

과제 3. 최종 보고서

201821479 황혜린

1. 주제 선정 배경

워크 앤 라이프 밸런스(Work and Life Balance), 워라밸을 아십니까? 워라밸이란 일과 삶의 균형을 의미하는 단어로, 이 말이 한국에 알려지기 시작한 것은 2018 년 즈음입니다. 그 전의 우리나라는 OECD 국가 중에서 상위권에 해당하는 긴 근무시간, 또한 거기에 더불어 잦은 야근까지 많은 직장인들은 하루 대부분의 시간을 노동에 할애하며 워라밸과 반대되는 삶을 살았습니다. 최근에서야 비로소 개인의 영역과 여유로운 삶을 중요하게 생각하는 사회적 분위기가 형성되면서 청년들은 워라밸에 관심을 가지게 되었습니다. 점점 관심이 고조되는 가운데 2019 년 통계청의 자료조사에서 "일·가정 모두 중요" 응답자가 "일이 우선" 응답자를 처음으로 넘으면서 "일을 우선시하던 사회에서 일과 가정생활의 균형을 중요시하는 사회로 변화하고 있다"고 분석할 만큼, 현재 워라밸은 직장인들이 중요시하는 요소가 되었습니다. 하지만 직장생활을 하는 한국인중에서 현실이 워라밸에 가깝다고 생각하는 사람은 노동인구의 9.5%밖에 되지 않아, 아직까지도 워라밸을 누리고 있는 직장인들은 얼마되지 않음을 알 수 있습니다. 따라서 저는 워라밸과 상호관계를 가지는 요소들이 무엇인지 파악하여 워라밸을 누리고 보장하고자 하는 개인과 기업이 어떻게 전보다 쉽게 목표에 도달할 수 있는지 분석해보고자 합니다. 근무시간 이외에 워라밸에 영향을 주는 요소는 무엇인지, 근로의 어떤 요소와 관련이 있는지 이번 프로젝트를 통해 알아볼 것입니다.

[참고]

<https://data.oecd.org/emp/hours-worked.htm>

<https://www.fnnews.com/news/201911251137478207>

<https://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=001&oid=277&aid=0004153687>

<https://namu.wiki/w/%EC%9B%8C%EB%9D%BC%EB%B0%B8#rfn-2>

2. 데이터 수집

마이크로 데이터 통합 서비스 사이트(<https://mdis.kostat.go.kr>)에서 “2019 년 사회조사_복지.사회참여.문화와여가.소득과소비.노동” 자료를 다운로드 하였습니다.

● 조사 목적

삶의 질과 관련된 국민의 사회적 관심사와 주관적 의식에 관한 사항을 파악하여
관련 정책의 수립 및 연구의 기초자료로 제공

- 조사 대상

약 27,336 표본가구의 만 13 세 이상 상주 가구원
(1,548 개 조사구, 조사구당 16~20 가구)

- 조사 시기

2019.5.15~5.30

- 조사 항목

- 기본항목(성별, 연령, 교육정도 등)과 10 개 부문을 2 년 주기로 5 개
부문씩 선정하여 조사
- 2019 년: 기본항목(12), 복지(17), 사회참여(17), 문화와 여가(11), 소득과
소비(6), 노동(11)

3. 데이터 정제

R studio 에서 파일을 불러옵니다. 원본 파일중에서 제가 사용할 컬럼만 선택하여
데이터를 다시 생성합니다. 데이터는 아직 머리글행이 없으므로 데이터와 함께 다운받은
조사표를 참고하여 머리글행 이름을 부여합니다.

다음 차례는 데이터가 가지는 속성에 맞게 형식을 변형해주는 것입니다. 저는 명목형
데이터를 가지는 컬럼을 factor 로 지정하고 컬럼에 해당하는 값들은 “남자”, “여자”와
같이 labels 를 통해 변환해 주었습니다.

직장을 가지지 않은 사람들에 의해 NA 로 표시된 값들은 분석과 연관이 없으므로
제거합니다. 또한 필수항목에서 NA 로 나타난 값은 불성실한 응답에 의한 것이므로
제거합니다.

3.1 요약 통계

일과 가정의 우선도를 묻는 문항에서 응답이 "일"에 많이 치우치긴 했지만, 둘 다
중요하다는 응답이 6763 으로 가장 많았습니다. 대부분의 문항들과는 다르게
“소득만족도”와 “근로여건만족도_임금”에 관한 항목에서는 응답자들이 만족하지 않는다고
답했습니다. 일자리에 대한 만족도는 만족과 보통에 많은 응답이 몰려 있으므로, 2019 년
사회조사에서는 보다 많은 사람들이 본인의 일자리에 만족하고 있음을 알 수 있습니다.

