

A Study on Health Conditions among Retired Travelling Ticket Examiners of South Coast Railway

K.V. Sobha Rani¹, K.V.Subba Rao²

¹Lecturer in Computer Applications, Pithapur Rajah's Government College (Autonomous), Kakinada 533001, Andhra Pradesh, India

²Retired Travelling Ticket Examiner, South Coast Railway, India

Abstract

The health and well-being of retired railway employees are important concerns due to the increasing prevalence of chronic diseases and changing healthcare requirements with advancing age. The present study examines the health conditions and healthcare-utilization patterns among retired Travelling Ticket Examiners (TTEs) of South Coast Railway, India, using data mining techniques. The study is based on primary data collected from 157 retired TTEs through a structured questionnaire covering demographic characteristics, years of service, years after retirement, lifestyle practices, health conditions, chronic diseases, hospital utilization, healthcare expenditure, railway medical facility usage, overall health status, and health satisfaction. Descriptive statistical analysis was used to examine the distribution of major health conditions and healthcare-related characteristics. In addition, Apriori association rule mining was employed to identify co-occurrence patterns among health conditions using support and confidence. K-Means clustering was used to identify groups of respondents with similar health and healthcare-utilization characteristics.

The findings indicate that diabetes, high blood pressure, and eye problems were among the frequently reported health conditions. Approximately 65% of the respondents reported their overall health status as good, while a substantial proportion reported chronic health conditions. The Apriori analysis identified meaningful associations among selected health conditions, particularly between diabetes and high blood pressure. The clustering analysis grouped the respondents into three broad health and healthcare-utilization profiles. The largest cluster represented a moderate healthcare-utilization group, while another cluster represented relatively lower healthcare utilization. The third cluster consisted predominantly of older respondents with a higher reported chronic-disease burden and greater hospital utilization.

Overall, the study demonstrates the usefulness of data mining techniques in identifying health patterns, associations, and distinct healthcare-utilization profiles among retired TTEs. The findings highlight the importance of regular health screening, chronic-disease monitoring and targeted healthcare support, particularly for older retirees with greater healthcare requirements. The study provides a data-driven basis for understanding the health needs of retired railway employees and may support the development of more focused post-retirement healthcare interventions.

Keywords: Retired TTEs, South Coast Railway, Health Conditions, Data Mining, Association Rules, Clustering, Chronic Diseases.

1. INTRODUCTION

Retirement marks an important period in one's life when health assumes importance as a determinant of life quality and well-being. Retired employees can face age-associated health problems, suffer from chronic illnesses, incur increasing expenses on healthcare and require more medical assistance. These aspects need careful consideration, especially with regard to occupations that have been marked by difficult working conditions while performing one's active service.

Travelling Ticket Examiners (TTEs) are an important category of railway employees who perform duties related to travel and ticket examinations. Working life of TTEs can include travel, shift work, long working hours and occupational stress. After retirement, combined effect of aging, personal habits and occupational experiences can affect the physical condition of employees and their healthcare needs.

Traditional statistical analysis is important to describe the situation concerning the state of health. However, methods of data mining allow for gaining new information through discovery of hidden connections and clustering of respondents with common health profile. Methods of association rule mining will help in identification of common disease and health attributes combinations.

2. OBJECTIVES OF THE STUDY

The objectives of this study include

- The investigation of demographic and service-related attributes of retired TTEs from South Coast Railway.
- The determination of the prevalence of major diseases and conditions among retired TTEs.
- The analysis of their lifestyle attributes healthcare usage and health expenditures.
- The establishment of associations between health conditions and certain lifestyle variables through association rule mining.
- The clustering of retired TTEs into health profiles using clustering technique.
- Making recommendations on how to enhance post-retirement healthcare of retired TTEs.

3. RESEARCH METHODOLOGY

In the current study, the primary source of data was mainly the primary data obtained directly from the retired Travelling Ticket Examiners. The primary data was collected through the use of a structured questionnaire covering the following areas: age, gender, years of service, years of retirement, lifestyle, illness, hospital visits, health expenditure, use of railway medical facilities, general health condition, and satisfaction level.

