (from step-1), you'll see that for the test data, we got the correct results 14 out of 20 times:

	A	B	C	D	E
1	attending	reason	chances	predicted	matching
2	No	Employment	No	yes	FALSE
3	Yes	Education	Yes	no	FALSE
4	No	Family reunification	No	yes	FALSE
5	Yes	Education	Yes	no	FALSE
6	Yes	Education	Yes	yes	TRUE
7	Yes	Education	Yes	yes	TRUE
8	Yes	Education	Yes	yes	TRUE
9	No	Employment	Yes	yes	TRUE
10	Yes	Employment	No	yes	FALSE
11	No	Employment	Yes	yes	TRUE
12	No	Employment	Yes	yes	TRUE
13	No	Employment	Yes	yes	TRUE
14	No	Education	No	yes	FALSE
15	Yes	Education	Yes	yes	TRUE
16	Yes	Education	Yes	yes	TRUE
17	Yes	Family reunification	Yes	yes	TRUE
18	No	Employment	Yes	yes	TRUE
19	No	Education	Yes	yes	TRUE
20	Yes	Employment	Yes	yes	TRUE
21	No	Employment	Yes	yes	TRUE

This is matching with the accuracy level of 70%.

To understand logistic regression, let's go over the odds of success.

Odds ( $\theta$ ) = Probability of an event happening / Probability of an event not happening

$$\theta = p / 1 - p$$

The values of odds range from zero to  $\infty$  and the values of probability lies between zero and one.

Consider the equation of a straight line:

$$y = \beta_0 + \beta_1 * x$$

Here,  $\beta_0$  is the y-intercept

$\beta_1$  is the slope of the line

x is the value of the x coordinate

y is the value of the prediction

Now to predict the odds of success, we use the following formula:

$$\log\left(\frac{p(x)}{1-p(x)}\right) = \beta_0 + \beta_1 x$$

Exponentiating both the sides, we have:

$$e^{\ln\left(\frac{p(x)}{1-p(x)}\right)} = e^{\beta_0 + \beta_1 x}$$

$$\left(\frac{p(x)}{1-p(x)}\right) = e^{\beta_0 + \beta_1 x}$$

Let  $Y = e^{\beta_0 + \beta_1 x}$

Then  $p(x) / 1 - p(x) = Y$

$$p(x) = Y(1 - p(x))$$

$$p(x) = Y - Y(p(x))$$

$$p(x) + Y(p(x)) = Y$$

$$p(x)(1+Y) = Y$$

$$p(x) = Y / 1+Y$$

$$p(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

```
In [1]: import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn import metrics
from sklearn import linear_model
import statsmodels.api as sm
from sklearn.metrics import mean_squared_error
import seaborn as sns
import matplotlib.pyplot as plt
data=pd.read_csv("C:/Users/vaibhavi/Desktop/logit.csv")
data
```

```
Out[1]:
```

	college/school	reason	chances
0	No	Employment	No
1	Yes	Education	Yes
2	No	Family reunification	No
3	Yes	Education	Yes
4	Yes	Education	Yes
...	...	...	...
395	No	Daily wages	Yes
396	No	Family reunification	No
397	No	Family reunification	No
398	No	Family reunification	No
399	No	Family reunification	No

400 rows × 3 columns

```
In [4]: data['college/school'].unique()
```

```
Out[4]: array(['No', 'Yes'], dtype=object)
```

```
In [5]: data['reason'].unique()
```

```
Out[5]: array(['Employment', 'Education', 'Family reunification', 'Marriage',
'Only for business related migration .', 'Children education',
'Retirement', 'Daily wages', 'Family migration', 'Covid 19'],
dtype=object)
```

```
In [6]: data['chances'].unique()
```

```
Out[6]: array(['No', 'Yes'], dtype=object)
```

```
In [7]: data['reason'].map({'Employment':0,'Education':1,'Family reunification':2,'Marriage':3,'Only for business related migration .':4,
```

```
Out[7]: 0    0
1    1
2    2
3    1
4    1
..
395  4
396  2
397  2
398  2
399  2
Name: reason, Length: 400, dtype: int64
```



```
In [8]: data['college/school'].map({'Yes':0,'No':1})
```

```
Out[8]: 0      1
        1      0
        2      1
        3      0
        4      0
        ..
       395     1
       396     1
       397     1
       398     1
       399     1
       Name: college/school, Length: 400, dtype: int64
```

```
In [9]: data['chances'].map({'Yes':0,'No':1})
```

```
Out[9]: 0      1
        1      0
        2      1
        3      0
        4      0
        ..
       395     0
       396     1
       397     1
       398     1
       399     1
       Name: chances, Length: 400, dtype: int64
```

```
In [10]: data['reason']=data['reason'].map({'Employment':0,'Education':1,'Family reunification':2,'Marriage':3,'Only for business related':4})
data['college/school']=data['college/school'].map({'Yes':0,'No':1})
data['chances']=data['chances'].map({'Yes':0,'No':1})
data.head()
```

```
Out[10]:
```

	college/school	reason	chances
0	1	0	1
1	0	1	0
2	1	2	1
3	0	1	0
4	0	1	0

```
In [11]: x=data[['reason','college/school']]
y=data['chances']
```

In [12]: x

Out[12]:

	reason	college/school
0	0	1
1	1	0
2	2	1
3	1	0
4	1	0
...	...	...
395	4	1
396	2	1
397	2	1
398	2	1
399	2	1

400 rows × 2 columns

In [13]: y

Out[13]:

```
0      1
1      0
2      1
3      0
4      0
..
395    0
396    1
397    1
398    1
399    1
```

Name: chances, Length: 400, dtype: int64

```
In [14]: import statsmodels.api as sm
x1=sm.add_constant(x)
logit_model=sm.Logit(y,x1)
result=logit_model.fit()
print(result.summary2())
```

Optimization terminated successfully.  
Current function value: 0.580053  
Iterations 5

#### Results: Logit

```
=====
Model:                Logit                Pseudo R-squared: 0.085
Dependent Variable:    chances                AIC:                470.0425
Date:                 2023-04-13 23:57        BIC:                482.0169
No. Observations:     400                    Log-Likelihood:     -232.02
Df Model:              2                      LL-Null:            -253.67
Df Residuals:          397                    LLR p-value:         3.9576e-10
Converged:             1.0000                  Scale:              1.0000
No. Iterations:        5.0000

-----
                Coef.  Std.Err.    z    P>|z|    [0.025  0.975]
-----
const          -1.8288    0.2257  -8.1030  0.0000   -2.2712  -1.3865
reason           0.3681    0.0780   4.7176  0.0000    0.2152   0.5210
college/school   0.8868    0.2457   3.6094  0.0003    0.4053   1.3684
=====
```

$$Y(\text{chances}) = \begin{cases} 1 & \text{if Yes} \\ 0 & \text{if No} \end{cases}$$