4. 데이터 탐색

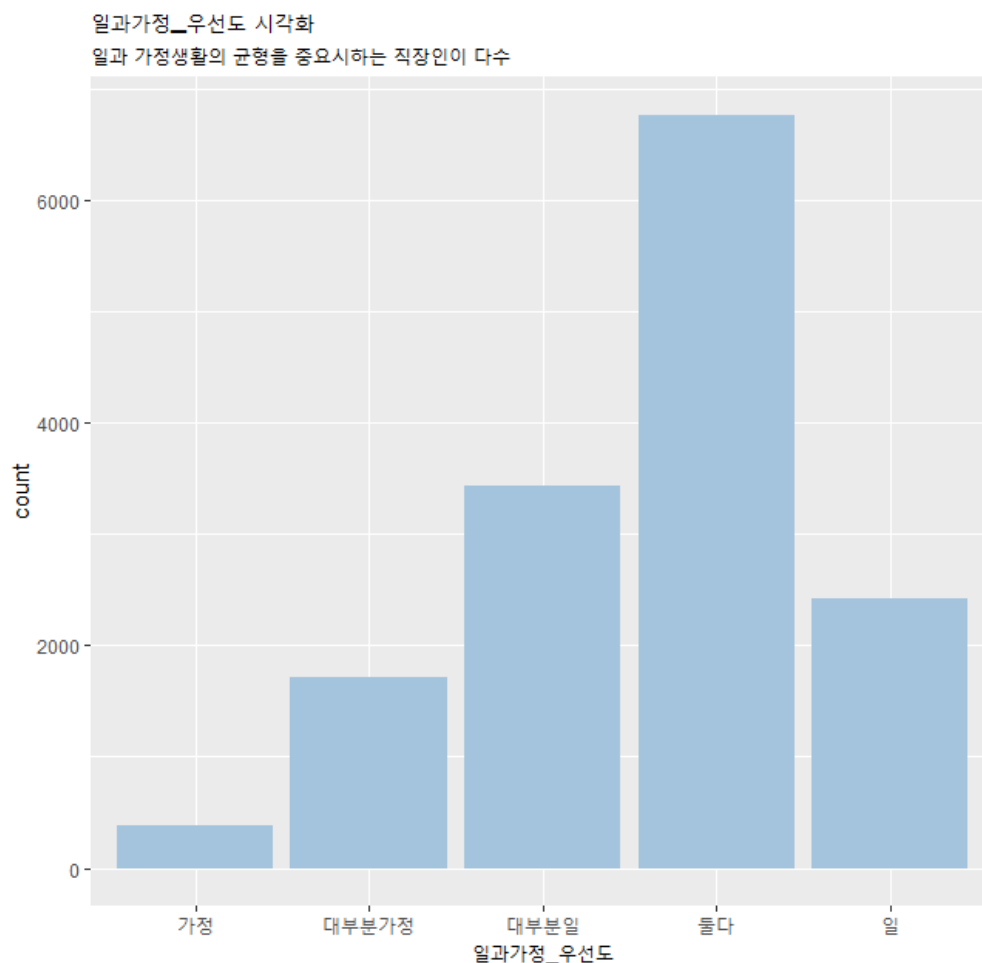
4.1 탐색 목적

수집한 데이터를 분석하기 전에 데이터를 시각화 하여 탐색하는 과정을 거쳤습니다.

주제와 관련된 변수들 간의 관계를 그래프로 시각화 하고자 하였습니다.

4.2 데이터 시각화 탐색

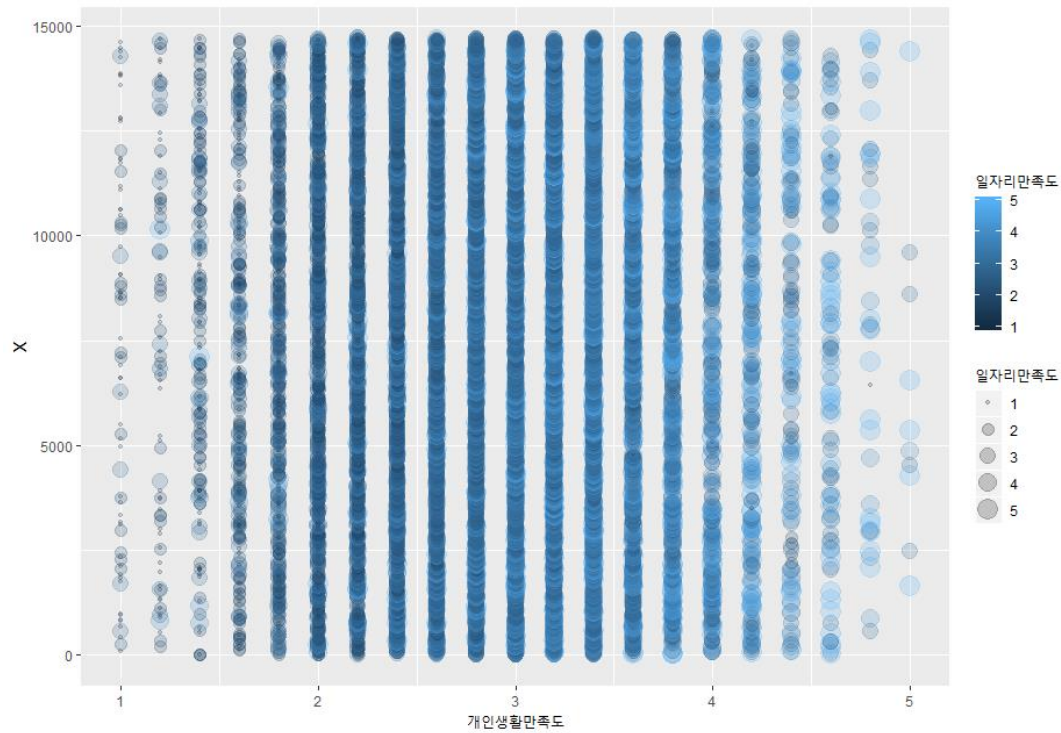
- 일과 가정_우선도 시각화



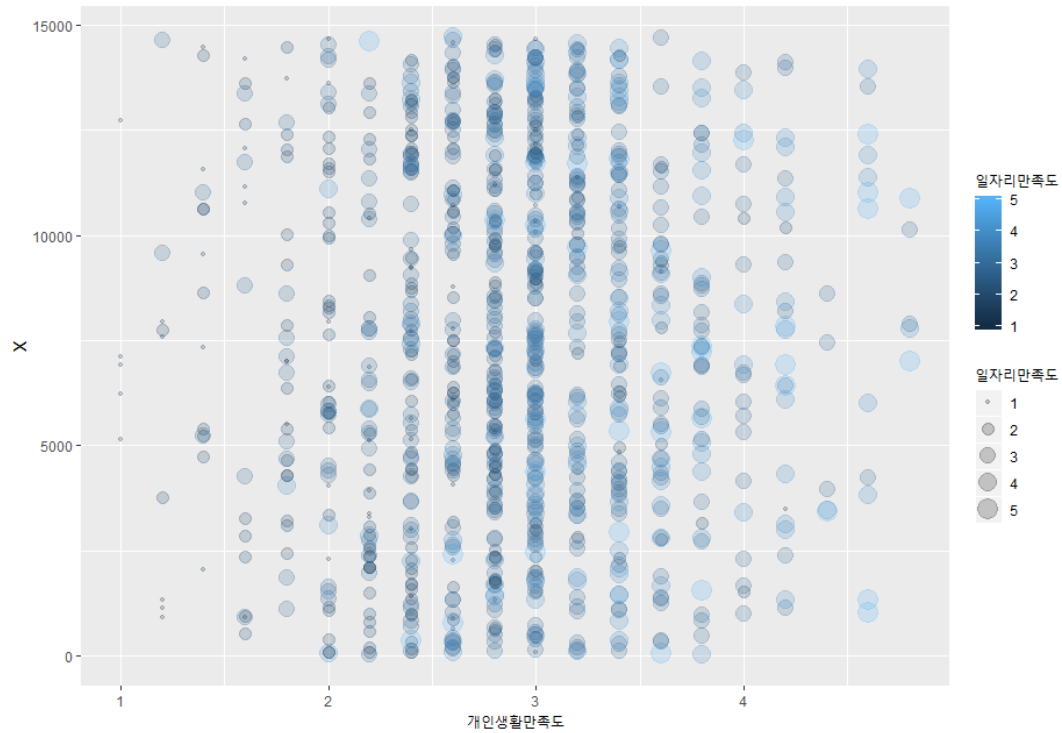
주제 선정 배경에서 "일.가정 모두 중요" 응답자가 "일이 우선" 응답자를 처음으로 넘으면서 "일을 우선시하던 사회에서 일과 가정생활의 균형을 중요시하는 사회로 변화하고 있다"고 서술한 내용을 직접 그래프로 나타내어 보았습니다. 일과 가정 둘 다 중요하다고 선택한 직장인의 수가 다른 항목보다 두배가량 높은 것을 알 수 있습니다.

