

경력단절여성을 위한 연령 및 학력별 취업률이 높은 교육직종 빅 데이터 시각화 분석

이정원¹ · 이충호^{2*}

Big Data Visualization Analysis of Education Occupations with High Employment Rates by Age and Educational Background for Career-Interrupted Women

Jeongwon Lee¹ · Choong Ho Lee^{2*}

¹Graduate Student, Department of Information Communication Engineering, Hanbat National University, Daejeon, 34158 Korea

^{2*}Professor, Department of Information Communication Engineering, Hanbat National University, Daejeon, 34158 Korea

요 약

경력단절여성들의 재취업을 위한 교육 훈련 등의 지원정책이 이루어지고 있으나, 효과적으로 취업과 연계되지 않고 있다. 또한, 고학력 경력단절여성의 경우에는 재취업을 위하여 이전 경력을 수정하거나 새로운 경력을 계획해야 하는 등 재진입이 어려운 상황이다. 기존 선행연구들에서는 취업기회가 높은 유망 직종 등 재취업을 위한 근본적인 문제를 해결하기 위한 연구가 부족하였다. 따라서 경력단절여성들을 위한 교육과정 개발 시 교육 대상자의 연령 및 학력별로 취업에 도움이 되는 교육직종을 선별할 필요성이 있다. 본 연구는 경력단절여성들의 직업훈련 교육 데이터를 활용하여 연령 및 학력별로 취업률이 높은 교육직종을 분석하였다.

ABSTRACT

Support policies such as education and training for re-employment of career-interrupted women are being implemented, but they are not being effectively employed. In addition, it is difficult for women with high educational background to re-enter, such as having to revise their previous careers or plan a new career for re-employment. In the previous studies, there was insufficient research to solve fundamental problems for re-employment, such as promising jobs with high employment opportunities. Therefore, when developing a curriculum for women with career interruptions, it was felt the need to select educational occupations that would be helpful in finding employment by age and educational background of the trainees. In this study, data on vocational training education of women with career interruptions were used to analyze the educational occupations with the highest employment rate by age and educational background.

키워드 : 빅데이터 분석, 경력단절, 취업, 직종, 시각화 분석

Keywords : Big data analysis, Career break, Employment, Occupation, Visualization analysis

Received 22 June 2021, Revised 1 July 2021, Accepted 14 July 2021

* **Corresponding Author** Choong Ho Lee(E-mail:chlee@hanbat.ac.kr, Tel:+82-42-821-1221)

Professor, Department of Information Communication Engineering, Hanbat National University, Daejeon, 34158 Korea

Open Access <http://doi.org/10.6109/jkiice.2021.25.8.1019>

print ISSN: 2234-4772 online ISSN: 2288-4165

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서론

우리나라는 여러 가지 이유로 경력이 단절된 여성의 직업능력 개발을 통한 창업 및 취업 시장의 진입 문제를 해결하기 위하여[1] 다양한 일자리 창출 프로그램 및 경제활동을 지원하고 있다. 그러나 고용 핵심 연령인 25~44세 여성 중 40% 정도가 미취업 상태로 국가적 손실이 매우 크다고 할 수 있다[2-3].

4차 산업혁명 시대에 융복합의 강점과 잠재력을 갖춘 고학력 과학기술인 여성의 수요가 증대되고 있으나, 경력단절 이후 30~40대 여성들이 재취업하는 직종 수준은 점차 낮아지고 있는 실정이다. 그럼에도 불구하고 고학력 경력단절여성의 재취업 및 창업의지는 매우 높게 나타나고 있다.[4-6].

이처럼 정책적으로는 정부지원 프로그램에 참여하는 경력단절여성의 특성을 분석하는데 어려움이 있기에 [7], 본 연구에서는 경력단절여성을 위한 교육과정 개발 시 교육 대상자의 연령 및 학력별로 취업에 도움이 되는 교육직종을 선별하고자 하였다. 이에 경력단절여성의 직업훈련 교육에 대한 데이터를 활용하여 연령 및 학력별로 취업률이 높은 교육직종을 분석하기 위한 연구를 수행하였다.

II. 분석 방법

본 연구는 파이썬 기반의 데이터 시각화 분석을 위하여 2013~2018년도 6년간의 전국 새일여성센터 경력단절여성을 위한 교육직종별 직업훈련 정형 데이터를 전자파일로 수집하여 활용하였다. 수집한 데이터는 결측치 데이터 및 이상 데이터를 처리하기 위한 데이터 크린징을 통하여 1차 가공을 하였다[8-9].

