Discussion of Structural Changes of Medical Cost through Data Mining Techniques – A Case of Certain Regional Hospital

Pei-Chen Sun[1] Hsieh,Chia-Jung [2] 孫培真 謝佳容

Professor, Graduate Institute of Information and Computer Education, National Kaohsiung Normal University

Graduate Institute of Information and Computer Education, National Kaohsiung Normal University
國立中山大學資訊管理博士
現任高雄師範大學教授

國防管理學院正期93年班
現任海軍左營醫院資訊官

Abstract

As the Bureau of National Health Insurance (BNHI) gradually reformed the payment system, to control the individual medical payment policy, the operating strategy had significant influence on survival and development of hospitals. Since traditional descriptive statistics were insufficient in analyzing the causal relationship of fee variations within operating performance, this research compared the trend and variations of medical fee in the same time period through clustering data mining, and the analysis started from the overall level to the fee types in each department, and conduct vertical in-depth explorations on appropriateness of medical resources usage, to provide effective predictions and supervision, and simultaneously avoid from overlooking the overall medical market trend due to being bound in past experiences. It was expected to provide managers in hospital references for operating strategy, planning and developing focuses.

1. Research Background

After launching National Health Insurance (NHI) since May 1995, since lacking of linked mechanism between revenue and fee, the insurance system was in deficit after three years of implementation. Besides, with the factors of increasing medical usage due to aging population, medical demands facilitated by income increases and raising medical fee due to technological development, the revenue of health insurance fee grew with

an average rate of 4.34% from 1995 to 2008, while the medical fee grew with an average rate of 5.34%, the growth of medical fee was even higher than growth of insurance revenue year by year. (Department of Health, Executive Yuan Taiwan, 2009) In order to relief the financial pressure, the BNHI devoted in actions of financial supervision and economizing on fees.

As for medical payment system, the BNHI implemented capitated payment system from July 2002, and cut down the prices of medicine payments in 2009, and starts to promote case payment system since 2010, with expectations to enhance the supervision of violations for certified hospitals and control the growth of medical fee through file analysis and project management.

However, it was impossible for hospital to satisfy all the requirements by patients with controlling the medical service quantity without the full payment by BNHI, under the current growing trend of medical service quantity. With the low fee of NHI system, it was hard for hospitals to maintain quality and reasonable labor terms. In addition, regional hospitals was inferior than medical center in supports, many middle or small-sized hospital or clinics were shut down or integrated in greater-scale hospitals due to economic of scale. Therefore, systematic management, inspecting the variation of medical fee structure and promotion of many implementations of budget economizing to serve as buffer for financial imbalance was a significant issue for on-going management of hospitals.

1.1 Research Motives

Under the transition of NHI system and variation of patients structure, hospitals were no longer more profitable with more visiting by patients. Rather, identifying the cost difference among departments, assisting doctors to plan better clinical pathways, let the resources be collectively integrated in all aspects, to elevate medical quality and reduce operating costs were main issues for hospitals. Hence, the IT personnel in medical industry faced huge challenges, to positively achieve the goal of cost management without conflicting medical morals.

Currently, the performance evaluation and fee supervision in most hospitals adopted the comparison with previous same-period data, total amount of impatient care, or proportions of medicine fees. However it was lacking of consideration on variation of overall fee structure. Some hospitals adopted hospital management support system, to determine strategic goal

through KPI (Key Performance Indices), whereas the determination of strategic goal or KPIs were most by experiences and recognitions by executives. (Chang Huai-Lu, 2004) This research adopted data mining techniques to analyze the hospitalization fee from each department and usage of NHI resources, at the same time discussed the variations between medical fee of hospitalized patients, to provide reference for decision makers of hospital and medical fee management at home and abroad.

