Development of Data Science Program
for Working Women

Maki Tokii[1], Chieko Mizoue[1], Hidehiko Hasegawa[1],

and Megumi Umeda[2]

[1] University of Tsukuba, 1-2 Kasuga, Tsukuba, Ibaraki, 305-8550, Japan
[2] IBM Japan Ltd., 19-21 Nihonbashi Hakozaki, Chuo-ku, Tokyo, 103-8510, Japan
mizoue@slis.tsukuba.ac.jp

Abstract. In recent years, data science has been paid attention in Japan, and data science education has begun to be implemented for working people. It can be said that with the progress of the knowledge-based society, the need for the ability to understand and analyze data is increasing in the modern society. On the other hand, the decrease in the population of working age in Japan is caused to activation of women. As these two factors cross each other, we believe that there is a strong demand for data science education for women working in companies that have just finished a family leave.

This study aims to develop a data science program for working women who finish a family leave aiming to identify some issues. The data science program was conducted four one and a half hour training classes in February 2019 developed by the University of Tsukuba, targeting employees of IBM Japan. Two types of questionnaire survey were conducted to measure the effectiveness of the program.

Many participants felt the pace was appropriate, while the differences among them could be seen depending on their level of understanding of statistics. As for understanding of the contents, they could understand in essence.

Regarding of usefulness, they chose to be useful. It can be said that this program succeeded in motivating their learning of data science. Furthermore, the fact that many requests for practical training and exercises can be said to indicate that extending the one and a half hour training time to two hours can lead to the acquisition of useful skills for work.

Keywords: Data Science, Working Women, Education Program

1. Introduction

1.1 Japan's Declining Birthrate and Aging Population

Japan, as one of the world's leading low birth rates and aging population, is facing problem with the declining population of the working-age from 15 to 64. While Japan's

total population is decreasing, elderly population aged 65 and over is increasing, and working-age population is rapidly decreasing. Against this background of demographic changes, Japan has high expectations for the use of elderly people and women for the maintenance and development of its society. Let us look at the current working situation of women in Japan.

The number of employed people in Japan in 2018 is 29.5 million women and 37.2 million men. The employment rate in the working-age population is 69.6% of women and 83.9% of men. Both of the number of employed women and its rate is lower than the ones of men (Cabinet Office, 2019).

Above all, it has been pointed out that the labor force participation rates of women in Japan by age group figures an "M-shaped curve" due to the burdens in the labor force participation of women around their thirties (Cabinet Office, 2019).   It means that the existence of women who resign due to life events such as marriage and childbirth, and return to work after finishing a family leave. In recent years, people who support life stages in Japan "married, have children but continue working" is 43.0%, while those who support life stages "married and have children, but once quitted for the opportunity of marriage or childbirth, work again after parenting" is 36.0% (Soken, 2019, p.62). Therefore, the period corresponding to the M-shaped valley is shortened and the bottom age class is increased (See Fig. 1).
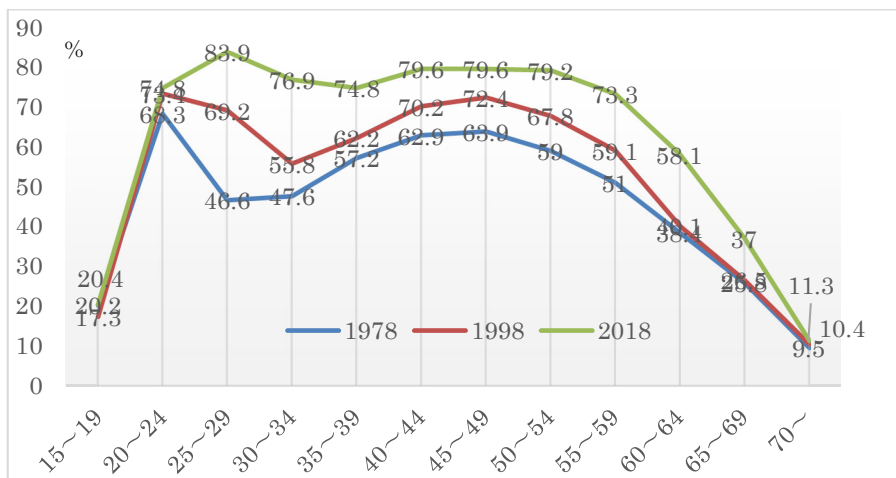

