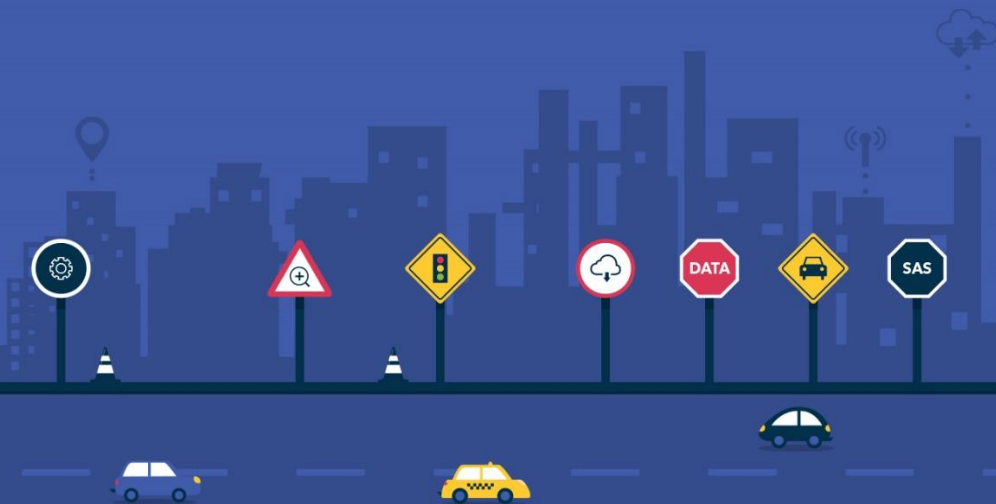




행정안전부와 함께하는

제16회 SAS 분석 챔피언십

분석 데이터 설명



주최 한국싸스소프트웨어(유) | 후원 행정안전부 | 자문 한국데이터마케팅학회


THE POWER TO KNOW.

분석 데이터

01



분석 참고 사항

02



분석 데이터 리스트

	카테고리	데이터 리스트	참고 사항
1	교통사고 데이터	16, 17년	전체 데이터: 서울시에 한정 사고 위치 (위도, 경도): 3개 구에 한정(강남, 영등포, 중구)
2	디지털 운행 기록(DTG)	서울시 택시 - 16, 17년 6-8월	초 단위 운행 기록 데이터 전체 데이터: 3개 구에 한정(강남, 영등포, 중구)
3	교통 시설물 데이터(11개)	도로 교차 점 위치정보	반드시 제공된 데이터로 사용하여야 함
		교량 위치정보	
		버스 정류장 위치정보	
		안전센터관할 위치정보	
		어린이 보호구역 위치정보	
		운행제한 시설물 위치정보	
		자동차 전용도로 위치정보	
		주거지역 위치정보	
		CCTV 위치정보	
		횡단보도 위치정보	반드시 제공된 데이터로 사용하여야 함
		힘프 위치정보	반드시 제공된 데이터로 사용하여야 함
4	기타 데이터 (5개)	도로 별 통행 속도	16, 17년 / [별첨] 링크 위치 정보
		도로 별 교통량	16, 17년 / [별첨] 교통량 지점정보
		시간 별 기상정보	16, 17년
		인구 수, 차량 수 정보	요약된 데이터임
		행정구역 위치정보	시군구 위치데이터, 읍면동 위치데이터

01 분석 데이터

1. 교통사고 데이터

index	column name	type	참고사항	sample1	sample2
1	사고번호	numeric	사고 구분 번호	2016010100100001	2016010100100002
2	발생일시	character	연월일시	2016년 1월 1일 00시	2016년 1월 1일 00시
3	발생요일	character		금요일	금요일
4	발생시군구	character		중구 봉래동2가	동대문구 답십리동
5	사고내용	character		경상사고	경상사고
6	사망자수	numeric		0	0
7	중상자수	numeric		0	0
8	경상자수	numeric		1	1
9	부상신고자수	numeric		0	0
10	사고유형	character		차대차 - 측면직각충돌	차대차 - 측면직각충돌
11	법규위반	character		신호위반	직진우회전진행방해
12	노면상태	character		포장 - 건조	포장 - 건조
13	기상상태	character		맑음	흐림
14	도로형태	character		교차로 - 교차로안	교차로 - 교차로부근
15	가해운전자차종	character		승용	승용
16	가해운전자성별	character		남	남
17	가해운전자연령	character		21세	55세
18	가해운전자상해정도	character		상해없음	상해없음
19	피해운전자차종	character		승용	승용
20	피해운전자성별	character		남	남
21	피해운전자연령	character		57세	53세
22	피해운전자상해정도	character		상해없음	경상
23	위도	numeric	강남구, 영등포구, 중구 데이터만 존재	37.55702	.
24	경도	numeric		126.9724	.

