

빅데이터 분석 결과 시각화

전국 주요 상권 현황 시각화

학습내용

- 전국 상권 현황 시각화
- 창업데이터 시각화
- 그래프 알고리즘 실습

학습목표

- 전국 주요 상권 현황 자료를 수집하고 시각화할 수 있다.
- 상권정보 사이트에서 창업데이터 시각화 자료를 수집하고 시각화 할 수 있다.
- 그래프 알고리즘을 만들 수 있다.

● 전국 상권 현황 시각화

1. 전국 주요 상권 현황 시각화란?

◆ 상권의 개념 및 분류

- 상권
 - 점포 창업 및 영업을 통해 고객을 흡수하고 그 결과 매출을 발생할 수 있는 모든 지역
 - 고객이 흡인되는 지리적 범위, 즉 해당 점포나 사무실을 이용하는 고객들의 거주지역
- 지역상권 : 특정지역 전체
- 지구상권 : 지역상권내에서 상업 지역 공간인 지구 전체
- 지점상권(점포상권) : 지구상권 내에서 특정 입지의 점포가 있는 지점 전체

1차 상권	<ul style="list-style-type: none">• 사업장 이용고객의 60% ~ 70%를 포함하는 범위• 점포인 경우는 약 500m 반경 이내의 지점(도보 5분 거리)
2차 상권	<ul style="list-style-type: none">• 사업장 이용고객의 14% ~ 25%를 포함하는 범위• 점포인 경우는 약 1km 반경 이내의 지점(도보 15분 이내)
3차 상권	<ul style="list-style-type: none">• 1차 상권, 2차 상권 이외의 고객을 포함하는 범위• 점포인 경우는 약 2km 반경 이외의 지구(도보 불가)

- 원료지향형 입지
- 수송지향형 입지
- 시장지향형 입지

● 전국 상권 현황 시각화

1. 전국 주요 상권 현황 시각화란?

◆ 상권의 개념 및 분류

- 특급 상권
 - 시간당 유동인구가 1만 명에 가까운 상권
 - 대표적인 예 : 종로, 명동, 신촌, 강남역 등
 - TV에서 보는 각종 브랜드가 포진하고 있음
 - 점포의 규모도 눈에 확연히 드러날 정도로 큼
 - 유동 인구도 남녀노소를 막론하고 포괄적임
- 1급 상권
 - 시간당 유동인구가 5천명 이상이 되는 상권으로 주로 대형 역세권을 포함함
 - 대표적인 예 : 건대 입구, 대학로, 신림동, 노량진, 신천, 양재, 돈암동 등
 - 1급 상권은 특급 상권과 마찬가지로 유명 브랜드가 입점해 있지만 규모가 다소 작고 뜨내기 유동인구와 단골 이동 인구가 비슷함
 - 소비성이 높은 젊은 층이 많음
- 2급(중형) 상권
 - 대단위 아파트 단지가 모인 지역과 지하철 역이 있는 지역 그리고 대학가가 곁들여 있는 지역
 - 시간당 약 3천명 이상이 이동하는 상권
 - 상권의 크기는 1급 상권과 거의 비슷하나 고정 인구가 유입 인구보다 많음
- 소형(동네) 상권
 - 시간당 이동량은 약 1천명 미만임
 - 외부 유입인구는 거의 없고 거주지 인구만이 이동하는 상권으로 표현함

● 전국 상권 현황 시각화

1. 전국 주요 상권 현황 시각화란?

◆ 상권 조사의 주요 항목

- 통계자료 조사
 - 인구수, 세대수, 가족구성원수, 주거형태(단독주택, 아파트, 복합형)
- 상권형태 및 규모파악
 - 주간상권, 야간상권, 고정상권, 유동상권
- 통행인구 조사
 - 성별, 연령별, 시간대별, 요일별 통행객 수를 관찰함
 - 통행객과 통행성격과 통행객의 수준을 파악함
- 통행차량 조사
 - 통행차량의 수와 어느 시간대에 많이 지나가는지 파악함
- 경쟁점포 조사
 - 예상되는 경쟁점포의 이용객 수, 계층, 제품의 가격대, 매장구성의 장단점을 파악함
- 상권의 향후 전망
 - 주변 상권의 확대 또는 축소 가능성, 대형 집객 시설의 개발정보, 주변 건물의 신축 또는 철거 계획 등을 파악함

● 전국 상권 현황 시각화

1. 전국 주요 상권 현황 시각화란?