- 개인생활 만족도와 일자리만족도의 관련성 시각화

시각화하기에 앞서 분석에 사용하기 위해 개인의 생활에 관련된 컬럼들의 평균을 내어 "개인생활만족도"라는 새로운 컬럼을 만들었습니다.



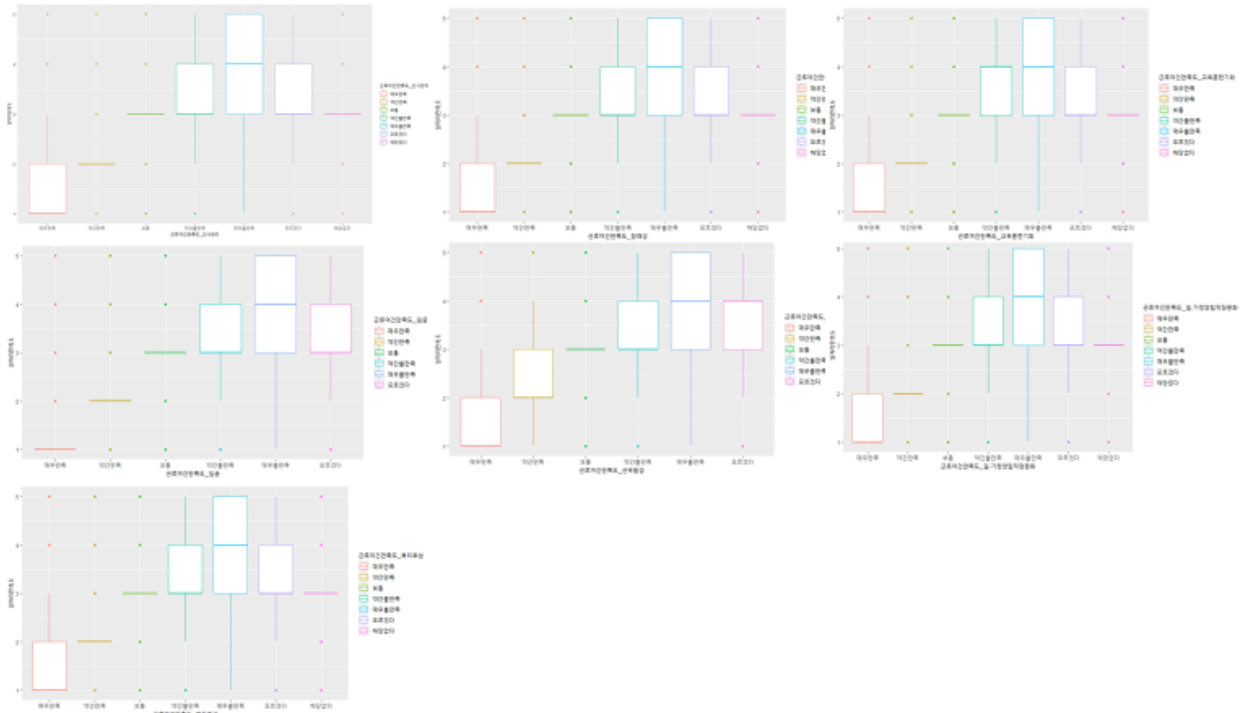
일자리만족도 버블의 색깔과 크기를 통해 개인생활만족도가 낮아질수록 일자리만족도가 낮아지는 경향을 확인할 수 있으나, 데이터 수가 너무 많아 분석하기 힘듭니다. 따라서 14713 개의 데이터 중에서 1000 개를 임의추출하여 그래프를 다시 그려보았습니다.



개인생활만족도에 해당하는 숫자가 커질수록 일자리만족도 버블의 색깔이 푸르게 변하고 크기가 커지는 경향을 확인할 수 있습니다. 따라서 개인생활 만족도와 일자리만족도는 연관이 있습니다.

● 근로 여건의 분야에 따른 근로 여건 만족도 분포

그래프를 분석할 때 한눈에 파악하기 위해 데이터를 추가로 정제하였습니다. 그래프의 X 축에 오는 데이터 값을 명목형으로 변환하였습니다. 그리고 분홍색으로 나타나는 박스플롯은 근로여건 만족도가 “해당없음”에 관한 항목이므로 분석에서 제외하였습니다.



위는 비슷한 분포를 가지는 그래프를 모아놓은 것입니다. 각각 인사관리, 임금, 복리후생, 장래성, 근무환경, 일·가정 양립 직장문화, 교육훈련기회에 해당하는 그래프입니다. 근로여건에 대한 만족도가 상승하면서 일자리만족도 또한 상승하지만, 근로여건만족도가 매우 불만족인 지점에서는 일자리만족도가 보통에서 매우 불만족 사이에 분포하는 것으로 나타났습니다.