Fig. 1 Women's Labor Force Participation Rate by Age Group in Japan

Source: Cabinet Office (2019). White Paper on Gender Equality: 2019

However, of the 2.37 million Japanese women who wish to work, the largest number for non-job-seeker reason is still "for childbirth/childcare" (32.6%) (Statistics Bureau of Japan, 2018, p.11). The "M-shaped curve" employment patterns

have not been completely wiped out compared to European countries (Cabinet Office, 2019).

So where do they acquire the knowledge and skills needed for work when they face a balance between home and work? According to a survey by the Cabinet Office in 2018, 80.3% of women are "(I have) done while working" (Soken, 2019, p.27), but in another survey, the largest number of training at work companies is "joining in the training at the time of employment" (67.4%), and the "training on career development for the future" is only 16.1% (The Japan Institute for Labor Policy and Training, 2018, p.363).

With the changing demographics of Japanese population and the aim of utilizing women in our society, especially in companies, now a need for training to contribute to women's career development is required while achieving a balance between home and work.

Additionally, from around 2010, data science has been paid attention in Japan. In recent years, not only in school education but also for working people, data science education has begun to be implemented. It can be said that with the progress of the knowledge-based society, the need for the ability to understand and analyze data is increasing in the modern society.

As these two factors cross each other, we have a need for data science education for women working in companies that returned from a family leave.

1.2 Why Data Science?

We need to consider what skills are useful for women workers who returned from a family leave. In some companies related to computers and software such as IBM Japan, most of them are college graduates, have general knowledge and skills, and are used to intellectual work. Then, we chose data science for this purpose.

There is no clear definition in data science. This article uses "data science" as a general methodology for effectively using data collected by using computer systems as an extension of the Internet, IT, and IoT concepts. The purpose of data science is to increase productivity and enrich society based on collected data. Although this purpose has been insisted, it could not have been achieved. Since big data, artificial intelligence, and machine learning have become very familiar recently, data science has also become very important again.

However, in Japanese universities, application fields such as data science have not been taught enough. The urgent need of society is for people with data science skills, and organizations that produce human resources with such skills. Roughly speaking, the concept of data science consists of the concepts of data analysis, database technology, and knowledge of application areas. It is necessary to improve data

analysis and database technology skills. Knowledge of application area is also important, but it is not easy to improve its skill in a general sense. Each area has different characteristics.

Computer and software companies such as IBM Japan, particularly women workers who returned from a family leave, have several advantages in dealing with data science.

1) There is no need to learn basic computer operations.
2) It is necessary to learn basic concepts of data analysis and database.
3) It is possible to learn continuously at home without time limit
4) No need for physical labor or commuting

This means they can improve their data science skills regardless of time and place. Acquiring these skills may not be easy, but is achievable. These skills are valuable and will not be useless in the future.

2 Literature Review

"Efficient analyzing and utilizing of those data may bring abundant living".

It is said that it is necessary to train data scientists who can find problems from various data and have the capacity to use analysis results to solve problems (Maruyama et al., 2015).

In order to provide much people with data science knowledge, online courses are effective and currently various courses are provided. The Statistics Bureau of Japan offers introductory course on data science for working people (Statistics Bureau of Japan, n.d.). The Ministry owns and publishes a wide range of data related to people's lives. If it is possible to interpret data from various viewpoints using knowledge obtained from the course through these open data, it can be expected that ideas for using more data will be born.

In 2017, the Faculty of Data Science was established at Shiga University, Japan (Takemura et al., 2018). In addition to techniques for analyzing data, fieldwork is also conducted to learn examples of data utilization in the business domain. The curriculum is designed to acquire problem solving skills. The University of Tokyo Extension has also set up a data science course, and provides an opportunity for working people in various fields to learn data science (The University of Tokyo Extension, n.d.).

