Elasticsearch with Python

ElasticSearch คือเครื่องมือค้นหาและวิเคราะห์ข้อมูลแบบกระจาย ที่มีพื้นฐานมาจาก Apache Lucene การรองรับภาษาต่างๆ ประสิทธิภาพที่สูง และเอกสาร JSON ที่ปราศจากสคีมาทำให้ Elasticsearch เป็นตัวเลือกที่เหมาะอย่างยิ่งสำหรับการวิเคราะห์บันทึกและกรณีใช้งานการค้นหาต่างๆ

Kibana คือเครื่องมือแสดงข้อมูลด้วยภาพและสำรวจข้อมูลที่ใช้สำหรับการวิเคราะห์บันทึกและอนุกรมเวลา การตรวจสอบแอปพลิเคชัน และการใช้งานความอัจฉริยะในการดำเนินการ ซึ่งมีคุณสมบัติประสิทธิภาพสูงแต่ใช้งานง่ายมากมาย เช่น ฮิสโตแกรม กราฟเส้น แผนภูมิวงกลม แผนภูมิความร้อน และการสนับสนุนภูมิสารสนเทศในตัว นอกจากนั้นยังมีการผสานรวมที่เหนียวแน่นกับเครื่องมือวิเคราะห์และค้นหายอดนิยมอย่าง Elasticsearch อีกด้วย ซึ่งทำให้ Kibana กลายเป็นตัวเลือกแรกๆ ในการแสดงข้อมูลที่อยู่ใน Elasticsearch ด้วยภาพ

วัตถุประสงค์:

- สามารถเขียน log ด้วย Python ไปยัง Elasticsearch ได้

กลุ่มเป้าหมาย:

- ผู้ดูแลระบบ
- Database Administrator (DBA)

ความรู้พื้นฐาน:

- พื้นฐานการออกแบบเว็บไซต์ด้วยภาษา HTML และ CSS
- เข้าใจแนวคิด โครงสร้างของเทคโนโลยี Python
- พื้นฐานการใช้งาน JSON

ระยะเวลาในการอบรม:

- 18 ชั่วโมง (3 วัน)

ราคาคอร์สอบรม:

- 12,500 บาท / คน (ราคานี้ยังไม่ได้รวมภาษีมูลค่าเพิ่ม)

วิทยากรผู้สอน:

- อาจารย์สนิทวงศ์ กมลภากรณ์

คอร์สที่ควรอบรมก่อนหน้า:

- Python Basic

เนื้อหาการอบรม:

Module 1: Introduction to Elasticsearch with Python

- Introduction to Elasticsearch
- Overview of the Elastic Stack
- Elasticsearch architectures

Module 2: Getting started

- Overview of installation options
- Running Elasticsearch & Kibana in Elastic Cloud
- Setting up Elasticsearch & Kibana on macOS & Linux
- Setting up Elasticsearch & Kibana on Windows
- Understanding the basic architecture
- Inspecting the cluster
- Sending queries with cURL
- Sharding and scalability
- Sharding
- Understanding replication
- Adding more nodes to the cluster
- Overview of node roles

Module 3: Managing Documents with Python

- Creating & deleting indices
- Indexing documents
- Retrieving documents by ID
- Updating documents
- Scripted updates
- Upserts
- Replacing documents
- Deleting documents
- Understanding routing
- How Elasticsearch reads data
- How Elasticsearch writes data
- Understanding document versioning
- Optimistic concurrency control
- Update by query
- Delete by query
- Batch processing
- Importing data with cURL

Module 4: Mapping & Analysis

- Introduction to Mapping
- Introduction to analysis
- Using the Analyze API
- Understanding inverted indices
- Introduction to mapping
- Overview of data types
- How the "keyword" data type works
- Understanding type coercion
- Understanding arrays
- Adding explicit mappings
- Retrieving mappings
- Using dot notation in field names
- Adding mappings to existing indices
- How dates work in Elasticsearch
- How missing fields are handled

- Overview of mapping parameters
- Updating existing mappings
- Reindexing documents with the Reindex API
- Defining field aliases
- Multi-field mappings
- Index templates
- Introduction to the Elastic Common Schema (ECS)
- Introduction to dynamic mapping
- Combining explicit and dynamic mapping
- Configuring dynamic mapping
- Dynamic templates
- Mapping recommendations
- Stemming & stop words
- Analyzers and search queries
- Built-in analyzers
- Creating custom analyzers
- Adding analyzers to existing indices
- Updating analyzers

Module 5: Searching for Data

- Introduction to searching
- Introduction to term level queries
- Searching for terms
- Retrieving documents by IDs
- Range searches
- Prefixes, wildcards & regular expression
- Querying by field existence
- Introduction to full text queries
- The match query
- Introduction to relevance scoring
- Searching multiple fields
- Phrase searches
- Leaf and compound queries
- Querying with boolean logic
- Query execution contexts
- Boosting query

- Disjunction max (dis_max)
- Querying nested objects
- Nested inner hits
- Nested fields limitations

Module 6: Joining Queries

- Introduction to Joining Queries
- Add departments test data
- Mapping document relationships
- Adding documents
- Querying by parent ID
- Querying child documents by parent
- Querying parent by child documents
- Multi-level relations
- Parent/child inner hits
- Terms lookup mechanism
- Join limitations
- Join field performance considerations

Module 7: Controlling Query Results

- A word on document types
- Specifying the result format
- Source filtering
- Specifying the result size
- Specifying an offset
- Pagination
- Sorting results
- Sorting by multi-value fields
- Filters

Module 8: Aggregations

- Introduction to aggregations
- Metric aggregations
- Introduction to bucket aggregations
- Document counts are approximate
- Nested aggregations
- Filtering out documents
- Defining bucket rules with filters
- Range aggregations
- Histograms
- Global aggregation
- Missing field values
- Aggregating nested objects

Module 9: Improving Search Results

- Introduction to this section
- Proximity searches
- Affecting relevance scoring with proximity
- Fuzzy match query (handling typos)
- Fuzzy query
- Adding synonyms
- Adding synonyms from file
- Highlighting matches in fields
- Stemming