

삼성 청년 SW 아카데미

데이터분석 이론 및 실습

〈알림〉

본 강의는 삼성 청년 SW아카데미의 콘텐츠로
보안서약서에 의거하여
강의 내용을 어떠한 사유로도 임의로 복사, 촬영,
녹음, 복제, 보관, 전송하거나
허가 받지 않은 저장매체를
이용한 보관, 제3자에게 누설, 공개,
또는 사용하는 등의 행위를 금합니다.

9월25일 실습 및 과제 알고리즘

실습 알고리즘

• 데이터 활용신청

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

공공데이터



교통

활용사례(갤러리) 등록

URL 복사

목록 이동

서울시 지하철호선별 역별 승하차 인원 정보

교통카드(선후불교통카드 및 1회용 교통카드)를 이용한 지하철호선별 역별(서울교통공사, 한국철도공사, 공항철도, 9호선) 승하차인원을 나타내는 정보입니다. (일단위)

※ Sheet 서비스는 마지막 한달치 데이터만 서비스 합니다. (* 데이터 적재는 매일 3일전 데이터를 갱신합니다.)

파일내려받기

* 파일에 이상이 있는 경우 '오류신고'를 통해 운영자에게 알려주세요.

오류신고

| NO | 항목 | 파일명 | 용량 (MB) | 수정일 | 내려받기 |
|----|-----|------------------------------|---------|-------------|---|
| 1 | 명세서 | CARD_SUBWAY_MONTH_202408.csv | 1.16 | 2024.09.10. |  |
| 2 | 명세서 | CARD_SUBWAY_MONTH_202407.csv | 1.2 | 2024.08.10. |  |
| 3 | 명세서 | CARD_SUBWAY_MONTH_202406.csv | 1.04 | 2024.07.10. |  |
| 4 | 명세서 | CARD_SUBWAY_MONTH_202405.csv | 1.2 | 2024.06.10. |  |
| 5 | 명세서 | CARD_SUBWAY_MONTH_202404.csv | 1.19 | 2024.05.10. |  |

[전체 파일보기](#)

• 데이터 활용신청

미리보기

닫힘 -

Sheet

Open API

▶ 샘플 URL

Open API 이용안내

인증키 신청

서식 다운로드

샘플 URL

2015년11월1일 지하철호선별 역별 승하차 인원
http://openapi.seoul.go.kr:8088/(인증키)/xml/CardSubwayStatsNew/1/5/20151101

예제

<?xml version="1.0" encoding="UTF-8"?>
<CardSubwayStatsNew>
<list_total_count>548</list_total_count>
<RESULT>
<CODE>INFO-000</CODE>
<MESSAGE>정상 처리되었습니다</MESSAGE>
</RESULT>
<row>
<USE_YMD>20151101</USE_YMD>
<SBWY_ROUT_LN_NM>1호선</SBWY_ROUT_LN_NM>
<SBWY_STNS_NM>서울역</SBWY_STNS_NM>
<GTON_TNOPE>53097</GTON_TNOPE>
<GTOFF_TNOPE>45172</GTOFF_TNOPE>

▶ 요청인자

| 변수명 | 타입 | 변수설명 | 값설명 |
|-------------|------------|--------|--|
| KEY | String(필수) | 인증키 | OpenAPI에서 발급된 인증키 |
| TYPE | String(필수) | 요청파일타입 | xml:xml,xml파일:xmlif,엑셀파일:xls,json파일:json |
| SERVICE | String(필수) | 서비스명 | CardSubwayStatsNew |
| START INDEX | INTGFR(필수) | 요청시작위치 | 적수 인력 (페이지 시작번호 인니다: 데이터 행 시작번호) |

인증키 신청

• 실행결과

| | USE_YMD | SBWY_ROUT_LN_NM | SBWY_STNS_NM | GTON_TNOPE | GTOFF_TNOPE | REG_YMD | |
|---|----------|-----------------|--------------|------------|-------------|---------|----------|
| 0 | 20230901 | | 1호선 | 서울역 | 61380.0 | 58750.0 | 20230904 |
| 1 | 20230901 | | 1호선 | 시청 | 28527.0 | 28773.0 | 20230904 |
| 2 | 20230901 | | 1호선 | 종각 | 43893.0 | 43866.0 | 20230904 |
| 3 | 20230901 | | 1호선 | 종로3가 | 29355.0 | 26756.0 | 20230904 |
| 4 | 20230901 | | 1호선 | 종로5가 | 27878.0 | 27493.0 | 20230904 |

