POST GRADUATE DIPLOMA IN BUSINESS ANALYTICS (PGDBA)

IIM Calcutta, ISI Kolkata, and IIT Kharagpur

Project Report: Human Resource Management

Option 1: Survey Data

Cross Sectional Analysis of Unorganised Sector Workers in Urban Indian

Group Members

S.No.	Name	Roll Number
1	Subhasis Panda	19BM6JP20
2	Vaibhav Anand	19BM6JP24
3	Prasad Bankar	19BM6JP25
4	Hemanthkumar E	19BM6JP27
5	Sakshi Nagori	19BM6JP42
6	Vineet Kumar	19BM6JP46
7	Paturu Harish	19BM6JP55

ABSTRACT

About 93% of the Indian workforce belongs to the unorganised sector as per the economic survey 2018-19. Thus, it directly affects the lives of a significant number of households in India. Moreover, the effect of the unorganised sector goes beyond these numbers, since a lot of firms in the organised sector rely upon the unorganised sector at different stages of their value chain. However, despite being so central to the well-being of the Indian economy and the populace, the sector does not receive enough protection via regulation and a lot of analysis is based upon theories developed for the organised sector. In this work we attempt to analyse the unorganised sector by means of a survey conducted on cab drivers, security guards, and housekeeping staff. We aim to study the relationships amongst different covariates captured via the survey questions as well as try to find the features that contain maximum information about the data.

Keywords: survey, unorganised sector, multivariate factor analysis, mixed data.

TABLE OF CONTENTS

List of Figures and Tables

1. INTRODUCTION

Surveys often comprise tick-box questions where respondents are asked to select one (or potentially more) of a fixed number of possible options resulting in what are referred to statistically as categorical data. These types of questions are often preferred to open-ended free-text questions as they are generally easier to analyse and, since they are quicker to fill in, can result in higher response rates. Analysing categorical data requires the use of a specialist set of statistical tools as they are not normally amenable to the standard tools available for continuous data. With this report, we aim to apply relevant statistical tools to understand the questionnaire and intricacies related to forming a questionnaire. This also focuses on the kind of analysis an individual can do with categorical and continuous responses to such survey data.

We also make an attempt to understand the reasons why an analysis might sometimes lead to mis-leading conclusions. Flawed data can guide even the greatest leaders to the wrong conclusions. When success hangs in the balance, one needs to be absolutely sure that they are gathering the right data with the right methods.

The rest of this report is organised as follows --- Section 2 examines related works in the area of labour studies. Section 3 summarises our main objective. In section 4 we briefly describe the dataset and our methodology. This section also details the statistical and quantitative analysis followed by qualitative inference after each subsection. Section 5 describes the conclusion and some broader insights for more elaborate study and analysis.

2. RELATED WORK

According to Guy Ryder [1], the ILO Director-General, "Decent work" puts money in the pockets of individuals and families that they can spend in the local economy. Decent work reduces inequality and increases resilience. The Agenda for 2030 is to achieve full and productive employment and decent work for all women and men, including for young people and persons with disabilities, and equal pay for work of equal value. It also focuses on taking immediate and effective measures to eradicate forced labor, end modern slavery and human trafficking. It is of utmost importance to protect labor rights and promote safe and secure working environments for all workers, including migrant workers, in particular women

migrants and those in precarious employment. It is believed that Decent work for all, including social protection, is the main route out of poverty for individuals, communities and countries. Also, Decent work, with its emphasis on a fair income, security in the workplace and social protection of individuals and families, is a direct means to reduce inequalities in income, wealth and economic influence.

High levels of segmentation and informality is prominent in the Indian labor market. Of the total employed in 2011–12, more than half (51.4 per cent, or 206 million people) were self-employed, and of the 195 million wage earners, 62 per cent (i.e. 121 million) were employed as casual workers. Employment in the organized sector has grown, but even in this sector many jobs have been casual or informal.

The wage inequality in India still remains very high:
The Gini coefficient for wages is 0.49 and the D9/D1 wage ratio is 6.7. Regional disparities in average wages have increased over time as wages rose more rapidly in high-wage states than in low-wage ones. The Fair Wages Committee defined three levels of wages: living wage, fair wage and minimum wage. Among the three, the living wage constituted the highest level: covering food, clothing, shelter, education of children, health expenditure and old age insurance. Fair wage, envisaged as between the living wage and the minimum wage, includes subsistence plus standard wage. It considers national income, productivity and the capacity to pay of the industry, in the determination of its level. It should be set in such a way as to ensure continuation and growth of employment, at a level which is comparable to levels in similar occupations/ activities. A minimum wage was defined as one necessary for sustenance of life and some measure of education, meeting medical requirements and for amenities for the preservation of the efficiency of the worker. It is the absolute minimum below which wages should not be set.

The potential for minimum wages to reach low-paid workers depends on the level at which the minimum wage is fixed, and the mechanism, the extent of coverage of workers; most importantly, on whether the enforcement machinery is in place. Minimum wage legislation by itself cannot ensure that minimum wages would reach low-paid workers and efforts would need to be made to make it effective and put enforcement mechanisms in place.

There are several possible options to improve the current minimum wage system. These recommendations have been based on a vast body of existing literature and the ILO standards

(particularly the Minimum Wage Fixing Convention, 1970, No.131) as well as the conclusions of a capacity-building workshop titled "Towards more effective wage policies in India", jointly organized by V.V. Giri National Labour Institute and the ILO, New Delhi in April 2015.

These are:

- Extending legal coverage to all workers in an employment relationship.
- Ensuring full consultation and, in so far as possible, the direct participation
- of social partners on a basis of equality in the establishment and operation of
- minimum wage systems.
- Undertaking more regular and evidence-based adjustments.
- Progressively consolidating and simplifying minimum wage structures.
- Taking stronger measures to ensure a more effective application of minimum
- Wage law.

The traditional trade unions and new forms of organizations try to secure basic rights for various groups of informal workers and support service workers by negotiating with different stakeholders.

Extending the coverage of minimum wages in India:

In India, the Minimum Wage Act of 1948 is perceived as being of great importance, particularly to the unorganized casual workers. The paper states that for two-thirds of all wage-earners and a total number of about 116 million workers, minimum wage is important. One important discussion has revolved around the question of what is the appropriate level of the minimum wage to prevent labour - exploitation and provide a decent standard of living. Another debate concerns the way to increase compliance by elevating the minimum wage to a fundamental right, even equating noncompliance with a form of forced labour.

According to the analysis in the paper [3], the minimum wage has no adverse effects on the demand for labour, the Gini coefficient for wage inequality would fall from a very high close to 0.50 in actuality to anywhere between 0.39 and 0.41, which is closer to the level of inequality found in other lower middle-income countries.

An important effect of an extended minimum wage would be a sharp reduction in the gender pay gap. The paper finds that if all workers would receive at least minimum wages, average

wages of women compared to men would increase from 84 percent to 90 percent for salaried workers and from 74 per cent to 92 per cent for casual workers. In India, over 50 Million people are employed as domestic helps across the country, with women constituting over 75% of this sector, as indicated by Ghosh et al [4]. Children too, are a part of this workforce; nearly 200,000 minors are estimated to be employed as help.

In the eyes of the law, domestic workers neither fall under the definition of "workmen," nor their workplace is considered an "establishment." This characterisation is often used by the Centre to explain the lack of regulation in the industry.

On 12 December 2013, more than a lakh workers predominantly from the unorganised sector marched on Parliament to demand a minimum living wage, social security measures and regularization of work. The call had been given by trade unions across the political spectrum and the participants came from all parts of the country.

One reason for the huge presence of informal sector workers was that many of the demands in the charter related to them. The immediate and most pressing demand was that of fixing the "minimum wage to not less than Rs 10,000 per month linked with the consumer price index and an amendment to the Minimum Wages Act to ensure universal coverage irrespective of schedules". The general opinion was that the amount had to be the minimum living wage across the country considering price rise and living conditions. The demand for "assured pension for the entire population found resonance with employees in the informal sector.