2. Literature Review

2.1 Data Mining

Knowledge discovery in database (KDD) (Fayyad, 1996) was defined as to discover the data with most knowledgeable value in database. The process of KDD included the dynamic cycle of data selection, preprocessing, data transformation, data mining, data interpretation and evaluation. Data clustering was a technique of data analysis seperating data into several similar clusters through statistical method, and then automatically identifying related clusters through algorithm according to data distribution. With clustering, the characteristics of each cluster would be popped out and seperated from other clusters, and then the searching and computation in certain cluster could be conducted upon appointed terms.

Jiawei Han and Micheline Kamber (Jiawei Han 2006) categorized five clustering methods: partitioned, hierarchical, density-based, grid-based and neural network clusterings. This research adopted partitioned clustering K-Means algorithm.

2.2 Clustering K-Means Algorithm

K-Means (MacQueen 1967) was introduced by MacQueen in 1967, who was the pioneer of clustering algorithm, proposing the most typical and center-of-gravity-based partitioned clustering algorithm. Before applying this method, the number of clusters should be pre-determined, that is, to define the value of k, it selected the centers of gravity from each cluster as the representative objects, and then distributed all the data point to the closest cluster according to distance data, and redetermined the new cluster center, and then reallocate all the data point based on the new cluster centers. The step above would be repeated until each center was fixed. The advantages of this method were: (1) concept simplicity, (2) great efficiency in greater data sets, (3) greater scalability (Huang 2003; Pham 2004; Chiu Yung-Hsin, 2006) while the disadvantages of K-means algorithm were: (1) the selection of innital clusteirng center would have significant influence on efficiency of this method,

(2) the final number of cluster should be determined by user, (3) the outcome of clustering could be easily altered by noises or outliers and (4) the non-conves clutering could not be identified. (T.S Chen, 2006) This research adopted this method since the number of sample was great and certain efficiency should be maintained.

2.3 Discussion of Strategic Management and Performance Dimension for Hospitals

Miles and Snow(1978) divided operating strategies for hospital into proactive, defensive, analytic and responding types. The research result showed that most middle or small-sized hospitals selected "analytic" style of operating strategy. (Chu Wen-Yang, 2001) The operating strategy had determining influence on survival and development of hospital. As mentioned by Harkey J. and Vraciu R. (1992), after elevation of medical quality, maintaining excellent relationship between doctors and patients and efficiency improvement (such as efficient treatment or shortening days of hospitalization), the hospital may achieve profitability goal (as Figure 2.1). Through complete and detailed process, the operating strategies of hospital could be improved through realizing the regional environment, trend of development of medical industry, changes on health insurance policy, analysis of demands of competitors and customers, and thus elevate production efficiency, seek and develop new opportunities, technology adaptability and reconciliation, to inspect the effectiveness and efficiency of resource devoted. With integration of these strategic dimensions, the performance of hospital could be improved.

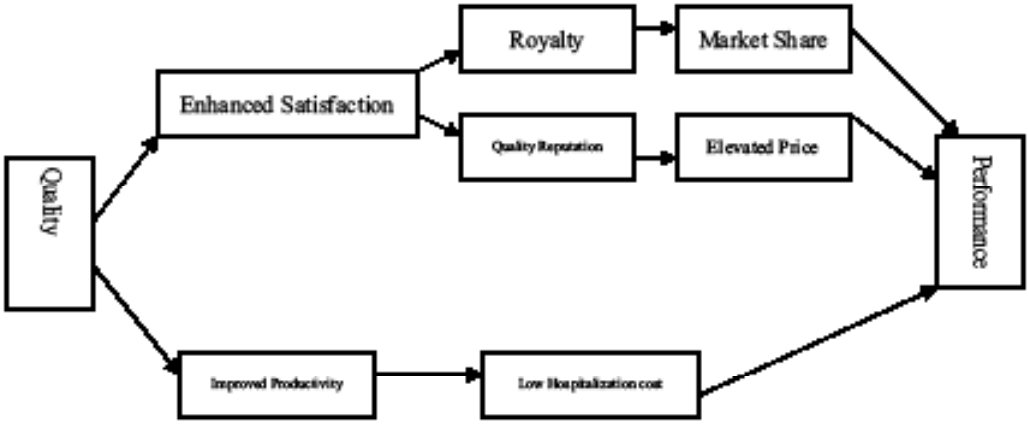


Figure 2-1 Model of Medical Quality--Profitability (Harkey J. 1992)

