

# **EFK를 활용한 로그 수집과 분석 (3일)**

**(Elasticsearch + Fluentd + Kibana)**

**2022 – 09**

# 3 일차: EFK, Kibana

- Kibana 시각화 및 탐색 도구의 이해
- Index와 Event 및 다양한 시각화
- Kibana를 활용한 Kubernetes 통합 실습: 메트릭 수집 및 분석
  - 1) 호스트, Docker, Application, Kubernetes Metric 수집
  - 2) Kubernetes 상태 Metric 및 이벤트 수집
  - 3) Kibana의 Metric과 Dashboard 활용 다양한 시각화

과정 목표

# 비정형 데이터의 정형화

과정 목표(2일차)

# 다양한 머신데이터와 인덱스패턴 (SourceType)

**과정 목표**

**시각화 (Visualization)**

## 교육 내용 및 일정 소개

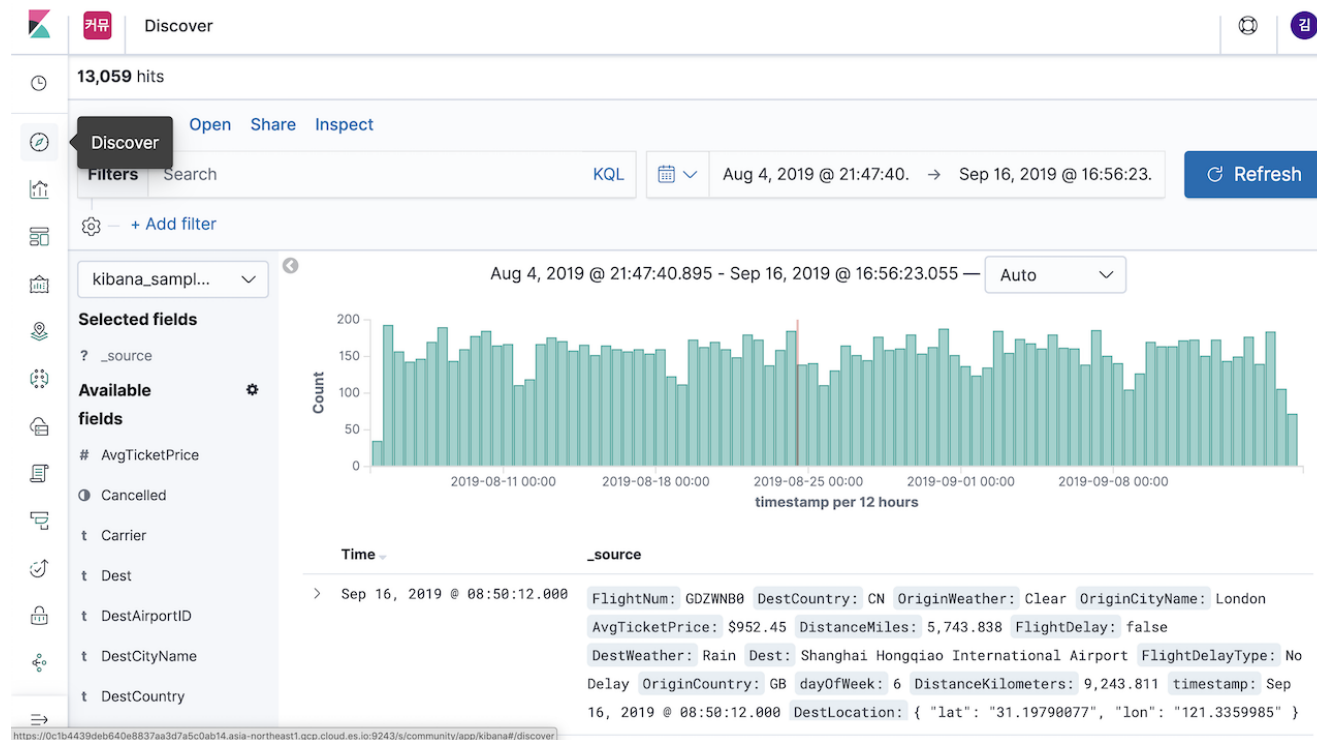
교육 내용	주제	시간	학습 내용
	Kibana 개요	1H	<ul style="list-style-type: none"> <li>▪ Kibana 설치 및 기본 환경 설정</li> </ul>
	Kibana 활용	1H	<ul style="list-style-type: none"> <li>▪ Kibana 활용 방법 – Discover / Dashboard / Visualization</li> </ul>
		1H	<ul style="list-style-type: none"> <li>▪ [실습] 수집된 데이터의 인덱스화 및 Devtools를 이용한 조회 실습</li> </ul>
	쿠버네티스 메트릭 수집 파이프라인 구축 및 Kibana를 통한 데이터 분석	1H	<ul style="list-style-type: none"> <li>▪ Kibana를 활용한 Kubernetes 메트릭 수집 파이프라인 실습 환경 구축</li> </ul>
		1H	<ul style="list-style-type: none"> <li>▪ [실습] 호스트, Docker, Application, Kubernetes Metric 수집</li> </ul>
		1H	<ul style="list-style-type: none"> <li>▪ [실습] Kubernetes 상태 Metric 및 이벤트 수집</li> </ul>
		1H	<ul style="list-style-type: none"> <li>▪ [실습] Kibana의 Dashboard 활용 다양한 시각화</li> </ul>