In this way, data science learning opportunities are beginning to be provided to various layers.

3 Purpose and Method of Research

3.1. Research Purpose
This study aims to develop a data science training program for working people. This time, we developed a pilot program for women who return from a family leave aiming to identify some issues.

3.2. Research Method
This pilot program was conducted four times in February 2019 for the data science program developed by the University of Tsukuba, targeting employees of IBM Japan Ltd. (IBM Japan) (See Table 1). The lecturers were member of the faculty of the Library Information and Media Science, University of Tsukuba.

Since this is a program for reinstated women employees, the opening time was set under company lunch break (from 12:00 to 13:30), held in IBM Hakozaki Office meeting room. Attendance was optional. 20 participants were recruited through public invitations by intranet and direct e-mail to employees who participated in the past year's parenting support seminar hosted by Diversity Division.

The affiliation of the participants was as follows:

Human Resources: 7, Finance: 6, Marketing:6, Engineering: 1

Among the participants, 13 have taken a family leave within the past 5 years. All participants had experiences to use Microsoft Excel, but there was a considerable difference in work experience.

In this program, two types of questionnaire were conducted to measure the effectiveness of the program. The first one was a questionnaire survey on information utilization which was conducted before and after the training. This questionnaire items were selected and modified the survey of the "Creating a scale of practical use of information and examining its reliability and validity" (Takahara et al., 2001). The second one was conducted after each class in order to extract improvement points in the content.

Table 1 Pilot Program Outline

Date		Contents	No. of Attendee
1st class	2019/02/12	What is date science	14
2nd class	2019/02/15	Understand data-1 （Analyze data)	11
3rd class	2019/02/19	Understand data-2 （Case study)	16
4th class	2019/02/26	Understand data-3 （Visualize results)	12

The syllabus of this program is on the following website: http://www.slis.tsukuba.ac.jp/~tokii.maki.ga/DataScience/

4 Outline of Program

We aimed to provide an opportunity for each participant to consider "relationship between own field and data science". We considered the program contents to achieve this purpose by introducing cases of various fields in the flow of "Collection", "Storage" and "Utilization" of data.

This program consisted of four classes and each class had 60-minute lecture and 30-minute exercise. By utilizing more familiar data, we devised to be able to acquire the analysis skill necessary for data science.

The goals for the four classes are as follows;

1st class: To know what data science is. Review of basic statistics
2nd class: To know the flow from data collection to utilization
3rd class: To know what you need to interpret the analyzed results
4th class: To know the flow visualizing analyzed results

4. 1. a) 1st class (lecture)

First, in order to consider "what could be done" by analyzing data, the lecture aimed to realize the necessity of recognizing the analysis result from various directions. The class provided explanation of the meaning of the numerical value which could be obtained from statistics widely. By showing concrete analysis examples from basic keywords of statistics, we got to know which part the participants did not understand from their reaction.

4.1. b) 1st class (exercise)

The lecturer put data in a table, made a scattergram, and performed an exercise to calculate basic statistics using Excel. It may be difficult to see how the statistics can be obtained by calculation and how the statistics be changed by the numerical value changes. Therefore, at first, analysis methods and analyzed results using several data were shown, and then data for analysis was increased to about 50. The consumption of food in each city published by the Statistics Bureau of Japan was used. As this data is open to public, the participants themselves acquired the same open data and they would have an additional learning opportunity by performing analysis other than the food shown as a sample.

Finally, for data in the next class, we conducted a questionnaire asking anonymous questions about food preferences with 5-point scale.

4.2. a) 2nd class (lecture)

The lecture started with the review element of statistics. Focusing on the fact that one statistics alone cannot see trends in the data, we have presented concrete examples

from the mean and variance of the data. By increasing the number of items, we presented the keywords of the "t-test", and lectured that it was possible to find various trends based on basic statistics.