3.1 Dataset

The dataset consists of 157 respondent records and 28 columns. The dataset contains demographic, occupational, lifestyle, and health-related information of retired TTEs. Table 1 shows the column names and description of each column.

Table 1 : Dataset columns and description

Sl.No	Column Name	Description
1	Respondent_Name	Name of the Respondent
2	Age	Current age
3	Gender	Male/Female
4	Years of Service	Total years worked in Railways
5	Years_After_Retirement	Number of years since retirement
6	Smoking_Habit	Yes/No
7	Alcohol_Consumption	Yes/No
8	Physical_Exercise	Daily/Weekly/Occasionally/Never
9	Blood_Pressure_Status	Normal/High
10	Diabetes	Yes/No
11	Heart_Disease	Yes/No
12	Arthritis_or_Joint_Pain	Yes/No
13	Sleep_Disorders	Yes/No
14	Stress_Level	Low/Moderate/High
15	Number_of_Chronic_Diseases	Total number of chronic diseases the respondent has
16	Kidney_Problem	Yes/No
17	Back_Pain	Yes/No
18	Eye_Problem	Yes/No
19	Gastric_Problem	Yes/No
20	Cancer	Yes/No
21	Brain_Stroke	Yes/No
22	Paralysis	Yes/No
23	Dietary_Pattern	Vegetarian/Non-Vegetarian
24	Hospital_Visits_Per_Year	Number of times visited hospitals
25	Health_Expenditure_Per_Month	Money spent on health per month
26	Railway_Medical_Facility_Usage	Frequently/Occasionally/Never
27	Overall_Health_Status	Excellent/Good/Average/Poor
28	Health_satisfaction	Dissatisfied/ Neutral(Average Satisfaction)/ Satisfied/ Highly Satisfied

3.2 Statistical and Data Mining Techniques

Descriptive Statistics methods such as Mean, Standard Deviation, Minimum, and Maximum were used for quantitative data. Percentage analysis was carried out on lifestyle factors, diseases, health status, and healthcare utilization. The association rules were analysed based on support, confidence. K-Means clustering method was used to classify respondents based on variables like age, years of service, years since retirement, chronic diseases, visits to hospitals, and healthcare expenses.

4. RESULTS AND FINDINGS

4.1 Descriptive Statistics

The mean age of the respondents was 69.72 years. The mean service period was 33.87 years, whereas

the mean period after retirement was 9.92 years. The mean number of chronic diseases reported by the respondents was approximately 0.97. The mean number of hospital visits per year was 3.66 visits. The mean monthly health care expenditure was approximately ₹2,159.24. The above results show that the research population comprised mainly of elderly retirees with extensive service periods and different healthcare needs.

4.2 Lifestyle Characteristics

The data reveal that

- 153 respondents (97.5%) were not habitual smokers.
- 144 respondents (91.7%) were not habitual drinkers
- 108 respondents (68.8%) did physical exercises daily
- 23 respondents (14.6%) never did physical exercises
- 112 respondents (71.3%) had a non-vegetarian diet
- 45 respondents (28.7%) had a vegetarian diet

High proportions reporting daily physical exercise along with low proportions of habitual smoking and drinking indicate favourable lifestyle characteristics.

4.3 Prevalence of Major Health Conditions

Table 2 shows the major health conditions with percentages. Diabetes emerged as the most frequently reported health condition, followed by eye problems and high blood pressure. These findings suggest the importance of regular screening and long-term monitoring of chronic and age-related health conditions.

Table 2: Major health conditions with percentages

Health Condition	Number of Respondents	Percentage
Diabetes	70	44.6
High Blood Pressure	46	29.3
Eye Problems	47	29.9
Heart Disease	29	18.5
Arthritis/Joint Pain	26	16.6
Gastric Problems	25	15.9
Sleep Disorders	20	12.7
Back Pain	11	7.0
Kidney Problems	5	3.2
Paralysis	4	2.5
Brain Stroke	2	1.3
Cancer	1	0.6

4.4. Association Rule Mining

The association-rule analysis identifies several important co-occurrence patterns among the health conditions of retired TTEs. Based on the analysis performed on 157-record TTE dataset, some of the generated association rules are shown in Table 3.