## Kibana는 Elasticsearch를 가장 쉽게 시각화 할 수 있는 도구

- Discover, Visualize, Dashboard 3개의 기본 메뉴와 다양한 App 들로 구성
- 플러그인을 통해 App의 설치가 가능

# Discover

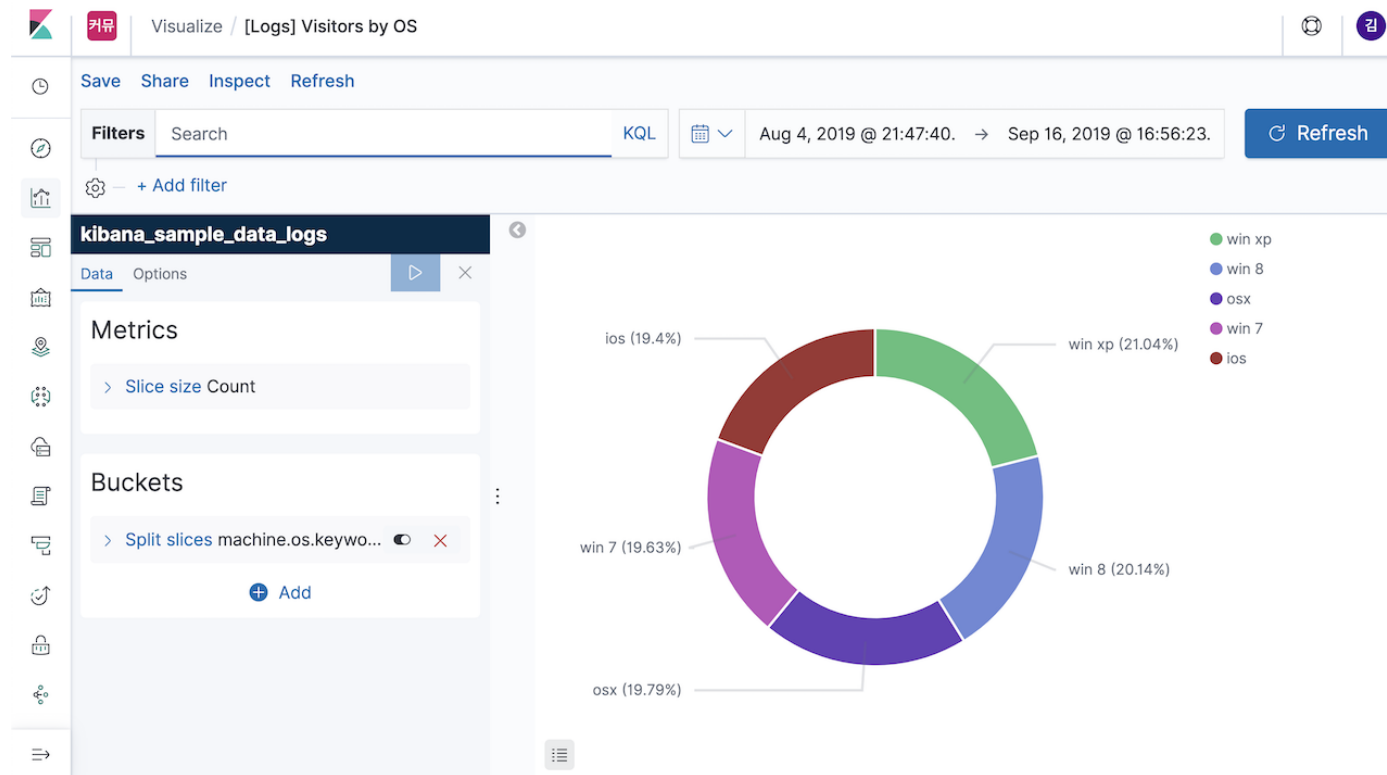
- Discover는 Elasticsearch에 색인된 소스 데이터들의 검색을 위한 메뉴입니다. 검색 창에 질의문을 통해 데이터를 간편하게 검색, 필터링 할 수 있으며, 검색된 데이터의 원본 문서를 확인하거나 보고 싶은 필드만 선택해서 테이블 형태로 조회가 가능합니다. 시계열(time series) 기반의 로그 데이터인 경우 시간 히스토그램 그래프를 통해 시간대별 로그 수도 표시됩니다.





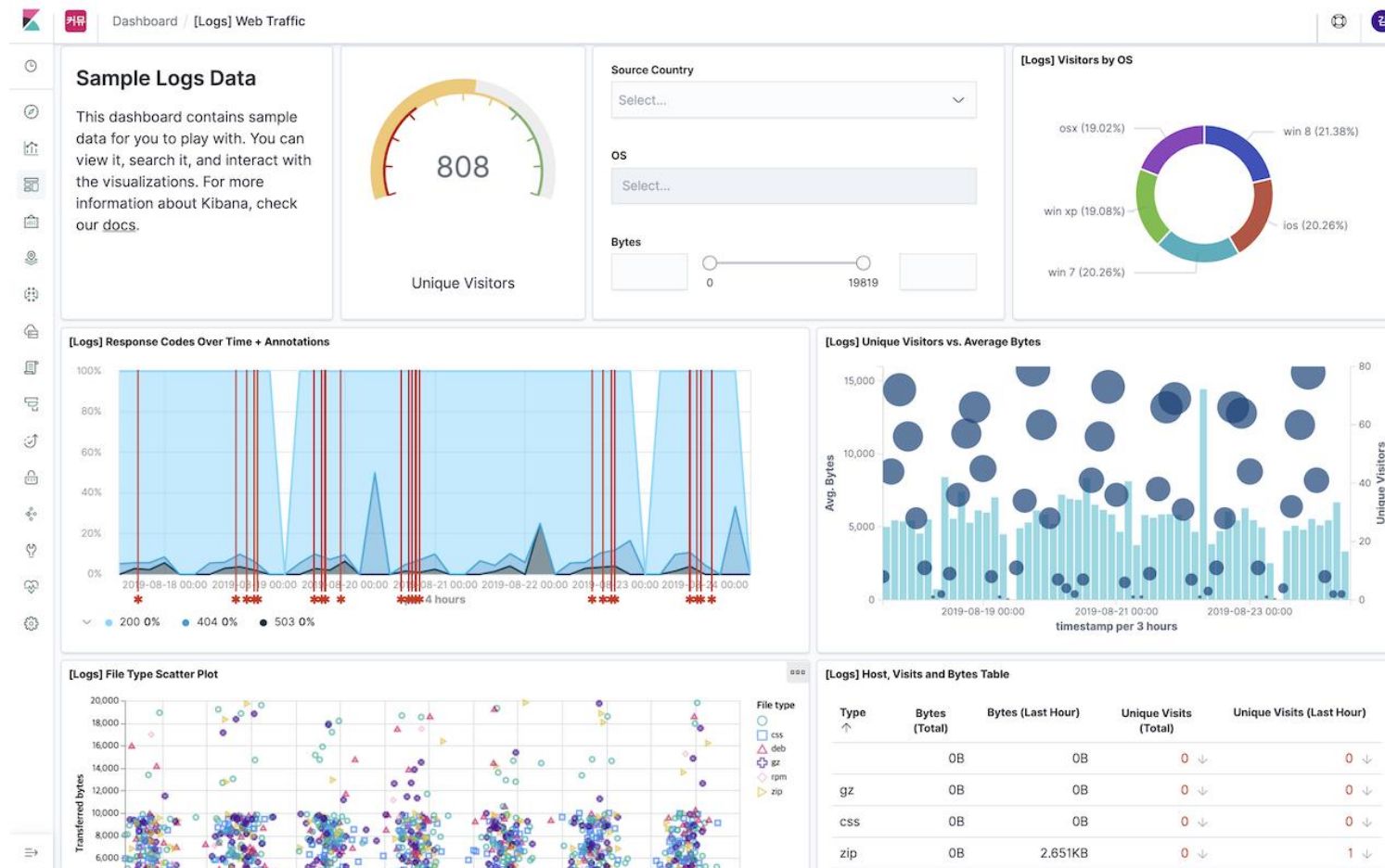
# Visualize

- Visualize는 aggregation 집계 기능을 통해 조회된 데이터의 통계를 다양한 차트로 표현할 수 있는 패널을 만드는 메뉴입니다. 영역차트, 바차트, 파이차트, 라인차트 등 다양한 시각화 도구들의 사용이 가능하며 여기서 만들어진 패널들을 조합해서 대시보드를 만들게 됩니다.



