

빅데이터 분석 결과 시각화

연도별 데이터 시각화

학습내용

- 시계열 분석 시각화
- 연도별 데이터 수집
- 연도별 데이터 시각화

학습목표

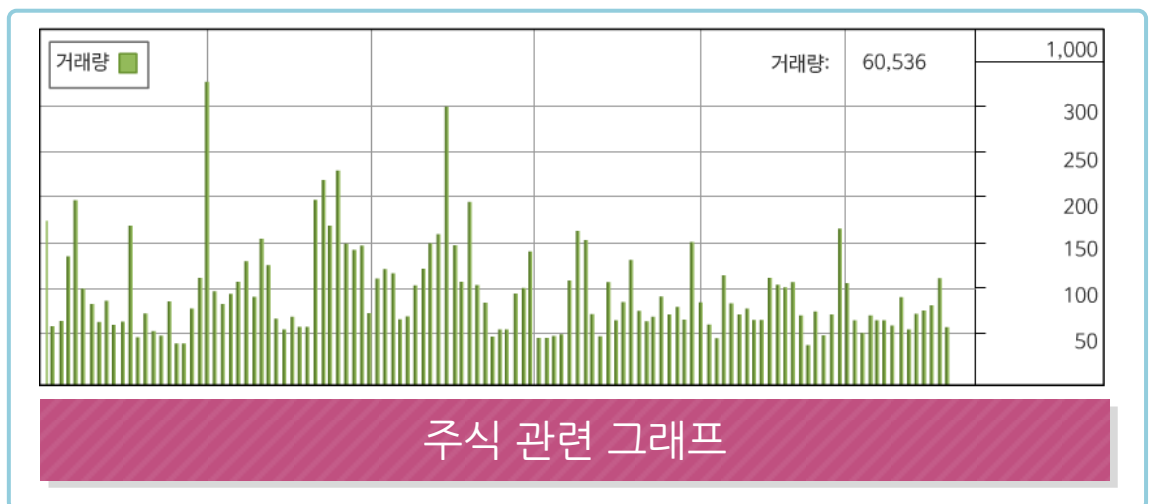
- 시계열 분석에 대한 데이터를 활용하여 시각화할 수 있다.
- 날짜와 시간 축 포맷에 대해 이해하고, 연도별 데이터를 수집할 수 있다.
- 수집한 연도별 데이터를 활용하여 시각화할 수 있다.

● 시계열 분석 시각화

1. 시계열 분석의 개념

◆ 시계열 분석

- 시계열(時系列, Time Series)
 - 일정 시간 간격으로 배치된 데이터들의 수열
- 시계열 해석(Time Series Analysis)
 - 시계열을 해석하고 이해하는 데 쓰이는 여러 가지 방법을 연구하는 분야
- 시계열 예측(Time Series Prediction)
 - 주어진 시계열을 보고 수학적 모델 만들어서 미래에 일어날 것들을 예측하는 것
 - 일반적으로 이런 방법들은 공학이나 과학계산, 혹은 금융시장에서의 주가 예측 등에서 많이 쓰임



시계열 분석 시각화

1. 시계열 분석의 개념

◆시계열 데이터 분석 목적

미래 값 예측	데이터의 특성 파악
<ul style="list-style-type: none"> 일주일간 주가 예측, 다음 달 매출액 예측 등 	<ul style="list-style-type: none"> 경향(Trend), 주기(Cycle), 계절성(Seasonality), 불규칙성(Irregular) 등



● 시계열 분석 시각화

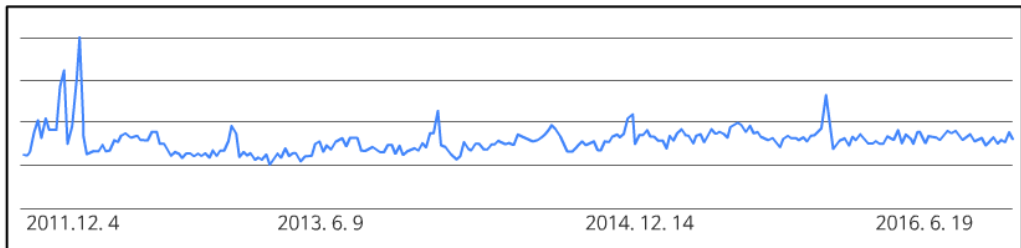
1. 시계열 분석의 개념

◆ 시계열 자료의 예

- 국민총생산, 물가지수, 주가지수 등 경제활동과 관련된 시계열
- 일일 강수량, 기온, 연간 지진 발생 수 등 물리적 현상과 관련된 시계열
- 상품판매량, 상품광고액, 상품재고량 등 경영활동과 관련된 시계열
- 총인구, 농가 수, 인구증가율 등 인구와 관련된 시계열
- 품질관리 등 생산관리와 관련된 시계열
- 월별 교통사고 건수, 월별 범죄발생 수 등 사회생활과 관련된 시계열