Table 3: Association Rules for Major Health Conditions among Retired TTEs

Association Rule	Support (%)	Confidence (%)
Rule 1: Gastric Problems → Eye Problems	12	72
Rule 2: High Blood Pressure → Eye Problems	15	50
Rule 3: High Blood Pressure → Diabetes	20	68

Rule 1: Gastric Problems → Eye Problems

This rule indicates that 12% of all respondents reported both gastric problems and eye problems. The confidence value of 72% indicates that among respondents who reported gastric problems, 72% also reported eye problems.

Rule 2: High Blood Pressure → Eye Problems

The rule High Blood Pressure → Eye Problems produced a support of 15% and a confidence of 50%. Thus, 15% of the retired TTE respondents reported both conditions, while half of the respondents reporting high blood pressure also reported eye problems. This pattern suggests that eye health may warrant attention during routine health monitoring of retirees with hypertension.

Rule 3: High Blood Pressure → Diabetes

The rule High Blood Pressure → Diabetes showed a support of 20% and a confidence of 68%. This indicates that 20% of the total respondents reported both high blood pressure and diabetes. Furthermore, among respondents with high blood pressure, 68% also reported diabetes. This represents an important co-occurrence pattern and suggests the value of integrated screening and monitoring of hypertension and diabetes among retired railway employees.

Table 4: Number of Association Rules at different minimum support thresholds

Support (%)	Number of Rules
5	24
10	9
15	5
20	3
25	1

Using the thresholds minimum support = 10% and minimum confidence = 50%, the analysis gives 9 qualifying association rules. Table 4 and Figure 1 show the number of association rules with minimum confidence of 50% at different support thresholds.

Figure 1: Number of Association Rules at different minimum support thresholds

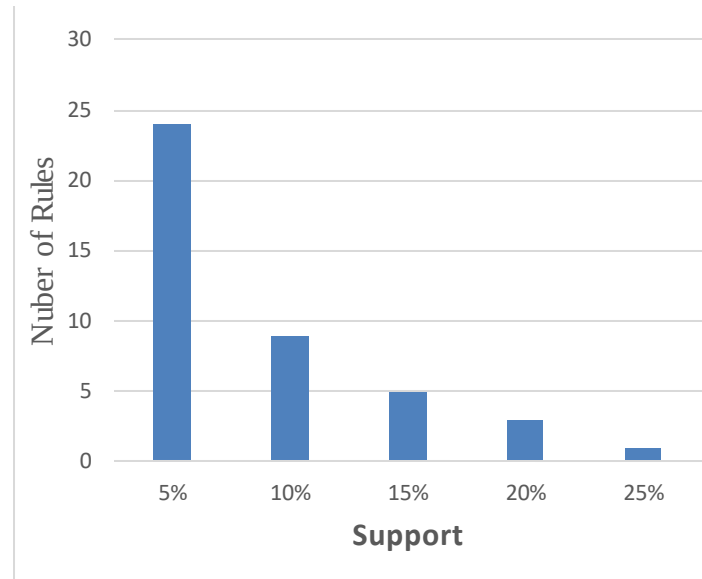
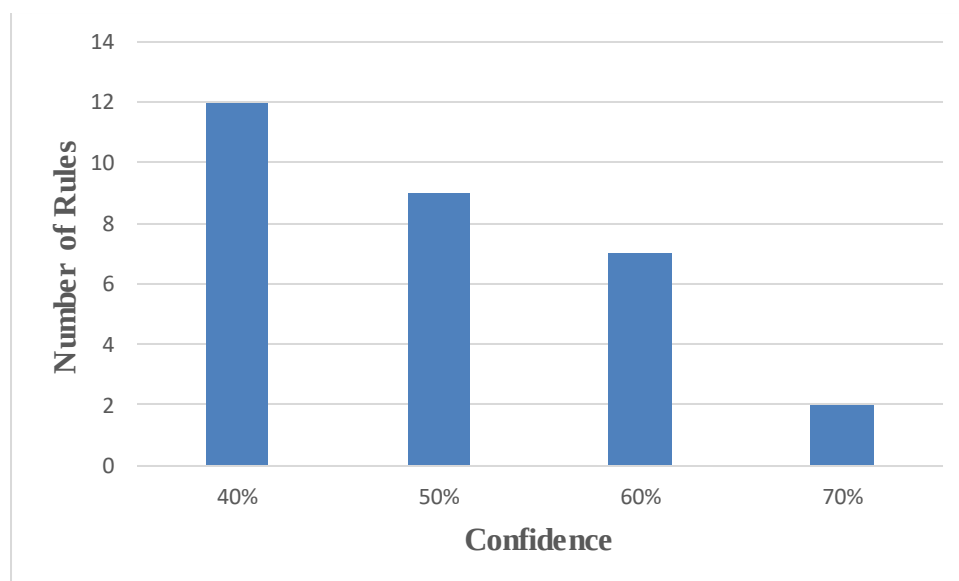