# Dashboard

- Visualize 메뉴에서 만들어진 시각화 도구들을 조합해서 대시보드 화면을 만들고 저장, 불러오기 등을 할 수 있는 메뉴



# Kibana 기능 소개

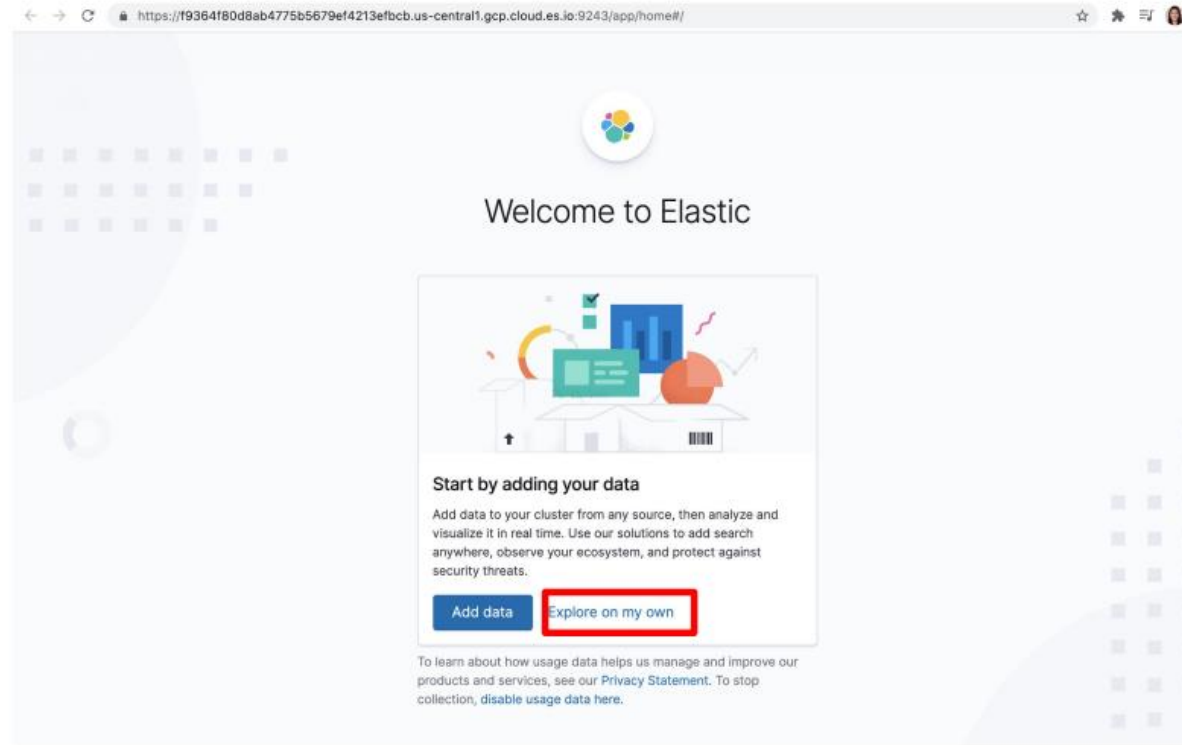
## Kibana 기능

Kibana는 Elasticsearch 데이터의 시각화와 Elastic Stack의 탐색을 지원합니다. Kibana를 활용하면 새벽 2시에 긴급 호출을 받은 이유를 신속하게 찾아내고 자연재해가 이번 분기 수치에 미치는 영향을 예측하는 등과 같은 작업을 수행할 수 있습니다.

탐색 및 시각화	관리 및 모니터링	솔루션
시각화	SECURITY	ELASTIC MAPS
Kibana Lens	보안 스페이스	지도 레이어
Time Series Visual Builder	암호화된 통신	벡터 타일
위치 기반 정보 분석	역할 기반 액세스 제어(RBAC)	사용자 정의 지역 지도
차트	익명 액세스 제어(공개 공유용)	Elastic Maps Service(확대/축소 수준)
Metrics	필드 및 문서 수준 보안	Elastic Maps Server
데이터 표	보안 영역	GeoJSON 업로드
Vega(사용자 정의)	싱글 사인온(SSO)	위치 기반 경보 기능
Kibana 플러그인	보안 API	셰이프파일 업로드
Canvas		
User Experience	관리	ELASTIC LOGS
Kibana 런타임 필드 편집기	다크 테마	Log 수집기(Filebeat)
	인덱스 패턴	Logs 대시보드
데이터 탐색	인덱스 수명 주기 관리	로그 속도 이상 징후 탐지
대시보드	스냅샷 수명 주기 관리	Logs 앱
발견	데이터 풀업 관리	
필드 통계	사용자 및 역할 관리	

<https://www.elastic.co/kr/kibana/features>

# Kibana 실습



## Kibana 열기

# Kibana 실습

The screenshot shows the Elastic Cloud console interface. The browser address bar displays the URL: <https://cloud.elastic.co/deployments/316be9b6fe7c451cba113e1b2cfe35c5>. The page title is "Beginner's Crash Course" for the deployment "us-central1 (Iowa)".

**Left Sidebar:**

- Deployments**
  - Beginner's Crash Course to t...
  - Edit
  - Elasticsearch
  - Snapshots
  - API console
  - Kibana
  - APM
  - Enterprise Search
  - Logs and metrics
  - Activity
  - Security
  - Performance
- Extensions**
- API keys**
- Traffic filters**
- Help**

**Main Content Area:**

## Beginner's Crash Course

**Get started with your deployment**

The next step is to ingest data and create visualizations in Kibana.