크린징 데이터의 처리결과를 이용하여 분석에 필요한 데이터의 추출 및 시각화 분석을 위한 빅데이터 분석 중 탐색적 데이터 분석(Exploratory Data Analysis : EDA) 기법을 사용하였다[10-11].

분석을 위하여 파이썬 및 Anaconda 플랫폼을 이용하였으며[12], 데이터의 추출 및 가공은 파이썬 라이브러리인 판다스(Pandas)를 사용하여 여러 데이터의 추출, 변환, 저장을 위한 프로그래밍을 하였다. Matplotlib 및 Seaborn 시각화 라이브러리를 활용하여 히트맵(Heatmap),

산점도(Pairplot) 시각화 분석을 통해 경력단절여성의 연령 및 학력별로 취업률이 높은 교육직종을 도출하였다.

분석 결과로는 연령별 취업률이 높은 상위 5개의 교육직종 분석, 학력별 취업률이 높은 상위 5개의 교육직종 분석, 마지막으로 연령 및 학력 전체 취업률 평균을 통한 취업률 상위 5개의 교육직종을 분석하였다.

2.1. 분석 절차

그림 1은 본 연구의 데이터 시각화 분석을 위한 절차이다.

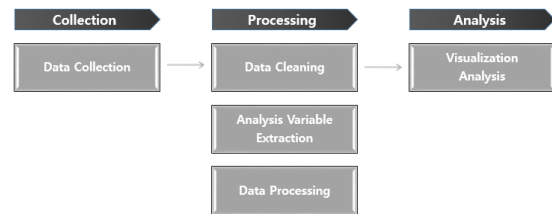


Fig. 1 Analysis Procedure

경력단절여성 직업훈련 교육 중 취업률 상위 교육직종에 대한 데이터 시각화 분석을 위하여 데이터 수집, 데이터 크린징, 분석 변수 추출, 데이터 가공, 데이터 시각화 분석 등의 절차를 통하여 연구를 진행하였다.

2.2. 분석 환경

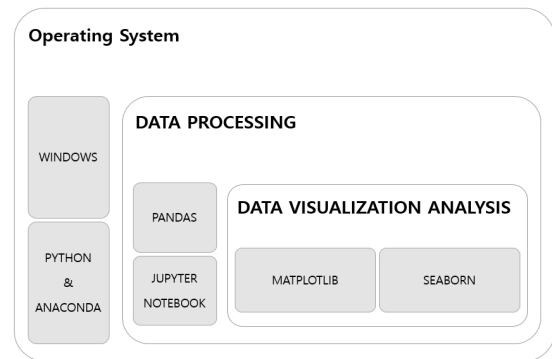


Fig. 2 Analysis Environment

그림 2는 데이터 수집, 데이터 크린징, 변수 추출, 데이터 가공 및 시각화 분석에 활용될 라이브러리로 파이썬 및 아나콘다 환경에서 판다스(Pandas) 라이브러리를 이용하여 데이터 추출 및 전처리를 하였으며, Seaborn 라이브러리를 활용하여 데이터 시각화 분석을 진행하였다.

III. 탐색적 데이터 분석

수집된 데이터는 연도별 시계열 Excel 데이터로 연도별 데이터들을 통합한 후 데이터가 없는 결측치 데이터를 확인하여 분석에 가능한 데이터로 변환하는 등의 크린징 처리를 하였다.

처리된 데이터를 이용하여 데이터 시각화 분석에 필요한 기준년도, 연령, 학력, 취업률, 교육직종 등의 데이터를 추출한 후 분석에 필요한 데이터를 추가하는 등의 변환 및 저장 작업을 진행하였다.

변환이 완료된 데이터를 통해 20대부터 60대까지의 연령별로 취업률이 가장 높은 교육직종 및 초등학교 졸업부터 대학원 박사 졸업까지의 학력별 취업률이 가장 높은 교육직종을 히트맵 시각화 분석을 통하여 분석하였다.

경력단절여성들이 참여하는 교육생들의 직종, 학력, 나이, 취업률 등의 연관성 등 상관관계를 종합적으로 분석하기 위하여 데이터의 구조적 특성을 알아내기 위한 통계적 기법으로 탐색적 데이터 분석(EDA; Exploratory Data Analysis)방법을 사용하였다.