In this lecture, we introduced cases using various data from data acquisition to analysis and utilization.

First, we showed how to create a questionnaire on the web site that was conducted to the participants at the 1st exercise. It was shown that we could share the acquired data with all the participants, quantify the questionnaire results and visualize the trends.

Next, as a concrete example, the results of the questionnaire survey on sports by the Statistics Bureau of Japan were shown, and the analyzing method with many items was introduced. These data were used to indicate the need to look at data from various perspectives, by items, year of data acquisition, age group, and even by geographical area. As for the data indicating the geographical area, we presented a method to visualize the data on a map. It was able to lead to the introduction of the field of GIS (Geographic Information System).

4.2. b) 2nd class (exercise)
Using the food preference data generated by the participants in the 1st class (exercise), the steps of "summarize in table", "quantify" and "determine basic statistics" were shown. We aimed to let the participants feel data science more familiar than the ever.

4. 3. a) 3rd class (lecture)
A 90-minute lecture on research was given on the subject of sensibility and data science. In order to analyze what people feel interesting, it was indicated which action of the person should be quantified and characteristic value to be used for analysis was narrowed down from the obtained numerical value showed that. It was shown that narrowing down the obtained feature quantity makes the result easier to interpret. By introducing research that deals with large-scale data, we aimed to convey the importance of preprocessing data.

4.3. b) 3rd class (exercise)
No additional exercises

4. 4. a) 4th class (lecture)
We introduced methods such as Principal Component Analysis and clustering with the theme of distance as to how to look at the data. By these methods, we showed that cases where the data presented in the previous exercises could be analyzed from another perspective.

First, we asked the participants, "Can you analyze that people who like A tend to like B as well?". We introduced a part of the analysis to see the co-occurrence relationship in Excel, then introduced the analysis software R that could be useful for the large-scale data, and tried to create opportunities for using specialized analysis software.

Finally, we described the process of putting together the obtained results and talked the necessity of selecting items to show in information visualization. By introducing information visualization technology, we aimed to have the participants reaffirm the importance of communicating in the workplace.

4. 4. b) 4th class (exercise)

For calculation of basic statistics to t-test, we provided a file that each participant could modify data on spreadsheet software for practice. A total review of the lectures was performed.

5. Result by Questionnaires

5. 1. Results of Questionnaire Conducted Each Class

In this pilot program, we conducted a questionnaire each class. The purpose of this questionnaire is to review the class contents of the next round and to utilize it for future program composition by knowing the understanding of the participants. The five 5-point scale items, one 3-point scale item and a free description were used. As Table 2 shows, in Q1 and Q3, answer "3" means that the progress of the class is appropriate. In Q2, Q4 and Q6 with 5-point scale, "5" is selected when satisfaction is very high. Q5 is 3-point scale and "3" is selected when satisfaction is high.

Table 2 Questionnaire Items Conducted Each Class

Question No.	Contents of question	
Q1	How was the class of the training by the instructor?	
Q2	Was the content easy to understand?	5-point scale
Q3	Was the length of time appropriate?	
Q4	Was this training helpful to you in the future?	
Q5	Would you recommend this training to your colleagues?	3-point scale
Q6	Please tell us about your overall satisfaction.	5-point scale

The mean and standard deviation (SD) are shown in Table 3.

Table 3 Questionnaire Results

Question No.	1st class (n=11)		2nd class (n=7)		3rd class (n=18)		4th class (n=11)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Q1 (Speed)	3.45	0.78	3.14	0.35	3.06	0.52	3.55	0.50
Q3 (Length)	3.00	0.43	3.00	0.00	2.78	0.42	3.18	0.39
Q2 (Understanding)	3.45	0.78	4.00	0.76	3.61	0.95	3.73	0.96
Q4 (Usefulness)	4.27	0.75	4.71	0.45	3.89	0.81	4.18	0.57
Q6 (Overall Satisfaction)	4.18	0.57	4.43	0.49	4.00	0.58	4.27	0.62
Q5 (Recommendation)	2.27	0.62	2.71	0.45	2.12	0.58	2.18	0.39

Table 3 shows that the mean of the "Speed" of the class (Q1) was from 3.06 to 3.55. It can be said that many participants felt that the pace was appropriate. In the 1st class, since the standard deviation was larger, differences among participants could be seen depending on the participants' level of understanding.