◆ 상권 분석 원칙

- 출퇴근 방향을 구분함
 - 출근과 퇴근 방향에 적합한 창업 업종은 다름
 - 편의점, 문방구, 스넥 등은 출근 방향
 - 주점, 식당, 제과점, 선물용품점 등은 퇴근 방향
 - 통행차량의 수와 어느 시간대에 많이 지나가는지 파악함
- 상권의 주입구를 결정함
 - 같은 역세권이라도 4개의 기본 출입구 중에서 가장 인구 이동이 빈번한 출입구가 있음
 - 주거지 또는 정류장, 횡단보도의 영향으로 주력 출입구가 결정되므로 반드시 체크해야 함
 - 상권의 주 출입구를 결정해야 점포 위치를 잡는데 도움이 됨
- 상권 분석의 원칙
 - 상권의 주 출입구가 분석되었다면 양쪽의 상가에서 왼편을 택하는 것이 좋음
 - 오른손잡이 오른쪽 통행에서 길들여진 시선 처리가 왼편에 집중
 - 가급적 상가는 상가 거리의 왼편을 구하는 것이 좋음
- 노점을 주시함
 - 노점상은 하루 벌이 생계형 창업자이기 때문에 노점상의 위치를 주시할 필요가 있음
 - 노점의 좌판이 많으면서 자연스레 흥미 위주의 거리가 형성되기 때문에 고객이 흘러가지 않고 머무는 혜택을 입을 수 있음
- 약국과 금은방 그리고 제과점을 주시함
 - 동네 상권에서 가장 좋은 자리는 대부분 이들 세 업종이 장악하고 있다는 통계가 있음
 - 때문에 이들 점포를 기준으로 분석을 할 필요가 있음

● 전국 상권 현황 시각화

1. 전국 주요 상권 현황 시각화란?

◆ 다변량 데이터 시각화 방법

- 변수 3개 이상의 다변량 데이터(Multivariate Data Set)를 2차원 평면에 효과적으로 시각화할 수 있는 방법
 - 레이더 차트(Radar Chart) or 거미줄 그림(Spider Plot)
 - 별 그래프(Star Graph) : 레이더 차트와 유사하나 중심점 차이 존재함
 - 평행 좌표 그림(Parallel Coordinate Plot)
 - 3차원 산점도(3 Dimensional Scatter Plot)
 - 체르노프 얼굴그림(Chernoff Faces)
 - 산점도 행렬(Scatter Plot Matrix)
 - 모자이크 그림(Mosaic Plot)