3.1. 경력단절여성 메타 데이터 통합

```
def create_total_data(year, gubun):
    df_total = pd.DataFrame()
    for year in range(2013, 2019, 1):
        for gbn in range(0, 2, 1):
            try:
                if gbn == 0 :
                    gubun = "First-Half"
                elif gbn == 1:
                    gubun = "Second-Half"
                temp = create_total_data(str(year),
                                         gubun)
                df_total = df_total.append(temp,
                                           ignore_index=True)
            except:
                pass
```

Fig. 3 Data Integration

년도 및 반기별로 분리되어있는 수집 데이터를 분석이 용이 하도록 하나의 데이터로 통합하기 위하여 그림 3과 같이 모든 Excel 데이터를 자동으로 통합하기 위한 전처리 함수(Function)를 생성하였다.

통합 데이터를 저장하기 위하여 DataFrame을 사용하여 2013년부터 2018년도 데이터까지 모든 데이터를 자동으로 통합하였다.

3.2. 연령 및 학력별 데이터 크린징

통합된 데이터에서 분석 데이터를 추출하기 전에 데이터가 없거나 잘못된 값이 포함되어 있는지를 확인하여 데이터 분석이 가능한 형태의 데이터로 처리하였다.

그림 4는 연령 및 학력 데이터의 취업률 데이터 중 데이터가 없는 결측 데이터 및 정수형태의 데이터에 NaN의 이상데이터가 존재하였으며, 이를 처리하기 위하여 Pandas의 fillna() 함수를 이용하여 결측치 및 이상데이터를 0으로 치환하는 크린징 처리를 진행하였다.

또한, 데이터 분석에 사용하기 어려운 항목 이름을 변환하였으며, 불필요한 항목은 삭제하는 등의 변환을 위하여 rename() 및 drop() 함수를 이용하여 처리하였다.

```
df_age_employ = df_age_employ.fillna(0)
age_all_Mean.rename(columns=(0:rename),
                    inplace=True)
age_all_Mean_finally.drop(['index'],
                          axis=1,
                          inplace=True)

df_edu_employ = df_edu_employ.fillna(0)
edu_all_Mean.rename(columns=(0:rename),
                    inplace=True)
edu_all_Mean_finally.drop(['index'],
                          axis=1,
                          inplace=True)
```

Fig. 4 Data Cleansing

3.3. 데이터 시각화 분석

연령 및 학력별 정제 데이터를 이용하여 연령별 취업률 상위 교육직종, 학력별 취업률 상위 교육직종에 대한 시각화 분석을 하였다.

3.3.1. 연령 및 교육직종별 취업률 시각화 분석

그림 5는 20 ~ 60대 중 40대 연령에 대한 연도별(Y축), 직종별(X축) 취업률(박스의 정수값으로 명함 진하기를 표시함)을 종합적으로 분석하기 위하여 히트맵 시각화를 수행한 결과이다.

20대의 경우 화학기계 및 화학제품제조 직종에 높은 취업률을 나타내고 있으며, 30대는 통신설비운용, 무역, 가구·목공예·도자기공예 및 사원층 직종에 높은 취업률을 보이고 있다.

40대의 경우 전체 직종에 취업률이 고르게 분포하고 있으며, 주로 운송장비 제조, 건축, 금융·보험, 사무, 자동차 및 프로그램 직종에 높은 취업률을 보이고 있다. 마지막으로 50대 및 60대의 경우 섬유가공, 음식료품 가공, 운전, 의료보조 등의 직종에 높은 취업률을 보이고 있다.