There were many methods for performance evaluation, the main shortages of conventional financial performance evaluation indices were over-focusing on outcomes and ignoring the process, the statistical analysis lacking of analysis and judgment of artificial intelligence was not helpful for managers to make operating decisions. Besides, it was lacking of predictive ability, sometimes even sacrificed long-term competitive advantages for saving short-term cost. (Li Shu-Hsing, 1995) There were three wildly used methods for evaluating organization performance: ratio analysis, regression analysis and data envelopment analysis. (DEA) In which ratio analysis disclosed the relationships between numerical data from each spreadsheet through ratio calculation and analysis, to reflect the financial situation and operating result of an organization, to provide valuable and rapid financial information for users. Regression analysis integrated all the factors influencing productivity of hospital into one or several regression models, and then calculated the productivity of hospital and its influential factor through R-squared and regression coefficient. DEA compared multiple-inputs and multiple-outputs of distinct units through linear planning, and identify the organizations with greater or poorer efficiency (Chu Wen-Yang, 2001) However in practice, when evaluating performance of hospitals, the financial performance indices adopted were activity, mobile force, return on asset (ROA) and profitability, operating performances indices were bed occupation rate, average hospitalization day, turnover rate of bed, and cost of unit production, and quality performance indices of mortality, infectious rate and satisfaction. (Chen Pei-Ni, 1996)

3. Materials and Methods

The research resources were database of NHI declaration from certain regional hospital in South Taiwan from January 2007 to September 2009, and the research subjects were hospitalization fees of NHI and database of order entry lists. The data of hospitalization fees were categorized through ID number of patients, and then arranged in separated departments into total amount of medical fees, times of patients, physician fee, room charges, tube feeding fee, inspection fee, radioactive ray fee, treatment fee, surgical fee, rehabitation fee, cost of blood plasma, blood dialysis fee, anesthesia fee, special materials fee, medicine fee, medicine service fee, psychiatric treatment fee, injection fee, and the fees were listed in chart 3.1. For the item with total medical fee of zero in NHI

declaration, they were deleted and 33,472 samples were gained. This research adopted the K-means model in SPSS Clementine to analyze the hospitalization fee, with setting the cluster from seasonal medical fee, to acquire the seasonal clusters and usage condition of health insurance resources of hospitalization service users.

4. Preliminary Results

In view of the departments in hospitals were categorized in a delicate sense, this research further arranged these departments into seven sections: internal medicine, surgical department, gynecology department, pediatrics, ophthalmology, E.N.T department, and psychiatric department. The seasonal medical fee setting for research samples were divided into 5 clusters, to observe through the clustering results from K-means in SPSS Clementine. (As in Figure 4-1)

The result showed that the Q1 hospitalization fee of 2007 for internal medicine was assigned in Cluster 1 (As in Figure 4-2), Q2 2007 in Cluster 4, and Q3 2007 in Cluster 4, with total data of 11 seasons. While the statistical results for internal medicine were in Figure 4-3 and external medicine were in Figure 4-4. The clustering results from Q1 2007 to Q3 2009 were connected and the variances of each medical fees were recorded, in order to identify the structural changes on medical fees.