두개의 변수가 서로 얼마나 상관관계를 갖는지 분석하기 위해 상관분석기법을 사용하였습니다. 가설은 다음과 같습니다.

1.	‘일·가정 양립 직장문화에 대한 만족도’는 ‘주관적만족감’과 상관관계를 갖는가
2.	‘일·가정 양립 직장문화에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
3.	‘개인생활만족도’는 ‘일자리만족도’와 상관관계를 갖는가
4.	‘하는 일에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
5.	‘인사관리에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
6.	‘임금에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
7.	‘복리후생에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
8.	‘장래성에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
9.	‘근무환경에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
10.	‘인간관계에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
11.	‘근로시간에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
12.	‘폭력방지노력에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가
13.	‘교육훈련기회에 대한 만족도’는 ‘일자리만족도’와 상관관계를 갖는가

1~2 에 대한 가설은 일·가정 양립 직장문화에 대한 만족도가 개인의 주관적만족감과 일자리만족도와 얼마나 상관관계가 있는지 파악하여 직장문화 개선에 대한 뒷받침을 하기 위해 설정하였습니다.

3 에 대한 가설은 데이터 탐색 과정에서 연관성을 발견한 두가지 항목에 대해 상관분석을 해보기 위해 설정하였습니다.

4~13 에 대한 가설은 근로여건 만족도에 해당하는 항목들이 일자리만족도와 양의 상관관계를 가지며, 어떤 항목이 보다 큰 상관관계를 가지는지 파악하여 위라벨을 높이고자 할 때 신경 써야 할 항목들을 분석하기 위해 설정하였습니다.

- 분석기법 적용 결과 분석

1. '일·가정 양립 직장문화에 대한 만족도'는 '주관적만족감'과 상관관계를 갖는가
2. '일·가정 양립 직장문화에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가

```
> corr.test(data$근로여건만족도_일.가정양립직장문화, data$주관적만족감)
Call:corr.test(x = data$근로여건만족도_일.가정양립직장문화, y = data$주관적만족감)
Correlation matrix
[1] 0.24
Sample Size
[1] 14713
Probability values adjusted for multiple tests.
[1] 0

To see confidence intervals of the correlations, print with the short=FALSE option
> corr.test(data$근로여건만족도_일.가정양립직장문화, data$일자리만족도)
Call:corr.test(x = data$근로여건만족도_일.가정양립직장문화, y = data$일자리만족도)
Correlation matrix
[1] 0.53
Sample Size
[1] 14713
Probability values adjusted for multiple tests.
[1] 0
```

일과 가정 모두 중요시하는 직장문화에 대한 만족도는 주관적만족감과 일자리만족도와 양의 상관관계가 있습니다. 앞서 그래프 1 번에서 일과 가정 모두 중요하다고 선택한 직장인의 수가 많았던 만큼 그것을 존중해주는 직장문화가 조성되면 개인의 주관적만족감과 일자리만족도가 높아질 것입니다.

3. '개인생활만족도'는 '일자리만족도'와 상관관계를 갖는가

```
> corr.test(data$개인생활만족도, data$일자리만족도)
Call:corr.test(x = data$개인생활만족도, y = data$일자리만족도)
Correlation matrix
[1] 0.49
Sample Size
[1] 14713
Probability values adjusted for multiple tests.
[1] 0
```

데이터 탐색과정에서 개인생활만족도와 일자리만족도의 연관성을 발견하였습니다. 따라서 두가지 요인을 상관분석 해본 결과 높은 양의 상관관계를 가지는 것으로 분석되었습니다. 이는 일자리만족도가 높아질 수록 개인생활만족도 또한 높아진다는 것을 의미합니다.

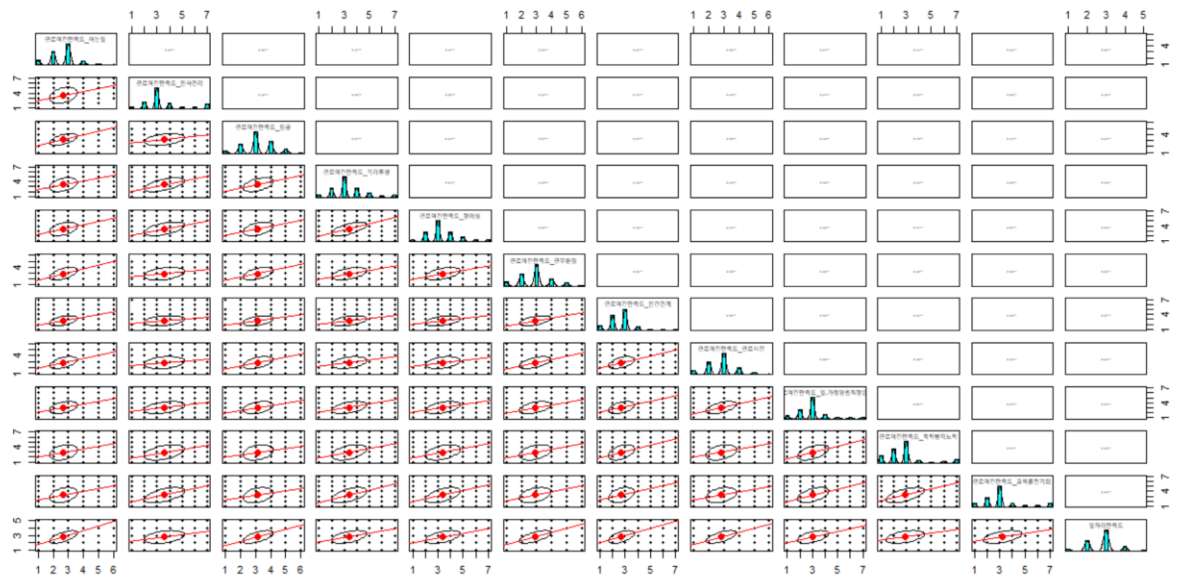
4. '하는 일에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
5. '인사관리에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
6. '임금에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
7. '복리후생에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
8. '장래성에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
9. '근무환경에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
10. '인간관계에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
11. '근로시간에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
12. '폭력방지노력에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가
13. '교육훈련기회에 대한 만족도'는 '일자리만족도'와 상관관계를 갖는가

```
> corr.test(data_verify[,c(2:9,11:13)])
Call:corr.test(x = data_verify[, c(2:9, 11:13)])
Correlation matrix
```

