

Cultural Industry
& Data Analytics



Data-driven Final Group Project

김상호 김성은 김유정 양신 전아영

June 2025

* Contents

1. 주제 소개
2. 데이터 정보 및 수집 과정
3. 데이터 선정 이유
4. 데이터 살펴보기
5. 가설 설정
6. 데이터 분석
7. 분석 결과
8. 도출 문제에 대한 보완 전략(시사점)

* 1 주제 소개

한국으로 온 결혼이민자의 인공지능 활용이
일상생활 만족도와 사회적 자본에 어떤 영향을 미치는지 탐색



* 1 주제 소개

사회적 자본: 결혼이민자의 사회적 네트워크와 그로부터 얻을 수 있는 정서적·실질적 자원을 의미

실제 사회적 자본 문항 10개

1. 내 문제를 해결하는 데 도움을 줄 수 있는 사람이 있다.
2. 매우 중요한 결정을 내릴 때 조언을 구할 수 있는 사람이 있다.
3. 나의 중요한 일을 맡길 수 있는 사람이 있다.
- ⋮
9. 긴밀한 사적인 문제에 대해 편안하게 이야기할 사람이 있다.
10. 내가 불의에 대하여 싸우는 것을 도와줄 사람들이 있다.

* 2 데이터 정보 및 수집 과정



- 1. 일반국민_파일설계서
- 2. 농어민_파일설계서
- 3. 장애인_파일설계서
- 4. 저소득층_파일설계서
- 5. 북한이탈주민_파일설계서
- 6. 결혼이민자_파일설계서
- 7. 고령층_파일설계서

- 1 일반국민_raw data
- 2 농어민_raw data
- 3 장애인_raw data
- 4 저소득층_raw data
- 5 북한이탈주민_raw data
- 6 결혼이민자_raw data
- 7 고령층_raw data

데이터 정보

- 출처: 한국지능정보사회진흥원, 2023년도 디지털 정보격차 실태조사
- 대상: 결혼이민자 700명
- 수집 방식: 성별, 연령, 출신국적, 권역별 비례할당표집 / 1:1 대면 면접
- 데이터 형식: 총 230개의 문항, 변수명은 설문 항목을 코드화한 형태 (Q1A1, Q30A6 등)

* 3 데이터 선정 이유

- 1) 인공지능 서비스 경험률에서 결혼이민자가 일반국민 보다 더 높은 수치를 보임

| 그림 233. 인공지능 서비스 경험률

(단위: %)

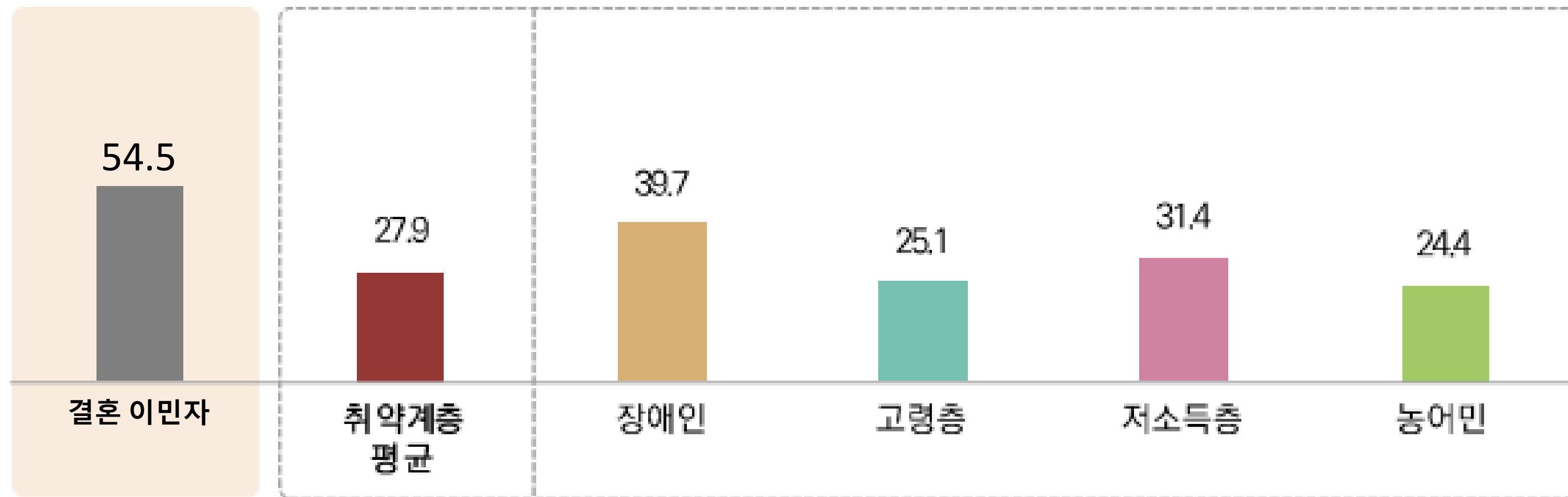


* 3 데이터 선정 이유

2) 특히 다른 취약 계층과 비교했을 때 결혼이민자의 인공지능 서비스 이용 경험률이 높은 것을 알 수 있음

| 그림 239. 인공지능(AI) 서비스 이용 경험

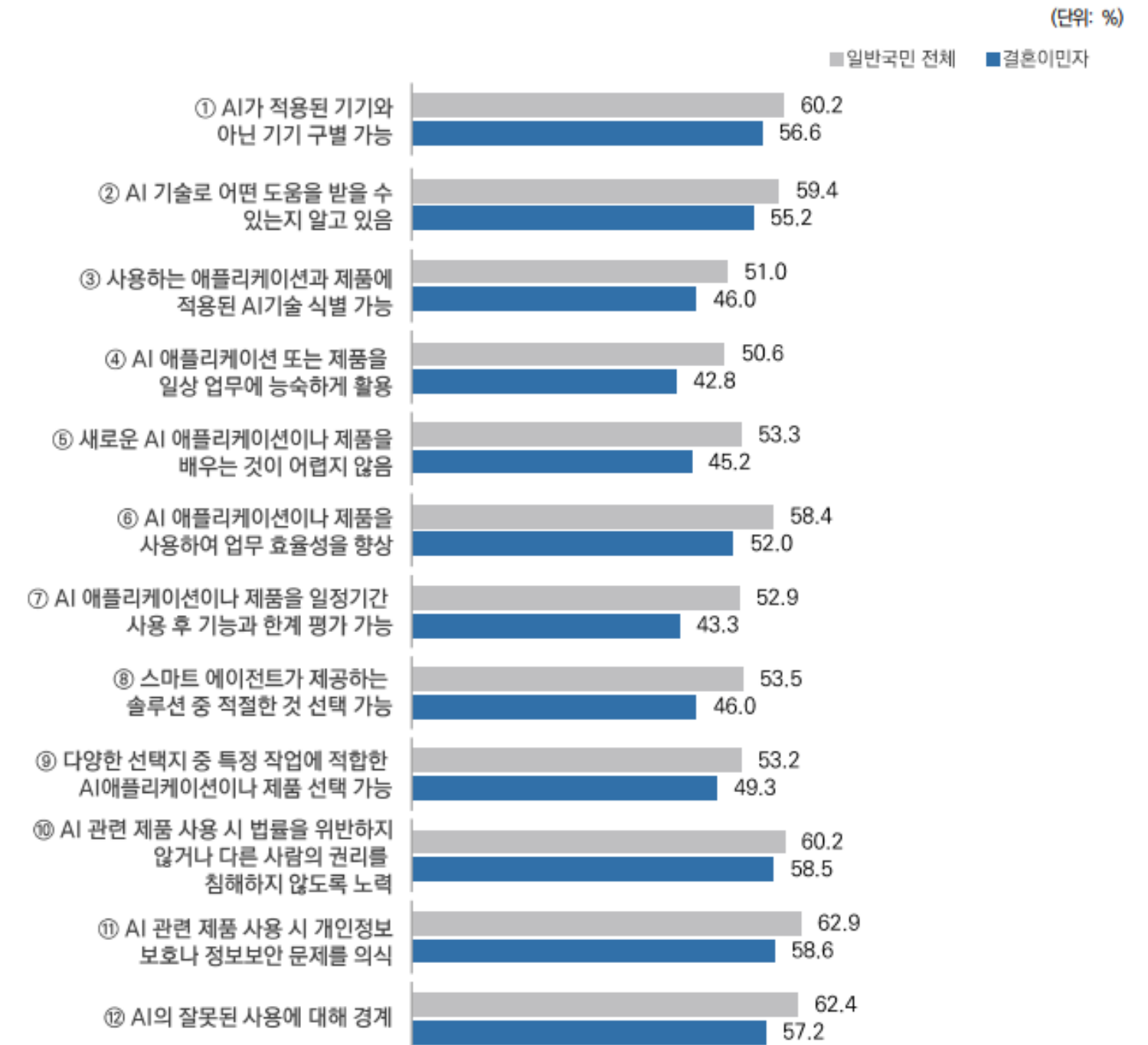
(단위: %)