1 - Family Reunification

2 - Marriage

$X_1(\text{reasons}) = 3$  - Others

4 - Employment

5 - Education

$$X_2(\text{attending school/college}) = \begin{cases} 1 & \text{if Yes} \\ 0 & \text{if No} \end{cases}$$

The estimated logistic regression model is,

$$Y = \frac{e^{b_0 + b_1x_1 + b_2x_2}}{1 + e^{b_0 + b_1x_1 + b_2x_2}}$$

From the output of the python code the estimated logistic regression is

$$Y = \frac{e^{-1.8288 + 0.3681*x_1 + 0.8868*x_2}}{1 + e^{-1.8288 + 0.3681*x_1 + 0.8868*x_2}}$$

$$P(y=1/X_1=5, X_2=1) = \frac{e^{-1.8288 + 0.3681*5 + 0.8868*1}}{1 + e^{-1.8288 + 0.3681*5 + 0.8868*1}} = 0.7106$$

From the above conditional probability, we can interpret that if a person's reason for migration is education and currently attending school / college then there is a 71.06% chance that the person is likely to migrate next time.

$$P(y=1/X_1=5, X_2=0) = \frac{e^{-1.8288 + 0.3681*5 + 0.8868*0}}{1 + e^{-1.8288 + 0.3681*5 + 0.8868*0}} = 0.5029$$

From the above conditional probability, we can interpret that if a person's reason for migration is education and currently not attending school / college i.e., the person may be indulged in other activity / job, then there are 50.29% chances that the person is likely to migrate next time.

Similarly,

$$P(y=1/X_1=1, X_2=1) = \frac{e^{-1.8288 + 0.3681*1 + 0.8868*1}}{1 + e^{-1.8288 + 0.3681*1 + 0.8868*1}} = 0.1883$$

From the above conditional probability, we can interpret that if a person's reason for migration is family reunification and currently attending school / college then there are 18.83% chances that the person is likely to migrate next time.

$$P(y=1/X_1=1, X_2=0) = \frac{e^{-1.8288 + 0.3681*1 + 0.8868*0}}{1 + e^{-1.8288 + 0.3681*1 + 0.8868*0}} = 0.3603$$

From the above conditional probability, we can interpret that if a person's reason for migration is family reunification and currently not attending school / college i.e., the person may be indulged in other activity / job, then there are 36.03% chances that the person is likely to migrate next time.

## Other findings:

### 1. To look for the reasons which are mainly responsible for rural-urban and urban-urban migration.

reason \* come from Crosstabulation

reason			come from		Total
			Rural - Urban	Urban - Urban	
Children education	Count		13	0	13
	Expected Count		8.0	5.0	13.0
	% of Total		3.3%	0.0%	3.3%
Covid 19	Count		12	0	12
	Expected Count		7.4	4.6	12.0
	% of Total		3.0%	0.0%	3.0%
Daily wages	Count		13	0	13
	Expected Count		8.0	5.0	13.0
	% of Total		3.3%	0.0%	3.3%
Education	Count		85	65	150
	Expected Count		91.9	58.1	150.0
	% of Total		21.3%	16.3%	37.5%
Employment	Count		56	70	126
	Expected Count		77.2	48.8	126.0
	% of Total		14.0%	17.5%	31.5%
Family migration	Count		1	0	1
	Expected Count		.6	.4	1.0
	% of Total		0.3%	0.0%	0.3%
Family reunification	Count		25	5	30
	Expected Count		18.4	11.6	30.0
	% of Total		6.3%	1.3%	7.5%
Marriage	Count		24	12	36
	Expected Count		22.1	14.0	36.0
	% of Total		6.0%	3.0%	9.0%
Only for business related migration.	Count		2	3	5
	Expected Count		3.1	1.9	5.0
	% of Total		0.5%	0.8%	1.3%
Retirement	Count		14	0	14
	Expected Count		8.6	5.4	14.0
	% of Total		3.5%	0.0%	3.5%
Total	Count		245	155	400
	Expected Count		245.0	155.0	400.0
	% of Total		61.3%	38.8%	100.0%

### Interpretation:

- From the above table, we observe that about 3.3% respondents out of total population are migrating from rural-urban for children's education. Whereas the percentage (%) for urban -urban for children's education is 0.

- About 3.0% respondents out of total population are migrating from rural-urban due to covid 19. Whereas the percentage (%) for urban -urban due to covid 19 is 0.
- 3.3% respondents out of total population are migrating from rural-urban for daily wages. Whereas the percentage (%) for urban -urban for daily wages is 0.
- Also, about 21.3% respondents out of total population are migrating from rural-urban for education purpose, whereas the percentage (%) for urban-urban for education purpose is 16.3%
- Also, about 14.0% respondents out of total population are migrating from rural-urban for employment, whereas the percentage (%) for urban-urban for employment is 17.5%
- For family migration, only 0.3% respondents have migrated from rural-urban which is very small count and for urban-urban the percentage (%) is 0.
- For family reunification, about 6.3% respondents have migrated from rural-urban and for urban-urban the percentage for family reunification is 1.3%.
- Moreover, for marriage, 6.0% respondents have migrated from rural-urban and for urban-urban the percentage (%) for marriage is 3.0%.
- About 0.5% of respondents are migrating for the business purpose from rural-urban whereas 0.8% of respondents are migrating from urban-urban for business purpose.
- For retirement, about 3.5% of respondents are migrating from rural-urban for a better life, whereas the percentage (%) for urban-urban for retirement is 0.

## 2. correlation between the present age the age when migrated.

One way to quantify the relationship between two variables is to use the Pearson correlation coefficient, which measures the linear association between two variables.

It always takes on a value between -1 and 1 where:

- -1 indicates a perfectly negative linear correlation
- 0 indicates no linear correlation
- 1 indicates a perfectly positive linear correlation

To determine if a correlation coefficient is statistically significant, we can calculate the corresponding t-score and p-value.

The formula to calculate the t-score of a correlation coefficient (r) is:

$$t = r * \sqrt{n-2} / \sqrt{1-r^2}$$

The p-value is then calculated as the corresponding two-sided p-value for the t-distribution with n-2 degrees of freedom.

