

0. 주제 : 저희 스파게티 팀은 "코로나19에 따른 생활 양상 변화"를 분석해보았습니다.

(1) 목차

- 1. 코로나 확진자 변화
- 2. 변화에 따른 생활 양상 변화를 크게 3가지로 분석해보겠습니다.
- 3. 프로젝트 중 발생한 주요 문제, 이슈
- 4. 앞으로 개선 사항으로 말씀 드리겠습니다.

(2) 배경 설명 : 코로나 확진자 수가 증가하면 집에서 보내는 시간이 증가하기에  
홈트레이닝 시간, 온라인 쇼핑, 재택 근무 시간 등이 증가할 것으

로

예상.

하지만, 코로나가 길어짐에 따라 확진자수 증감에 사람들이 점차  
둔감해 질 것으로 예상.

(3) 목표 : 실제로 코로나 확진자 수 변화의 영향이 얼마나 되는 지 확인.

(다음 슬라이드)

1. 코로나 확진자 수의 증감을 보여주는 그래프입니다.

확진자 수의 증감으로는 판단하기에 부족하다 생각해서

전'달'에 비해 얼마나 증가 했는지가 사람들의 심리를 잘 보여준다 생각해

비율 그래프를 준비했습니다.

코로나 초창기인 '3월경'과 여름이 다가오는 '8월경'에 큰 증가율을 보임을 알 수  
있었습니다.

(8월에 큰 증가율은 그 당시 코로나가 높은 온도에 취약하다는 뉴스가 나오고,  
초기 방역에 성공적이라는

뉴스에 사회 전반적인 방심이 있었기 때문으로 추정합니다. 다만 근거가 되는  
뉴스 크롤링은 아직 시간

상 진행이 되지않아 확인이 안되는 상황입니다.)

(다음 슬라이드)

2. 분석 내용

2-1) 공공 자전거

자료가 방대하여 서울시로 한정하여 분석해보았습니다.

2-1-1) 서울시 공공 자전거 이용객 수를 보여주는 그래프입니다.

2020년 8월부터 급증하는 코로나 확진자 추세로 인해 거리두기 단계 격상이 시

행됨에 따라

자전거 이용수가 감소함을 보여줍니다. (매월 사회적 거리두기 단계 시행은 검색 중 입니다.

(다음 슬라이드)

2-1-2) 전체 이용객 수로는 여러번 사용하는 사람을 파악하기 어려워 한 사람의 평균 이용건수로 파악해보고자

사람당 평균 이용건수 그래프를 활용해보겠습니다.

그래프에서 보이듯이 8월부터 평균 이용건수도 감소함을 보여줍니다. 하루에 여러번 타는 헤비 유저의

비율이 많이 감소 한 것으로 보입니다.

(다음 슬라이드)

2-1-3) 데이터셋에서 보니 이용건수는 많은데 이용시간은 없거나 반대인 경우도 있어

이용건수보다 이용시간이 실제로 사용한 비율을 잘 보여줄 것으로 판단해

사람당 평균 이용시간, 이동거리로 판단해 보겠습니다.

이용건수가 감소함에 따라 이용시간도 감소했습니다.

(다음 슬라이드)

2-1-4) 다만 이동거리는 이용건수의 경향을 따르지 않아 영향을 미치는 다른 인자가

있는지 고민 중에 있습니다. 현재는 이동거리 데이터의 문제를 의심하고 있습니다.

(다음 슬라이드)

2-2) 지하철

역시 마찬가지로 서울시로 한정했습니다.

2-2-3) 코로나에 대한 사회의 인식과 통제 상황에 따라 대중교통 이용률을 보고자 합니다.

그 중에서 서울시 지하철로 분석하였습니다.

코로나 초창기에는 많은 이용객 수의 감소를 보입니다.

문제는 대중교통은 계절에도 많은 영향을 받기에 코로나의 영향만 보기에는 부족합니다.

(다음 슬라이드)

2-2-4) 코로나로 인한 영향을 보고자하는 취지에 맞지 않아, 다른 연도 같은 달의 이용객 수의 변화 비율로 판단하고자 합니다.