**Open Kibana** (button highlighted with a red rectangle)

Forgot to save your credentials?  
[Reset your deployment password](#)

**Deployment details:**

- Deployment name:** Beginner's Crash Course (with [Edit](#) link)
- Deployment ID:** 316be9b
- Deployment version:** v7.10.1
- Deployment status:** Healthy
- Actions:** [Open Kibana](#) and [Manage](#) (dropdown)

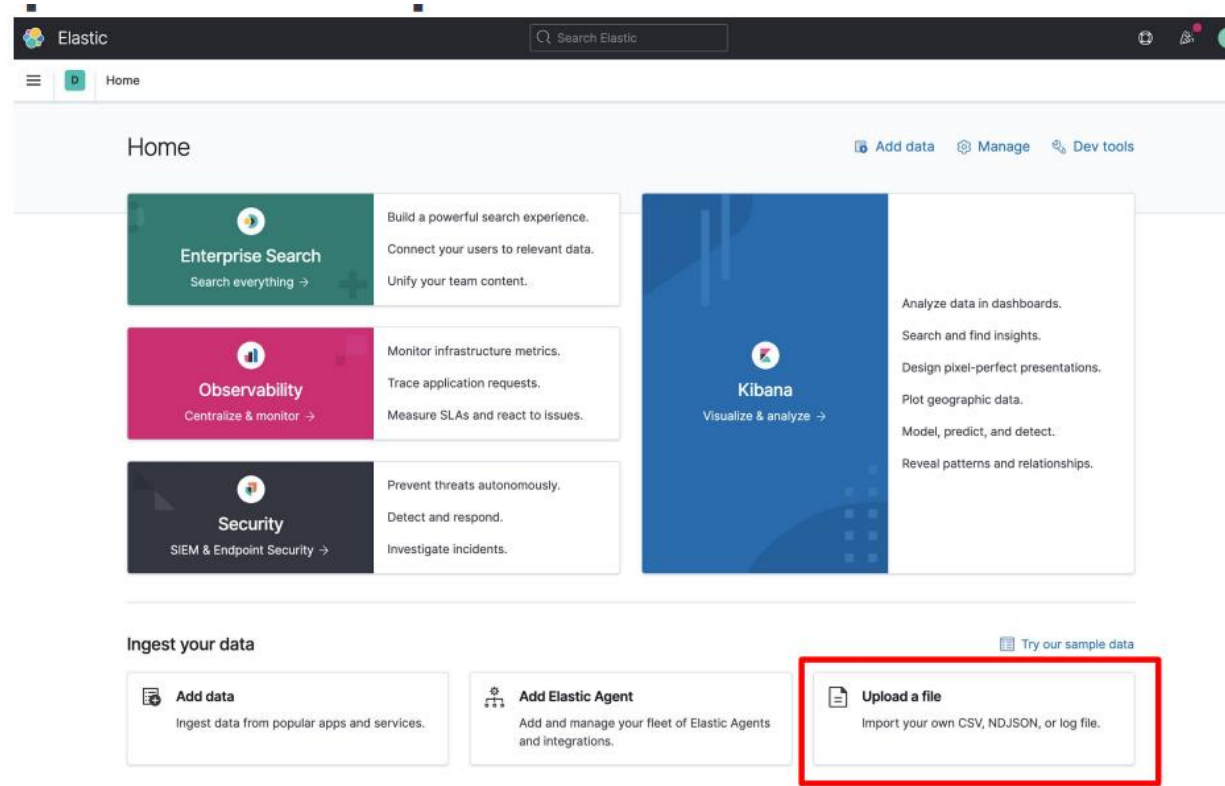
**Preview of Kibana interface:**

The preview shows the Kibana dashboard with sections for:

- Observability:** APM, Logs, Metrics, and Security.
- Visualize and Explore Data:** APM, App Search, and Canvas.
- Manage and Administer the Elastic Stack:** Settings, Alerts, and Security Settings.