### # code for Correlation Test in Python

```
In [1]: import pandas as pd
import numpy as np
import scipy.stats as ss
from scipy.stats import chi2_contingency
import seaborn as sns
import matplotlib.pyplot as plt
data=pd.read_excel("C:/Users/vaibhavi/Desktop/correlation age.xlsx")
data
```

Out[1]:

	Present age	Age when migrated
0	46 - 55	36 - 45
1	16 - 25	16 - 25
2	16 - 25	36 - 45
3	16 - 25	Below 15
4	16 - 25	16 - 25
...	...	...
395	36 - 45	36 - 45
396	36 - 45	16 - 25
397	36 - 45	Below 15
398	36 - 45	Below 15
399	36 - 45	16 - 25

400 rows × 2 columns

```
In [2]: data['Present age'].map({'16 - 25':1,'26 - 35':2,'36 - 45':3,'46 - 55':4,'56 and above':5})
```

```
Out[2]: 0      4
        1      1
        2      1
        3      1
        4      1
        ..
        395    3
        396    3
        397    3
        398    3
        399    3
        Name: Present age, Length: 400, dtype: int64
```

```
In [3]: data['Age when migrated'].map({'Below 15':0,'16 - 25':1,'26 - 35':2,'36 - 45':3,'46 - 55':4,'56 and above':5})
```

```
Out[3]: 0      3
        1      1
        2      3
        3      0
        4      1
        ..
        395    3
        396    1
        397    0
        398    0
        399    1
```

```
In [4]: data['Present age']=data['Present age'].map({'16 - 25':1,'26 - 35':2,'36 - 45':3,'46 - 55':4,'56 and above':5})
        data['Age when migrated']=data['Age when migrated'].map({'Below 15':0,'16 - 25':1,'26 - 35':2,'36 - 45':3,'46 - 55':4,'56 and above':5})
        data.head()
```

```
Out[4]:
```

	Present age	Age when migrated
0	4	3
1	1	1
2	1	3
3	1	0
4	1	1

```
In [5]: x=data['Present age']
        y=data['Age when migrated']
```

```
In [6]: x
```

```
Out[6]: 0      4
        1      1
        2      1
        3      1
        4      1
        ..
        395    3
        396    3
        397    3
        398    3
        399    3
        Name: Present age, Length: 400, dtype: int64
```



```

In [7]: y
Out[7]: 0      3
        1      1
        2      3
        3      0
        4      1
        ..
        395    3
        396    1
        397    0
        398    0
        399    1
        Name: Age when migrated, Length: 400, dtype: int64

In [8]: np.corrcoef(x,y)
Out[8]: array([[1.          , 0.54173681],
               [0.54173681, 1.          ]])

In [9]: np.corrcoef(x,y)[0,1]
Out[9]: 0.5417368115654785

In [10]: from scipy.stats import pearsonr
         pearsonr(x,y)
Out[10]: PearsonRRResult(statistic=0.5417368115654774, pvalue=6.938718237981657e-32)

```

## Output:

Pearson correlation coefficient (r): 0.5417

Two-tailed p-value: 6.93e-32

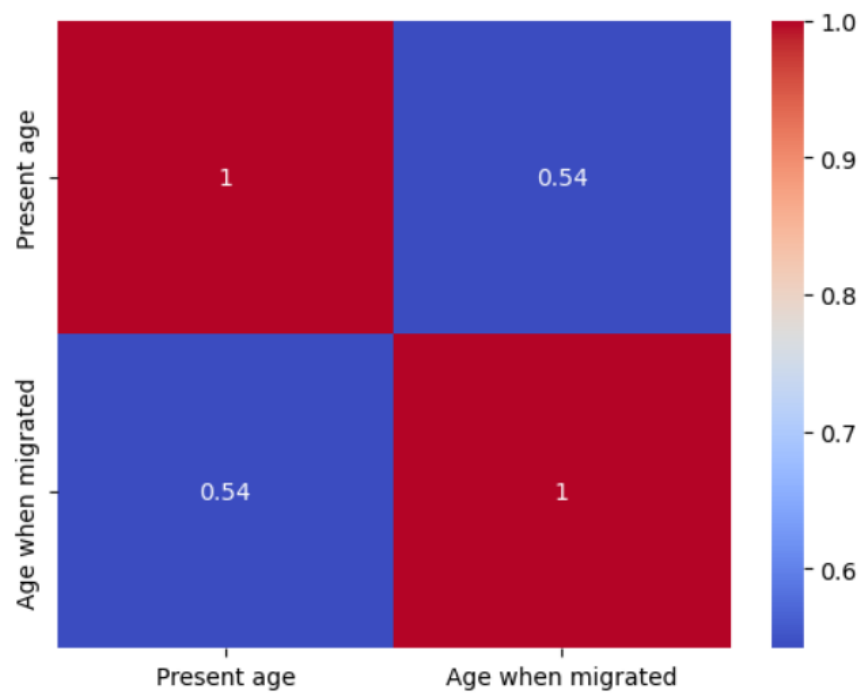
Since the correlation coefficient is close to 1, this tells us that there is a strong positive association between the two variables.

And since the corresponding p-value is less than 0.05, we conclude that there is a statistically significant association between the two variables.

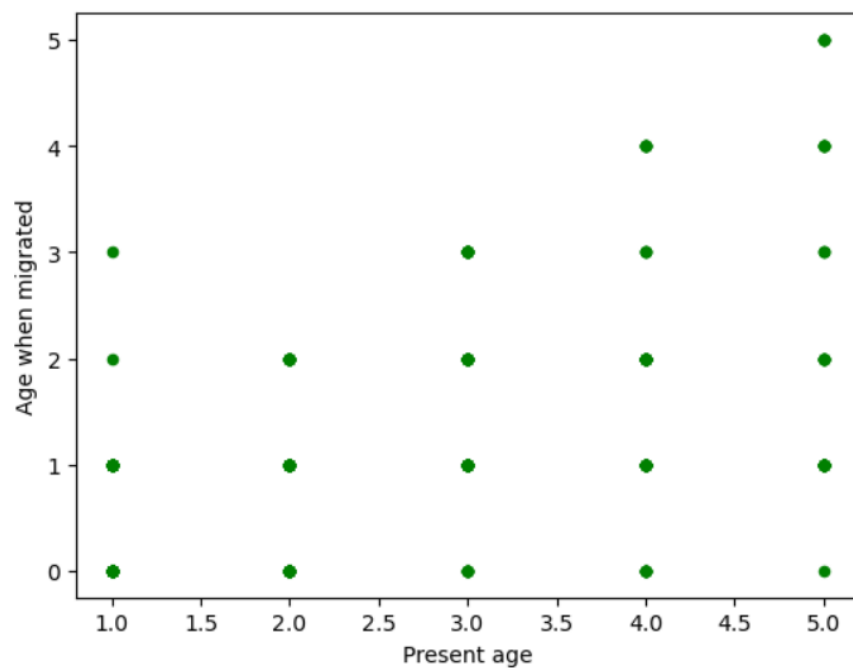
Following is the heatmap and scatter plot for the graphical representation for the correlation between the variables

```
In [11]: import seaborn as sns
sns.heatmap(data.corr(), annot = True, cmap= 'coolwarm')
```

Out[11]: <AxesSubplot:>



```
In [12]: scatter_plot=data.plot.scatter(x="Present age", y="Age when migrated", color="green")
```



# Chapter 5

## *Conclusion*



The survey was conducted on a sample of 400 respondents which consisted of different age groups where the respondents were migrating for various reasons. It was observed that 31.50% of respondents have migrated for employment, 9% of respondents have migrated because of marriage, 37.50% of respondents have migrated for education purposes, 7.50% of respondents have migrated because of family reunification while 14.5% of respondents have other reasons where others include the reasons like Covid-19(3%), Only for business related migration (1.25%), Children education (3.25%), Retirement (3.50%), Daily wages (3.25%), Family Migration (0.25%).

