

Elastic Stack 을 활용한 Data Dashboard 만들기

Week 2 - Kibana를 더 잘 사용해보자



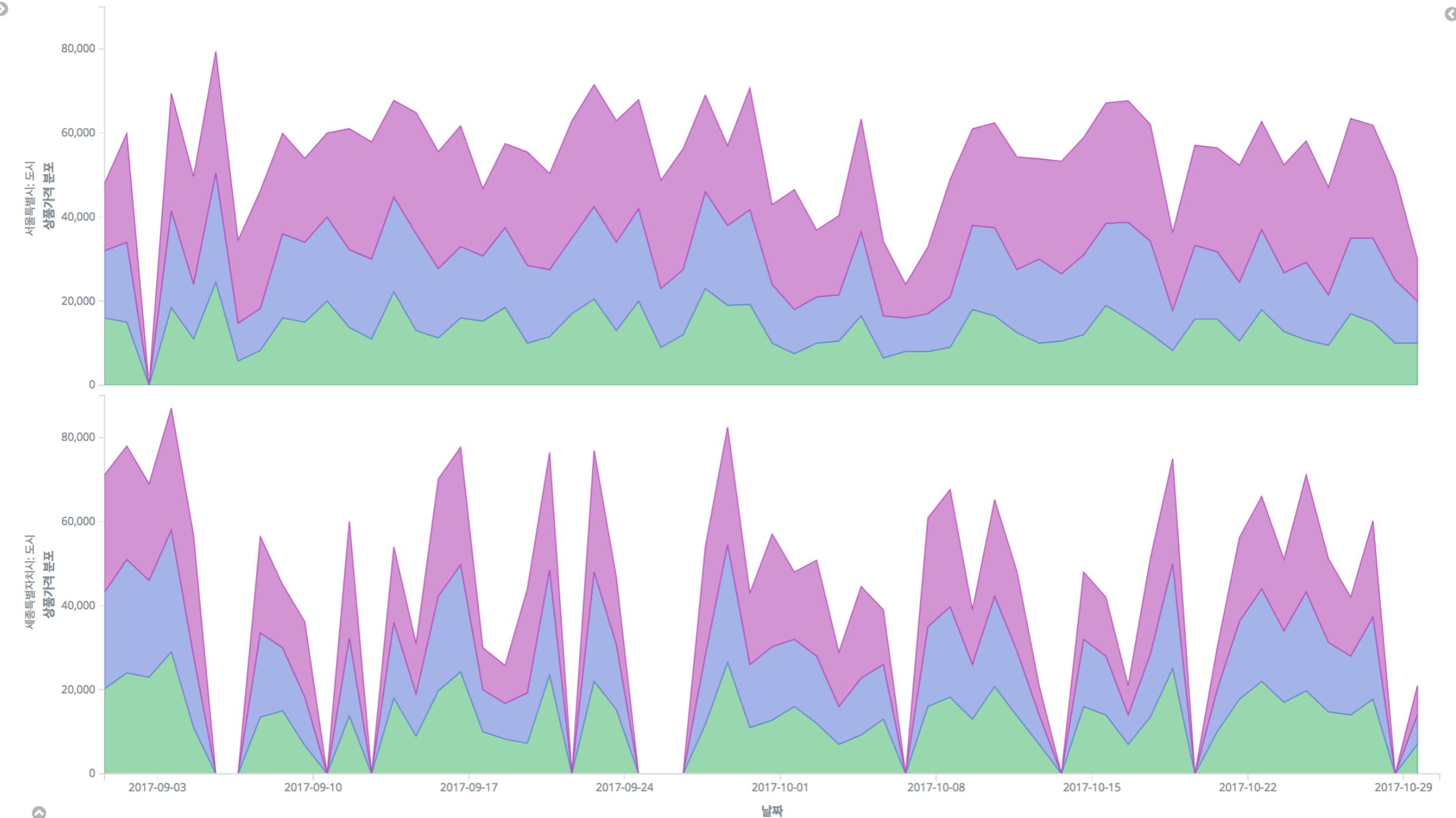
Fast Campus

목차

- 실전 Visualize
 - Area Map 4
 - Gauge 8
 - Heat Map 12
 - Data Table 16
- 실전 Dashboard 20
- 몇 가지 고민들
 - data format 변경 31
 - csv 출력 56
 - JSON Input 61
 - Scripted Field 79
- 데이터 조회
 - Filtering by Field 101
 - Lucene Query 114

문제 1 - Area Map 🔍 ⚒

Visualize - Area Map



Visualize - Area Map

데이터

- Index : shopping
- Time Range : 2017년 9월 1일 ~ 2017년

문제

- “상품가격”의 합이 가장 큰
- “고객주소_시도” 2개의
- “상품가격”의 25분위, 50분위, 95분위를
- “주문시간”을 기준으로 daily로 표시하고
- week2_{id}_area로 저장하세요

사용한 Aggregation

- Metrics (Y-Axis) : **Percentiles** Aggregation
- Buckets
 - Split Chart (Rows) : **Terms** Aggregation
 - X-Axis : **Date Histogram** Aggregation



순서에 주의!!

순서가 바뀌면 결과가 왜 바뀌는지 생각해보자

사용 필드

- 상품가격
- 고객주소_시도
- 주문시간

Visualize - Area Map

metrics

Y-Axis
Aggregation
Percentiles

Field
상품가격

Percents
25



+ Add

Custom Label
상품가격 분포

buckets

Split Chart
Rows Columns

Aggregation


Field
고객주소_시도

Order By
Custom Metric

Aggregation


Field
상품가격

Order
Descending

Size


Custom Label


X-Axis

Sub Aggregation
Date Histogram

Field


Interval

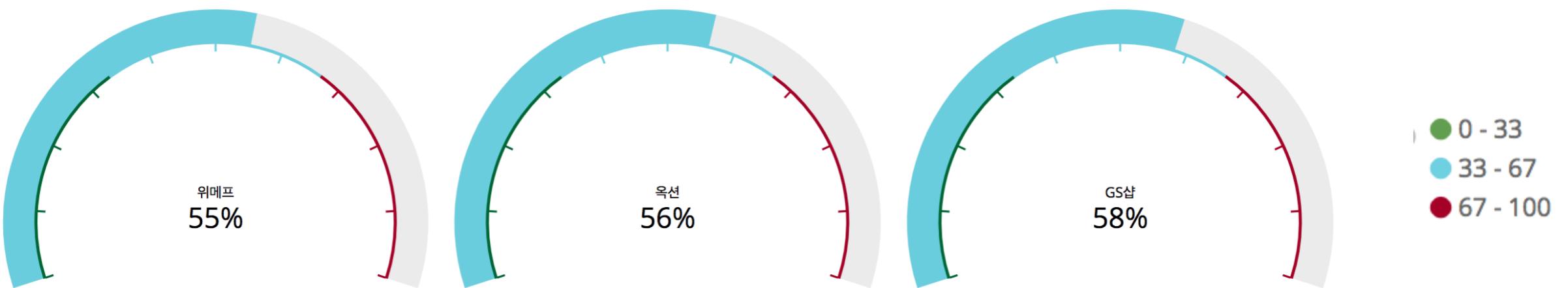

Custom Label
날짜

Add sub-buckets

Advanced

문제 2 - Gauge 🔪 ⚒

Visualize - Gauge



일별 상품가격 평균의 전체 평균

Visualize - Gauge

데이터

- Index : shopping
- Time Range : 2017년 9월 1일 ~ 2017년

문제

- “판매자평점”의 평균이 큰
- “구매사이트” 3개를
- “주문시간” 을 기준으로 daily로 나누고
- daily “상품가격”의 평균을 낸 후
- daily “상품가격”의 평균의 전체 평균을 표시하고
- week2_{id}_gauge로 저장하세요

단계별 목표 구간

- 0 ~ 9999
- 10000 ~ 19999
- 20000 ~ 29999

사용한 Aggregation

- Metrics
 - Aggregation : Average Bucket Aggregation
 - Bucket : Date Histogram Aggregation
 - Metric : Average Aggregation
- Buckets
 - Aggregation : Terms Aggregation

단순히 Average Aggregation 사용하는 것과 뭐가 다를까?
또는 어떤 상황에서 두 값의 차이가 크게 날까?

사용 필드

- 주문시간
- 상품가격
- 구매사이트
- 판매자평점

Visualize - Gauge

metrics

Metric

Aggregation

Average Bucket

Bucket

Aggregation

주문시간

Field

주문시간

Interval

Daily

Custom Label

◀ Advanced

Metric

Aggregation

상품가격

Field

상품가격

Custom Label

◀ Advanced

Custom Label

일별 상품가격 평균의 전체 평균

buckets

Split Group

Aggregation

Terms

Field

주문시간

Order By

Custom Metric

Aggregation

Average

Field

주문시간

◀ Advanced

Order

주문시간

Size

3

Custom Label

Data Options X

Gauge Type Arc

Percentage Mode

Vertical Split

Show Legend

Show Labels

Sub Text

Auto Extend Range

Ranges

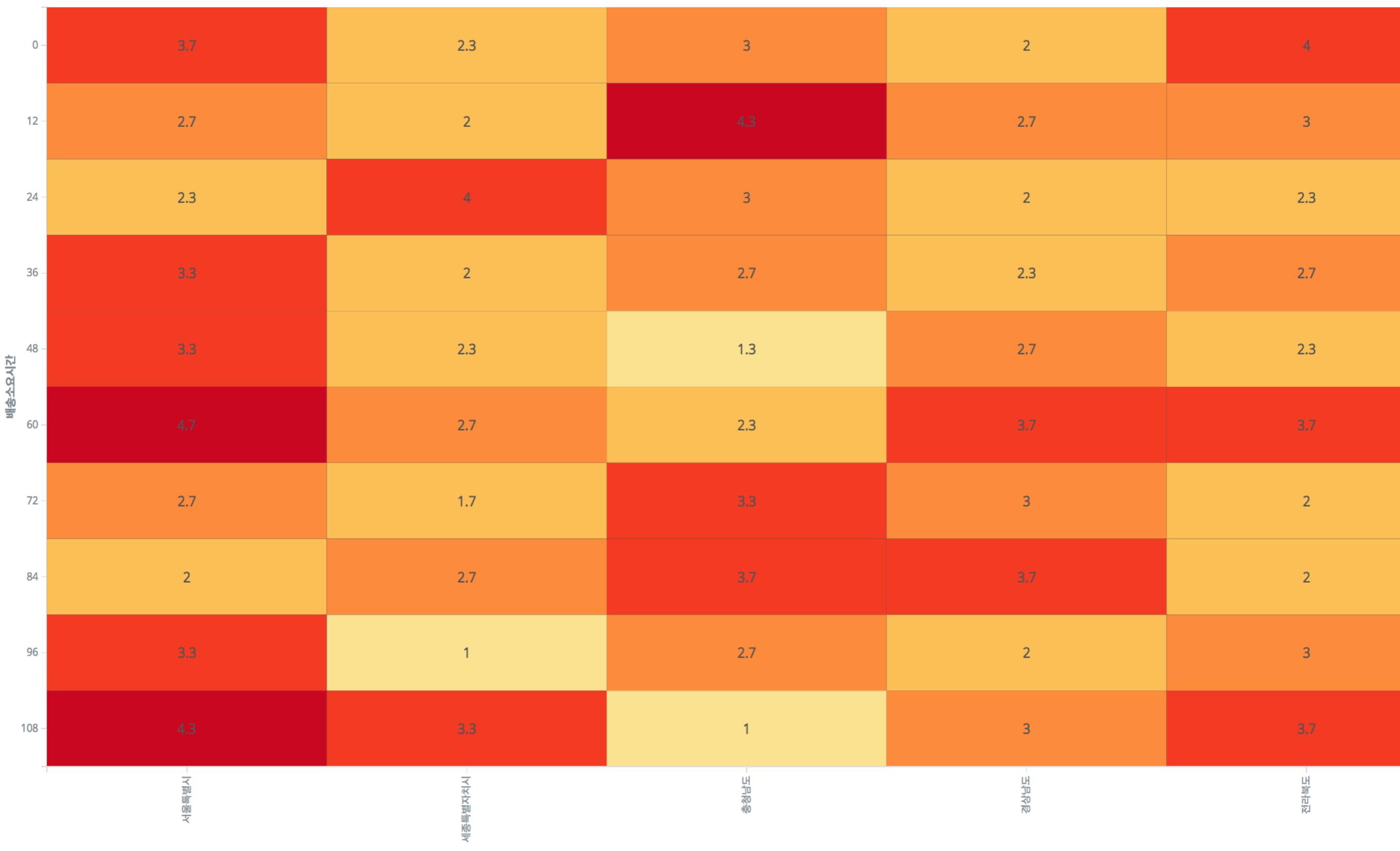
From	To
0	주문시간
10000	19999
주문시간	29999

Add Range

Note: colors can be changed in the legend

문제 3 - Heat Map 🔍 ⚒

Visualize - Heat Map



Visualize - Heat Map

데이터

- Index : shopping
- Time Range : 2017년 9월 1일 ~ 2017년

문제

- “상품가격”의 합이 큰 “고객주소_시도” 5개를 X축으로 하고
- “배송소요시간”을 12시간 간격으로 나누어 Y축으로 한 후
- “고객나이”가 가장 작은 3명의
- “판매자평점”의 평균을 표시하고,
- week2_{id}_heatmap으로 저장하세요

사용한 Aggregation

- Metrics
 - Aggregation : Top Hit Aggregation
- Buckets
 - X-Axis : Terms Aggregation, Sum Aggregation
 - Y-Axis : Histogram Aggregation

사용 필드

- 판매자평점
- 고객주소_시도
- 배송소요시간
- 상품가격
- 고객나이

Visualize - Heat Map

shopping

Data Options  

metrics

Value

Aggregation

Top Hit

Field 

Aggregate With  Size 

Average  

Sort On

고객나이 

Order

Custom Label



shopping

Data Options  

Advanced

buckets

X-Axis   

Aggregation 

Field 고객주소_시도

Order By Custom Metric

Aggregation 

Field 상품가격

Order  Size  5

Custom Label 주소 

Y-Axis   

Sub Aggregation Histogram

Field 

Interval 12

shopping

Data Options  

Basic Settings

Show Tooltips

Highlight

Legend Position right

Heatmap Settings

Color Schema Yellow to Red 

Reverse Color Schema

Color Scale linear 

Scale to Data Bounds

Percentage Mode

Number of colors 6

Custom Ranges 

Show Labels

문제 4 - Data Table



Visualize - Data Table

날짜 ◆	인기 Top 3 ◆	매출 ◆	매출 증감 ◆	매출 누적 ◆
09월01일	가디건, 스웨터, 원피스	331,000	-	331,000
09월02일	티셔츠, 가디건, 가디건	419,000	88,000	750,000
09월03일	원피스, 청바지, 블라우스	407,000	-12,000	1,157,000
09월04일	팬츠, 청바지, 니트	443,000	36,000	1,600,000
09월05일	셔츠, 가디건, 남방	433,000	-10,000	2,033,000
09월06일	점퍼, 남방, 니트	341,000	-92,000	2,374,000
09월07일	코트, 셔츠, 가디건	360,000	19,000	2,734,000
09월08일	셔츠, 니트, 남방	428,000	68,000	3,162,000
09월09일	가디건, 남방, 가디건	358,000	-70,000	3,520,000
09월10일	가디건, 코트, 셔츠	364,000	6,000	3,884,000

Visualize - Data Table

데이터

- Index : shopping
- Time Range : 2017년 9월 1일 ~ 2017년

문제

“상품가격”이 가장 비쌌던 “상품분류” 3개를 “주문시간” 기준으로 일별로 선별하고 다음을 각각 표시하고 week2_{id}_datatable로 저장하자.

- “상품가격”의 합
- “상품가격”의 합의 일별 증감
- “상품가격”의 일별 누적합

사용한 Aggregation

- Metrics
 - Aggregation
 - Top Hit Aggregation
 - Sum Aggregation
 - Derivative Aggregation
 - Cumulative Sum Aggregation
- Buckets
 - (Split Rows) Aggregation : Date Histogram Aggregation

사용 필드

- 상품가격
- 상품분류
- 고객나이

Visualize - Data Table

shopping

Data Options ▶ ×

Metric ○ ↑ ×

Aggregation
Top Hit ↑

Field ◀ ↑ ×

Field ◀ ↑ ×

Aggregate With i Size i
Concatenate ↑ ◀ ↑ ×

Sort On ◀ ↑ ×

Order ◀ ↑ ×

Custom Label
인기 Top 3

Metric ○ ↑ ×

Aggregation
Sum ↑

Field ◀ ↑ ×

Custom Label
매출

Metric ○ ↑ ×

Aggregation
Derivative ↑

Metric ◀ ↑ ×

Custom Label
매출 증감

Advanced ◀

Metric ○ ↑ ×

Aggregation
Cumulative Sum ↑

Metric ◀ ↑ ×

Custom Label
매출 누적

buckets

Split Rows ○ ×

Aggregation
Date Histogram ↑

Field ◀ ↑ ×

Interval ◀ ↑ ×

Custom Label
날짜

Dashboard 

Kibana - dashboard

Kibana에 접속해서 Dashboard 화면으로 가보자 !

The screenshot shows the Kibana interface. On the left, there is a sidebar with the following options:

- Discover
- Visualize
- Dashboard** (highlighted with a red dashed box and a red hand icon pointing to it)
- Timeline
- Dev Tools
- Management

The main area is titled "Dashboard". It features a search bar at the top right with the placeholder "Search...". Below the search bar is a table with the following columns:

Name	Description
<input type="checkbox"/> data	
<input type="checkbox"/> higee	
<input type="checkbox"/> shopping	
<input type="checkbox"/> test_copy	
<input type="checkbox"/> week1	

A red box highlights the "data" row, and a red hand icon points to it with the text "기존에 생성한 dashboard 조회할 경우 선택". A red box also highlights the blue "+" button in the top right corner of the dashboard list area, with the text "처음 생성하는 경우 선택".

