

국제보건 데이터 마이닝 그룹: DHS-8 데이터 분석 가이드

- 조혼과 여성 생식건강 데이터 분석 사례를 중심으로 -

신보혜 강유미 김하나 서희연 신동혁 임유미 최지선
[국제보건 데이터 마이닝 그룹]

1. 서론

1) 연구 배경

국제 보건 분야는 최근 보건 서비스의 불평등 문제, 감염병 확산, 비감염성 질환의 증가 등 다양한 요인으로 인해 그 중요성이 더욱 부각되고 있습니다. 특히, 저소득 및 중저소득 국가(LMICs)는 기본적인 보건 서비스조차 부족한 상황에서 전염병과 비전염성 질환의 이중 부담을 겪고 있으며, 이러한 문제를 해결하려면 지역별 특성을 반영한 맞춤형 중재 전략이 필요합니다.

그러나 국내에서는 이들 국가의 보건 문제를 체계적으로 분석하는 연구가 드문 편입니다. 국내 연구 현황을 분석한 결과, 한국과학기술정보연구원(KISTI)의 과학기술 지식인프라 ScienceON에서 'DHS 데이터'를 키워드로 검색한 바, 국내 연구자들이 해당 데이터를 활용한 논문 수가 제한적임을 확인할 수 있었습니다. 이는 한국의 연구자들이 국제보건 연구에 중요한 자원으로 자리 잡은 DHS 데이터 활용에 비교적 덜 집중하고 있음을 시사합니다. 또한, 행정안전부의 '공공부문 데이터 분석·활용 우수사례집'에서 다양한 공공데이터 활용 사례를 제시하고 있지만, DHS 데이터를 활용한 사례는 포함되지 않은 점도 이를 뒷받침합니다. NGO의 데이터 활용 현황을 다룬 자료에 따르면, 국내 다수의 단체들이 공공데이터를 적극 활용하고 있으나, 국제 데이터를 활용한 사례, 특히 DHS 데이터를 사용하는 경우는 드문 편입니다. 이는 국내 NGO 활동가들이 DHS 데이터의 존재와 활용 방법에 대한 인식이 낮거나 접근성이 부족할 가능성을 시사합니다.

이에 국제보건 데이터 마이닝 그룹은 미국 국제개발처(USAID)에서 제공하는 DHS (Demographic and Health Surveys) 데이터를 활용하여 LMICs의 보건 문제를 분석하고, 국가별 맞춤형 중재 전략을 도출하려는 목표로 본 연구 모임을 시작하게 되었습니다.

DHS 데이터는 주로 인구, 건강, 영양 지표를 포함하는 방대한 자료로, LMICs의 모성 건강, 아동 사망률, 영양 상태 등을 자세히 다루고 있습니다. 이는 WHO, UNICEF, USAID 등 국제기구에서 사용되는 공신력 있는 데이터로서 각국의 보건 정책 수립과 평가에 필수적인 정보를 제공합니다. 본 연구모임의 시작은 LMICs의 보건 문제를 국제적으로 비교하고 분석할 수 있는 자료로서 DHS 데이터의 잠재력에 주목한 데 있습니다. 국내에서는 보건 데이터를 주로 국내 연구에 한정하여 활용해 왔기 때문에 국제적 데이터를 분석하고 활용할 기회가 부족했던 상황을 개선하고자 모임을 통해 DHS 대규모 데이터 전처리와 기본 분석을 돕기 위해 추진되었습니다.

2) 연구 필요성

LMICs의 보건 문제는 보건 서비스와 자원의 불균형이 주요 요인으로 작용하고 있으며, 국제적으로 공조하여 해결해야 할 과제입니다. 특히 출산 관련 합병증, 아동 사망률, 영양 결핍과 같은 기본적인 보건 지표가 여전히 낮은 수준에 머물러 있는 국가들이 많습니다. DHS 데이터는 이러한 국가들의 보건 문제를 체계적으로 이해하고, 구체적인 개선 방안을 마련하는 데 유용한 자료임에도 불구하고, 국내 연구자들에게는 접근성과 분석의 어려움이 존재합니다.

따라서, 국제보건 데이터 마이닝 그룹은 DHS 데이터를 활용해 국가별 보건 상황을 비교하고 분석할 수 있는 역량을 개발하는 것을 목표로 했습니다. 이를 통해 각국의 보건 문제에 대한 인식을 높이고, LMICs의 보건 개선을 위한 정책 수립에 기여할 수 있는 실질적 근거를 제공하고자 합니다. 또한, 국제 보건 데이터를 활용하는 국내 연구자 및 학생들이 증가할 수 있도록 데이터

활용 방법을 익히고 공유하는 것이 필요했습니다.

3) 연구 목적

본 연구 모임의 주된 목적은 각국의 최신 DHS 데이터(Phase-8)를 분석하여 국가별 주요 보건 이슈를 파악하고, 이를 바탕으로 구체적인 중재 전략을 수립하는 데 필요한 근거 자료를 제공하는 것입니다. 연구모임은 초보 연구자도 데이터를 쉽게 다루고 각 변수의 의미를 이해할 수 있도록, DHS 데이터의 접근성과 이해도를 높이고자 합니다. 이를 통해 연구자들이 국가별 보건 지표를 활용한 분석을 수행하고, 이를 기반으로 한 중재 전략 수립에 능숙해질 수 있도록 지원하고자 합니다.

또한, 본 연구 모임은 국제 보건 지표에 대한 초보 연구자들의 이해를 증진하고, LMICs의 다양한 보건 이슈에 체계적으로 접근할 기회를 제공합니다. 연구자들은 최신 데이터를 기반으로 보건 이슈를 과학적이고 종합적으로 분석하며, 그 결과를 바탕으로 효과적인 중재 전략을 제안할 수 있습니다. 이 연구 모임의 활동은 학문적 차원을 넘어, 향후 연구와 정책 개발에 실질적으로 기여하며 국제 보건 발전을 촉진하는 데 목적이 있습니다.

2. 본론

1) 연구모임 과정

국제보건 데이터 마이닝 그룹의 연구모임은 다양한 이력 배경을 가진 간호사 전공자와 통계학 전공의 연구자들이 모여 각자의 전문 지식을 공유하고, DHS 데이터 분석에 필요한 기초 지식을 습득하는 과정으로 시작되었습니다. 2024년 3월 16일에 진행된 1차 연구모임에서는 각 연구자가 자신의 연구 배경과 목표를 소개하고, 앞으로의 모임 운영 계획과 역할 분담을 논의했습니다. 모임 초기에는 각자가 연구에 필요한 DHS 데이터에 대한 이해를 높이고, 통계 분석에 필요한 기초 자료를 준비하는 데 집중했습니다.

특히 연구 초기 단계에서는 각 분야의 연구자들이 주제를 선정하여 발제를 맡았습니다. 예를 들어, 방글라데시와 아이티 등 다양한 국가에서 시행된 기존 연구 사례를 함께 검토하여, 각국의 보건 이슈가 다른 배경에서 어떻게 접근되고 해결되는지를 파악할 수 있었습니다. 이를 통해 국제 보건 문제에 대한 이해를 넓히고, DHS 데이터 분석 시 유용한 시사점을 도출했습니다. 그 중 하나는 청소년 대상의 식수위생 사업과 양질의 보건교육 제공 문제와 여성의 건강 지표를 악화시키는 조혼 제도가 만연한 아시아/아프리카 지역의 여성청소년 성교육 및 생식 건강 중재에 대한 논의가 활발한 점이 본 연구 모임을 통해 분석하고자 하는 주요 변수들이었습니다.

또한, 외부 전문가 초청 특강을 통해 논문 작성법과 데이터 분석 사례를 학습하며 모임을 발전시켰습니다. 전문가의 연구 경험과 논문 작성 노하우는 연구자들이 DHS 데이터를 보다 효과적으로 분석하고 결과를 체계적으로 정리하는 데 큰 도움이 되었습니다. 이를 통해 모임 초반에는 이론적 배경과 실무적인 데이터 분석 기술을 강화한 후, 각자의 연구 주제에 맞는 데이터 분석에 집중할 수 있는 기반을 마련했습니다. 국제보건 데이터 마이닝 그룹 연구모임은 다방면의 지식을 결합하여 국제 보건 문제에 대한 심층적인 이해를 도모하고, 데이터 분석에 필요한 역량을 축적하는 중요한 초석이 되고자 이와 같이 시작하여 진행했습니다.

2) 연구방법

본 연구는 DHS 데이터를 활용하여 LMICs의 다양한 보건 지표를 비교 분석하고, 국가별로 최적화된 중재 전략을 수립하는 것을 목표로 합니다. DHS 데이터는 각국의 여성 건강 지표(IR: Individual recode)를 중심으로 인구, 건강, 영양 상태를 심도 있게 다루고 있어 국제 보건 연구

에 필수적인 자료입니다. 그러나 방대한 데이터의 특성상 단순히 데이터를 다운로드 해서 엑셀과 같은 기존 MS office 프로그램에서 분석하는 것만으로는 효율적인 분석을 수행하기에 한계가 있습니다.

