

## 데이터 분석을 활용한 프로젝트 계획서

프로젝트 주제[5점]	지역별 소득 수준과 기대수명 간의 상관관계 분석
학번	이름
3329	조성민

### 1. 문제 설정

주제를 생각한 배경 및 동기(300자 이상)[20점]
<p>현재 기술의 발달과 사람들의 생활 수준이 향상되면서 사회 전체적으로 기대수명이 증가하고 있습니다. 기대수명은 사람이 얼마나 오래 살 수 있는지 예측하는 수치로, 사람의 건강과 직결되는 수치입니다. 하지만 기대수명이 증가한 지금 사회 내에서도 의료 접근성, 주거환경, 스트레스 수준 등에 따라 다르게 나타날 수 있습니다. 이에 따라 저는 지역별 소득과 기대수명 간의 연관성을 분석하여 지역 간 건강 불균형 수준을 알아보고, 이를 바탕으로 어떻게 불균형을 줄이고, 어떤 방향의 지원이 필요할지 생각해보려 합니다. 이를 바탕으로 지역별 불균형 해소와 동시에 올바른 방향의 지원으로 평균 기대수명이 증가하기를 기대합니다.</p>
주제에 대한 자세한 설명(300자 이상)[15점]
<p>이 프로젝트는 국내 시구신고 평균소득월액과 지역별 기대수명지표를 바탕으로 지역별 평균 소득과 기대수명 간의 상관관계를 분석하는 것이 목표입니다. 특히 기대수명지표의 경우 첨부되어있는 보험료 분위 별 기대수명 및 기대수명 격차도 활용하여 같은 지역 내에서의 기대수명 격차가 각 지역별로 어떻게 나타나는지도 파악하려고 합니다. 이를 바탕으로 지역 내의 격차와 소득 간의 관계도 도출하여 지역별 비교도 시행할 예정입니다. 또한 지도를 사용한 시각화를 통해 지역 간 격차를 직관적으로 파악할 수 있게 하며, 지역별 정책적 개입 필요성 및 방향성을 도출하는 것을 목적으로 합니다.</p>

2. 데이터 수집

수집하고자 하는 데이터에 대한 설명(200자 이상)[15점]
<p>첫 번째로 지역별 평균 소득 수준을 알기 위해 국민연금공단_자격 시구신고 평균소득월액을 사용할 것입니다. 해당 데이터는 기준년월, 시군구, 평균소득월액으로 구성되어있습니다.</p> <p>두 번째로 지역별 기대수명을 알기 위해 국민건강보험공단_지역별 기대수명지표를 사용할 것입니다. 해당 데이터는 성별, 시도, 시군구, 평균 기대수며으 보험료1분위, 보험료2분위, 보험료3분위, 보험료4분위, 보험료5분위, 기대수명격차로 구성되어 있습니다.</p>
데이터 수집 방법(200자이상)[15점]
<p>데이터 수집의 경우 공공데이터포털을 사용할 것입니다. 공공데이터포털에서 국민연금공단_자격 시구신고 평균소득월액(<a href="https://www.data.go.kr/data/3046077/fileData.do">https://www.data.go.kr/data/3046077/fileData.do</a>)과 건강보험공단_지역별 기대수명지표(<a href="https://www.data.go.kr/data/15039778/fileData.do">https://www.data.go.kr/data/15039778/fileData.do</a>)를 제공하고 있습니다. 각각 csv 형태로 정보가 제공되고 있으며, 전자의 경우 2025년 1월 기준 데이터가, 후자의 경우 전국은 1년, 시도는 4년, 시군구는 6년분의 자료를 바탕으로한 결과를 제공하고 있습니다.</p>

3. 데이터 분석

데이터 분석을 통해 얻고자 하는 결과[10점]
<p>소득수준과 기대수명의 상관관계 정도를 파악하여 그 강도와 방향을 통계적으로 도출해내고자 합니다. 그리고 이를 바탕으로 낮은 소득수준이 야기하는 건강 불균형으로 이어지는 사회의 구조적 문제를 드러내고자 합니다.</p>
예상 결과[5점]
<p>일반적으로 소득수준이 높은 지역일수록 기대수명도 높을 것으로 예상됩니다. 왜냐하면 소득수준이 높은 지역일수록 건강과 관련된 인프라가 다른 지역에 비해 잘 되어있고, 건강에 투자할 수 있는 시간과 부담이 적으며, 평소 스트레스의 정도가 적을 가능성이 높기 때문입니다.</p>

#### 4. 앞으로의 적용 계획

프로젝트를 위해 추가로 알아야 할 지식[10점]
프로젝트를 진행하기 위해선 데이터를 읽을 때 사용할 pandas, 지도 시각화를 위한 folium Choropleth, 데이터 시각화를 위한 matplotlib과 seaborn 등의 지식을 보유하고 있어야 합니다.
수업 시간에 다뤘으면 하는 분야[5점]
공공데이터를 API요청으로 불러와서 분석하는 예제, xlsx, csv외의 다른 형태의 데이터 분석하는 방법 등 여러 가지 형태의 데이터들을 불러오고 분석하는 방법을 알아보면 좋을 것 같습니다.