Kibana - dashboard

기존에 만든 visualization을 추가하자

2. 모두 추가한 후 week2_{id}로 저장하자

1. Add를 선택하고 visualization을 하나씩 추가하자 ↗

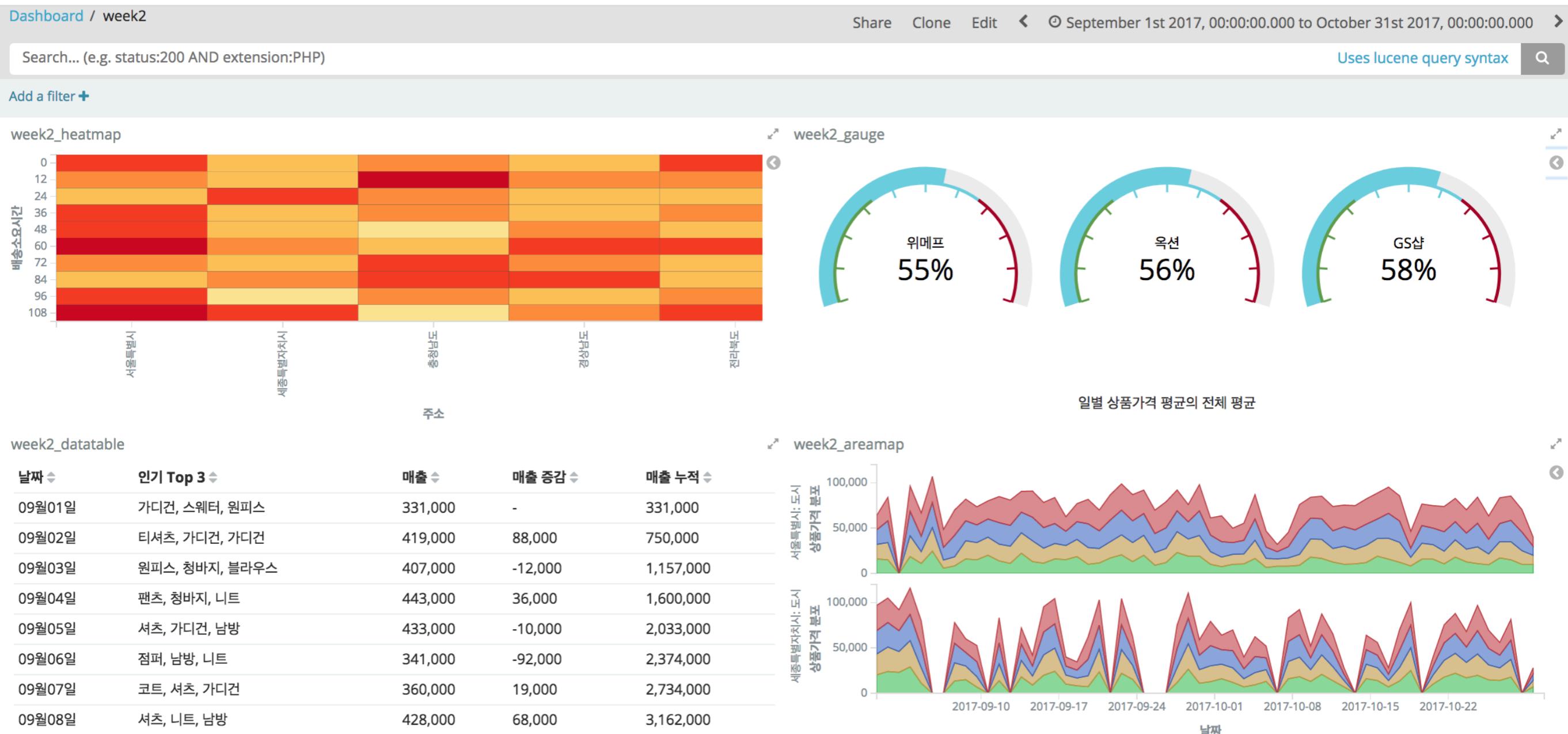
This dashboard is empty. Let's fill it up!

Click the **Add** button in the menu bar above to add a visualization to the dashboard. If you haven't set up any visualizations yet, visit the [Visualize app](#) to create your first.

- week2_{id}_area
- week2_{id}_gauge
- week2_{id}_heatmap
- week2_{id}_datatable

Kibana - dashboard

앞에서 Add를 했다면 아래와 같은 형태로 생성될 것이다



Kibana - dashboard

Auto Refresh 기능을 이용하자

Dashboard / week2

Share Clone Edit **Auto-refresh** September 1st 2017, 00:00:00.000 to October 31st 2017, 00:00:00.000

Refresh Interval

Off 5 seconds 1 minute 1 hour
10 seconds 5 minutes 2 hour
30 seconds 15 minutes 12 hour
45 seconds 30 minutes 1 day

Search... (e.g. status:200 AND extension:PHP)

Add a filter +

1. Auto-refresh 선택

2. Refresh 간격 설정

week2_heatmap

주소

날짜

인기 Top 3

매출

매출 증감

매출 누적

상품가격 분포

날짜

위메프 55%

옥션 56%

GS샵 58%

일별 상품가격 평균의 전체 평균

week2_gauge

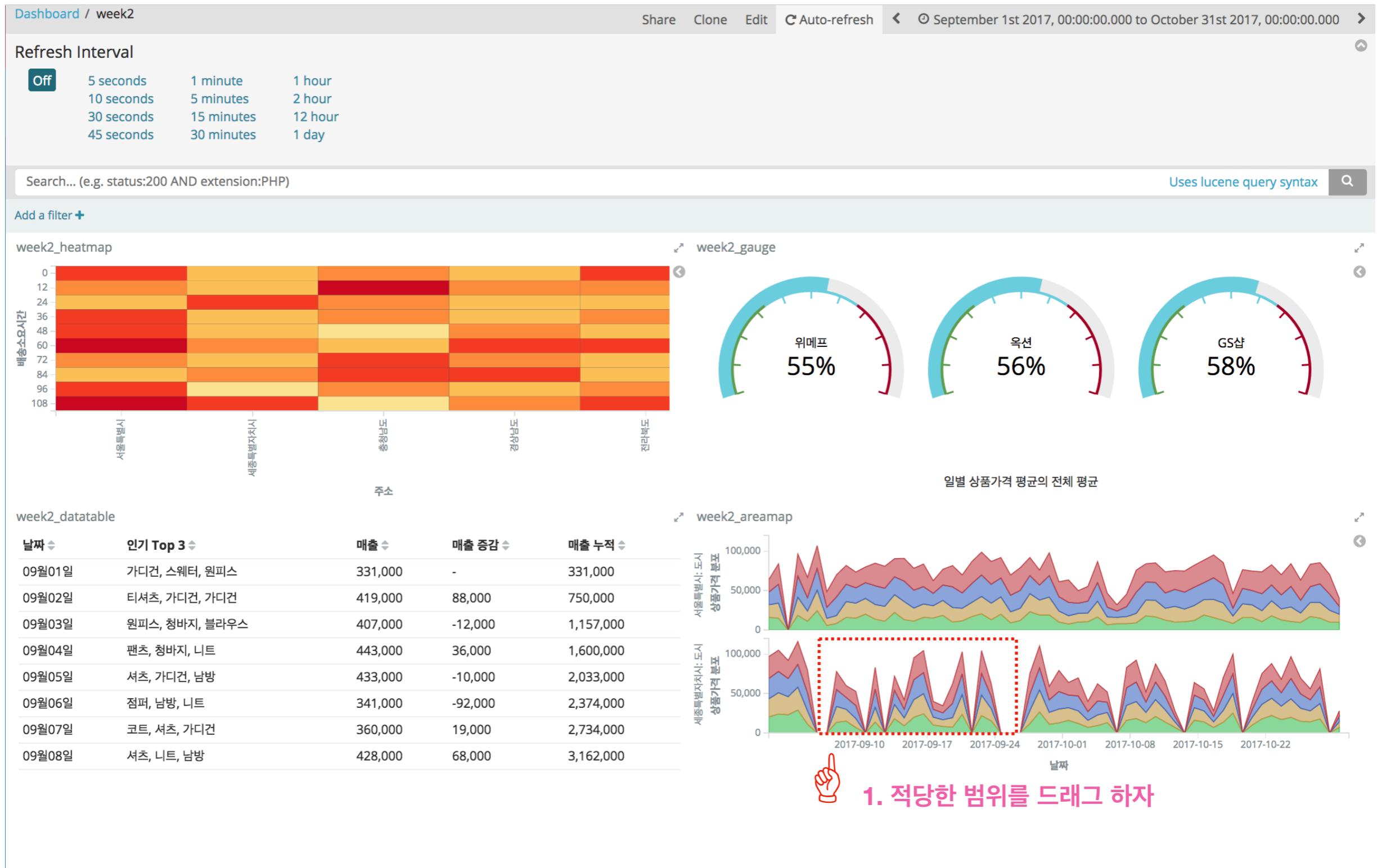
week2_datatable

week2_areamap

날짜	인기 Top 3	매출	매출 증감	매출 누적
09월01일	가디건, 스웨터, 원피스	331,000	-	331,000
09월02일	티셔츠, 가디건, 가디건	419,000	88,000	750,000
09월03일	원피스, 청바지, 블라우스	407,000	-12,000	1,157,000
09월04일	팬츠, 청바지, 니트	443,000	36,000	1,600,000
09월05일	셔츠, 가디건, 남방	433,000	-10,000	2,033,000
09월06일	점퍼, 남방, 니트	341,000	-92,000	2,374,000
09월07일	코트, 셔츠, 가디건	360,000	19,000	2,734,000
09월08일	셔츠, 니트, 남방	428,000	68,000	3,162,000

Kibana - dashboard

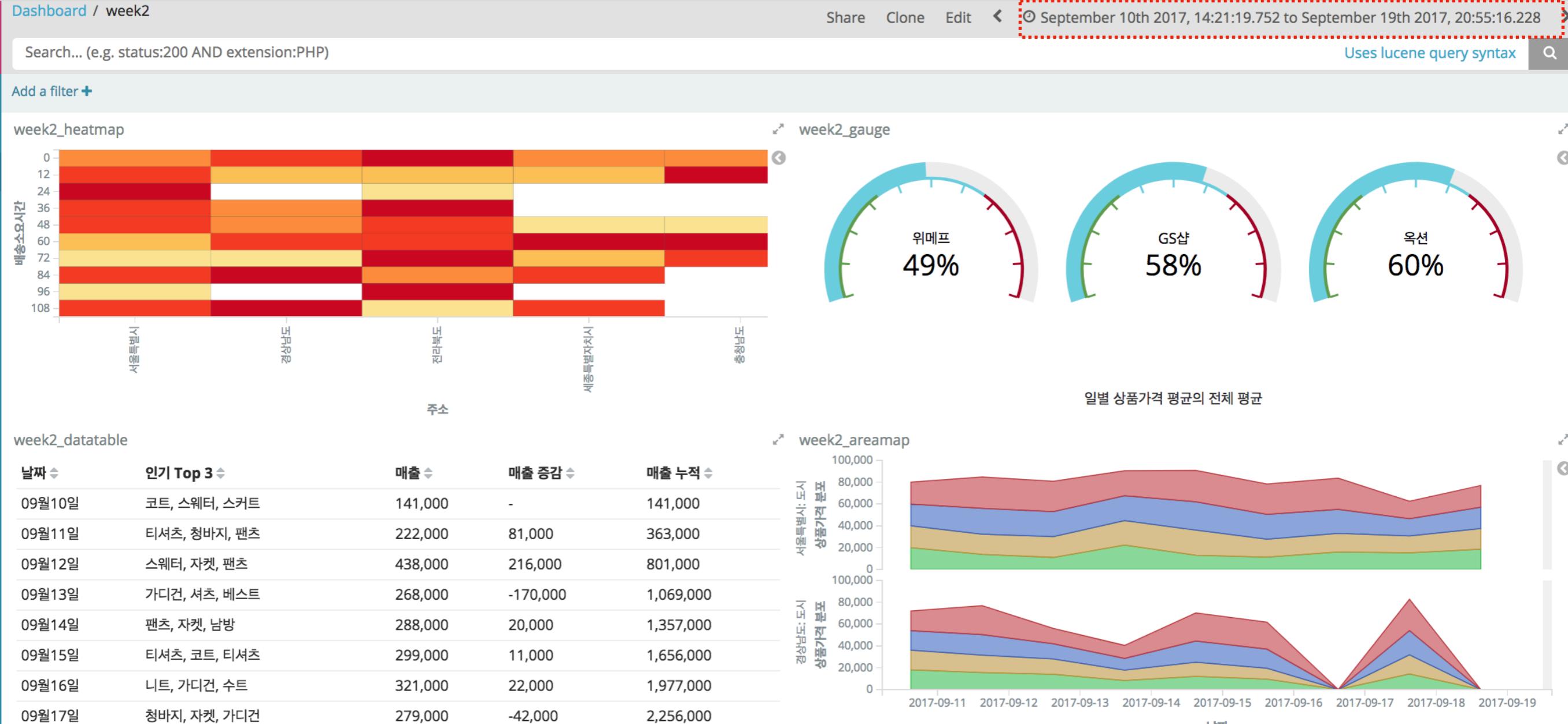
Interactive Dashboard 기능을 이용하자



Kibana - dashboard

선택한 조건에 맞게 dashboard에 filter가 적용되었다

2. 날짜 filter를 적용했기에 날짜가 변경되었다



Kibana - dashboard

dashboard를 공유하자

공유 전 변경 공유 후 변경

saved dashboard	반영 o	반영 o
Snapshot	반영 o	반영 x

Share 선택

Dashboard / week2

Share saved dashboard

You can share this URL with people to let them load the most recent saved version of this dashboard.

Embedded iframe

<iframe src="http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPl0SIAlpN750?embed=true&_g=(refreshInterval:(display:5000))&_a=(id:1)">

Add to your HTML source. Note that all clients must be able to access Kibana.

Link

[http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPl0SIAlpN750?_g=\(refreshInterval%3A\(display:5000\)\)&_a=\(id:1\)](http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPl0SIAlpN750?_g=(refreshInterval%3A(display:5000))&_a=(id:1))

Search... (e.g. status:200 AND extension:PHP)

Add a filter +

week2_heatmap

week2_gauge

일별 상품가격 평균의 전체 평균

Share Snapshot

Snapshot URLs encode the current state of the dashboard in the URL itself. Edits to the saved dashboard won't be visible via this URL.

Embedded iframe

<iframe src="http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPl0SIAlpN750?embed=true&_g=(refreshInterval:(display:5000))&_a=(id:1)">

Add to your HTML source. Note that all clients must be able to access Kibana.

Link

[http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPl0SIAlpN750?_g=\(refreshInterval:\(display:5000\)\)&_a=\(id:1\)](http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPl0SIAlpN750?_g=(refreshInterval:(display:5000))&_a=(id:1))

We recommend sharing shortened snapshot URLs for maximum compatibility. Internet Explorer has URL length restrictions, and some wiki and markup parsers don't do well with the full-length version of the snapshot URL, but the short URL should work great.

Uses lucene query syntax

Kibana - dashboard

dashboard를 공유하자

공유 전 변경 공유 후 변경

saved dashboard 반영 o 반영 o

snapshot 반영 o 반영 x

1. Share 선택

Dashboard / week2

Share saved dashboard

You can share this URL with people to let them load the most recent saved version of this dashboard.

Embedded iframe

<iframe src="http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPloSIAlpN750?embed=true&_s=1">

Add to your HTML source. Note that all clients must be able to access Kibana.

Link **dashboard 수정해도 url 변하지 않음**

http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPloSIAlpN750?_g=(refreshInterval%3A(display:(shortUrl:1)))&_s=1

Search... (e.g. status:200 AND extension:PHP)

Add a filter +

week2_heatmap

주소

서울특별시 경상남도 전라북도 세종특별자치시 충청남도 배송소요시간

week2_gauge

위메프 49%

GS샵 58%

옥션 60%

일별 상품가격 평균의 전체 평균

Share Clone Edit < ⏴ September 10th 2017, 14:21:19.752 to September 19th 2017, 20:55:16.228 >

Share Snapshot

Snapshot URLs encode the current state of the dashboard in the URL itself. Edits to the saved dashboard won't be visible via this URL.

Embedded iframe

<iframe src="http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPloSIAlpN750?embed=true&_s=1">

Add to your HTML source. Note that all clients must be able to access Kibana.

Link **dashboard 수정하면 url 변함**

http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPloSIAlpN750?_g=(refreshInterval:(display:(shortUrl:1)))&_s=1

We recommend sharing shortened snapshot URLs for maximum compatibility. Internet Explorer has URL length restrictions, and some wiki and markup parsers don't do well with the full-length version of the snapshot URL, but the short URL should work great.

Uses lucene query syntax

Kibana - dashboard

Embedded iFrame을 공유하자

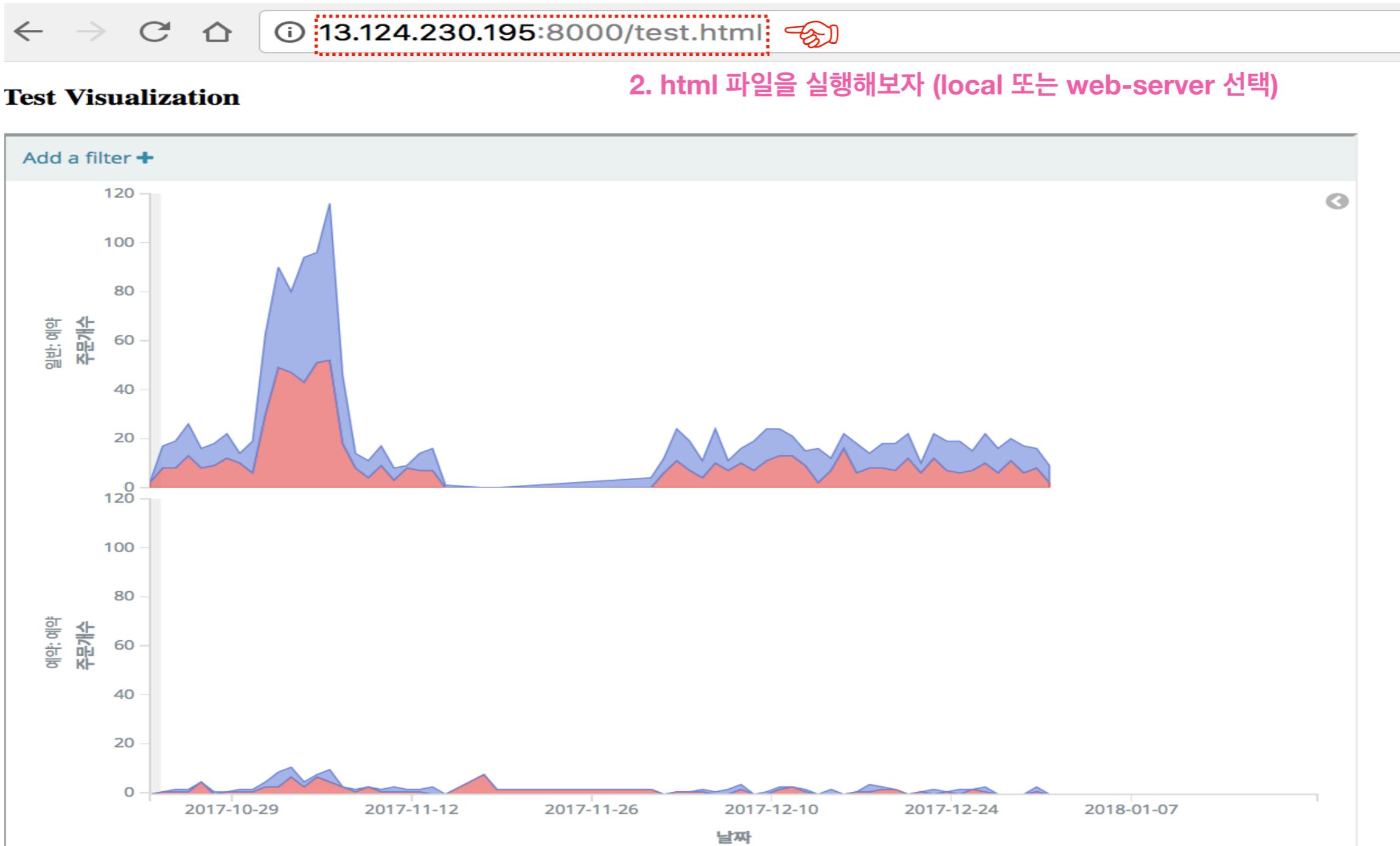
1. 아래 url를 활용한 간단한 html 파일을 만들자 🎉

The screenshot shows the Kibana dashboard interface with the following details:

- Left Sidebar:** Includes icons for Discover, Visualize, Dashboard, Timelion, Dev Tools, and Management.
- Header:** Shows "Dashboard / week2".
- Share Section:** Contains "Share saved dashboard" and "Share Snapshot".
- Embedded iframe:** A section with a red box highlighting the "Embedded iframe" URL. A hand icon points to this area. The URL is:
<iframe src="http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPloSIAlpN750?embed=true&_g=(refreshInterval:display:60s,refreshInterval:search:60s)">
- Link:** Provides a direct link to the dashboard:
http://13.124.230.195:5601/app/kibana#/dashboard/AWEYQjbSPloSIAlpN750?_g=(refreshInterval%3A(display%3A60s,refreshInterval%3Asearch%3A60s))
- Search Bar:** "Search... (e.g. status:200 AND extension:PHP)".
- Add a filter +**
- Visualizations:**
 - week2_heatmap:** A heatmap visualization showing data across various locations and time intervals.
 - week2_gauge:** Three gauge charts showing metrics: 위메프 (49%), GS샵 (58%), and 옥션 (60%).
- Bottom:** "일별 상품가격 평균의 전체 평균"