2. 디지털 운행 기록(DTG)

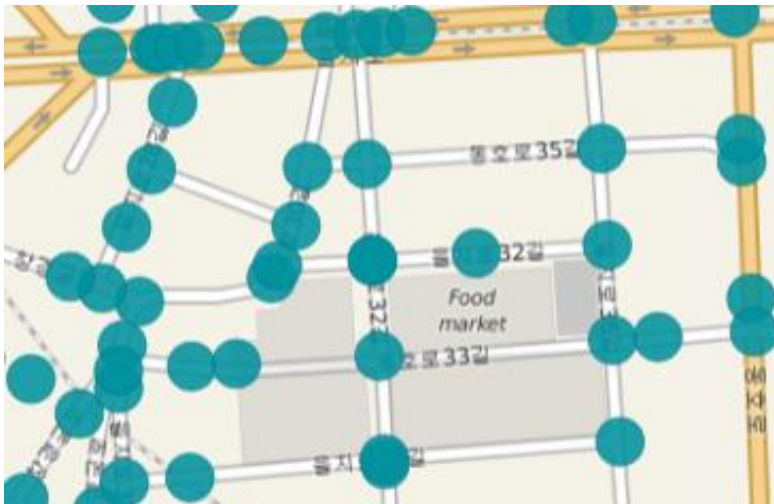
index	column name	type	데이터 정의	참고 사항	sample
1	trip_key	character	운행 기록 key 값	'C 차량 등록 번호 운행일자' 로 구성	C-58586289116061300000000
2	car_no	character	차대 번호	암호화된 코드	XXXXXXXXXP9A702880
3	car_drno	character	차량 등록 번호	암호화된 코드	-585862891
4	saup_no	character	사업자 등록 번호	암호화된 코드	XXXXXX02886
5	driver_code	character	운전자 코드	암호화된 코드	0
6	daily_dis	numeric	일일 운행 거리	하루 동안의 누적 운행 거리	2
7	total_dis	numeric	누적 운행 거리	현재까지의 총 누적 운행거리	504881
8	speed	numeric	차량 속도	Km/h	0
9	rpm	numeric	엔진 회전 수		700
10	break	binary	브레이크 작동 여부	브레이크 작동 = 1, 미작동 = 0	1
11	spot_x	numeric	차량 위치 X 좌표 값	100,000으로 나누어 경도로 사용	126929256
12	spot_y	numeric	차량 위치 Y 좌표 값	100,000으로 나누어 위도로 사용	37485020
13	gis	numeric	GIS 방위각	북쪽 0=360 기준, 차량 진행 방향을 나타내는 방위각	57
14	Vx	numeric	가속도 Vx	X축 기준 가속도	7.7
15	Vy	numeric	가속도 Vy	Y축 기준 가속도	-0.7
16	company_code	character	운송회사 코드	암호화된 코드	XX822
17	time	numeric	시간	YYMMDDHHmmss00 의 형태	16061300051700
18	sig_cd	numeric	시군구 코드	강남구, 영등포구, 중구	11560