	근로여건만족도_하는일	근로여건만족도_인사관리	근로여건만족도_임금	근로여건만족도_복리후생	근로여건만족도_장래성	
근로여건만족도_하는일	1.00	0.35	0.56	0.37	0.42	
근로여건만족도_인사관리	0.35	1.00	0.36	0.58	0.55	
근로여건만족도_임금	0.56	0.36	1.00	0.47	0.43	
근로여건만족도_복리후생	0.37	0.58	0.47	1.00	0.63	
근로여건만족도_장래성	0.42	0.55	0.43	0.63	1.00	
근로여건만족도_근무환경	0.54	0.34	0.51	0.48	0.48	
근로여건만족도_인간관계	0.47	0.31	0.41	0.39	0.39	
근로여건만족도_근로시간	0.50	0.26	0.50	0.36	0.33	
근로여건만족도_폭력방지노력	0.32	0.44	0.32	0.44	0.41	
근로여건만족도_교육훈련기회	0.34	0.56	0.32	0.55	0.52	
일자리만족도	0.66	0.38	0.63	0.49	0.49	
근로여건만족도_근무환경	0.54	0.47	0.50	0.32	0.34	
근로여건만족도_인사관리	0.34	0.31	0.26	0.44	0.56	
근로여건만족도_임금	0.51	0.41	0.50	0.28	0.32	
근로여건만족도_복리후생	0.48	0.39	0.36	0.44	0.55	
근로여건만족도_장래성	0.48	0.39	0.33	0.41	0.52	
근로여건만족도_근무환경	1.00	0.54	0.55	0.38	0.40	
근로여건만족도_인간관계	0.54	1.00	0.50	0.42	0.39	
근로여건만족도_근로시간	0.55	0.50	1.00	0.35	0.34	
근로여건만족도_폭력방지노력	0.38	0.42	0.35	1.00	0.61	
근로여건만족도_교육훈련기회	0.40	0.39	0.34	0.61	1.00	
일자리만족도	0.65	0.57	0.63	0.44	0.48	
근로여건만족도_하는일	0.66					
근로여건만족도_인사관리	0.38					
근로여건만족도_임금	0.63					
근로여건만족도_복리후생	0.49					
근로여건만족도_장래성	0.49					
근로여건만족도_근무환경	0.65					
근로여건만족도_인간관계	0.57					
근로여건만족도_근로시간	0.63					
근로여건만족도_폭력방지노력	0.44					
근로여건만족도_교육훈련기회	0.48					
일자리만족도	1.00					

일자리만족도와 각 항목별 근로여건만족도는 모두 양의 상관관계를 가지는 것으로 분석되었습니다. 일자리만족도와 보다 높은 상관관계를 가지는 근로여건만족도는 순서대로 하는 일, 근무환경, 폭력방지노력=임금, 근로시간, 인간관계, 복리후생=장래성, 교육훈련기회, 인사관리입니다. 기업에서는 분석한 자료를 참고하여 근로여건을 개선한다면 직장인들의 일자리만족도를 향상시키는데 도움이 될 것입니다. 앞서 일자리만족도는 개인생활만족도와 양의 상관관계를 가진다는 것을 분석하였으므로 기업이 일자리만족도를 향상시키면 개인생활만족도도 향상되어 전보다 높은 워라벨을 누릴 수 있습니다.

● 상관계수 그래프



위의 상관분석에서 얻은 상관계수 그래프입니다. 모든 항목에서 높은 상관관계를 가짐을 알 수 있습니다.

● 편상관분석

```
> partial.r(cor(data_verify), c(2:13), 1)
partial correlations
```

	근로여건만족도_하늘일	근로여건만족도_인사관리	근로여건만족도_임금	근로여건만족도_복지후생	근로여건만족도_장래성
근로여건만족도_하늘일	1.00	0.24	0.43	0.24	0.30
근로여건만족도_인사관리	0.24	1.00	0.25	0.52	0.49
근로여건만족도_임금	0.43	0.25	1.00	0.36	0.31
근로여건만족도_복지후생	0.24	0.52	0.36	1.00	0.57
근로여건만족도_장래성	0.30	0.49	0.31	0.57	1.00
근로여건만족도_근무환경	0.44	0.24	0.40	0.39	0.39
근로여건만족도_인간관계	0.37	0.23	0.29	0.30	0.30
근로여건만족도_근로시간	0.41	0.16	0.39	0.27	0.23
근로여건만족도_일, 가정양립직장문화	0.31	0.34	0.28	0.39	0.38
근로여건만족도_폭력방지노력	0.22	0.38	0.17	0.37	0.35
근로여건만족도_교육훈련기회	0.23	0.50	0.20	0.49	0.46
일자리만족도	0.56	0.27	0.51	0.38	0.39

	근로여건만족도_근무환경	근로여건만족도_인간관계	근로여건만족도_근로시간	근로여건만족도_일, 가정양립직장문화
근로여건만족도_하늘일	0.44	0.37	0.41	0.31
근로여건만족도_인사관리	0.24	0.23	0.16	0.34
근로여건만족도_임금	0.40	0.29	0.39	0.28
근로여건만족도_복지후생	0.39	0.30	0.27	0.39
근로여건만족도_장래성	0.39	0.30	0.23	0.38
근로여건만족도_근무환경	1.00	0.47	0.48	0.37
근로여건만족도_인간관계	0.47	1.00	0.43	0.37
근로여건만족도_근로시간	0.48	0.43	1.00	0.49
근로여건만족도_일, 가정양립직장문화	0.37	0.37	0.49	1.00
근로여건만족도_폭력방지노력	0.31	0.36	0.28	0.47
근로여건만족도_교육훈련기회	0.31	0.32	0.26	0.45
일자리만족도	0.57	0.49	0.56	0.45