After applying the suitable tests on the data and analyzing it, the following results came to be known.

✚ The objective is to examine the proportions for males and females who are migrating 1<sup>st</sup> time. So, to study this, it should be checked whether the proportions of males and females who are migrating 1<sup>st</sup> time are same or not. Then it was found that out of 400 respondents, 67% respondents are migrating 1<sup>st</sup> time. After analysing this data, It was found that the proportions of males and females who are migrating 1<sup>st</sup> time are not the same.

✚ One of the objectives is to check whether there is an association between gender and reasons for migration. After analysing the data, it was found that there is an association between gender and reasons for migration. The findings from 2011 census show that marriages made up for 46 percent of the total migrations in India, of which 97 percent are women. Marriage is the biggest reason why women migrate from one place to another. But here the scenario changes completely for female migration. Through this survey, it was found that females are migrating more due to education. Increased education seems to encourage more migration of

women. Higher education levels have been found to be positively associated with increased migration for women. Also, employment is found to be the main reason for migration among men. As for the male migrants, economic reasons are the single dominant and the most important cause behind the migration in all of the four streams of migration.

✚ Later, it was found that reason for migration is related to a person's age. For example, migration motivated by health reasons is a phenomenon characteristic of old people, whereas education-related migration is predominantly associated with young people. Thus, in order to understand better why people move it is important to disaggregate cause-specific migration data by age and by sex. The age at which youths are most likely to first migrate is in the age of 16 – 25 for the educational purposes, reflecting a youthful migration profile. Moreover, above 25, migration is mostly related to employment, business or job transfers for males and marriage for females.

✚ Also, it was found that there is no association between the duration of stay and gender. Long-term migration or short-term migration is independent of gender.

✚ Moreover, from the results of the logistic model, it was found that if a person's reason for migration is education and currently attending school / college then there is a 71.06% chance that the person is likely to migrate next time.

Also, if a person's reason for migration is family reunification and currently attending school / college, then there are 18.83%% chances that

the person is likely to migrate next time. Thus, the result varies for different conditions and different types of reasons for migration.

- ✚ Through this survey, it was found that rural-urban migration has taken place in more percentage (%) as rural to urban areas pertains to the movement of people from one place to another to seek better opportunities. Various facilities like educational institutions, better healthcare facilities, financial institutions, and improved financial independence influence the people living in rural areas to move to urban areas and lead a better life.

## **Summary:**

- ❖ Urban migration is a common situation in which people from rural areas move to urban areas due to attraction towards employment, education, health, and some other factors. Migration mainly happens from rural areas to urban areas for many reasons. The lack of employment in rural areas is the main reason for the migration from rural areas to urban areas. In urban areas, there will be more employment opportunities due to factories and industries. Another reason for the migration from the rural areas to urban areas is for the attraction of better education facilities in the urban areas. Schooling in rural areas is minute and will lack in the quality of teaching.
- ❖ Urban to urban migration is when a person moves from an urban locality to another urban locale. This kind of migration is mostly related to securing better employment opportunities, better facilities or a better standard of living.

# Chapter: 6

## Reference





- Gupta S C, Kapoor V K (1970): Fundamental of Mathematical Statistics, first edition: Sultan Chand and Sons.
- Darren George, Paul Mallery: SPSS for Windows Step by Step (6<sup>th</sup> edition).
- S.L. Gupta, Hitesh Gupta: Research Methodology (2<sup>nd</sup> edition).
- Sudha G. Purohit, Sharad D. Gore, Shailaja R. Deshmukh: Statistics Using R.
- Statistics made simple Do It Yourself on PC (2<sup>nd</sup> edition) : K.V.S. Sarma
- online videos on statistical analysis
- online videos on logistic regression using python.

Websites:

<https://realpython.com/logistic-regression-python/>

<https://www.shethepeople.tv/news/marriage-common-reason-migration-women/>

[https://en.m.wikipedia.org/wiki/Main\\_page](https://en.m.wikipedia.org/wiki/Main_page)

<https://www.researchgate.net/>

<https://www.simplilearn.com/tutorials/machine-learning-tutorial/logistic-regression-in-python>

[\(PDF\) Structural Dynamics of Various Causes of Migration in Jaipur \(researchgate.net\)](#)

# Annexure

## A). Coding.

Sr.no	Question	Variable Name	Category	Code
1		Name/ E-mail ID		
2	Q.2	Gender	Female	0
			Male	1
3	Q.3	Age Group	16 - 25	0
			26 - 35	1
			36 – 45	2
			46 - 55	3
			56 and above	4
4	Q.4	Are you, Student?	Yes	0
			No	1
5	Q.5	Qualification	Primary	0
			Secondary	1
			Diploma	2
			Graduate	3
			Post graduate	4
			Doctorate	5
6	Q.6	Do you have any occupation	Yes	0
			No	1

7	Q.7	Occupation	Business	0
			Government service	1
			Private service	2
			Others (daily wages / labor/part time/housewife	3
8	Q.8	Salary	50,000	0
			1 lakh	1
			2 lakhs	2
			4 lakhs	3
			5 lakh and above	4
9	Q.9	Marital status	Single	0
			Married	1
			Divorcee	2
			Widow	3
10	Q.10	Family structure	Single	0
			Joint family	1
			Nuclear family	2
11	Q.11	Reasons for migration	Employment	0
			Marriage	1
			Education	2
			Family Reunification	3
			Others	4

12	Q.12	Migrating 1 <sup>st</sup> time	Yes	0
			No	1
13	Q.13	Type of move	Urban – Urban	0
			Rural - Urban	1
14	Q.14	Age in which 1 <sup>st</sup> leave their place of residence	Below 15	0
			16 – 25	1
			26 – 35	2
			36 – 45	3
			46 - 55	4
			56 and above	5
15	Q.15	Duration of stay	6 months (short term migration)	0
			1 year and more (long term migration)	1
16	Q.16	Chance of migration	Yes	0
			No	1
			Maybe	2

