빅데이터 분석 결과 보고서

2021년 00월 00일.

과목 : 빅데이터

담당 : 허준 교수

학번 : 201744015

이름 : 김현우

목차

1. **분석 개요**
   1. 분석 대상
   2. 분석 목적
   3. 활용 데이터
   4. 분석 수행절차
   5. 분석 결과 요약
2. **분석 절차**
   1. 활용 데이터 수집 대상 및 방법
   2. 활용 데이터 가공 방법
   3. 활용 데이터 분석 도구 및 방법
3. **분석 방법**
   1. 분석 도구
   2. 분석 방법
   3. 분석 알고리즘
   4. 분석 결과 시각화 및 설명
4. **시사점, 한계점 및 소감**
   1. 시사점
   2. 한계점
   3. 본 프로젝트 수행 후 소감

**1. 분석 개요**

1.1 분석 대상

1.2 분석 목적

1) 지하철 역사 내 점포의 면적 별 임대료를 지하철 이용객 수와 대응하여 가격 메리트를 확인하고자 함

2) 지하철 이용객 수 비교는 환승 인원 수 비교, 지하철 승차 및 하차 총 승객 수를 이용하여 파악하고자 함

3) 지하철 승차 및 하차 총 승객 수는 평일과 주말의 편차 및 월별 편차가 있기 때문에 일년간 데이터의 전체 평균값을 비교한다.

1.3 활용 데이터

1) 서울교통공사 지하철 상가 현황

2021/08/31일자 데이터를 활용했다. 서울교통공사의 상가임대 현황(면적,업종,계약일,임대료 등) 데이터입니다. 해당 데이터는 상가 유형, 호선, 역사명, 상가번호, 면적(㎡), 영업업종, 계약시작일자, 계약종료일자, 월임대료 정보를 포함하고 있습니다.

출처 : [https://www.data.go.kr/data/15071329/fileData.do]

|  |  |  |  |
| --- | --- | --- | --- |
| 순번 | 항목명 | 항목 설명 | 형식 |
| 1 | 상가유형 | \*상가유형 | string |
| 2 | 호선 | 호선 | numeric |
| 3 | 역사명 | 역사명 | string |
| 4 | 상가번호 | 상가번호 | numeric |
| 5 | 면적(제곱미터) | 면적(제곱미터) | numeric |
| 6 | 영업업종 | 영업업종 | string |
| 7 | 계약시작일자 | 계약시작일자 | date |
| 8 | 계약종료일자 | 계약종료일자 | date |
| 9 | 월임대료 | 월임대료 | numeric |

\* 상가유형

- 67일괄: 6,7호선 상업공간 일괄 임대차 사업 추진 상가, 현재 공실

- 네트워크: 다수역에 다수점표를 일괄 임대차하여 운영하는 상가

- 소송상가: 계약이 종료되었으나 명도거부하여 명도소송 중인 상가

- GS: GS리테일이 공사와 계약하여 다수역에 다수 점포를 조성하여 운영한 상가

2) 서울시 지하철 호선별 역별 시간대별 승하차 인원 정보

교통카드(선후불교통카드 및 1회용 교통카드)를 이용한 지하철 호선별 역별(1~9호선, 서울시 관할 운송기관에 한함) 시간대별 승하차인원을 나타내는 정보입니다.

출처 : [https://data.seoul.go.kr/dataList/OA-12914/S/1/datasetView.do]

|  |  |  |  |
| --- | --- | --- | --- |
| 순번 | 항목명 | 항목 설명 | 형식 |
| 1 | 사용일자 | 승객 수 조사 일자 | date |
| 2 | 노선명 | 1~9호선, 분당 등 모든 노선 | string |
| 3 | 역명 | 역명 | string |
| 4 | 승차총승객수 | 승차총승객수 | numeric |
| 5 | 하차총승객수 | 하차총승객수 | numeric |
| 6 | 등록일자 | 등록일자 | date |

3) 서울교통공사\_환승역\_환승인원정보

역별 요일별 (평일/토요일/일요일) 관할 환승역 환승인원을 나타내는 정보(2019년)입니다.

출처 : [https://www.data.go.kr/data/15062858/fileData.do]

|  |  |  |
| --- | --- | --- |
| 항목명 | 항목 설명 | 형식 |
| 역명 | 환승 가능한 역 이름 | string |
| 평일(일평균) | 일평균 환승 인원 정보 | numeric |
| 토요일 | 토요일 환승 인원 정보 | numeric |
| 일요일 | 일요일 환승 인원 정보 | numeric |

# 카카오 지도 api 요약 : https://apis.map.kakao.com/web/guide/#whatlibrary

목록 나열

1.4 분석 결과 요약

노선을 1~9호선에 한정할 지, 경의선 분당선 등 다양한 노선을 추가할지 고민했지만, 1~9호선부터 시작하기로 했다.

**2. 분석 절차**

* 1. 활용 데이터 수집 대상 및 방법

1) 서울교통공사 지하철 상가 현황

출처 : [https://www.data.go.kr/data/15071329/fileData.do]

2) 서울시 지하철 호선별 역별 시간대별 승하차 인원 정보

출처 : [https://data.seoul.go.kr/dataList/OA-12914/S/1/datasetView.do]

3) 서울교통공사\_환승역\_환승인원정보

출처 : [https://www.data.go.kr/data/15062858/fileData.do]

2.2 활용 데이터 가공 방법

* 1. 활용 데이터 분석 도구 및 방법

**3. 분석 방법**

* 1. 분석 도구
  2. 분석 방법
  3. 분석 알고리즘

3.4 분석 결과 시각화 및 설명

1. **시사점, 한계점 및 소감**
   1. 시사점
   2. 한계점
   3. 본 프로젝트 수행 후 소감

본인은 2021년 2학기 빅데이터 과목의 빅데이터 분석 결과보고서를 제출합니다.