

AGENCY FOR INTERNATIONAL DEVELOPMENT
WASHINGTON, D. C. 20523
BIBLIOGRAPHIC INPUT SHEET

FOR AID USE ONLY

Batch 70

1. SUBJECT CLASSI- FICATION	A. PRIMARY Social sciences	SD00-0000-6696
	B. SECONDARY Political science	

2. TITLE AND SUBTITLE
Historical data and computers: the Japanese diet, 1890-1970

3. AUTHOR(S)
Kuroda, Yasumasa

4. DOCUMENT DATE 1972	5. NUMBER OF PAGES 9p.	6. ARC NUMBER ARC
--------------------------	---------------------------	----------------------

7. REFERENCE ORGANIZATION NAME AND ADDRESS
Hawaii

8. SUPPLEMENTARY NOTES (*Sponsoring Organization, Publishers, Availability*)
(Presented at 1st USA-Japan Computer Conf., 1972)

9. ABSTRACT

10. CONTROL NUMBER PN-AAE-974	11. PRICE OF DOCUMENT
12. DESCRIPTORS Computer programs Data processing Japan Legislatures	13. PROJECT NUMBER
	14. CONTRACT NUMBER CSD-3293 211(d)
	15. TYPE OF DOCUMENT

**THIS DOCUMENT HAS BEEN EVALUATED AS SUBSTANDARD COPY FOR
ROUTINE REPRODUCTION. EFFORTS IN AID/W TO OBTAIN A MORE
ACCEPTABLE COPY OF THE DOCUMENT HAVE NOT BEEN SUCCESSFUL.
DESPITE THIS DISADVANTAGE, WE HAVE CHOSEN TO REPRODUCE THE
DOCUMENT BECAUSE OF THE SUBJECT TREATED AND TO MAKE THE
DISCERNIBLE INFORMATION AVAILABLE.**

HISTORICAL DATA AND COMPUTER: THE JAPANESE DIET, 1990-1970

Yasumasa Kuroda

(University of Hawaii, Honolulu, Hawaii)

INTRODUCTION

Preceded by articles on the use of computer for historical research by such authors as Dollar (1) and Murphy (2), the beginning of the 1970s is marked by the publication of several introductory quantitative methodology books for historians, e.g., Aydelotte (3), Dollar and Jensen (4), Shorter (5), and Swierenga (6). The increasing popularity of the use of quantitative methods not only in history but also in other branches of humanities is manifested in the emergence of several journals in the 1960s that are exclusively devoted to the use of scientific methods in the humanities such as Computers and the Humanities, Computer Studies in the Humanities and Verbal Behavior, and Historical Methods Newsletter.^{*} In spite of the initial resistance to the use of quantitative techniques by political scientists and other social scientists, most social scientists in the U.S. today accept the use of the computer as one of the most useful tools for their work. Likewise, the 1970s probably will mark the beginning of a more widespread use of the computer by historians. A recent article presenting a syllabus for quantitative techniques in historical analysis is indicative of the advent of the computer age for historians (Watts (7)). What is to be reported here represents an effort to make historical documents and data more useful for government officials as well as for scholars.

Approximately two pages of information on the social and economic backgrounds for each member of the House of Representatives (elected from the first General Election in 1890 through 1963) were gathered by Professor George Akita, Department of History, University of Hawaii. Subsequently a portion of his data was coded and placed on IBM cards under my direction.² The data placed on the cards were limited to those parts which could readily be quantified. We are in the process of collecting additional data on more recent election results in order to make our data more complete to include all those who have been elected since 1963 through 1970.

When all the data are gathered and cleaned, we will be able to do a number of things hitherto considered Sisyphean tasks. Hopefully, we will have more consistent records of elected Japanese representatives than now available. The paper is designed to demonstrate how we placed all the data on a tape and what we may be able to do with the data through the use of computers. The paper ends with a plea to

establish a National Data Library in Japan for scholars and students in the humanities and social sciences and for any other interested persons.

THE SIZE AND ARRANGEMENT OF THE DATA

The size of the data is probably the largest for this type of study. Studies of social backgrounds of decision-makers by Dogan (8) or Frey (9), for example, have stopped at no more than several thousand cases over a shorter period of time. Our study includes all those who have been elected to the House of Representatives from 1890 when the first General Election was ever held in Japan, through the 32nd General Election, and includes those who were elected in special elections. When the first General Election was held, the number of representatives was limited to 300. Gradually the number has increased over the years to the current figure of 491.