Kibana - dashboard

Embedded iFrame을 공유하자

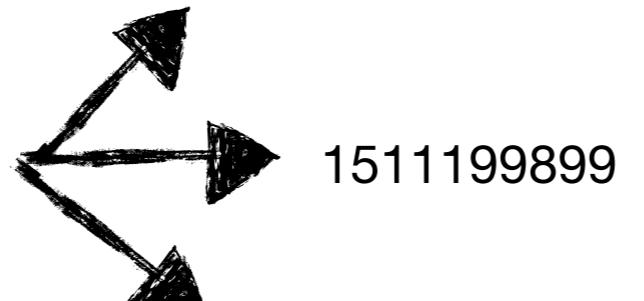


Managing Field 

Kibana - management

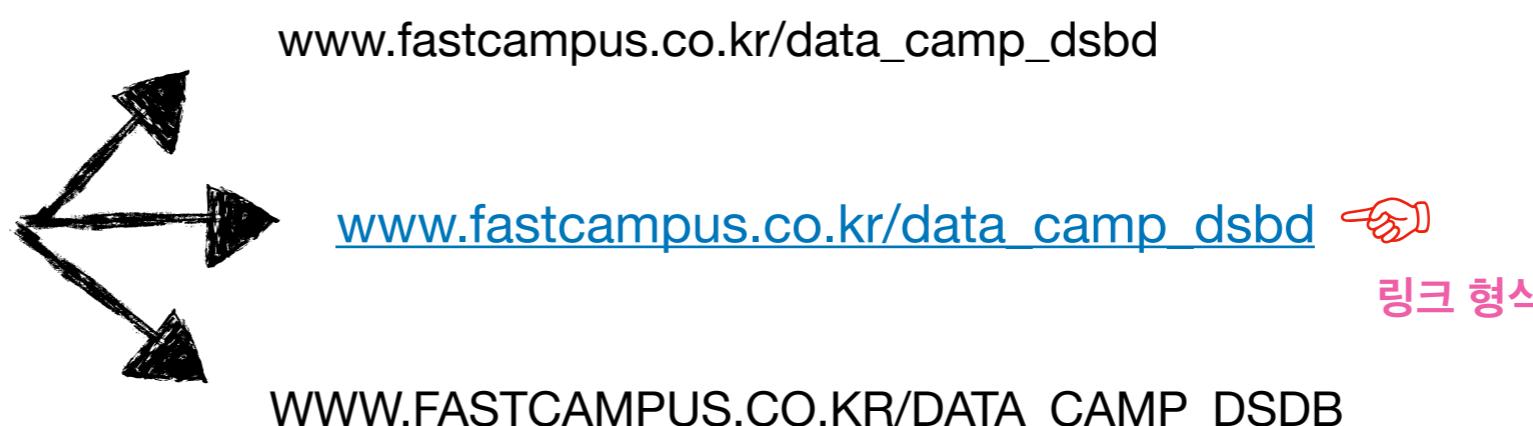
2017년 11월 20일 17시 44분

Date를 다르게 표현할 수 없나?



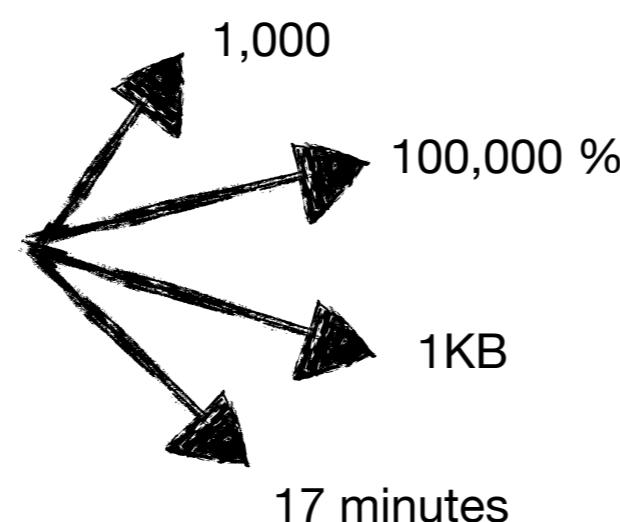
11/20/2017 5:44pm

String를 다르게 보여줄 수 없나?



링크 형식

Number를 다르게 보여줄 수 없나?



17 minutes

주의!!

Data Format이 변하는 것이지 **Data Type**이 변하는 것이 아니다.
그러므로 Elasticsearch에 저장된 데이터 자체는 변하지 않는다!

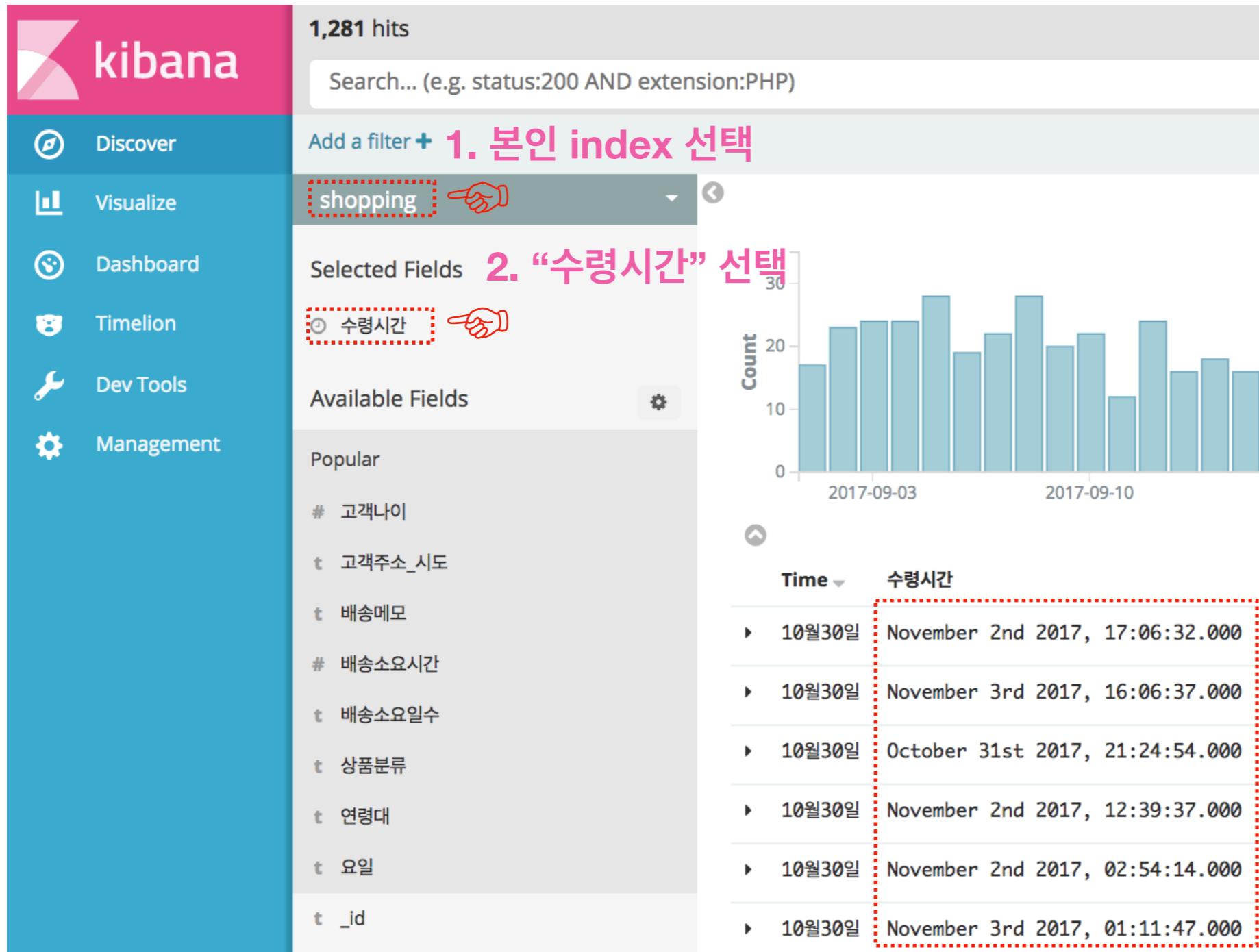
Date Type의 Format 전환 실습을 위해 Index를 등록하자 

Index : week2_{id}

Time Filter Field : 주문시간

Kibana - management

그리고 **Discover**에 가서 **Index** 선택 후 “수령시간”을 표시해보자



mission : Date Format 변경

Kibana - management

Management - Index Patterns - Index 선택하자

The screenshot shows the Kibana Management interface with the following steps highlighted:

- 1. Management 선택**: A hand icon points to the "Management" button in the sidebar.
- 2. Index Patterns 선택**: A hand icon points to the "Index Patterns" tab in the top navigation bar.
- 3. 본인 Index 선택**: A hand icon points to the "shopping" index pattern listed in the main content area.

The main content area displays the "shopping" index pattern with the following details:

- Time Filter field name:** 주문시간
- Fields:** 24 (number of fields)
- Scripted Fields:** 7
- Source Filters:** 0

name	type	format	searchable	aggregatable	excluded	controls
판매자평점	number		✓	✓		
주문시간	date	Date	✓	✓		
접수시간	date		✓	✓		
접수번호	string		✓	✓		
예약여부	string		✓	✓		
수령시간	date		✓	✓		

Date Format 변경하려는 Field의 Control 선택

The screenshot shows the Kibana Management interface with the 'Management / Kibana' header. The left sidebar has icons for Discover, Visualize, Dashboard, Timelion, Dev Tools, and Management, with 'Management' being the active tab. The main area shows the 'shopping' index pattern. A green banner at the top says 'Time Filter field name: 주문시간'. Below it, a message states: 'This page lists every field in the **shopping** index and the field's associated core type as recorded by Elasticsearch. While this list allows you to view the core type of each field, changing field types must be done using Elasticsearch's [Mapping API](#)'. There are three tabs: 'fields (24)', 'scripted fields (7)', and 'source filters (0)'. A search bar with 'Filter' and a dropdown 'All field types' are also present. The table lists fields with columns for name, type, format, searchable, aggregatable, excluded, and controls. The '수령시간' field is highlighted with a red dashed border and a red callout pointing to it with the Korean word '선택' (Select).

name	type	format	searchable	aggregatable	excluded	controls
판매자평점	number		✓	✓		
주문시간	date	Date	✓	✓		
접수시간	date		✓	✓		
접수번호	string		✓	✓		
예약여부	string		✓	✓		
수령시간	date		✓	✓		

Format을 변경하자

The screenshot shows the Kibana Management interface with the following details:

- Left Sidebar:** Discover, Visualize, Dashboard, Timelion, Dev Tools, Management.
- Top Bar:** Management / Kibana / Indices / shopping / Field
- Header:** Index Patterns (selected), Saved Objects, Advanced Settings
- Create Index Pattern:** + Create Index Pattern
- Index Pattern List:** ★ shopping
- Selected Field:** ★ shopping (수령시간)
- Type:** date (수정 안됨 - Not Editable)
- Format (Default: Date):** - default - (수정 가능 => Date 선택 - Editable, Set Date Selection)
- Popularity:** 1 (+, -)
- Buttons:** Update Field, Cancel

Annotations in pink text and arrows highlight the 'Type' and 'Format' fields, indicating they are not editable.

적절한 Format Pattern으로 수정하자

The screenshot shows the Kibana Management interface with the following details:

- Left Sidebar:** Discover, Visualize, Dashboard, Timelion, Dev Tools, Management (selected).
- Top Bar:** Management / Kibana / Indices / shopping / Field
- Index Patterns:** shopping (selected)
- Field Configuration:**
 - Name:** ★ shopping
 - Type:** date
 - Format (Default: Date):** moment.js format pattern (Default: "MMMM Do YYYY, HH:mm:ss.SSS")
Input: MMMM Do YYYY, HH:mm:ss.SSS
 - Samples:**

Input	Formatted
1516536304175	January 21st 2018, 21:05:04.175
1514732400000	January 1st 2018, 00:00:00.000
1546268399999	December 31st 2018, 23:59:59.999
 - Popularity:** 1 (with + and - buttons)
- Buttons at the bottom:** Update Field, Cancel

Annotations:

- A pink box highlights the "moment.js format pattern" input field with a red border and a pink hand icon pointing to it. To its right, the text "1. 아래와 같이 입력하자" (Input it like this) is written in pink.
- A pink box highlights the "1" popularity input field with a red border and a pink hand icon pointing to it. To its right, the text "MM월DD일" (MM month DD day) is written in pink.
- A pink box highlights the "Update Field" button with a red border and a pink hand icon pointing to it. To its left, the text "2. 선택하자" (Select it) is written in pink.

Kibana - management

Discover에 돌아가서 확인하자

The screenshot shows the Kibana Management interface. On the left, there's a sidebar with icons for Discover, Visualize, Dashboard, Timelion, Dev Tools, and Management. The Management icon is highlighted with a red box and a hand cursor. The main area is titled "Discover" and shows "1,281 hits". It has a search bar with placeholder text "(e.g. status:200 AND extension:PHP)". Below the search bar is a dropdown menu set to "shopping". A "Selected Fields" section contains a field labeled "수령시간" with a red dashed box and a hand cursor. An "Available Fields" section lists various fields like 고객나이, 고객주소_시도, 배송메모, etc. To the right, there's a bar chart titled "Count" showing data from 2017-09-03, and a table with columns "Time" and "수령시간". The "수령시간" column for the first row is also highlighted with a red dashed box and a hand cursor.

Discover

Visualize

Dashboard

Timelion

Dev Tools

Management

1,281 hits

Search... (e.g. status:200 AND extension:PHP)

Add a filter +

shopping

Selected Fields

수령시간

Available Fields

Popular

고객나이

t 고객주소_시도

t 배송메모

배송소요시간

t 배송소요일수

t 상품분류

t 연령대

t 요일

Count

Time

수령시간

Time	수령시간
10월30일	11월02일
10월30일	11월03일
10월30일	10월31일
10월30일	11월02일
10월30일	11월02일

Kibana - management

아래를 참고해서 “주문시간”을 Unix Timestamp (Millisecond)로 표시해보자 ✎ 🤴

날짜 단위	문법	예시	설명
Year	YYYY	2014	4자리 표시
Year	YY	14	2자리 표시
Month	M	1	1~2자리 표시
Month	MM	01	2자리 표시
Day	D	1	1~2자리 표시
Day	DD	01	2자리 표시
Day	Do	1st	며칠째인지 표시
Hour	H	1	1자리 표시 (24시)
Hour	HH	01	1~2자리 표시 (24시)
Hour	h	1	1자리 표시 (12시)
Hour	hh	01	1~2자리 표시 (12시)
a	h	am/pm	소문자 표시
A	hh	AM/PM	대문자 표시
Minute	m	1	1자리 표시
Minute	mm	01	1~2자리 표시
Second	s	1	1자리 표시
Second	ss	01	1~2자리 표시
Second	X	1410715640.579	Unix Timestamp 초
Millisecond	x	1410715640579	Unix Timestam 밀리초

String Type의 Format 전환 실습을 위해 Index를 등록하자 

Index : week2_url_{id}

Time Filter Field : 없음

Kibana - management

그리고 Discover에 가서 “full_url”과 “partial_url”을 표시해보자

3 hits

Search... (e.g. status:200 AND extension:PHP)

Discover Add a filter +

week2_url_higee 1. 본인 index 선택

Selected Fields

t full_url 2. “full_url”, “partial_url” 선택

t partial_url

full_url	partial_url
http://higee.io/221117841117	221117841117
http://higee.io/221189884818	221189884818
http://higee.io/221175500228	221175500228

mission : 두 Field 모두 링크로 변경

Kibana - management

(앞에서 했던 것처럼) Date Format 변경하려는 Field의 Control 선택

The screenshot shows the Kibana Management interface for the 'Index Patterns' tab. The left sidebar has 'Management' selected. The main area displays the fields for the index pattern 'week2_url_higee'. The 'full_url' field is highlighted with a red dashed box. A red hand icon is pointing at the 'controls' column for the 'full_url' row. The table columns are: name, type, format, searchable, aggregatable, excluded, and controls.

name	type	format	searchable	aggregatable	excluded	controls
_id	string		✓			
_index	string		✓	✓		
_score	number					
_source	_source					
_type	string		✓	✓		
full_url	string		✓	✓		
partial_url	string		✓	✓		

[“<http://higee.io/221117841117>”](http://higee.io/221117841117) 와 같이 full_url 값을 가진 field를 url화 하는 작업

Kibana - management

Format을 변경하자

The screenshot shows the Kibana Management interface with the following details:

- Management / Kibana / Indices / week2_url_higee / Field**
- Index Patterns Saved Objects Advanced Settings**
- + Create Index Pattern**
- week2_url_higee**
- full_url**
- Type**: string (highlighted with a red dashed box; a pink hand icon and the text "수정 안됨" are placed to its right)
- Format (Default: String)**: - default - (highlighted with a red dashed box; a pink hand icon and the text "수정 가능 => URL 선택" are placed to its right)
- Popularity**: 0 (+ -)
- Buttons**: Update Field, Cancel

Url Templates에 적절한 값을 입력하자

week2_url_higee
full_url

Type
string

Format (Default: String) ⚠ Warning
Url

Type
Link

Url Template Url Template Help
 1. 다음과 같이 입력하자

Label Template Label Template Help
{{rawValue}}

Samples

Input	Formatted
john	john
/some pathname/asset.png	/some pathname/asset.png
1234	1234

Popularity
0 + -

Update Field Cancel

2. Update Field 선택

Discover에 돌아가서 확인하자

The screenshot shows the Kibana interface with the 'Discover' tab selected. The search bar contains the query 'week2_url_higee'. The results table has two columns: 'full_url' and 'partial_url'. Three results are listed, each with a red box around the 'full_url' value:

full_url	partial_url
http://higee.io/221117841117	221117841117
http://higee.io/221189884818	221189884818
http://higee.io/221175500228	221175500228

A red hand icon with the text '클릭해보자' (Click here) points to the first URL in the list.

Kibana - management 📄

그렇다면 이번에는 “partial_url” field를 url data format으로 변형해보자.
단, url_template 부분에는 다음과 같이 입력해야 한다.

<http://higee.io/{{value}}>

Discover에 돌아가서 확인하자

The screenshot shows the Kibana interface with the sidebar on the left and the Discover view on the right.

Sidebar:

- Discover (selected)
- Visualize
- Dashboard
- Timelion
- Dev Tools

Discover View:

3 hits

Search... (e.g. status:200 AND extension:PHP)

Add a filter +

Selected Fields: week2_url_higee

full_url	partial_url
http://higee.io/221117841117	221117841117
http://higee.io/221189884818	221189884818
http://higee.io/221175500228	221175500228

A red dashed box highlights the 'partial_url' column, and a pink hand icon with the text '클릭해보자' (Click here) points to the first item in that column.