Chart 3-1 Hospitalization Fees of NHI (data sources: arranged by this research)

IDB_AMT_1 Physician Fee	IDB_AMT_7 Surgical Fee	IDB_AMT_13 Medicine Fee
IDB_AMT_2 Room Charges	IDB_AMT_8 Rehabitation Fee	IDB_AMT_14 Medicine Service Fee
IDB_AMT_3 Tube Feeding Fee	IDB_AMT_9 Cost of Blood Plasma	IDB_AMT_15 Psychiatric Treatment Fee
IDB_AMT_4 Inspection Fee	IDB_AMT_10 Blood Dialysis Fee	IDB_AMT_16 Injection Fee
IDB_AMT_5 Radioactive Ray Fee	IDB_AMT_11 Anesthesia Fee	IDB_TREAT_TOT_FEE Total Medical Fee
IDB_AMT_6 Treatment Fee	IDB_AMT_12 Special Material Fee	Records Item #

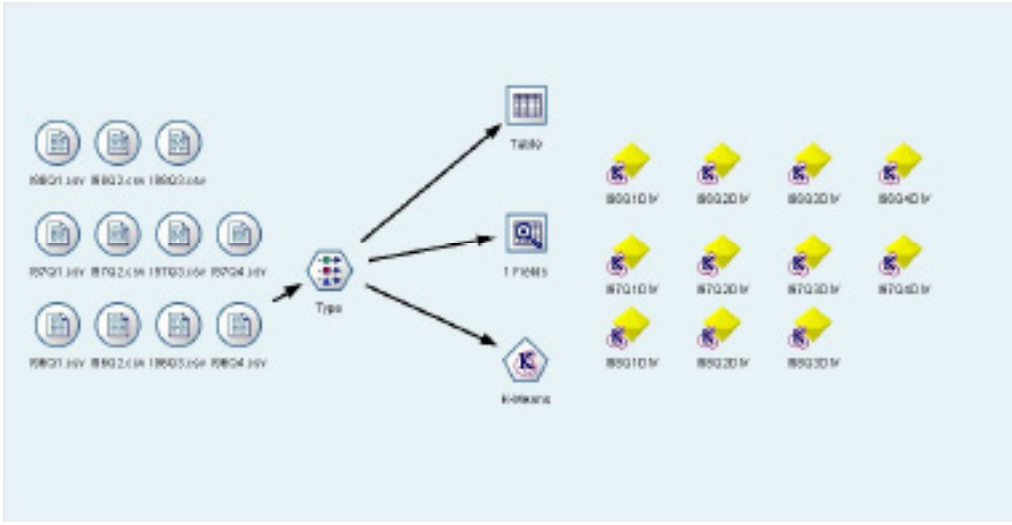
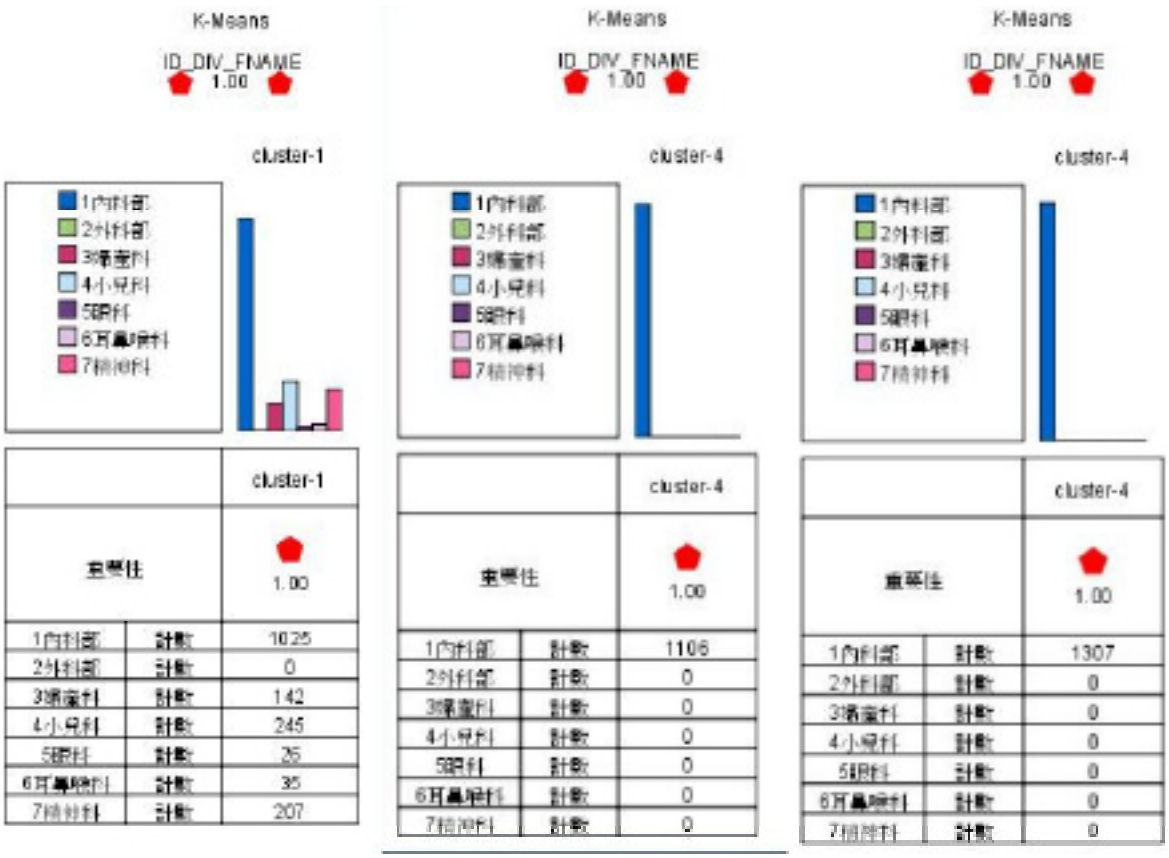