<표 1> 국제보건 데이터 마이닝 그룹 모임 전반기 발제 및 특강

<h3>1회차 모임</h3> <p>3월 16일 (토)</p> <p>[연구 모임 운영 및 DHS 개황]</p> <ul style="list-style-type: none"> • 대표: 신보혜 — 품씨연구회 전체 공지 전달, 모임일정 확인, 장소 예약 및 공지, 중간공유(8월)/결과공유(11월) 참석, 최종보고서 취합 및 제출 • 준비: 임유미 — 매 활동 발제자 준비사항 확인, 외부연자 준비사항 확인, 학회 참석 & IRB 심의면제 준비 • 예산: 김하나 — 11/15까지 정산보고서 제출, 매 활동 예산 내역 정리, 추가 증빙자료 요청 및 취합, 예산변경 필요 시 전체 회의 후 변경 작성 후 대표에게 제출 • 중앙: 강유미 — 매 활동 모임 사진 촬영, [사유서] 증명 필요한 경우 직접 작성 및 해당자에게 작성 요청 • 기록: 서희연 — 매 활동 기록 술과난눔-품씨연구회4기 폴더에 작성하고 업로드 활동 기록 양식 • 자료: 최지선 — 뉴스레터 작성, 중간공유 자료 작성, 결과보고서 편집, 결과보고 발표 자료 작성 • 데이터: 신동혁 — 최종 산출물 연동된 데이터 N개 • 산출: 정지은 — 데이터 분석표, 중재 전략 보고서, 모임 과정 단행본 	<h3>DHS Themes</h3> <p>여성/아동 분야 데이터</p> <ul style="list-style-type: none"> • Child Health • Nutrition of Children and Adults • Malaria • HIV/AIDS • HIV prevalence • Women's empowerment • Water, Sanitation and Hygiene • Domestic Violence • Adult and Maternal Mortality • Female Genital Cutting • Fistula <ul style="list-style-type: none"> • Population and housing • Respondents' Characteristics • Marriage and Sexual Activity • Fertility • Fertility Preferences • Family Planning • Infant and Child Mortality • Maternal Healthcare
<h3>2회차 모임</h3> <p>4월 14일 (일)</p> <p>[DHS 국가 선정 준비, 개인 별개 진행]</p> <ul style="list-style-type: none"> • DHS 국가 선정 준비 • 아시아 3개국, 아프리카 3개국으로 DHS 데이터 신청 원료 (GPT 사용) • 발제 진행 1) 임유미: IRB와 학위논문 과정 2) 김하나: 임상연구와 프로토콜 이해 3) 강유미: 초보자의 연구 논문 검색 방법 및 의료기기 임상연구 방법 	<h3>2회차 모임</h3> <p>임유미: IRB와 학위논문 과정</p> <h4>연구방법</h4> <ul style="list-style-type: none"> 1차: 직접 DATA 수집: 설문조사, 각종 임상시험 (표호프 등) in Depth interview, Focus Group Interview 2차: 공개된 DATA 활용: 각종 공문 데이터 (국가 건강정책연구소, 청소년건강성태조사, 한국복지패널조사, DHS 등등) 문헌: 문헌 자료 활용: 공공기관 발표물, 정책 발표물, 실행 연구 논문, 뉴스 기사 등
<h3>2회차 모임</h3>  <p>김하나: 임상연구와 프로토콜 이해</p> <p>우연의 연속이 이어지면?</p> <p>우리가 살릴 데이터에서도 이런 우연을 만날지도 기</p>	<h3>2회차 모임</h3> <p>강유미: 초보자의 연구 논문 검색 방법 및 의료기기 임상연구 방법</p> 
<h3>3회차 모임</h3> <p>5월 18일 (토)</p> <p>[DHS 국가 선정 준비, 개인 별개 진행]</p> <ul style="list-style-type: none"> • 국제보건의료학회 참석 (5/10) • 국가별 데이터 논문 참조 • 특강 <ul style="list-style-type: none"> - 고대협 (외부 연사): 체계적 문헌고찰과 학술논문 출판 • 발제 진행 <ul style="list-style-type: none"> - 신동혁: Jamovi, R 프로그램을 이용한 데이터 세팅 연습 	<h3>3회차 모임</h3> <p>국제보건의료학회 참석 (5/10)</p> 

3회차 모임

고대협 (외부 연가): 체계적 문헌고찰과 학술논문 출판

Methodology - PICO 선정

핵심질문의 구성요소

Focus (Question/Problem)

- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)

Intervention (Intervention/Problem)

- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)

Comparison

- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)

Outcome

- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)

핵심질문 작성

- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)
- 학술 논문 (Key question)은: 문헌고찰(Clinical question)

3회차 모임

김동혁: Jamovi, R 프로그램을 이용한 데이터 세팅 연습

Latest Features

최신 버전을 선택해주세요. 추천 버전에서는 호환성 문제가 없습니다.



2.3.28 solid
2.5.4 current

반도분석

실습파일들을 불러옵니다.



4회차 모임

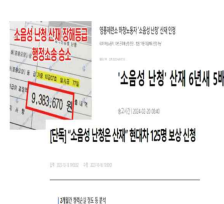
6월 22일 (목)

[개념 공개 진행]

- 발제 진행
 - 신동혁: DHS 데이터 전처리와 통계패키지 사용법-2
 - 서희연: 산업보건 현장에서 시 활용
- 특강
 - 김유나 (외부 연사): 국제 보건 DHS 데이터 연구 및 수집 경험 공유

4회차 모임

서희연: 산업보건 현장에서 AI 활용



소용성 남창 신재 6년새...
3,393,600 원

출처

출판

소용성 남창 신재가 저술한 '산업보건 현장에서 AI 활용'은 다음과 같습니다.

- 산업보건 현장에서 AI 활용을 위한 기초 지식과 AI 활용 사례를 소개합니다.
- 산업보건 및 부상 관련 주요 이슈가 소용성 남창 신재에 실려 있습니다. 저자 소개 및 출판 정보도 포함되어 있습니다.
- 산업보건 현장에서 AI 활용을 위한 기초 지식과 AI 활용 사례를 소개합니다. 저자 소개 및 출판 정보도 포함되어 있습니다.
- 산업보건 현장에서 AI 활용을 위한 기초 지식과 AI 활용 사례를 소개합니다. 저자 소개 및 출판 정보도 포함되어 있습니다.
- 산업보건 현장에서 AI 활용을 위한 기초 지식과 AI 활용 사례를 소개합니다. 저자 소개 및 출판 정보도 포함되어 있습니다.

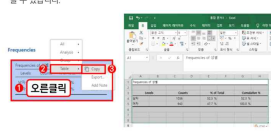
이 책은 산업보건 현장에서 AI 활용을 위한 기초 지식과 AI 활용 사례를 소개합니다. 저자 소개 및 출판 정보도 포함되어 있습니다.

4회차 모임

김동혁: DHS 데이터 전처리와 통계패키지 사용법-2

결과표 작성

결과표 위에서 오른쪽이면 결과표를 복사할 수 있고, 역방향 그대로 붙여넣을 수 있습니다.



결과표 작성

변수 설명, 성별, 휴대전화 보유 여부, 휴대전화 종류 등 빈도분석한 변수명을 순서대로 입력합니다.

변수	구분	N	%
성별			
휴대전화 보유			
휴대전화 종류			
휴대전화 보유			
휴대전화 종류			
휴대전화 보유			
휴대전화 종류			

4회차 모임

김유나 (외부 연가): 국제 보건 DHS 데이터 연구 및 수집 경험 공유


지역 소개

리코나보섬 현황

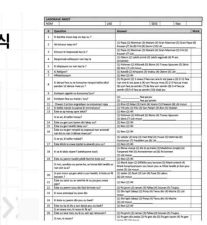
인프라의 용리

수인성 질병 예방

극심한 가뭄



설문지 양식



따라서 본 연구팀은 R과 Jamovi라는 두 가지 도구를 통해 데이터 정제 및 분석 과정을 단계적으로 진행했습니다. 이러한 방법론을 선택한 이유는 데이터의 신뢰성과 분석의 정확성을 높이기 위함이며, 특히 복잡한 데이터를 체계적으로 다루어 실질적인 정책 제안을 위한 기초 자료를 마련하고자 했기 때문입니다.

(1) DHS 데이터 획득 과정

DHS 데이터는 미국 국제개발처(USAID)에서 제공하며, 각국의 보건 및 인구 통계를 집계한 고품질 자료입니다. 하지만 데이터의 특성상 민감한 정보가 포함되어 있어 연구자는 데이터를 단순히 별도의 절차 없이 클릭 한 번으로 다운로드 할 수 없으며, 적합한 연구 계획서 제출 및 승인 절차를 거쳐야만 접근할 수 있습니다.