◆ 전국 주요 상권 현황 시각화 실습

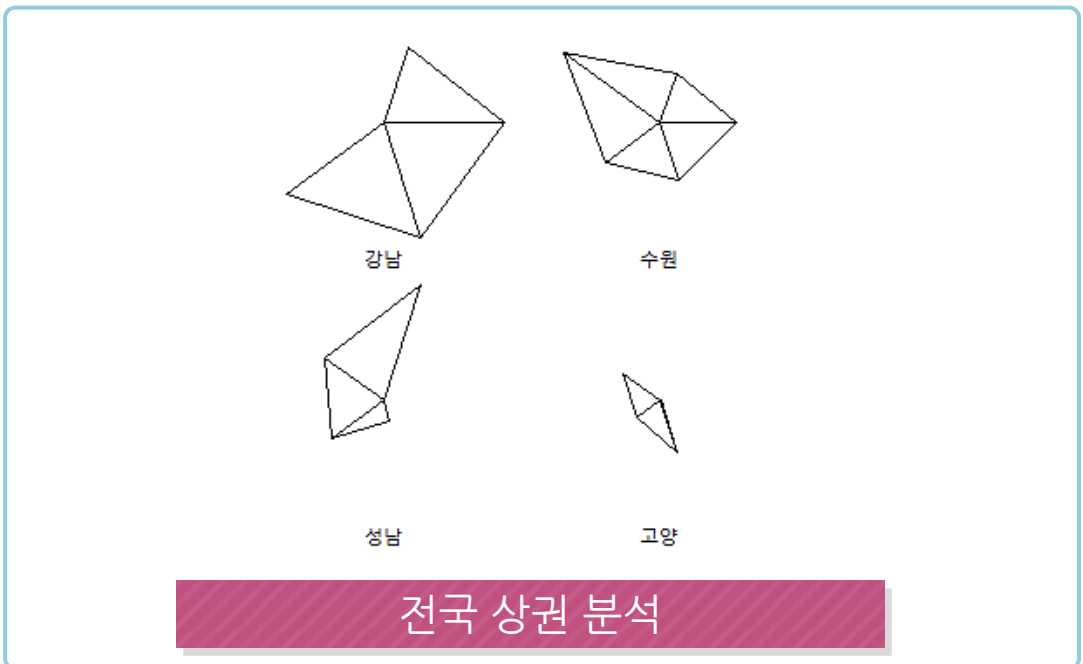
- 별 그래프(Stars 함수)
 - 몇 개의 축을 그리고 중앙으로부터의 거리를 통해 수치를 나타내는 그래프
 - 축구선수 능력치 또는 상권분석에 필요한 지수, 영화평을 나타내는 그래프 등에 활용함
- 별 그래프(Stars 함수) 옵션
 - stars() 함수는 dataframe 이나 matrix 형태의 데이터셋을 사용함
 - scale = TRUE 옵션 사용
 - (minimum value)0 ~ (maximum value)1 사이의 값으로 다변량 변수들의 값을 표준화해줌

● 전국 상권 현황 시각화

1. 전국 주요 상권 현황 시각화란?

◆ 분석 및 시각화에 적용되는 기능들

- 별 그래프(Stars 함수)의 기본원리
 - 중심점(Center Point)으로부터
각 관측치별 / 각 변수별로 거리(Distance) 혹은 반지름(Radius)이 얼마나 떨어져있는가를 시각화함
 - draw.segments=TRUE 옵션을 주면 나이팅게일 차트가 됨



- 소상공인포털 사이트 <http://www.sbiz.or.kr>



● 전국 상권 현황 시각화

2. 전국 주요 상권시각화 실습

◆ 전국 주요 상권시각화 실습 순서

- ① 데이터 준비하기
- ② 데이터 수집하기
- ③ 상권분석.csv 파일 불러오기
- ④ 행번호 이름 변경하기
- ⑤ 별(Stars) 그래프 그리기
- ⑥ 나이팅게일 차트 그리기