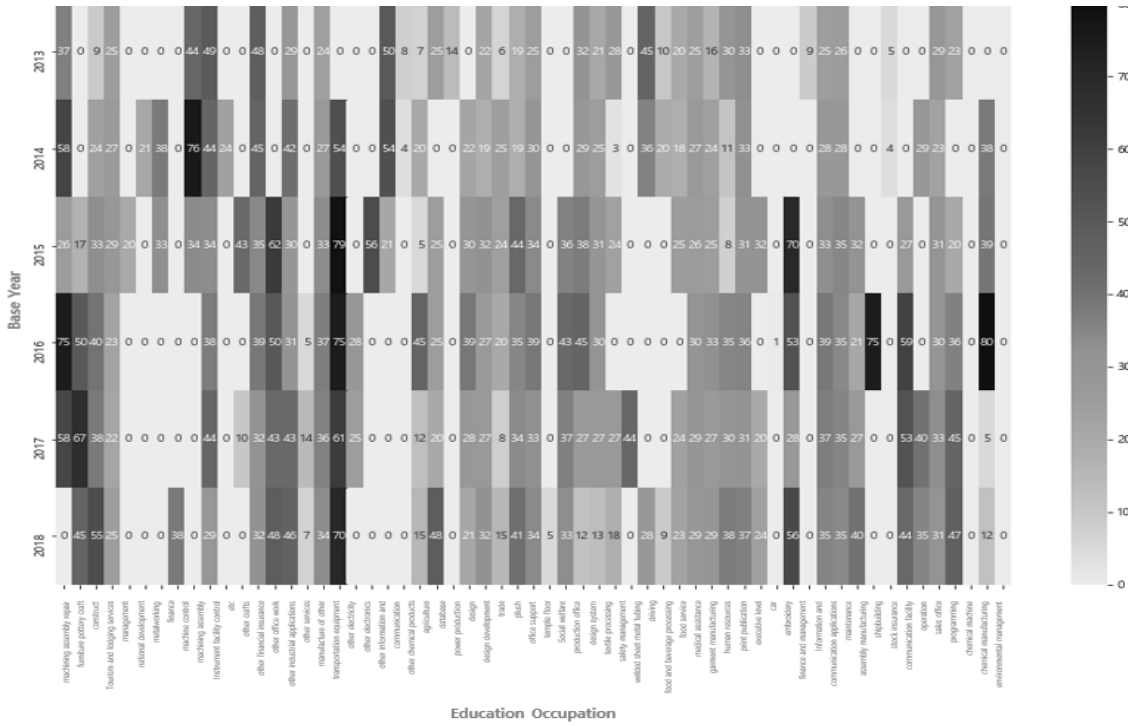


Fig. 5 Employment rate by occupation in their 40s

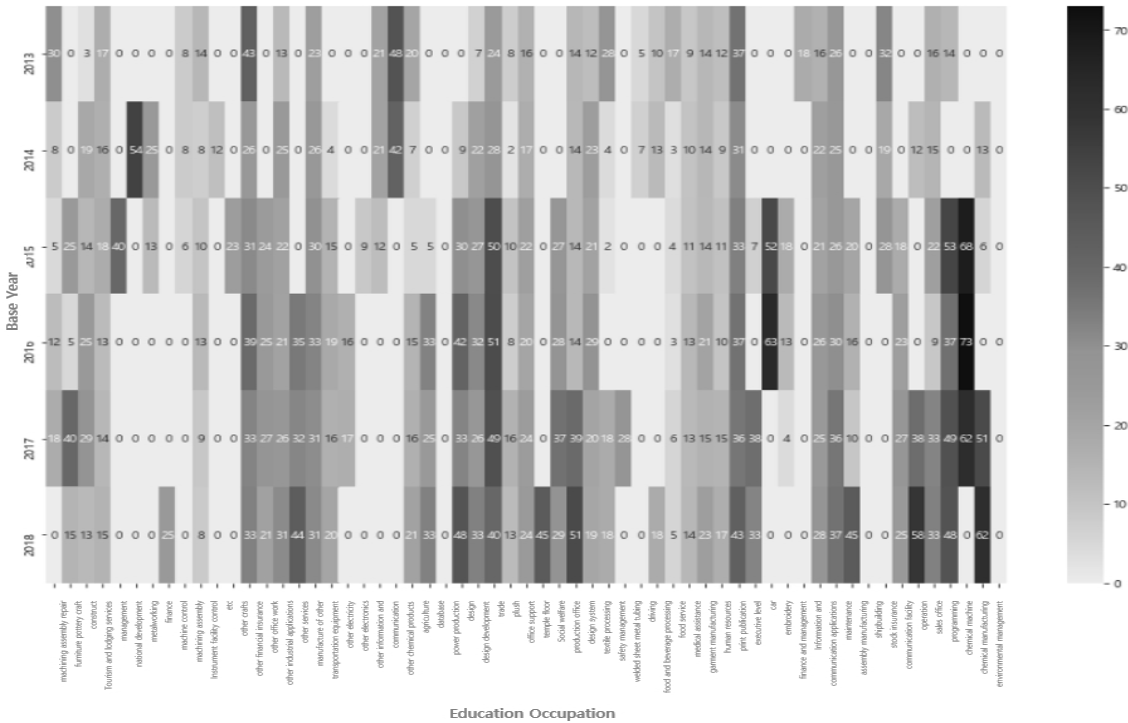


Fig. 6 employment rate by job type according to educational background (4-year college graduation)

3.3.2. 학력 및 교육직종별 취업률 시각화 분석

그림 6은 초등학교 졸업부터 대학원 박사학위를 취득한 학력 중 대학교(4년)에 대한 연도별(Y축), 직종별(X축) 취업률(박사의 정수값으로 명함 진하기를 표시함)을 종합적으로 분석하기 위하여 히트맵 시각화를 수행한 결과이다.

