

# ICT 융합 특론 실습

## Data Visualization

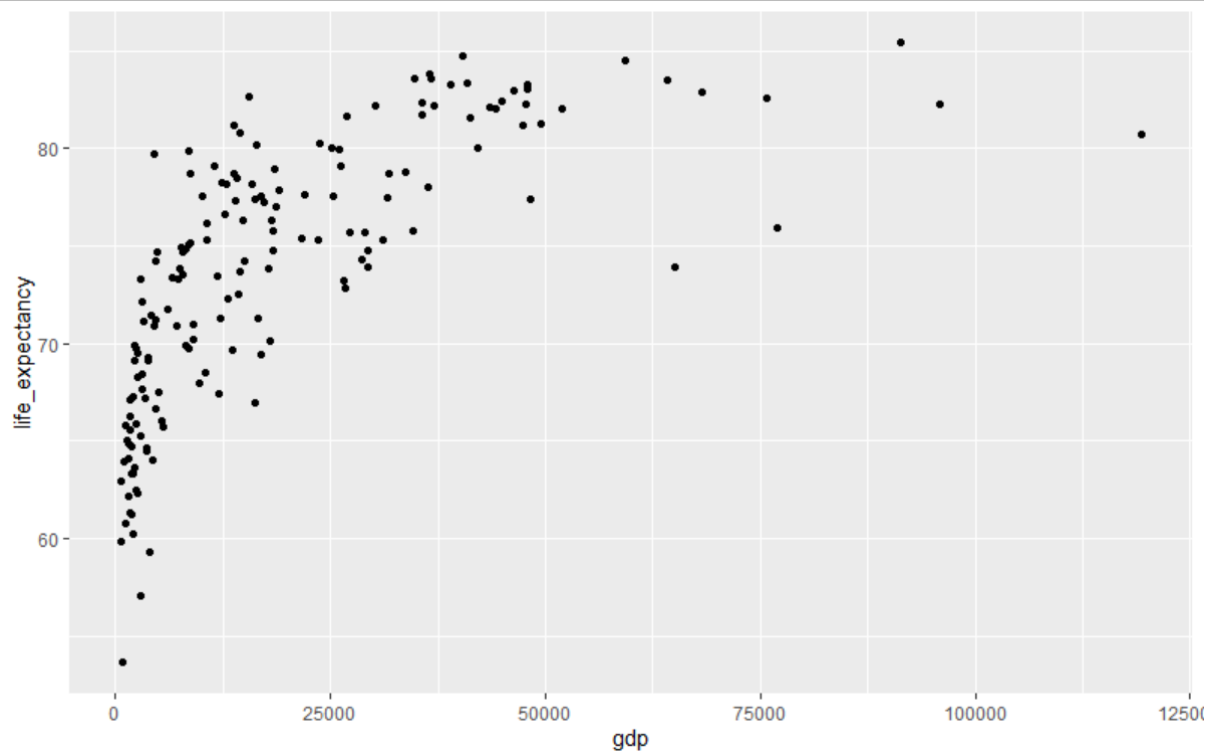
데이터 설명

covid19_210903.csv	2021년 5월 27일까지 국가/지역별 코로나 확진자/사망자/완치자 통계 (일별)
population_total.csv	국가별 인구통계
GM-GDP per capita - Dataset - v26 - data-GDP-per-capita-in-columns.csv	국가별 GDP per capita (인구 1인당 GDP) 통계
GM-Life Expectancy- Dataset - v11 - data-for-countries-etc-by-year.csv	국가별 기대 수명 통계
country_code_with_region.csv	국가별로 지역분류 및 국가코드 등의 정보가 포함된 table
GM-GDP per capita - Dataset - v26.xlsx	GDP 관련 자세한 내용(참고용)
GM-Life Expectancy- Dataset - v11.xlsx	기대수명 관련 자세한 내용(참고용)