* 3 데이터 선정 이유

3) 그러나 결혼이민자의 인공지능 서비스 이용 역량은 모든 문항에서 일반 국민보다 비교적 낮은 것을 확인할 수 있음

∴ 그림 246 인공지능 서비스 이용 역량



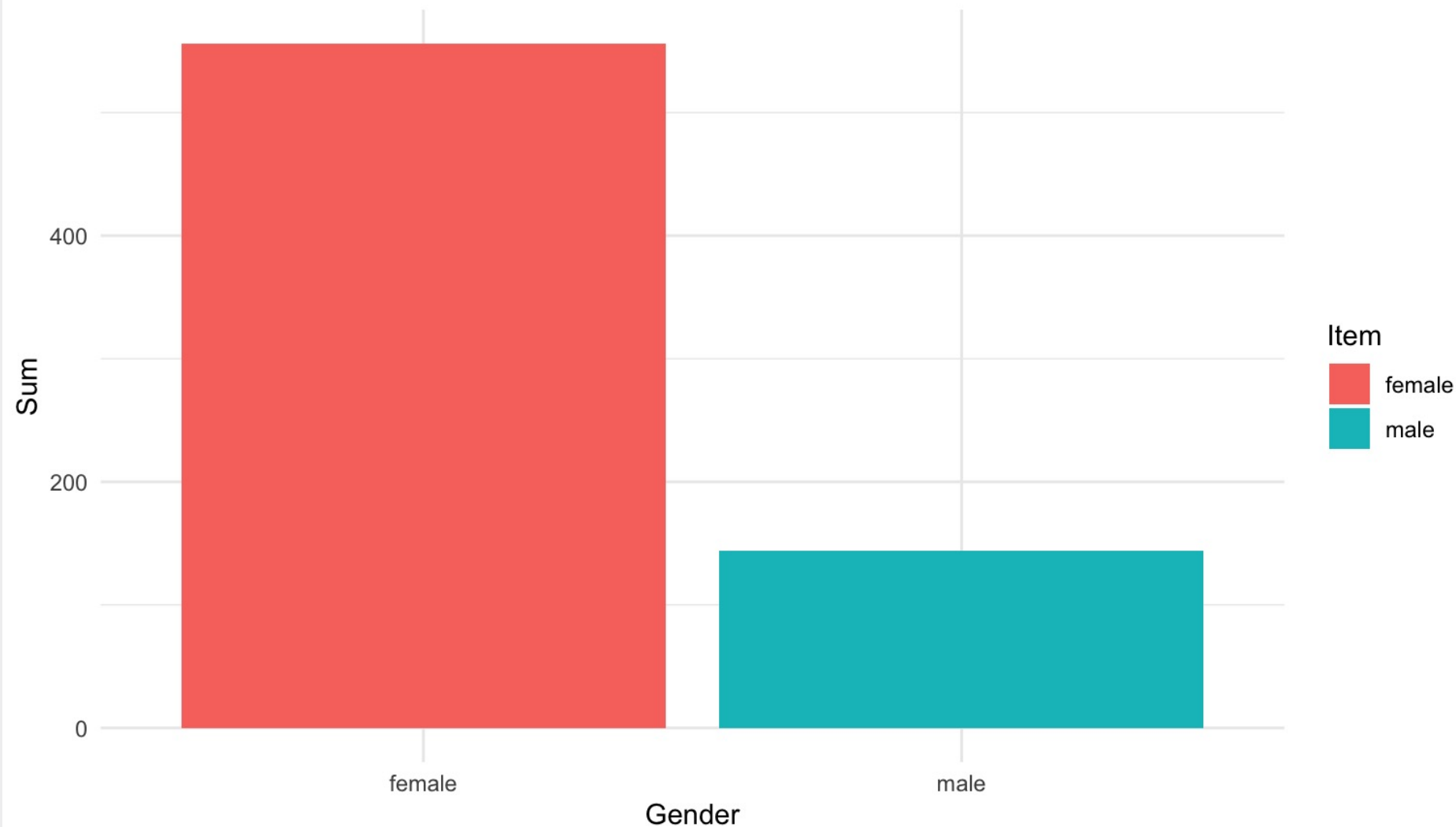
※ 각 항목에 대해 긍정(그런 편이다 + 매우 그렇다) 응답 비율

* 4 데이터 살펴보기