Number Type의 Format 전환 실습을 위해 Index를 등록하자 

Index : week2_number_{id}

Time Filter Field : 없음

Kibana - management

그리고 Discover에 가서 “percent”, “bytes”, “duration”을 표시해보자

The screenshot shows the Kibana interface with the 'Discover' tab selected. The search bar contains the query 'week2_number_higee'. The results table has three columns: 'byte', 'duration', and 'percent'. A red box highlights the 'Selected Fields' section, which includes '# byte', '# duration', and '# percent'. Red numbers 1 and 2 with arrows point to the 'Selected Fields' section and the 'Selected Fields' list respectively. The table data is as follows:

byte	duration	percent
133	33,535	0.3
135,351	335,315	0.7
8,191	10,005	0.5
13	3,535	0.1
13	335	0.2
71	9,315	0.5
13,535,139	33	0.4
1,351	35,315	0.5
791	9,005	0.5
81,910	135,305	0.5
133,333	3,335	0.2



Kibana - management

duration Field Format을 변경하자

week2_number_higee

duration

Type

number

Format (Default: Number)

Duration  선택

Input Format

Milliseconds  원본 시간 단위

Output Format

Seconds  변환하려는 시간 단위

Decimal Places

2

 소수점 자리수

Samples

Input

Formatted

-123

-0.12

1

0.00

12

0.01

123

0.12

658

0.66

1988

1.99

3857

3.86

123292

123.29

923528271

923528.27

Popularity

1

+

-

 Update Field Cancel

 선택

Kibana - management

Discover에 돌아가서 확인하자

The screenshot shows the Kibana Discover interface with the following details:

- Selected Fields:** week2_number_higee
- Available Fields:** _id, _index, _score, _type
- Selected Fields (List):** byte, duration, percent
- Results:** 11 hits
- Search Bar:** Search... (e.g. status:200 AND extension:PHP)
- Table Headers:** byte, duration, percent
- Data Rows:** (Listed below)

byte	duration	percent
133B	33.53	0.3
132.18KB	335.31	0.7
8KB	10.01	0.5
13B	3.54	0.1
13B	0.34	0.2
71B	9.31	0.5
12.91MB	0.03	0.4
1.32KB	35.31	0.5
791B	9.01	0.5
79.99KB	135.31	0.5
130.21KB	3.33	0.2



seconds 단위로 변환됐다!
access log 수집할 때 유용하다

Kibana - management

percent Field Format을 변경하자

week2_number_higee

percent

Type

number

Format (Default: Number)

Percentage



선택

Numerical.js format pattern (Default: "0,0.[000]%)")

0,0.[000]%



입력

Samples

Input	Formatted
0.1	10%
0.99999	99.999%
1	100%
100	10,000%
1000	100,000%

Popularity

1



Update Field

Cancel



선택

Discover에 돌아가서 확인하자

The screenshot shows the Kibana Discover interface with the following details:

- Left Sidebar:** Includes icons for Discover, Visualize, Dashboard, Timelion, Dev Tools, and Management. The Management icon is highlighted.
- Header:** Shows "11 hits" and a search bar with placeholder text "Search... (e.g. status:200 AND extension:PHP)".
- Selected Fields:** A dropdown menu set to "week2_number_higee". Below it, under "Selected Fields", are "# byte", "# duration", and "# percent".
- Available Fields:** A list including "_id", "_index", "# _score", and "_type".
- Table:** Displays 11 rows of data with columns: byte, duration, and percent. The "percent" column is highlighted with a red dotted border.

byte	duration	percent
133B	33.53	30%
132.18KB	335.31	71.1%
8KB	10.01	51.1%
13B	3.54	10%
13B	0.34	20%
71B	9.31	51.1%
12.91MB	0.03	37.11%
1.32KB	35.31	51.1%
791B	9.01	51.1%
79.99KB	135.31	51.1%
130.21KB	3.33	23%

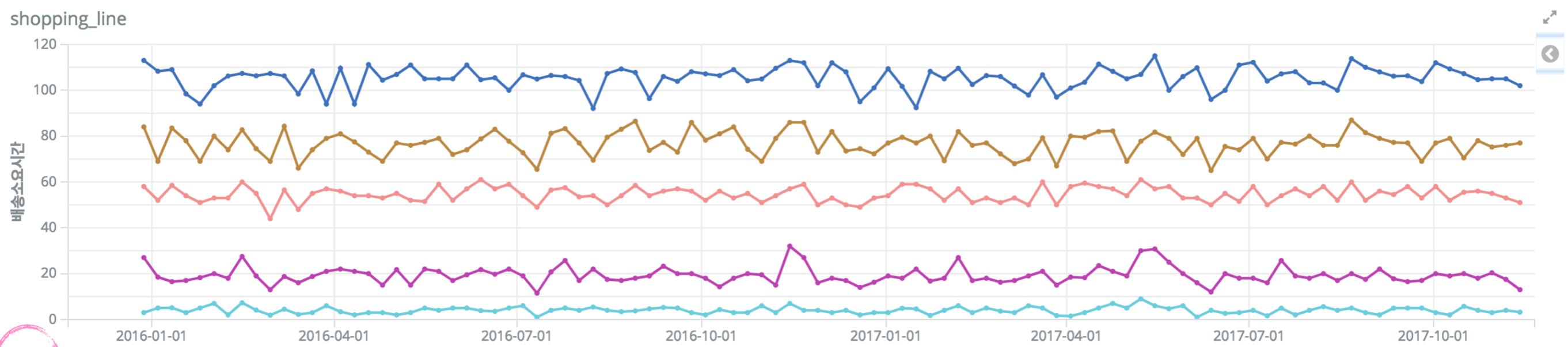
% 형태로 표시됐다



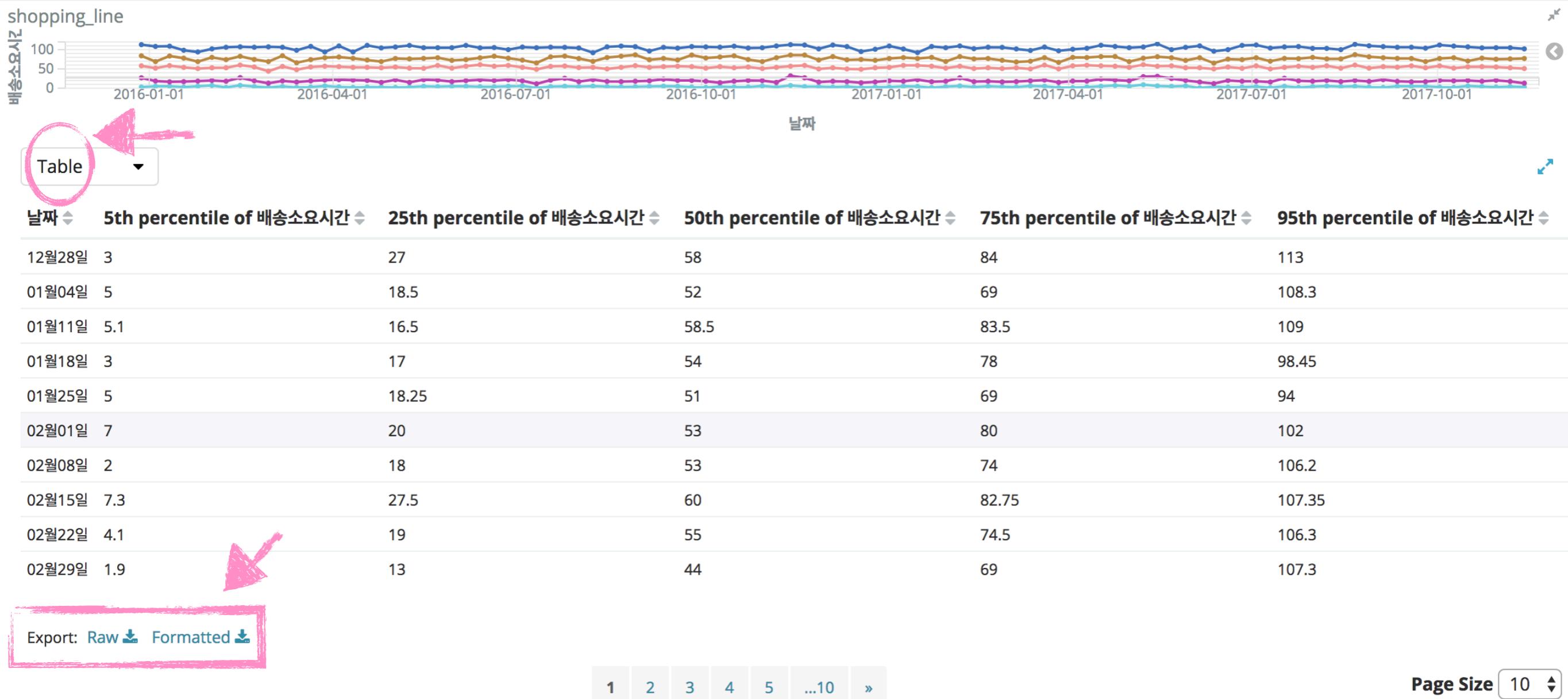
Visualization Spy

Visualization 결과를 csv로 다운 받을 수 없나?

Visualization Spy



Visualization Spy



Visualization Spy

Raw

	A	B	C
1	날짜	5th percentile of 배송소요시간	25th percentile of 배송소요시간
2	1.45126E+12	3	27
3	1.45187E+12	5	18.5
4	1.45247E+12	5.1	16.5
5	1.45308E+12	3	17
6	1.45368E+12	5	18.25
7	1.45428E+12	7	20
8	1.45489E+12	2	18
9	1.45549E+12	7.3	27.5
10	1.4561E+12	4.1	19
11	1.4567E+12	1.9	13
12	1.45731E+12	4.5	18.75
13	1.45791E+12	2.2	16
14	1.45852E+12	3	18.75
15	1.45912E+12	6	21
16	1.45973E+12	3.4	22

Elasticsearch에 저장된 Data Format

Format

	A	B	C
1	날짜	5th percentile of 배송소요시간	25th percentile of 배송소요시간
2	12월28일	3	27
3	01월04일	5	18.5
4	01월11일	5.1	16.5
5	01월18일	3	17
6	01월25일	5	18.25
7	02월01일	7	20
8	02월08일	2	18
9	02월15일	7.3	27.5
10	02월22일	4.1	19
11	02월29일	1.9	13
12	03월07일	4.5	18.75
13	03월14일	2.2	16
14	03월21일	3	18.75
15	03월28일	6	21
16	04월04일	3.4	22

Kibana에서 변경된 Data Format

JSON Input 

JSON Input

missing data가 있을 때 특정한 값으로 대체할 수 없나?

이동평균을 구할 때 window size를 변경할 수 없나?

Term Aggregation 시 5개 이하 Bucket은 제외할 수 없나?

⋮

JSON Input

missing data가 있을 때 특정한 값으로 대체할 수 없나?



이동평균을 구할 때 window size를 변경할 수 없나?

Term Aggregation 시 5개 이하 Bucket은 제외할 수 없나?

다음과 같은 **Visualization**을 만들자

데이터

- Index : shopping
- Time Range : 2017년 11월 18일 03시~ 2017년 11월 18일 23시

문제

- “주문시간”을 기준으로 시간당
- “상품가격”의 합을 표시하자

사용한 Visualization

- Data Table

사용한 Aggregation

- Metrics : Sum Aggregation
- Buckets : Date Histogram

사용 필드

- 상품가격
- 주문시간

~~힌트는 다음 페이지~~

JSON Input

Missing Value 확인

Kibana Visualize / New Visualization (unsaved) November 18th 2017, 03:00:00.000 to November 18th 2017, 23:00:00.000

Search... (e.g. status:200 AND extension:PHP)

Add a filter +

shopping

Data Options

metrics

Metric Sum of 상품가격

Add metrics

buckets

Split Rows

Aggregation

Date Histogram

Field

주문시간

Interval

Hourly

Custom Label

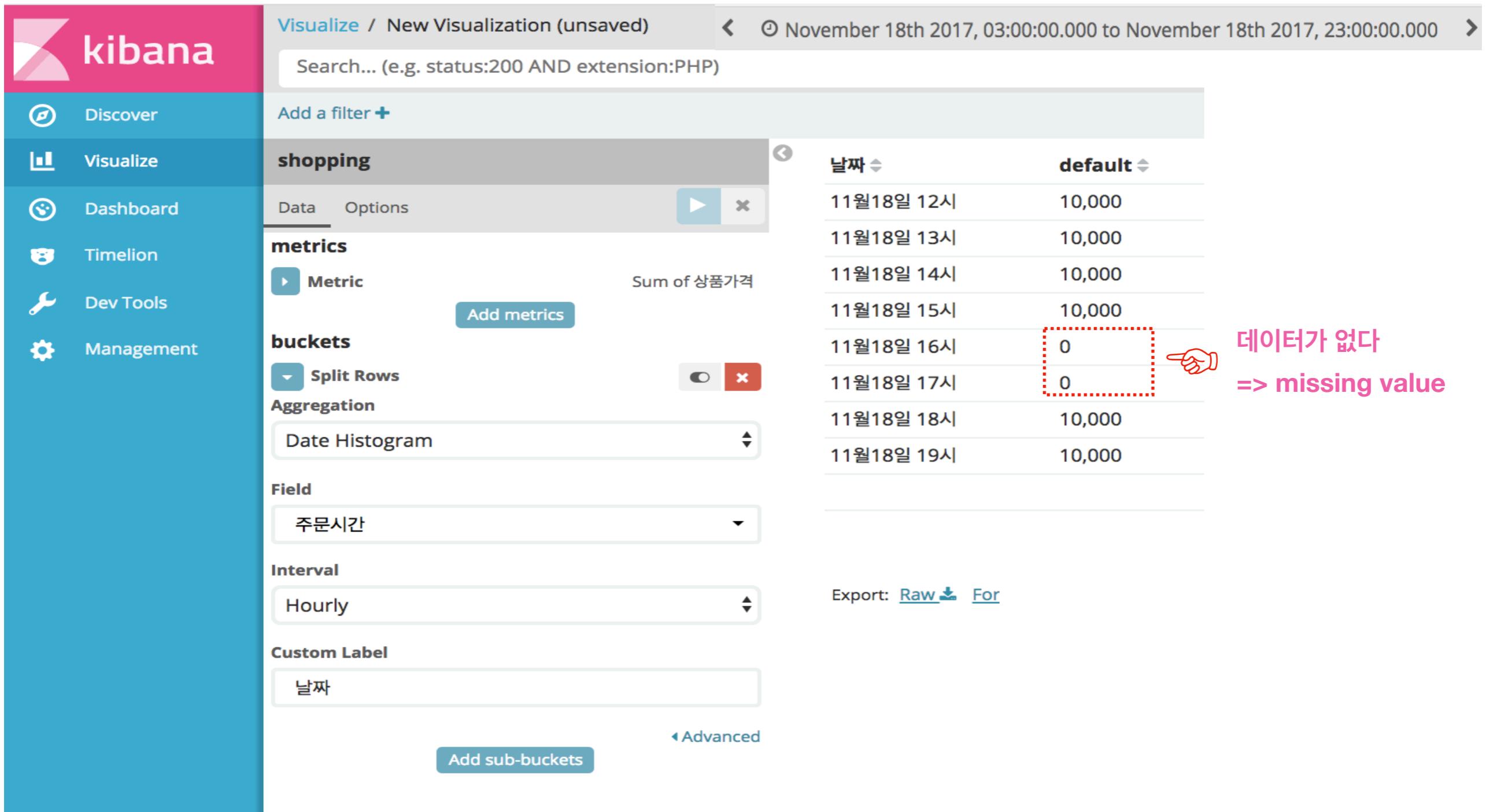
날짜

Add sub-buckets

Advanced

날짜	default
11월18일 12시	10,000
11월18일 13시	10,000
11월18일 14시	10,000
11월18일 15시	10,000
11월18일 16시	0
11월18일 17시	0
11월18일 18시	10,000
11월18일 19시	10,000

데이터가 없다
=> missing value



JSON Input

Missing Value 처리

Visualize / New Visualization (unsaved) ⏪ ⓘ November 18th 2017, 03:00:00.000 to November 18th 2017, 23:00:00.000 ⏩

Search... (e.g. status:200 AND extension:PHP)

Discover Visualize Dashboard Timelion Dev Tools Management

shopping

3. 선택

Add a filter +

metrics

Metric Metric

Aggregation Sum

Field 상품가격

Custom Label missing value 처리

JSON Input ⓘ {"missing":100}

Advanced

날짜 default missing value 처리

날짜	default	missing value 처리
11월18일 12시	10,000	10,000
11월18일 13시	10,000	10,000
11월18일 14시	10,000	10,000
11월18일 15시	10,000	10,000
11월18일 16시	0	100
11월18일 17시	0	100
11월18일 18시	10,000	10,000
11월18일 19시	10,000	10,000

Export: Raw

1. 선택

2. 다음과 같이 입력

{“missing” : 100}

JSON Input

무슨 일이 벌어진걸까? Visualization Spy의 Request를 보자

Visualize / New Visualization (unsaved) Save Share Refresh ⏪ ⏩ November 18th 2017, 03:00:00.000 to November 18th 2017, 23:00:00.000 > Search... (e.g. status:200 AND extension:PHP) Uses lucene query syntax

Add a filter +

shopping

Data Options

metrics

Metric Sum of 상품가격 Metric Metric

Aggregation

Sum

Field

상품가격

Custom Label

missing value 처리

Advanced

JSON Input

{"missing":100}

Add metrics

buckets

Split Rows 주문시간 per hour Add sub-buckets

날짜	default	missing value 처리
11월18일 12시	10,000	10,000
11월18일 13시	10,000	10,000
11월18일 14시	10,000	10,000
11월18일 15시	10,000	10,000
11월18일 16시	0	100
11월18일 17시	0	100
11월18일 18시	10,000	10,000
11월18일 19시	10,000	10,000

Export: Raw Formatted Request

2. Request 선택

Elasticsearch request body

```
_source: {  
    "excludes": []  
},  
"aggs": {  
    "2": {  
        "date_histogram": {  
            "field": "주문시간",  
            "interval": "1h",  
            "time_zone": "Asia/Tokyo",  
            "min_doc_count": 1  
        },  
        "aggs": {  
            "1": {  
                "sum": {  
                    "field": "상품가격"  
                }  
            }  
        }  
    }  
},  
"3": {  
    "sum": {  
        "field": "상품가격",  
        "missing": 100  
    }  
}
```

1. 선택

3. Sum Aggregation Parameter에 반영된 거 확인
정확히 무슨 역할을 하는지는 공식 문서 확인

JSON Input

Q. 어떤 Aggregation에 어떤 parameter가 쓰이는지 외워야 되나?

A.

- 1) 사용하려는 Aggregation 검색  (ex: Moving Average)
- 2) Parameter List 확인
- 3) Kibana JSON Input에 비슷하게 넣어보기
- 4) Visualization Spy - Request에서 확인하고 디버깅하기

JSON Input

JSON Input 넣고 확인하기

Elasticsearch

```
POST /_search
{
  "size": 0,
  "aggs": {
    "my_date_histo": {
      "date_histogram": {
        "field": "date",
        "interval": "1M"
      },
      "aggs": {
        "the_sum": {
          "sum": { "field": "price" }
        },
        "the_movavg": {
          "moving_avg": {
            "buckets_path": "the_sum",
            "window": 30,
            "model": "simple"
          }
        }
      }
    }
  }
}
```

Kibana JSON

shopping

Data Options

metrics

Metric

Aggregation

Moving Avg

Custom Metric

Count

Advanced

Custom Label

JSON Input

```
{"windo":10}
```

buckets

Split Rows

주문시간 per 3 hours

Add metrics

Add sub-buckets

시간
11월01일 00시
11월01일 03시
11월01일 06시
11월01일 09시
11월01일 12시
11월01일 15시
11월01일 18시
11월01일 21시
11월02일 00시
11월02일 03시

Export: Raw Formatted

```
  "size": 0,
  "_source": {
    "excludes": []
  },
  "aggs": {
    "2": {
      "date_histogram": {
        "field": "주문시간",
        "interval": "3h",
        "time_zone": "UTC",
        "min_doc_count": 0
      }
    },
    "aggs": {
      "4": {
        "moving_avg": {
          "buckets_path": "_count",
          "windo": 10
        }
      }
    }
  }
}
```

JSON Input

missing data가 있을 때 특정한 값으로 대체할 수 없나?

이동평균을 구할 때 window size를 변경할 수 없나?



Term Aggregation 시 5개 이하 Bucket은 제외할 수 없나?