## B). Data sheet

Q.2	Q.3	Q.4	Q.5	Q.6	Q.7	Q.8	Q.9	Q.10	Q.11	Q.12	Q.13	Q.14	Q.15	Q.16
0	3	1		0	1	3	1	2	0	0	0	3	1	1
1	0	0	4	0	3	4	0	2	2	1	0	1	1	0
0	0	1		0	0	4	0	2	3	1	1	3	0	1
1	0	0	4	1			0	2	2	1	1	0	1	2
0	0	0	4	1			0	1	2	1	1	1	1	0
0	0	0	4	1			0	1	2	1	1	1	1	0
0	0	0	4	0	3	1	0	0	2	0	1	1	1	0
1	0	1		0	2	4	0	0	0	0	0	1	1	2
0	0	0	4	1			0	0	0	0	0	1	0	1
1	0	1		0	2	4	0	2	0	1	1	1	1	0
0	0	1		0	2	1	0	2	0	1	0	1	1	0
0	0	1		0	2	3	0	2	0	0	1	0	0	2
0	0	1		1			0	2	2	0	1	0	0	1
1	0	0	3	1			0	0	2	0	1	0	1	0
1	0	0	3	1			0	0	2	0	0	1	0	2
1	0	0	4	1			0	2	3	0	0	1	1	2
0	0	1		0	2	2	0	1	0	0	1	1	1	0
0	0	1		1			0	2	2	0	1	1	1	2
0	0	0	4	1			0	2	0	0	1	1	0	2
0	0	1		0	1	2	0	2	0	0	0	1	1	0
0	0	0	4	1			0	1	2	0	1	1	1	2
1	0	0	3	1			0	2	2	0	1	1	0	0
0	0	0	4	1			0	1	2	0	1	1	0	1
0	0	0	4	1			0	2	2	0	0	1	1	1
0	0	0	1	1			0	1	2	0	1	0	1	2
0	0	1		0	2	3	0	0	3	0	1	1	1	2
0	0	1		0	2	3	0	2	0	0	0	1	0	2
0	1	0	3	1			0	0	2	0	1	1	0	2
0	0	0	4	1			0	1	2	0	0	0	1	2
1	0	0	3	1			0	0	2	1	1	1	1	0
1	0	0	3	1			0	0	0	0	0	1	1	2
1	0	0	3	1			0	0	0	1	0	1	1	0
0	0	0	4	1			0	2	2	0	1	0	1	2
0	1	1		1			1	2	1	0	0	2	1	0
1	0	0	4	1			0	2	2	1	0	1	1	0
1	0	1		1			0	1	2	0	0	1	1	1
1	0	0	4	1			0	1	2	1	0	0	0	2
1	0	0	3	1			0	1	2	0	1	1	0	0
0	0	0	4	1			0	2	2	1	1	1	1	0
1	0	0	4	0	2	3	0	0	2	1	0	1	1	0
1	0	0	3	1			0	2	2	1	1	1	1	0
1	0	1		0	2	4	0	2	0	0	0	1	0	2
0	0	0	4	1			0	1	2	0	0	1	1	2
1	0	1		1			0	0	2	0	0	0	1	0
1	0	1		1			0	2	2	0	0	1	0	0
1	0	1		1			0	0	2	0	1	1	0	0
0	0	0	4	1			0	1	2	0	1	1	1	2
0	0	0	3	0	1	1	0	1	2	0	1	0	1	2
0	0	0	4	1			0	2	2	0	0	1	1	0
0	0	0	4	1			0	2	2	0	1	1	1	0
1	0	0	4	1			0	1	2	0	0	0	1	0
0	0	1		0	2	4	0	2	0	1	1	1	1	0
0	0	0	4	1			0	2	2	1	1	0	1	0
0	0	0	4	1			0	1	2	0	0	1	1	0
1	0	1		0	2	3	0	2	0	0	0	1	1	2
1	0	1		1			0	0	2	0	0	1	0	0
1	0	0	4	0	3	0	0	0	2	0	1	1	1	2
0	0	0	4	1			0	2	2	0	1	1	1	2
0	0	0	1	1			0	2	2	0	0	1	0	2
1	0	0	4	1			0	1	2	0	1	1	1	0

0	0	0	3	1			0	0	2	0	1	1	1	2
0	0	0	4	1			0	0	2	1	1	0	1	2
1	0	0	3	1			0	2	2	0	0	1	0	2
0	1	1		1			0	2	2	1	0	1	1	0
1	0	1		1			0	1	2	1	1	0	1	2
0	0	0	4	1			0	2	2	1	0	0	1	1
0	0	0	4	1			0	1	0	0	1	1	0	1
0	0	0	4	1			0	2	2	1	0	1	1	0
1	0	0	3	1			0	1	0	0	0	1	1	1
0	0	0	3	1			0	2	2	1	0	1	1	0
0	0	1		0	2	0	0	0	2	1	1	0	0	2
1	0	0	3	1			0	0	2	1	0	1	1	0
1	1	1		0	2	1	0	2	2	0	1	1	1	2
1	0	1		1			0	0	2	0	0	1	1	2
0	0	0	4	1			0	1	2	0	1	1	1	2
0	0	0	4	1			0	2	2	1	1	0	0	1
1	0	0	3	1			0	0	0	0	1	0	1	2
1	1	1		0	1	4	0	0	0	0	0	1	1	0
1	1	0	5	1			0	2	0	1	0	0	1	0
0	1	1		0	0	4	0	2	4	1	0	0	0	2
1	0	0	5	1			0	0	2	1	0	1	1	0
0	1	0	2	0	0	1	1	2	2	0	1	1	1	2
0	0	0	4	1			0	0	2	0	0	1	1	0
1	1	1		0	0	4	1	2	0	1	0	0	1	0
0	0	0	4	1			0	2	2	0	0	1	1	2
0	0	0	4	1			0	2	2	0	1	1	1	2
0	1	0	4	1			0	0	2	0	0	2	1	0
0	0	0	4	1			0	2	2	0	1	1	0	0
1	0	0	3	1			0	0	2	0	0	1	0	2
0	1	1		0	2	2	1	1	2	1	0	2	0	2
0	0	0	3	1			0	0	2	1	0	1	1	0
1	0	0	2	1			0	1	2	1	0	0	0	2
0	0	0	4	1			0	1	2	0	1	1	0	0
0	0	0	3	1			0	2	2	0	0	1	1	0
1	1	0	5	0	2	2	0	0	2	0	1	0	1	2
0	0	0	3	1			0	1	2	0	1	0	0	2
1	0	0	4	1			0	0	2	0	0	1	0	0
1	3	1		0	1	4	1	1	0	0	1	1	1	1
1	3	1		0	1	4	1	1	0	0	0	0	1	1
0	0	1		0	2	3	0	2	0	0	0	1	1	2
0	1	1		0	2	1	1	2	1	0	0	2	0	1
0	0	1		0	2	3	0	2	2	1	0	1	1	0
0	0	1		1			0	1	2	0	0	1	0	2
0	0	1		1			0	0	2	0	0	1	0	0
1	0	1		1			0	0	0	1	0	2	0	1
1	0	0	4	1			0	0	2	1	1	1	1	2
1	0	1		0	2	3	0	0	2	0	1	0	1	0
0	0	1		0	2	0	0	1	0	0	0	0	0	1
1	0	0	1	1			0	0	2	0	0	1	0	0
0	1	1		0	2	1	0	1	0	0	0	1	1	2
0	0	0	3	1			0	2	2	0	1	1	1	0
0	0	0	4	1			0	1	2	0	0	1	0	1
1	0	1		0	2	4	0	0	0	0	0	1	0	2
0	2	1		1			1	2	1	0	1	1	1	1
0	2	1		0	3	0	3	2	2	0	1	3	1	0
0	2	0	4	0	1	4	1	1	2	0	1	1	1	2
0	0	0	4	0	2	2	0	1	0	0	0	1	1	2
0	0	0	4	1			0	2	2	0	1	1	1	0
1	1	1		0	2	4	1	1	0	1	1	1	1	0
1	1	0	3	0	2	2	0	1	2	0	1	1	0	1
1	0	0	3	1			0	2	0	1	0	0	0	2
0	0	0	4	1			0	2	2	0	1	1	1	1
1	0	0	4	1			0	0	2	1	0	0	1	1
1	4	1		0	1	3	1	1	0	1	1	5	1	2
1	0	0	4	1			0	0	2	0	0	1	0	0