About the "Length" of the class (Q3), it can be understood that many participants felt the length of class was moderate (Mean was from 2.78 to 3.18). As for "Understanding" of the contents (Q2), many participants could understand almost (every means was over 3.45). And regarding of "Usefulness" (Q4), many participants chose to be useful since every means was over 3.89. Q6's "Overall Satisfaction" was relatively high (every means was over 4.00). The second and fourth classes, when familiar examples were introduced, were much satisfying.

Regarding "Recommendation" with 3-point scale (Q5), many participants would like to recommend to others, since every means was over 2.12.

5. 2. Results of Pre-post Comparison Questionnaire Survey

"A Scale for the skills of practical use of Information" by Takahira et al. (2001) was modified and conducted for this survey. The same questionnaire was used twice pre and post attending.

Two or three questions were selected for each of the six skills shown in the previous research. In this study, a total of 13 were used as question items. Since it was a questionnaire for school students in Takahira's research, some question items were changed to questions related to workers. The questions and results are shown in Table 4.

In this study, since the 4-point scale was used, an item showing three or larger values are an item selected as "applicable".

20 people answered pre-attending, and ten answered post attending. In Table 4, the mean and standard deviation (SD) were shown, and the mean value indicating a

value of 3 or more and the standard deviation indicating a value of 1 or more are underlined.

Table 4 Pre-post Comparison Results of Information Utilization

Question item	Pre-attending （n= 20)		Post attending (n=10)	
	Mean	SD	Mean	SD
① Collection skill				
Q11: I am a type of person who reads a talked about book/journal.	2.40	0.92	2.80	0.75
Q12: I am a type of person who does not read a newspaper and TV news.	2.30	0.84	2.30	0.90
② Judgment				
Q13: I am a type of person who believes a press/TV reporting.	2.35	0.73	2.40	<u>1.02</u>
Q14: I am paying attention to whether work-related data is old.	<u>3.30</u>	0.64	<u>3.00</u>	0.89
③ Expression				
Q15: I have rewritten a figure/table by myself to understand the writing.	<u>3.00</u>	<u>1.00</u>	<u>3.40</u>	0.92
Q16: I am a type of person who classifies data by contents when many of them are collected.	<u>3.50</u>	0.50	<u>3.40</u>	0.66
④ Processing skill				
Q17: I can summarize data well, even it is various.	2.45	0.74	2.80	0.75
Q18: I am not good at understanding what the author/speaker wants to say.	2.55	0.80	2.00	0.63
Q19: I am good at drawing similarities among a plenty of data.	2.40	0.58	2.70	0.78
⑤ Imagination				
Q20: I am a type of person who considers from a different point of view.	2.65	0.65	2.80	0.75
Q21: I am not good at creating a new/different thing.	2.85	0.79	2.80	0.75
⑥ Outgoing ability				
Q22: I am good at making the easy-understanding documents/materials.	2.45	0.86	2.60	0.66
Q23: I am not good at explaining my idea/thought to others.	2.70	0.78	2.80	0.75

As Table 4 shows, it was understood that many participants selected "applicable" (Q14, Q15 and Q16). The value of the standard deviation pre-attending showed that there was a difference between the participants (Q11, Q15, and Q22). In the data science of Q15 and Q22, it is considered to be close to the item that asks the ability concerning visualization of data.