1~2월을 보시면 아직은 코로나에 대한 사회의 경각심이 있어서 전년도에 비해 많이 감소했음이 보입니다.

하지만 21년 3월부터는 전년3월에 비해 코로나 확진자 수가 증가했음에도 경각심이 약해져 승객이 많이

증가하는 추세를 보임을 알 수 있습니다.

(다음 슬라이드)

2-3) 온라인 상품군별 거래액(국가통계포털)

2-3-1) 코로나가 길어지면서 집에 있는 시간이 많아지면서 온라인 거래가 늘었을 것으로 생각해

온라인 쇼핑몰 상품군별 거래액으로 판단해보고자 합니다. 단위는 백만원입니다.

대표적으로 연관있어 보이는 5가지 상품군을 선별해 그래프로 만들었습니다.

먼저 코로나 확진자 수가 증가한 20년 3월에 여행 및 교통서비스 그래프가

거의 1조에 가까운 거래액 감소를 보여 코로나의 영향이 큼을 보여줍니다.

그에 비해 컴퓨터 및 주변 기기 거래액은 온라인 수업, 재택근무로 늘었을 것으로 예상했으나

많은 증가를 보이진 않아 원인을 고민하고 있습니다.

스포츠 및 레저용품에 대한 거래액 또한 큰 변화가 없어 원인을 고민하고 있습니다.

큰 변화가 없는 컴퓨터 상품군은 카테고리가 커서 노트북 같은 특정 상품의 변화는 잘 반영이 되지

않았을 꺼라 예상합니다.

스포츠 및 레저용품은 홈트레이닝 관련 상품의 거래액은 늘었지만 야외활동이 어려워지는 시기로 인해

관련 상품 거래액은 감소해 서로 상쇄되어 거래액이 큰 변화를 보여주지 못했다 생각됩니다.

추후 관련 내용을 찾아 원인을 파악하려 합니다.

그외 2가지 음식료품과 가전, 전자, 통신기기는 거래액이 상승하는 경향을 보여 예상과 부합하는

것으로 보입니다.

+ 분석 결과 : 코로나 초기에는 사회의 경각심이 높아 사회활동이 감소했음을 보여줍니다.

시간이 경과함에 따라 경각심이 낮아져 점차 사회활동이 초기에 비해 감소되는 정도가 약해지는

것으로 보입니다.

### 3. 진행 과정

#### (1) 주요 이슈

+ 코로나 확진자 수(국가 통계포털, 코로나 라이브)

- 확진자 수 일별로 존재, 데이터 상 숫자가 누적값이라 처리 필요.

- 해결) 해당 월 마지막 일을 잡아 차로 매달 확진자수 계산해 데이터

셋

을 만들어 해결.

- 만들어진 데이터셋 값을 코로나 라이브의 월별 확진자 수와 대조 비

교 확인.

+ 공공 자전거(공공데이터)

- 데이터셋에 결측값과 //N으로 표현된 값이 존재해 처리 에러.

- 해결) map을 이용해 특정 문자만 처리시 지정되지 않은 자료는 전부 NAN값으로 변하는 문제 발생. replace를 이용해 해결.

+ 지하철 승객수(공공데이터)

- 데이터가 일별로 정리되어 있어 월별로 정리가 필요.

- 해결) 날짜를 조건으로 지정하여 새로운 데이터셋을 만들어 해결.
- + 온라인 거래액(국가 통계포털)
  - 그래프 y값이 계속 증가하는 그래프만 그려지는 문제.
  - 해결) y축 값을 확인하니 오름차순이 아님을 발견. 문자형으로 인식  
된다 판단. info, dtype을 이용해 확인한 뒤 자료형을 변환해 해

결.

(다음 슬라이드)

#### 4. 앞으로 개선사항

- (1) 변화를 파악하기 쉽도록 코로나 확진자 그래프와 각 항목 그래프를 한 그래프로 완성.
- (2) 각 파트별 더 많은 분석 요인 추가.
- (3) 코로나와 매출 품목별 변화 요인을 분석할 수 있는 데이터 추가.
- (4) 그래프 변화 시점 주요 뉴스 제목을 시각화

(다음 슬라이드)

#### 5. 질문

(끝)