Table 5 and Figure 2 show the number of association rules with minimum support of 10% at different support thresholds.

Table 5: Number of Association Rules at different minimum support thresholds

Confidence (%)	Number of Rules
40	12
50	9
60	7
70	2

Figure 2: Number of Association Rules at different minimum confidence thresholds



4.5. Clustering Analysis

K-Means clustering was applied using age, years of service, years after retirement, number of chronic di-

seases, hospital visits, and monthly health expenditure. The respondents were grouped into three clusters. The main objective of the clustering analysis was to identify homogeneous groups of retired TTEs with similar health-related and healthcare-utilization characteristics.

The clustering analysis specifically attempted to identify groups based on

- Age
- Years of service as TTE
- Years after retirement
- Number of chronic diseases
- Number of hospital visits per year
- Monthly health expenditure

Thus, instead of examining each variable independently, K-Means clustering provides an overall profile of groups of respondents having similar characteristics.

4.5.1. Results of K-Means Clustering

The analysis produced three distinct clusters. Table 6 shows mean values of each cluster as shown below

Table 6: Mean Values of Each Cluster

Cluster	Number of TTEs	Mean Age	Mean Years of service	Mean Years after Retirement	Mean Chronic Diseases	Mean Hospital Visits per Year	Mean Monthly Health Expenditure
Cluster 1	36(22.9%)	67.03	32.25	7.56	0.31	2.00	₹16.67
Cluster 2	84(53.5%)	66.86	35.44	6.92	0.83	3.04	₹2,928.57
Cluster 3	37(23.6%)	78.84	31.86	19.05	1.95	6.68	₹2,497.30

Cluster 2 represents the largest group, containing approximately 53.5% of the respondents. Table 7 shows comparison of three clusters

Table 7: Comparison of three clusters

Cluster	General Profile	Key Characteristics
Cluster 1	Lower healthcare utilization	Fewer chronic diseases and fewer hospital visits
Cluster 2	Moderate healthcare utilization	Largest group; intermediate chronic disease burden and hospital utilization
Cluster 3	Higher healthcare utilization	Highest age, chronic diseases and hospital visits

The most notable finding is that the third cluster consists of older retirees with more chronic diseases and hospital visits.

5. CONCLUSION

The study concludes that retired TTEs of South Central Railway exhibit different health and healthcare-utilization patterns. Data mining techniques identified important associations among health conditions and three distinct health profiles through K-Means clustering. The findings particularly highlight the need for regular health screening, chronic-disease monitoring and targeted support for older retirees with

greater healthcare needs. Overall, the study demonstrates that data mining can be effectively used to understand health patterns and support better healthcare planning for retired railway employees.

6. REFERENCES

1. Agrawal, R., & Srikant, R. Fast algorithms for mining association rules. Proceedings of the 20th International Conference on Very Large Data Bases, 487–499.
2. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
3. Han, J., Kamber, M., & Pei, J. *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
4. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.
5. World Health Organization. *World Report on Ageing and Health*. Geneva: WHO
6. United Nations. *World Population Ageing*. United Nations Department of Economic and Social Affairs.