Looking at the values of Q14 and Q15 shown in Table 4, 0.1 point is down. Therefore, we divided it into a group who answered both questionnaires and a group who answered only one questionnaire. The mean is as follows:

   Q14: Pre-attending 3.30 (2.89, 3.64), Post attending 3.00(2.88, 4.00)

   Q16: Pre-attending 3.50 (3.33, 3.64), Post attending 3.40(3.33, 4.00)

Value in parentheses: (Answered both questionnaires, Answered pre-questionnaire only)

The value of the group answering both questionnaires shows little change. It shows that those who answered both questionnaires have lower self-assessment than those who answered the pre-attending only.   People who felt the need for data knowledge tended to participate actively in this program. However, it is said that skill improvement has not been realized in the four classes. To meet the needs of the participants, it is necessary to evaluate their skill improvement and understanding of the program contents for each class, and to analyze the results.

Furthermore, Table 5 showed the results of nine people who answered both questionnaires.

Underlining was applied to items whose mean value has risen by 0.2 or more.

Table 5 Pre-post Comparison Results of Information Utilization of Both Attending (n=9)

Question No	Pre		Post	
	Mean	SD	Mean	SD
Q11	2.56	0.83	2.78	0.79
Q12	2.22	1.03	2.33	0.94
Q13	2.33	0.67	2.33	1.05
Q14	2.89	0.57	2.89	0.87
Q15	3.22	0.79	3.33	0.94
Q16	3.33	0.47	3.33	0.67
Q17	2.44	0.68	2.67	0.67
Q18	2.44	0.50	2.11	0.57
Q19	2.22	0.63	2.56	0.68
Q20	2.67	0.67	2.89	0.74
Q21	2.89	0.74	2.89	0.74
Q22	2.56	0.83	2.44	0.50
Q23	2.67	0.67	3.00	0.47

It is difficult, on the one hand, to see changes in the practical skills of information utilization due to a short period of time, but small changes were seen in processing skill (Q17 and Q19). There were also small changes in the imagination (Q20).

On the other hand, the means of Q18 and Q22 were decreasing. We supposed this might have been influenced by the contents of the 4th lecture. In this lecture, the previous lectures were summarized and the various analysis methods were introduced. A participant said "I wanted to hear more slowly." So, the participants might assess their understanding skill as low. In the 4th lecture, we also introduced infographics. Since many participants have thought the importance of information in the workplace, they were interested in infographics. After learning the techniques to make clear expressions, however, they might lose their confidence of making documents/materials.

6 Consideration

Based on the above results and free description of the questionnaire conducted after each class, we considered construction of future programs.

6.1 Handling Various Level of Participants' Understanding
The understanding level of the participants was varying. Since the mean of the speed of the class was from 3.06 to 3.55, many participants felt that the speed was appropriate. However, in the 1st class, since the standard deviation was larger, differences among them could be seen depending on the participants' level of understanding of statistics. And, for example, a participant said "I wanted to confirm the understanding level on the way". This program was widely introduced from basic knowledge of statistics to utilized cases of data science. After the second class, we responded as needed to supplement the items that appear to have differences in the level of understanding based on the content of the questions and requests from the participants. Because the participants were working in various fields, the scope of interest was wide and it was difficult to explain everything in the lecture. Therefore, we established a web page for the class and introduced reference books and web pages as needed.

Many participants were using Excel in their work. The participant said that learning data analysis methods using Excel anew was a hint for solving business problems. So, it can be said that using Excel as an exercise is appropriate in an entry tool of data science education.

6.2 Providing the Familiar Data
Using open data and showing concrete examples of analysis, there was an opinion "I could get an opportunity to consider that data science would use the cases related to my

duty." The overall satisfaction of the second and forth class was relatively high. Since we introduced the familiar examples in these classes, we supposed that the familiar examples made the participants be satisfied. By using familiar and easily accessible data, we can provide the participants to improve the motivation for learning about data science and create opportunities for learning.

6.3 Showing the Leading-edge Research Activities

By showing the research of the quantification of sensibility, we succeeded in getting the possibility of data science felt by quantifying a variety of information. Some participants wanted to know more detailed analysis methods, and some were interested in the analysis process as well as the results. We thought it was a good opportunity to consider what could be done with data science. In addition, some participants considered the possibility of data science to be able to add value to data by showing academic research. We can successfully provide a learning opportunity to the participants who think about data science from various perspectives from business side and daily life.