Social security:

Social security in India was traditionally taken care of by the set up of family/community in general. With the rapid industrialization/urbanization beginning during the early 20th century resulting to an extent the break up of the family set up the need for institutionalized and State-cum-society regulated social security arrangement has been felt necessary. The problem has been aggravated further with the ageing of the society and embarking towards market economy

The Employee's Provident Funds and Miscellaneous Provisions Act, 1952 is a welfare legislation enacted for the purpose of instituting a Provident Fund for employees working in factories and other establishments. The Act aims at providing social security and timely

monetary assistance to industrial employees and their families when they are in distress and/or unable to meet family and social obligations and to protect them in old age, disablement, early death of the breadwinner and in some other contingencies.

The Act is applicable to factories and other classes of establishments engaged in specific industries, classes of establishments employing 20 or more persons. The Act, however does not apply to cooperative societies employing less than 50 persons and working without the aid of power.

There are multiple Acts which try to empower the working class. Some of them are noted.

The Payment of Gratuity Act, 1972: The Act provides for a scheme of compulsory payment of gratuity to employees engaged in factories, mines, oil fields, plantations, ports, railway companies, motor transport undertakings, shops or other establishments.

The Maternity Benefit Act, 1961: The Maternity Benefit Act, 1961 is a piece of social legislation enacted to promote the welfare of working women. The Act prohibits the working of pregnant women for a specified period before and after delivery. It also provides for maternity leave and payment of certain monetary benefits for women workers during the period when they are out of employment on account of their pregnancy. The Maternity Benefit Amendment Act of 2017 has increased the duration of paid maternity leave available for women employees from the existing 12 weeks to 26 weeks.

The Factories Act, 1948 regulates the working conditions in the factories and ensures that basic minimum requirements for the safety, health and welfare of the factory workers are provided. The Act also envisages to regulate the working hours, leave, holidays, overtime, employment of children, women and young persons.

Industrial Disputes Act, 1947: Chapter VA deals with those industrial establishments (Factory, Mines and units of plantations) which are not of seasonal character and in which at least 50 workmen on an average is employed for a continuous period of at least one year.

Lay-off: In case an employee is Laid-off, the employer shall pay compensation for the days laid-off at the rate of 50% of the (basic Salary plus Dearness Allowance) subject to the Maximum of 45 days. If any employee is laid-off beyond 45 days the Employer can retrench such employee after paying retrenchment compensation.

Social Insurance Schemes:

93% of India's workforce belongs to the unorganized sector. Few schemes listed below cover them for social security.

Janshree Bima Yojana (JBY): Janshree Bima Yojana (JBY) a group insurance scheme implemented by LIC is available to persons between ages 18 and 60 years and are living below or marginally above the poverty line.

Universal Health Insurance Scheme (UHIS): The Government has also launched a heavily subsidized Universal Health Insurance Scheme (UHIS) for BPL families.

State governments have their own schemes for social security of unorganized workers.

Overtime Rules and regulations in India: There are numerous overtime rules and procedures in India stipulating different periods of working hours under the labor law. Section 51 and Section 59 of the Factories Act, 1948 states, "No employee is supposed to work for more than 48 hours in a week and 9 hours in a day. Any employee who works for more than this period is eligible for overtime remuneration prescribed as twice the amount of ordinary wages." Additionally, Section 14 of the Minimum Wages Act, 1948 states, "When the minimum wages of an employee are fixed for a particular period of time and the employee works beyond that period, then the employee has to be paid overtime wages for the extra time."

Every state in India has its own Shops and Establishment Act (SEA) which also lays down overtime rules and procedures for workmen employed in different institutions. The SEA is applicable to all managerial and non-managerial positions, in every Indian shop and establishment. The overtime rules also state that the employees must be provided at least one break for half an hour between the working hours and the entire working period per day must be calculated in such a way that no working period exceeds 5 hours without an interval.

The total number of working hours in India in a day must be 12 and a half hours, limiting the maximum number of overtime hours in a day at 2 hours. Overtime rules under the Factories Act, 1948 also specifies the punishment in case an employer violates these provisions.

Any employer found to be contravening these provisions would be liable for punishment of imprisonment up to 2 years and fine of up to Rs. 1 lakh or both. If the employer continues to violate the provisions after conviction, a fine of Rs. 1000 per day is levied for each day of violation.

- Under the overtime payment rules in India, overtime wages are paid in two ways — either on a per hour rate or a per piece rate in factories.

- In a per hour or hourly rate, the per hour wage of an employee is calculated and double the amount is paid for every extra working hour.

- In per piece method, an employee is paid overtime for every extra piece made during the overtime period.

3. OBJECTIVE

To understand the associations amongst the variables assessed by means of the survey and explore the variability in the data.

4. ANALYSIS

4.1 Methodology

Since the available data was of mixed type, having numeric as well as qualitative features, Factor Analysis for Mixed Data (FAMD) was used. FAMD is a principal component method dedicated to explore data having both continuous as well as categorical variables. It can be roughly seen as a hybrid of Principal Component Method (PCA) and Multiple Correspondence Analysis (MCA).

Under FAMD, the continuous variables are scaled to unit variance, while the categorical variables are transformed into a disjunctive data table (crisp-coding) and then scaled using the specific scaling of MCA. This ensures that the influence of both continuous and categorical variables in the analysis is balanced and hence both kinds of variables are on an equal footing to determine the dimensions of variability. Using FAMD we can study similarities among individuals taking into account mixed variables as well as explore the relationships amongst all the variables. This can be done graphically using a correlation circle for continuous variables, categorical representations for categorical variables, and graphs for exploring associations.

We also tried to study the association between categorical features using Fisher's Exact Test (since the sample sizes were small), which essentially tests the independence of rows and columns in a contingency table with fixed marginals.

The experiments were performed in R version 3.6.3.

4.2 Cab Drivers Dataset

The survey of cab-drivers consists of workers catering to international ICT-ITES firms in Mumbai. The survey records wages, demographics and multiple HR related responses from respondents.

4.2.1 Factor Analysis of Mixed Data

Dimension	Eigenvalue	% Variance	% Cumulative Variance
1	9.409336	6.231348	6.231348
2	7.755498	5.136092	11.367440
3	6.660833	4.411147	15.778587
4	6.136400	4.063841	19.842428
5	5.510886	3.649593	23.492022

Table 4.1: Eigenvalues Obtained from FAMD Analysis
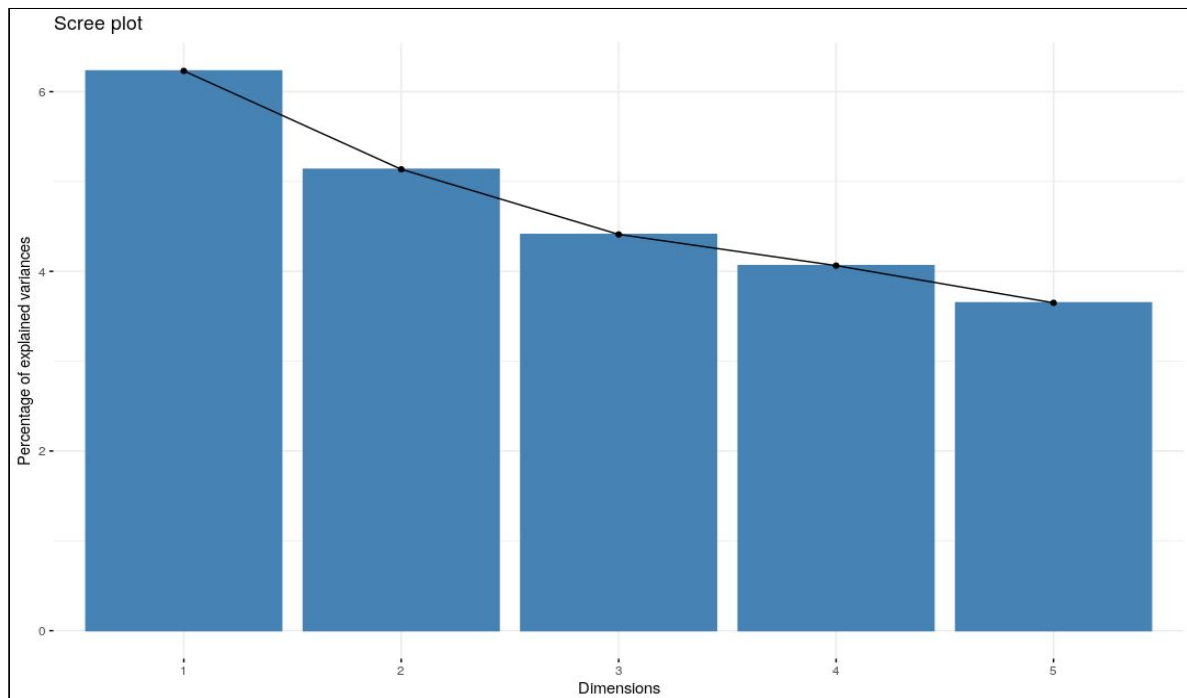