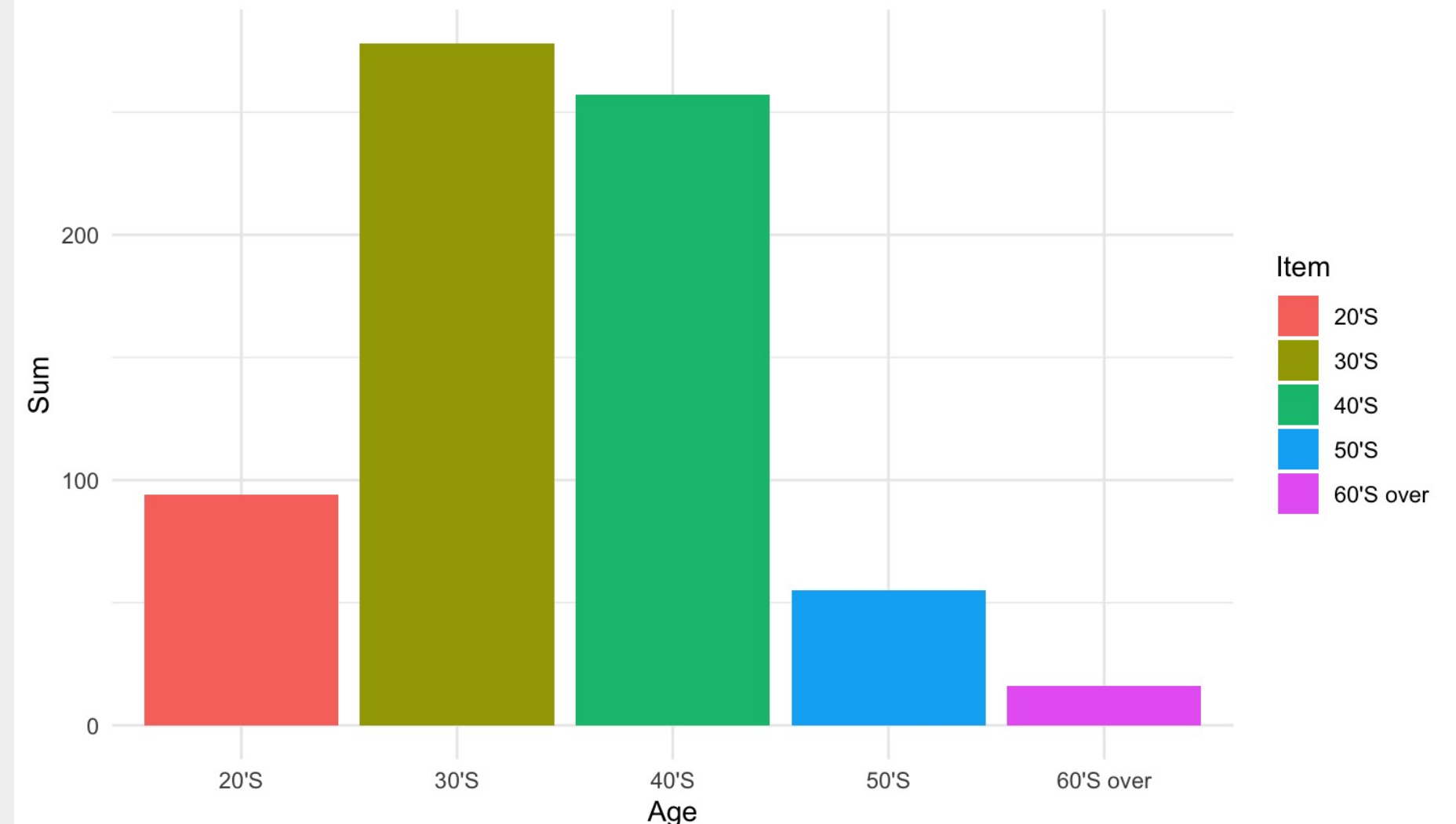
1) 데이터 개요: 성별 & 나이 분포

...with ggplot & tidyverse!

Gender Distribution



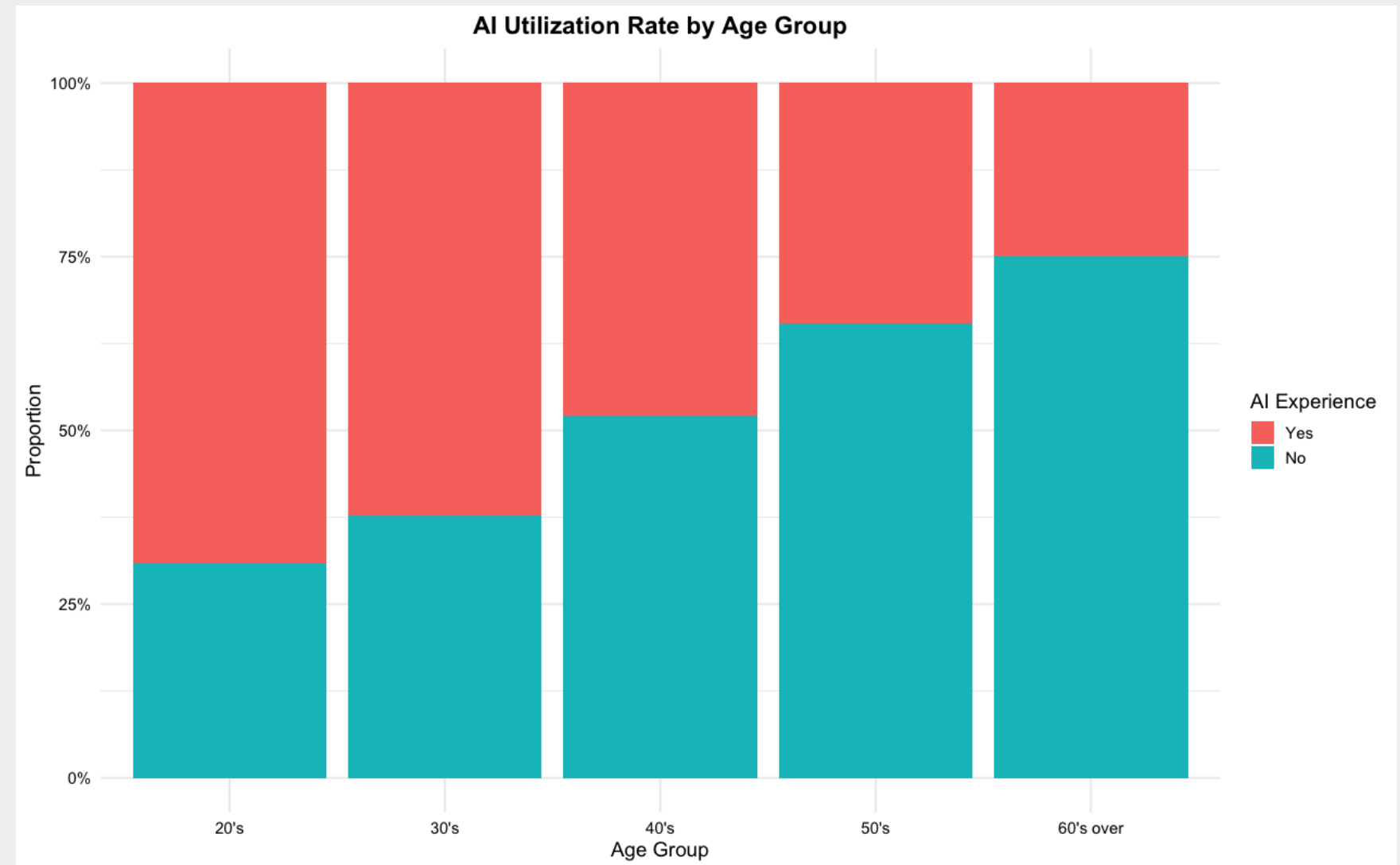
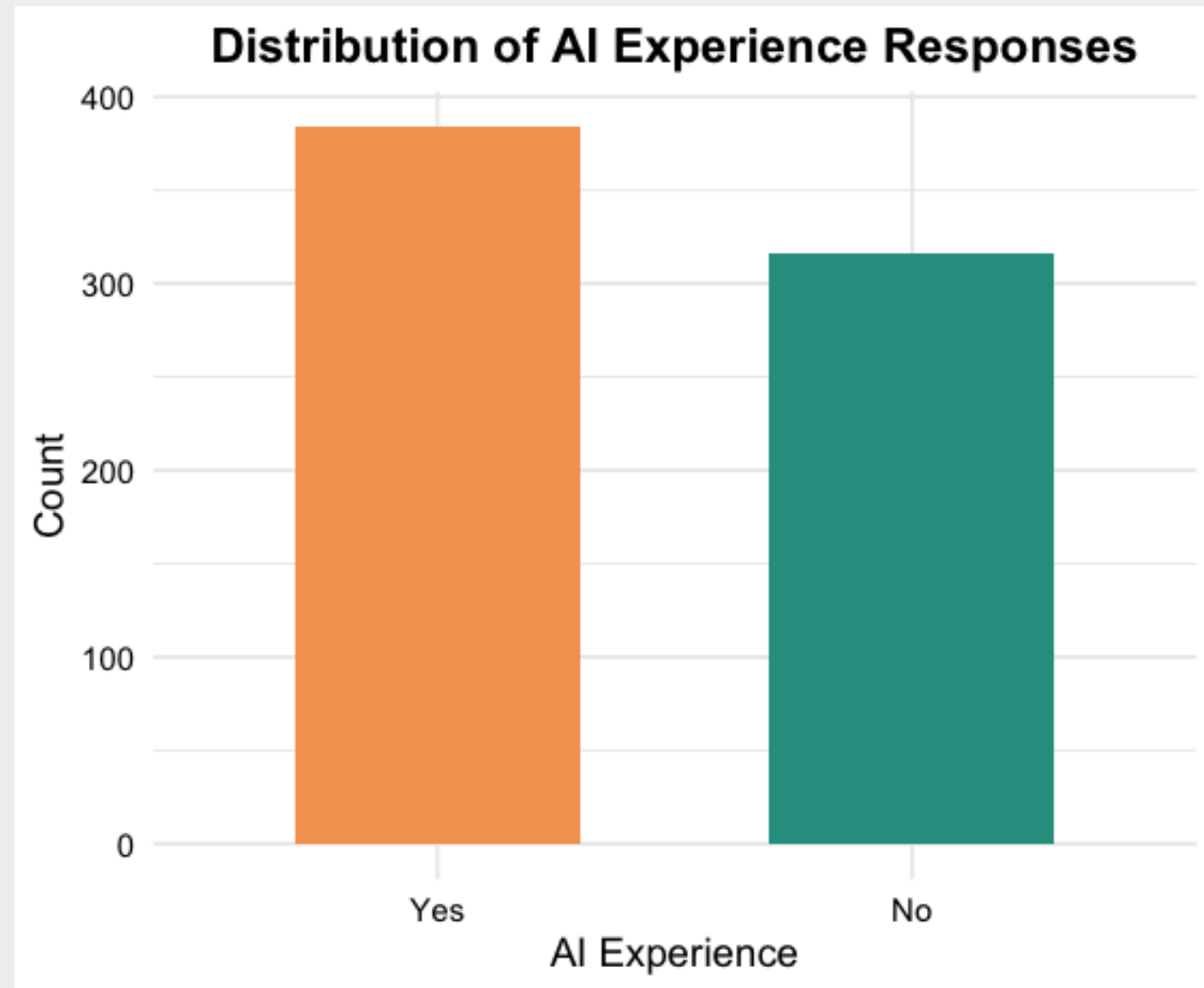
Age Distribution