● 창업데이터 시각화

1. 창업데이터 시각화 자료준비

① <http://sg.sbiz.or.kr> 사이트 접속

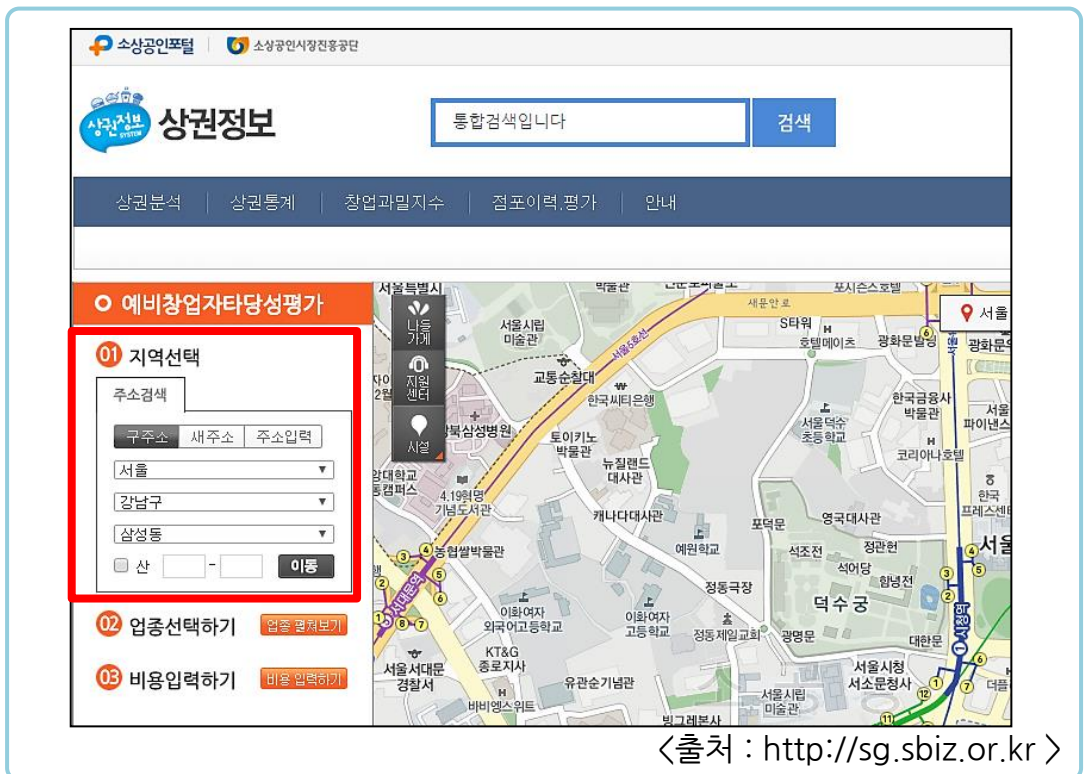
● 예비창업자 타당성 평가 탭 선택



② 지역선택으로 창업지 선정

● 자신이 창업하고자 하는 주소 선택

예) 서울 강남구 삼성동



● 창업데이터 시각화

1. 창업데이터 시각화 자료준비

③ 업종 선택하기

- 제과점 선택 후 확인 버튼 클릭

업종 분류표

업종검색

검색

음식업

서비스업

소매업

간이음식

☐ 제과점

☐ 커피전문점/카페/다방

☐ 패스트푸드

☐ 피자전문

한식/분식

☐ 도시락전문

☐ 떡볶이전문

☐ 라면김밥분식

☐ 죽전문점

☐ 한식/백반/한정식

외식

☐ 양식

☐ 일식/수산물

☐ 중국음식/중국집

최대1개의 업종 선택이 가능합니다.

✓ 확인

취소

〈출처 : <http://sg.sbiz.or.kr>〉

④ 창업 비용 입력

● 비용항목 입력

물건정보					
임대면적	30	평	99	제곱미터	
실 면적	24	평	79	제곱미터	
초기 투자비용					
권리금	6,465	만원	보증금	6,000	만원
대출금/이자(연간)	11,000	만원	%		
기타 투자비	22,545	만원	3년 ※필수입력		
<div>건축비</div> <div>인테리어비</div> <div>설비비</div> <div>가동관련</div> <div>기타비용</div>	<div>480</div> <div>3,600</div> <div>3,600</div> <div>0</div> <div>2,400</div>	만원	※ 투자비/ 감가상각 기간 (※ 감가상각 기간은 인테리어나 설비, 집기 등 기간이 지날수록 가치가 떨어져 고쳐해야 하는 시기까지의 기간을 의미)		
운영비용					
월세	142	만원	인건비	295	만원
재료비	0	만원	기타비용	2,000	만원
※ 객 단가 (※ 1명의 고객이 1회 결제 시 이용하는 평균금액)				11,000 원 ※필수입력	
<div style="border: 1px solid #ccc; padding: 2px 10px; margin-right: 10px;">확인</div> <div style="border: 1px solid #ccc; padding: 2px 10px;">취소</div>					