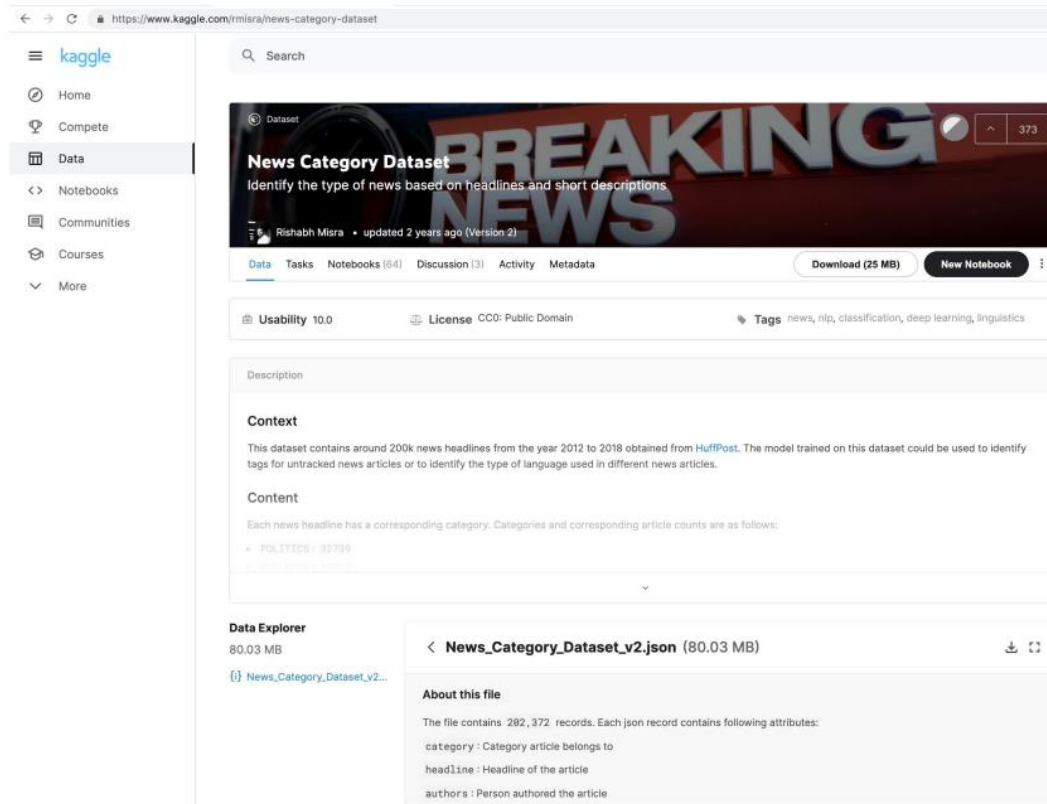
my own option 클릭

# Kibana 실습



Upload file option 선택

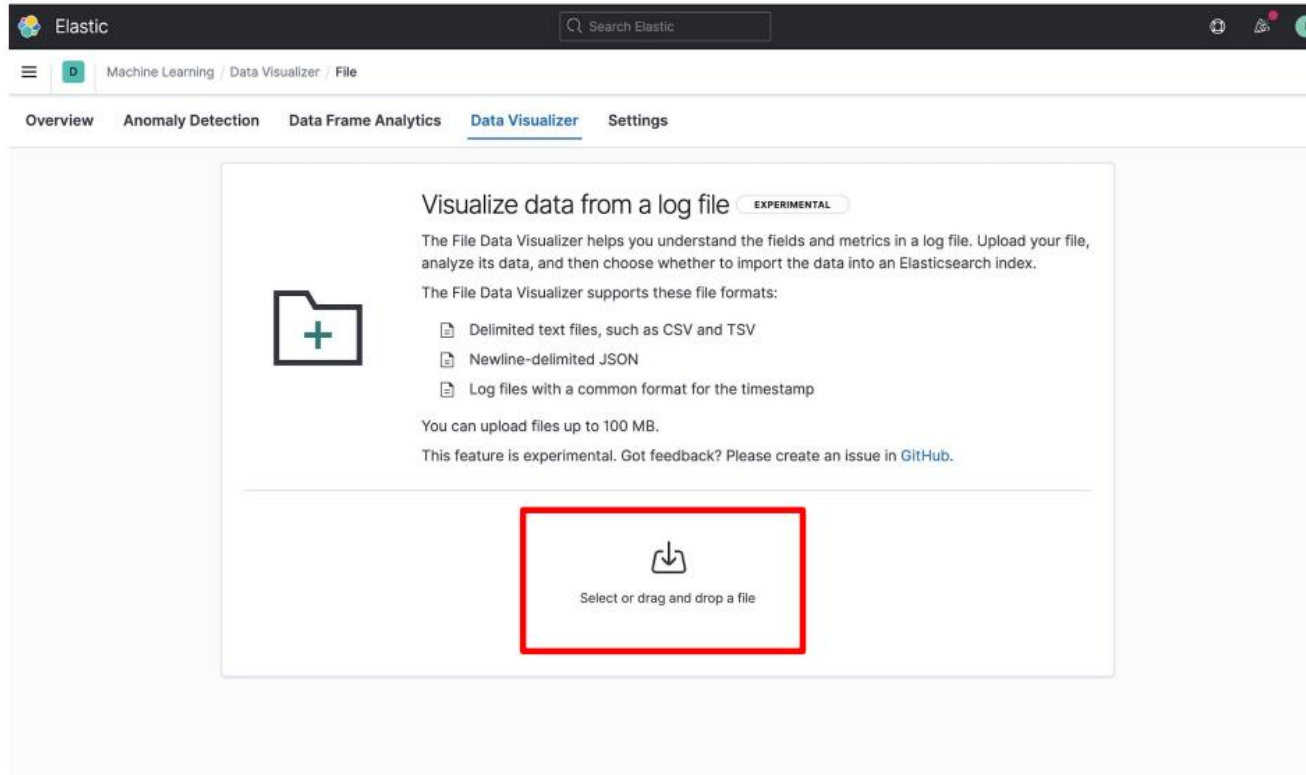
# Kibana 실습



The screenshot shows the Kaggle website interface for the 'News Category Dataset'. The left sidebar contains navigation links: Home, Compete, Data (selected), Notebooks, Communities, Courses, and More. The main content area features a search bar, a dataset card for 'News Category Dataset' by Rishabh Misra (updated 2 years ago, Version 2), and a 'BREAKING NEWS' banner. Below the card, there are tabs for Data, Tasks, Notebooks (64), Discussion (3), Activity, and Metadata. A 'Download (25 MB)' button and a 'New Notebook' button are visible. The dataset's Usability is 10.0, and the License is CC0: Public Domain. Tags include news, nlp, classification, deep learning, and linguistics. The Description section includes a 'Context' paragraph stating the dataset contains around 200k news headlines from 2012 to 2018, and a 'Content' section listing categories and counts: POLITICS (39739), SPORTS (10000), and OTHER (10000). At the bottom, the 'Data Explorer' shows the file 'News\_Category\_Dataset\_v2.json' (80.03 MB) with a download icon. A pop-up window titled 'About this file' provides details: 'The file contains 282,372 records. Each json record contains following attributes: category : Category article belongs to, headline : Headline of the article, authors : Person authored the article'.

## Kaggle News Category Dataset 선택 & Download

# Kibana 실습



**Kaggle News Category Dataset upload**



# Kibana 실습

The screenshot shows the Kibana Data Visualizer interface. The top navigation bar includes the Elastic logo, a search bar, and tabs for Machine Learning, Data Visualizer, and File. The 'Data Visualizer' tab is active, showing a list of visualizations. The 'News\_Category\_Dataset\_v2.json' visualization is selected, displaying the 'File contents' of the dataset. The 'File contents' section shows the first 1,000 lines of the JSON file, with a red box highlighting the first five lines. Below this, the 'Summary' section provides a table of dataset statistics, with a blue box highlighting the 'Format', 'Time field', and 'Time format' rows.

**File contents**  
First 1,000 lines

```
1 {"category": "CRIME", "headline": "There Were 2 Mass Shootings In Texas Last Week, But Only 1 On TV", "authors": "Melissa Jeltsen", "link": "https://www.huffingtonpost.com/entry/texas-amanda-painter-mass-shooting_us_5b081ab4e4b0802d69caad89", "short_description": "She left her husband. He killed their children. Just another day in America.", "date": "2018-05-26"}
2 {"category": "ENTERTAINMENT", "headline": "Will Smith Joins Diplo And Nicky Jam For The 2018 World Cup's Official Song", "authors": "Andy McDonald", "link": "https://www.huffingtonpost.com/entry/will-smith-joins-diplo-and-nicky-jam-for-the-official-2018-world-cup-song_us_5b09726fe4b0fdb2aa541201", "short_description": "Of course it", "date": "2018-05-26"}
3 {"category": "ENTERTAINMENT", "headline": "Hugh Grant Marries For The First Time At Age 57", "authors": "Ron Dicker", "link": "https://www.huffingtonpost.com/entry/hugh-marries_us_5b09212ce4b0568a880b9a8c", "short_description": "The actor and his longtime girlfriend Anna Eberstein tied the knot in a civil ceremony.", "date": "2018-05-26"}
4 {"category": "ENTERTAINMENT", "headline": "Jim Carrey Blasts 'Castrato' Adam Schiff And Democrats In New Artwork", "authors": "Ron Dicker", "link": "https://www.huffingtonpost.com/entry/jim-carrey-adam-schiff-democrats_us_5b0950e8e4b0fdb2aa53e675", "short_description": "The actor gives Dems an ass-kicking for not fighting hard enough against", "date": "2018-05-26"}
5 {"category": "ENTERTAINMENT", "headline": "Julianna Margulies Uses Donald Trump Poop Bags To Pick Up After Her Dog", "authors": "Ron Dicker", "link": "https://www.huffingtonpost.com/entry/julianna-margulies-trump-poop-bags_us_5b093ee7a4b0fdb2aa534f70", "short_description": "The 'Dietland' actress said using the bag is a 'really cathartic th
```

**Summary**

Number of lines analyzed	1000
Format	ndjson
Time field	date
Time format	ISO8601

[Override settings](#) [Analysis explanation](#)