* 4 데이터 살펴보기

1) 데이터 개요: AI 활용 여부 및 연령대별 분포

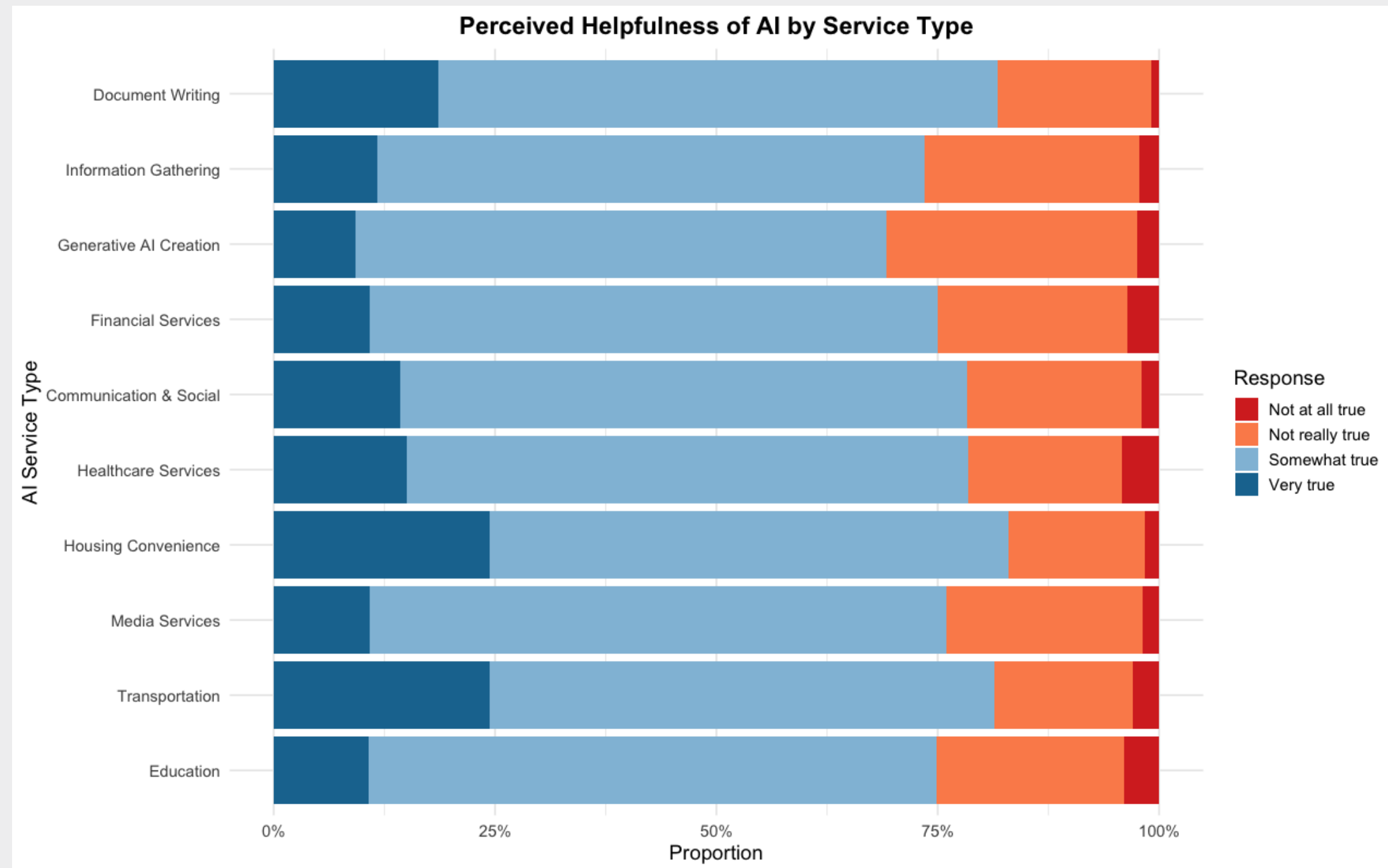
...with ggplot & tidyverse!



* 4 데이터 살펴보기

1) 데이터 개요: 일상생활 만족도 응답 분포

...with ggplot & tidyverse!



* 4 데이터 살펴보기

2) 데이터 전처리

1. 결혼 이민자들의 실제 AI 사용 여부 (True/False)

Q21	116	문21. 인공지능 서비스 이용 경험 여부	5	오른쪽 맞춤	F1	F1
-----	-----	------------------------	---	--------	----	----

1	예
2	아니오

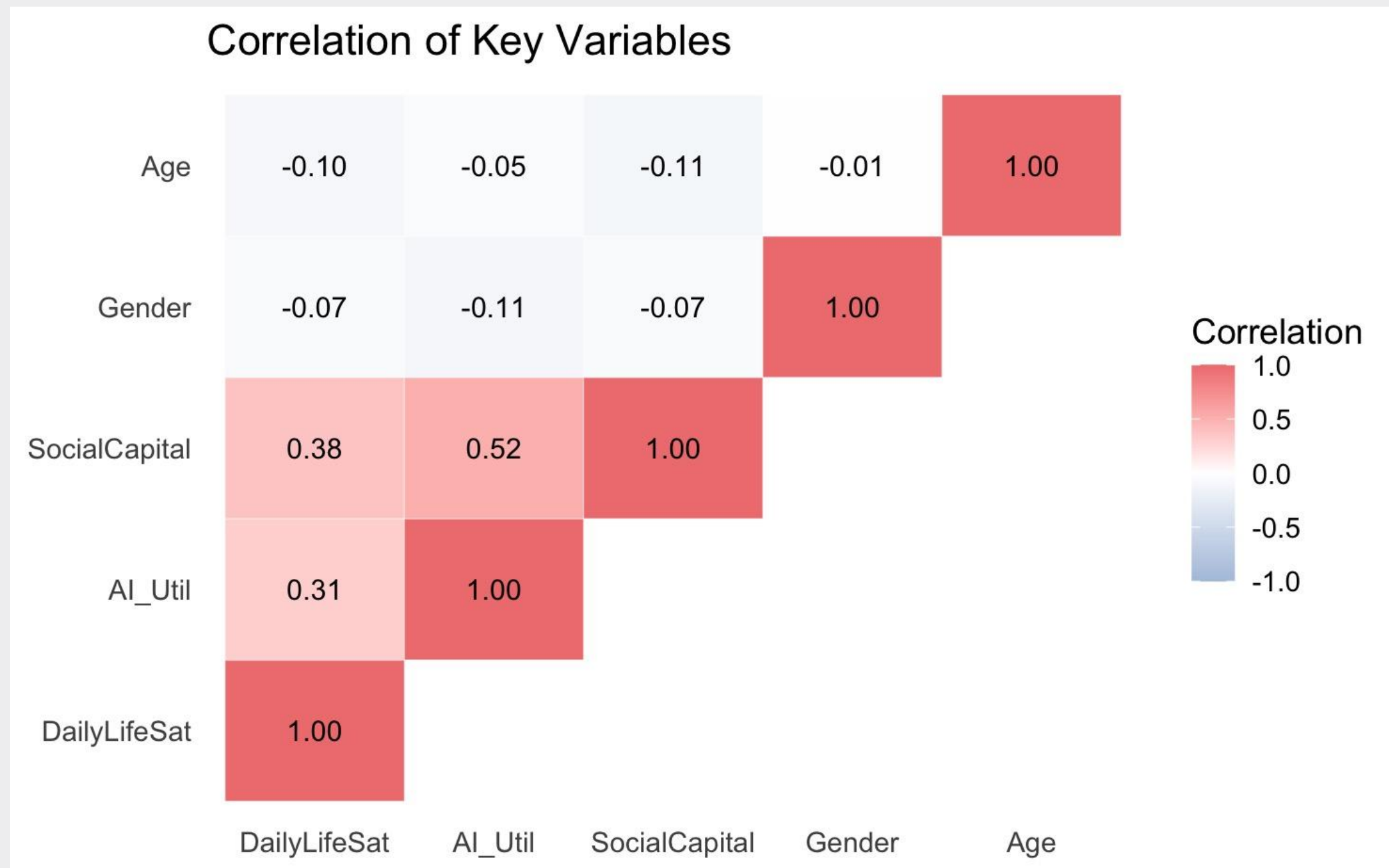
2. 결혼 이민자들의, AI 활용으로 발생할 일상 생활 증진에 대한 인식