The span of time covered is eighty years. Approximately 13,335 persons were elected during this period. A total of thirty-eight attributes have been gathered for these members of the Diet. All these numbers may be better represented in pictorial form as presented in Figure 1. The vertical axis represents the members of the House of Representatives elected at a given period while the thirty-eight variables are represented by the horizontal axis. This set of data is stored for each election held from 1890 through 1970, as shown in Figure 1.

Since there are a number of M.P.s who are elected more than once, the total number of 13,335 does not mean that there have been 13,335 individual citizens elected. The actual number of people elected only once is less than one-half of the total number of cases. However, we collected data on all individuals for each election because most of them had different occupations at a given time of election or held new positions since last elected. Mr. Yukio Ozaki was elected more times than anyone else in the history of the Japanese Diet, being elected 27 times. Ken Inukai is second being elected 17 times. There are many who have been elected five or six times. Although we should have 508,130 bits of information on our tape, in reality the amount of information we have is significantly less than one-half million for the reasons given above.

Basically there are three different methods in the humanities and social sciences of placing the

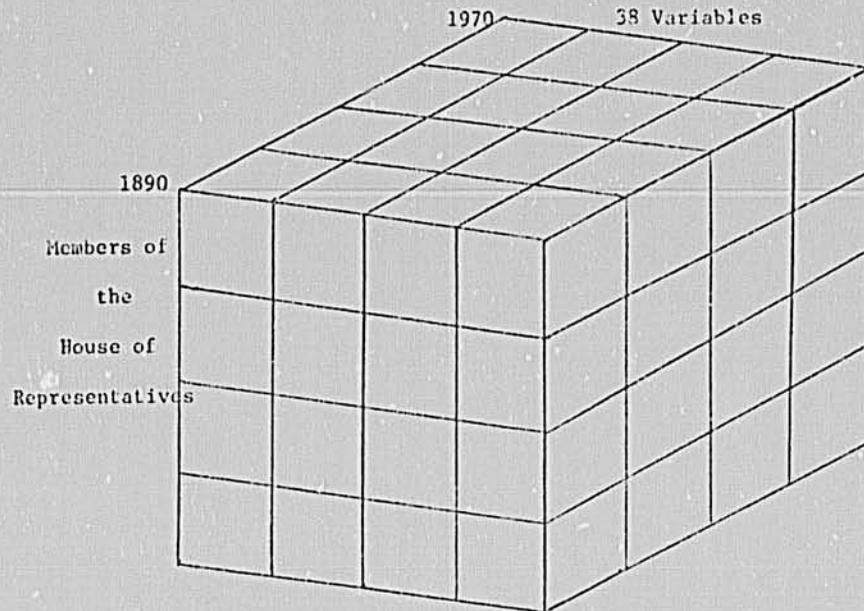


Fig. 1: Data Size

data on the cards: (1) One could theoretically place every bit of information collected on the cards by spelling out everything; this would take an enormous amount of key-punch time but one would have all the data. There are a few serious problems with this method. The cost in terms of manpower, machine time, and materials would be prohibitive for a data collection of this size. Furthermore, the retrieval of the data would be complicated and time-consuming. (2) The second method is the traditional method whereby one individual subject would have his background information placed on the cards. Each column or a combination of a few columns would represent a variable. For example, the date of birth may take three columns and written in such a manner as follows: 972 for 1972. Column five may contain information on sex. A punch on row one on column five may indicate female while a punch on row two signifies male. No punch on column five would denote N.A. (not ascertainable). (3) The third model is a program called BEAR (the Berkeley Elites Automated Retrieval System) developed at Berkeley.³ The BEAR is essentially a combination of the two methods described above. It attempts to

store as much information as the first method without incurring the usual cost of the first method, but it is sufficiently different from the traditional method mentioned above in that it does not restrict itself by using a few columns for one variable but using more than one card to accommodate the need to preserve as much information as possible.

We opted for the traditional method for several reasons.⁴ The amount of resources available for our use was rather limited. Second, it was not until 1970 that we became aware of the BEAR system by which time we had already begun punching the cards. To change the whole set of the data was deemed to be too costly. However, attempts were made to keep as much data as possible.

