**<위도, 경도와 국가 지표 간의 상관관계 분석>**

PROJECT BY GO GO STUDY

팀장★ : 201411098 강도형

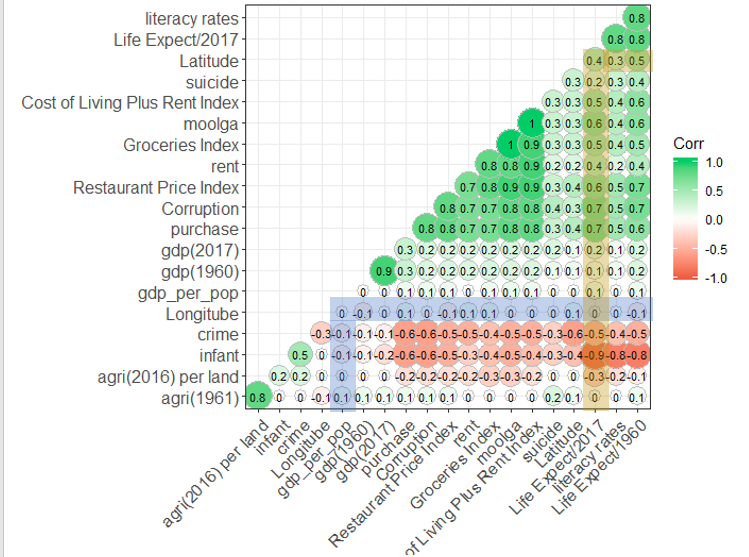
팀원 1 : 201411102 강재성

팀원 2 : 201411149 이준서

재레드 다이아몬드의 ‘총, 균, 쇠’를 읽어보면 문명이 발생하는데 필수적인 조건으로 기후와 온도가 등장한다. 기후와 온도는 위도와 아주 밀접한 관계가 있는데 이는 4대 문명의 발상지가 모두 비슷한 위도에 있다는 것으로 증명된다. 우리는 여기서 물음을 던진다. 그렇다면 국가가 가지는 다양한 지표들과 위도, 경도는 상관관계를 가지지 않을까?

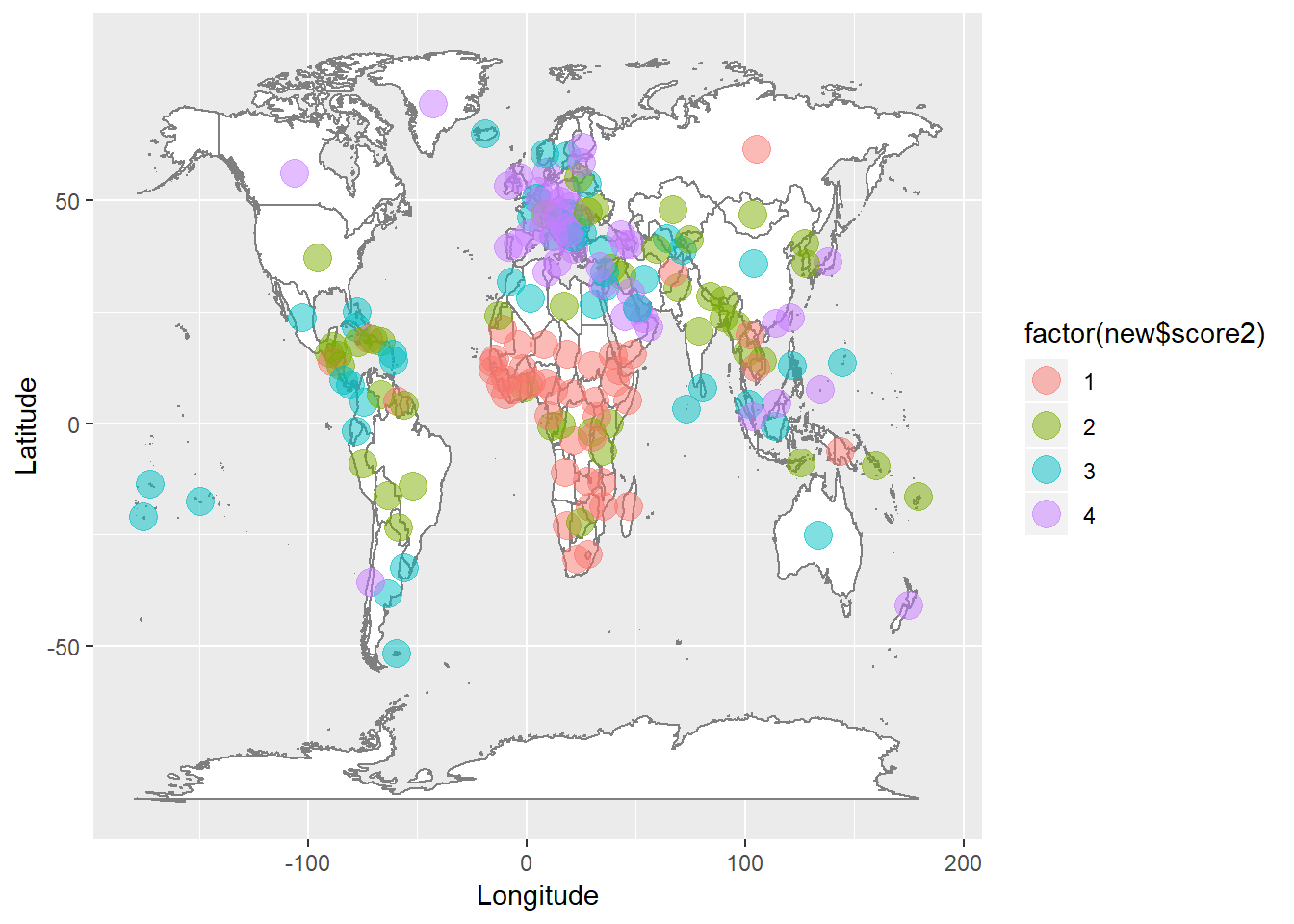
우선 국가의 위도와 경도로 데이터셋의 뼈대를 갖춘 뒤 다양한 지표들을 변수로 입력했다. 국가별 생활지수, 경제지수, 범죄율, 등을 조사하였다. 국가의 위도, 경도 좌표를 입력한 데이터셋에 인구수, 석유 매장량, 문해율, GDP(1960/2017), 인구 10만명당 자살인구, 부패인식지수, 인구 10만명 당 유아사망수, 인구 10만명 당 범죄 범죄 발생 수, 기대수명(1960/2017), 물가 지수, 임대료 지수, 식료품 가격 지수, 소비자 구매력 지수, 국토의 농경지 비율(1961/2016), 1인당 GDP를 나열했다. 결측값은 knnImputation을 통해 비슷한 위도와 경도를 가진 국가의 지표들의 평균값으로 대체했다.