Q27A1	171	문27. 인공지능 인식 - 내 삶을 편리하게 할 것	7	오른쪽 맞춤	F1	F1
Q27A2	172	문27. 인공지능 인식 - 경제적으로 많은 기회가 생길 것	7	오른쪽 맞춤	F1	F1
Q27A3	173	문27. 인공지능 인식 - 더 나은 정보 서비스를 받게 해줄 것	7	오른쪽 맞춤	F1	F1
Q27A4	174	문27. 인공지능 인식 - 더 좋은 정보를 얻게 될 것	7	오른쪽 맞춤	F1	F1
Q27A5	175	문27. 인공지능 인식 - 일자리가 줄어들 것	7	오른쪽 맞춤	F1	F1
Q27A6	176	문27. 인공지능 인식 - 개인정보 침해나 프라이버시 침해 우려가 커질 것	7	오른쪽 맞춤	F1	F1
Q27A7	177	문27. 인공지능 인식 - 허위 조작 정보가 심각해질 것	7	오른쪽 맞춤	F1	F1
Q27A8	178	문27. 인공지능 인식 - 사람들의 관계를 더 좋게 만들 것	7	오른쪽 맞춤	F1	F1

1	전혀 그렇지 않다
2	그렇지 않은 편이다
3	그런 편이다
4	매우 그렇다

* 4 데이터 살펴보기

3) 주요 변수들 간의 상관관계 heatmap



* 5 가설 설정

H1: 결혼이민자의 AI 활용은 일상생활 만족도에 긍정적인 영향을 미친다.

H2: 결혼이민자의 AI 활용은 사회적 자본에 긍정적인 영향을 미친다.

* 6 데이터 분석

H1: 결혼이민자의 AI 활용은 일상생활 만족도에
긍정적인 영향을 미친다.

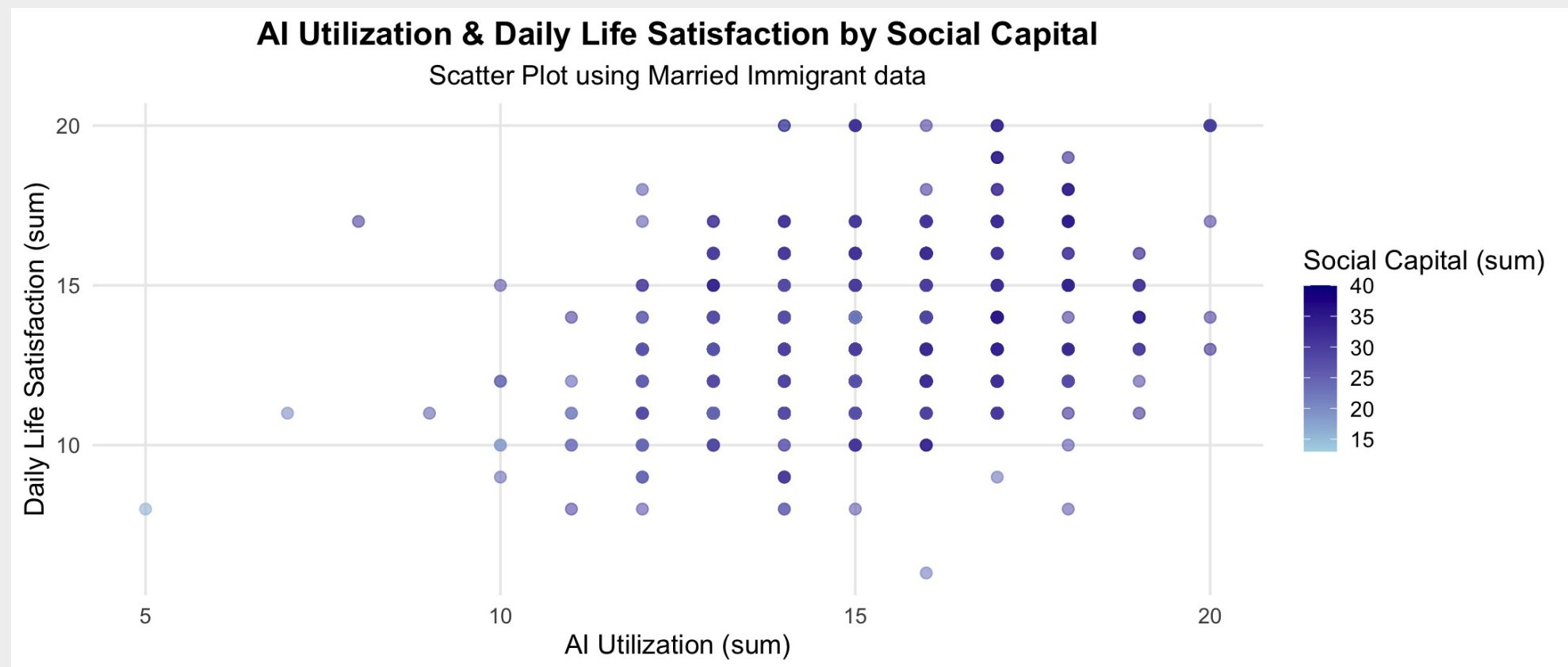
AI 활용도 증진 → 일상생활 만족도 상승 경향

특히, 결혼이민자 분포는 AI 활용에 대한 긍정적인 답변에 정비례하여
높은 일상생활 만족도 분포를 보이고 있음 (우상향)

시사점: 결혼이민자들의 AI 활용에 대한 인식은
일상생활 만족도와 양의 관계를 지니고 있음

[실행 코드]

```
90 p <- ggplot(df, aes(x = AI_Util, y = DailyLifeSat, color = SocialCapital)) +
91   geom_point(alpha = 0.6, size = 2.5) +
92   scale_color_gradient(low = "lightblue", high = "darkblue") +
93   labs(
94     title = "AI Utilization & Daily Life Satisfaction by Social Capital",
95     subtitle = "Scatter Plot using Married Immigrant data",
96     x = "AI Utilization (sum)",
97     y = "Daily Life Satisfaction (sum)",
98     color = "Social Capital (sum)"
99   ) +
100   theme_minimal(base_size = 14) +
101   theme(
102     plot.title = element_text(face = "bold", hjust = 0.5),
103     plot.subtitle = element_text(hjust = 0.5),
104     legend.position = "right",
105     panel.grid.minor = element_blank()
106   )
107
108 print(p)
```