① 회원 가입 및 연구 계획서 작성

연구자는 먼저 DHS 웹사이트에서 회원 가입을 한 후, 연구 목적과 구체적인 데이터 사용 계획을 포함한 연구 계획서를 제출해야 합니다. 이 과정에서 연구의 주제와 목적, 필요한 변수, 예상 분석 방법 등을 명시해야 하며, 이를 통해 DHS 데이터가 연구의 적합성을 충족하는지 검토받습니다. 이처럼 연구 계획서를 제출하여 승인받는 이유는 데이터의 민감성과 그 중요성에 기인합니다.

다. 특히, DHS 데이터에는 모성 사망률, 아동 건강, 가족계획과 같은 민감한 정보가 포함되어 있어, 데이터를 오용하지 않도록 하기 위함입니다. 승인이 완료되면 회원가입 했던 이메일 계정으로 승인 레터가 오면 간직하고 있다가 향후 연구 시작 후 논문 작성 시, ‘심의면제’ IRB(기관생명윤리위원회) 제출 문서에 첨부하면 됩니다.

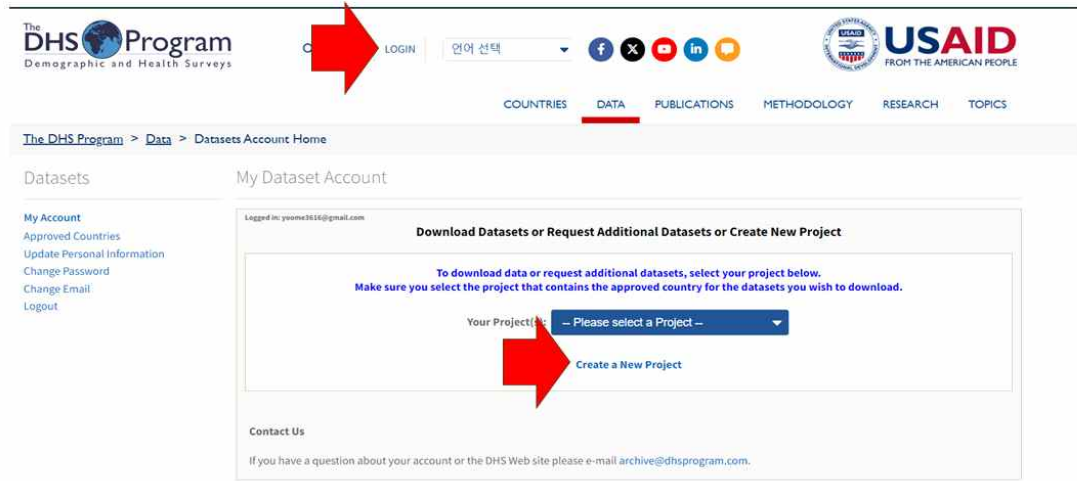
② 데이터 승인 후 다운로드

연구 계획서가 승인되면, 연구자는 DHS 웹사이트에서 필요한 국가별 데이터를 다운로드 할 수 있습니다. DHS 데이터는 국가별로 구분되어 있으며, SPSS, STATA 등 여러 통계패키지에 불러들일 수 있는 ‘.sav’, ‘.dta’ 와 같은 확장자로 제공됩니다. 각 국가의 보건 상황에 대한 대규모 데이터를 제공하기 때문에, 연구자는 필요한 국가와 변수를 미리 파악하여 효율적으로 데이터를 선택해야 합니다. 데이터가 광범위한 만큼, 각 연구자가 추구하는 분석 목적에 맞는 데이터를 정확히 선택하고 필터링하는 것이 중요한 이유이기도 합니다.



<그림 15> DHS dataset AuthLetter 승인 레터

연구 과정 https://dhsprogram.com/data/dataset_admin/



<그림 16> DHS 로그인 및 create a new project 클릭하여 승인받으면 상단의 your project 클릭하여 원하는 국가의 데이터 다운로드 페이지로 이동할 수 있다.

Household Recode	Pregnancy and Postnatal Care Recode
<input type="checkbox"/> GHHR8BDT.ZIP	<input type="checkbox"/> GHNR8BDT.ZIP 1.83 MB Stata dataset (.dta)
<input type="checkbox"/> GHHR8BFL.ZIP	<input type="checkbox"/> GHNR8BFL.ZIP 1.99 MB Flat ASCII data (.dat)
<input type="checkbox"/> GHHR8BSD.ZIP	<input type="checkbox"/> GHNR8BSD.ZIP 2.47 MB SAS dataset (.sas7bdat)
<input type="checkbox"/> GHHR8BSV.ZIP	<input type="checkbox"/> GHNR8BSV.ZIP 2.13 MB SPSS dataset (.sav)
Individual Recode	Household Member Recode
<input type="checkbox"/> GHIR8B.ZIP	<input type="checkbox"/> GHPR8BDT.ZIP 3.89 MB Stata dataset (.dta)
<input type="checkbox"/> GHIR8BDT.ZIP	<input type="checkbox"/> GHPR8BFL.ZIP 4.39 MB Flat ASCII data (.dat)
<input type="checkbox"/> GHIR8BFL.ZIP	<input type="checkbox"/> GHPR8BSD.ZIP 5.50 MB SAS dataset (.sas7bdat)
<input type="checkbox"/> GHIR8BSD.ZIP	<input type="checkbox"/> GHPR8BSV.ZIP 6.19 MB SPSS dataset (.sav)
<input type="checkbox"/> GHIR8BSV.ZIP	Geographic Datasets
Children's Recode	<input type="checkbox"/> GHG8AFL.ZIP 75.0 KB Shape file (.shp)
<input type="checkbox"/> GHKR8BDT.ZIP	Geospatial Covariates
<input type="checkbox"/> GHKR8BFL.ZIP	<input type="checkbox"/> GHG8AFL.ZIP 835 KB Comma delimited data (.csv)
<input type="checkbox"/> GHKR8BSD.ZIP	
<input type="checkbox"/> GHKR8BSV.ZIP	
Men's Recode	
<input type="checkbox"/> GHMR8B.ZIP	
<input type="checkbox"/> GHMR8BDT.ZIP	
<input type="checkbox"/> GHMR8BFL.ZIP	
<input type="checkbox"/> GHMR8BSD.ZIP	
<input type="checkbox"/> GHMR8BSV.ZIP	

<그림 17> 여러 Recode가 있으며 코드북에서 Recode 별 설명도 있다. 기본 분석에는 Individual Recode를 내려받아서 사용하면 된다. 여성 개인별 CASEID를 기준으로 변수들이 가로로 붙어있다.

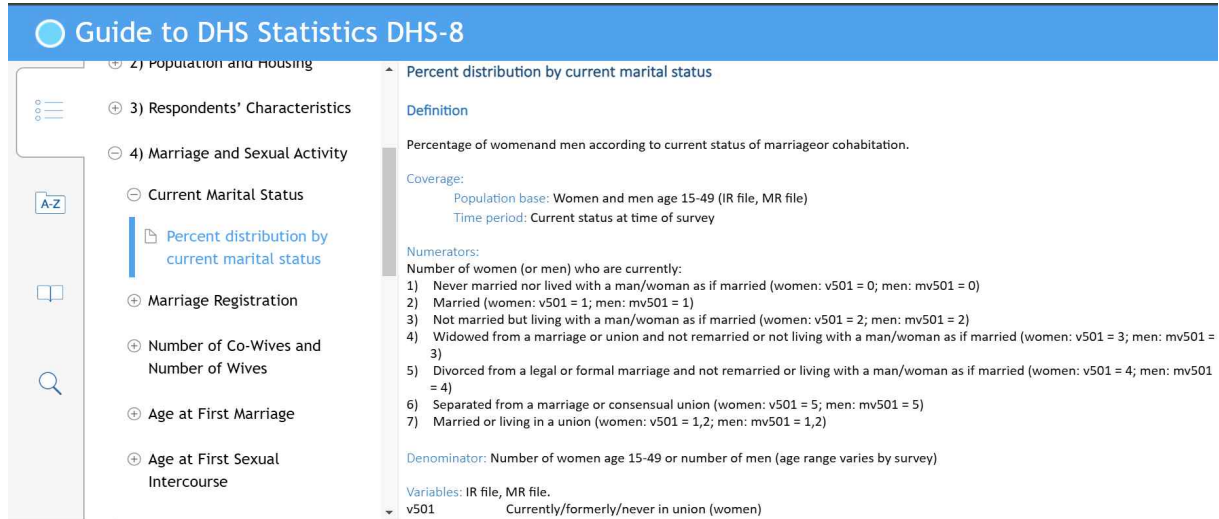
③ DHS 데이터 코드북 활용

데이터를 다운로드 한 후에는 코드북을 사용하여 데이터 형태와 변수명을 확인합니다. DHS 코드북은 데이터 변수에 대한 설명, 코드화된 변수값, 측정 단위 등을 상세히 제공합니다. 코드북을 통해 변수의 의미와 측정 단위, 데이터 수집 방법 등을 파악함으로써, 연구자는 각 국가와 변수의 특성에 따른 보정 작업을 수행할 수 있습니다. 예를 들어, 특정 국가에서는 가족계획에 대한 응답이 문화적 요인에 따라 다르게 나타날 수 있기 때문에 이러한 요소를 사전에 파악하고 해석하는 것이 데이터 분석의 신뢰성을 높이는 데 필수적입니다.