Figure 4-1 SPSS Clementine K-means Clustering Design (Data Resources: Arranged in this Research)



		cluster-1
重要性		1.00
1内科部	計数	1025
2外科部	計数	0
3婦産科	計数	142
4小児科	計数	245
5眼科	計数	26
6耳鼻咽喉科	計数	35
7精神科	計数	207

		cluster-4
重要性		1.00
1内科部	計数	1106
2外科部	計数	0
3婦産科	計数	0
4小児科	計数	0
5眼科	計数	0
6耳鼻咽喉科	計数	0
7精神科	計数	0

		cluster-4
重要性		1.00
1内科部	計数	1307
2外科部	計数	0
3婦産科	計数	0
4小児科	計数	0
5眼科	計数	0
6耳鼻咽喉科	計数	0
7精神科	計数	0

Figure 4-2 Clustering Results for Internal Medicine Department from Q1 to Q3 2007 (Data Resources: Arranged in this Research)

Figure 4-3 Clustering Trend of Hospitalization Fee for Department of Internal Medicine (Data Resources: Arranged in this Research)

Figure 4-4 Clustering Trend of Hospitalization Fee for Surgical Department of Internal Medicine (Data Resources: Arranged in this Research)

According to the professional supervision standard provided by BNHI, it was to restrict the unduly growth of medical fee through respecting the professional self-management of hospital by making use of their medical resources adequately, maintaining medical service quality and reducing inappropriate medical usage.

This research applied data mining clustering techniques to categorize fee structure for hospital, to analyze the structural changes of medical fee, rather than applying statistical analysis requiring definitions of preconditions and analytic dimensions. The results showed the fee structure of certain regional hospital in South Taiwan in the past three years, and there was a decreasing trend of medical fee in the main cluster. In the same period data, Q1 in 2007, 2008 and 2009, taking 2007 as base period, the medical fee decreased 8.9% in 2008 and 20.5% in 2009, which showed the growth of medical fee was controlled gradually.