1000줄 데이터 및 데이터 세트 요약 제공

# Kibana 실습



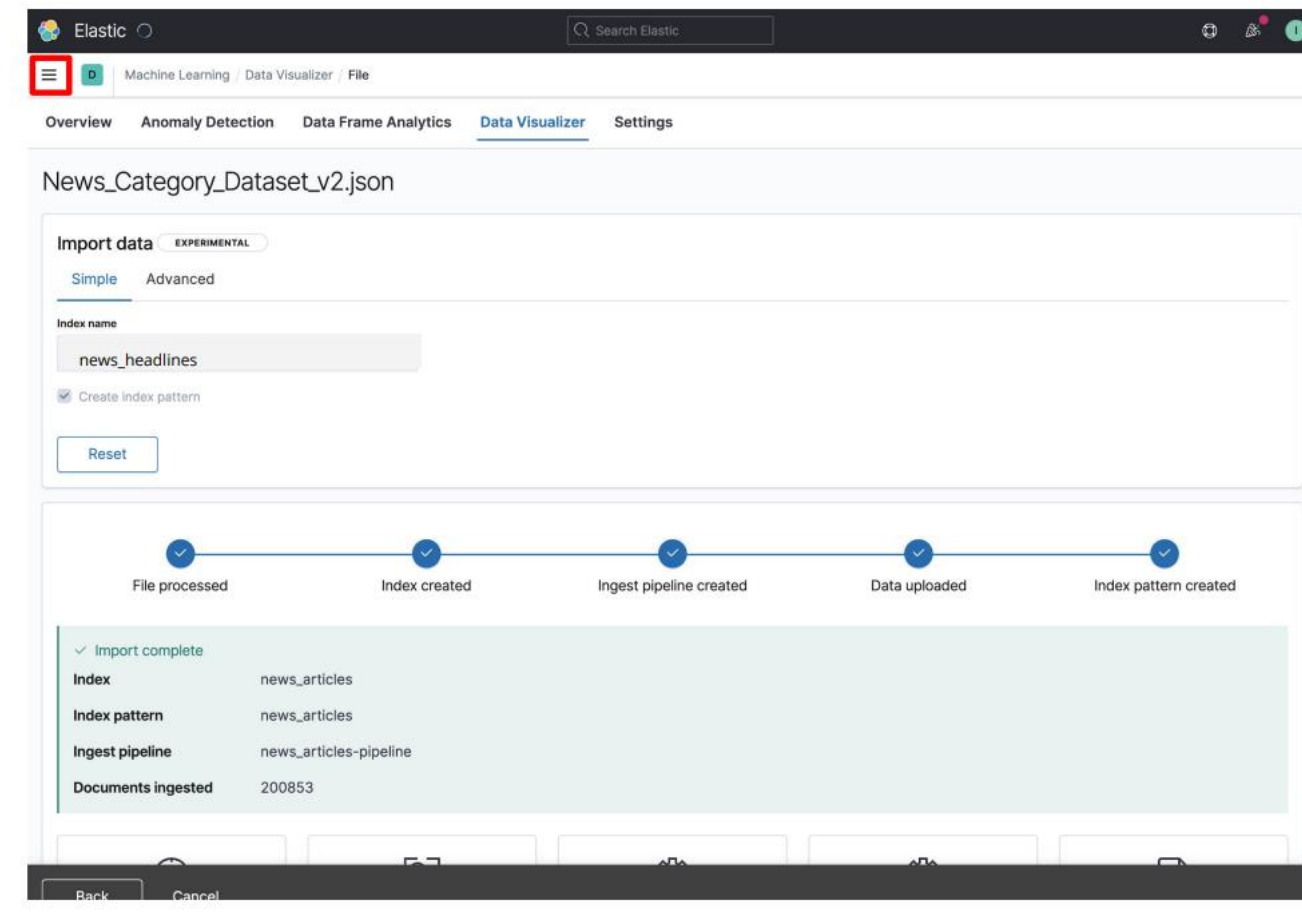
필드 섹션에는 식별된 필드, 높은 수준의 통계, 및 상위 발생 값 표시

# Kibana 실습



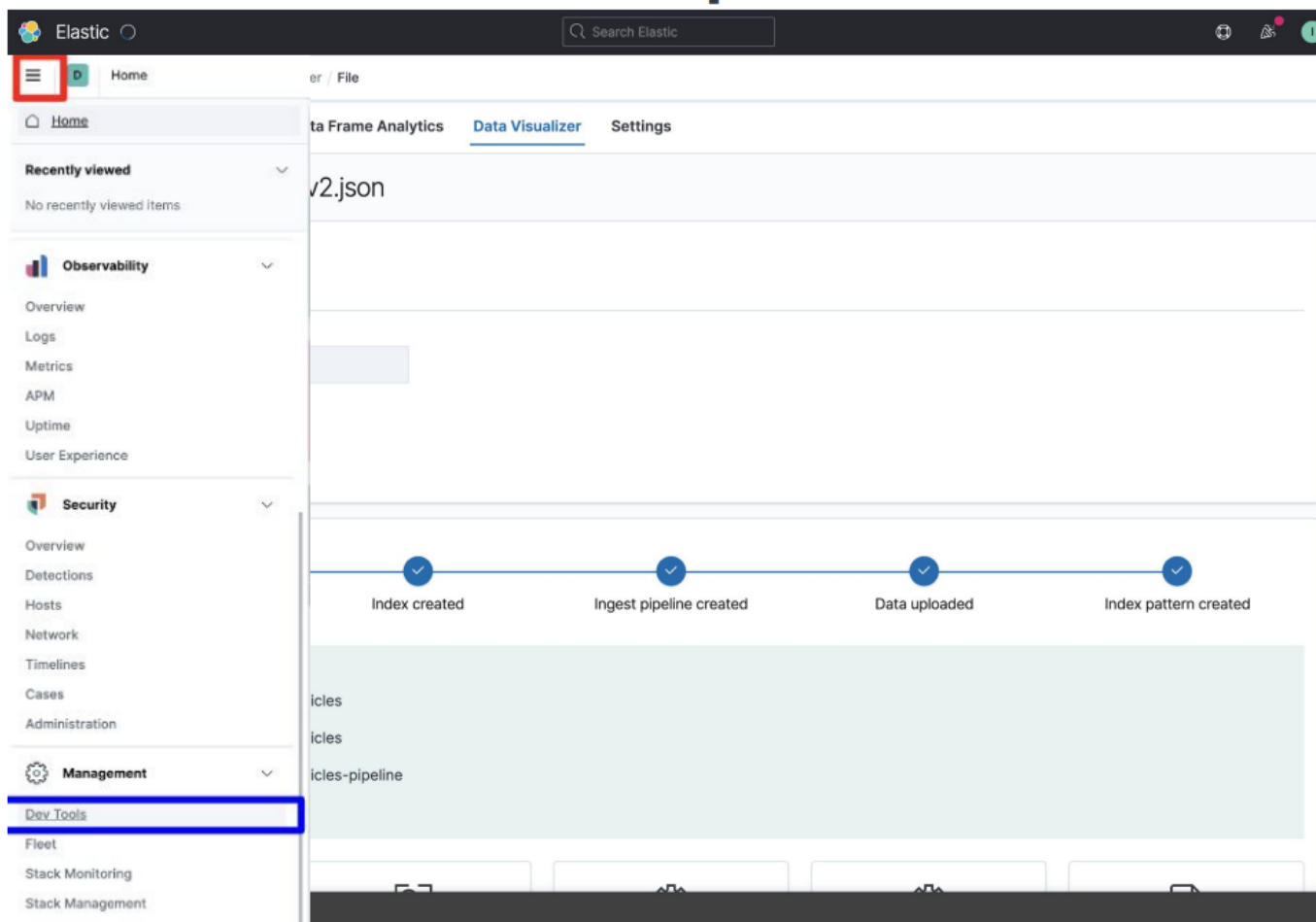
import button 클릭

# Kibana 실습 – Index 생성



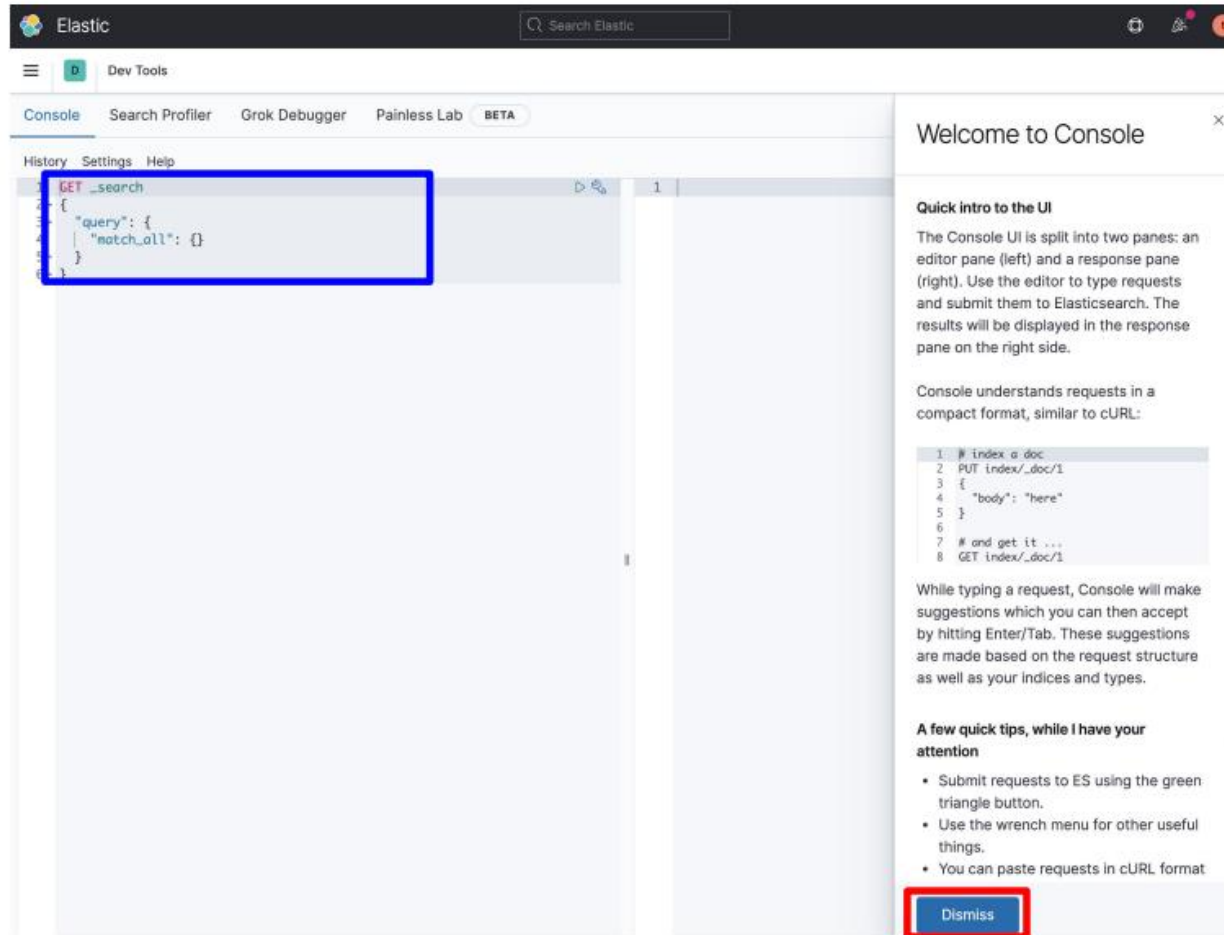
Index 생성

# Kibana 실습 – Index 생성



햄버거 메뉴를 누르고 Dev Tools 메뉴를 클릭

# Kibana 실습

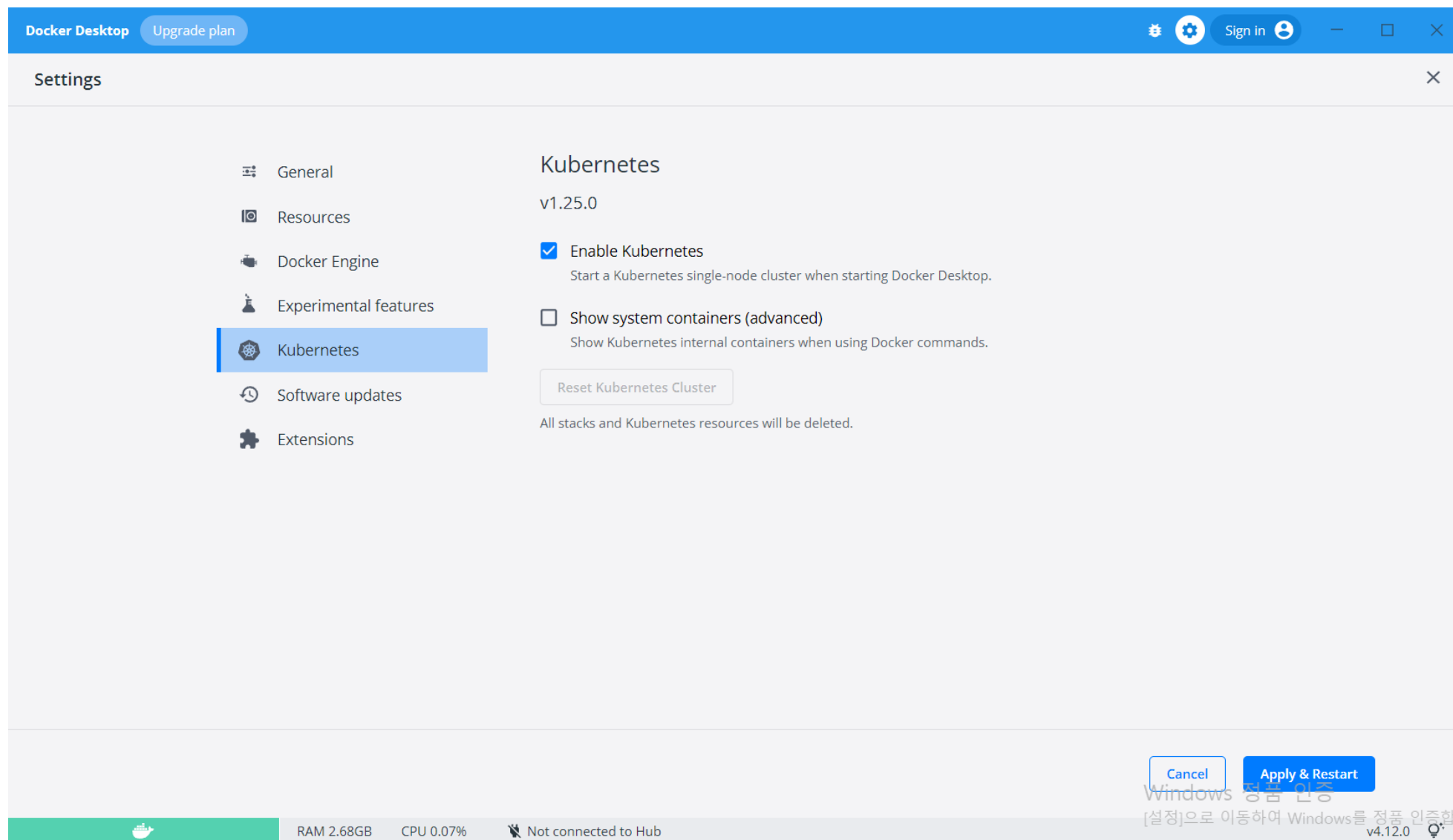


Explore를 사용하여 Data검색

## [실습] Kubernetes에서 메트릭 수집 및 분석

- Kubernetes 로그의 대상이 이동하는 것과 비슷하게, Kubernetes에서 메트릭을 수집하는 것은 다음과 같음
  - 1) Kubernetes는 CPU, 메모리, 디스크 활용도, 디스크 및 네트워크 I/O와 같은 메트릭을 수집하여 모니터링해야 하는 다른 호스트에서 구성 요소를 실행합니다.
  - 2) 미니 VM의 일종인 Kubernetes 컨테이너도 자체 메트릭 세트를 생성합니다.
  - 3) 애플리케이션 서버와 데이터베이스 양쪽 모두 Kubernetes 포드로 실행될 수 있지만, 각 기술에는 관련 메트릭을 보고할 수 있는 고유한 방법이 있습니다.

# [실습] Docker Desktop에서 Kubernetes 설치





# Kubernetes Monitoring : Metrics

- **Host**

- Kubernetes Cluster를 구성하는 모든 Node의
- Host에 대한 CPU, 메모리, 디스크, 네트워크, 사용량과
- Node의 OS와 커널에 대한 모니터링을 함.

- **Container**

- Node에서 실행 중인 Container의
- CPU,메모리, 디스크, 네트워크 사용량 모니터링 함.

- **Application**

- Container에서 구동되는 개별 어플리케이션에 대한 모니터링.