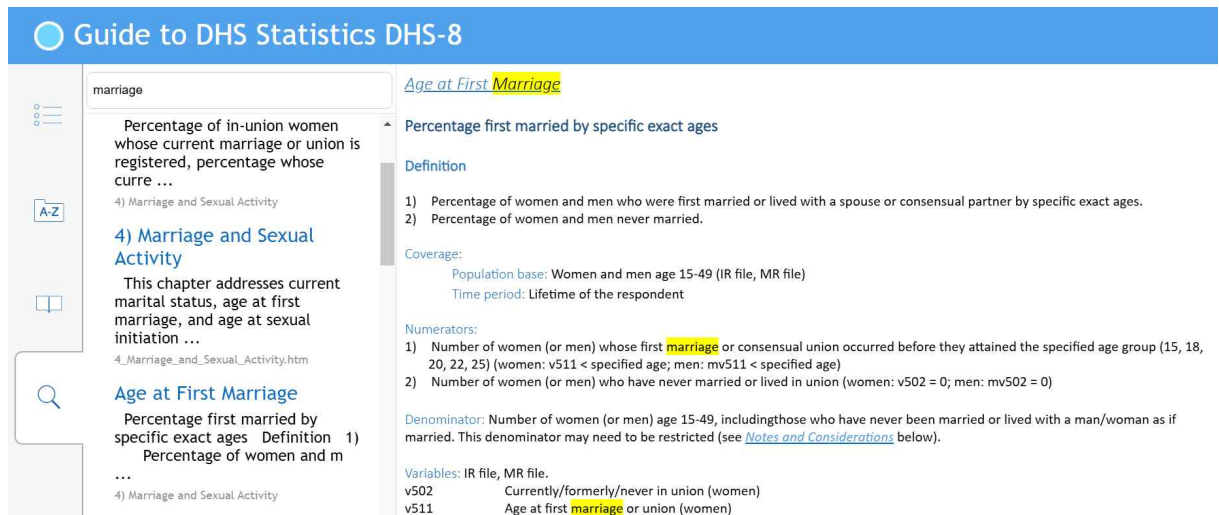
또한 DHS 데이터 코드북을 PDF로 다운받지 않고 사용할 수 있는 방법도 있습니다. 아래 링크로 접속하면 각 데이터셋마다 분석에 필요한 변수명과 변수 설명과 결과가 나와 있습니다. 아동

청소년의 혼인 연령과 교육 수준을 본다고 한다면 관련 키워드를 검색하여 보거나 관련 카테고리를 클릭하여 변수를 찾아서 분석할 수 있습니다.

<https://www.dhsprogram.com/Data/Guide-to-DHS-Statistics/index.cfm>



<그림 18> 카테고리에서 변수를 찾는 화면



<그림 19> 키워드로 검색하여 변수를 찾는 화면

(2) R을 사용한 데이터 전처리 및 정제

DHS 데이터는 원시 데이터 형태로 제공되며, 이를 연구 목적에 맞게 가공하고 분석하기 위해 R을 활용한 데이터 전처리 과정을 거쳤습니다. R은 데이터 분석에 최적화된 프로그래밍 언어로, 대규모 데이터를 다루기 적합한 기능을 제공합니다. 모든 분석 과정은 아래 설명과 사진을 통해 따라하실 수 있으며 본 자료를 활용한 챗GPT 프롬프트 제작을 하였기에 유료로 구독하는 연구자님들은 챗GPT의 “재단법인 숲과나눔 DHS prompt”를 검색해서 사용하실 수도 있습니다.

① 데이터 불러오기

데이터를 불러올 때는 `haven` 패키지를 이용하여 `.sav` 파일 형식을 R 환경에서 사용할 수 있도록 하였습니다. 이를 통해 데이터를 R로 불러온 후, `head(data)`, `dim(data)`, `View(data)`

명령어를 사용하여 데이터의 형태와 구조를 파악했습니다. 이 과정에서 데이터의 열(변수)과 행(개별 응답자 수)을 확인함으로써 연구에 필요한 데이터 구조를 명확히 이해할 수 있습니다. DHS 데이터는 다차원적이고 방대한 정보를 포함하고 있어, 데이터를 효율적으로 관리하고 필요한 정보만을 선별하는 작업이 필수적입니다.

② 필요 변수 추출 및 정제

본 연구는 국가별 보건 이슈와 관련된 특정 변수를 다루기 때문에, 필요 없는 변수를 제외하고 원하는 변수만 추출했습니다. jamovi가 직관적인 인터페이스가 장점인 반면, 무료이면서 초보자가 사용하기 좋은 통계패키지다 보니 대용량 변수가 있는 데이터를 불러오기에는 오류가 자주 생기는 단점이 있습니다. 그래서 R에서 주요 변수를 추출하거나 가중치를 계산하는 작업을 1차 처리하고 .csv로 저장한 뒤 jamovi로 불러오기 하여 분석에 사용했습니다. 예를 들어, CASEID와 주요 변수를 선택하여 새로운 데이터셋을 만들었으며, 이를 `select()` 함수로 선별해 관리했습니다. 이처럼 데이터에서 필요한 변수만을 추출함으로써, 데이터를 보다 효율적으로 분석하고 의미 있는 결과를 도출할 수 있는 기반을 마련하였습니다. 연구에서 필요한 데이터만을 별도로 저장하고 관리함으로써, 향후 분석 과정에서 불필요한 처리 시간을 줄일 수 있었습니다.

<표 2> 조혼과 여성건강 지표를 중심으로 추출할 변수

IR file	Description
v024	Region
v025	Type of place of residence
v044	Selected for Domestic Violence module
v106	Highest educational level (women)
v133	Education in single years (women)
v155	Literacy (women)
v169a	Owens a mobile telephone (women)
v170	Has an account in a bank or other financial institution (women)
v212	Age at first birth
v312	Current contraceptive method
v313	Current use by method type
v501	Marital status: 0 -no / 1-5 -yes
v502	Number of currently married women (v502 = 1)
v504	whose current marriage is registered with civil authority (v542 = 1 or v544 = 1)
v511	Age at first marriage or union
v743a	Person who usually makes decisions about: respondent's health care (women)
v743b	Person who usually makes decisions about: large household purchases (women)
v743d	Person who usually decides on visits to family or relatives (women)
v754cp	Reduce risk of getting HIV: always use condoms during sex (women)
v754dp	Reduce risk of getting HIV: have 1 sex partner only, who has no other partners (women)

```

1 #데이터 불러올 패키지 설치
2 install.packages("haven")
3 install.packages("readxl")
4 #excel파일은 "readxl", csv는 따로 설치할 필요가 없음
5 #패키지 로드 필수
6 #매번 새로 열때마다 library는 ctrl+Enter로 할당해야 함함
7 library(haven)
8 library(readxl)
9
10 #데이터 불러오기
11 kenya_ind <- read_sav("D:\\6_기타활동\\2024_2월
-11월_풀씨연구회_4기_국보마\\DHS_Data
handling\\Kenya_individual\\KEIR8BFL_부분.sav")
12
13 head(kenya_ind)
14 dim(kenya_ind)
15 view(kenya_ind)
17 #Main package dplyr
18 install.packages("dplyr")
19 library(dplyr)
20 install.packages("tidyverse")
21 library(tidyverse)
22
23 kenya_ind$
24
25 Kenya_data1<-kenya_ind %>%
26   select(CASEID, V012, V013, V024, V025, V105, V113, V116,
V119, V120, V121, V122, V123, V124, V125, V127,V128, V129,
V130, V155, V157, V158, V159, V160, V161, V171A, V171B, V176,
V177, V190)
27
28 library(dplyr)
29 library(haven)
30
31 Kenya_data1 <- Kenya_data1 %>%
32   mutate(across(where(~ inherits(., "haven_labelled")), as
.character))
33
34 write.csv(Kenya_data1, "D:\\6_기타활동\\2024_2월
-11월_풀씨연구회_4기_국보마\\DHS_Data
handling\\모임내용\\kenya_ind_1.csv")

```

<그림 20> R로 원본 DHS 데이터 변수 추출하여 저장하기 (#초록색은 설명 부분)

③ 결측치 및 이상치 처리

DHS 데이터는 국가별, 또는 질문 응답에 따라 결측치나 이상치가 발생할 수 있습니다. 예를 들어, 나이와 같은 변수에서 음수 값이 나타나는 경우 이는 데이터 설계의 결측치를 의미합니다. 결측치를 파악함으로써 데이터의 신뢰성과 유효성을 확보할 수 있습니다. 기술통계를 통해 이상치를 확인한 후, R을 사용하여 이상치를 결측치로 처리하거나 적절한 대체 값을 입력하는 방법을 선택합니다. 다만, 전처리 단계에서 모든 변수를 일일이 결측값을 채워 넣는 것은 어렵기 때문에 기본 분석(빈도, 기술통계)을 통해 NA값을 확인합니다. 이후 관계나 인과성을 파악하는 심화 분석을 시행할 때 기존 문헌을 고찰하여 공동 연구진들이 토론하여 결정하여 결측치를 처리합니다. 이 작업은 분석 결과의 왜곡을 방지하고, 데이터의 일관성을 유지하는 데 중요한 역할을 합니다.