초등학교 및 중학교 졸업 여성의 경우 가구·목공예·도자기공예, 운송장비 제조 및 음식료가공 직종에 높은 취업률을 나타내고 있으며, 고등학교 졸업 여성의 경우에는 가공·조립·수립, 섬유가공, 음식료품가공 및 자동차 직종에 높은 취업률을 보이고 있다.

대학(2~3년) 졸업자의 경우에는 건축, 금융·보험, 운송장비 제조, 안전관리 등의 직종에 높은 취업률을 나타내고 있으며, 대학교(4년) 졸업 여성의 경우에는 통신설비운용, 화학제품제조, 산업응용, 생산사무, 디자인, 프로그래밍 및 인사노무 등의 사무직 직종에 높은 취업률을 보이고 있다.

대학원 석사 및 박사 졸업 여성의 경우에는 다소 차이는 있지만, 인사노무, 데이터베이스, 프로그래밍, 사회복지 등의 직종에 높은 취업률을 보였으나, 고학력에 따른 다양한 직종에 취업이 어려움이 있는 것으로 나타났다.

IV. 분석 결과

4.1. 연령별 취업률 상위 교육직종 분석

20 ~ 60대 연령별 취업률 교육직종 시각화 분석에 따라 각 연령별로 취업률이 높은 교육직종 분석을 하였으며, 전체 연령 중 취업률이 가장 높은 교육직종은 표 1과 같이 분석되었다.

Table. 1 Top 5 education occupations with the highest employment rate among all age groups

Ranking	Education Occupation	Employment Rate
1	Manufacture of other transportation equipment	57.0
2	Machining, assembly, repair	42.0
3	Machining, assembly	40.0
4	Other crafts	38.0
5	Other office work	37.0

4.2. 학력별 취업률 상위 교육직종 분석

초등학교 졸업에서 대학원 박사 졸업까지의 학력별 취업률 교육직종 시각화 분석에 따라 각 학력별로 취업률이 높은 교육직종에 대한 분석을 하였으며, 전체 학력 중 취업률이 가장 높은 교육직종은 표 2와 같이 분석되었다.

Table. 2 Top 5 education occupations with the highest employment rate among all educational groups

Ranking	Education Occupation	Employment Rate
1	Machining assembly	49.0
2	Machining, assembly, repair	48.0
3	Manufacture of other transportation equipment	45.0
4	Trade	40.0
5	Sewing	40.0

4.3. 경력단절여성 취업률 상위 교육직종 분석

연령 및 학력 전체에 대한 교육직종별 취업률을 평균한 결과 기타 운송장비 제조 직종이 가장 높은 취업률을 나타내는 것으로 분석되었으며, 다음으로 기계 가공·조립, 자동차, 기타 사무 및 공예 순으로 분석되었다.

표 3은 연령 및 학력 전체에 대한 취업률 상위 교육직종에 대한 분석 결과이며, 그림 7은 취업률에 따른 교육직종에 대한 산점도 분석 결과이다.

Table. 3 Top 5 education occupations with the highest employment rate for all ages and educational background

Ranking	Education Occupation	Employment Rate
1	Manufacture of other transportation equipment	51.0
2	Machining assembly	44.0
3	Car	39.0
4	Other office work	37.0
5	Other crafts	31.0

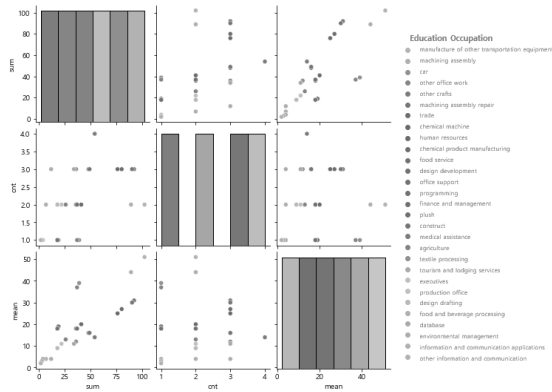


Fig. 7 Results of scatterplot analysis of education occupations with the highest employment rate for all age and educational background

V. 결 론

본 연구는 경력단절여성을 위한 연령 및 학력별 취업률이 높은 교육직종에 대한 빅데이터 시각화 분석을 위하여 경력단절여성을 위한 교육과정의 취업 직종 데이터를 활용하여 분석하였다.

