

교육 과정 소개서.

한 번에 끝내는 데이터 엔지니어링 초격차 패키지 Online.

안내.

해당 교육 과정 소개서는 모든 강의 영상이 촬영하기 전 작성되었습니다.

* 커리큘럼은 촬영 및 편집을 거치며 일부 변경될 수 있으나, 전반적인 강의 내용에는 변동이 없습니다.

아래 각 오픈 일정에 따라 공개됩니다.

- 1차 : 2022년 05월 13일
- 2차 : 2022년 06월 17일
- 3차 : 2022년 07월 15일
- 4차 : 2022년 08월 19일
- 5차 : 2022년 09월 16일
- 6차 : 2022년 10월 14일
- 7차 : 2022년 11월 18일
- 8차 : 2022년 12월 16일
- 9차 : 2023년 01월 20일
- 최종 : 2023년 1분기

최근 수정일자 2023년 1월 27일



강의정보

강의장	온라인 강의 데스크탑, 노트북, 모바일 등
수강 기간	평생 소장
상세페이지	https://fastcampus.co.kr/data_online_engineering
담당	패스트캠퍼스 고객경험혁신팀
강의시간	80시간 예정 (* 사전 판매 중인 강의는 시간이 변경될 수 있습니다.)
문의	고객지원 : 02-501-9396 강의 관련 문의: help.online@fastcampus.co.kr 수료증 및 행정 문의: help@fastcampus.co.kr

강의특징

나만의 속도로	낮이나 새벽이나 내가 원하는 시간대에 나의 스케줄대로 수강
원하는 곳 어디서나	시간을 쪼개 먼 거리를 오가며 오프라인 강의장을 찾을 필요 없이 어디서든 수강
무제한 복습	무엇이든 반복적으로 학습해야 내것이 되기에 이해가 안가는 구간 몇번이고 재생



강의목표

- 데이터 엔지니어링 업무의 작은 조각을 보고 큰 그림을 그리고 이해하게 됩니다.
- 과제가 주어졌을 때 어떤 도구로 어떻게 처리하면 될 지 스스로 분석과 설계를 할 수 있습니다.
- 원천 데이터를 보고 데이터의 특징을 파악하고 특징에 따라 처리에 유의할 점을 알게 됩니다.
- 실시간 데이터 처리에서의 대표적인 문제와 해결방법을 이해할 수 있습니다.
- 대용량 데이터 분석시스템의 대표적인 문제와 해결방법을 이해할 수 있습니다.
- 시스템을 운영하기 위해서 필요한 지식과 도구를 배워 스스로 세팅하고, 활용할 수 있습니다.

강의요약

- Elasticsearch, Hadoop, HBase, Kafka, Storm, Flink, Presto, Druid 등의 툴들을 제대로 이해하고, 기술 스택을 구성하고 있는 원천 기술에 대한 원리도 함께 배울 수 있습니다.
- 데이터 엔지니어가 되기 위한 이론과 실무를 함께 배우고, 실제 빅테크기업에서 사용하는 도구, 상용 서비스에서 고민하는 Use Case 중심의 실습을 진행합니다.
- Data Engineering 기술은 기본적으로 서버사이드(백엔드) 기술을 포함합니다.
- 데이터 엔지니어도 알아야 하는 서버 엔지니어링에서의 기본적인 내용을 학습합니다.
- 수업에 사용했던 소프트웨어가 LTS(long term support) 기한이 지난 경우, 버전 업데이트에 맞춰 새롭게 리뉴얼 됩니다. (22.05.13~24.03.27)
- 하루 20억건 이상 대용량 데이터의 실시간 처리 파이프라인과 수십 페타바이트 데이터의 분석시스템을 개발하는 강사님께 강의를 듣고 질의 응답(22.05.13~24.03.27)을 할 수 있습니다.
- 스타트업부터 대기업까지 바로 활용할 수 있는 기술 스택과 활용방법 제시 뿐만 아니라 실시간 스트리밍, 대용량 분석시스템 기술 등 방대한분량을 가성비 있는 가격으로 평생 소장 할 수 있습니다.