Fig 4.1: Scree Plot of Eigenvalues from FAMD Analysis
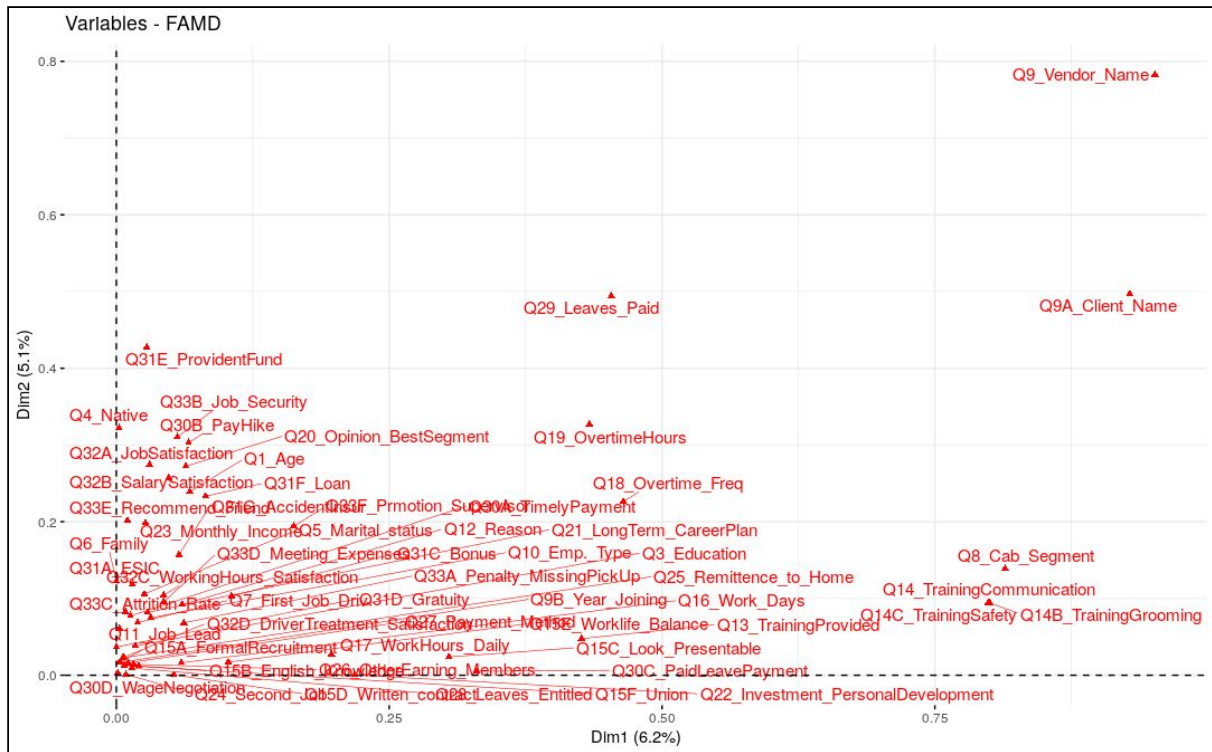
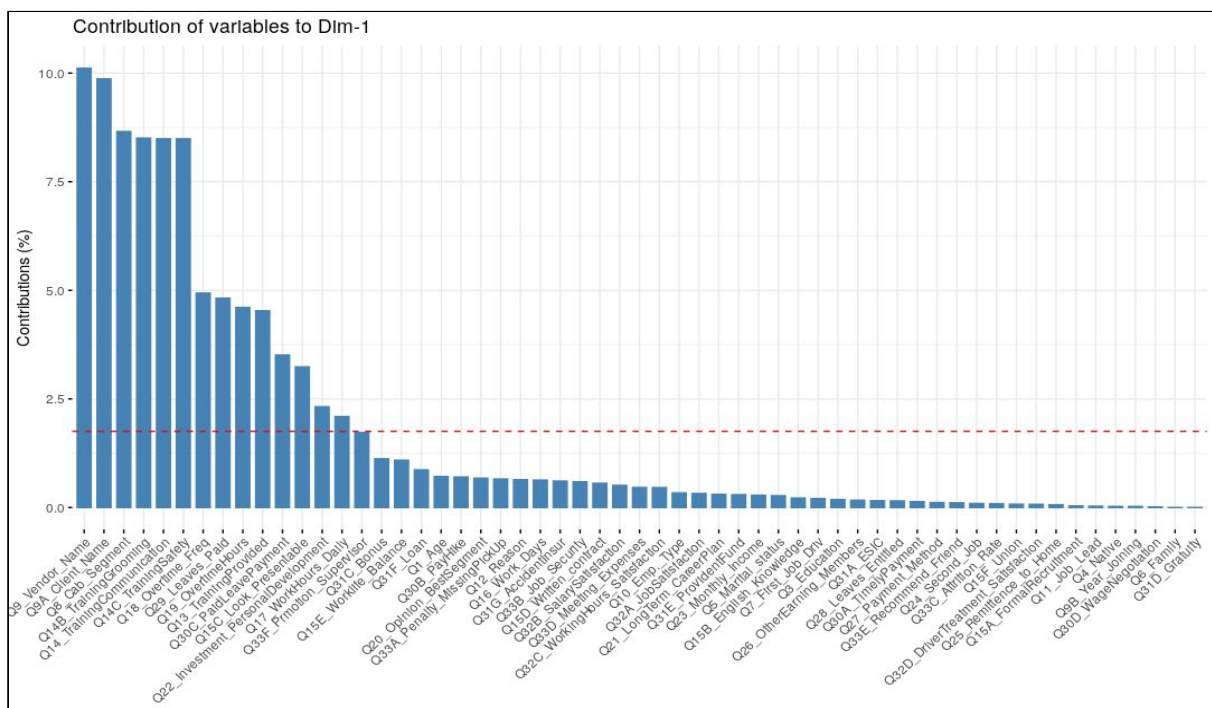**Fig 4.2: Illustration of Features Forming the First Two Dimensions**