Q20A1	111	문20. 삶에 대한 만족도_1) 대부분의 경우에서 내 삶은 내 이상과 가깝다
Q20A2	112	문20. 삶에 대한 만족도_2) 내 삶의 조건들은 매우 우수하다
Q20A3	113	문20. 삶에 대한 만족도_3) 나는 내 삶에 만족한다
Q20A4	114	문20. 삶에 대한 만족도_4) 지금까지 나는 인생에서 내가 원하는 중요한 것들을 얻었다
Q20A5	115	문20. 삶에 대한 만족도_5) 내 삶을 다시 산다면, 나는 거의 아무것도 바꾸지 않을 것이다

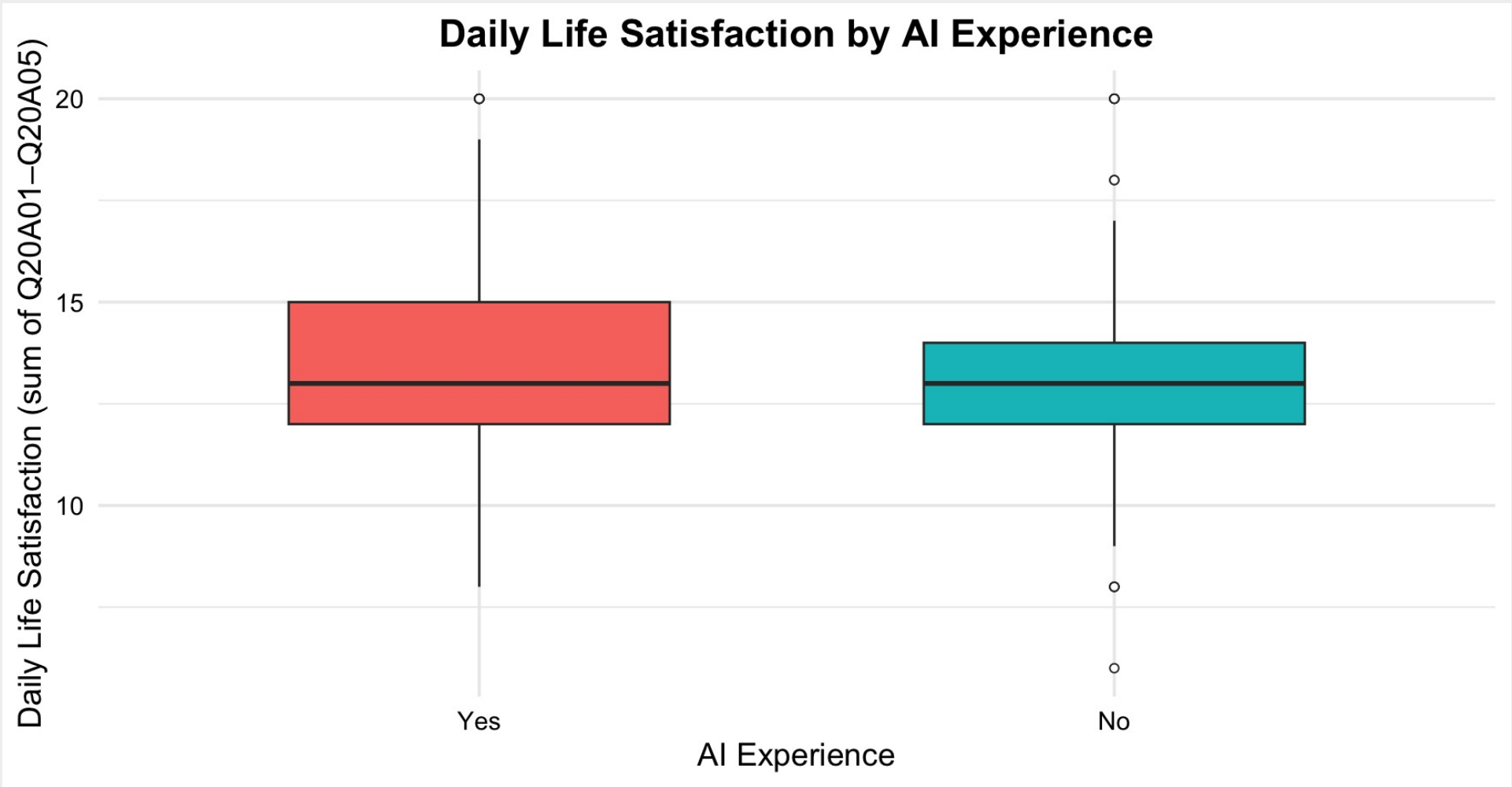
＊ 6 데이터 분석

H1: 결혼이민자의 AI 활용은 일상생활 만족도에
긍정적인 영향을 미친다.

실제로 AI를 경험한 결혼이민자들의 일상생활 만족도가
경험하지 않은 사람들의 일상 만족도보다
비교적 높고 넓은 분포를 보이고 있음

[실행 코드]

```
37 ~ # 6) Bar plot -----
38 ggplot(data, aes(x = AI_experience, y = DailyLifeSat, fill = AI_experience)) +
39   geom_boxplot(width = 0.6, alpha = 0.85, outlier.shape = 1) +
40   labs(
41     title = "Daily Life Satisfaction by AI Experience",
42     x = "AI Experience",
43     y = "Daily Life Satisfaction (sum of Q20A01-Q20A05)"
44   ) +
45   scale_fill_manual(values = c("Yes" = "#e76f51", "No" = "#2a9d8f")) +
46   theme_minimal(base_size = 14) +
47   theme(
48     plot.title = element_text(face = "bold", hjust = 0.5),
49     axis.text = element_text(color = "black"),
50     legend.position = "none"
51   )
```



Q20A1	111	문20. 삶에 대한 만족도_1) 대부분의 경우에서 내 삶은 내 이상과 가깝다
Q20A2	112	문20. 삶에 대한 만족도_2) 내 삶의 조건들은 매우 우수하다
Q20A3	113	문20. 삶에 대한 만족도_3) 나는 내 삶에 만족한다
Q20A4	114	문20. 삶에 대한 만족도_4) 지금까지 나는 인생에서 내가 원하는 중요한 것들을 얻었다
Q20A5	115	문20. 삶에 대한 만족도_5) 내 삶을 다시 산다면, 나는 거의 아무것도 바꾸지 않을 것이다