④ DHS 데이터 가중치(Weight)

표본 설계가 복잡하고 다단계로 이루어진 DHS 데이터에서는 가중치를 적용하지 않으면 분석

결과가 왜곡될 수 있습니다. DHS는 데이터의 대표성을 확보하기 위해 가중치를 제공하며, Phase-8 데이터에서도 이 가중치를 사용해야만 신뢰성 있는 분석 결과를 얻을 수 있습니다. 일반적으로 DHS의 가중치 값은 100만을 기준으로 스케일링되어 있습니다. 따라서 가중치를 적용하기 전에 100만으로 나누어야 합니다. Jamovi에서는 직접 가중치를 적용하는 기능이 제한적이기 때문에, R이나 Stata 같은 다른 소프트웨어를 사용하여 가중치를 적용한 후, 그 결과를 Jamovi로 가져오는 방법입니다. R에서 select() 함수로 변수를 추출하여 가져온 것처럼 survey 패키지나 srvyr 패키지를 사용하여 가중치를 적용한 후 결과를 내보낼 수 있습니다.

데이터 가이드북에 따르면 가중치 변수가 지정되어 있어 가구 recode는 hv005, 개인 recode는 v005가 있습니다. R에서 survey 패키지를 설치한 뒤 가중치 조정을 위해 df\$weight <- df\$v005/1000000을 입력합니다. 이때 100만이라는 숫자는 DHS에서 지정한 스케일링 단위 값입니다. DHS 데이터는 다단계 층화 표본 설계를 따르므로, 분석에 가중치와 함께 층화 변수(stratum)와 집락 변수(cluster)를 함께 지정해 표본 설계를 반영해야 합니다.

```
# survey 디자인 객체 생성
survey_design <- svydesign(
  id = ~v021,      # 집락 변수(cluster)
  strata = ~v023, # 층화 변수(stratum)
  weights = ~weight, # 가중치
  data = df,
  nest = TRUE
)

# 가중치를 적용한 빈도 분석 (v106 변수의 빈도)
svytable(~v106, design = survey_design)
```

<그림 21> Survey 패키지를 이용한 가중치 변수 적용

(3) Jamovi를 통한 데이터 분석 및 시각화

R을 통해 정제된 데이터는 Jamovi를 통해 기본적인 통계 분석을 수행하고 시각화하여 연구 결과를 보다 직관적으로 해석할 수 있도록 하였습니다. Jamovi는 통계 초보자도 쉽게 사용할 수 있는 GUI 기반의 통계 프로그램으로, 연구팀에서 통계 분석의 접근성을 높이기 위해 사용되었습니다.

① Jamovi 설치 및 환경 설정

Jamovi는 무료로 제공되며, 공식 웹사이트에서 다운로드 한 후 설치할 수 있습니다. 설치 후 프로그램을 실행하면 데이터 불러오기와 분석 준비가 완료됩니다. GUI 환경에서 데이터를 입력하고 바로 분석할 수 있어 R에 익숙하지 않은 연구자도 쉽게 분석에 참여할 수 있었습니다. 이처럼 Jamovi의 간단한 설치와 사용자 친화적인 인터페이스는 본 연구모임처럼 통계 패키지 연구 초보자들이 초기 투자 시간에 비해 많은 분석을 효율적으로 다룰 수 있도록 돕는 도구로 선택하였습니다.

② 빈도 분석 및 기술통계

데이터를 불러온 후, 빈도 분석과 기술통계를 통해 연구 대상의 특성을 파악했습니다. Jamovi의 `Exploration` 메뉴에서 `Descriptives` 옵션을 사용하여 분석하고자 하는 변수를 선택한 후,

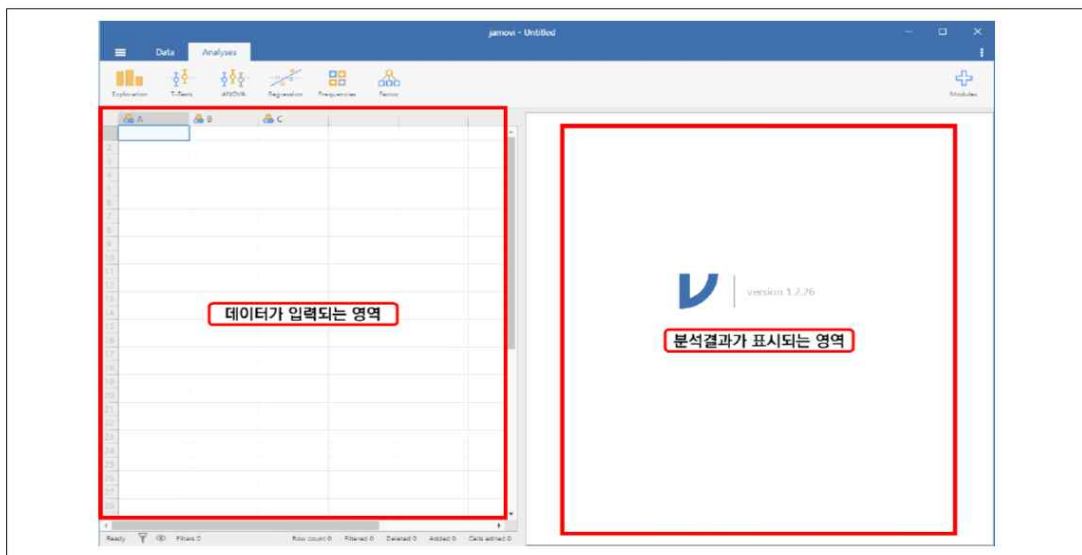
빈도표와 기술 통계량을 산출하였습니다. 이 과정에서 데이터의 기본적인 특성을 파악하고, 연구자가 예상했던 변수의 분포와 일치하는지 확인할 수 있었습니다. 빈도 분석은 주로 각 변수의 분포를 이해하는 데 중요한 역할을 하며, 기술통계를 통해 평균, 표준편차, 최솟값 및 최댓값 등 주요 통계량을 계산하여 데이터의 일반적 특성을 더 구체적으로 이해할 수 있었습니다.

③ 추가 분석: T-검정 및 교차 분석

Jamovi에서는 빈도 분석 외에도 독립표본 T-검정, 교차 분석을 활용해 변수 간의 차이나 상관 관계를 분석했습니다. 예를 들어, 조혼(18세 미만에 사실혼^{cohabitation}인 경우=1)과 주요 사회경제적 수준의 차이를 확인하기 위해 빈도표와 막대그래프로 시각화 자료를 생성하고 T-검정을 수행했습니다. 이러한 추가 분석은 데이터에서 의미 있는 패턴을 발견하고, 각국의 조혼과 여성건강 지표에 대한 심층적인 이해를 돕기 위한 중요한 과정이었습니다.

데이터 입력

jamovi 프로그램 창에서 왼쪽은 데이터 영역, 오른쪽은 분석결과 영역입니다.

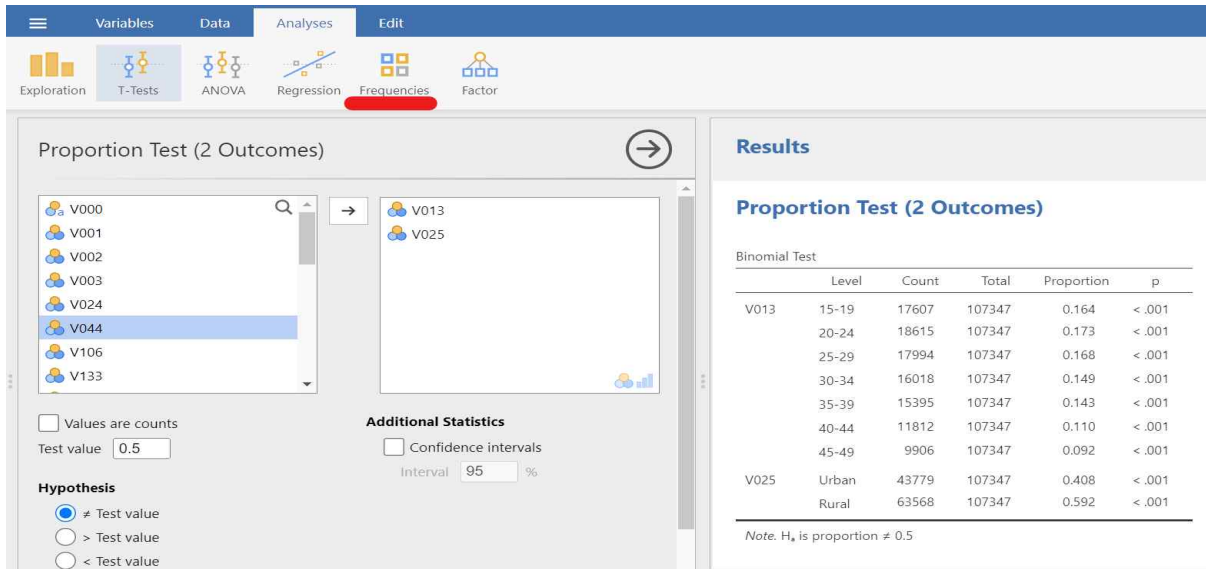


<그림 22> Jamovi 분석 기본 화면

④ 시각화를 통한 결과 도출

Jamovi는 분석 결과를 쉽게 시각화할 수 있는 기능을 제공하여, 연구자는 데이터를 보다 직관적으로 해석할 수 있었습니다. 막대그래프와 선그래프 등을 통해 분석 결과를 시각적으로 표현하여 최종 보고서에 포함시켰습니다. 이러한 시각 자료는 연구 결과를 보다 효과적으로 전달하고, 데이터의 특성을 명확히 보여줌으로써 독자가 쉽게 이해할 수 있도록 돕습니다.