◆ 시계열 분석의 통계 그래프

- 구글 트렌드에서 검색량을 보면 어떤 데이터의 흐름에 따라 관심도의 변화를 보고 이후 통계량을 예측할 수 있음



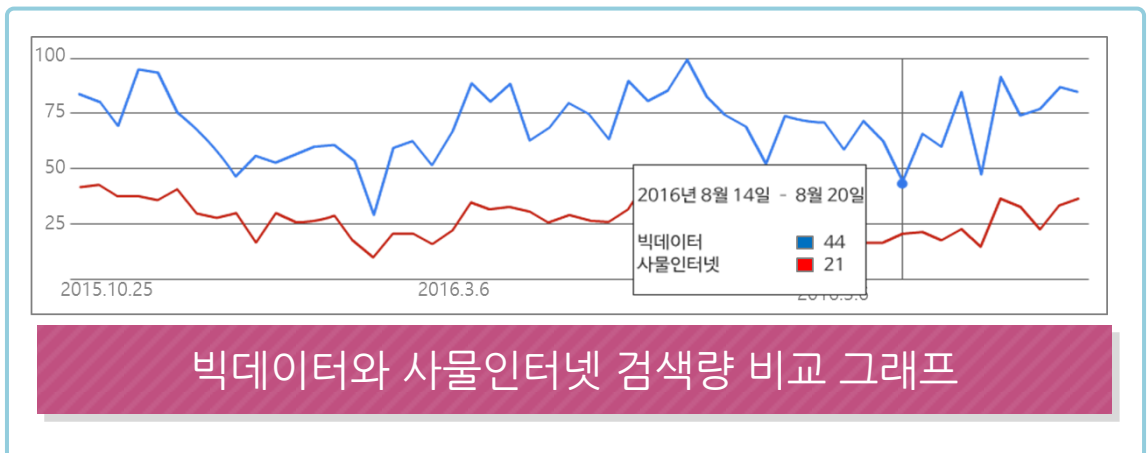
5년 간 '치킨' 검색량에 따른 관심도 변화 그래프

● 시계열 분석 시각화

1. 시계열 분석의 개념

◆ 예측 기법

- 예측이 왜 필요한가?
- 누가 예측치를 사용할 것인가?
- 가용한 데이터의 특징은 무엇인가?
- 어느 정도의 기간을 예측할 것인가?
- 최소한의 데이터 필요는 어느 정도인가?
- 어느 정도의 정확도를 요하는가?
- 예측에 수반되는 비용은 얼마인가?



● 시계열 분석 시각화

2. 시계열 분석 시각화 실습

◆ 시계열 분석 시각화 실습 순서

- ① 데이터 준비하기
- ② 날짜 포맷 핵심 코드 입력하기
- ③ 축 표현하기
- ④ 결과화면 확인하기

● 연도별 데이터 수집

1. 날짜 / 시간 축 포맷

◆ 타임 포맷 설정하기

```
var margin = {top: 20, right: 20, bottom: 100, left: 50},
```

```
svg.append("g")
  .attr("class", "axis")
  .attr("transform", "translate(0, "+ height + ")")
  .call(d3.axisBottom(x)
    .tickFormat(d3.timeFormat("%Y-%m-%d"))))
.selectAll("text")
  .style("text-anchor", "end")
  .attr("dx", "-.8em")
  .attr("dy", ".15em")
  .attr("transform", "rotate(-65)");
```

↑ 날짜와 시간에 따른 축의 값 표현

● 연도별 데이터 수집

1. 날짜 / 시간 축 포맷

◆ 타임 포맷의 종류

종류	의미
%a	요일 이름을 약어로 표현
%A	전체 요일 이름
%b	월 이름을 약어
%B	전체 월 이름
%c	날짜 및 시간, “%a %b %e %H:%M:%S %Y”
%d	진수 [01, 31]와 달의 0이 들어간 일
%e	진수 [1, 31] 같은 달의 일
%H	시간 (24시간) 진수 [00,23] 등
%I	진수 [01, 12]와 같은 시간 (12시간 시계)
%j	진수 [001, 366] 등 한 해의 날
%m	진수 [01, 12] 같은 달
%M	진수 [00, 59]와 같은 분
%p	AM 또는 PM 중 하나
%S	진수 [00, 61]로 초
%U	진수 [00, 53]로 올해 (일요일 주일의 첫날로)의 주 번호
%w	십진수로 요일 [0(일), 6]
%W	진수 [00, 53]로 올해 (월요일주의 첫 번째 날 등)의 주 번호
%x	날짜 “%m/%d/%y”
%X	시간 “%H:%M:%S”
%y	진수 [00,99]와 같은 세기 없이 해
%Y	십진수로 세기와 해
%Z	시간대는 “-0700”으로 오프셋