1	0	0	3	0	3	1	0	1	0	1	1	0	1	0
1	0	0	4	1			0	2	2	1	1	1	1	0
1	1	1		0	2	4	0	1	0	1	0	1	1	0
1	1	0	4	0	0	1	0	1	2	1	1	0	1	0
1	1	1		0	2	2	1	1	0	0	0	1	0	0
0	1	1		1			1	0	1	0	0	1	0	0
0	0	0	2	1			0	0	2	0	1	0	0	0
1	0	1		0	2	4	0	0	0	0	0	1	1	0
0	0	0	1	1			0	1	2	0	1	0	0	0
1	2	1		0	1	4	1	0	0	1	0	1	1	0
0	1	1		1			1	0	1	1	1	0	1	2
0	1	1		1			1	1	0	1	0	1	1	0
0	0	0	4	1			0	2	2	0	0	1	1	0
1	0	1		1			0	1	2	1	1	1	1	2
1	0	0	3	1			0	1	0	1	0	0	1	2
0	0	0	4	1			0	2	2	1	0	1	1	0
1	0	0	4	1			0	2	2	0	1	0	1	0
0	0	1		0	3	1	0	1	2	0	0	1	1	2
1	1	1		0	0	4	1	1	1	1	1	0	1	2
1	2	1		0	1	1	1	0	0	0	0	0	1	0
0	0	0	4	1			0	2	2	1	1	0	0	2
1	0	0	4	1			0	1	2	0	1	1	1	0
1	0	0	4	0	0	3	0	1	2	0	1	1	1	2
0	3	1		0	2	4	3	2	0	0	1	2	1	1
1	3	1		0	1	4	1	1	0	0	1	1	0	1
0	0	1		0	3	1	1	2	1	0	1	1	1	2
1	1	1		0	1	2	1	2	3	1	1	1	1	2
1	1	1		0	1	3	1	2	0	0	1	2	0	2
1	3	1		0	1	3	1	1	0	0	1	3	0	1
1	3	1		0	1	4	1	2	0	1	0	2	1	2
0	2	1		0	1	4	1	2	0	0	0	2	1	2
0	2	1		0	2	2	1	2	0	0	1	1	1	1
1	3	1		0	1	4	1	2	0	1	0	2	1	1
1	4	1		0	3	1	1	2	4	0	1	3	1	1
0	0	1		0	2	1	1	2	0	1	0	1	1	0
1	3	1		0	1	4	1	2	0	1	0	1	1	0
1	3	1		0	0	4	1	0	0	0	1	1	1	1
1	4	1		0	3	1	1	2	0	1	0	1	1	0
1	3	1		0	3	0	1	1	1	0	1	0	1	1
1	2	1		0	2	3	1	2	3	1	0	1	1	1
1	0	0	4	1			0	0	2	0	0	1	1	1
1	0	1		0	2		3	0	0	0	0	1	1	0
0	0	0	3	1			0	2	2	0	1	0	1	0
1	0	1		0	2	3	0	2	0	0	0	1	1	1
1	0	1		1			0	0	2	0	0	0	1	1
0	0	1		0	2	4	0	2	0	1	0	1	1	0
0	0	1		0	2	3	0	0	2	0	0	1	1	2
1	0	0	4	1			0	0	2	0	1	1	1	2
1	0	1		0	0	4	0	1	0	0	0	1	0	1
0	0	0	3	1			0	2	2	0	0	1	1	0
0	0	0	3	1			0	1	2	0	0	1	0	0
1	3	1		0	1	3	1	1	0	1	0	3	1	0
0	3	1		0	2	1	1	1	3	1	0	4	0	0
1	0	1		0	2	3	0	0	0	0	0	1	0	2
1	0	1		0	3	1	0	0	2	1	0	1	1	2
1	0	0	4	1			0	1	0	0	0	1	1	2
1	0	1		1			0	2	2	0	0	1	1	0
1	0	0	3	1			0	0	2	1	0	1	1	0
1	3	0	5	0	1	4	1	0	0	1	1	2	1	2
0	2	1		0	2	2	3	0	0	1	1	1	1	0
1	0	0	0	1			0	1	2	0	1	0	0	0
0	1	0	5	1			0	0	2	0	0	1	0	0
1	1	0	4	1			0	0	2	1	0	1	1	0
1	3	1		0	2	3	1	2	0	0	0	2	1	0
1	2	1		0	2	3	0	0	3	1	1	2	1	2
1	4	1		1			1	0	0	1	1	3	1	2