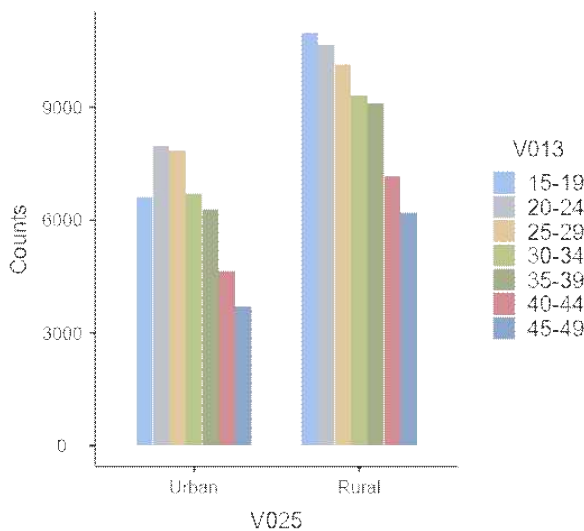
본 연구모임은 DHS 데이터를 획득하여 R을 통해 정제하고, Jamovi를 사용하여 다각도로 분석하는 방식을 통해 국가별 보건 지표를 보다 심층적으로 이해할 수 있었습니다. 데이터를 체계적으로 다루고 분석하는 이 과정은 연구팀이 목표로 하는 보건 정책 개발을 위한 실질적이고 신뢰성 있는 근거 자료를 확보하고자 아래와 같은 분석 결과를 도출했습니다.



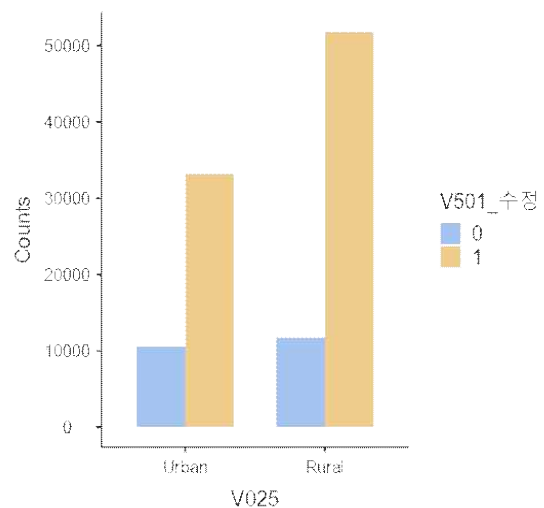
<그림 23> Frequencies 클릭으로 범주형 변수인 설문 항목에 대한 빈도를 확인함

3) 분석 결과

DHS에서 사용하는 도시(Urban)와 농촌(Rural) 지역의 정의는 각 국가의 통계청 및 조사 기관이 설정한 기준에 따라 결정됩니다. 일반적으로 도시 지역은 인구 밀도가 높고 기반 시설, 행정 서비스, 상업 시설 등이 잘 발달한 지역을 의미하며, 농촌 지역은 상대적으로 인구 밀도가 낮고 농업에 의존하는 생활 형태가 주를 이루는 지역으로 정의됩니다. 이러한 분류는 조사국의 사회경제적 특성에 따라 다를 수 있습니다. 연령 분포와 대상자의 혼인 분포를 도시/시골 지역에 따라 분석했을 때 고르게 분포하는 것을 알 수 있으며 시골 지역의 대상자 연령군이 더 어리며 혼인 비율이 도시보다 높은 것을 확인할 수 있습니다. 또한, 5개 국가에서 초산모 연령 평균을 조혼 기준 나이인 18세 미만으로 그룹을 나눴을 때 해당 그룹별 평균연령이 18세 미만은 16세 후반, 18세 이상은 이보다 더 분포가 넓은 18세 후반에서 20세 후반으로 각각 다르게 나타났습니다.



<그림 24> 도시/시골 지역에 따른 대상자 연령 분포

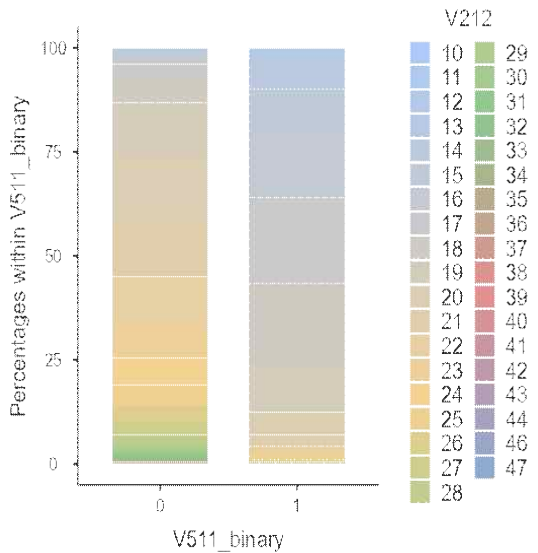


<그림 25> 도시/시골 지역에 따른 혼인 분포

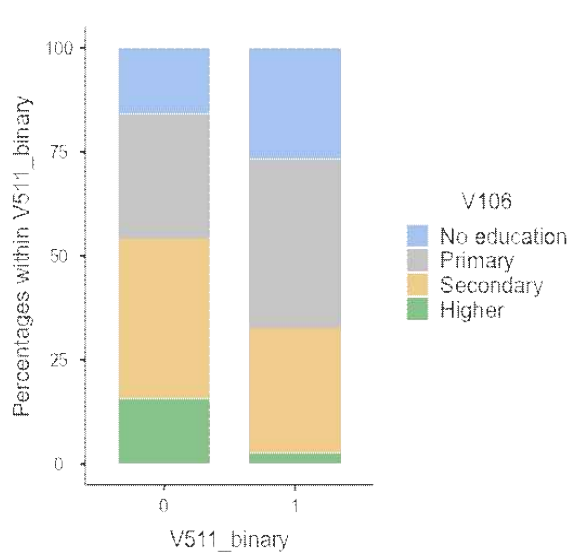
<표 3> 국가별 초산모 연령의 만 18세 기준 그룹별 연령 평균과 표준편차

국가명	나이기준	N	평균	표준편차	평균의 표준오차
방글라데시	18세 미만	1,109	16.62	1.43	0.043
	18세 이상	25,824	18.94	3.686	0.023
가나	18세 미만	325	16.75	1.307	0.072
	18세 이상	10,031	20.94	4.223	0.042
케냐	18세 미만	836	16.55	1.487	0.051
	18세 이상	22,507	19.98	3.955	0.026
네팔	18세 미만	310	16.94	1.277	0.073
	18세 이상	10,425	20.02	3.449	0.034
탄자니아	18세 미만	462	16.71	1.377	0.064
	18세 이상	10,654	19.76	3.685	0.036

본 연구모임에서 조혼을 정의하는 변수는 사실혼 관계를 의미하는 Cohabitation 변수인 V511를 사용하였습니다. 대부분의 국가에서 채택하는 만 18세 이상 법적 혼인 연령을 기준으로 하여 18세 미만인 경우를 0, 18세 이상에 사실혼인 경우를 1로 코딩하여 결과를 확인했습니다. 이를 기준으로 조혼을 한 경우와 조혼을 하지 않은 그룹으로 첫 출산 여성의 평균 연령을 확인했고, 교육 수준도 확인했습니다. 조혼을 한 경우 하지 않은 그룹보다 10대에 출산하는 빈도가 높았고, 최