다음과 같은 Visualization을 만들자

데이터

- Index : shopping
- Time Range : 2017년 10월 1일~ 2017년 10월 10일

문제

- “주문시간”을 기준으로 일별로(daily)
- “상품가격”의 합과,
- “상품가격”의 합의 이동평균을
- 각각 표시하자

사용한 Visualization

- Data Table

사용한 Aggregation

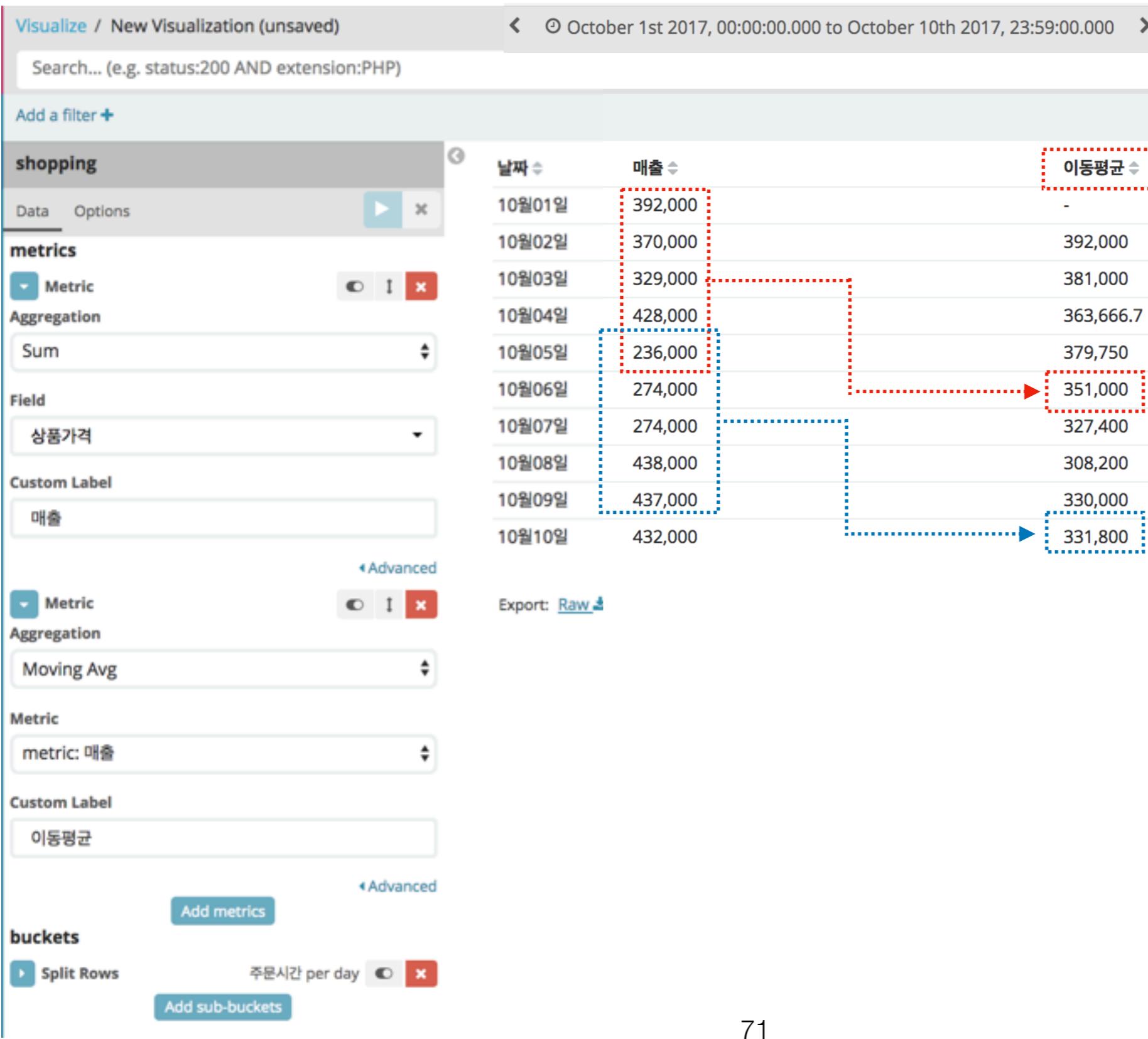
- Metrics : Sum Aggregation, Moving Average Aggregation
- Buckets : Date Histogram

사용 필드

- 상품가격
- 주문시간

JSON Input

Moving Average 확인



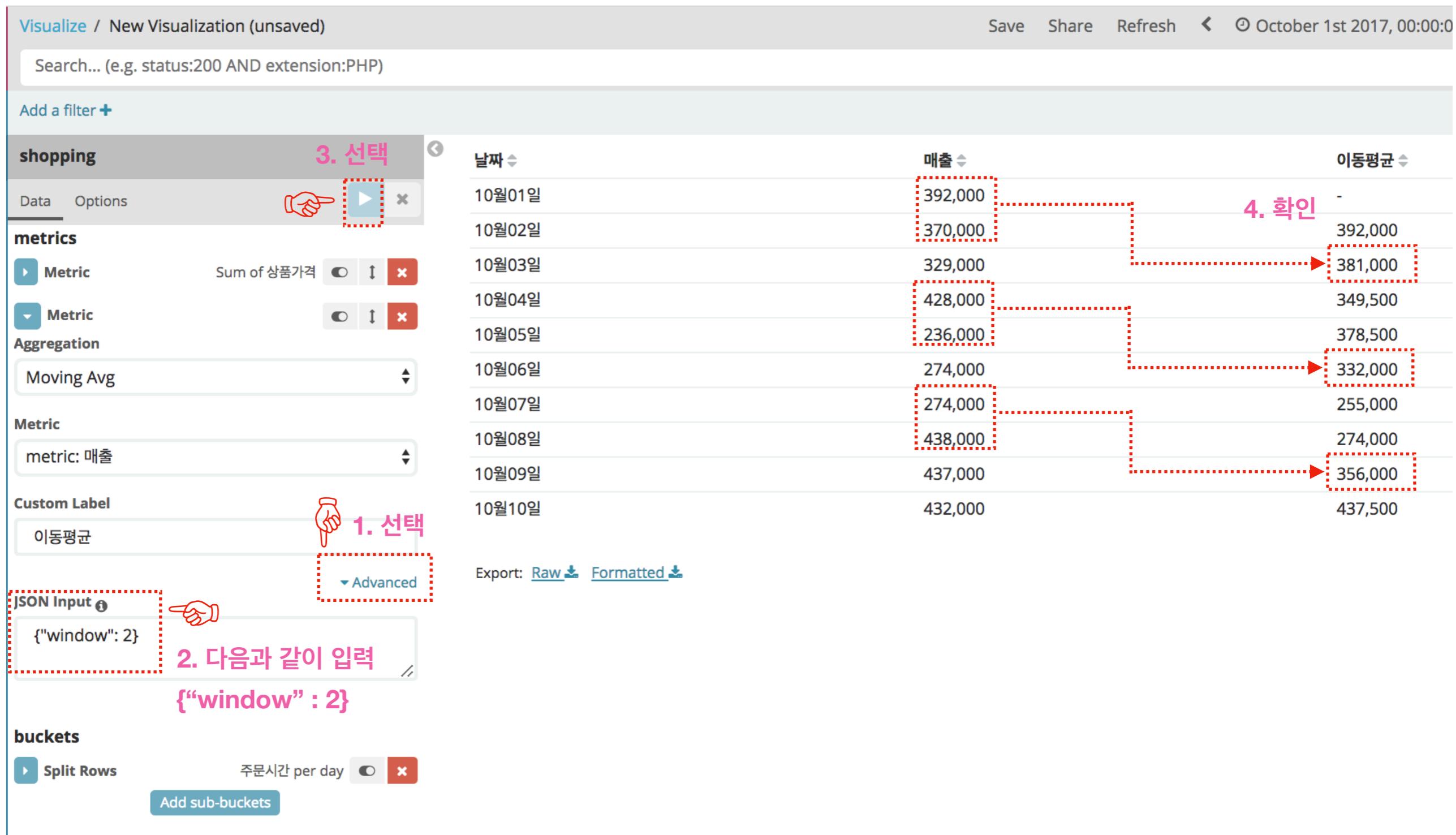
☞

이동평균은 어떻게 구할까? 👑

=> default window size : 5

JSON Input

Moving Average window size 변경



JSON Input

missing data가 있을 때 특정한 값으로 대체할 수 없나?

이동평균을 구할 때 window size를 변경할 수 없나?

Term Aggregation 시 5개 이하 Bucket은 제외할 수 없나?



다음과 같은 Visualization을 만들자

데이터

- Index : shopping
- Time Range : 2017년 10월 1일~ 2017년 10월 10일

문제

- 개수가 가장 많은 “상품분류”를 10개 선별하고
- 그 개수를 바탕으로 Tag Size를 표시하자

사용한 Visualization

- Tag Cloud

사용한 Aggregation

- Metrics : Count Aggregation
- Buckets : Terms Aggregation

사용 필드

- 상품분류

JSON Input

Visualization 확인

The screenshot shows the Kibana visualization interface. At the top, there are buttons for Save, Share, Refresh, and a date range from October 1st, 2017, to October 10th, 2017. Below that is a search bar with placeholder text "Search... (e.g. status:200 AND extension:PHP)". A "Add a filter +" button is also present.

The main area is titled "shopping". It contains sections for "metrics" and "buckets".

- metrics:** Set to "Tag Size". Under "Aggregation", "Count" is selected. There is a "Custom Label" input field and an "Advanced" button.
- buckets:** Set to "Tags". Under "Aggregation", "Terms" is selected. The "Field" dropdown is set to "상품분류" (Product Category). "Order By" is set to "metric: Count" with "Order" as "Descending" and "Size" as "10". There is also a "Custom Label" input field and an "Advanced" button.

On the right side, there is a large red dashed box containing Korean text:

티셔츠 자켓
남방 가디건 셔츠
니트 코트 청바지
스웨터 스커트

A pink note next to the box says:

만약 이 중에서 Count가 너무 작은 상품이 있다면?
즉, Count의 최소값을 설정하고 싶다면?

JSON Input

Data 확인

Visualize / New Visualization (unsaved) Save Share Refresh ⌘ October 1st 2017, 00:00:00.000 to Oct
Search... (e.g. status:200 AND extension:PHP)

Add a filter +

shopping

Data Options  

metrics

 Tag Size

Aggregation

Count

Custom Label



buckets

 Tags

Aggregation

Terms

Field

상품분류

Order By

metric: Count

Order **Size**

Descending 10

Custom Label



Table

상품분류: Descending

가디건 20
코트 20
자켓 19
청바지 18
티셔츠 18
니트 16
셔츠 15
스웨터 14
남방 13
스커트 12

Export: Raw  Formatted 

1. 선택 

2. 확인 

3. 최소 16은 필요하다고 의사결정 



티셔츠 자켓
남방 가디건 셔츠
니트 코트 청바지
스웨터 스커트

JSON Input

Bucket 별 최소값 설정

Visualize / New Visualization (unsaved) Save Share Refresh ⌘ October 1st 2017, 00:00:00.000 to October 10th 2017, 23:59:00.000 > Search... (e.g. status:200 AND extension:PHP) Uses lucene query syntax

Add a filter +

shopping

Data Options

Aggregation

Count

Custom Label

buckets

Tags

Aggregation

Terms

Field

상품분류

Order By

metric: Count

Order Size

Descending 10

Custom Label

Exclude

Include

Advanced

Table

상품분류: Descending

가디건
코트
자켓
청바지
티셔츠
니트

Count

20
20
19
18
18
16

1. 선택

2. 다음처럼 입력

{“min_doc_count” : 16}

3. 확인

JSON Input ⓘ

{“min_doc_count” : 16}

Export: Raw Formatted

Page Size 10

Scripted Field  

Scripted Field

String Concatenation 연산은 안되나?

Field 간 연산은 안되나?

Date Field에서 연/월/일/시/분/초 등 접근 안되나?

기존 Field에 조건을 적용해서 새로운 Field를 만들 수 없나?

Scripted Field - String Concat

String Concatenation 연산은 안되나?



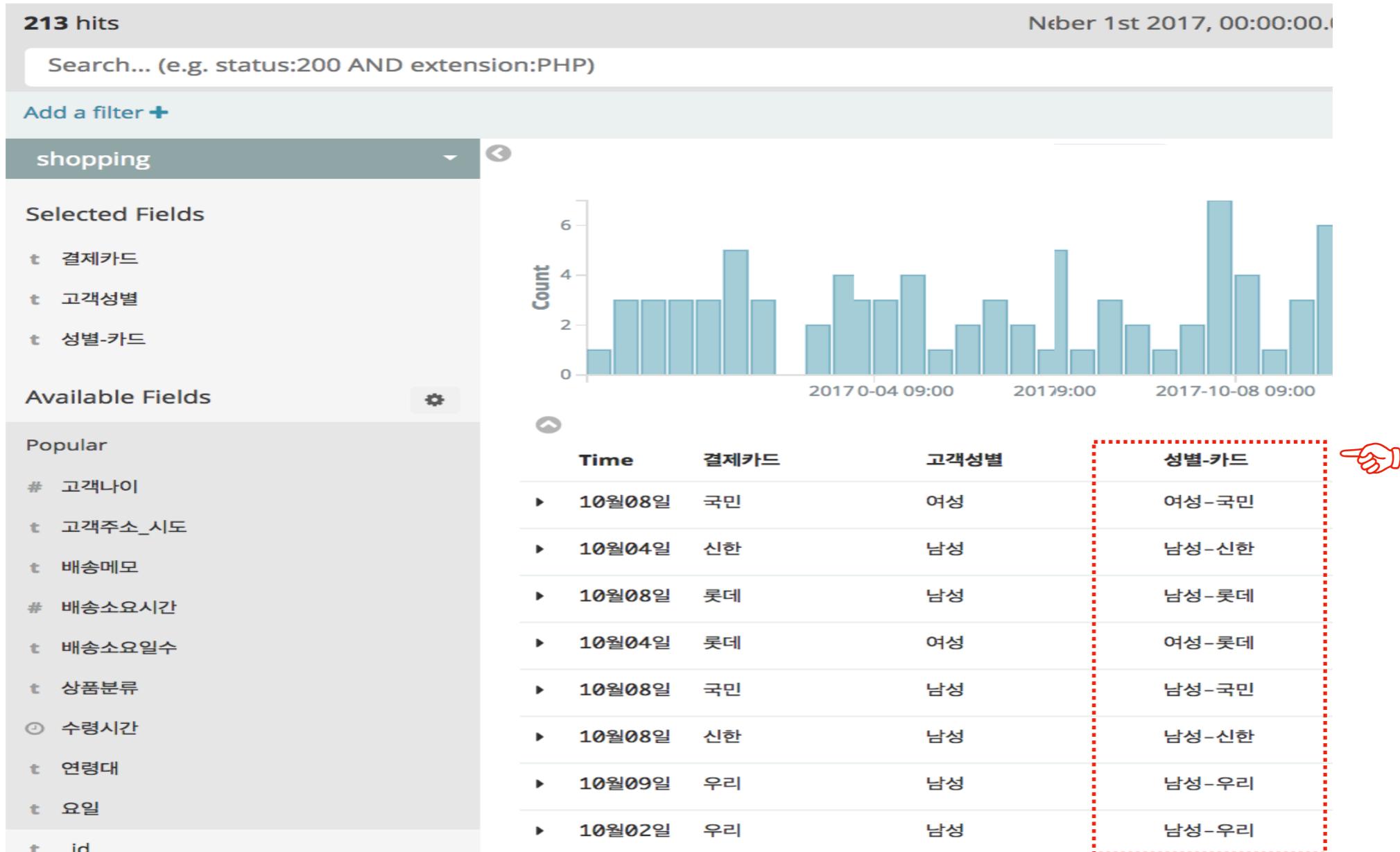
Field 간 연산은 안되나?

Date Field에서 연/월/일/시/분/초 등 접근 안되나?

기존 Field에 조건을 적용해서 새로운 Field를 만들 수 없나?

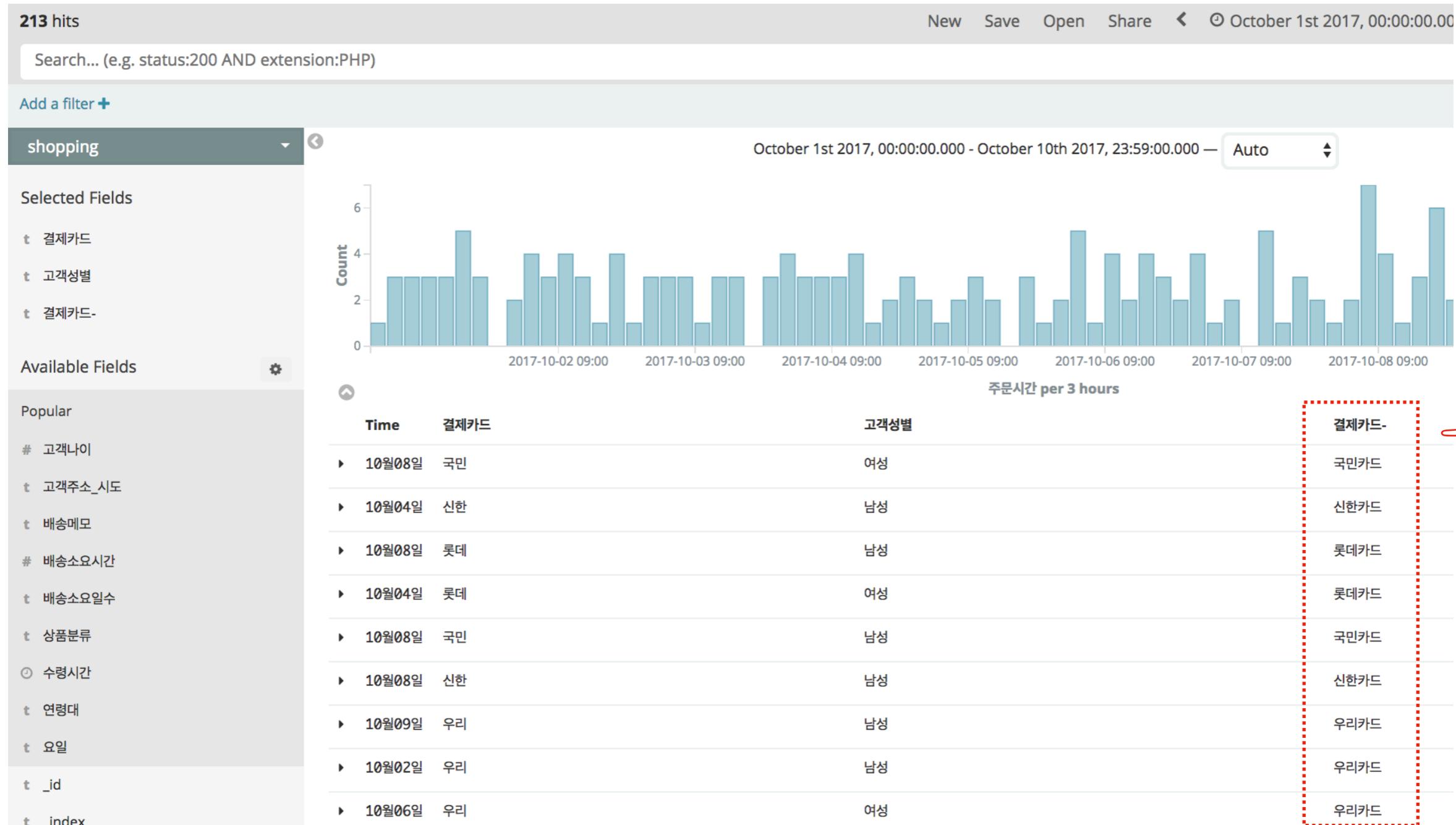
Scripted Field - String Concat

특정한 두 개 혹은 그 이상의 Field를 합쳐서 하나의 Field를 만들고 싶으면 어떻게 해야할까?