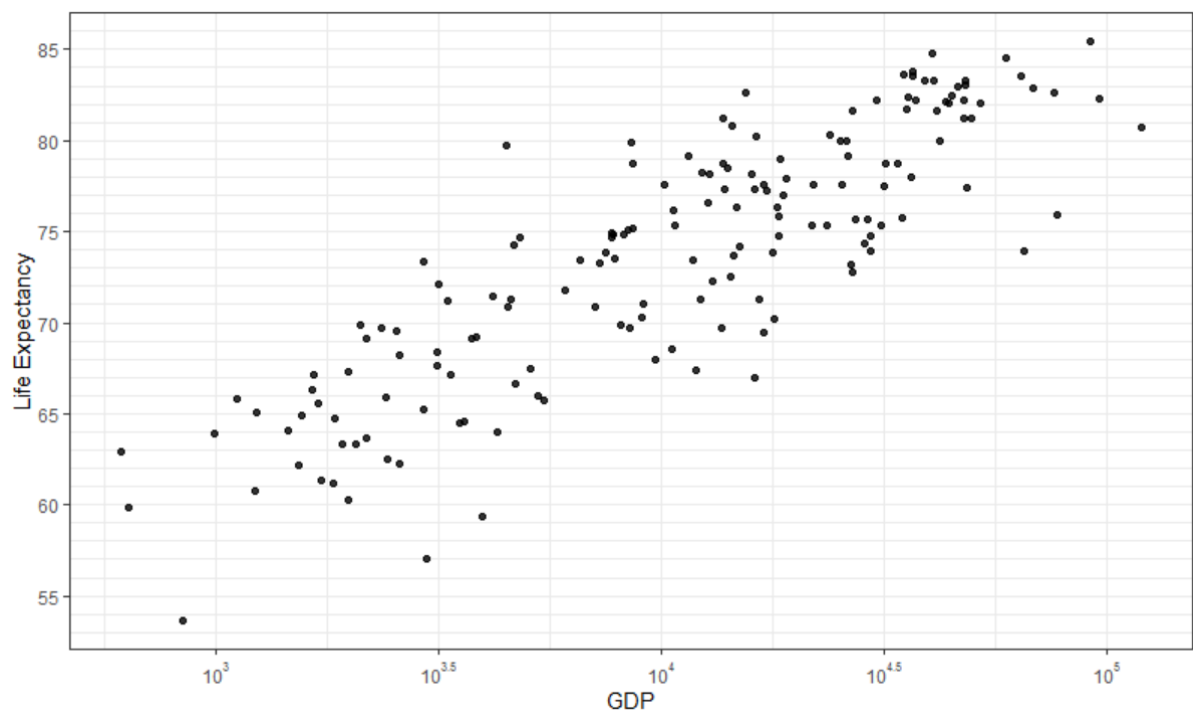
위 데이터를 사용해 아래 Task를 수행하시오.

(시각화 결과물이 예시 그래프와 동일하지 않아도 됩니다. Task의 요구사항을 달성하는 선에서 다르게 그리거나 더 잘 그려도 좋습니다.)