3. 교통 시설물 데이터 (1) 도로 교차 점 위치정보

index	column name	type	데이터 정의	참고 사항	sample
1	SIG_CD	numeric	도로1 시군구 코드		11110
2	RDS_MAN_NO	numeric	도로1 도로구간일련번호		2933
3	RN	character	도로1 도로명		통일로14길
4	RN_CD	numeric	도로1 도로명코드		4100479
5	ENG_RN	character	도로1 영문도로명		Tongil-ro 14-gil
6	WDR_RD_CD	numeric	도로1 광역도로구분코드	행안부=1, 시도=2, 시군구=3	3
7	ROA_CLS_SE	numeric	도로1 도로위계기능구분	고속도로=1, 대로=2, 로=3, 길=4	4
8	ROAD_BT	numeric	도로1 도로폭	미터 단위 기준	8
9	ROAD_LT	numeric	도로1 도로길이	미터 단위 기준	217
10	BSI_INT	numeric	도로1 기초간격	미터 단위 기준	10
11	SIG_CD_2	numeric	도로2 시군구 코드		11410
12	RDS_MAN_NO_2	numeric	도로2 도로구간일련번호		2993
13	RN_2	character	도로2 도로명		통일로
14	RN_CD_2	numeric	도로2 도로명코드		3000008
15	ENG_RN_2	character	도로2 영문도로명		Tongil-ro
16	WDR_RD_CD_2	numeric	도로2 광역도로구분코드	행안부=1, 시도=2, 시군구=3	1
17	ROA_CLS_SE_2	numeric	도로2 도로위계기능구분	고속도로=1, 대로=2, 로=3, 길=4	3
18	ROAD_BT_2	numeric	도로2 도로폭	미터 단위 기준	42
19	ROAD_LT_2	numeric	도로2 도로길이	미터 단위 기준	4880
20	BSI_INT_2	numeric	도로2 기초간격	미터 단위 기준	20
21	lat	numeric	위도		37.57432
22	lon	numeric	경도		126.9581

3. 교통 시설물 데이터 (1) 도로 교차 점 위치정보

- 각 도로가 교차하는 교차 점의 위, 경도 값
- 도로1의 특징, 도로2의 특징 모두 활용
- 중복되는 위, 경도 존재



3. 교통 시설물 데이터 (2) ~ (5)

(2) 교량 위치정보

index	column name	type	sample
1	시군구코드	numeric	11260
2	교량일련번호	numeric	4
3	교량명_한글	character	구목동교
4	위도	numeric	37.618097
5	경도	numeric	127.078856

(3) 버스 정류장 위치정보

index	column name	type	sample
1	정류소ID	numeric	100000001
2	정류소번호	numeric	1001
3	정류소명	character	종로2가사거리
4	시군구 코드	numeric	11110
5	시군구_영어명	character	Jongno-gu
6	시군구_한글명	character	종로구
7	위도	numeric	37.56976
8	경도	numeric	126.98775

(4) 안전센터관할 위치 정보

index	column name	type	sample
1	고유번호	numeric	1
2	센터ID	numeric	1121102
3	센터명	character	신정119안전센터
4	위도	numeric	37.520698
5	경도	numeric	126.851035

(5) 어린이 보호구역 위치 정보

index	column name	type	sample
1	대상시설명	charcater	새롬학교
2	소재지도로명주소	charcater	관악구 보라매로 44
3	소재지번주소	charcater	관악구 보라매동 722-19
4	관리기관명	charcater	서울특별시 관악구청
5	관할경찰서명	charcater	서울관악경찰서
6	CCTV설치여부	charcater	Y
7	CCTV설치대수	numeric	1
8	보호구역도로폭	numeric	20 (미터 단위 기준)
9	위도	numeric	37.493126
10	경도	numeric	126.927367

3. 교통 시설물 데이터 (6) ~ (8)

(6) 운행제한 시설물 위치정보

index	column name	type	참고	sample
1	관리번호	numeric		86
2	구분	character		육교
3	관리기관	character		구로
4	시설명	character		이씨의다리
5	위치	character	구로 구분	구로구
6	폭제한	numeric	미터 단위 기준	
7	높이제한	numeric	미터 단위 기준	4
8	길이제한	numeric	미터 단위 기준	
9	실제높이	numeric	미터 단위 기준	4.5
10	출발 위도	numeric	시설물 시작 위도	37.511526
11	출발 경도	numeric	시설물 시작 경도	126.878215
12	도착 위도	numeric	시설물 도착 위도	
13	도착 경도	numeric	시설물 도착 경도	