● 연도별 데이터 수집

1. 날짜 / 시간 축 포맷

◆ 타임 포맷의 예시

타임포맷	결과화면
d3.time.format("%Y-%m-%d")	2026-05-28
d3.time.format("%m/%d/%Y")	05/28/1986
d3.time.format("%H:%M")	10:19
d3.time.format("%H:%M %p")	10:19 AM
d3.time.format("%B %d")	January 28
d3.time.format("%d %b")	28 Jan
d3.time.format("%d-%b-%y")	28-Jan-86
d3.time.format("%S s")	13 s
d3.time.format("%M m")	39 m

● 연도별 데이터 수집

2. 연도별 데이터 수집 실습

◆ 연도별 데이터 수집 실습 순서

- ① 데이터 준비하기
- ② CSS 설정하기
- ③ 날짜 포맷 작성하기
- ④ 요일별 선거범죄 그래프 결과 확인하기

● 연도별 데이터 시각화

1. 연도별 범죄 데이터 만들기

① 범죄 발생 시간별 데이터 수집

범죄 발생시간													
													단위 : 건(발생원표)
구 분	계	00:00-02:5	03:00-05:5	06:00-08:5	09:00-11:5	12:00-14:5	15:00-17:5	18:00-20:5	21:00-23:5	미상			
죄 중													

② 지능범죄 데이터를 새로운 데이터로 만듦

	공갈	4,310	164	108	150	338	378	641	673	478	1,380	
	손괴	57,813	5,957	4,850	3,445	4,512	4,005	5,725	10,224	12,599	6,496	
지능범죄	소계	329,439	7,976	5,339	3,392	55,880	29,908	21,391	14,495	13,122	177,936	
	적부유기	555	3	4	2	161	51	22	12	6	294	
	직권남용	325	9	5	10	60	25	19	17	26	154	
	조소리	253	1	0	1	40	17	15	7	11	161	

③ 범죄 발생시간을 date, 지능범죄수를 close로 바꿈

	A	B	C
1	범죄발생시간	지능범죄수	
2	00:00-02:59	7,976	
3	03:00-05:59	5,339	
4	06:00-08:59	3,392	
5	09:00-11:59	55,880	
6	12:00-14:59	29,908	
7	15:00-17:59	21,391	
8	18:00-20:59	14,495	
9	21:00-23:59	13,122	
10			

crime.csv

	A	B	C
1	date	close	
2	0:00	7976	
3	3:00	5339	
4	6:00	3392	
5	9:00	55880	
6	12:00	29908	
7	15:00	21391	
8	18:00	14495	
9	21:00	13122	
10			

crime2.csv

④ 타임설정 및 축 범위 설정

● 날짜 시간 분석

```
var parseTime = d3.timeParse("%H:%M"); // 시간과 분을 분석
```

● 축 범위설정

```
var x = d3.scaleTime().range([0, width]);
var y = d3.scaleLinear().range([height, 0]);
```

● 연도별 데이터 시각화

1. 연도별 범죄 데이터 만들기

⑤ 데이터 가져오기

- 데이터 가져오기

```
d3.csv("crime2.csv", function(error, data) {  
  if (error) throw error;
```

- 데이터 date, close 값 가져오기

```
data.forEach(function(d) {  
  d.date = parseTime(d.date);  
  d.close = +d.close;  
});
```

⑥ X축 좌표 만들기

- X축의 좌표에 글자 넣기

```
svg.append("g")  
  .attr("class", "axis")  
  .attr("transform", "translate(0, "+ height +)")  
  .call(d3.axisBottom(x)  
    .tickFormat(d3.timeFormat("%H:%M")))  
  .selectAll("text")  
    .style("text-anchor", "end")  
    .attr("dx", "-.8em")  
    .attr("dy", ".15em")  
    .attr("transform", "rotate(-65)");
```

← 시간과 분을 나타냄

● 연도별 데이터 시각화

2. 연도별 범죄 데이터 시각화 실습

◆ 연도별 범죄 데이터 시각화 실습 순서

- ① 데이터 준비하기
- ② CSS 설정하기
- ③ 시간을 설정하는 축의 값 표현하기
- ④ 결과화면 확인하기

1. 시계열 분석 시각화

■ 시계열 분석이란?