연령 및 학력에 대한 취업률이 가장 높은 교육직종을 종합적으로 분석한 결과, 기타 운송장비 제조 직종이 가장 높은 취업률을 보였으며, 기계 가공 조립, 자동차, 기타 사무 및 공예 순으로 분석되었다.

본 연구 결과를 활용하여 경력단절여성을 위한 교육과정 편성 시 교육 대상 여성의 연령 및 학력을 고려하여 취업률이 높은 직종에 대한 교육과정을 편성한다면 좀 더 효과적인 일자리창출이 될 것으로 기대된다.

본 연구의 시각화 분석 방법을 활용하여 정부지원의 취업지원과정 내 청소년, 고경력자, 재취업자 등 다양한 교육집단에 적용하여 취업률에 영향을 미치는 요인을 분석하고자 한다.

ACKNOWLEDGEMENT

This research was supported by the research fund of Hanbat National University in 2021.

REFERENCES

- [1] M. A. Kook, "Career-Interrupted women, why did they quit their jobs?," *Journal of Women's Studies*, vol. 28, no. 1, pp. 251-282, Apr. 2018.
- [2] N. R. Bae, "A study on the support for the re-employment of women with career break in life cycle," *Journal of Korea Academia-Industrial cooperation Society*, vol. 19, no. 8, pp. 268-275, 2018.
- [3] S. Y. Jang, H. K. Eun, and O. M. Hong, "A study on promising jobs for reentry women to labor market," *Korea Research Institute for Vocational Education and Training*, vol. 12, no. 1, pp. 163-187, Dec. 2019.
- [4] R. G. Park and J. Y. Yun, "Development of group counseling program to support for woman scientists and technicians with career breaks: a preliminary study," *The Journal of the Korea Contents Association*, vol. 20, no. 2, pp. 636-648, 2020.
- [5] N. J. Kim, "Analysis of survey on economic activities of career-break women," Korean Women's Development Institute, Research Report, 2019.
- [6] J. H. Kim and S. H. Byun, "The relationship of among of vocational education, entrepreneurial intention, re-employment need of career interrupted woman," *Asia-Pacific Journal of Business Venturing and Entrepreneurship*, vol. 11, no. 4, pp. 237-245, 2016.
- [7] Y. K. Kim, "Research on the way to enhance the employment rate of the vocational training institution," *The Korean Society for the Study of Vocational Education*, vol. 23, no. 3, pp. 165-180, 2004.
- [8] J. A. Jeong, K. H. Lee, and H. K. Jung, "Prediction Model for Unpaid Customers Using Big Data," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 24, no. 7, pp. 827-833, Jul. 2020.
- [9] Y. M. Kang, J. G. Yoo, K. S. Han, and J. B. Kim, "A Study of Business Model based on Big Data," *Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, vol. 6, no. 5, pp. 1-10, May. 2016.
- [10] M. H. Huh and J. W. Jung, "Software review of statistical package programs on EDA aspects," *The Korean Journal of Applied Statistics*, vol. 3, no. 2, pp. 17-25, 1990.
- [11] J. Y. Jung, "Big data analysis on 112 report data : Focusing on the EDA technique," *Korean Security Journal*, vol. 66, pp. 71-92, Mar. 2021.
- [12] Anaconda individual edition [Internet]. Available: <http://www.anaconda.com/distribution/>.

경력단절여성을 위한 연령 및 학력별 취업률이 높은 교육직종 빅데이터 시각화 분석



이정원(Jeongwon Lee)

공주대학교 교육대학원 컴퓨터교육과 교육학석사
공주대학교 대학원 컴퓨터교육과 박사수료
한밭대학교 정보통신전문대학원 정보통신공학과 박사과정
※관심분야 : 데이터 분석, 데이터베이스, 인공지능, 응용소프트웨어, 인적자원개발



이충호(Choong Ho Lee)

연세대학교 전자공학과 공학사
연세대학교 대학원 전자공학과 공학석사
도호쿠대학 대학원 정보과학연구과 공학박사
KT 멀티미디어연구소 전임연구원
현)한밭대학교 정보통신공학과 교수
※관심분야 : 디지털신호처리, 영상처리, 패턴인식, 인공지능, 응용소프트웨어