Fig 4.3: Contribution of Features to the First Dimension
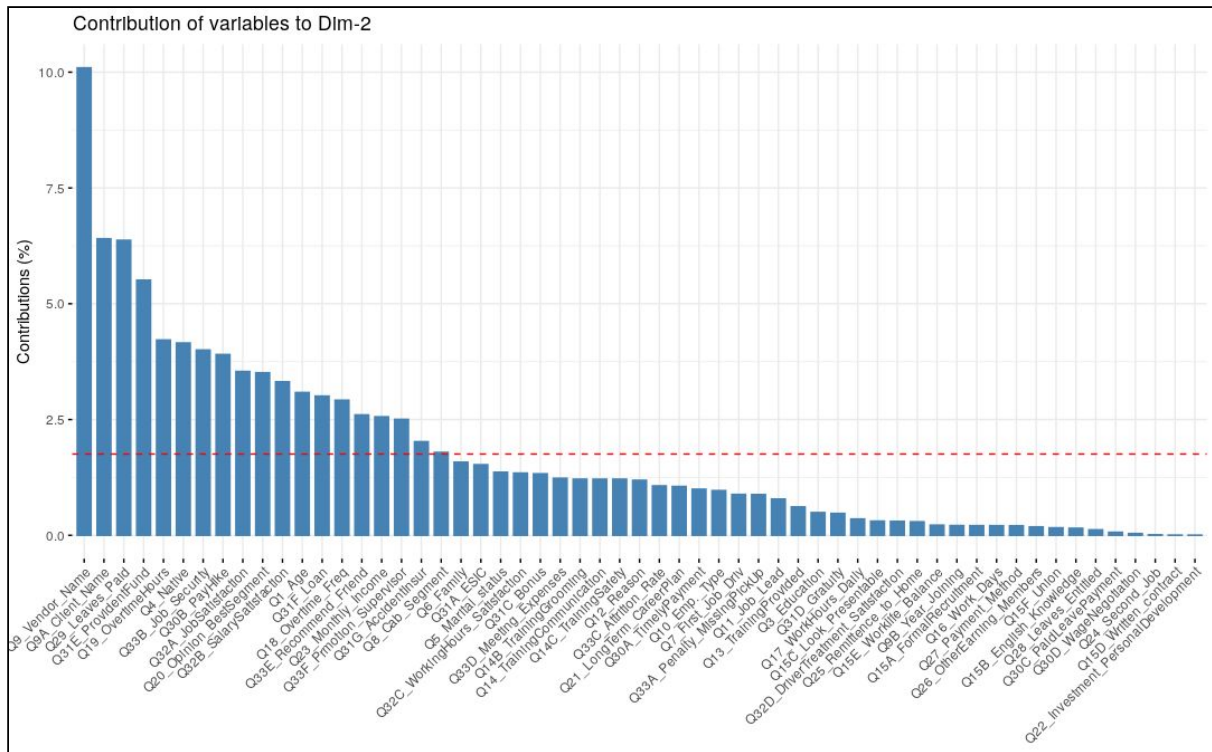
Fig 4.4: Contribution of Features to the Second Dimension
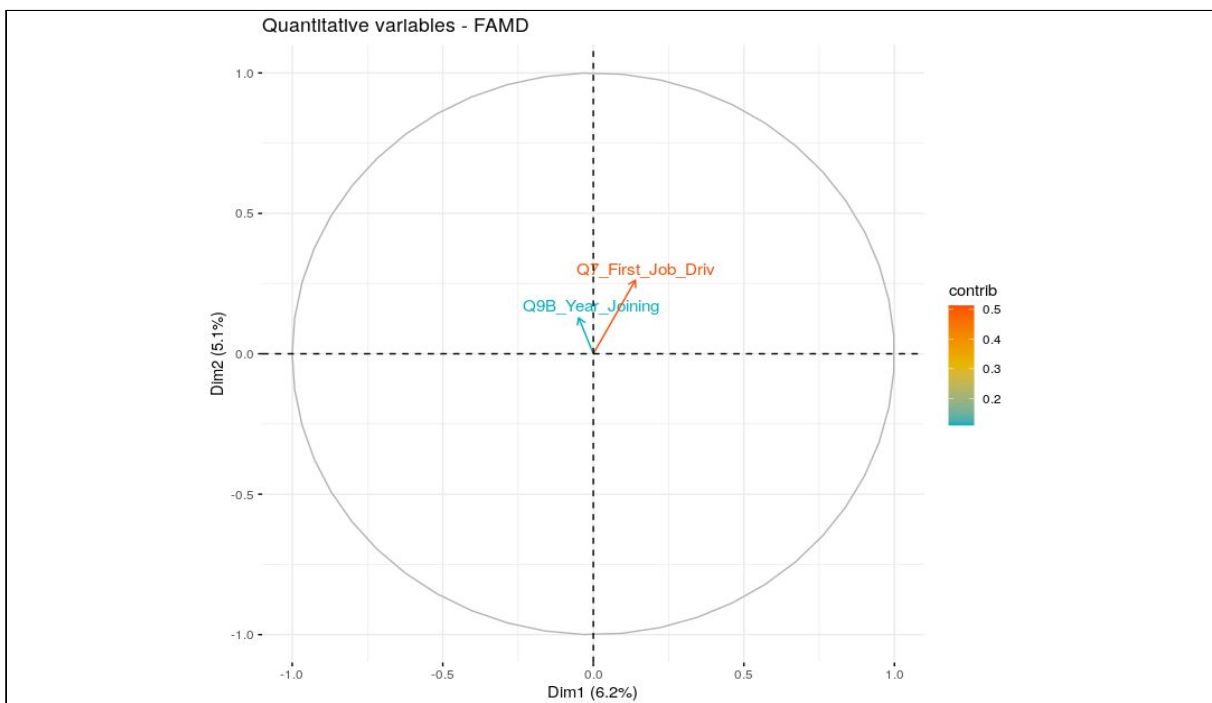
Fig 4.5: Contribution of Quantitative Features to the First Two Dimensions

4.2.2 Association Analysis Using Fisher's Exact Test

4.2.2.1 Different education levels vs opinions (Q3 vs Q20)

p-value = 0.6203 ; alternative hypothesis: two.sided

Conclusion: No association

4.2.2.2 Different education levels vs long term career plans (Q3 vs Q21)

```
                Not sure about it Remain in transport sector Will switch to other sector
10th pass               1                   16                          0
12th pass               0                   11                          1
Graudate/diploma        1                    2                          0
Stand. 1-5              0                    5                          1
Stand. 6-9              1                   12                          0

                Will switch to personal business
10th pass                       3
12th pass                       1
Graudate/diploma                0
Stand. 1-5                      2
Stand. 6-9                      2
```

p-value = 0.522 ; alternative hypothesis: two.sided

Conclusion: No association

4.2.2.3 Education and monthly income(Q3 vs Q23)

p-value = 0.3928 ; alternative hypothesis: two.sided

Conclusion: No association

4.2.2.4 Education and personal development (Q3 vs Q22)

p-value = 0.514 ; alternative hypothesis: two.sided

Conclusion: No association

4.2.2.5 Age and long-term career (Q1 and 21)

p-value = 0.6895 ; alternative hypothesis: two.sided

Conclusion: No association

4.2.2.6 Age and monthly income (Q1 and 23)

p-value = 0.5741 ; alternative hypothesis: two.sided

Conclusion: No association

4.2.2.7 Age and personal development(Q1 vs Q22)

p-value = 1 ; alternative hypothesis: two.sided

Conclusion: No association

4.2.2.8 Education and personal business as long-term goal

H_0: people with graduation/ diploma holders choose to go for personal business in long term career

Converting this hypothesis in terms of odds ratio:

H_0: odds ratio=1 where odds of ppl without graduation/diploma getting personal business = odds of people with graduation/diploma getting into personal business

H_1 : odds ratio<1 where odds of ppl without graduation/diploma a getting personal business < odds of people with graduation/diploma getting into personal business

p-value = 0.6406 ; 95 % confidence interval: [0.00000, 11.56691]
sample estimates: odds ratio = 0

Hence, we cannot reject the null hypothesis. So, there is no difference between people with diploma/ graduation and lower education level towards starting personal business.

4.2.2.9 Age 20-25 and personal business as long-term goal

H_0: the people of age 20-25 choose to go for personal business in long term career.