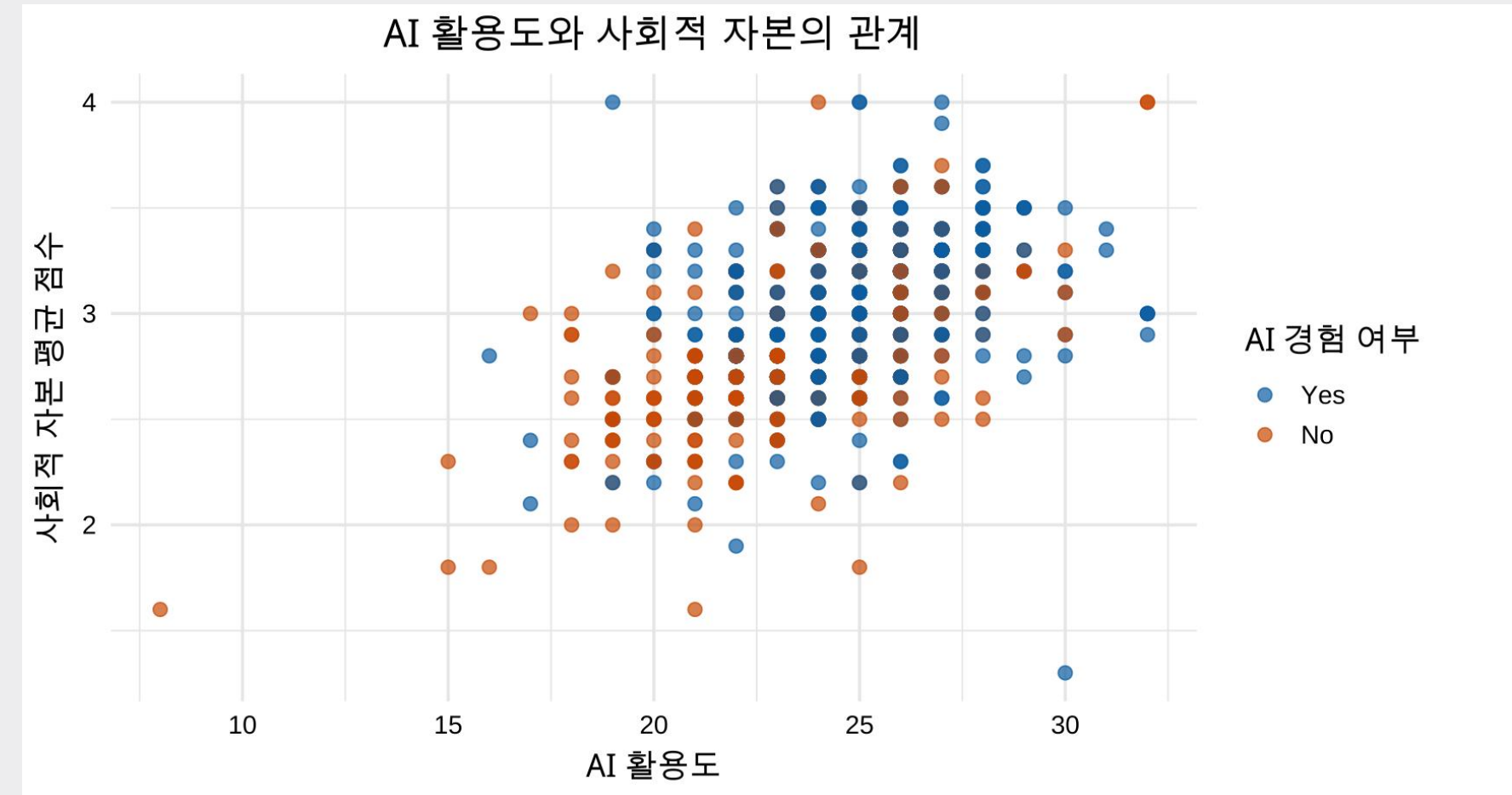
* 6 데이터 분석

H2: 결혼이민자의 AI 활용은 사회적 자본에 긍정적인 영향을 미친다

AI 활용도 증진 → 자본 점수 상승 경향

특히, AI 경험 집단 ●은 평균적으로 더 높은 AI 활용도와 사회적 자본 수준을 나타냄

시사점: 단순한 AI 경험 여부를 넘어, 얼마나 능동적으로 활용하는지가 사회적 자본 형성에 긍정적인 영향을 미칠 수 있음



[실행 코드]

```
# 2. 시각화
p <- ggplot(df, aes(x = AI_Util, y = SocialCapital, color = AI_experience)) +
  geom_point(alpha = 0.7, size = 2.5) +
  scale_color_manual(values = c("Yes" = "#0072B2", "No" = "#D55E00"), name = "AI 경험 여부") +
  labs(
    title = "AI 활용도와 사회적 자본의 관계",
    x = "AI 활용도",
    y = "사회적 자본 평균 점수"
  ) +
  theme_minimal(base_size = 14) +
  theme(
    plot.title = element_text(face = "bold", hjust = 0.5),
    plot.subtitle = element_text(hjust = 0.5),
    legend.position = "right",
    axis.text = element_text(color = "black")
  )
```

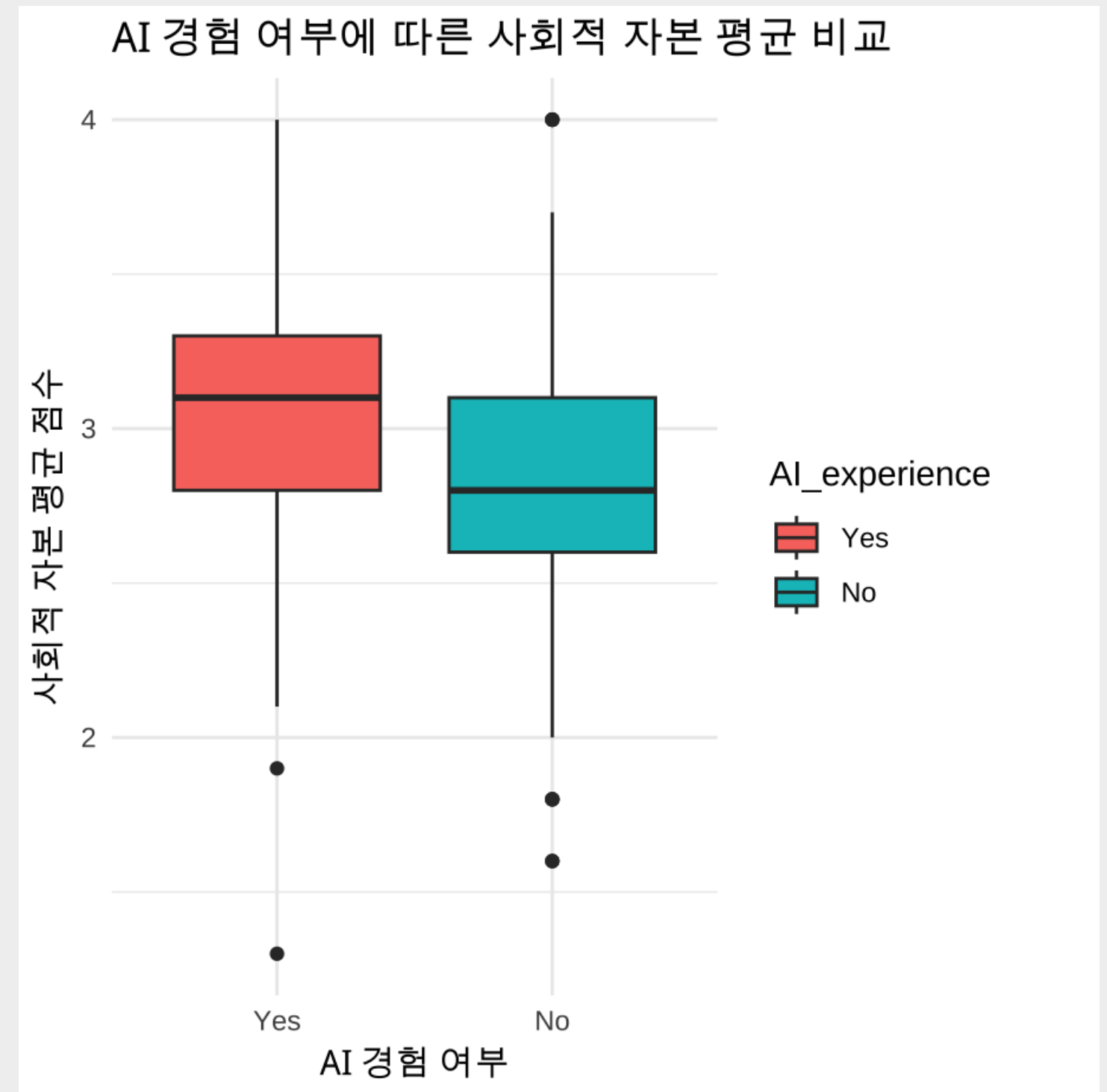
* 6 데이터 분석

H2: 결혼이민자의 AI 활용은 사회적 자본에 긍정적인 영향을 미친다

AI를 경험한 결혼이민자들이 실제로 더 높은 사회적 자본 수준을 보이며, 이는 통계적으로 유의미한 결과

[실행 코드]

```
# 7. 시각화 (Boxplot)
library(ggplot2)
ggplot(data, aes(x = AI_experience, y = social_capital, fill = AI_experience)) +
  geom_boxplot() +
  labs(
    title = "AI 경험 여부에 따른 사회적 자본 평균 비교",
    x = "AI 경험 여부",
    y = "사회적 자본 평균 점수"
  ) +
  theme_minimal()
```