- ex) 컨테이너에서 기동 되는 node.js기반의 애플리케이션의 http 에러 빈도 등

- **Kubernetes**

- Container를 컨트롤 하는 Kubernetes 자체에 대한 모니터링
- (Kubernetes의 자원인 Service, POD, 계정 정보 등이 이에 해당)

## **[실습 예제] Kubernetes Monitoring 실습**

**[https://github.com/JSJeong-me/EFK/tree/main/03\\_Day#readme](https://github.com/JSJeong-me/EFK/tree/main/03_Day#readme)**

# 호스트, Docker, Kubernetes 메트릭 수집

- 시스템(호스트) 메트릭 구성

```
system.yml: |-
- module: system
  period: 10s
  metricsets:
    - cpu
    - load
    - memory
    - network
    - process
    - process_summary
    - core
    - diskio
    # - 소켓
  processes: ['.*']
  process.include_top_n:
    by_cpu: 5      # CPU별 상위 5개 프로세스 포함
    by_memory: 5   # 메모리별 상위 5개 프로세스 포함
- module: system
  period: 1m
  metricsets:
    - filesystem
    - fsstat
  processors:
  - drop_event.when.regexp:
      system.filesystem.mount_point: '^(/sys|cgroup|proc|dev|etc|host|lib)($/|/)'
```

## ● Docker 메트릭 구성

```
docker.yml: |-
- module: docker
  metricsets:
    - "container"
    - "cpu"
    - "diskio"
    - "event"
    - "healthcheck"
    - "info"
    # - "이미지"