From the data of medical fee summarized, the Drill-Dow algorithm was applied to track further data from departments, and this research discovered the difference between two fee structure of internal medicine and surgical department. The hospitalization fee, physician fee, room charges, tube feeding fee, treatment fee, cost of blood plasma, medicine fee and injection fee all had a significant level less than 0.05 in two-tail test, while in surgical department, physician fee, room charges, inspection fee, treatment fee, surgical fee, anesthesia fee, medicine fee and injection fee were significantly correlated, and this phenomenon was resulted from the characteristics of operating differences in these two departments. From the clustering data in internal medicine department, in the same period data of Q1 in 2007, 2008 and 2009 with base period of 2007, the medical fee raised 0.4% in 2008 and 5% in 2009. As for in surgical department, with 2007 as base period, the medical fee decreased 5% in 2008 and decreased 27.8% in 2009. Hence, the gradual decreases of overall medical fee were attributed to influences from surgical department.

When the payment standard was based on quantity, the payment to hospital would vary with the number of patients. However, the current NHI policy was capitated payment system, the payment from BNHI to hospital would not vary with the fluctuation of patient number. Hence, in view of internal medicine department discussed in this research, the medical fee only varied slightly (0.4% growth in 2008) from 2007 to 2008, the proportion of room charges to overall medical fee decreased from 36.56% in Q1 2007 to 30.47% in Q1 2008, which showed though the hospitalized patients were fewer, the hospital would like to maintain fixed payment received from NHI system, and might pursuit income through elevating the severity of disease and service density, and therefore the inspection fee at the same period increased from 13.26% in Q1 2007 to 20.31% in Q1 2008.

The operating strategy of medical institution usually tended to increase number of declaration, or expand service area to achieve performance purpose. In view of changes of medical fee in department of internal medicine, as the local population density decreased gradually, or competitors appeared, doctors would ask patients do more inspections under NHI payment, which was not only easily accepted, but elevate the operating performance of hospital, as well as the salary of doctor. However, the BNHI elevated the decreased rate of medical fee

declaration, and simultaneously promoted the implementation of captitated payment system, DRG payment institution and enhance the policy of coinsurance rate to deal with the enormous pressure resulted from skyrocketing medical fee. Thus, the income of medical institution reduced comparatively, as well as the salary of doctors. The operating policy of hospital then drive the doctors to add more self-paid items and amount on patients, such as: making recommendations of the usage of medicine, inspection not covered in NHI or elevating the registration fee.

In the short run, hospitals gained revenue from self-paid items and the overall declaration amount to NHI was not influenced under this policy. Nevertheless, in long-term observation, as the amount of self-paid items increased, the attendance rate decreased accordingly. The inference above was consistent with the data mining results in this research: decreasing hospitalized patients in department of internal medicine, and increasing inspection fee year by year.

Assume the overall variation of medical fee from hospitalization was simultaneously influenced by individual attendance behavior and severity of illness, and then the individual average days of hospitalization represented the attendance behavior, while the average fee spent on hospitalization reflected the fluctuation of severity of disease, if there was no significant variation of average days of hospitalized individuals, then the medical fee was only influenced by severity of disease. In view of the changes of medical fee in surgical clustering, the room charges to overall medical fee maintain 25% in average each year, and surgical fees and treatment fees were not altered significantly, but the overall medical fee presented a decreasing trend, which could be resulted from the aging of local population and geographical relationship of industrial area. Since the alteration of patient number was slight, the fee should only be influenced by the severity of disease. Under the standard of capitated payment system of BNHI, the hospital should maintain certain payment by NHI to support the operating objective, and the policy should focus on decreasing medicine cost, and therefore the medicine cost to overall medical fee was decreased from 8.14% in 2007 to 6.66% in 2009, which was consistent with the surgical clustering data mining in this research.

