

빅데이터 기술을 이용한 ‘사회적 거리두기 4단계 지속 기간’과 ‘음원 스트리밍 횟수’의 관계 분석

2021320322 데이터과학과 윤민서

1. 연구 배경

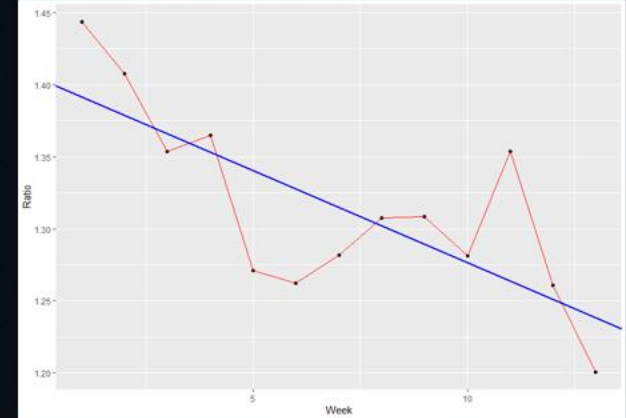
- 지난 7월부터 현재까지 수도권의 사회적 거리두기 단계가 4단계인 채로 3개월 동안 유지 중
- 코로나19로 인해 각종 시장이 침체기를 맞은 상황 속에서 음원 시장의 경우를 집중적으로 연구해 보기 위함
- 공통 동영상에서 빅데이터 기술로부터 얻을 수 있는 장점 또한 강조

2. 연구 방법

- 수도권 사회적 거리두기 4단계가 시행된 직후의 데이터로써 존재하는 7월 12일부터 가능한 최근의 데이터인 10월 10일까지 1주일 단위로 음원 스트리밍 횟수를 조사
- 많은 고정적인 팬들을 가진 가수가 음원을 출시했을 때 차트 최상위권의 음원 스트리밍 횟수가 많아지면서 전체적인 재생 횟수가 상승하는 경향이 있기 때문에 보정을 실시. 즉, 1~5등의 스트리밍 횟수의 변동성에 의해 하위 순위의 평균 스트리밍 횟수가 정해지는 경향이 큼
- 차트 31~40등의 스트리밍 횟수를 모두 더한 것을 1~5등의 횟수를 더한 것으로 나누는 것으로 산출
- 사회적 거리두기 4단계 지속 기간과 위에서 서술한 비율에 대한 상관관계와 선형 회귀 방정식을 도출
- 빅데이터 분석 언어로 R을 이용하였고 분석을 진행하기 위한 개발 환경으로 RStudio를 이용

3. 연구 결과

B	C	D	E
Date	Top 31-40	Top 1-5	D / C
2021-07-12 ~ 2021-07-18	10,325,427	14,907,542	1.443770025
2021-07-19 ~ 2021-07-25	10,394,057	14,634,200	1.407939171
2021-07-26 ~ 2021-08-01	10,130,625	13,713,098	1.353628034
2021-08-02 ~ 2021-08-08	9,943,891	13,574,647	1.365124276
2021-08-09 ~ 2021-08-15	10,764,787	13,681,218	1.270923243
2021-08-16 ~ 2021-08-22	10,342,152	13,057,267	1.262529017
2021-08-23 ~ 2021-08-29	10,100,570	12,947,580	1.281866271
2021-08-30 ~ 2021-09-05	10,073,948	13,170,919	1.307423763
2021-09-06 ~ 2021-09-12	9,985,339	13,064,163	1.308334449
2021-09-13 ~ 2021-09-19	9,983,177	12,789,381	1.281093283
2021-09-20 ~ 2021-09-26	8,706,272	11,784,413	1.353554426
2021-09-27 ~ 2021-10-03	9,235,496	11,646,539	1.261062644
2021-10-04 ~ 2021-10-10	9,543,829	11,458,633	1.200632681



4. 결론 및 제언

- 사회적 거리두기 4단계 지속 기간과 음원 스트리밍 횟수 간의 상관관계수는 -0.74 로 매우 강한 음의 상관관계를 보여주고 있고 파란 선으로 표현된 선형 회귀 방정식의 기울기 또한 -0.01 로 뚜렷하게 음수 값을 보여주고 있다.
- 수도권에서 사회적 거리두기 4단계가 지속됨에 따라 한국의 음원 시장도 침체되는 경향을 보이고 있다.
- 현재 바이러스의 변이로 인해 감염 시 위험성도 매우 낮아진 상태에서 대부분의 시장이 침체되는 것을 지켜보기만 해야 되는가?

5. 참고자료

<https://xn--o39an51b2re.com/> 음원 스트리밍 횟수

<https://journals.sagepub.com/doi/pdf/10.1177/875647939000600106>

피어슨 상관관계수의 해석