- 시계열(時系列, Time Series)은 일정 시간 간격으로 배치된 데이터들의 수열을 말함
- 시계열 해석(Time Series Analysis)라고 하는 것은 이런 시계열을 해석하고 이해하는 데 쓰이는 여러 가지 방법을 연구하는 분야임
- 시계열 예측(Time Series Prediction)이라고 하는 것은 주어진 시계열을 보고 수학적인 모델을 만들어서 미래에 일어날 것들을 예측하는 것을 뜻함
- 일반적으로 이런 방법들은 공학이나 과학계산, 혹은 금융시장에서의 주가 예측 등에서 많이 쓰임

■ 시계열 데이터 분석 목적

- 가장 중요한 목적 : 미래 값을 예측하는 것
- 시계열 데이터의 특성 : 경향(Trend), 주기(Cycle), 계절성(Seasonality), 불규칙성(Irregular) 등

■ 시계열 자료의 예

- 국민총생산, 물가지수, 주가지수 등 경제활동과 관련된 시계열
- 일일 강수량, 기온, 연간 지진 발생 수 등 물리적 현상과 관련된 시계열
- 상품판매량, 상품광고액, 상품재고량 등 경영활동과 관련된 시계열
- 총인구, 농가 수, 인구증가율 등 인구와 관련된 시계열
- 품질관리 등 생산관리와 관련된 시계열
- 월별 교통사고 건수, 월별 범죄발생 수 등 사회생활과 관련된 시계열

핵심요약

1. 시계열 분석 시각화

■ 시계열 분석의 통계 그래프

- 구글 트렌드에서 검색량을 보면 어떤 데이터의 흐름에 따라 관심도의 변화를 보고 이후 통계량을 예측할 수 있음

■ 예측 기법

- 예측이 왜 필요한가?
- 누가 예측치를 사용할 것인가?
- 가용한 데이터의 특징은 무엇인가?
- 어느 정도의 기간을 예측할 것인가?
- 최소한의 데이터 필요는 어느 정도인가?
- 어느 정도의 정확도를 요하는가?
- 예측에 수반되는 비용은 얼마인가?

2. 연도별 데이터 수집

■ 타임 포맷의 종류

- %a - 요일 이름을 약어로 표현
- %A - 전체 요일 이름
- %b - 월 이름을 약어
- %B - 전체 월 이름
- %c - 날짜 및 시간, "%a %b %e %H:%M:%S %Y"
- %d - 진수 [01,31]와 달의 0이 들어간 일
- %e - 진수 [1,31] 같은 달의 일
- %H - 시간 (24 시간) 진수 [00,23] 등
- %I - 진수 [01,12]와 같은 시간 (12 시간 시계)
- %j - 진수 [001,366] 등 한 해의 날
- %m - 진수 [01,12] 같은 달
- %M - 진수 [00,59]와 같은 분
- %p - AM 또는 PM 중 하나
- %S - 진수 [00,61]로 초
- %U - 진수 [00,53]로 올해 (일요일 주일의 첫날로)의 주 번호
- %w - 십진수로 요일 [0 (일), 6]
- %W - 진수 [00,53]로 올해 (월요일주의 첫 번째 날 등)의 주 번호
- %x - 날짜 "%m/%d/%y"
- %X - 시간, "%H:%M:%S"
- %y - 진수 [00,99]와 같은 세기 없이 해
- %Y - 십진수로 세기와 해
- %Z - 시간대는 "-0700"으로 오프셋

2. 연도별 데이터 수집

■ 타임 포맷의 예시

타임포맷	결과화면
<code>d3.time.format("%Y-%m-%d")</code>	2026-05-28
<code>d3.time.format("%m/%d/%Y")</code>	05/28/1986
<code>d3.time.format("%H:%M")</code>	10:19
<code>d3.time.format("%H:%M %p")</code>	10:19 AM
<code>d3.time.format("%B %d")</code>	January 28
<code>d3.time.format("%d %b")</code>	28 Jan
<code>d3.time.format("%d-%b-%y")</code>	28-Jan-86
<code>d3.time.format("%S s")</code>	13 s
<code>d3.time.format("%M m")</code>	39 m

3. 연도별 데이터 시각화

■ . 연도별 범죄 데이터 만들기

- ① 범죄 발생 시간별 데이터 수집
- ② 지능범죄 데이터를 새로운 데이터로 만들기
- ③ 범죄 발생시간을 date, 지능범죄수를 close로 바꿈
- ④ 타임설정 및 축 범위 설정
- ⑤ 데이터 가져오기
- ⑥ X축 좌표 만들기