Scripted Field - String Concat

혹은 특정 Field에 글자를 덧붙이고 싶다면?



Scripted Field - String Concat

scripted field를 추가하자

The screenshot shows the Kibana Management interface for the 'shopping' index pattern. A red box labeled '1. 선택' highlights the 'Management' tab in the sidebar. Another red box labeled '2. 선택' highlights the 'Index Patterns' tab in the top navigation bar. A red box labeled '3. week2_{id} 선택' highlights the 'week2_{id}' field in the list of fields. A red box labeled '4. 선택' highlights the 'scripted fields (7)' tab. A red box labeled '5. 선택' highlights the '+ Add Scripted Field' button. Hand icons with arrows point from each numbered label to its corresponding highlighted element.

1. 선택

2. 선택

3. week2_{id} 선택

4. 선택

5. 선택

format controls

Page Size 25

Scroll to top

Collapse

Scripted Field - String Concat

script를 작성하자

Management / Kibana / Indices / shopping

Index Patterns Saved Objects Advanced Settings

+ Create Index Pattern

★ shopping

★ shopping Create Scripted Field

Name 1. Field 이름 임의로 생성

Language

Type 2. String 선택

Format (Default: String) 3. String 선택

Transform

Samples

Input	Formatted
A Quick Brown Fox.	A Quick Brown Fox.
STAY CALM!	STAY CALM!
com.organizations.project.ClassName	com.organizations.project.ClassName
hostname.net	hostname.net
SGVsbG8gd29ybGQ=	SGVsbG8gd29ybGQ=

Popularity + -

Script 4. 다음과 같이 script 작성

doc['고객성별'].value + '-' + doc['결제카드'].value

Scripted Field - String Concat

Discover에 가서 확인하자

213 hits November 1st 2017, 00:00:00..

Search... (e.g. status:200 AND extension:PHP)

Add a filter +

shopping

week2_{id} 선택

Selected Fields

- t 결제카드
- t 고객성별
- t 성별-카드

Available Fields

- # 고객나이
- t 고객주소_시도
- t 배송메모
- # 배송소요시간
- t 배송소요일수
- t 상품분류
- ⌚ 수령시간
- t 연령대
- t 요일
- t _id

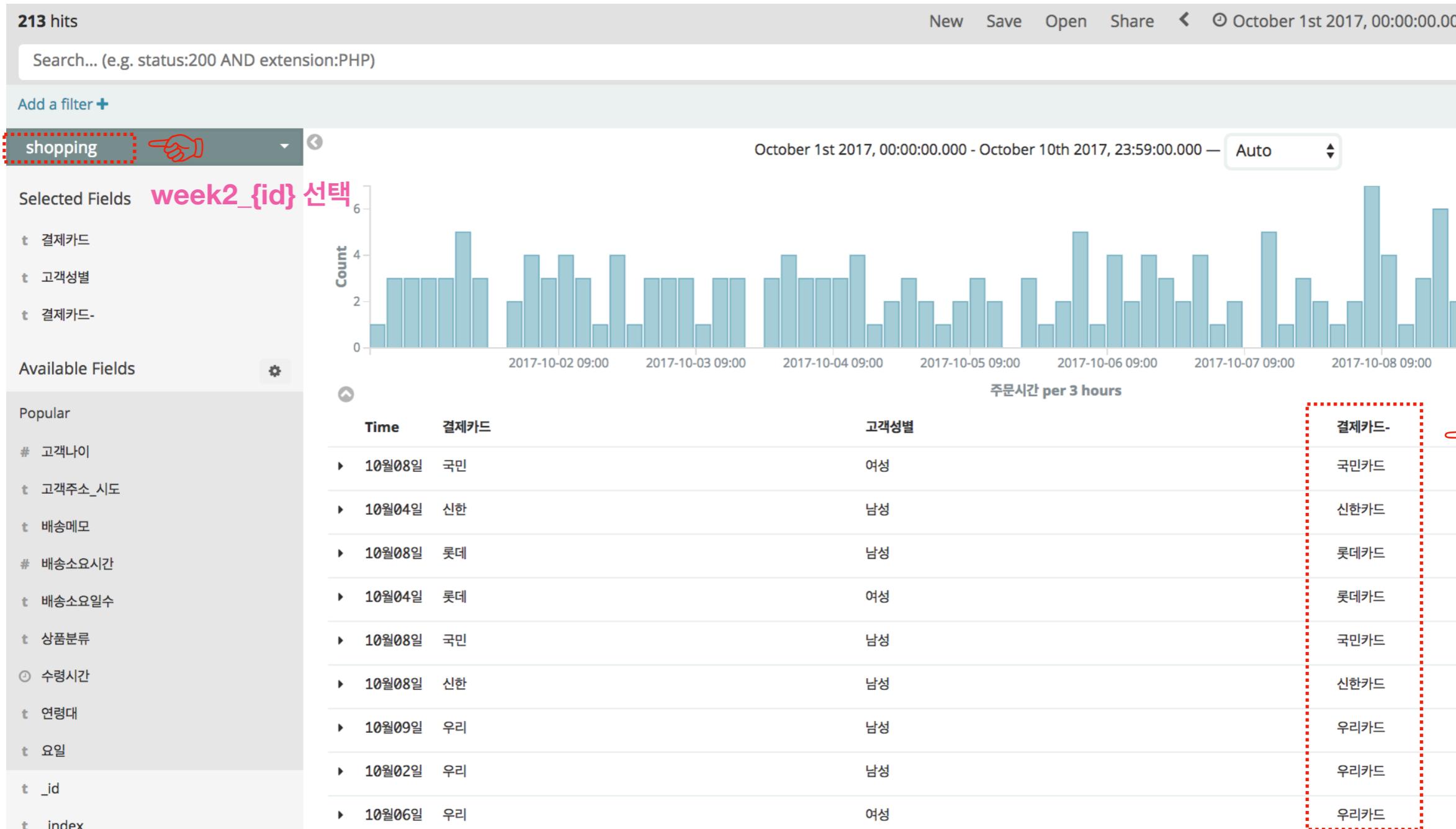
Count

Time

Time	결제카드	고객성별	성별-카드
10월08일	국민	여성	여성-국민
10월04일	신한	남성	남성-신한
10월08일	롯데	남성	남성-롯데
10월04일	롯데	여성	여성-롯데
10월08일	국민	남성	남성-국민
10월08일	신한	남성	남성-신한
10월09일	우리	남성	남성-우리
10월02일	우리	남성	남성-우리

Scripted Field - String Concat

이번에는 아래와 같은 Field를 생성해보자 ✎



Scripted Field - arithmetic operation

String Concatenation 연산은 안되나?

Field 간 연산은 안되나?

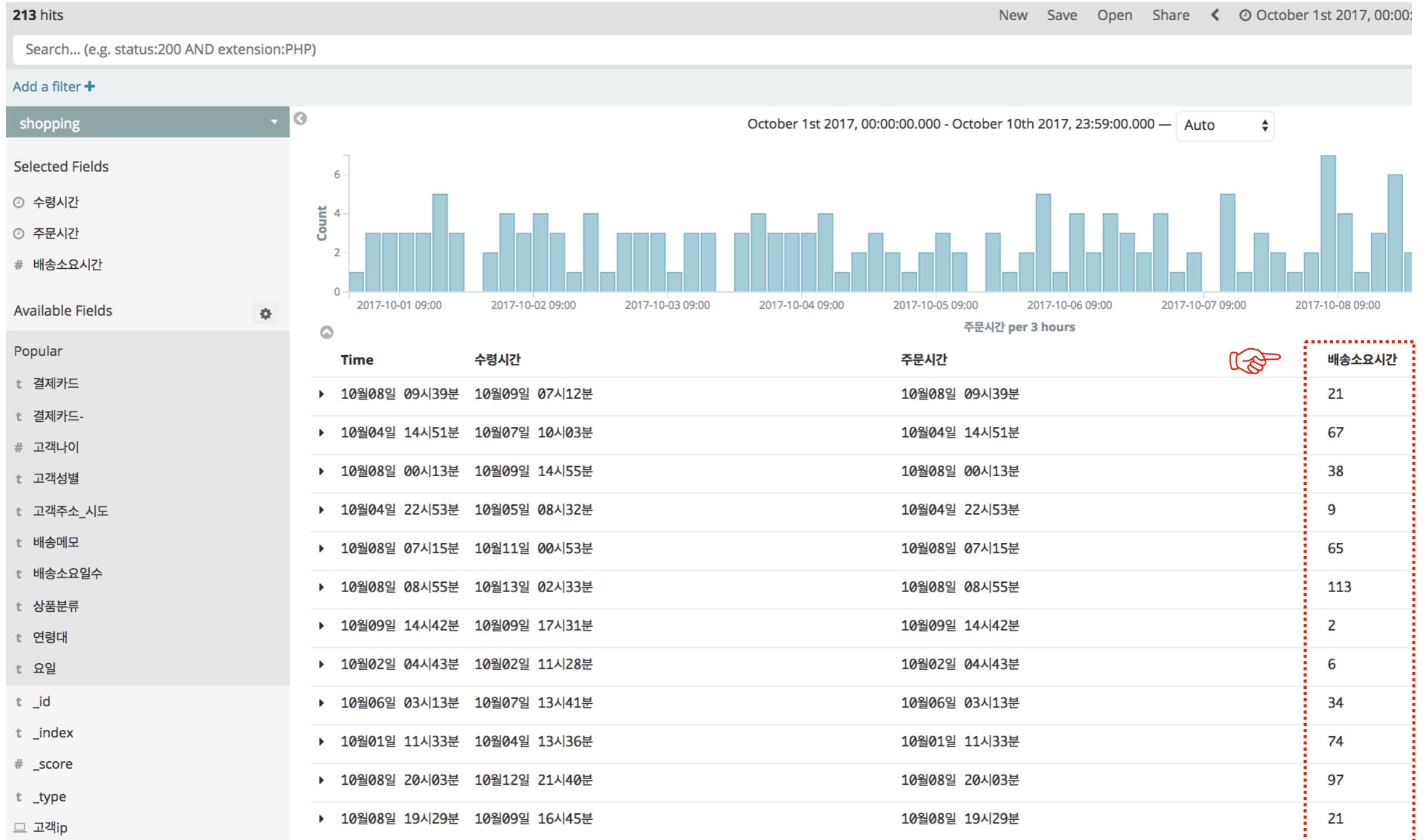


Date Field에서 연/월/일/시/분/초 등 접근 안되나?

기존 Field에 조건을 적용해서 새로운 Field를 만들 수 없나?

Scripted Field - arithmetic operation

2개의 Date Field의 차이를 통해 특정한 event의 소요시간을 구하고 싶다면?



Scripted Field - arithmetic operation

scripted field를 추가하자

The screenshot shows the Kibana Management interface for the 'shopping' index pattern. A red hand icon points to the 'Management' tab in the sidebar, labeled '1. 선택'. Another red hand icon points to the 'Index Patterns' tab in the top navigation bar, labeled '2. 선택'. A third red hand icon points to the 'shopping' index pattern in the list, labeled '3. week2_{id} 선택'. A fourth red hand icon points to the 'scripted fields (7)' tab in the top navigation bar, labeled '4. 선택'. A fifth red hand icon points to the '+ Add Scripted Field' button, labeled '5. 선택'. The page lists 24 fields, 7 of which are scripted fields. The right side of the screen shows the 'Scripted fields' section with a table for managing them.

1. 선택

2. 선택

3. week2_{id} 선택

4. 선택

5. 선택

format controls

Page Size 25

Scroll to top

Collapse

Scripted Field - arithmetic operation

script를 작성하자

Management / Kibana / Indices / shopping

Index Patterns Saved Objects Advanced Settings

+ Create Index Pattern

★ shopping

★ shopping
Create Scripted Field

Name

배송소요시간-



1. Field 이름 임의로 생성

Language

painless

Type

number



2. Number 선택

Format (Default: Number)

Number



3. Number 선택

Numerical.js format pattern (Default: "0,0.[0]")

0,0.[0]



4. 적당한 Format 입력

Samples

Input

10000

Formatted

10,000

12.345678

12.3

-1

-1

-999

-999

0.52

0.5

Popularity

0



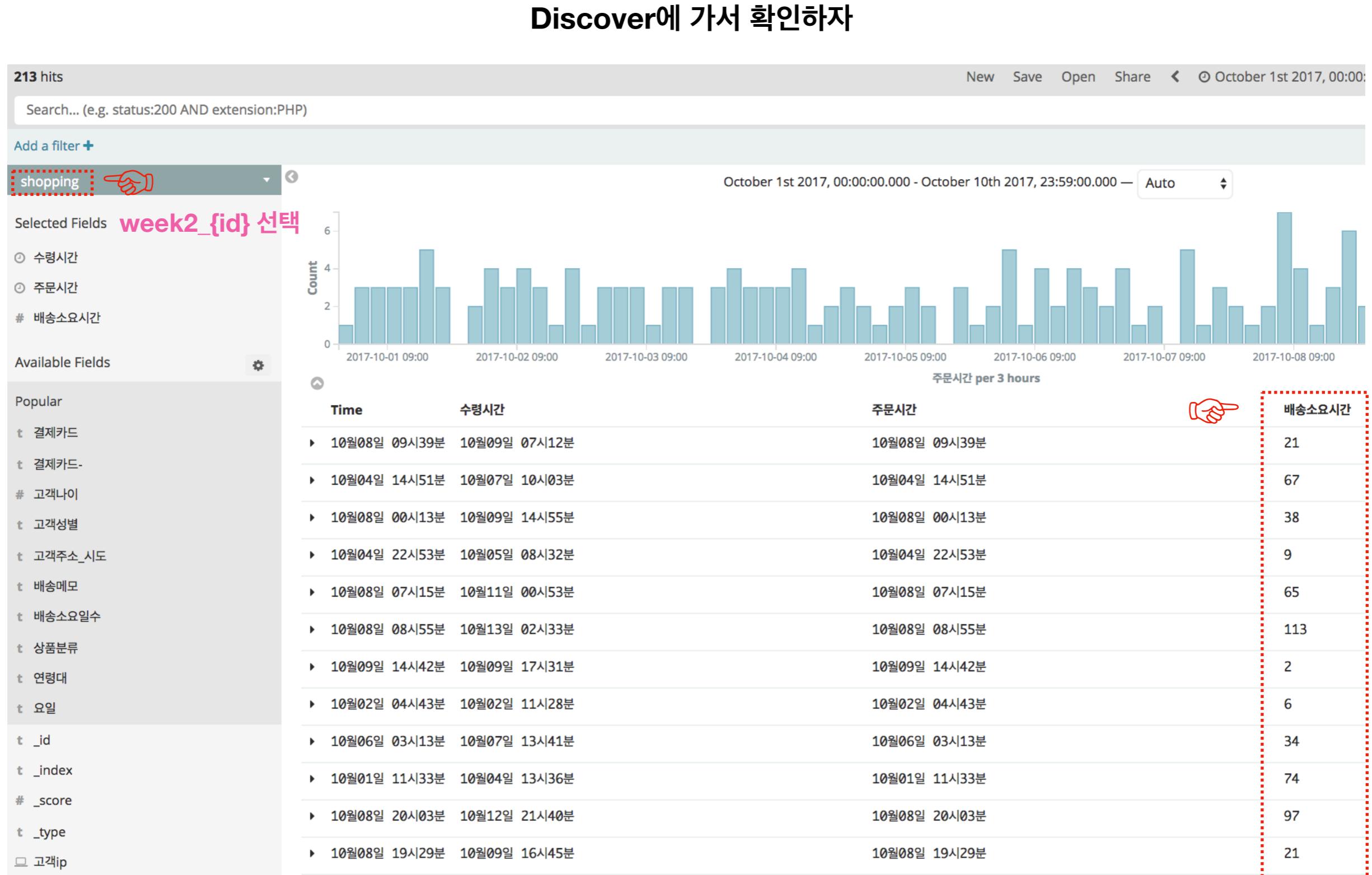
Script

(doc['수령시간'].value-doc['주문시간'].value)/1000/60/60

5. 다음과 같이 script 작성

(doc['수령시간'].value-doc['주문시간'].value)/1000/60/60

Scripted Field - arithmetic operation



Scripted Field - Date

String Concatenation 연산은 안되나?

Field 간 연산은 안되나?

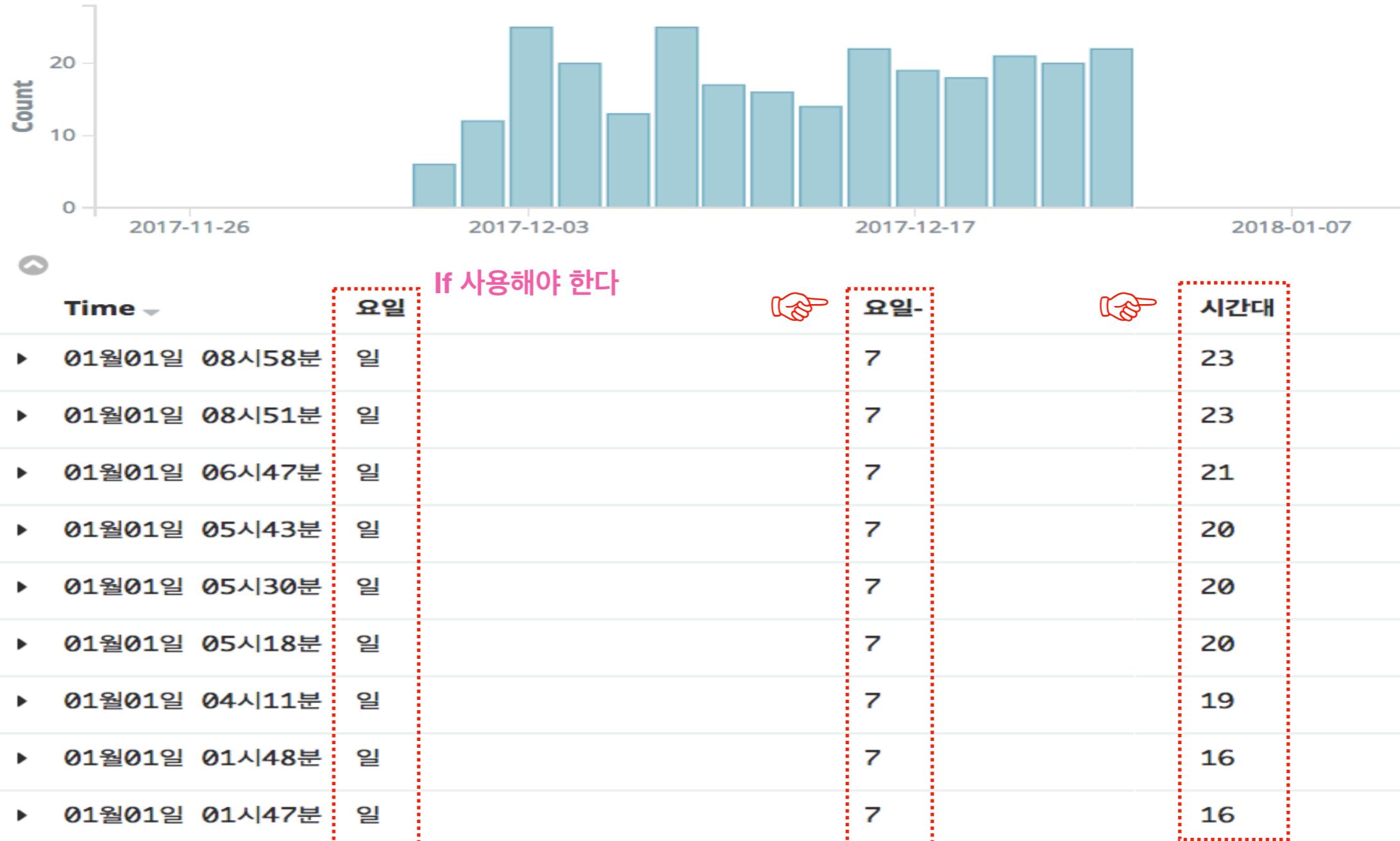
Date Field에서 연/월/일/시/분/초 등 접근 안되나?