5. Conclusions and Recommendations

Medical service is directly related to the health and level of living quality, the gradual increases of medical fee has become the universal problem faced by hospitals. The increases of medical fee were most from raising price of medical service and change of medical service types, and not exactly had relationship with the increase of average individual medical service quantity, and hence the traditional descriptive statistics was hard to achieve the objective of this research. This research applied data mining and comparison between the same period to conduct medical fee analysis, to realize the growing trend and variation of fees, and further analyze the cause of variations through discussing the average individual medical behaviors and the fluctuation of severity of disease in each hospitalization. With addition to influences from departments and types, this research made effective prediction and supervision of the reasonableness of medical fee, and the analysis was from the overall aspect to details in each department, to make appropriate, vertical and in-depth discovery upon appropriateness of usage of medical resources. The practical data could be reference for decision makers, and was helpful for avoiding from binding in the past experience and overlooking the trend of overall medical market. The subject in this research was a regional teaching hospital in North Kaohsiung, and most patients (more than 50%) were elders older than 55 years old. Hence this research could be referred to relational research of aging population, severity of disease and medical fee of NHI. In the future, this research method was expected to provide more concrete evidence for hospitals in operating management and strategy.

6. Reference

Wen-Yang Chu (2001). "Strategic Management and Hospital Performance in small and Medium Hospitals: From the Viewpoint of Balanced Scorecard." EMBA Master Essay, NSYSU.

Department of Health, Executive Yuan (2009), “Policy Planning” (2010~2013)

Lee, Shu-xing (1995). “The Practical Innovative Strategic Performance Evaluation." Accounting Research Monthly Taiwan, 1994, 113: 15-23

Tung-Shou Chen; Yu-Lin Chen; Ming-Shan Liou; Wen-Shou Hsu); Chih-Chiang Lin; and Yung-Hsing Ch (2006). " A New Two-Phase Clustering Algorithm Based on K-means and Hierarchical Clustering with Single-Linkage Agglomerative Method"

Computer Journal 17: PP.65-76.

Chang Huai-Lu (2004). "Adopting BSC on Managing Strategy Implementation in National Defense Medical Institution—A case of certain Nationl Army Hospital." EMBA Essay, NSYSU

Chen Pei-Ni (1996). " A Study on the Relationship between Quality of Care and Operating Performance in Taipei Medical Region Hospitals" Master Essay, NTU

Liu Hsing-Kuan (2002). "Hospital Strategy Management" Taipei: Putien Limited Co.

Yien-Yu-Hua, Hsu, Fung & Sun(2009). " Healthcare Fee Forecast Model for a Global Budget System--A Case Study in a Teaching Hospital " Cheng Vhing Medical Journal

Berry, M. J. A., and G. Linoff (1997). "Data Mining Technique for Marketing." Sale, and Customer Support, Wiley Computer.

Berry, M. J. A., and G. Linoff (2000). "Mastering Data Mining: The Art and Science of Customer Relationship Management." Sale, and Customer Support, Wiley Computer.

Fayyad, U., G. Piatetsky, and P. Smith (1996). "From Data Mining to Knowledge Discovery in Databases." AI Magazine,1996, pp. 37-54.

Harkey J., V. R. (1992). "Quality of health care and financial performance: Is there a link? ." Health Care Manage Rev 17(4): 55-63.

Huang, Z. a. N., M.K. (2003). "A Note on K-Modes Clustering." Journal of Classification 20: pp.257-261.

Jiawei Han, M. K. (2006). "Data Mining:Concepts and Techniques." 770.

MacQueen, J. B. (1967). "Some Methods for Classification and Analysis of Multivariate Observations." Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 281–297.

Pham, D. T., Dimov, S.S. and Nguyen, C.D. (2004). "An Incremental K-Means algorithm." In Proceedings of the IMECH E Part C Journal of Mechanical Engineering Science 218: pp.783-795.

T.S. Chen, C. C. L., Y.H. Chiu and R.C. Chen (2006). "Combined Density- and Constraint-based Algorithm for Clustering." In Proceedings of 2006 International Conference on Intelligent Systems and Knowledge Engineering.