Translating it to odds ratio, we get:

H_0: odds ratio=1 where odds of age group 25+ getting personal business = odds of people in age group 20-25 getting into personal business

H_1 : odds ratio<1 where odds of age group 25+ getting personal business < odds of people in age group 20-25 getting into personal business

p-value = 0.5358 ; 95 % confidence interval: [0.000000, 3.723659]
sample estimates: odds ratio = 0.7328958

Hence, we cannot reject the null hypothesis. So, there is no difference between people of younger age and other age group towards starting personal business

4.2.2.10 Age 20-25 and personal business as long-term goal

H_0: the people of age 26-30 choose to go for personal business in long term career
Translating it to odds ratio, we get:

H_0: odds ratio=1 where odds of age group 30+ getting personal business = odds of people in age group 26-30 getting into personal business

H_1 : odds ratio<1 where odds of age group 30+ getting personal business < odds of people in age group 26-30 getting into personal business

p-value = 0.6235 ; 95 % confidence interval: [0.000000, 4.259902]
sample estimates: odds ratio = 0.9311232

Hence, we cannot reject the null hypothesis. So, there is no difference between people of younger age and other age group towards starting personal business

4.2.2.11 10th education level and best segment

p-value = 0.798 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.2.2.12 12th education level and best segment

p-value = 0.6186 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.2.2.13 Graduate education level and best segment

p-value = 1 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.2.2.14 Primary school education and best segment

p-value = 0.0743; alternative hypothesis: two.sided
Conclusion: Strong association

4.2.2.15 High school education and best segment

p-value = 0.4833 ; alternative hypothesis: two.sided

Conclusion: No association

4.2.2 Discussions

We find a number of mutually unassociated features. For instance, we find that education level and long-term career plans were not strongly associated. What is particularly striking is that there is no appreciable association between the desire to have a personal business and the highest education level of the cab drivers. This indicates presences of an entrepreneurial streak within the cab drivers. Further, it is interesting to observe that this desire to build a

personal business is not merely limited to the young cab drivers, but pervades across age groups. However, to convert this desire to reality, a lot of capital is needed and perhaps, the drivers are trying to accumulate enough capital to make that switch.

Another observation is that related to monthly income. Usually, it is believed that the younger drivers will tend to drive more and tend to have a higher monthly income. Similarly, cab drivers who are better educated might be able to negotiate better fare as well as identify routes with higher revenue per trip. However, our analysis reveals this conventional wisdom to be flawed. In fact, this can be understood from the fact that the act of driving is not helped by the degree and it is the experience and knowledge of ground realities that counts. Hence, perhaps the older cab drivers get an edge by virtue of their experience. Similarly, the less educated cab drivers might work extra hard and pick up skills that help them perform as well as the educated drivers.

The FAMD analysis shows that we can explain $\sim \frac{1}{8}^{th}$ of the total variation by the first 6 components. Further, we see that out of all the available features, only a few contribute to the two components. The contribution of vendor name and client name is highest for the first two dimensions. Hence, they are important discriminants of variability in the data.

4.3 Security Guards Dataset

The data comes from a survey of security guards that captures the work and employment experiences of different segments of the security guards in Mumbai.

4.3.1 Factor Analysis of Mixed Data

Dimension	Eigenvalue	% Variance	% Cumulative Variance
1	33.144886	8.100913	8.100913
2	16.392610	4.006504	12.107417
3	9.902826	2.420341	14.527758
4	8.318785	2.033187	16.560945
5	7.614687	1.861099	18.422044

Table 4.2: Eigenvalues Obtained From FAMD Analysis of Security Guards Data
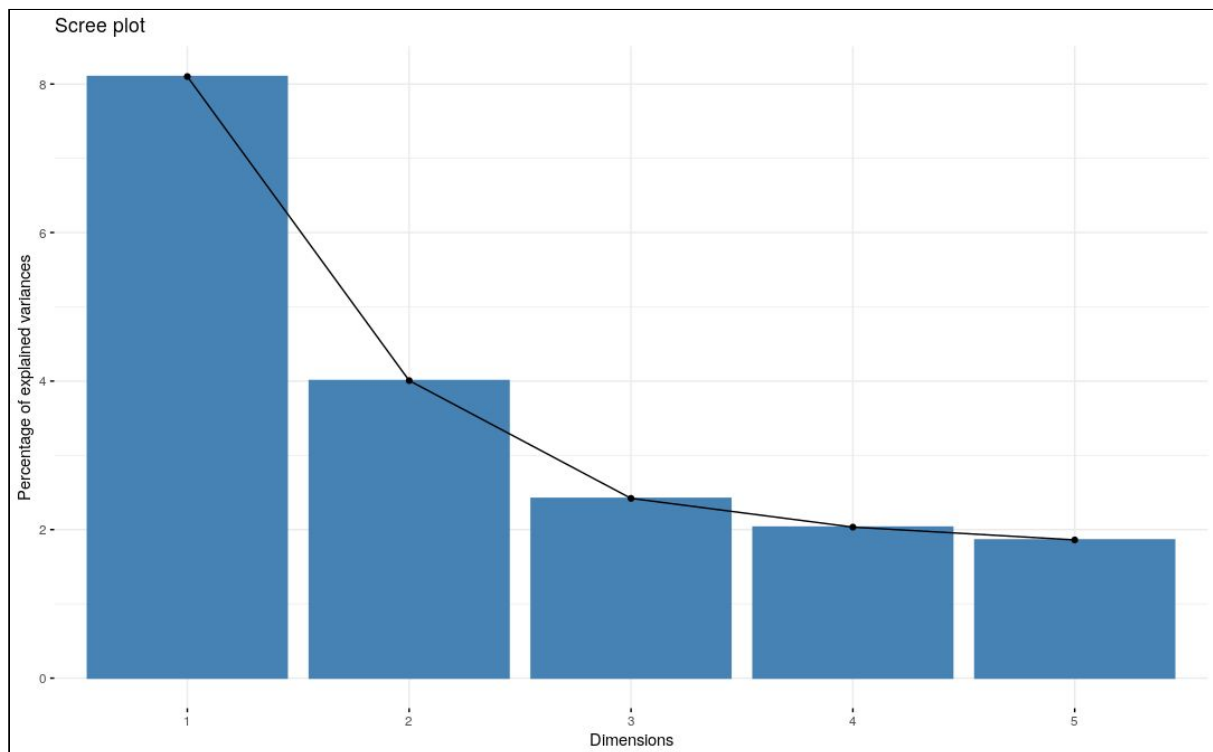