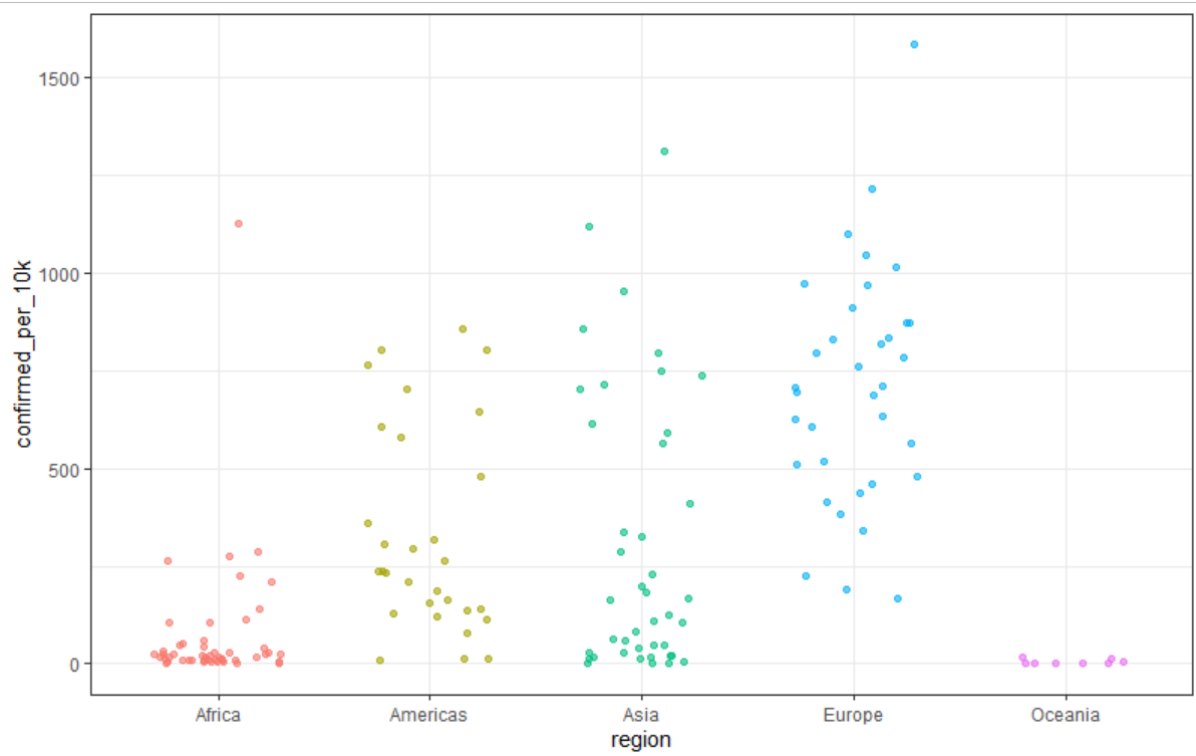
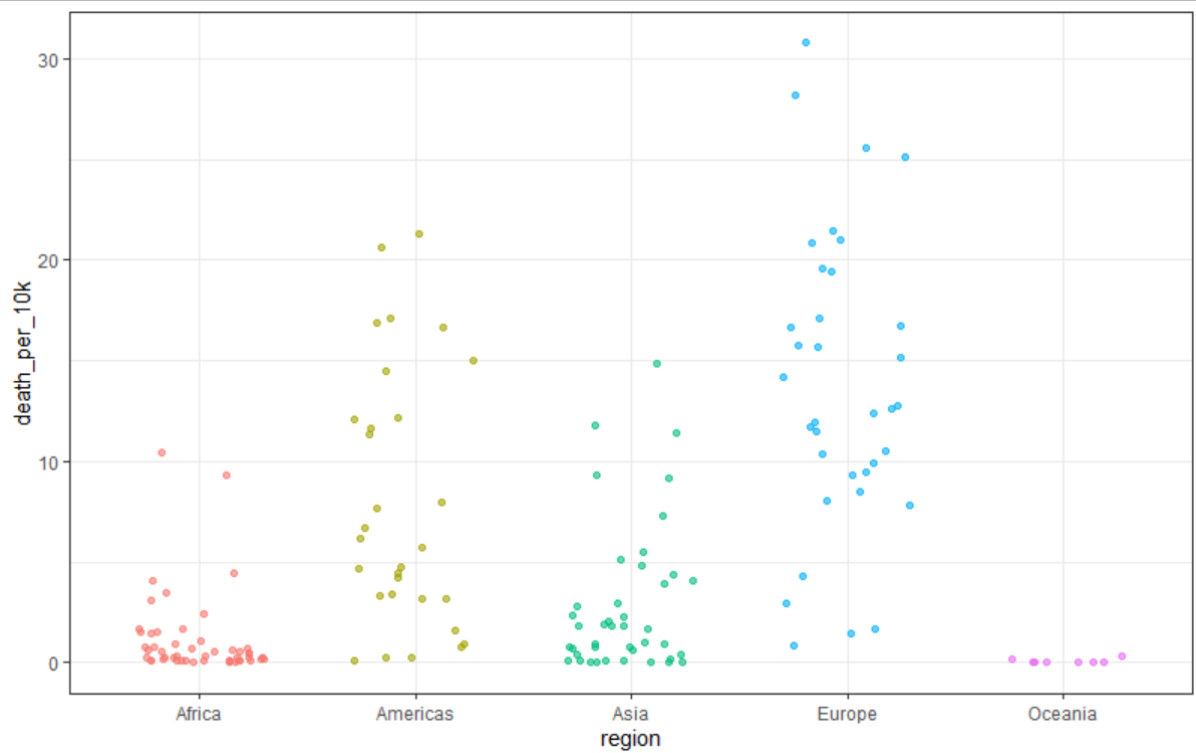
1-1 GDP와 기대수명(life expectancy) 사이의 상관관계를 나타내기 위한 그래프를 아래와 같이 그리라. 피어슨 상관계수를 계산하여 보고, 어떤 상관관계가 있는지 설명하여보라



1-2 GDP와 같이 monetary value에 관한 변수는 일반적으로 log-normal 분포를 갖는 것으로 알려져 있다. GDP에 log를 적용하여 기대수명(life expectancy)과의 상관관계를 그래프로 표현하여 보고, 상관계수를 통해 상관관계를 설명하여보라.



2-1 만명당 확진자 수, 사망자 수를 계산하여 대륙별(region)로 분포를 나타내는 시각화를 수행하고 비교하시오.



3-1 위 1,2 번 문항과 상관없이 국가별 COVID19 상황을 비교하기 위해서 (To understand the dataset better) exploratory visualization을 수행하시오. (그래프 2개 이상) 어떠한 사실을 발견할 수 있었는지 설명하시오. (open question)

3-2 위 3-1문항에서 발견한 사실을 효과적으로 설명하기 위해서 explanatory visualization을 수행하시오. (그래프 1개만)

Tips) 이 수업에서는 시각화를 전문적으로 배우는 것이 아니기 때문에 여러분들이 시각화에 필요한 R의 기능들에 대해서 익숙하지 않을 수 있습니다. 표현하고자 하는 그래프가 있는데 그 방법을 잘 모를 때, 교수님께 적극적으로 얘기를 하면 도와주겠습니다.