	근로여건만족도_폭력방지노력	근로여건만족도_교육훈련기회	일자리만족도
근로여건만족도_하늘일	0.22	0.23	0.56
근로여건만족도_인사관리	0.38	0.50	0.27
근로여건만족도_임금	0.17	0.20	0.51
근로여건만족도_복지후생	0.37	0.49	0.38
근로여건만족도_장래성	0.35	0.46	0.39
근로여건만족도_근무환경	0.31	0.31	0.57
근로여건만족도_인간관계	0.36	0.32	0.49
근로여건만족도_근로시간	0.28	0.26	0.56
근로여건만족도_일, 가정양립직장문화	0.47	0.45	0.45
근로여건만족도_폭력방지노력	1.00	0.57	0.36
근로여건만족도_교육훈련기회	0.57	1.00	0.39
일자리만족도	0.36	0.39	1.00

개인생활만족도가 일자리만족도와 상관관계가 있음을 알게 되었으므로 근로여건 만족도와 일자리만족도를 상관분석할 때 개인생활만족도를 통제하여 편상관분석을 시행하였습니다. 상관계수가 작아졌지만 편상관분석을 거쳐도 개인생활만족도와 일자리만족도는 양의 상관관계에 있음을 알 수 있습니다.

5.2 연관분석

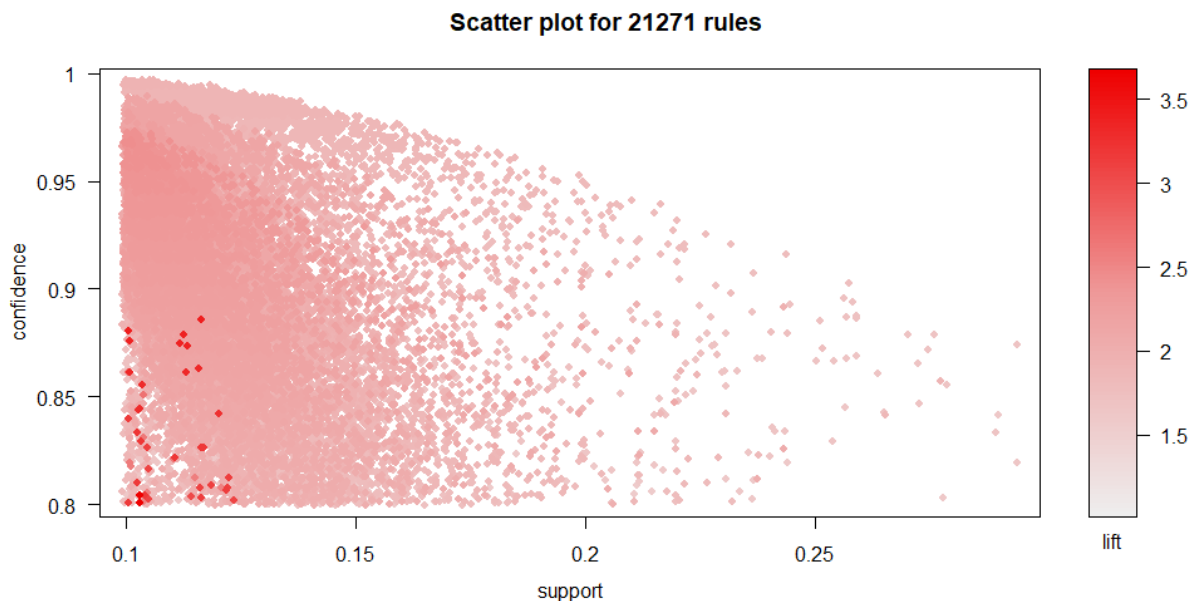
연관분석을 하기에 앞서 분석할 데이터를 범주형으로 변환하고 분석절차에 들어갔습니다.

```
> rule <- apriori(adata, parameter=list(minlen=2))
```

apriori 함수를 사용하여 연관분석을 시행하였습니다. 파라미터 값을 따로 설정하지 않은 이유는 디폴트값으로 분석을 시행하였을 때 응답규칙이 적게 나올 것을 우려하여 support=0.025, confidence=0.2 로 설정하여 실행해보았으나 해당 파라미터 값은 응답규칙의 수가 너무 많아 코드를 실행하는데 오랜 시간이 걸리고 디폴트 값으로 파라미터를 설정해도 충분히 많은 응답규칙을 찾을 수 있기 때문입니다.

```
> rule.df <- as(rule, 'data.frame')
```

분석결과를 쉽게 파악하기 위해 data frame 형태로 응답규칙을 저장하였습니다.



plot 함수를 사용하여 그래프로 나타내 보았습니다. 데이터의 수가 많기 때문에 그래프에 많은 연관규칙들이 표현되었습니다. lift 값이 큰 응답규칙은 지지도가 다소 낮으며 신뢰도는 0.8 과 0.9 사이에 분포되어 있음을 알 수 있습니다.

- 지지도가 높은 응답규칙

{근로여건만족도_일.가정양립직장문화=보통,근로여건만족도_교육훈련기회=보통} => {일자리만족도=보통}

지지도가 가장 크게 나타나는 응답규칙은 위와 같습니다. 일.가정양립직장문화와 교육훈련기회에 대해 보통의 만족도를 가진다는 응답과 일자리만족도에 보통의 만족도를 가진다는 응답을 동시에 응답한 비중이 약 30%입니다. 지지도가 50%미만으로 다소 작은 값을 보이는데, 이는 응답자가 선택해야 할 항목이 다양하고 만족도 또한 다양한 범위로 분포되어있어 응답이 퍼져있기 때문이라고 판단됩니다.

- 신뢰도가 높은 응답규칙

{근로여건만족도_하는일=보통,근로여건만족도_인사관리=보통,근로여건만족도_임금=보통,근로여건만족도_근무환경=보통,근로여건만족도_근로시간=보통,근로여건만족도_폭력방지노력=보통,근로여건만족도_교육훈련기회=보통} => {일자리만족도=보통}

가장 높은 신뢰도는 99.7%로 아주 높은 값으로 나타났습니다. 하는일과 인사관리, 임금, 근무환경 그리고 근로시간, 폭력방지노력, 교육훈련기회에 대해 보통의 만족도를 가진 응답자들이 일자리 만족도에 보통을 체크했을 경우는 약 99.7%입니다.