Figure 2 is constructed to demonstrate the entire process of the data acquisition and data analysis. First, pertinent information concerning social backgrounds of the members of the House of Representatives was gathered and all records were noted in Japanese. Subsequently, a large portion of the data in Japanese

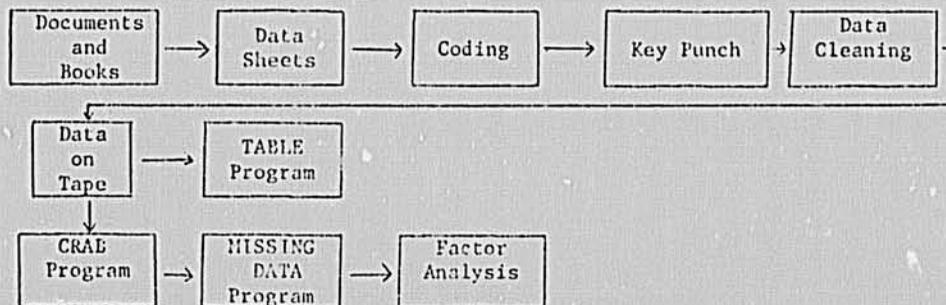


Fig. 2: Data Processing

was translated into numerical values and coded on coding sheets, i.e., coded. The data were then punched onto the IBM cards and verified. They were then tabulated and cleaned, i.e., obvious logical errors were corrected. For example, error was obvious in the case of those who were recorded as female and elected before 1946, since no women were elected until 1946. After all corrections were made, the data were transferred to a tape which could now be used directly to cross-tabulate the data. We have made several analysis data decks to analyze the data for specific purposes. For example, we transformed the data whenever possible from qualitative (nominal and ordinal) to quantitative or metric data by creating dummy variables through the use of the CRAB program, designed to perform recoding and reformatting of data. If needed, a missing data program was used prior to the employment of multi-variate analysis techniques, such as canonical analysis, factor analysis and regression analysis.

UNIQUE DIFFICULTIES IN THE JAPANESE DATA

Although difficulties we encountered may not be any more serious than those encountered by others who have attempted similar projects, there are a few difficulties which are more or less unique to the Japanese materials. First, the pronunciation of Japanese names presents a serious problem. Chinese characters, which are used to describe the members' family and given names, can be pronounced in several different ways. Women's names are often easier in this respect in view of the fact that their names are often written in either "katakana" or "hiragana" (Japanese syllabary) and not in Chinese characters. But they constitute a small minority in the Japanese Diet. Thus, on many occasions, we just did not know how the name should be pronounced. Once pronunciation is established we could transfer the Japanese sounds into the Roman alphabet. There are some books which specify the proper pronunciation of the members of the Diet, but occasionally there are differences among the documents. Furthermore, this problem is complicated by the traditional Japanese family system. Japanese men sometimes change their family name when they marry the heiress of another family. Boys are sometimes adopted into a family without any son in order to perpetuate the family name. Thus, Premier Kishi Nobusuke and Premier Sato Eisaku whose names are different but who in fact are brothers.

A second major problem we encountered is that the Japanese electoral system had gone through a number of changes. Thirdly, there have been times when all representatives elected from a district have been declared to have violated electoral campaign rules and thus the election was considered null and void. A special election then was held and some were re-elected while others were not. Also, there were cases when no further election was held but those who received the highest vote among the losers were automatically considered winners and seated in the House of Representatives.

Minor problems included such items as political party membership. Not only did some members switch their party membership from one election to another but there are so many political parties in Japan that it is difficult to keep track of minor parties.

Obviously, the situation would have been much easier were we dealing with the American party system where two major parties have been in existence for a long time without any significant change. The problem of those postwar Conservative politicians who were elected without the Liberal Democratic party's endorsement and who usually joined the Liberal Democratic party following the General Election, were handled by using the political party identification of each candidate at the time of the election.⁵

Another problem of major significance is that notwithstanding the Japanese meticulousness in maintaining historical government documents, there are discrepancies among the documents, including government publications.⁶ Whatever discrepancies appeared among different sources of the data, we attempted to look for a consistency. Whenever possible we relied on documents which contained more detailed information on individuals and circumstances involved. For example, one book stated that candidate A was elected, while other documents stated that he was not. We consulted with documents which listed not only winners but also losers, accompanied by the number of votes each received.