<그림 26> 첫아이 출생 시 여성의 평균 연령

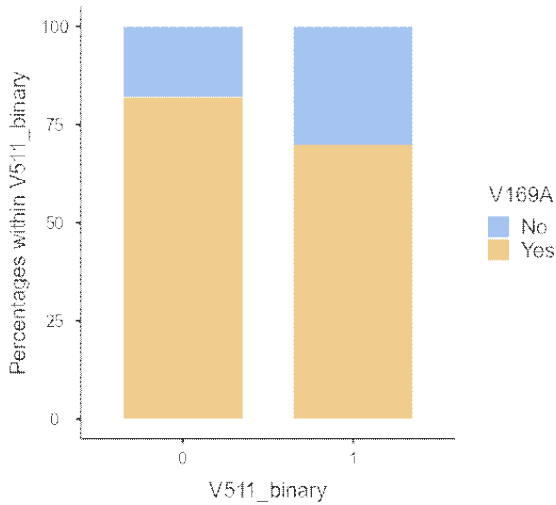


<그림 27> 조혼 여부에 따른 교육 수준

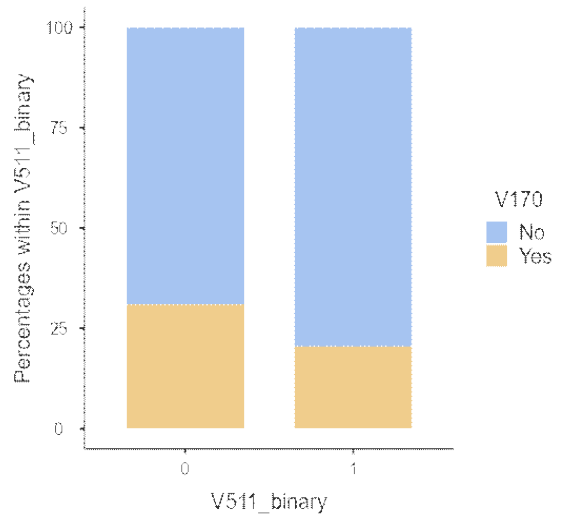
고학력이 초등학교 졸업 수준의 비율이 더 많았습니다.

조혼에 따른 휴대폰 소유 및 개인 계좌를 개설한 비율은 큰 차이가 없었습니다. 하지만 유사한 경향으로 조혼을 한 경우의 대상자들의 소유 수준이 다소 낮았습니다. 이는 다른 매개변수와 더불어 독립표본 t-검정이 필요할 것으로 추후 분석 계획으로 세웠습니다.

조혼에 따른 문맹 수준에서도 조혼을 한 그룹의 문맹 비율이 높은 것으로 나타났습니다. 이는 교육 수준의 분포와 동일한 경향을 보이는 것이며 기존 참고문헌을 통해 교육수준과 조혼, 그리고 이를 해결했을 경우의 여성의 건강 지표 개선이 어떤 영향을 주는지 확인할 필요가 있는 것으로 결론지었습니다. 이는 추후 교육으로 연결지어 DHS 연구 계획으로 세웠습니다.



<그림 28> 조혼 여부에 따른 휴대폰 소유 분포



<그림 29> 조혼 여부에 따른 개인 계좌 소유 분포

<표 4> 18세 미만에 사실혼(V511)을 한 경우 Literacy(V155) 차이

V511	V155	Counts	% of Total
Cohabitation 18≤	Cannot read at all	9,194	12.3 %
	Able to read only parts of sentence	3,995	5.3 %
	Able to read whole sentence	28,110	37.5 %
	No card with required language	28	0.0 %
	Blind/visually impaired	33	0.0 %
Cohabitation <18	Cannot read at all	11,909	15.9 %
	Able to read only parts of sentence	4,830	6.4 %
	Able to read whole sentence	16,797	22.4 %
	No card with required language	26	0.0 %
	Blind/visually impaired	15	0.0 %

4) 문헌고찰

최근 DHS 데이터를 활용한 연구들은 여성의 사회경제적 요소가 아동 건강과 관련된 다양한 요인을 다각도로 분석하며, 특히 저소득 국가에서 아동 사망률 및 질병 발생에 영향을 미치는 요인을 규명하는 데 주력하고 있습니다. 단순 여성 건강의 결과만을 비교하기 보단 어린이와 청소년이 되는 아동의 건강 위기에 주목하고 있습니다. 본 보고서에서는 주요 연구 주제들을 카테고리별로 정리하고, 각 연구에서 다른 주요 변수들을 개괄적으로 설명합니다.

① 아동 사망률과 영양 상태

아동 사망률은 저소득 국가에서 여전히 중요한 공중보건 문제로 자리하고 있으며, 이와 관련된 인구통계학적 및 환경적 요인들이 연구의 중심에 있습니다. 예를 들어, 방글라데시에서 실시된 연구에서는 아동 사망률이 모의 교육 수준, 출산 간격, 출생 순서와 같은 변수를 통해 영향을 받는다는 점이 밝혀졌습니다. 특히, 모의 교육 수준이 높을수록 아동 생존율이 증가하며, 짧은 출산 간격은 영아의 사망 위험을 높이는 것으로 나타났습니다. 이는 모성 교육과 가정 내 출산 계획이

아동 건강 개선에 중요한 역할을 할 수 있음을 시사합니다.

② 기후 요인과 아동 건강

최근 급격한 기후 변화가 아동 건강에 미치는 영향에 대한 관심이 높아지면서, 극단적 기후 조건이 아동의 감염성 질환 발생에 미치는 영향을 분석한 연구들이 늘어나고 있습니다. 예를 들어, 서부 아프리카 지역을 대상으로 한 연구에서는 극단적인 강수량이 특정 열대 기후 지역에서 아동의 급성 호흡기 감염(ARI) 발생 위험을 증가시키는 것으로 나타났습니다. 또한, 다른 연구에서는 온도 변동성이 아프리카 지역에서 저체중아 출산 위험과 연관이 있다는 점이 확인되었습니다. 높은 온도 변동성은 아동의 출생 체중에 부정적인 영향을 미쳐, 조산 및 저체중 출산을 유발할 가능성이 큽니다. 이러한 연구들은 기후 변화가 아동 건강에 미치는 잠재적 위험 요소로서 기능할 수 있음을 시사하며, 기후-건강 위험 평가의 필요성을 강조합니다.

③ 가정 내 환경과 호흡기 감염

가정 내 환경 요인, 특히 조리 시 사용하는 연료와 조리 장소가 아동의 호흡기 감염 발생에 미치는 영향에 대한 연구가 아프리카 여러 국가에서 수행되었습니다. 실내에서 연기를 발생시키는 연료를 사용하는 가정의 아동은 호흡기 감염에 걸릴 위험이 높아지는 것으로 보고되었습니다. 예를 들어, 탄자니아를 대상으로 한 연구에서는 실내에서 연기를 발생시키는 연료를 사용하는 가정에서 아동의 호흡기 감염 위험이 크게 증가하는 것으로 나타났습니다. 이러한 결과는 가정 내 연료 선택과 조리 환경 개선이 아동 건강에 중요한 역할을 할 수 있음을 보여줍니다.

④ 사회적 요인과 전염성 질환

아동의 건강에는 사회적, 경제적 요인 역시 큰 영향을 미칩니다. 가나에서 실시된 연구에 따르면, 다처제 가정에서 자라는 아동이나 모유 수유를 받지 않는 아동이 설사병에 걸릴 가능성이 더 높았습니다. 또한, 사회경제적 지위가 낮은 가정의 아동들이 전염성 질환에 더 취약한 것으로 나타났습니다. 이는 가정의 경제적 자원과 가족 구조가 아동의 전염성 질환 발생에 중요한 변수로 작용할 수 있음을 시사하며, 이러한 가정 환경을 개선하기 위한 사회적 지원 정책의 필요성을 강조합니다.

⑤ 위생 및 보건 시설과 설사병

설사병은 아동 사망의 주요 원인 중 하나로, 특히 위생 상태가 열악한 국가에서 높은 발병률을 보이고 있습니다. 방글라데시와 캄보디아를 대상으로 한 연구에서는 비위생적인 화장실 시설과 아동 배설물의 부적절한 처리가 설사병 발병과 밀접한 관련이 있음을 밝혔습니다. 이들 연구에서는 위생 시설 개선과 아동 배설물 관리에 대한 교육이 설사병 예방에 효과적일 수 있음을 강조하였으며, 정부와 국제기구의 협력으로 이러한 문제를 해결할 수 있는 정책적 노력이 필요하다는 점을 제기했습니다.

5) 중재전략 도출

이번 연구모임은 DHS 데이터를 활용해 조혼, 여성의 생식 건강, 아동 건강 문제를 포함한 국제 보건 이슈를 심층적으로 분석하고, 이를 해결하기 위한 구체적인 중재 전략을 도출하는 데 중점을 두었습니다. 초보 연구자가 접근하기 쉬운 데이터 핸들링과 DHS 데이터 구득에 대한 방법론을 알려주고자 하는 것이 주된 목표였기에 더 많은 건강 관련 변수들을 분석하지 못한 것이 향후 과제로 남았습니다. 모든 변수들을 심도있게 분석하고 고찰할 순 없었지만 사용할 수 있는 연구 주

제를 탐구한 문헌고찰을 활용해 다음과 같은 국제보건 DHS 데이터를 통한 인사이트를 갖고 중재 전략을 정리했습니다. 다음은 본 연구모임의 분석 결과 및 문헌고찰에서 제안하는 중재 전략입니다.