- 연관성이 높은 응답규칙

{근로여건만족도_인간관계=약간만족,근로여건만족도_근로시간=약간만족,일자리만족도=약간만족} => {근로여건만족도_일.가정양립직장문화=약간만족}

가장 높은 lift 값은 3.667253 입니다. 인간관계, 근로시간, 일자리만족도에 약간만족이 선택된 응답규칙과 일.가정양립직장문화에 약간만족이 선택된 응답규칙 사이에 연관성은 약 3.7로 높은 연관성을 보이고 있습니다.

5.2.2 연관분석을 통해 분석하고자 하는 응답규칙

분석하고자 하는 응답규칙을 말하기에 앞서 이 프로젝트의 목표인 워라밸에 대해 다시 설명해보겠습니다.

워라밸이란 일과 가정생활에 대한 균형입니다.

가정생활과 일, 두가지를 동시에 만족해야 워라밸을 누릴 수 있습니다.

앞서 상관분석에서 일자리만족도와 개인생활만족도는 양의 상관관계에 있음을 알아냈습니다.

따라서 일자리만족도가 증가할수록 개인생활만족도 또한 증가하고 워라벨을 누릴 확률이 높아집니다.

제가 연관분석에서 할 일은 일자리만족도와 높은 연관성을 가지는 근로요건을 분석하는 것입니다.

이로써 전보다 높은 워라벨을 보장하고자 하는 기업에게 도움을 주고자 하는 이 프로젝트의 목표를 달성할 수 있습니다.

● 일자리 만족도에 영향을 주는 근로요건 분석

일자리 만족도와 강한 연관관계를 가지는 응답규칙을 발견하기 위하여 일자리만족도를 매우만족에 두고 연관분석을 하였습니다.

lhs		rhs	support	confidence	coverage	lift	count
{근로여건만족도_하는일=매우만족,근로여건만족도_인간관계=매우만족}	=>	{일자리만족도=매우만족}	0.04050839	0.779085	0.05199483	12.220338	596
{근로여건만족도_하는일=매우만족,근로여건만족도_근로시간=매우만족}	=>	{일자리만족도=매우만족}	0.0407123	0.7759067	0.0524706	12.170486	599
{근로여건만족도_근무환경=매우만족,근로여건만족도_근로시간=매우만족}	=>	{일자리만족도=매우만족}	0.04057636	0.7733161	0.0524706	12.12985	597
{근로여건만족도_근무환경=매우만족,근로여건만족도_인간관계=매우만족}	=>	{일자리만족도=매우만족}	0.04044043	0.7569975	0.05342214	11.873884	595
{근로여건만족도_일.가정양립직장문화=매우만족,근로여건만족도_폭력방지노력=매우만족}	=>	{일자리만족도=매우만족}	0.04234351	0.7364066	0.05750017	11.550907	623
{근로여건만족도_인간관계=매우만족,근로여건만족도_근로시간=매우만족}	=>	{일자리만족도=매우만족}	0.04159587	0.7346939	0.0566166	11.524042	612
{근로여건만족도_하는일=매우만족,근로여건만족도_폭력방지노력=매우만족}	=>	{일자리만족도=매우만족}	0.04186774	0.7333333	0.05709237	11.502701	616
{근로여건만족도_근로시간=매우만족,근로여건만족도_일.가정양립직장문화=매우만족}	=>	{일자리만족도=매우만족}	0.04227554	0.7300469	0.05790797	11.451152	622
{근로여건만족도_근무환경=매우만족,근로여건만족도_폭력방지노력=매우만족}	=>	{일자리만족도=매우만족}	0.04078026	0.7075472	0.0576361	11.098232	600
{근로여건만족도_근로시간=매우만족,근로여건만족도_폭력방지노력=매우만족}	=>	{일자리만족도=매우만족}	0.04363488	0.6708464	0.06504452	10.522562	642
{근로여건만족도_인간관계=매우만족,근로여건만족도_폭력방지노력=매우만족}	=>	{일자리만족도=매우만족}	0.04322708	0.6625	0.06524842	10.391644	636
{근로여건만족도_일.가정양립직장문화=매우만족}	=>	{일자리만족도=매우만족}	0.04621763	0.6557377	0.07048189	10.285574	680
{근로여건만족도_폭력방지노력=매우만족,근로여건만족도_교육훈련기회=매우만족}	=>	{일자리만족도=매우만족}	0.04411065	0.6264479	0.07041392	9.826149	649
{근로여건만족도_근무환경=매우만족}	=>	{일자리만족도=매우만족}	0.04723714	0.5806182	0.08135662	9.107288	695
{근로여건만족도_교육훈련기회=매우만족}	=>	{일자리만족도=매우만족}	0.0473051	0.5771144	0.08196833	9.052329	696
{근로여건만족도_근로시간=매우만족}	=>	{일자리만족도=매우만족}	0.04934412	0.5219267	0.09454224	8.186681	726
{근로여건만족도_인간관계=매우만족}	=>	{일자리만족도=매우만족}	0.04859648	0.5125448	0.09481411	8.039522	715

위는 lift 값에 대해 내림차순으로 정렬한 표입니다. lift 값은 8 에서 12 사이로 높은 값을 가집니다. 좌측 위에서부터 차례대로 “일자리만족도=매우만족”과 연관성이 높은 규칙입니다. 기업은 전보다 높은 하는일과 인간관계, 근로시간, 근무환경을 먼저 재정비한다면 전보다 높은 워라벨을 보장할 수 있을 것입니다.

5.3 텍스트분석

- 사용한 텍스트에 대한 설명

제가 사용한 텍스트는 고용노동부에서 운영하는 블로그의 “[걸어서 워라밸 속으로] 워라밸 지수 1 위 네덜란드”입니다. OECD 국가중에서 워라밸지수 1 위인 네덜란드의 비결이 무엇인지 서술한 글입니다. 어떻게 높은 워라밸지수를 가지게 되었는지 궁금하여 텍스트분석을 진행해보았습니다.