DATA UTILIZATION⁷

How have the data been used so far? The first type of use has been to describe the social background of the members of the House of Representatives through cross-tabulation. Second, factor analysis has been used to uncover major dimensions of the data on selected General Elections. Third, an exploratory attempt has been made to factor analyze the members of the House of Representatives who were elected in 1946 by treating each member of the House as a variable. This attempt turned out to be the largest factor analysis ever done at the University of Hawaii. The analysis produced a huge 400 by 400 correlation matrix. A difficulty was experienced in our attempt to identify or name factors because of the size of variables. Fourth, attempts also have been made to answer the question of what makes some of the Diet members more successful than others in the sense that they become Cabinet members, party officials, Diet officials and win re-election over and over. Fifth, a Ph.D. dissertation is being written under my direction to answer the question of career patterns in the process of recruitment of the Diet members. Sixth, another dissertation is being prepared in order to predict through the use of simulation techniques who among the members of the Diet are most likely to be recruited into Cabinet positions. The data such as ours should, obviously, not be limited to our own use after we complete our major work. The data will be deposited at as many data libraries as possible for use by any interested party. This raises the question of a data library.

HISTORICAL DATA AND DATA LIBRARY IN JAPAN

How can we maximize the use of the data we gather? Many a data library or archive has been established in the United States to store and maximize the use of public opinion and other data in recent years, e.g., Inter-University Consortium for Political Research at the University of Michigan and the International Data Library and Reference Service at

Berkeley. Unfortunately, however, no data library has yet been established in Japan. Attempts have been made by both scholars and some organizations to set up a data library or data libraries in Japan. For example, Japan Academy of Sciences has established an ad hoc committee for the purpose of creating a data library for the social sciences. I have made a brief survey of major survey research organizations in Japan including major newspapers' polling sections and of some individual scholars in the spring of 1971.⁶ The major newspapers' public opinion survey sections have begun to store their survey results on cards since around 1967. They as well as others are reluctant to release the data for others' use for a number of understandable reasons. The government (the Cabinet Survey Research Office) has been publishing an annual which compiles each year's major public opinion surveys by various organizations in Japan, but much of the original cards are scattered and perhaps discarded by the various organizations engaged in research activities. The polling data can be of excellent use for the future historians interested in Japanese attitudes. At the moment the best place to obtain Japanese survey data is at the Berkeley International Data Library and Reference Service where forty-two sets of the survey data on Japan are located.

Survey data as well as aggregate data in Japan ought to be better preserved at some centrally located data library in Japan. I would like to end my paper by urging scholars, citizens, and government officials in Japan to make concerted efforts toward the establishment of a National Data Library in Japan, where the presently available data can be at least stored if such organizations as newspapers do not wish their data to be used by anyone at this time or in the near future. A considerable amount of polling data conducted prior to the middle of the 1960s seems to be already lost forever. We should not lose any more data than we already have.

One of the reasons for scholars and research organizations in Japan to cling tenaciously to their data may lie in their attitude toward libraries and books in Japan. The Japanese would rather buy their own books than to borrow from a library which is often poorly serviced and equipped. Perhaps a psychological change on the part of the scholars and the research organizations may be necessary before any large size data bank can be built in Japan.

Scholars in the humanities and the social sciences who still believe that the computer is nothing more than a chimera are declining rapidly. Future historians of Japan will find historical data preserved in the form of tapes and disks at data libraries in Japan a gold mine. The creation of such a data bank or library will result in better education for students and better research opportunities for scholars and research organizations.

FOOTNOTES

¹An increasing popularity of the use of the computer in the humanities is manifested in several American journals exclusively devoted to the use of computers in the humanities, e.g., Computers and the Humanities, Computer Studies in the Humanities and Verbal Behavior, and Historical Methods Newsletter. The situation in Japan seems to be lagging. For example, a series of publications on computers published by Sangyo Tosho Publisher in Tokyo does not include any book exclusively written for the humanities or social sciences although a few books such as Jyoho Shori to Tokai Suri (Data Management and Mathematical Statistics) by Chikio Hayashi et. al. (10), contain some valuable information for social scientists.