(7) 자동차 전용도로 위치정보

index	column name	type	sample
1	순번	numeric	30
2	시군구 코드	numeric	11740
3	도로 번호	numeric	21
4	전용도로_한글명	character	판교-구리고속국도
5	위도	numeric	37.546229
6	경도	numeric	127.180980

(8) 주거지역 위치정보

index	column name	type	sample
1	고유번호	numeric	68
2	도시계획사항 코드	character	UQA122
3	도시계획사항 명칭	character	제2종일반주거지역
4	시군구 코드	numeric	11110
5	시군구_영어명	character	Jongno-gu
6	시군구_한글명	character	종로구
7	위도	numeric	37.57794
8	경도	numeric	127.01595

3. 교통 시설물 데이터 (9) ~ (11)

(9) CCTV 위치정보

index	column name	type	sample
1	관리기관명	character	관악구청
2	소재지도로명주소	character	관악구 봉천로 243
3	소재지지번주소	character	관악구 보라매동 971-16
4	설치목적구분	character	다목적
5	카메라대수	numeric	1
6	설치년월	character	2013-01
7	위도	numeric	37.48934417
8	경도	numeric	126.9287941

(10) 횡단보도 위치정보

index	column name	type	데이터 정의	sample
1	MGRNU	character	횡단보도 관리번호	06-0000000077
2	HOL	numeric	가로길이(m)	5
3	VEL	numeric	세로길이(m)	8
4	AW_SN_QUA	numeric	화살표시수량	1
5	N_LEN_X_CDE	numeric	화살표시길이	1
6	GU_CDE	numeric	구코드	380
7	Lat	numeric	위도	37.58395
8	Lon	numeric	경도	126.8915

(11) 험프 위치정보 - 과속 방지턱

index	column name	type	데이터 정의	sample
1	MGRNU	character	험프 관리 번호	16-0000010563
2	GU_CDE	numeric	구 코드	680
3	lat	numeric	위도	37.50148
4	lon	numeric	경도	127.0534

4. 기타 데이터 (1) 도로 별 통행 속도

Index	column name	Type	Sample
1	링크 아이디	numeric	1000002800
2	링크 세부 구분	numeric	1
3	시군구 코드	character	11140
4	시군구_영어명	character	Jung-gu
5	시군구_한글명	character	중구
6	위도	numeric	37.56793
7	경도	numeric	126.98768

- 링크 아이디를 key 값으로 join 가능
- 상세한 데이터 구축 과정은 [서울시 교통정보 시스템](#) 웹사이트 '속도 정보' 메뉴 참고

[별첨] 링크 위치 정보

index	column name	type	Sample
1	도로명	character	4.19로
2	링크 아이디	numeric	1080012200
3	시점명	character	아카데미하우스
4	종점명	character	국립4.19묘지
5	방향	character	상행
6	거리	numeric	1,179
7	차선수	numeric	1
8	01시~24시	numeric	48

4. 기타 데이터 (2) 도로 별 교통량

index	column name	type	Sample
1	일자	numeric	20161101
2	요일	character	화
3	지점명	character	성산로(금화터널)
4	지점번호	character	A-01
5	유입유출	character	유입
6	방향	character	봉원고가차도->독립문역
7	0시~23시	numeric	740

- 지점번호 및 링크 아이디를 key 값으로 join 가능
- 상세한 데이터 구축 과정은 [서울시 교통정보 시스템](#) 웹사이트 '교통량 정보' 메뉴 참고

[별첨] 교통량 지점 정보

index	column name	type	Sample
1	지점번호	character	A-01
2	지점명칭	character	성산로(금화터널)
3	위도	numeric	37.568752
4	경도	numeric	126.948938
5	주소	character	서대문구 신촌동 1-142
6	유입 링크 아이디	numeric	1120006400
7	유출 링크 아이디	numeric	1120006300
8	유입 방향	character	봉원고가차도->독립문역
9	유출방향	character	독립문역->봉원고가차도