0	0	0	1	1			0	1	2	0	1	1	1	0
0	0	1		1			0	1	0	0	1	1	1	0
1	1	0	4	1			0	1	2	0	0	1	0	0
0	0	0	3	1			0	0	2	0	0	1	1	0
1	0	0	3	1			0	0	2	1	1	1	1	0
0	2	1		0	2	4	1	2	0	0	0	3	1	2
0	1	0	5	1			1	2	1	0	0	1	0	1
0	1	1		1			0	1	2	1	1	0	0	1
0	4	1		0	3	0	1	2	1	0	1	1	1	1
1	4	1		0	1	3	1	2	0	1	1	2	1	1
0	2	1		1			1	1	3	0	1	2	0	1
0	3	1		0	3	0	1	2	2	0	1	0	1	1
1	4	1		0	3	3	1	2	4	1	1	5	1	1
0	0	1		0	3	0	1	1	1	0	1	1	1	2
0	2	1		0	3	0	1	1	1	0	0	1	1	2
1	3	1		0	3	0	1	1	0	1	1	2	0	0
0	0	0	3	1			0	0	2	0	0	1	1	0
1	3	1		0	3	0	1	1	0	0	0	4	0	2
1	1	0	5	0	2	1	0	1	2	1	1	1	1	0
0	2	1		0	3	1	1	1	1	1	1	1	1	1
1	1	1		0	1	1	2	1	4	1	1	1	1	2
0	0	0	4	1			0	0	2	0	1	1	1	2
1	1	0	5	0	3	1	0	2	3	1	1	1	1	2
0	2	1		0	1	4	1	1	0	0	1	3	1	2
0	0	1		0	3	1	0	2	0	0	0	1	0	2
0	4	1		0	3	0	1	2	2	0	1	0	1	1
1	4	1		1			1	2	2	0	1	1	1	1
0	3	1		0	3	0	1	2	1	0	1	2	1	1
0	0	1		0	3	1	0	2	0	0	0	1	0	2
0	1	1		0	2	0	0	1	2	0	1	1	0	1
0	3	1		0	3	0	1	0	1	0	1	1	1	2
0	3	1		0	3	0	1	2	0	0	0	2	1	2
1	1	1		0	2	4	1	2	0	0	1	0	1	0
1	4	1		0	1	4	1	1	0	0	1	1	1	0
1	4	1		0	3	1	1	0	1	0	1	1	1	2
1	4	1		0	3	3	1	2	4	0	1	3	1	1
1	0	1		0	3	1	0	2	0	0	0	1	0	2
1	0	0	4	0	3	0	0	1	3	0	1	1	1	2
1	1	0	3	0	2	0	1	1	0	1	1	1	0	0
0	1	1		0	2	0	0	2	0	0	1	1	1	0
0	2	1		0	3	0	1	2	1	0	1	1	1	1
1	1	1		0	2	2	0	2	0	1	1	1	1	0
0	3	1		1			1	1	1	0	0	2	1	1
0	2	1		0	2	1	1	2	0	0	1	2	1	2
0	0	0	4	1			0	2	2	0	0	1	1	0
0	2	1		0	3	1	2	0	1	0	1	2	1	1
1	2	1		0	3	3	1	1	4	0	1	1	1	1
1	2	1		0	2	1	2	0	4	0	1	1	1	1
1	2	1		0	3	2	1	0	4	1	1	2	1	0
1	2	1		0	2	3	2	0	4	1	1	1	0	0
1	1	1		0	3	0	2	0	4	1	1	0	0	0
0	0	0	4	1			0	0	2	0	1	1	0	1
0	0	0	4	1			0	0	0	0	1	1	0	1
1	1	1		1			0	0	0	0	0	2	0	2
0	0	0	4	1			0	2	2	0	0	1	0	2
0	0	1		0	2	3	0	2	0	0	0	1	0	0
0	1	1		1			1	1	1	0	1	1	1	2
1	2	1		0	1	4	1	2	0	1	0	1	1	0
1	3	1		0	1	4	1	1	2	0	1	0	0	2
0	3	1		0	1	1	1	1	0	0	0	1	1	2
1	2	1		0	1	4	1	2	0	1	0	1	1	2
0	1	1		0	1	4	1	1	0	0	0	2	1	0
0	3	1		0	2	2	1	1	2	0	0	0	0	2
0	0	0	4	1			0	2	3	1	1	1	0	1
1	2	0	4	1			1	1	0	1	1	1	1	2
1	4	1		0	1	4	1	2	0	1	0	1	1	0



1	3	1		0	0	3	1	1	0	0	0	2	1	1
0	1	1		0	3	1	1	1	0	0	1	1	1	0
1	2	0	0	1			0	1	2	0	1	1	0	0
1	2	1		0	1	4	1	1	2	0	1	0	1	1
0	1	1		0	1	4	0	2	2	0	1	0	1	0
0	1	1		0	1	4	0	0	0	1	1	0	1	1
1	4	1		0	1	0	1	1	0	0	1	1	1	1
1	1	1		0	1	2	1	2	0	0	0	1	1	1
1	3	1		0	1	2	1	2	0	0	1	1	1	1
1	2	1		0	1	4	1	1	0	1	0	2	1	0
1	4	1		0	1	4	1	0	0	0	1	5	1	1
0	2	0	3	1			1	2	0	0	1	2	1	1
0	2	1		0	1	0	1	2	0	0	1	3	1	1
0	1	1		0	1	0	1	1	0	0	1	1	1	2
0	2	1		0	3	0	1	1	1	0	1	1	1	2
0	4	1		1			3	1	4	1	1	1	1	1
0	1	1		0	1	4	1	1	1	0	0	2	1	2
1	0	0	4	1			0	1	2	1	1	1	1	2
1	0	0	3	1			0	0	0	0	1	0	0	2
0	0	1		0	2	4	0	1	0	0	0	0	0	0
1	1	1		0	1	4	0	0	2	0	0	1	1	2
0	0	1		0	2	2	0	1	2	0	0	1	1	0
1	1	1		0	2	1	1	2	0	1	1	1	1	0
0	1	1		0	2	1	1	2	1	0	0	2	1	2
0	0	0	3	1			0	0	2	1	1	1	0	0
1	0	0	3	1			0	1	0	0	1	1	1	0
0	0	0	3	1			0	1	2	1	0	0	0	2
1	0	0	3	1			0	0	2	0	1	1	1	2
1	0	0	3	1			0	0	2	0	1	1	1	2
0	0	0	3	1			0	1	2	0	0	1	1	2
1	1	1		0	2	1	0	1	2	1	1	1	1	0
1	0	1		1			0	1	2	1	1	0	1	0
0	4	1		1			3	0	4	1	1	1	1	1
0	4	1		1			3	2	1	0	1	1	1	1
1	3	1		0	0	4	1	1	4	0	1	0	1	1
0	4	1		0	2	4	0	0	0	1	1	1	1	2
1	4	1		0	0	3	2	1	4	0	1	2	1	1
1	4	1		0	3	3	1	2	0	1	1	1	1	1
0	2	1		0	1	4	1	0	2	0	0	1	1	2
0	2	1		0	1	4	1	2	1	0	1	2	1	2
0	2	1		0	1	3	1	1	0	0	1	2	1	0
1	2	1		0	1	4	1	2	0	0	0	3	1	0
1	3	1		0	1	0	1	1	0	0	0	2	1	0
1	1	1		0	1	4	1	2	0	0	1	2	1	0
1	2	1		0	1	4	1	2	0	0	1	1	1	0
1	3	1		0	1	4	1	2	0	1	0	1	1	1
1	3	1		0	1	4	1	1	0	0	1	2	1	2
1	1	1		0	1	4	1	2	0	0	0	0	1	0
1	2	1		0	1	3	1	2	0	0	0	2	1	2
1	2	1		0	1	3	1	1	0	0	0	2	1	2
1	1	1		0	1	2	1	1	0	0	0	2	1	2
0	2	1		0	1	4	1	2	3	0	0	3	1	0
1	2	1		0	1	4	1	1	3	0	1	1	0	1
1	2	1		0	1	4	1	1	0	0	1	3	1	0
1	1	1		0	1	4	1	1	0	0	0	2	0	0
1	3	1		0	1	4	2	0	0	1	1	1	1	1
0	1	0	5	1			0	1	2	1	1	1	1	0
1	3	1		0	1	3	1	2	0	0	1	2	1	1
1	4	1		1			1	2	4	1	1	1	1	1
0	2	1		0	3	0	3	1	1	0	1	2	1	1
0	2	1		0	3	0	1	2	1	0	1	2	1	1
0	3	1		0	0	4	2	0	0	1	1	1	1	1
1	4	1		0	3	2	2	1	3	1	1	2	1	1
1	3	1		0	3	2	1	2	0	0	1	1	1	1
0	3	1		0	3	0	3	1	1	1	1	1	1	1
1	3	1		0	2	4	2	0	0	0	1	4	0	0