Fig 4.6: Scree Plot of Eigenvalues from FAMD

Fig 4.7: Contribution of Features to the First Dimension

Fig 4.8: Contribution of Features to the Second Dimension
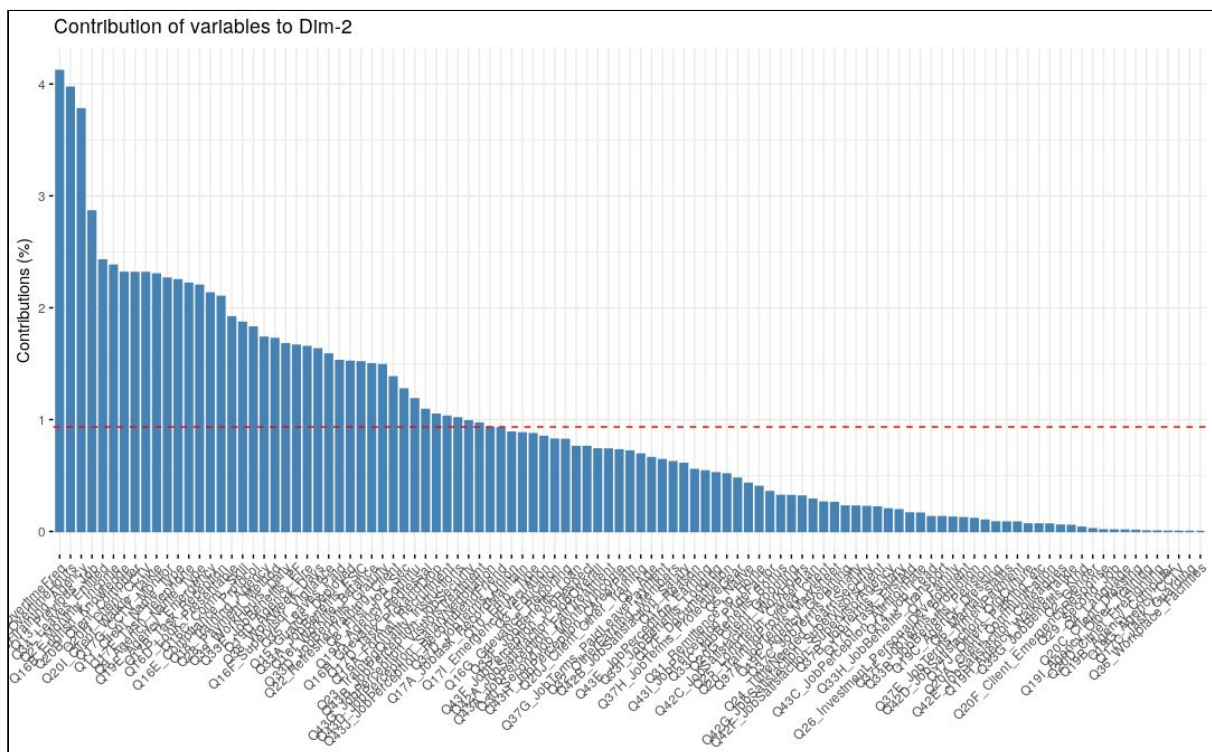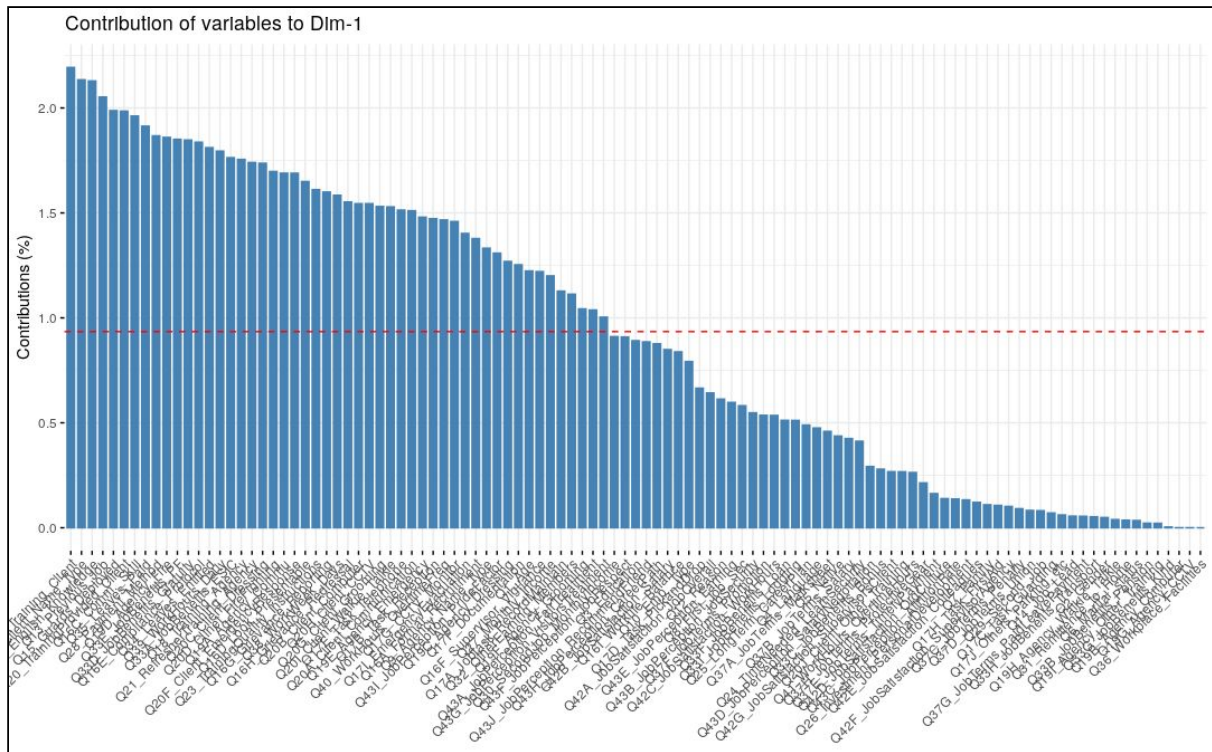
4.3.2 Association Analysis Using Fisher's Exact Test

4.3.2.1 Marital status and agency name

p-value = 5.944e-07 ; alternative hypothesis: two.sided

Conclusion: Strong association

4.3.2.2 Education and personal business as long-term plan

H_0: the people with graduation/ diploma holders choose to go for personal business in long term career

Translating it in terms of odds ratio, we get:

H_0: odds ratio=1 where odds of ppl without graduation/diploma getting personal business = odds of people with graduation/diploma getting into personal business

H_1: odds ratio<1 where odds of ppl without graduation/diploma a getting personal business < odds of people with graduation/diploma getting into personal business

p-value = 0.966 ; 95 % confidence interval: [0.0000, 3.8025]
sample estimates: odds ratio = 1.8985

Hence, we cannot reject null hypothesis and there is no difference between diploma and non-diploma holders towards starting personal business.

4.3.2.3 Age 26-30 and personal business as long-term plan

p-value = 0.5343 ; alternative hypothesis: true odds ratio is less than 1

95 percent confidence interval:[0.000000, 1.883546]

sample estimates: odds ratio = 0.950931

Hence, there is no difference between people of younger age and other age group towards starting personal business.

4.3.2.4 Job perception (fair treatment) and education (1-5)

p-value = 0.01641 ; alternative hypothesis: two.sided
Conclusion: Strong association

4.3.2.5 Job satisfaction (salary) and education (6-9)

p-value = 0.2119 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.3.2.6 Job satisfaction (salary) and education (1-5)

p-value = 0.1729 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.3.2.7 Job satisfaction (salary) and education (Graduate/diploma)

p-value = 1 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.3.3 Discussions

We find a strong association between the marital status of the security guards and the agency. This is probably a result of agency practices, that make one agency more suitable for guards with a particular marital status. Again, just like cab-drivers, we find no association between age and long-term plans of personal business. Similarly, we find no association between education and long-term plans of personal business. Interestingly, when it comes to job perception in terms of fair treatment, we find it to be strongly associated with the level of basic education, but not associated with higher levels of education. However, this association between job perception does not translate to association between job satisfaction in terms of salary. Hence, we can conclude that although salary might be a major factor for security guards, it is not the only factor that drives their perception of being fairly treated.

In contrast with the cab-drivers, the security guard data yields an even lesser explained variability using FAMD analysis. However, it should be noted here that the variability in case of the security guards is not arising due to only a small number of features, but is being contributed by a lot of features. Thus, the datasets might appear to be similar on surface, but are quite different as we dig deeper.

4.4 Housekeeping Staff Dataset

This survey captures information related to housekeeping staff. These workers mostly work for international ICT-ITES firms in Mumbai.