위험운전 행동분석 기준

위험운전행동		정의	참고 사항
과속 유형	과속	도로 제한속도 대비 20km/h 초과 운행	연속적인 행동은 1건으로 분석 행동별 간격 3초 이내는 1건, 4초 이후는 별도 건
	장기과속	도로 제한속도 대비 20km/h 초과, 3분 이상 운행	
급가속 유형	급가속	6.0km/h 이상 속도에서 초당 8km/h 이상 가속 운행	
	급출발	5.0km/h 이하 속도에서 출발 초당 10km/h 이상 가속 운행	
급감속 유형	급감속	초당 14km/h 이상 감속 운행 속도가 6.0km/h 이상인 경우	연속적인 급감속 행동은 1건으로 분석 급감속 행동별 간격 3초 이내는 1건, 4초 이후는 별도 건
	급정지	초당 14km/h 이상 감속 속도가 5.0km/h 이하가 된 경우	
급회전 유형	급좌우회전	속도가 30km/h 이상 3초 안에 좌/우측으로 급회전	좌/우측 급회전 누적 회전각은 60~120° 범위
	급U턴	속도가 25km/h 이상 6초 안에 좌/우측으로 운행	160~180° 운행 범위
급진로 변경유형	급앞지르기	30km/h 이상속도에서 진행방향이 좌/우측 10°sec 이 상으로 차로 변경, 5초 동안 누적각도가 $\pm 2^\circ/\text{sec}$ 이하, 가감속이 초당 $\pm 2\text{km/h}$ 이하인 경우	
	급진로변경	30km/h 이상속도에서 진행방향이 좌/우측 10°sec 이 상으로 차로 변경, 5초 동안 누적각도가 $\pm 2^\circ/\text{sec}$ 이하, 가감속이 초당 3km/h 이상인 경우	
연속운전		4시간 이상 운행, 10분 이하 휴식	11대 위험운전행동에 포함되지 않음

본 분석에 대한 가이드

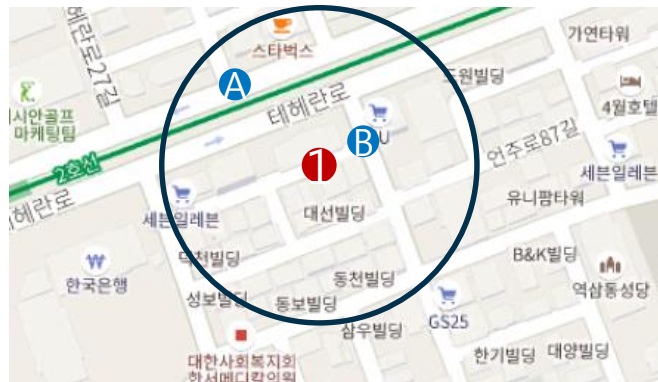
- 모든 위, 경도 데이터는 좌표계 WGS84를 따름
- 도로와 시설물 등에 포함된 폭, 길이 데이터의 단위는 모두 미터(m)
- 연관된 공공데이터를 찾아 활용 가능: 단, 공공데이터만 가능, 출처 표기 필수
- 주어진 모든 데이터를 활용할 필요는 없음, 분석가의 판단 하에 유연하게 활용 가능

02 분석 참고 사항

본 분석에 대한 가이드



- 위, 경도는 같지만 도로 특성이 전혀 다른 두 도로 구간 유의하여 분석 진행



- 구간 ① 에서 발생한 사고에 A 도로 구간의 특징이 반영 되면 안됨

[참고] 상세한 분석 가이드는 1차 결과 발표 후 안내 예정

SAS 분석 솔루션

- 모든 과정은 SAS 소프트웨어를 이용해야 함
- 1. 데이터 처리
 - ✓ SAS® Enterprise Guide®
 - ✓ SAS® Studio
- 2. 데이터 탐색 및 시각화: SAS® Home
 - ✓ SAS® Visual Analytics
 - ✓ SAS® Visual Statistics
 - ✓ SAS® Visual Data Mining and Machine Learning
- 3. 예측 모델링 : SAS® Studio 또는 SAS® Home
 - ✓ SAS® Visual Statistics
 - ✓ SAS® Visual Data Mining and Machine Learning



행정안전부와 함께하는

제16회 SAS 분석 챔피언십

감사합니다



주최 한국싸스소프트웨어(유) | 후원 행정안전부 | 자문 한국데이터마케팅학회


THE POWER TO KNOW.