강사

조이Joey

과목

- 데이터 엔지니어링

약력

- 현) SNS 대기업 Data Engineer
- 전) 티맥스 소프트 - 기업용 Java Framework 연구개발

CURRICULUM

오픈 분량.

Part 1 강의 로드맵	전체 오픈
Part 2 유틸리티 강의	Ch. 01 개발환경 세팅 ~ Ch. 05 Python 기초, Ch 07. JVM 이해하기 까지 오픈되었습니다. Ch 06. Scala 기초는 추후 업로드될 예정입니다.
Part 3 웹서비스와 데이터 기초	Ch. 01 웹 서비스와 데이터의 이해 까지 업로드되었습니다. Ch. 02 강의에서 활용할 예제 데이터 이해하기(필수 시청) 은 추후 업로드될 예정입니다.
Part 4 Server(Backend) Engineering 기초	전체 오픈
Part 5 Observability	전체 오픈
Part 6 분산 시스템 Fundamental	전체 오픈
Part 7 ElasticSearch(OpenSearch)	Ch 01. OpenSearch 구성 요소와 동작 방식 ~ Ch 04. OpenSearch의 인덱스와 샤드 관리, Ch 06. OpenSource Software와 기업관계 이해하기 전체 오픈되었습니다. Ch 05. 엔지니어링 설계와 의사결정 series - (2) 중견 기업 규모의 데이터 분석환경을 빠르게 구성하는 방법 중 (Ch 05_02. Opensearch 로 중견 기업 규모의 데이터 인프라 구성하기) 까지 오픈되었습니다.
Part 8. Hadoop. Hive. HBase	~ Ch 05. Hive 까지 오픈되었습니다.
Part 9 Batch, ETL, Workflow Scheduler	~ Ch 02. Batch와 ETL의 개념과 사례까지 오픈되었습니다.
Part 10. Distributed Streaming Processing	업로드 예정
Part 11. Spark	업로드 예정
Part 12. BigData OLAP	업로드 예정.
Part 13. 대용량 데이터 에코 시스템 구축하기	업로드 예정
Part 14. Engineer Way	업로드 예정

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

01.

초격차 데이터 엔지니어링 강의 로드맵

데이터 엔지니어 이해하기

채용공고로 살펴보는 데이터 엔지니어 직무 (1) - 대기업의 전사 시스템 수준의 데이터 엔지니어링

채용공고로 살펴보는 데이터 엔지니어 직무 (2) - 중견기업의 데이터 엔지니어

채용공고로 살펴보는 데이터 엔지니어 직무 (3) - 스타트업의 데이터 엔지니어

데이터 엔지니어링의 분야 소개

데이터 엔지니어 직무 F&Q

강의의 구성과 활용방법 (필수 시청)

강의의 특징, 수강 대상 안내

파트별 강의의 특징과 활용법

강의에서 사용하는 소프트웨어의 버전에 관한 정책

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

02. 유틸리티 강의

개발환경 세팅
java 개발환경
scala 개발환경
python 개발환경
python 개발환경 - pyenv와 virtualenv
Linux 기본 활용
Linux filesystem - 파일, 디렉토리 찾아보기 ls, mkdir,
Linux filesystem - inode, symbolic link
vi editor: 파일 만들고 수정하기
find, grep, pipe, tail, history: linux에서 원하는 파일 찾기, 로그파일 실시간으로 보기, 로그파일에서 원하는 내용 찾기
Linux OS: process, network 상태보기
환경변수란
shell과 script, 데이터 엔지니어가 shell script를 알아야 하는 이유
sh, bash shell - script 만들기
sh, bash shell - script에 파라미터 넘기기
AWS 활용 기초
실습에서 EC2 인스턴스 환경에서 하는 이유
EC2 인스턴스 만들기, key 파일 저장하기
EC2 인스턴스 접속하기 - Mac, Linux 사용자
EC2 인스턴스 접속하기 - Windows 사용자
예상 요금 계산하기
Java 기초
Java의 변수와 기본 자료형
Java의, class, method(함수) - 언어의 특징
Java 프로그램 컴파일, 실행하기
Java의 빌드 시스템 - maven
Java의 빌드 시스템 - gradle
Jar 파일 실행하기
Java 프로그램에 파라미터로 옵션주기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

02.