기존 Field에 조건을 적용해서 새로운 Field를 만들 수 없나?

Scripted Field - Date

Date Field에서 요일, 시간대 등을 뽑고 싶으면?



Scripted Field - Date

아래를 참고해서 “주문시간”의 요일과 시간대를 추출하자 ↗

Expression	Description
doc['field_name'].date.centuryOfEra	Century (1-2920000)
doc['field_name'].date.dayOfMonth	Day (1-31), e.g. 1 for the first of the month.
doc['field_name'].date.dayOfWeek	Day of the week (1-7), e.g. 1 for Monday.
doc['field_name'].date.dayOfYear	Day of the year, e.g. 1 for January 1.
doc['field_name'].date.era	Era: 0 for BC, 1 for AD.
doc['field_name'].date.hourOfDay	Hour (0-23).
doc['field_name'].date.millisOfDay	Milliseconds within the day (0-86399999).
doc['field_name'].date.millisOfSecond	Milliseconds within the second (0-999).
doc['field_name'].date.minuteOfDay	Minute within the day (0-1439).
doc['field_name'].date.minuteOfHour	Minute within the hour (0-59).
doc['field_name'].date.monthOfYear	Month within the year (1-12), e.g. 1 for January.
doc['field_name'].date.secondOfDay	Second within the day (0-86399).
doc['field_name'].date.secondOfMinute	Second within the minute (0-59).
doc['field_name'].date.year	Year (-292000000 - 292000000).
doc['field_name'].date.yearOfCentury	Year within the century (1-100).
doc['field_name'].date.yearOfEra	Year within the era (1-292000000).

Scripted Field - If Condition

String Concatenation 연산은 안되나?

Field 간 연산은 안되나?

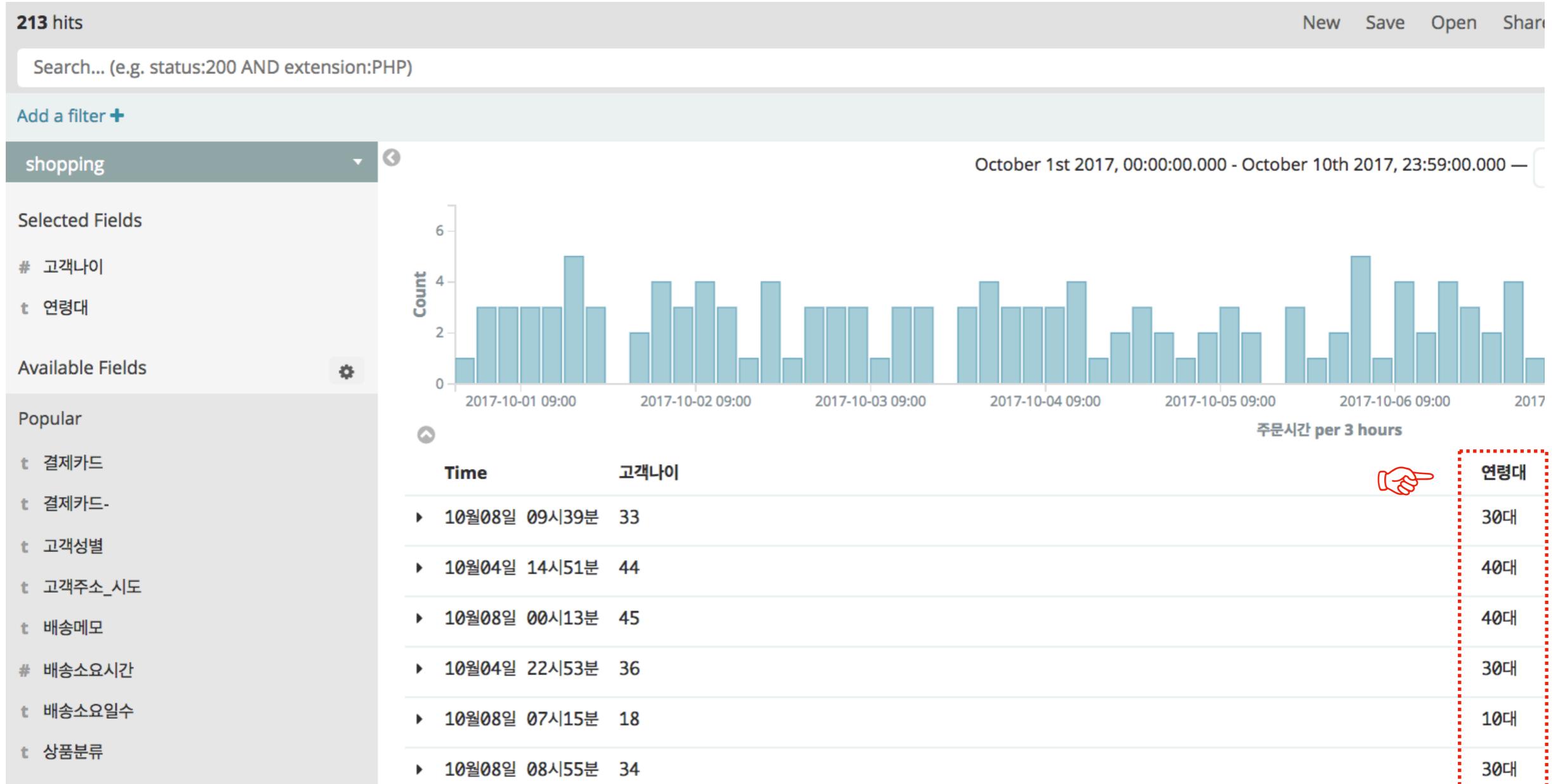
Date Field에서 연/월/일/시/분/초 등 접근 안되나?

기존 Field에 조건을 적용해서 새로운 Field를 만들 수 없나?



Scripted Field - If Condition

특정 조건에 따라 Value를 설정하고 싶다면?



Scripted Field - If Condition

script를 작성하자 

★ shopping

Create Scripted Field

Name

연령대- 

1. Field 이름 임의로 생성

Language

painless

Type

string 

2. String 선택

Format (Default: String)

String 

3. String 선택

Transform

Popularity

0  

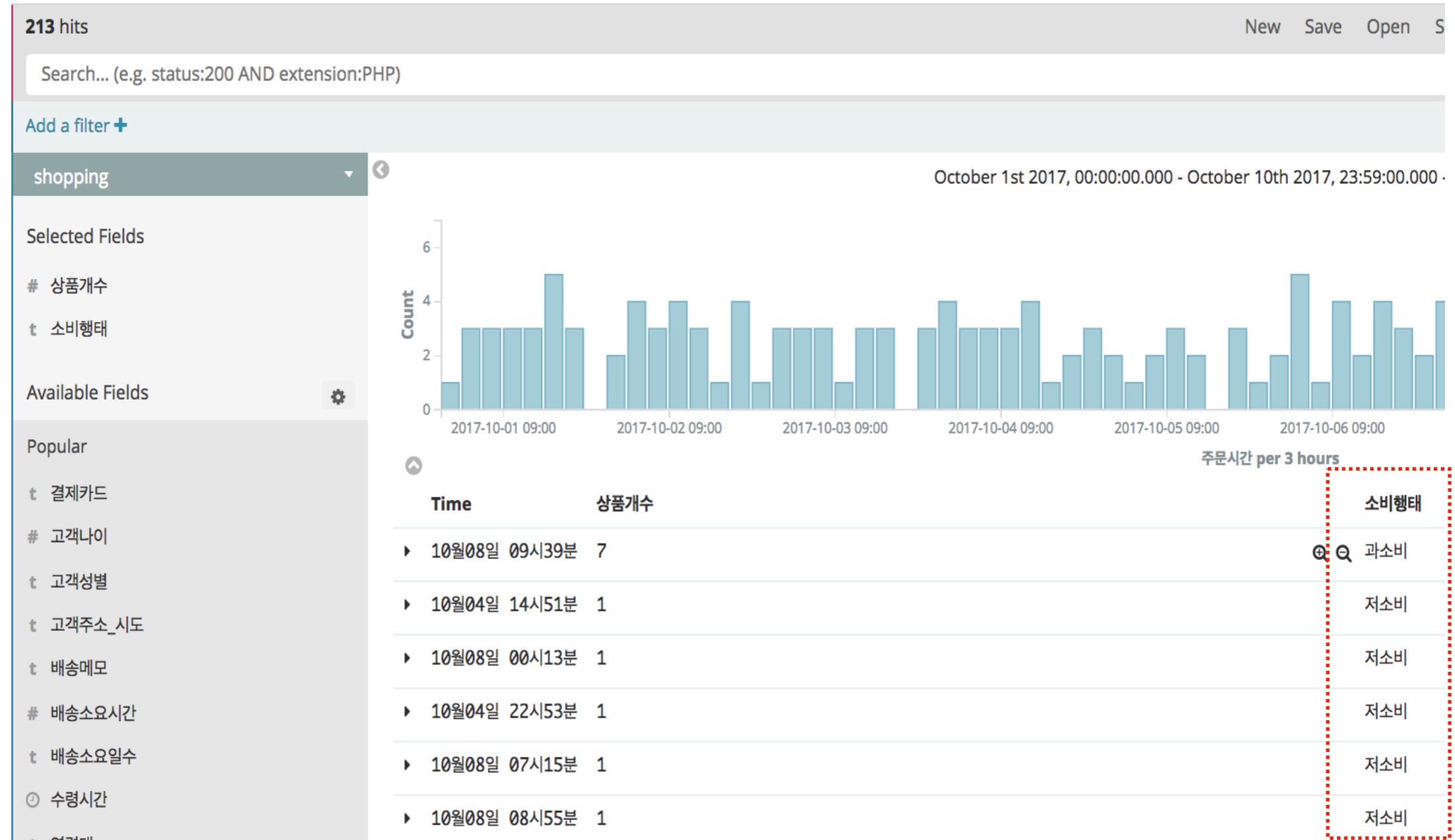
Script

```
if (doc['고객나이'].value < 20) {  
    return "10대"  
}  
else if (doc['고객나이'].value < 30) {  
    return "20대"  
}  
else if (doc['고객나이'].value < 40) {  
    return "30대"  
}  
else if (doc['고객나이'].value < 50) {  
    return "40대"  
} else {  
    return "50대 이상"  
}
```

 4. Script 작성

Scripted Field - If Condition

아래와 같은 Field를 생성해보자 ↗



"상품개수"에 따라 카테고리화 하기

- 1~2 : 저소비
- 3~5 : 평균
- 6~7 : 과소비



Scripted Field

단, Scripted Field는 **Search**는 안된다!!

그러므로 Scripted Field를 검색/조회할 때는 **Filter**를 이용해야 한다

Filtering by Field 

**Dashboard는 만들었는데
원하는 조건의 데이터만 보고 싶으면?**

Filter by Field

우선 Dashboard를 열자

The screenshot shows the Kibana interface. On the left, there is a sidebar with the following options:

- Discover
- Visualize
- Dashboard** (highlighted with a red dashed box and a hand icon)
- Timeline
- Dev Tools
- Management

The main area is titled "Dashboard" and contains a list of fields:

Name	Description
<input type="checkbox"/> data	
<input type="checkbox"/> higee	
<input type="checkbox"/> shopping	
<input type="checkbox"/> test_copy	
<input type="checkbox"/> week1	
<input type="checkbox"/> week2	

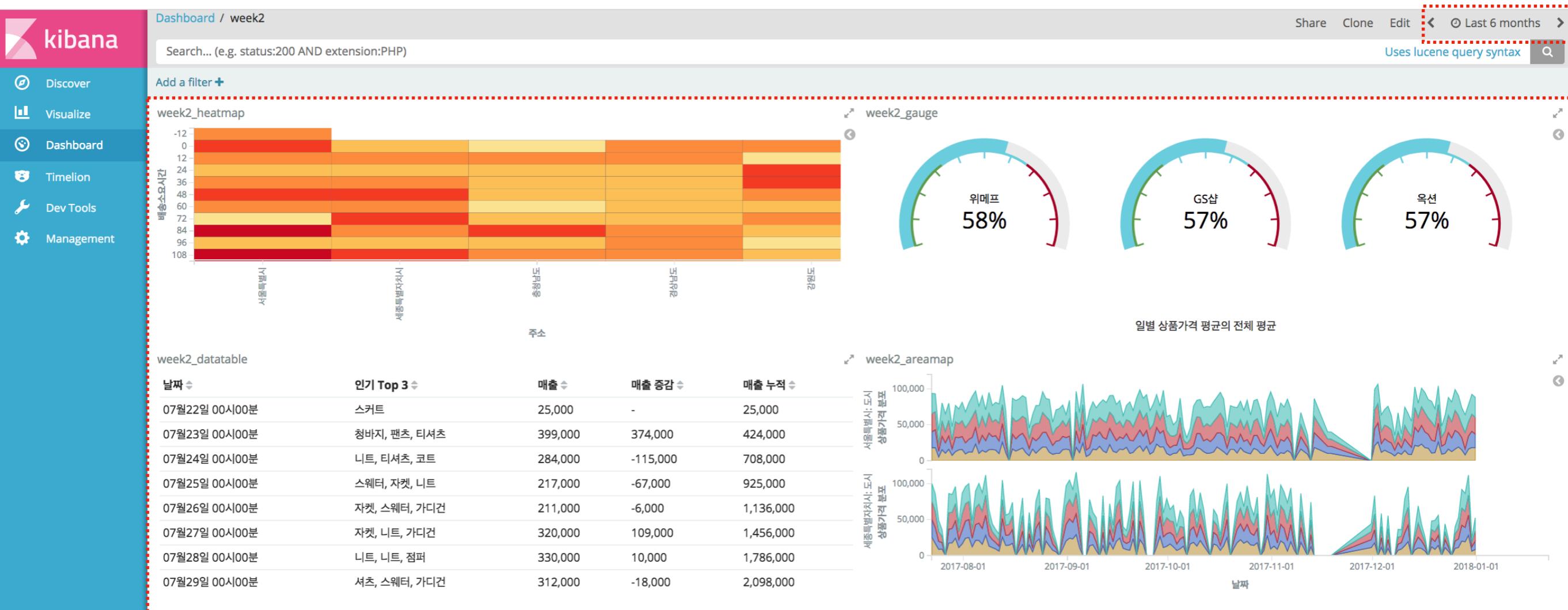
Two annotations are present:

1. 선택 (Select) - points to the "data" checkbox.
2. week2_{id} 선택 (없으면 week2 선택) (Select week2_{id} (if not selected, select week2)) - points to the "week2" checkbox.

Filter by Field

Dashboard를 확인하자

Time Picker 



전체 Documents 중에서 Time Picker 구간에 속한 Documents만 보여준다.

만약에 다른 조건을 추가하고 싶다면? 예를 들어, 20대의 데이터만 보고 싶으면 어떻게 할까?

Filter by Field

Filter를 이용하면 특정 조건을 만족하는 데이터만 선별하여 Dashboard에 시각화할 수 있다.

Filter by Field

Filter를 실행하자

Dashboard / week2

Share Clone Edit < ⏴ Last 6 months >

Search... (e.g. status:200 AND extension:PHP) Uses lucene query syntax

Add a filter + 선택

Add filter

Filter Fields...

Label Optional

Cancel Save

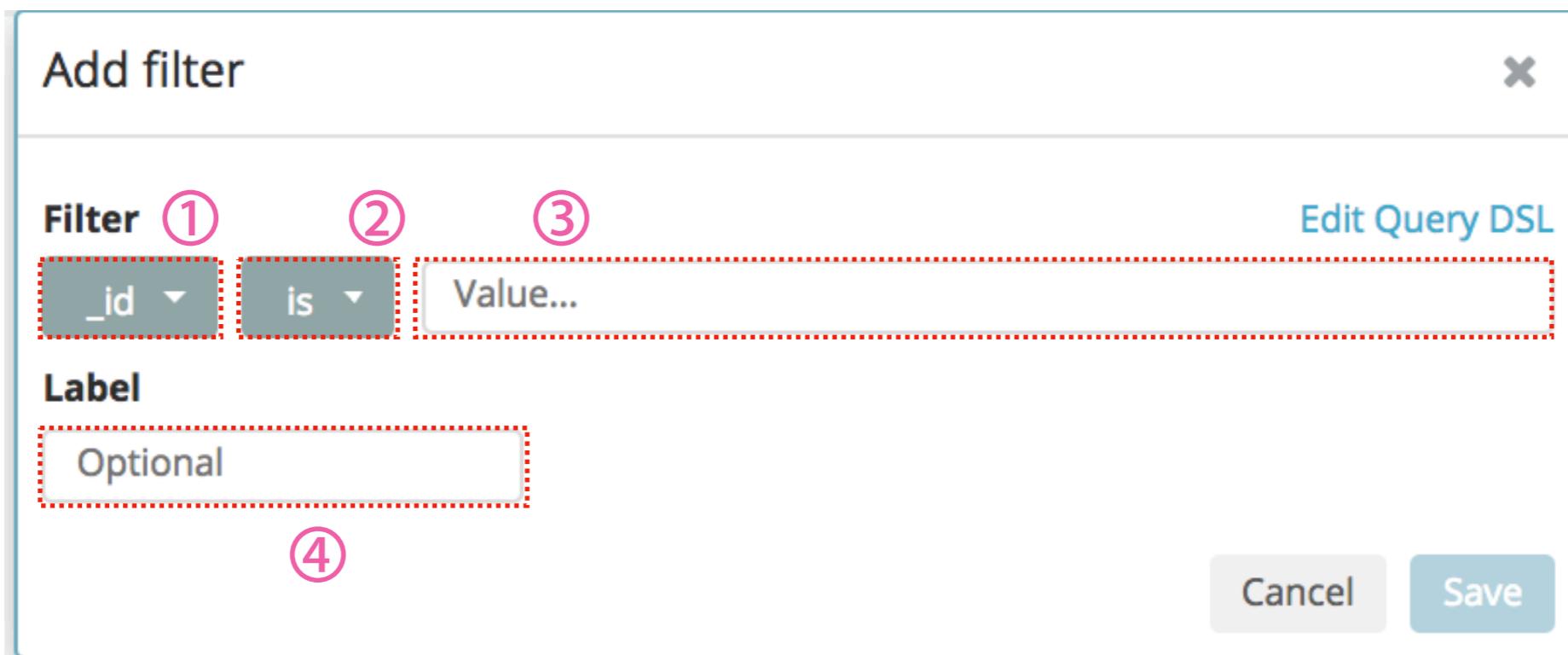
Edit Query DSL

날짜 인기 Top 3 매출 매출 증감 매출 누적

날짜	인기 Top 3	매출	매출 증감	매출 누적
07월22일 00시00분	스커트	25,000	-	25,000
07월23일 00시00분	청바지, 팬츠, 티셔츠	399,000	374,000	424,000
07월24일 00시00분	니트, 티셔츠, 코트	284,000	-115,000	708,000
07월25일 00시00분	스웨터, 자켓, 니트	217,000	-67,000	925,000
07월26일 00시00분	자켓, 스웨터, 가디건	211,000	-6,000	1,136,000
07월27일 00시00분	자켓, 니트, 가디건	320,000	109,000	1,456,000
07월28일 00시00분	니트, 니트, 점퍼	330,000	10,000	1,786,000
07월29일 00시00분	셔츠, 스웨터, 가디건	312,000	-18,000	2,098,000

Filter by Field

Filter의 사용법을 익하자



- ① Filter 적용할 Field 선택
- ② 적용할 Operator 선택 (다음 페이지 참조)
- ③ Filter에 적용하려는 Value 입력
- ④ (여러 Filter 구분하기 위한) 이름 입력