4.4.1 Factor Analysis of Mixed Data

Dimension	Eigenvalue	% Variance	% Cumulative Variance
1	10.64417	8.724731	8.724731
2	8.798084	7.211544	15.93628
3	7.528512	6.170912	22.10719
4	6.53932	5.360099	27.46729
5	5.875909	4.816319	32.28361

Table 4.3 Eigenvalues Obtained from FAMD Analysis
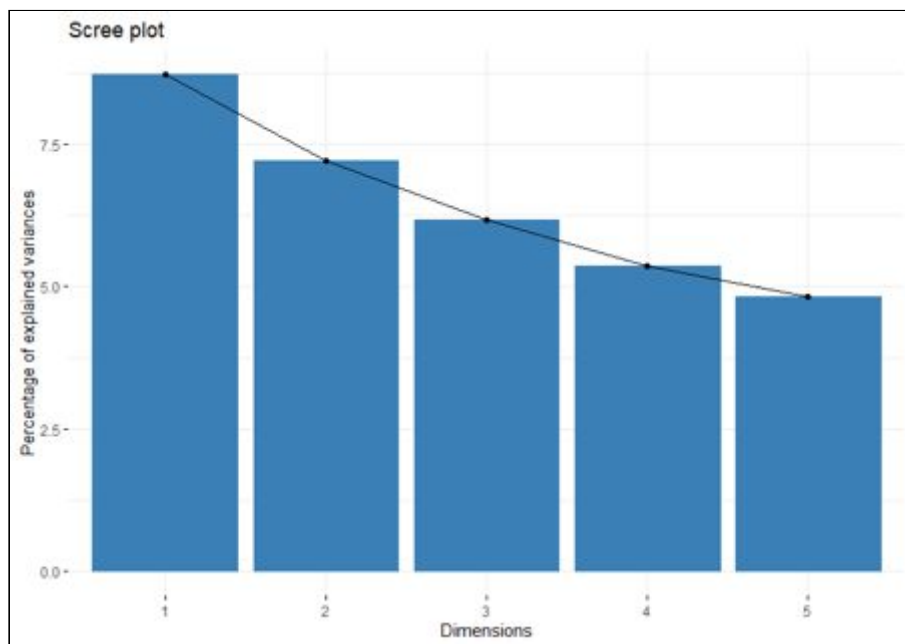
Fig 4.9: Scree Plot of Eigenvalues from FAMD Analysis

Fig 4.10: Illustration of Features Forming the First Two Dimensions
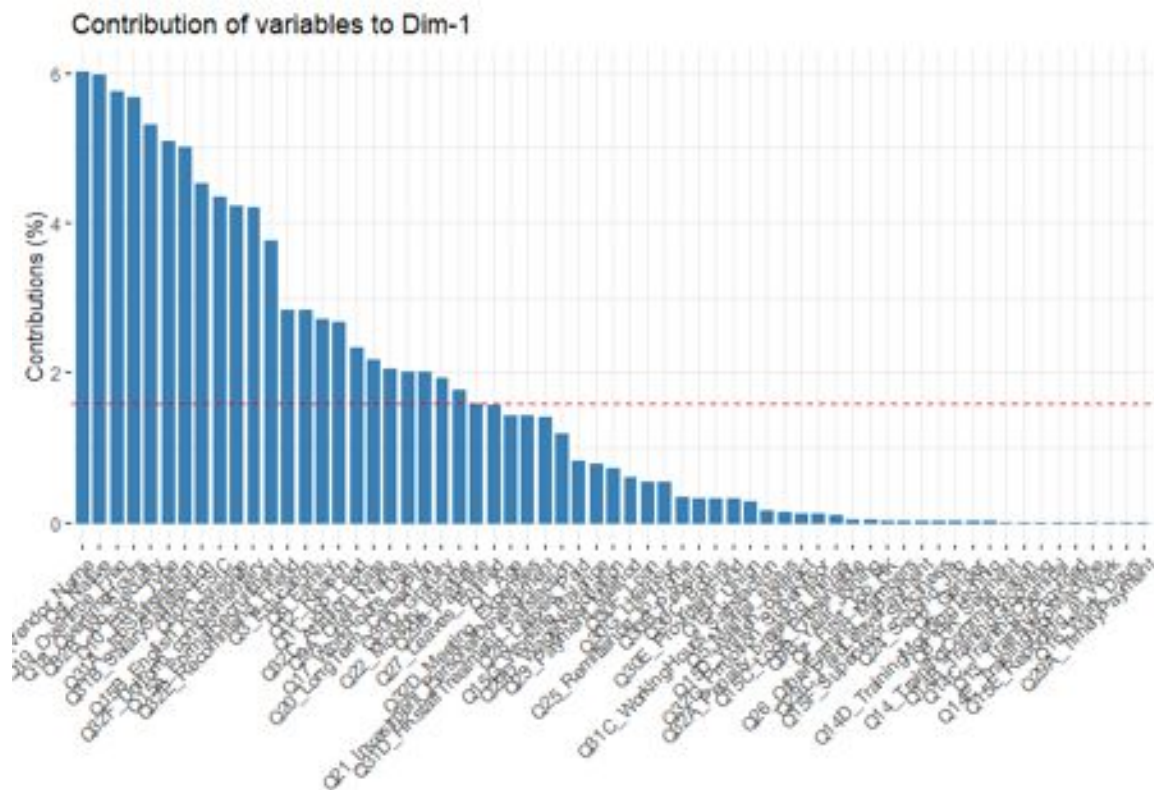
Fig 4.11: Contribution of Features to the First Dimension
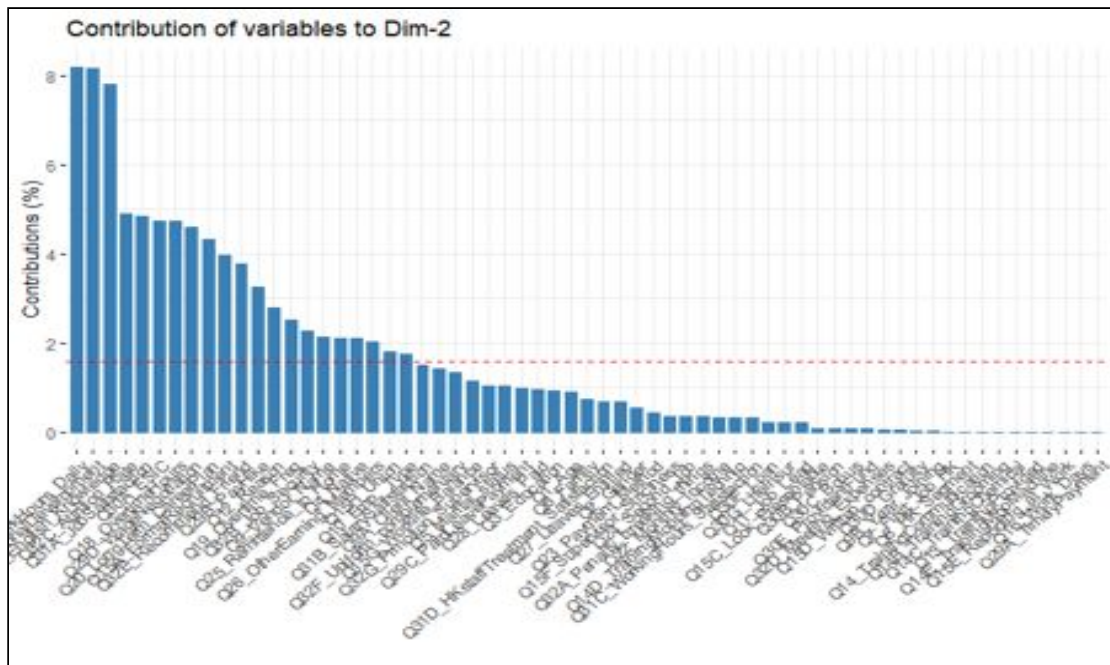
Fig 4.12: Contribution of Features to the Second Dimension
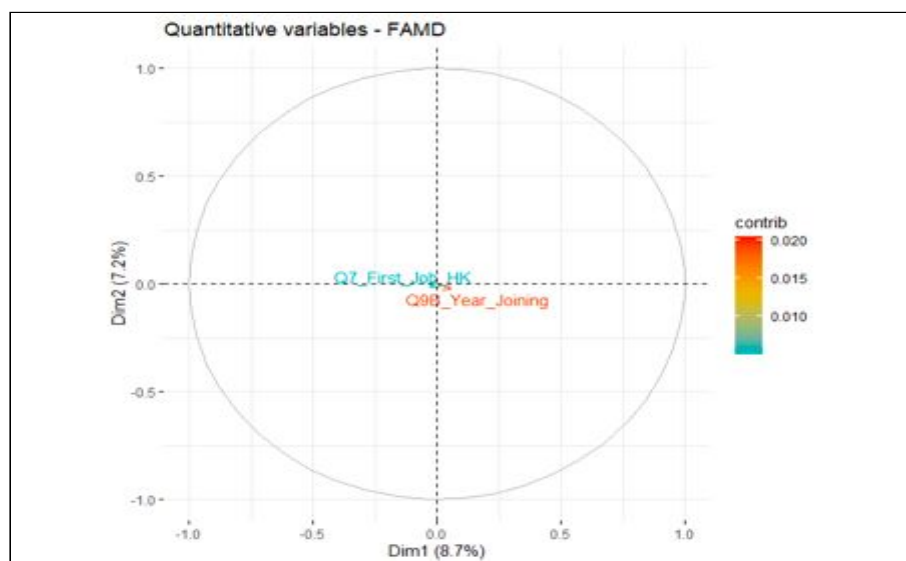
Fig 4.13: Contribution of Quantitative Features to the First Two Dimensions