[출처]

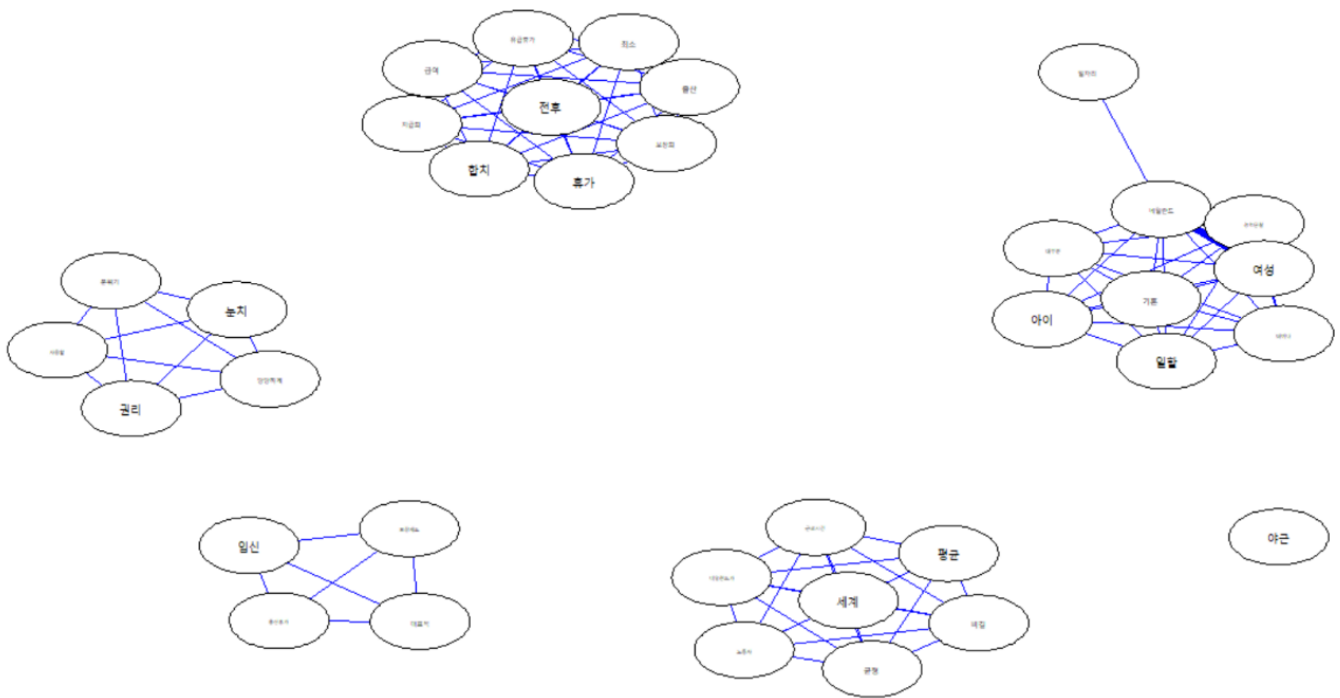
https://blog.naver.com/life_n_work/221354189587

5.3.1 워드클라우드 시각화



네덜란드에 대한 텍스트인만큼 네덜란드가 가장 크게 나타났습니다. 일자리, 근로시간, 경력단절, 여성, 야근, 균형이 그 다음으로 크게 나타났습니다. 네덜란드는 워라밸이 높으므로 부정적인 키워드인 경력단절과 야근 등은 현재 없애고 있고, 긍정적인 키워드인 일자리, 균형은 키우고 있을 것을 유추해 볼 수 있습니다.

5.3.2 네트워크 시각화



달락으로 이뤄진 글이기 때문에 네트워크 시각화가 덩어리들로 표현된 것 같습니다. 진한 선을 위주로 설명해보자면 오른쪽 상단의 네트워크 시각화에서 “네덜란드”와 “경력단절”, 그리고 “여성”이 진한 선으로 연결되어 있습니다. 이 3 가지 키워드는 강한 상호연관관계를 가집니다. 그리고 “여성”이 “기혼”, “일할”이라는 키워드와 연결되어 있는 것으로 보아 기혼여성들의 경력단절을 없애고 일할 기회를 준다는 의미로 해석할 수 있습니다. 또한 중앙 하단의 네트워크 시각화에서 근로시간, 세계, 균형이 진한 선으로 연결되어 있습니다. 이 3 개의 키워드 또한 강한 상호연관관계를 가지며 네덜란드는 균형 있는 근로시간을 갖고 있음을 유추해 볼 수 있습니다.

6. 프로젝트 성과 및 기대효과

프로젝트를 통해 여러 분석과정을 거치면서 워라밸이 어떤 요소와 관련이 있는지 알아보았습니다. 시각화를 통해 연관성을 발견한 개인생활 만족도와 일자리만족도는 상관분석을 거치며 서로 양의 상관관계를 가짐을 알게 되었고, 그 사실을 토대로 연관분석에서 워라밸을 보장하기 위해 일자리만족도와 높은 연관성을 가지는 근로요건을 찾아야 한다는 분석방향을 설정할 수 있었습니다.

상관분석과 연관분석에서 동시에 일자리만족도와 높은 연관성을 가지는 근로요건은 “하는 일, 근무환경, 폭력방지노력=임금, 근로시간, 인간관계”입니다. 이 근로요건을 개선함으로써 기업은 전보다 워라벨을 보장할 수 있을 것입니다.

7. Lessons Learned

이 프로젝트를 처음에 시작할 때 두려운 점이 많았습니다. 데이터를 분석하면 내 생각과는 다른 결과가 나와 당황하는 일이 많기 때문입니다. 데이터를 수집하는 과정에서도 보통은 정제 되어있는 데이터를 쓸 때가 많은데, 이번에는 직접 코드집을 참고하여 데이터를 정제해야 했기에 생각보다 시간이 많이 소요되었습니다. 하지만 이번 기회를 통해 내가 관심있는 주제인 워라벨을 직접 분석해보고 실제로 워라벨이 근로의 요소들과 관계가 있음을 알아냈을 때 보람이 있었습니다. 분석하는 과정은 힘들지만 결과물이 나왔을 때 뿌듯함을 느끼는 것이 데이터 분석의 묘미인 것 같습니다.