〈출처 : <http://sg.sbiz.or.kr>〉

● 창업데이터 시각화

1. 창업데이터 시각화 자료준비

⑤ 예비창업자 평가 상권 선택하기



출처 : <http://sg.sbiz.or.kr>

⑥ 예비창업자 평가 상권 선택하기



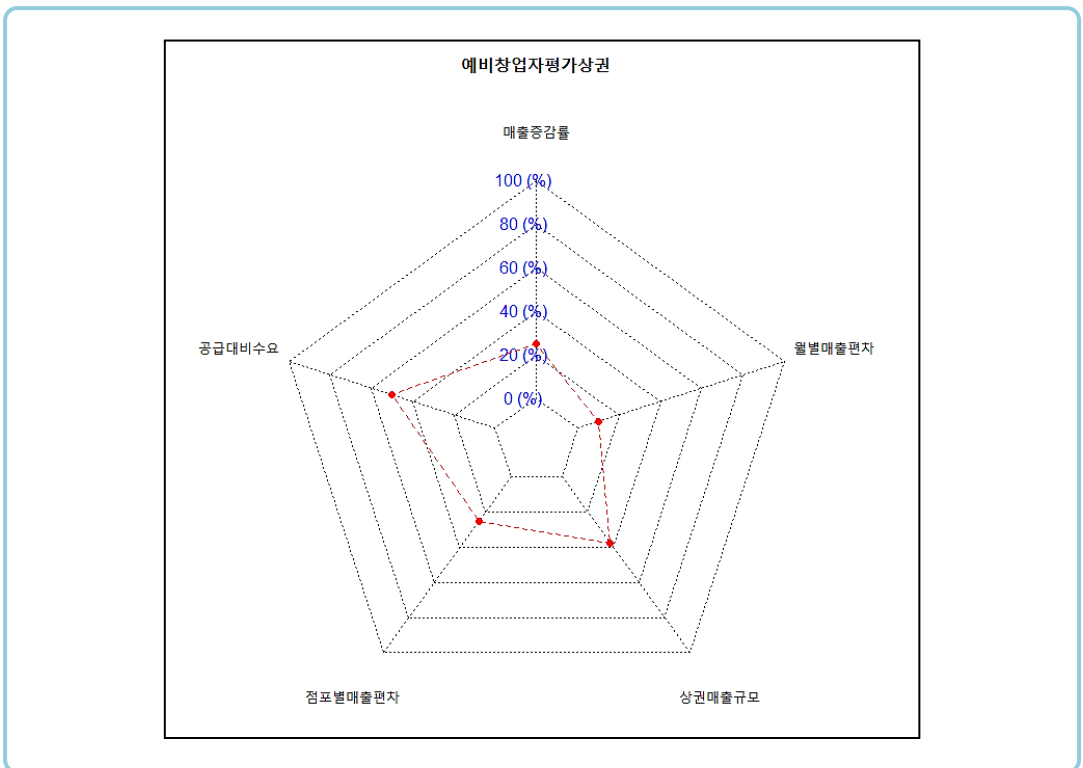
출처 : <http://sg.sbiz.or.kr>

● 창업데이터 시각화

1. 창업데이터 시각화 자료준비

⑦ 레이더 차트 출력하기

- 레이더 차트는 FMSB(Functions for Medical Statistics Book) 패키지를 사용하여 의학관련 다변량데이터 분석에 필요함
- 레이더 차트(Radar Chart)나 거미줄 그림(Spider Plot)
 - 하나의 공통된 Location을 중심점으로 하여 관측치들을 중첩하여 그린 다변량 그래프임



```
> radarchart(data,  
> axistype=1,  
> seg=4,  
> pcol="red",  
> plty=2,  
> title="예비창업자평가상권")
```

- data : 창업평가상권 데이터
- axistype : 그래프의 형태
- seg : 거미줄 개수
- pcol : 컬러지정
- plty : 선의 종류
- title : 제목 표시

● 창업데이터 시각화

2. 창업데이터 시각화 실습

◆ 창업데이터 시각화 실습 순서

- ① FMSB패키지 설치하기
- ② 데이터 파일 만들기
- ③ 열의 이름 바꾸기(Colnames) 및 rbind로 데이터 합치기
- ④ 레이더차트 그리기
- ⑤ 레이더차트에 색상 그리기