먼저 위도, 경도가 다른 지표들과 선형관계를 가진다는 가정을 했다. shapiro.test를 통해 각 지표들이 정규분포를 따르는 지를 확인해보았지만 대다수의 변수가 정규분포를 따르지 않아 선형관계에 있다는 가정을 세울 수 없게 되어 회귀 모형에 적합하는 것은 포기했다.

이후 우리는 ggcorrplot을 통해 모든 변수 간의 상관관계를 파악했다.

그 결과 일부 경제지표만이 0.6정도의 약한 상관관계를 보이고 있었고 우리는 다시 유의미하고 여겨지는 변수 몇 가지를 골라 Scoring을 진행했다

Scoring에 사용된 변수는 ‘문해율’, ‘부패인식지수’, ‘기대수명’, ‘2017년 1인당 GDP’, ‘뉴욕 대비 식료품 가격 지수’, ‘인구 10만명당 자살 인구’, ‘범죄발생수’, ‘유아 사망수’이다. 긍정적인 지표는 +, 부정적인 지표는 – 하여 Score 변수에 더했다. 모든 지표는 표준화한 뒤 진행했다.

그 결과를 Fivenum을 기준으로 범주형 자료로 변경 한 뒤 세계지도에 mapping했다..

우리가 설계한 Score와 경도와는 상관관계를 찾을 수 없었고 위도와의 상관 계수는 약 0.56으로 약한 상관관계를 찾을 수 있었지만 크게 의미가 있다고 볼 수는 없다.

세계지도에 Score를 mapping 한 자료를 보면 국가지표들은 위도에 따른 상관관계보다는 대륙에 따른 상관관계가 더 큰 것으로 보인다. 이는 세계화된 사회에서 국가간 교류 혹은 단절로 인해 위도나 경도의 의미가 희미해진 것으로 추정된다. 따라서 세계화가 진행되기 이전의 데이터를 이용해서 분석을 한다면 더 명확한 상관관계가 나올 것으로 예측할 수 있다. 또한 우리가 Score로 설정한 변수들의 개수가 워낙 적었기 때문에 의미 있는 결과를 도출해내지 못한 것 같다. 따라서 위도, 경도와 국가지표는 큰 상관관계를 가지지 않는다는 결론에 도달했다. 국가 지표에 영향을 끼치는 요소는 위도나 경도보다는 종교나 문화, 혹은 경제정책 등으로 예상된다.