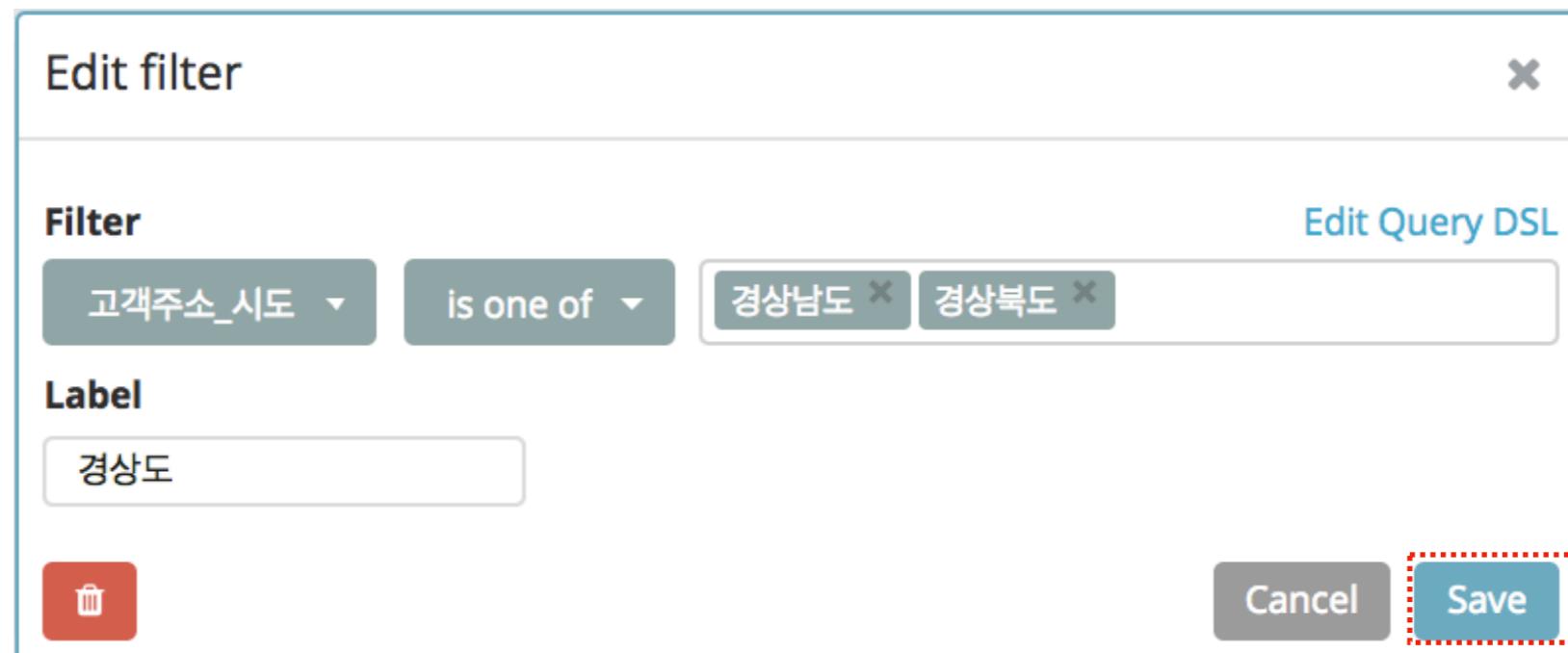
Operator 설명

Operator	역할
is	Field의 Value가 입력한 값과 일치하는 Documents 선택
is not	Field의 Value가 입력한 값과 일치하지 않는 Documents 선택
is one of	Field의 Value가 입력한 값 중에 존재하는 Documents 선택
is not one of	Field의 Value가 입력한 값 중에 존재하지 않는 Documents 선택
exists	Field가 적어도 한 개의 non-null 값을 가지는 Documents 선택
does not exist	Field가 존재하지 않거나 null 값만 가지는 Documents 선택
is between	Field의 Value가 입력한 값 사이에 존재하는 Documents 검색
is not between	Field의 Value가 입력한 값 사이에 존재하지 않는 Documents 검색

Filter by Field

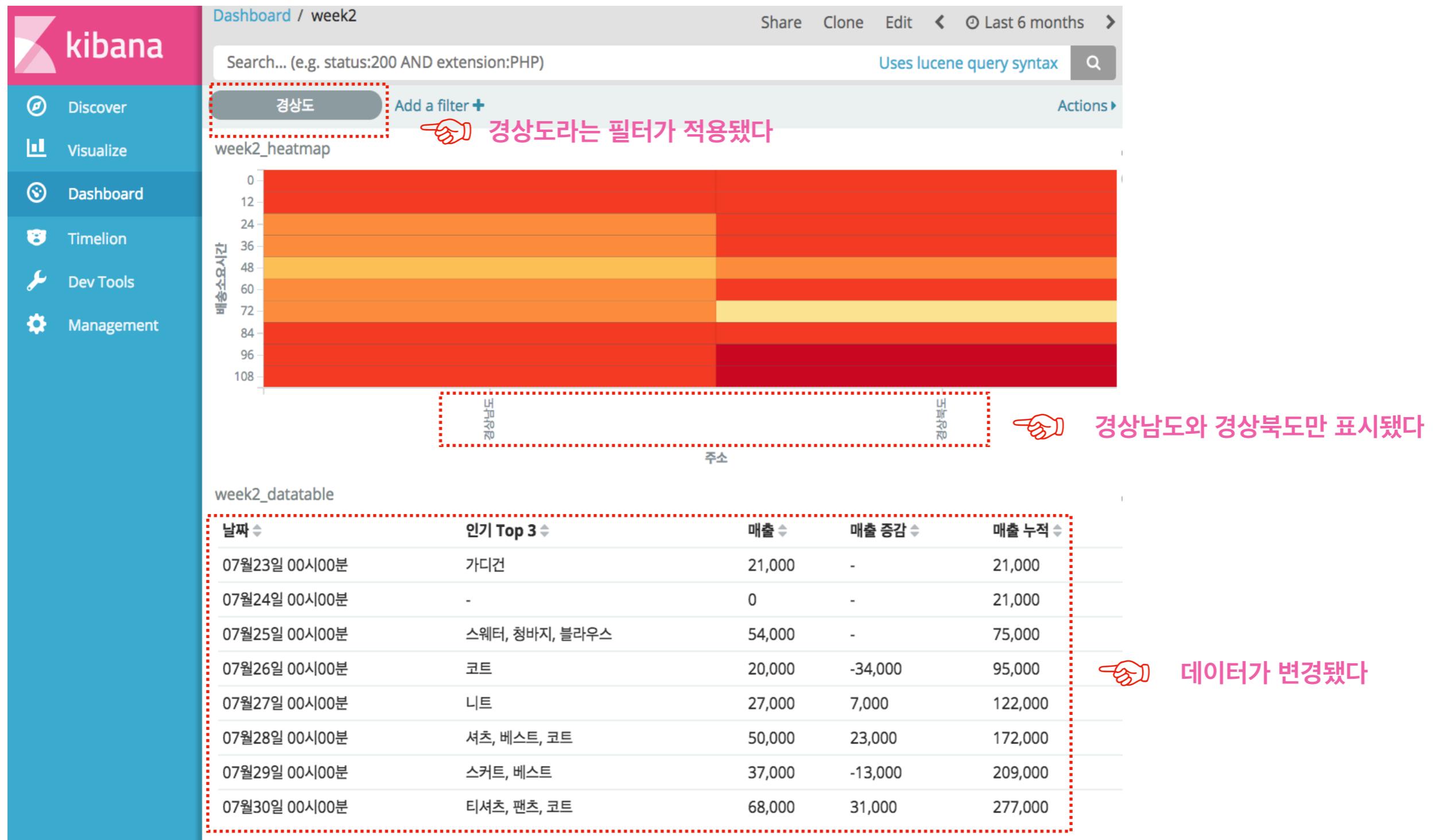
실 사례를 보자 ↗

“고객주소_시도” 가 “경상남도” 또는 “경상북도” 인 Documents를 선택하는 Filter를 만들고 “경상도”라고 하자



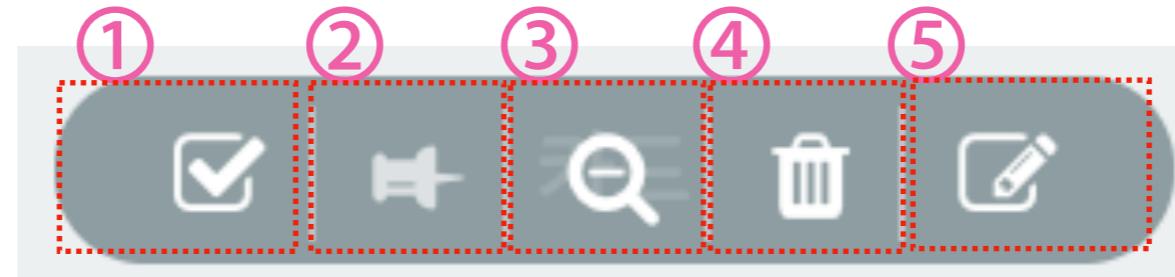
Filter by Field

다시 Dashboard를 보자



Filter by Field

필터에 마우스오버하면...



- ① 필터 적용 <=> 필터 적용 해제
- ② 필터 고정 (Discover, Visualize, Dashboard)
- ③ 필터 효과 적용 <=> 필터 효과 반대 적용
- ④ 필터 삭제
- ⑤ 필터 수정

Filter by Field

Discover에서 Filter를 적용해보자

Filter 적용

Kibana Discover interface showing a histogram of purchase times per day and three detailed document snippets.

Selected Fields: _source

Available Fields: (Listed under Popular)

- t 결제카드
- # 고객나이
- t 고객성별
- t 배송메모
- # 배송소요시간
- t 배송소요일수
- ① 수령시간
- t 연령대
- t 요일
- ② 주문시간
- t _id
- t _index
- # _score
- t _type
- t 결제카드-

Filter: 경상도 (highlighted with a red box and a hand icon)

Search Results:

- 01월01일 06시47분: 고객주소_시도: 경상남도 접수번호: 11258 주문시간: 01월01일 06시47분 수령시간: 01월03일 10시08분 예약여부: 일반 배송메모: 상품 이상 고객ip: 48.247.19.201 고객성별: 여성 고객나이: 30 물건좌표: 36.87039710269179, 128.3913144753862 구매사이트: 옥션 판매자평점: 5 상품분류: 블라우스 상품가격: 11,000 상품개수: 7 결제카드: 우리 _id: AV-iEwZcRJy4v-Hns3-K _type: shopping _index: shopping _score: - 배송소요일수: 모레 배송소요일수_sort: 2 시간대: 21 연령대: 30대 소비행태: 과소비 배송소요시간: 51 요일_sort: 7 요일: 일 성별-카드: 여성-우리 배송소요시간-: 51 결제카드-: 우리카드
- 12월31일 18시58분: 고객주소_시도: 경상남도 접수번호: 4402 주문시간: 12월31일 18시58분 수령시간: 01월04일 10시50분 예약여부: 일반 배송메모: 상품 이상 고객ip: 25.210.231.143 고객성별: 남성 고객나이: 24 물건좌표: 35.50457351682122, 129.24849426336837 구매사이트: 위메프 판매자평점: 4 상품분류: 청바지 상품가격: 7,000 상품개수: 7 결제카드: 삼성 _id: AV-iDfJcRJy4v-Hns2TB _type: shopping _index: shopping _score: - 배송소요일수: 나흘 이상 배송소요일수_sort: 3 시간대: 9 연령대: 20대 소비행태: 과소비 배송소요시간: 87 요일_sort: 7 요일: 일 성별-카드: 남성-삼성 배송소요시간-: 87 결제카드-: 삼성카드
- 12월31일 15시19분: 고객주소_시도: 경상북도 접수번호: 8095 주문시간: 12월31일 15시19분 수령시간: 01월02일 04시27분 예약여부: 일반 배송메모: 시간 내에 배송 못함 고객ip: 171.46.43.135 고객성별: 남성 고객나이: 30 물건좌표: 35.98496142746445, 129.21908843787452 구매사이트: GS샵 판매자평점: 1 상품분류: 청바지 상품가격: 26,000 상품개수: 7 결제카드: 우리 _id: AV-iDyCDR-Jy4v-Hns3Mu _type: shopping _index: shopping _score: - 배송소요일수: 다음날 배송소요일수_sort: 1 시간대: 6 연령대: 30대 소비행태: 과소비 배송소요시간: 37 요일_sort: 7 요일: 일 성별-카드: 남성-우리 배송소요시간-: 37 결제카드-: 우리카드

Filter된 Documents 하나하나를 볼 수 있다

Filter by Field

아래와 같은 Filter를 Dashboard에 적용해보자 ↗

Dashboard : week2_{id}

Time Range : Last 6 months

1. “고객성별”이 여성인 Data
2. “결제카드”가 우리 또는 국민인 Data
3. “고객성별”이 남성이면서 “연령대”가 20대
4. “구매사이트”가 쿠팡이거나 “상품개수”가 1~3인 Data
5. “결제카드”가 “우”로 시작하는 모든 Data
6. “구매사이트”가 22번가(오타 아니에요)와 유사한 Data

Lucene Query 

Filter는 사용하기 간단하나 기능이 제한적이다

+

더 스마트한 검색!!

우선 Dashboard를 열자

The screenshot shows the Kibana interface. On the left, there is a sidebar with the following options:

- Discover
- Visualize
- Dashboard** (highlighted with a red dashed box and a hand icon)
- Timeline
- Dev Tools
- Management

The main area is titled "Dashboard" and contains a list of items:

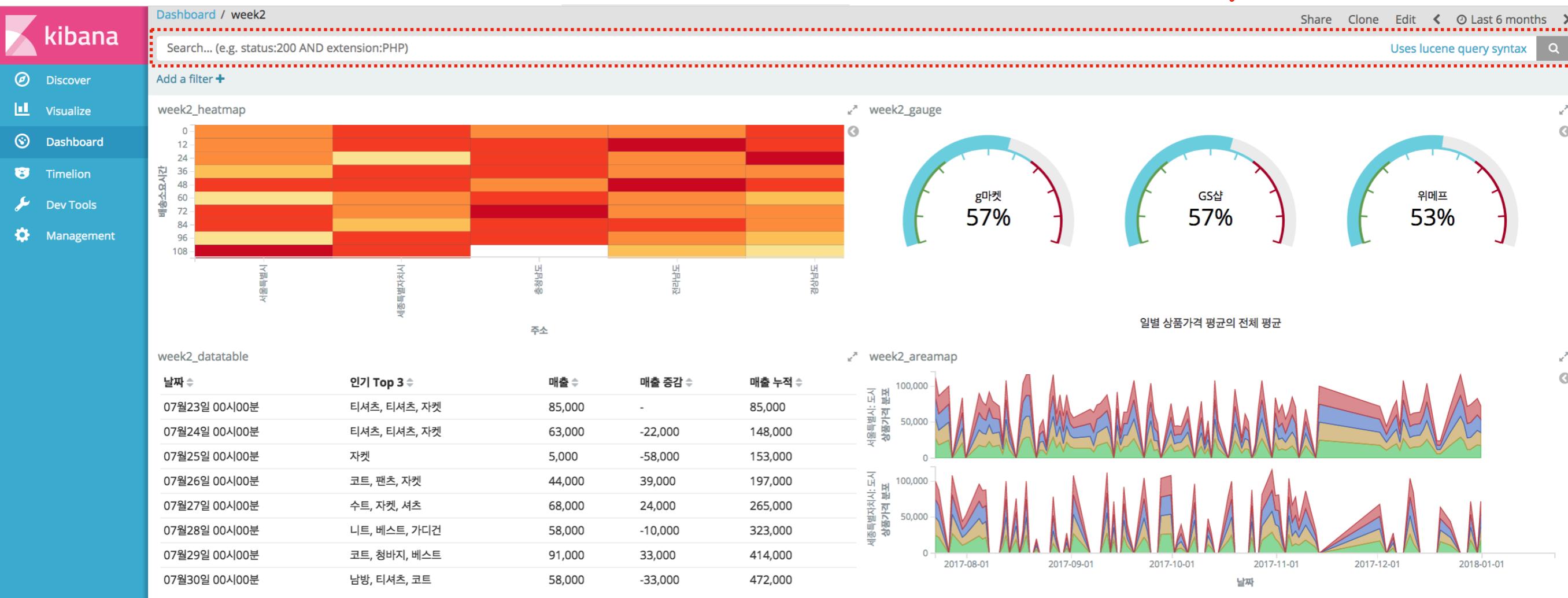
Name	Description
<input type="checkbox"/> data	
<input type="checkbox"/> higee	
<input type="checkbox"/> shopping	
<input type="checkbox"/> test_copy	
<input type="checkbox"/> week1	
<input type="checkbox"/> week2	

Two steps are highlighted with pink boxes and arrows:

1. 선택 (Select) - Points to the "data" item in the list.
2. week2_{id} 선택 (없으면 week2 선택) (Select week2_{id} (or week2 if not present)) - Points to the "week2" item in the list.

Lucene Query

Query Bar를 확인하자



Query Bar에 뭐라고 검색을 해야될까?

Lucene Query의 사용법을 익히자 ↗

종류	기능	Query 예시
Keyword 검색	Field에 상관없이 Value 일치하는 Documents 검색	여성
Field Match 검색	특정 Field의 Value가 일치하는 Documents 검색	고객성별:여성
Exact Field Match 검색	특정 Value가 정확히 모두 일치하는 Documents 검색	배송메모:"상품 이상"
Must be 검색	특정 Field가 존재하는 Documents 검색	_exists_:구매사이트
Must not be present 검색	특정 Field가 존재하지 않는 Documents 검색	_missing_:구매사이트
AND 검색	특정 조건들을 모두 만족하는 Documents 검색	고객성별:여성 AND 상품분류:셔츠
OR 검색	특정 조건들 중 적어도 1개를 만족하는 Documents 검색	고객성별:남성 OR 상품분류:셔츠
NOT 검색	특정 조건을 만족하지 않는 Documents 검색	NOT 구매사이트:옵션
Term 검색	조건 중 적어도 하나라도 만족하는 Documents 검색	상품분류: (니트 코트)
Fuzzy 검색	검색어와 유사한 Documents 검색	경상복도~
Proximity 검색	검색어의 순서를 변경해서 찾을 수 있는 Documents 검색	배송메모:"내에 시간 배송 못함"~2
Numeric Value 검색	Numeric Field Value로 Documents 검색	상품가격:>5000
Range 검색	Field의 Value가 입력한 값 사이에 존재하는 Documents 검색	고객나이 : [10 TO 30]
Wildcard ? 검색	Wildcard ?(한글자)를 활용해서 Documents 검색	서?특별시
Wildcard * 검색	Wildcard *(모든글자)를 활용해서 Documents 검색	쿠*

아래와 같은 Query를 Dashboard에서 검색해보자 ↗

Dashboard : week2_{id}

Time Range : Last 6 months

1. “고객성별”이 여성인 Data
2. “결제카드”가 우리 또는 국민인 Data
3. “고객성별”이 남성이면서 “연령대”가 20대
4. “구매사이트”가 쿠팡이거나 “상품개수”가 1~3인 Data
5. “결제카드”가 “우”로 시작하는 모든 Data
6. “구매사이트”가 22번가(오타 아니에요)와 유사한 Data

Filter와 Search를 비교해보자

		Filter	Search
“고객성별”이 여성인 Data		✓	✓
“결제카드”가 우리 또는 국민인 Data		✓	✓
“고객성별”이 남성이면서 “연령대”가 20대	SCRIPTED FIELD	✓	
“구매사이트”가 쿠팡이거나 “상품개수”가 1~3인 Data	OR 연산		✓
“결제카드”가 “우”로 시작하는 모든 Data	WILDCARD 검색		✓
“구매사이트”가 22번가(오타 아니에요)와 유사한 Data	FUZZY / PROXIMITY 검색		✓