● 그래프 알고리즘 실습

1. 조건문과 반복문 이해

◆ 조건문

- if ~ else if 문 사용하기

```
> if ( 조건식 1 ) {  
  조건식 1일 때 실행될 문장  
}
```

```
> else if ( 조건식 2 ) {  
  조건식 1이 아니고 조건식 2일 경우 실행될 문장  
}
```

```
> else {  
  조건식 1도 아니고 조건식 2도 아닐 경우 실행될 문장  
}
```

◆ 조건문 예제

- 입력된 숫자가 0보다 크거나 같으면 “사랑”을 출력하고 0보다 작을 경우 “이별”값을 출력하는 문장 만들기

```
> feeling <- function(x) {  
+   if ( x >= 0 ) {  
+     x <- “사랑”  
+     return(x) }  
+   else {  
+     x <- “이별”  
+     return(x) }  
+ }
```

- 결과 값

```
feeling(5)  
[1] "사랑"  
  
> feeling(0)  
[1] "사랑"  
  
> feeling(-3)  
[1] "이별"
```

● 그래프 알고리즘 실습

1. 조건문과 반복문 이해

◆ 반복문

```
> for ( 변수 in 횟수 ) {  
  < 반복할 식 >  
}
```

◆ 반복문 예제

- 사용자에게 숫자를 입력 받은 후 1 부터 그 숫자까지 연속적인 숫자 출력하기

```
> for(i in 1:5){  
  + print(i)  
  + }  
[1] 1  
[1] 2  
[1] 3  
[1] 4  
[1] 5
```


● 그래프 알고리즘 실습

2. 그래프 알고리즘 실습

◆ 그래프 알고리즘 실습 순서

- ① 데이터 준비하기
- ② For문을 이용하여 레이더 차트 컬러와 제목 바꾸기
- ③ 레이더차트 그리기

1. 전국 상권 현황 시각화

■ 상권의 개념 및 분류

- 상권이란 고객이 흡인되는 지리적 범위, 즉 해당 점포나 사무실을 이용하는 고객들의 거주지역을 일컫음
- 지역상권 : 특정지역 전체
- 지구상권 : 지역상권내에서 상업 지역 공간인 지구 전체
- 지점상권 : 지구상권 내에서 특정 입지의 점포가 있는 지점 전체
- 이러한 상권 정보를 시각화하는 것을 상권 현황 시각화라 함

■ 상권 조사의 주요 항목 및 분석 원칙

■ 상권 조사의 주요 항목

- ① 통계자료 조사 ② 상권형태 및 규모파악 ③ 통행인구 조사
- ④ 통행차량 조사 ⑤ 경쟁점포 조사 ⑥ 상권의 향후 전망

■ 상권 분석의 원칙

- ① 출퇴근 방향을 구분함
- ② 상권의 주입구를 결정함
- ③ 도보 라인의 왼편 상가를 공략함
- ④ 노점을 주시함
- ⑤ 약국과 금은방 그리고 제과점을 주시함

2. 창업데이터 시각화

■ 창업데이터 시각화 자료준비

- <http://sg.sbiz.or.kr> 사이트 활용
- 지역선택으로 창업지 선정
- 업종 선택하기
- 창업 비용 입력
- 예비창업자 평가 상권 선택하기
- 레이더 차트 출력하기

3. 그래프 알고리즘 실습

■ 조건문

- if ~ else if 문 사용하기

```
> if ( 조건식 1 ) {  
    조건식 1일 때 실행될 문장  
}
```

```
> else if ( 조건식 2 ) {  
    조건식 1이 아니고 조건식 2일 경우 실행될 문장  
}
```

```
> else {  
    조건식 1도 아니고 조건식 2도 아닐 경우 실행될 문장  
}
```

■ 반복문

```
for ( 변수 in 횟수 ) {  
    <반복할 식>  
}
```