²Clerical assistance was provided by the Social Science Research Institute at the University of Hawaii. I am indebted to a number of people at the Institute and to Professor George Akita who made the study possible.

³For further details see David Nasatir (11).

⁴For those who are not familiar with this method, see Appendix A which demonstrates how the data are arranged and any part of it can be retrieved.

⁵Since there are no primary elections in Japan, each party nominates a certain number of candidates for each district. The crux of the problem in the nomination process is that the Japanese electoral system has it that three to five candidates will be elected from a district while a voter is allowed to vote for only one candidate. Thus, if a political party nominates more candidates than it can hope to have elected, votes may be dispersed to such an extent that more of its members will lose while the opposition parties may win. Consequently, some party members do not get nominated. Some of them, rather than give up their aspirations, run as Independents. And if elected, in most cases, they join the Liberal-Democratic party even if the party did not endorse their candidacy, hoping that they will be endorsed the next time.

⁶Sources of documents and books used for the data collection appear in Appendix B.

⁷Papers read at professional associations to date include: Yasumasa Kuroda, "Patterns of Recruitment: Japanese Diet, 1946-1963" (12), "Patterns of Recruitment: Japanese Diet, 1890-1945" (13), and "Patterns of Recruitment: Japanese Diet, 1890-1963" (14).

⁸For a detailed report on this survey trip, see Yasumasa Kuroda (15).

Appendix A 1

Card 1

Col. Items

1-4 Identification Number

5 Card Number (1)

6 Blank

7-30 Name

31-32 Father's Occupation

(00) N. A.

(01) Public officials (官僚, 中央政治關係)

(02) Proprietors and managers (民間, 合作社社長, 重役, 商業, 銀行, entrepreneur)

(03) Elected or appointed local government officials, prefecture, city, etc.

(04) Lawyer 弁護士

(05) Legal profession 裁判所關係, 司法官, 調停委員會

(06) Military service

(07) Engineer

(08) Educators 校長, 教授, 先生

(09) Writer

(10) Journalists and others involved in mass media work

(11) Physician

(12) Professionals not mentioned elsewhere

(13) Organizational officials not mentioned elsewhere 團體役員

(14) Labor union officials

(15) Farmers organizational official

(16) Party (political) officials

(17) Religious leaders

(18) Secretary 秘書

(19) Farmer and Fisherman

(20) Clerical personnel

(21) Sales personnel (non-managerial)

(22) Mine workers (non-managerial)

(23) Communications and transportation personnel (non-managerial)

(24) Other skilled, semi-skilled, and unskilled laborers 技能工, 半雇工, 從業者

(25) Security and service personnel

(26) Impossible to classify elsewhere, e.g., housewife

33 Sex

(1) Male

(2) Female

34-37 Birth Date (Year only)

(0000) N. A.

38 Blank

Appendix A 2

GENERAL ELECTION 30

LIBERAL DEMOCRATIC PARTY

HOKKAIDO PREFECTURE	SASAKI HIDEYO
	SHITKUMA SABURO
	SHINODA KOSAKU
	TANAKA MASAMI
	NANJO TOKUO
	HONNA TAKESHI
	MATSUURA SHUTARO
	SATO TAKAYUKI
	SUHARA SHOICHI
AOMORI PREFECTURE	CHIZAKI USABURO
	NAKAGAWA ICHIRO
	MORITA JUJIRO
IWATE PREFECTURE	KUMAGAI YOSHIO
	TAZAWA KICHIRO
	ISURUGI MICHIOYUKI
MIYAGI PREFECTURE	OZAWA SAKKI
	SHIGA KENJIRO
	SHIINA EISUSABURO
	SUZUKI ZENKO
	NOHARA HASAKATSU
	AICHI KIICHI
AKITA PREFECTURE	UTSUMI YASUKICHI
	OISHI EITICHI
	HASEGAWA TAKASHI
	HOSHINA ZENSHIRO
YAMAGATA PREFECTURE	ISHIDA HIROHIDE
	SASAYAMA SHIGETARO
	NEMOTO RYUTARO
	SASAKI YOSHITAKE
FUKUSHIMA PREFECTURE	IKEDA MASANDSUKE
	KATO SEIZO
	KIMURA TAKEO
	KUROGANE YASUMI
	MATSUURA TOSUKE
FUKUSHIMA PREFECTURE	MATSUZAWA YUZO
	AMANO KOSEI
	ITO MASAYOSHI
	KIMURA MORIE
	SAITO KUNIKICHI
	HATTA SADAYOSHI
	KAMEOKA TAKAO
	SHIRUYA NAJZO
MINATO TETSURO	
ROMIYAMA HIDE	

Appendix B

人學與信新報「人學與信錄」東京 人學與信新, Vol 1~23, 1903~1965

木戶鏡之助「日本國國會史」東京 1890.