Process finished with exit code 0

- 실행결과

```
데이터 수집 완료  
중복 및 이상치 제거 완료
```

```
Process finished with exit code 0
```


• 실행결과

```
데이터 수집 완료
```

```
결측치 개수:
```

```
USE_YMD      0
```

```
호선명        0
```

```
지하철역      0
```

```
승차총승객수  1
```

```
하차총승객수  1
```

```
날짜          0
```

```
dtype: int64
```

```
결측치 제거 후 결측치 개수:
```

```
USE_YMD      0
```

```
호선명        0
```

```
지하철역      0
```

```
승차총승객수  0
```

```
하차총승객수  0
```

```
날짜          0
```

```
dtype: int64
```

```
Process finished with exit code 0
```

• 실행결과

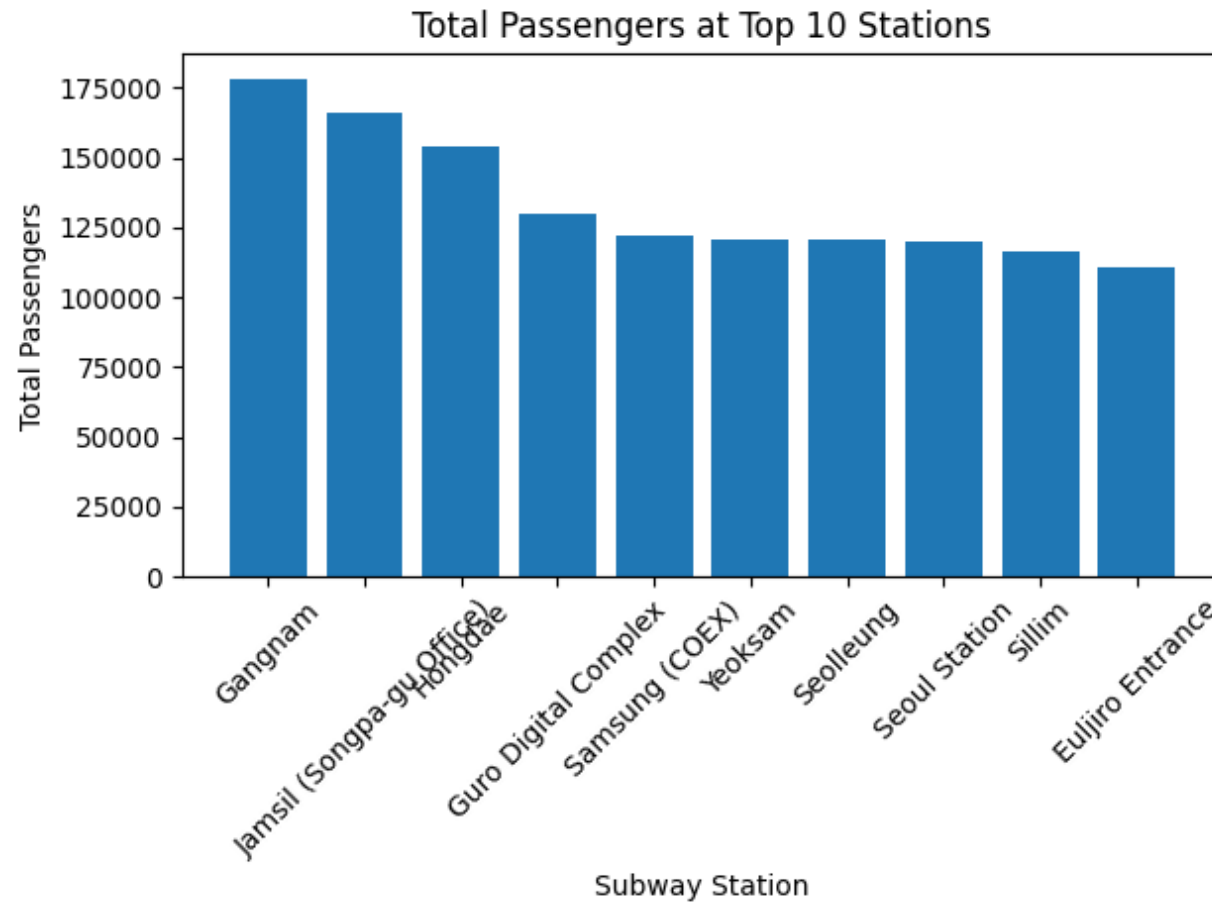
데이터 수집 완료

새로운 피쳐 '호선_역명' 생성 완료

| | USE_YMD | 호선명 | 지하철역 | GTON_TNOPE | GTOFF_TNOPE | REG_YMD | 호선_역명 |
|---|----------|-----|------|------------|-------------|----------|----------|
| 0 | 20230901 | 1호선 | 서울역 | 61380.0 | 58750.0 | 20230904 | 1호선_서울역 |
| 1 | 20230901 | 1호선 | 시청 | 28527.0 | 28773.0 | 20230904 | 1호선_시청 |
| 2 | 20230901 | 1호선 | 종각 | 43893.0 | 43866.0 | 20230904 | 1호선_종각 |
| 3 | 20230901 | 1호선 | 종로3가 | 29355.0 | 26756.0 | 20230904 | 1호선_종로3가 |
| 4 | 20230901 | 1호선 | 종로5가 | 27878.0 | 27493.0 | 20230904 | 1호선_종로5가 |

Process finished with exit code 0

• 실행결과



과제 알고리즘

• 실행결과

```
stationName      암호화된_정류소_이름
0   Seoul Station  Vhrxo Vwdwlrq
1   Gangnam Station Jdqjqdp Vwdwlrq
2       Jongno3ga   Mrqjqr3jd

Process finished with exit code 0|
```

• 데이터 활용 신청