6.4 Increasing Practice Time

Many participants showed more interest than expected in the area of infographics used to analyze and interpret data and communicate the results. In the future, we will consider to provide a visualization class that introduces visualizing results in various areas at first, then introduces analysis methods. This will make it easier to create an opportunity to think in conjunction with the work.

By showing the process from data collection to analysis and expression, some participants linked with their works, considered what could be done with data science, and asked questions. While many participants felt the length of class was moderate, there were many opinions to increase practice time using specific data. It was difficult to increase length due to time restriction. In the future class, we should provide some resources that the participants can be able to analyze them at home as a kind of homework.

7 Conclusion

From around 2010, data science has been paid attention in Japan. In recent years, not only in college students but also for working people, data science education has begun to be implemented. It can be said that with the progress of the knowledge-based society, the need for the ability to understand and analyze data is increasing in the modern society.

At the same time, considering the Japanese society, the decrease in the population of working age is caused to activation of women. As these two factors cross each other, we believe that the skill up in data science is the best for women working companies that have returned from a family leave, because it can be improved regardless of time and place. These skills are valuable and will not be useless in the future. The purpose of this project is to introduce the basic concepts of data science and to invite them to more advanced concepts for women working in companies that have returned from a family leave.

Needless to say again, in the four-time series of programs, it was not possible to reach the stage of examining the educational effect of the program. However, according to the results of questionnaire, as indicated by the free description, such as "it became a hint to solve business problems by learning how to analyze data", they showed strong interest in the concepts of data science. So, it can be said that this program succeeded in motivating their learning. Furthermore, the fact that many requests for practical training and exercises can be said to indicate that extending the a half hour exercise time to one hour can lead to the acquisition of useful skills for work or require self-training such as home work.

Based on the results of this pilot program, we would like to advance the development of more practical programs in data science that are much useful to working women in the future.

References

1. Cabinet Office (2018). *Annual Report on the Ageing Society: 2018* (Japanese). Retrieved from https://www8.cao.go.jp/kourei/whitepaper/w-2018/html/zenbun/index.html.
2. Cabinet Office (2019). *White Paper on Gender equality: 2019* (Japanese). Retrieved from http://www.gender.go.jp/about_danjo/whitepaper/r01/zentai/html/honpen/b1_s02_01.html.
3. The Japan Institute for Labor Policy and Training (2018). *Tayo na Hatarakikata to Jinzai manejimento no arikata ni kansuru tyosa* (Japanese). Retrieved form https://www.jil.go.jp/institute/research/2018/documents/184_02.pdf.
4. Maruyama, Akimichi et.al (2015). Developing Data Analytics Skills in Japan: Status and Challenge. *Journal of Japan Industrial Management Association* 65(4), pp.334-339.
5. Soken (2019). *Report on Learning to enable a variety of choices* (Japanese). Retrieved from http://www.gender.go.jp/research/kenkyu/pdf/select_research/08.pdf.
6. Statistics Bureau of Japan (2018). Labor Force Survey 2018 Yearly Average Results (Japanese). Retrieved from https://www.stat.go.jp/data/roudou/sokuhou/nen/dt/pdf/ndtindex.pdf.

7. Statistics Bureau of Japan (n.d.). Online course for data scientist. Retrieved from http://gacco.org/stat-japan/.
8. Takahira, Mieko et.al (2001). Construction of a Scale for the Skills of Practical Use of Information and Its Reliability and Validity (Japanese). *Japan journal of educational technology*, 24(4), pp.247-256.
9. Takamura, Akimichi et.al.(2018). Shiga University Model of Data Science Education. *Proceedings of the Institute of Statistical Mathematics*, 66(1), pp.63-78.
10. The University of Tokyo Extension (n.d.). University of Tokyo Extension Data Science School. Retrieved from https://www.utokyo-ext.co.jp/utex/asp-webapp/jsp/web/utex/base/school/index.jsp