「天誅後行軍及所借利債」Vol 1, 2, 1890.

「大日本」(詳見)東京:東京經濟雜誌社, 9th edition, 1921:

「人學名鑑」東京 Vol 1, 2, 1932.

「明治元朝打馬大論」

「明治讀史」, Vol. 7, 8 1890, 1892

「開國三十年史」

伊尾常吉「近代官制」

「日本政治學」

「國民學」, Vol. 1915~1930

「省日與港」 1931, 1940

「人學名鑑」, Vol 1, 2, 1932.

「大東人學錄」, Vol 1, 1940..

「地方自治與行政學」(靜岡・愛知).

大信濃

「日本紳士錄」 1930, 1936, 1941, 1950, 1954

「取真錄」 1930, 1931, 1933, 1935,

1949, 1952, 1956, 1957, 1960, 1962, 1963, 1965

「官界名鑑」 1952

「人學新報」 1947.

「勞働年鑑」, Vol 1, 2 1958, 1960

衆議院參議院「議會制度七十年史」政党内務部 東京 印刷局, 1961

" " 貴族院參議院議長名鑑 "

" " 衆議院議員名鑑 "

(Inco, etc)

REFERENCES

- (1) Dollar, Charles M. "Innovation in Historical Research." Computers and the Humanities, Vol. 3, No. 3, pp. 139-151. January, 1969.
- (2) Murphy, George G. S. "Historical Investigation and Automatic Data Processing Equipment." Computers and the Humanities, Vol. 3, No. 1, pp. 1-13. September, 1968.
- (3) Aydelotte, William O. Quantification in History. Addison Westley Publishing Co., Reading, Mass., 1971.
- (4) Dollar, Charles M. and Jensen, Richard J. Historian's Guide to Statistics. Holt, Rinehart & Winston, New York, 1971.
- (5) Shorter, Edward. The Historian and the Computer: A Practical Guide. Prentice-Hall, Englewood Cliffs, New Jersey, 1971.
- (6) Swierenga, Robert P., ed., Quantification in American History: Theory and Research. Atheneum, New York, 1970.
- (7) Watts, Eugene J. "Quantitative Methods in Historical Analysis: A Syllabus." Historical Methods Newsletter. Vol. 5, No. 2, pp. 59-67. March, 1972.
- (8) Dogan, Mattei. "Political Ascent in a Class Society: French Deputies 1870-1958." In Dwaine Marvick (ed.), Political Decision-Makers: Recruitment and Performance. The Free Press, New York, 1961, pp. 57-90.
- (9) Frey, Frederick W. The Turkish Political Elite. MIT Press, Cambridge, 1965.
- (10) Hayashi, Chikio, et al. Jyoho Shori to Tokai Suri (Data Management and Mathematical Statistics). Sangyo Tosho, Tokyo, 1970.
- (11) Nasatir, David. "Developing an Archive of Machine Readable Biographical Data." Paper delivered at the Annual Meeting of the American Political Science Association, Los Angeles, September, 1970.
- (12) Kuroda, Yasumasa. "Patterns of Recruitment: Japanese Diet, 1946-1963." Paper delivered at the Annual Meeting of the American Political Science Association, 1970.
- (13) Kuroda, Yasumasa. "Patterns of Recruitment: Japanese Diet, 1890-1945." Paper delivered at the 28th International Congress of Orientalists, Canberra, Australia, 1971a.
- (14) Kuroda, Yasumasa. "Patterns of Recruitment: Japanese Diet, 1890-1963." Paper delivered at the Annual Meeting of the Association for Asian Studies, Washington, D.C., 1971b.
- (15) Kuroda, Yasumasa. "Public Opinion Survey and Data Library in Japan." Unpublished paper, Social Science Research Institute, University of Hawaii, 1971c.