

[EDA이론]

기초통계량 + 시각화 둘 다 중요!

탐자분(탐색적자료분석)??

데이터로부터 유의미한 인사이트를 도출하는 프로젝트?

유의미한?

-신뢰도 높고

-정확도 높고

-목표가 있어야한다.

-임팩트가 큰가?

-실현가능한가?

-중요도가 낮지않은가? (의미없는 것을 깊게 팔지 말라) <- 재미 다리..?

문제에서 기반

당연하지 않는 걸

당연한걸 말로 표현해서 명시해주는 것 <- 공감을 불러일으키는 새로운 생각

- 0. 이상치 제거가 필요할 듯!
- 0. 표준화 정규화과정
- 0. 시각화
- 0. 사후처리
- 0. 데이터수집 (반복)...

EDA : 인사이트를 보면서 도출한다.

CDA : 가설을 세운 다음에 가설을 검정한다.

데이터에 초점을 하지 말고, 분석에 초점을 맞추자!

[주제선정]

0. 구글트렌드/네이버트렌드 vs 인기동영상 추이

늦은시간이지만 생각나서 말씀을 드리자면, 한가지 아이디어로써 유튜브 채널명을 구글트렌드랑 네이버트렌드에 검색해서 비교해볼 생각입니다. 유튜브 채널은 인기동영상에 대한 시계열 데이터이고 구글/네이버 트렌드는 검색량에 대한 시계열 데이터로 구해질 텐데 이 경향성이 비슷한지 알아보고자 합니다. 시간이 되면 텍스트마이닝 올려주신 걸로 유튜브 제목을 따오면 좋을 것 같은데 일단은 이전거라도 한번 해보겠습니다.

막상해보니 데이터를 얻는 것은 csv로 다운밖에 안되는 것 같다. 과연 이것들을 하는 의미를 찾아야하는데... 그냥 네이버랑 구글 검색창에 입력하고 뉴스나 동영상을 크롤링하는 것이 더 의미있을 순 있겠다.

유튜브 구독자수 추이와 조회수 추이가 인기동영상 추이와 동일한지 마찬가지로 네이버 구글 트렌드도 분석

0. 기타 크롤링 아이디어

그 영상을 클릭했을 때 옆에 뜨는 영상 중에 같은 채널에서 만든 영상의 갯수는?
그 영상들의 인기동영상 선정 날짜의 조회수 증가분 혹은 구독자수 증가분 혹은 댓글 수는?

유튜브 인기동영상 선정이 되면 떡상이 되는가..?

광고주의 경우에는 떡상하는 유튜버들에게 광고 외주를 맡기려고 한다. 마찬가지로 기존 default된 광고가 비슷하면 좋을 것이다. 광고 갯수가 적어도 좋을까..? 이걸 모르겠다..

0. Possion Process

홈화면이랑 구독자화면 인기동영상화면 외부링크화면 아니 노드들을 따로 추가 필요
탈출노드도 필요

스토케스틱 프로세스 포아송 프로세스 느낌

알 수 있는 것

오랜시간뒤에 이용자는 어떤 채널의 연상을 시청할 확률이 높은지..?

그럼 시청시간이 지수분포를 따르는가..?

주제

주제1 : 앞광고를 추진하는 광고회사들을 위한 맞춤 유튜버 추천

주제2 : 샌드박스의 유튜버 발굴 작업

떡상하는 유튜버

꾸준한 유튜버 / 팬덤이 있는 유튜버

기존 광고가 xx한 유튜버

카테고리가 맞는 유튜버

광고수나 시간 대비 조회수가 많이 나오는 유튜버

내가 가장 궁금했던 것은 어떤 영상을 봐야할까?(현명한 유튜브 사용법..?) 올바른 유튜브
플랫폼의 방향성 = 조회수 빨이가 아닌 양질의 동영상 비중이 증가해야한다..?

구독을 할 것인가? 말 것인가? / 알고리즘에 의한 효과 vs 구독에 의한 효과
어떤 키워드로 검색했을시에 구독자수와 조회수 이외에 어떤 정보를 더 추가해야할까
(유튜브 검색 옵션과 그 이외에 추가해볼 것은..?)

인기동영상은 대중들의 수요를 반영하는가 혹은 공급 중에 수요만을 반영하는가?

추천동영상을 내가 직접 맞춤으로 한다면 어떤 옵션들이 있어야할까?

가장 최신 댓글 날짜

주제1

유튜브내 광고를 진행하는 광고주 입장에서 카테고리별로 어떤 시점에 요일에 혹은 어떤 특징을 갖는 섹터에 두어야 효과가 극대화되는지?

- 효과의 극대화(조회수높음 / 그 기간에 그 카테고리가 많음)
- 어떤 특징(유튜브 광고 섹터옵션과 유사성을 비교)

주제2

유튜브내 광고를 진행하는 광고주 입장에서 카테고리별로 어떤 시점에 요일에 혹은 어떤 특징을 갖는 섹터에 두어야 효과가 극대화되는지?

- 효과의 극대화(조회수높음 / 그 기간에 그 카테고리가 많음)
- 어떤 특징(유튜브 광고 섹터옵션과 유사성을 비교)