    - "memory"
    - "network"
  hosts: ["unix:///var/run/docker.sock"]
  period: 10s
  enabled: true
```

## ● Kubernetes 메트릭 구성

```
system.yml: |-
- module: system
  period: 10s
  metricsets:
    - cpu
    - load
    - memory
    - network
    - process
    - process_summary
    - core
    - diskio
    # - 소켓
  processes: ['.*']
  process.include_top_n:
    by_cpu: 5      # CPU별 상위 5개 프로세스 포함
    by_memory: 5   # 메모리별 상위 5개 프로세스 포함
- module: system
  period: 1m
  metricsets:
    - filesystem
    - fsstat
  processors:
  - drop_event.when.regex:
      system.filesystem.mount_point: '^(/sys|cgroup|proc|dev|etc|host|lib)($/)'
```

# Kubernetes 상태 메트릭 및 이벤트 수집

- Kube-state-metrics API와 통합되어 Kubernetes가 관리하는 객체의 상태 변화를 모니터링

```
kubernetes.yml: |-
  - module: kubernetes
    metricsets:
      - state_node
      - state_deployment
      - state_replicaset
      - state_pod
      - state_container
      # k8s 이벤트를 얻으려면 이 주석 처리 제거:
      - event
    period: 10s
    host: ${NODE_NAME}
    hosts: ["kube-state-metrics:8080"]
```

- Kubernetes 메트릭 Kibana 대시보드(예시)

## Kubernetes