1	3	1		0	2	2	2	0	0	1	1	0	1	1
0	2	1		0	1	3	1	2	1	1	1	2	1	0
0	3	1		0	3	0	1	2	1	0	0	1	1	2
1	4	1		1			1	1	4	1	1	1	1	1
1	3	1		0	0	4	1	2	4	0	1	3	1	1
1	3	1		0	2	3	2	0	0	0	0	4	0	2
0	2	1		0	1	3	1	0	1	0	0	0	1	2
1	4	1		0	1	4	1	2	4	1	1	1	1	1
0	2	1		0	3	0	1	1	1	0	1	1	1	2
0	4	1		0	1	4	1	2	4	1	1	2	1	1
0	3	1		0	3	0	1	1	1	0	0	2	1	2
1	2	1		0	2	3	1	1	4	0	1	3	1	2
1	2	1		0	0	4	0	0	4	0	1	2	0	0
0	2	1		0	2	2	1	1	1	0	0	2	0	0
1	4	1		0	1	4	2	0	4	1	1	2	1	1
1	2	1		0	3	2	1	1	4	0	1	3	0	0
1	4	1		1			1	1	4	0	1	5	1	1
0	3	1		0	3	0	1	2	4	1	1	2	1	1
1	3	1		0	0	3	1	1	4	0	0	4	0	2
0	4	1		0	3	0	2	1	1	0	1	2	1	1
1	1	1		0	2	2	0	1	4	0	1	2	0	2
1	4	1		0	1	4	1	2	4	1	1	2	1	1
0	4	1		0	2	2	3	1	3	0	0	5	0	1
1	3	1		0	0	3	1	1	4	1	0	2	1	1
0	3	1		0	1	4	1	2	1	1	1	2	1	1
1	1	1		0	2	4	0	0	0	0	0	2	0	2
1	3	1		0	3	2	1	2	4	1	1	4	0	2
1	3	1		0	0	4	1	2	3	1	1	1	1	1
1	0	0	0	1			0	0	2	0	1	0	1	2
1	0	0	1	1			0	0	2	0	1	0	0	0
1	0	0	1	1			0	0	2	0	1	0	0	0
1	0	0	1	1			0	1	3	1	1	0	0	1
1	0	0	1	1			0	2	3	1	1	0	1	1
1	0	0	0	1			0	2	3	1	1	0	1	0
1	0	0	2	0	3	0	0	0	4	0	1	0	1	1
1	0	0	2	0	3	0	0	0	4	1	1	0	1	1
1	0	0	2	1			0	2	3	1	1	0	1	2
1	0	0	2	1			0	2	3	1	1	0	1	0
1	0	0	2	0	3	0	0	0	4	1	1	0	1	1
1	0	0	1	1			0	0	3	0	1	0	1	1
1	0	0	0	1			0	1	3	0	1	0	1	1
1	0	0	1	1			0	2	3	1	1	0	1	2
1	0	0	0	1			0	0	3	1	1	0	1	1
1	0	0	1	1			0	2	2	0	1	0	1	1
1	0	0	4	0	0	4	0	0	2	0	0	1	1	1
1	0	0	0	1			0	0	2	0	1	0	0	0
1	0	0	0	1			0	2	4	0	1	0	1	0
1	0	0	0	1			0	2	4	0	1	0	0	0
1	0	0	0	1			0	2	4	0	1	0	0	0
1	0	0	0	1			0	2	4	0	1	0	1	0
1	1	1		0	0	4	2	0	4	0	1	0	0	0
1	0	1		0	0	1	1	2	4	0	1	0	0	0
1	1	1		0	0	3	2	0	4	0	1	0	0	0
1	1	1		0	0	4	2	0	4	0	1	0	0	0
1	1	1		0	0	4	2	0	4	0	1	0	0	0
1	1	1		0	0	4	2	0	4	0	1	0	0	0
1	0	0	0	1			0	0	3	0	1	0	0	0
1	1	1		0	0	3	0	1	4	0	1	0	0	0
1	1	0	5	0	0	4	2	0	4	0	1	0	0	0
1	4	1		1			1	2	4	1	1	4	0	0
1	4	1		1			1	1	4	1	1	4	0	0
1	4	1		1			1	1	4	1	1	4	0	0
1	4	1		1			1	1	4	1	1	4	0	0
1	4	1		1			1	2	4	0	1	4	0	0
1	4	1		1			1	1	4	1	1	4	0	0
1	2	1		0	0	4	1	2	4	0	1	0	0	0
1	2	1		0	0	4	1	2	4	0	1	1	0	0

1	2	1		0	0	4	1	2	4	0	1	1	1	0
1	2	1		0	0	4	1	1	4	1	1	0	1	0
1	2	1		0	0	4	1	0	4	0	1	3	0	0
1	2	1		0	0	4	1	0	4	0	1	3	0	0
1	2	1		0	0	4	1	1	4	0	1	3	0	0
1	2	1		0	0	3	1	2	3	0	1	1	1	1
1	2	1		0	0	3	1	2	3	0	1	0	1	2
1	2	1		0	0	3	1	1	3	1	1	0	0	2
1	2	1		0	0	4	1	2	3	0	1	1	0	1