* 6 데이터 분석

H2: 결혼이민자의 AI 활용은 사회적 자본에 긍정적인 영향을 미친다

Welch Two Sample t-test

```
data: social_capital by AI_experience
t = 7.4563, df = 661.07, p-value = 2.803e-13
alternative hypothesis: true difference in means between group Yes and group No is not equal to 0
95 percent confidence interval:
 0.1529459 0.2622967
sample estimates:
mean in group Yes  mean in group No
      3.039583      2.831962
```

두 집단의 유의미 결과 기준: $p\text{-value} < 0.05$

실제 분석 값: **p-value 2.803e-13**

-> AI 사용 여부와 사회적 자본 사이에 실제 관련성이 존재할 가능성이 매우 높음

[실행 코드]

```
# 6. 그룹 간 평균 비교
# 정규성과 등분산성이 만족되면 t-test, 그렇지 않으면 wilcox.test 사용
t.test(social_capital ~ AI_experience, data = data, var.equal = FALSE)
```

* 7 분석 결과

Model	Term	B	Std. Error	Beta	t	P
AI Utilization	Constant	7.240	0.658	NA	11.01	<.001
	DailyLifeSat	0.151	0.030	0.176	5.03	<.001
	SocialCapital	0.193	0.018	0.383	10.71	<.001
	AIExperience	0.464	0.132	0.121	3.52	<.001
	GenderFemale	-0.219	0.154	-0.046	-1.42	0.156
	Age	0.000	0.007	-0.002	-0.07	0.948
F = 54.76, Adj. R² = 0.278						

해당 표를 통해 기존에 세운 가설 1과 2 모두 입증함

개선점:

결혼 이민자들은 AI 활용률이 다른 계층에 비해 월등히 높고 이는 그들의 삶의 만족도, 사회적 자본에 긍정적 영향을 미침
반면 결혼 이민자들의 인공지능 이용 역량은 낮음
이를 보완하고자 함

* 8 도출 문제에 대한 보완 전략

맞춤형 어플리케이션

한국 결혼이민자를 위한 AI 비서 <다운이>

다문화가정'을 뜻하는 다(多) + '따뜻함'과 '포용'을 상징하는 온(溫)

1. 다문화 가정 도우미

2. 한국어 도우미

3. AI 리터러시 퀴즈



* 8-1 타켓 페르소나



리나 왕 (Lina Wang) / 34세 / 한국 체류 3년 / 한국어: 중급

목표: 아이 교육과 병원 예약 스스로 진행, 한국 직장(파트타임)에서 경제적 자립 시도

문제 상황: 공공기관 서류의 어려운 법률·행정 용어 이해 어려움 / 번역 앱 사용 시에도 맥락에 맞는 번역 불가

기술 수용도: 스마트폰 기본 기능 숙지 / 고급 기능 활용에는 낮은 숙련도

행동 패턴: 카카오톡 사용 빈도 높음 / 전문용어 변환 기능 사용 어려움 / 앱 설치, 로그인 등 초기 설정에 도움필요



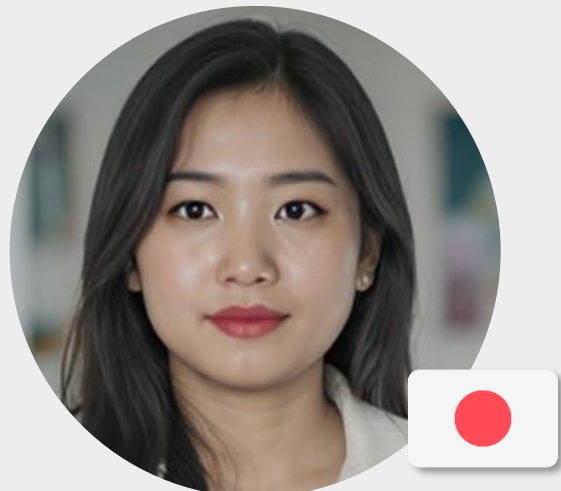
티엔 응우옌 (Thien Nguyen) / 29세 / 한국 체류 2년 / 한국어: 초급

목표: 지역 커뮤니티 참여를 통해 친구 사귀기, 온라인 한국어 학습으로 구직 준비

문제상황: 병원 예약 시 직접 전화 통화에 대한 두려움 / 병원 안내나 약 설명 이해에 어려움

기술 수용도: 스마트폰 카메라 및 SNS 기능 숙지

행동 패턴: 페이스북·카카오톡 오픈채팅으로 베트남 커뮤니티 접속 / 스마트폰 카메라로 사진 찍어 번역 앱 활용



사야카 타나카 (Sayaka Tanaka) / 31세 / 한국 체류 4년 / 한국어: 중급

목표: 한국 지역 문화 행사에 참여하여 한국인 친구 사귀기, 온라인으로 손쉽게 비자·외국인 등록 정보 조회

문제상황: AI 활용이 서툴러 정보 검색, 번역, 문의 방법을 몰라 시간이 많이 소요됨 / 직장 내 활동에 소극적

기술 수용도: 한국 포털, SNS 콘텐츠 탐색 능력 보유

행동 패턴: 네이버 블로그, 카카오톡스토리 통해 한국 정보 탐색 / 한국인 친구에게 카톡 활용하여 질문

The image displays four screenshots of a web application interface for a Korean AI chatbot, arranged in a 2x2 grid. Each screenshot represents a different user scenario.

Top-left screenshot: The header reads '다온이' (Daoni) and '한국 결혼이민자를 위한 AI 비서' (AI Secretary for Korean Marriage Immigrants). The sidebar shows '언어 선택' (Language Selection) set to '한국어' (Korean), '거주 지역' (Residence Area) set to '서울 종로구' (Seoul Jongno-gu), and '기능 선택' (Function Selection) set to 'AI 다문화가정도우미' (AI Multicultural Family Helper). The chat area shows a conversation where the AI asks for a question and the user asks about the chatbot's location.

Top-right screenshot: The header is the same. The sidebar shows '언어 선택' set to '한국어', '거주 지역' set to '서울 종로구', and '기능 선택' set to 'AI 한국어 선생님' (AI Korean Teacher). The chat area shows the AI asking for a question and the user asking for help with Korean.

Bottom-left screenshot: The header is the same. The sidebar shows '언어 선택' set to '한국어', '거주 지역' set to '서울 종로구', and '기능 선택' set to 'AI 리터러시 퀴즈' (AI Literacy Quiz). The chat area shows a conversation about a quiz.

Bottom-right screenshot: The header is the same. The sidebar shows '언어 선택' set to '베트남어' (Vietnamese), '거주 지역' set to '서울 종로구', and '기능 선택' set to 'AI 한국어 선생님'. The chat area shows a conversation in Vietnamese about a quiz.

Each screenshot includes a '질문을 입력하세요' (Enter your question) input field and a '보내기' (Send) button at the bottom.

* 8-3 페르소나 경험 여정 지도

	사용 전	온보딩	주요 사용	사용 후
<div>리나 왕</div> <div></div>	행정 서류 번역 실패 -> 관공서 방문 지연	기본 설정: 중국어 / 서울	기능 1 ‘AI 다문화가정 도우미’ 실행 -> ‘서류 번역 도와줘’ 채팅 -> 서류 이미지 업로드	번역본 PDF 다운로드 -> 다음 서류 번역 알림
<div>티엔 응우옌</div> <div></div>	병원 예약 전화 두려움 -> 진료 미루기	기본 설정: 베트남어 / 수원	기능 2 ‘AI 한국어 도우미’ 실행 -> 병원 예약 스크립트+오디오 제공	예약 성공 후 ‘후기 작성’ -> 커뮤니티 탭 이동
<div>사야카 타나카</div> <div></div>	AI 활용 능력 부족 -> 대화 참여 주저	기본 설정: 일본어 / 부산	기능 3 ‘AI 리터러시 퀴즈’ 실행 -> 프롬프트, 검색법 학습	퀴즈 배지 획득 -> 문화 행사 추천 알림 -> 링크 클릭 후 행사 참석



감사합니다.

Cultural Industry & Data Analytics

Data-driven Final Group Project

김상호 김성은 김유정 양신 전아영