① 조혼 예방을 위한 교육 프로그램 강화 및 법적 보호

조혼은 여성의 교육 기회 박탈과 생식건강에 악영향을 미치는 주요 요인입니다. LMICs 국가에서 조혼을 예방하고 여성의 권리를 보호하기 위해 지역 사회와 학교를 중심으로 성교육 및 건강 교육 프로그램을 확대해야 합니다. 특히 조혼의 위험성을 강조하고, 여성의 생식건강에 미치는 영향을 교육하는 프로그램을 개발하여 청소년과 부모가 조혼의 위험성을 충분히 인식할 수 있도록 돕는 것이 필요합니다. 법적 보호 장치를 강화하여 법적 결혼 연령을 준수하고, 결혼 연령에 대한 인식을 높이는 캠페인을 통해 지역 사회의 인식 변화를 유도합니다.

② 여성 건강 증진을 위한 보건 서비스 접근성 강화

조혼을 한 여성들이 조기 출산과도 깊은 연관을 가짐으로 인해 여성의 건강에 큰 부담을 줄 수 있으며, 이는 산모 사망률을 높이는 요인으로 작용할 수 있습니다. 지역 사회의 의료 접근성을 강화하여 여성들이 생애 초기부터 건강 검진과 상담을 받을 수 있도록 하는 중재 전략을 제안합니다. 임신 중인 여성과 청소년을 대상으로 무료 또는 저비용의 건강 검진, 영양 상담, 심리 상담을 제공함으로써 조기 출산을 예방하고 산모 건강을 향상시킬 수 있습니다.

③ 위생 환경 개선을 통한 아동 질병 예방

LMICs의 가정 내 환경, 특히 조리 연료와 실내 환경은 아동의 호흡기 질환에 큰 영향을 미칩니다. 가정 내에서 연기 발생이 적은 청정 연료를 사용할 수 있도록 보조금을 지원하고, 지역사회와 협력하여 안전한 조리 환경을 제공하는 것이 필요합니다. 또한, 아동에게 호흡기 감염의 주요 원인과 예방책을 교육하여 건강을 스스로 관리할 수 있도록 돕는 방안도 고려해야 합니다.

④ 기후 변화와 보건 위험에 대한 대응

기후 변화는 특히 아동 건강에 다양한 부정적 영향을 미치고 있으며, 고온이나 극단적 강수량은 질병 발생 위험을 증가시킬 수 있습니다. 본 연구는 기후 변화로 인한 보건 위험에 대응하기 위해, 기후 변화가 보건에 미치는 영향을 주기적으로 평가하고 이에 대한 대응책을 마련할 필요가 있음을 제안합니다. 이를 위해, 기후 변화가 건강에 미치는 영향을 이해할 수 있는 교육 자료와 대응 전략을 지역 사회에 제공하여 기후 적응력을 높이는 것이 중요합니다.

3. 결론

본 연구모임은 풀씨연구회 4기 활동의 일환으로 진행되었으며, IRB 심의면제 승인을 받아 향후 학술 논문으로 발전시킬 예정입니다. DHS 데이터를 활용하여 LMICs의 주요 보건 이슈를 심층적으로 분석하고, 이를 해결하기 위한 맞춤형 중재 전략 수립을 위한 기초 자료를 제공하였습니다. 연구 과정에서 조혼, 여성의 생식건강, 아동의 호흡기 질환 등의 변수를 통해 분석 가능한 보건 문제를 확인하였으며, 이를 통해 교육 프로그램 강화, 의료 서비스 접근성 확대, 위생 환경 개선, 기후 변화 대응과 같은 중재 전략을 제안하였습니다.

특히, 조혼 예방을 위한 교육 프로그램과 법적 보호 강화는 여성 건강과 청소년 발달에 긍정적 영향을 미칠 수 있음을 시사하였습니다. 의료 서비스 접근성 향상 정책은 출산 관련 합병증 감소에 기여할 것으로 기대되며, 아동의 호흡기 질환 예방을 위한 가정 내 환경 개선의 필요성이 확인되었습니다. 또한 기후 변화가 아동 건강에 미치는 영향에 대한 분석은 보건과 환경의 밀접한 연

관성을 재확인하였습니다.

본 연구는 DHS 데이터의 활용 방안을 확대하여 초보 연구자들의 복잡한 데이터 분석 접근성을 향상시켰으며, 대규모 빅데이터를 통한 국제 보건 연구의 진입장벽을 낮추는데 기여하고자 했습니다. 본 연구모임의 산출물은 단순한 데이터 분석을 넘어 실질적인 보건 중재 전략을 도출하고, LMICs의 보건 문제 해결에 기여할 수 있는 가능성을 제시하였다는 점에서 추후 연구에서 관련 연구자들이 활용할 수 있는 데 영향을 미칠 것이라 생각합니다.

참고문헌

- “결혼동의·최저결혼연령·혼인신고에 관한 협약(Convention on Consent to Marriage, Minimum Age for Marriage and Registration of Marriages)(1962)”, 세계법제정보센터,
https://world.moleg.go.kr/web/wli/lgsInfoReadPage.do?CTS_SEQ=4948&AST_SEQ=560
- “공공데이터 개방 및 활용 현황 분석”, KDI 경제정보센터,
<https://eiec.kdi.re.kr/policy/domesticView.do?ac=0000173832>, (2024년 11월 3일).
- “ODA 연구보고서”, KOICA 홈페이지, https://www.koica.go.kr/koica_kr/8136/subview.do, (2024년 11월 3일).
- Ahamad, M.G., & Tanin, F., 2021, Estimation of household smoke-exposure risk using Demographic and Health Survey (DHS) data, *MethodsX*, 8, 101390, DOI: <https://doi.org/10.1016/j.mex.2021.101390>.
- Amugsi, D.A., Aborigo, R.A., Oduro, A.R., Asoala, V., Awine, T., & Amenga-Etego, L., 2015, Socio-demographic and environmental determinants of infectious disease morbidity in children under 5 years in Ghana, *Global Health Action*, 8(1), 29349, DOI: <https://doi.org/10.3402/gha.v8.29349>.
- Balk, D., Pullum, T., Storeygard, A., Greenwell, F., & Neuman, M., 2004, A spatial analysis of childhood mortality in West Africa, *Population, Space and Place*, 10(3), 175-216, DOI: <https://doi.org/10.1002/psp.328>.
- Das, A.C., 2015, Childhood mortality and child nutritional status of Bangladesh: A review on Demographic and Health Survey, *Journal of Current and Advance Medical Research*, 2(2), 42-46, DOI: <https://doi.org/10.3329/jcamr.v2i2.26246>.
- Pisey, V., Banchonhattakit, P., & Laohasiriwong, W., 2021, The association of socio-demographic and environmental factors on childhood diarrhea in Cambodia, *F1000Research*, 9, 303, DOI: <https://doi.org/10.12688/f1000research.23246.5>.
- Rahman, M.M., Haider, M.R., Moinuddin, M., Rahman, A.E., Ahmed, S., & Khan, M.M., 2018, Determinants of caesarean section in Bangladesh: Cross-sectional analysis of Bangladesh Demographic and Health Survey 2014 Data, *PLoS ONE*, 13(9), e0202879, DOI: <https://doi.org/10.1371/journal.pone.0202879>.
- Roy, S., & Hossain, S.M.I., 2017, Fertility differential of women in Bangladesh demographic and health survey 2014, *Fertility Research and Practice*, 3(16), DOI: <https://doi.org/10.1186/s40738-017-0043-z>.
- R Core Team (2021). R: A Language and environment for statistical computing. (Version 4.1) [Computer software]. Retrieved from <https://cran.r-project.org>. (R packages retrieved from MRAN snapshot 2022-01-01).
- The jamovi project (2022). jamovi. (Version 2.3) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- Traoré, N., Singhal, T., Millogo, O., Sié, A., Utzinger, J., & Vounatsou, P., 2024, Relative effects of climate factors and malaria control interventions on changes of parasitaemia risk in Burkina Faso from 2014 to 2017/2018, *BMC Infectious Diseases*, 24(1), 166, DOI: <https://doi.org/10.1186/s12879-024-08981-2>.
- Uttajug, A., Ueda, K., Seposo, X., & Francis, J.M., 2023, Association between extreme rainfall and acute respiratory infection among children under-5 years in sub-Saharan Africa: An analysis of Demographic and Health Survey data, *BMJ Open*, 13(4), e071874, DOI: <https://doi.org/10.1136/bmjopen-2023-071874>.
- Wang, P., O'Donnell, K.J., Warren, J.L., Dubrow, R., & Chen, K., 2023, Temperature variability and birthweight: Epidemiological evidence from Africa, *Environment International*, 173, 107792, DOI: <https://doi.org/10.1016/j.envint.2023.107792>.