출처:
<https://data.seoul.go.kr/dataList/OA-13314/A/1/datasetView.do>

공공데이터


교통

활용사례(갤러리) 등록

URL 복사

목록 이동

서울시 교통량 지점 정보

교통량 API를 사용하기 위한 지점정보 입니다.

데이터 정보

| | | | |
|-------|---|--------|---------------------|
| 공개일자 | 2016.08.17. | 최신수정일자 | 2024.09.21. |
| 갱신주기 | 비정기(자료변경시) | 분류 | 교통 |
| 원본시스템 | 서울시 교통정보 시스템 | 저작권자 | 서울특별시 |
| 제공기관 | 서울특별시 | 제공부서 | 도시교통실 교통기획관 미래첨단교통과 |
| 담당자 | 안재성 (02-2133-4980) | | |
| 원본형태 | DB | 제3저작권자 | 없음 |
| 라이선스 |  공공누리 1유형:출처표시(상업적 이용 및 변경 가능) | | |
| 관련태그 | 통제, 교통상황, 소통, 돌발 | | |

• 실행결과

수집한 데이터:

| | SPOT_NM | GRS80TM_X | GRS80TM_Y |
|---|----------------|---------------|---------------|
| 0 | 자하문로(자하문터널) | 197216.855046 | 454350.990432 |
| 1 | 율곡로(안국역) | 198645.671347 | 452937.216603 |
| 2 | 창경궁로(서울여자대학교) | 199825.89671 | 453668.322568 |
| 3 | 대학로(한국방송통신대학교) | 200185.014636 | 453144.745427 |
| 4 | 종로(동묘앞역) | 201515.783975 | 452624.900832 |

정제 및 피처 엔지니어링 적용 후 데이터:

| | SPOT_NM | GRS80TM_X | GRS80TM_Y | 지역구 |
|---|----------------|---------------|---------------|-----|
| 0 | 자하문로(자하문터널) | 197216.855046 | 454350.990432 | 서북권 |
| 1 | 율곡로(안국역) | 198645.671347 | 452937.216603 | 서북권 |
| 2 | 창경궁로(서울여자대학교) | 199825.89671 | 453668.322568 | 서북권 |
| 3 | 대학로(한국방송통신대학교) | 200185.014636 | 453144.745427 | 동북권 |
| 4 | 종로(동묘앞역) | 201515.783975 | 452624.900832 | 동북권 |

Process finished with exit code 0

내일 방송에서 만나요!

삼성 청년 SW 아카데미