-

**A MINI PROJECT REPORT**

**ON**

**“**Employee Identification for Health Insurance**”**

**BY**

Avesh Shaikh B12053

Puja Tapadiya B12060

Kirtikumar Waykos B12066



**Department of Computer Engineering**

**ZEAL College of Engineering and Research, SAVITRIBAI PHULE PUNE UNIVERSITY**

**2018-2019**

**CONTENTS**

# TITLE

# 

1. **INTRODUCTION**

**2 PROBLEM STATEMENT**

**3 OBJECTIVE**

**4 PROPOSED SYSTEM**

**5**  **CONCLUSION**

**INTRODUCTION**

In today’s competitive environment, understanding the customers better, especially, most customer groups and the groups that have the biggest potential is the biggest challenge. By K-means customers based on their behavior, we can better target their actions, such as launching tailored products, target one-to-one marketing and to meet the customer expectations. However, the problem often is that the data regarding customer behavior is available in several different sources and analyzing the large data set is exhaustive and time consuming.

Elaborates the use of clustering to segment customer profiles of a retail store. The study concluded that the K-Means clustering allows retailers to increase customer understanding and make knowledge-driven decisions in order to provide personalized and efficient customer service. Using the insurance company’s claims database, Customer extracts three attributes for 547 randomly selected individuals. The three attributes are the insured’s weight in Pounds as recorded on the person’s most recent medical examination, their last cholesterol level determined by blood work in their doctor’s lab, and their gender.

As is typical in many data sets, the gender attribute uses 0 to indicate Female and 1 to indicate Male. We will use this sample data from customer’s database to build a cluster model with the help of k-means the health insurance policy holders, appear to group together on the basis of their weights, genders and cholesterol levels.

**PROBLEM STATEMENT**

There is a correlation between weight, gender and cholesterol which leads to coronary heart disease. Also along with heart there disease it also affect the overall performance of individual. Hence to eliminate, an organization identify the groups of individuals who are at high risk and offer and conduct weight and cholesterol management programs to individuals who receives health insurance through organization.

**OBJECTIVE**

To search natural group of individuals who are most at risk for high weight and high cholesterol and if there are such groups where the natural dividing lines between the groups occurs.

**PROPOSED SYSTEM**

Proposed model was a model for clustering time series of clinical markers obtained from routine visits in order to identify homogeneous customers subgroups. In this system we will look at how to predict defaulting, machine learning. We also look at how to summarize the data using classifier technique. The proposed approach used the K-means clustering with statistical analysis (ANOVA) for general selection and SVM to classify the diseases.

Data Mining is the process of uncovering patterns inside large sets of structured data to predict future outcomes. Structured data is data that is organized into columns and rows so that it can be accessed and modified efficiently. Using a wide range of machine learning algorithms, you can use data mining approaches for a wide variety of use cases to increase revenues, reduce costs, and avoid risks.  If you are looking to analyze unstructured data (e.g. data from essays, articles).

**RESULT:**

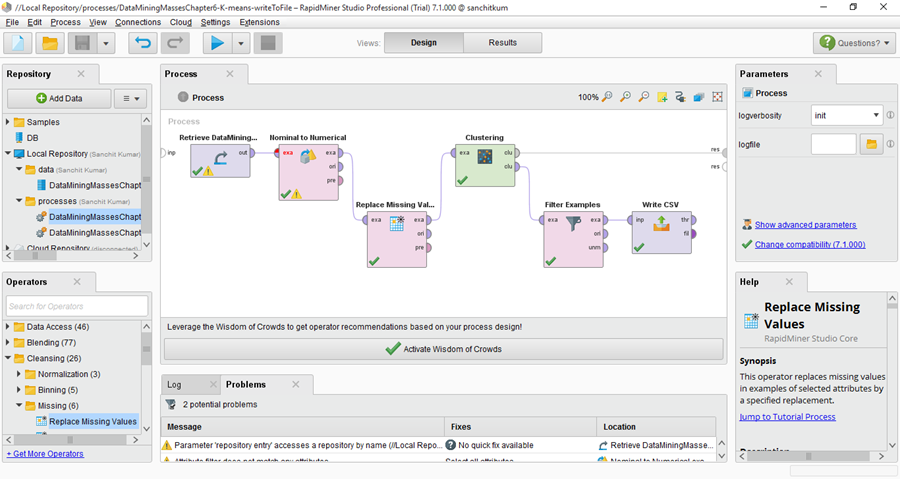


Figure 1. Process View

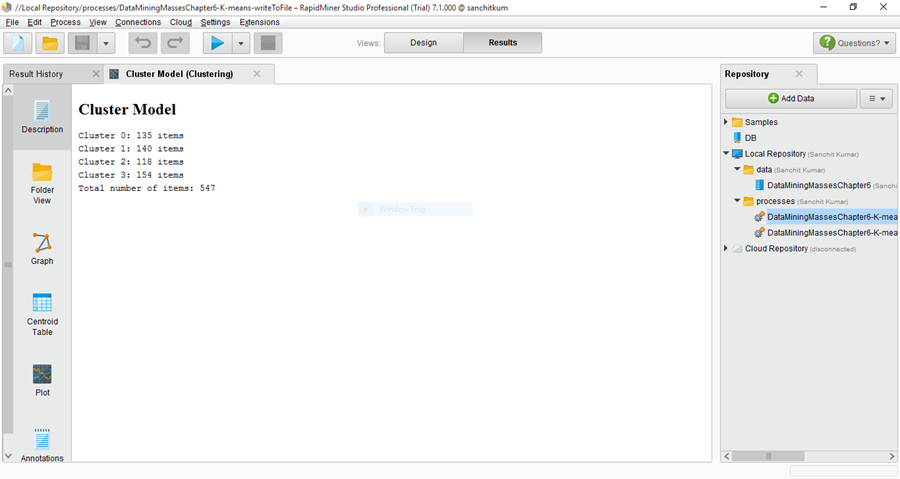


Figure 2. Cluster Model

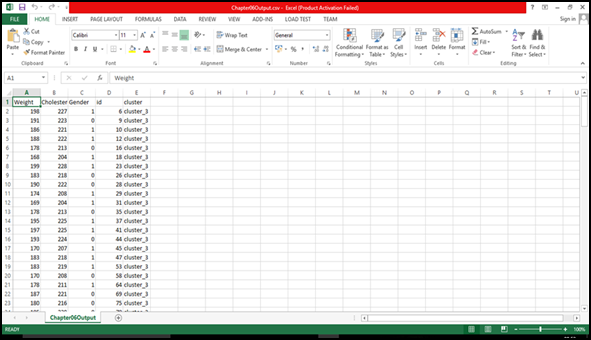


Figure 3. Filtered View of data

**CONCLUSION :**

We have successfully implemented the concept of identification, analysis of Data sets by using Data Mining concepts and tools such as K- mean cluster algorithm, Rapid Miner. Here we also are able to identify and predict the outcome as possible employee who can be eligible and satisfy (threshold value) criteria of considerate health for health Insurance of an Organization.

This work can be refer to identify such individual in health firms or organization.