4.4.2 Association Analysis Using Fisher's Exact Test

4.4.2.1 Different education levels vs long term career plans (Q3 vs Q20)

```
          Not sure about it Prepraing for better career elsewhere Remain in HK sector
10th pass         4                          4                                0
12th pass         1                          2                                0
Stand. 1-5        9                          2                                2
Stand. 6-9       13                          1                                0

          Will switch to other sector Will switch to personal business
10th pass              0                             0
12th pass              0                             0
Stand. 1-5            11                             6
Stand. 6-9             6                             3
```

p-value = 0.01118 ; alternative hypothesis: two.sided

Conclusion: strong association

4.4.2.2 Education and monthly income (Q3 vs Q22)

```
            5001-7000 7001-10000
10th pass        5          3
12th pass        2          1
Stand. 1-5      27          3
Stand. 6-9      17          6
```

p-value = 0.1413 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.4.2.3 Education and personal development (Q3 vs Q21)

```
            I plan to do so in future No Yes
10th pass                          0  3   5
12th pass                          0  0   3
Stand. 1-5                         2 21   7
Stand. 6-9                         1 10  12
```

p-value = 0.05171 ; alternative hypothesis: two.sided

Conclusion: Weakly associated

4.4.2.4 Age and long-term career plan (Q1 vs Q20)

```
          Not sure about it Preparing for better career elsewhere Remain in HK sector
20-25              5                            2                         0
26-30              6                            3                         0
31-35              4                            1                         0
36-40              9                            0                         0
41-45              0                            0                         1
46-50              1                            0                         0
Above 50           0                            0                         1
Below 20           2                            3                         0

          Will switch to other sector Will switch to personal business
20-25              9                            0
26-30              5                            5
31-35              1                            2
36-40              1                            1
41-45              0                            0
46-50              0                            1
Above 50           0                            0
Below 20           1                            0
```

Too small sample size

4.4.2.5 Age and monthly income (Q1 vs Q20)

```
          5001-7000 7001-10000
20-25         10          6
26-30         16          3
31-35          7          1
36-40         10          1
41-45          1          0
46-50          1          1
Above 50       0          1
Below 20       6          0
```

p-value = 0.1483 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.4.2.6 Personal development and age (Q1 vs Q21)

```
                I plan to do so in future No Yes
20-25                                    0  9   7
26-30                                    1 10   8
31-35                                    2  3   3
36-40                                    0  6   5
41-45                                    0  0   1
46-50                                    0  1   1
Above 50                                 0  1   0
Below 20                                 0  4   2
```

p-value = 0.8118 ; alternative hypothesis: two.sided

Conclusion: Not associated

4.4.2.7 Marital status and Long-term career plan (Q5 vs Q20)

```
        Not sure about it Prepraing for better career elsewhere Remain in HK sector
Married          23                                  1                       2
Single            4                                  8                       0

        Will switch to other sector Will switch to personal business
Married              8                                8
Single               9                                1
```

p-value = 9.354e-05 ; alternative hypothesis: two.sided

Conclusion: Associated

4.4.2.8 Long-term career plan and job satisfaction (Q20 vs Q31A)

```
                                    Not sure about it Prepraing for better career elsewhere
niether satisfied nor dis-satisfied        14                                  1
no                                         10                                  6
yes                                         3                                  2

                                    Remain in HK sector Will switch to other sector
niether satisfied nor dis-satisfied          2                       2
no                                           0                      14
yes                                          0                       1

                                    Will switch to personal business
niether satisfied nor dis-satisfied              3
no                                               5
yes                                              1
```

p-value = 0.0242 ; alternative hypothesis: two.sided

Conclusion: Strong association

4.4.3 Discussions

The housekeeping staff data indicates that there is a strong association between education levels and long-term career plans. This is different from earlier analysis for cab drivers.

Similarly, there is a weak association between education and personal development. This is expected in light of the fact that education is guiding their long-term plans. It is interesting to observe that this is not limited to a particular age, since we do not find substantial association between age and long-term career plans or even age and personal development. Thus, housekeeping staff appears to be aspirational across age groups. The importance of long-term career plans and education can be gauged from the fact that it is the primary driver of job satisfaction here.

The FAMD analysis explains ~ 32% of the variability in the data. Just like security guards data, however not as accentuated, we find that a number of features are driving the variability and contributing to the two principal dimensions.

5. CONCLUSIONS

In this work we have broadly analyzed data collected via three surveys, viz., cab drivers, housekeeping staff, and security guards. To achieve this purpose, we have used specific statistical tools and techniques. The multivariate factor analysis of mixed data (categorical and continuous) has given us enough insights as to which covariates of questionnaire have more information/variance. The next step of analysis chosen by us was to test multiple hypotheses from the data; consider for example the association between educational level and job satisfaction. The analysis helped us to refute multiple claims/biases that one may have. For instance, one might argue that younger workers might want to do something different in future leaving their current housekeeping job but the questionnaire data did not provide substantial evidence in favour of the claim. Multiple such hypotheses were analyzed throughout the report. However, there is ample scope for further enriching the insights by performing a deeper analysis.

Although we have analyzed the three surveys separately and validated multiple hypotheses, yet a possible extension of work can be to compare various statistical measures across the three datasets. In fact, the behaviour exhibited by the covariates under multivariate factor analysis of mixed data indicates that there is substantial scope of finding similarities and differences. For this purpose, the data can be enriched through the use of external datasets, leading to generation of new interesting finding of general interest.

REFERENCES

[1] Donato, Nicola. Decent work and the 2030 Agenda: a path towards a sustainable development. BS thesis. Università Ca'Foscari Venezia, 2017.

[2] Anand, Ishan. "Book review: International Labour Organisation, India Wage Report: Wage Policies for Decent Work and Inclusive Growth." (2019): 117-119.

[3] Belser, P. and Rani, U., 2011. Extending the coverage of minimum wages in India: Simulations from household data. Economic and Political Weekly, pp.47-55.

[4] Ghatak, M., and P. Ghosh. "The Land Acquisition Bill: A Critique and a Proposal: Making a Show: The Black Money Bill." Economic and Political Weekly 50 (2015): 7-8.

[5] "Report of the Working Group on Social Security" https://niti.gov.in/planningcommission.gov.in/docs/aboutus/committee/wrkgrp11/wg11_rpsoc.pdf

[6] Deakin, Simon, and Prabirjit Sarkar. Indian labour law and its impact on unemployment, 1970-2006: a leximetric study. Cambridge: Centre for Business Research, University of Cambridge, 2011.