유틸리티 강의

Python 기초
스크립트 언어란
python의 변수와 기본 자료형
python의 function과 class
python 프로그램 설치하고 실행하기
Scala 기초
Scala와 Java의 공통점과 차이점
Scala의 변수와 기본 자료형
Scala의 class, method
Scala 스크립트로 사용하기
Scala 컴파일, 실행하기
Scala의 빌드시스템 - sbt
JVM 이해하기
JVM이란, JVM이 사용되는 이유
데이터 엔지니어링에서 Java, JVM이 중요한 이유
JVM의 메모리 구조
JVM의 GC
JVM 모니터링 - jmx exporter, jmx prometheus exporter
JVM 모니터링 - thread, thread pool
JVM 모니터링 - memory
JVM 모니터링 - GC
JVM 모니터링 - heap dump 뜨고 분석하기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

03.

웹서비스와 데이터 기초

웹 서비스와 데이터의 이해
인터넷과 웹
server-client 모델, API
웹서비스에서 발생하는 데이터 - use cases
서버 엔지니어가 다루는 데이터 - transaction, metadata
데이터 엔지니어가 다루는 데이터 - event, log, aggregation
IT회사가 데이터를 활용하는 방법
강의에서 활용할 예제 데이터 이해하기(필수 시청)
예제 시나리오
이벤트 데이터
로그 데이터

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

04.

Server(Backend) Engineering 기초

백엔드 엔지니어 되어보기 - 강의소개
백엔드 엔지니어 되어보기 - 강의소개
백엔드 엔지니어 되어보기 - RDBMS 사용하기
RDBMS(MySQL)를 이용해서 데이터 관리하기 - 이론, 예제
RDBMS(MySQL)를 이용해서 데이터 관리하기 - 실습 세팅
RDBMS(MySQL)를 이용해서 데이터 관리하기 - C
RDBMS(MySQL)를 이용해서 데이터 관리하기 - R
RDBMS(MySQL)를 이용해서 데이터 관리하기 - U
RDBMS(MySQL)를 이용해서 데이터 관리하기 - D
RDBMS(MySQL)를 이용해서 transaction 처리하기 - 이론
RDBMS(MySQL)를 이용해서 transaction 처리하기 - 실습 1
백엔드 엔지니어 되어보기 - NoSQL DB 사용하기
NoSQL(MongoDB)를 이용해서 데이터 관리하기 - 이론, 예제
NoSQL(MongoDB)를 이용해서 데이터 관리하기 - C
NoSQL(MongoDB)를 이용해서 데이터 관리하기 - R
NoSQL(MongoDB)를 이용해서 데이터 관리하기 - U
NoSQL(MongoDB)를 이용해서 데이터 관리하기 - D
백엔드 엔지니어 되어보기 - Cache DB로 응답시간을 빠르게 만들기
Cache DB (Redis)를 이용해서 데이터 관리하기 - 이론, 예제
Cache DB (Redis)를 이용해서 데이터 관리하기 - 실습 세팅
Cache DB (Redis)를 이용해서 데이터 관리하기 - Set, Get, Exist
Cache DB (Redis)를 이용해서 데이터 관리하기 - Sorting
DNS로 서비스 노출하기
DNS가 필요한 이유, DNS의 원리
AWS에서 DNS 사용하기
백엔드 엔지니어 되어보기 - LoadBalancer로 요청 분산하기
LoadBalancer가 필요한 이유
LoadBalancing 방법
AWS LoadBalancer 사용하기
LoadBalancer의 동작 원리, LoadBalancer 사용하기 전에 개발자가 해야할 일

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

04.

Server(Backend) Engineering 기초

백엔드 엔지니어 되어보기 - 웹서버 사용하기
웹서버가 필요한 이유
Nginx로 서비스별 요청 분산하기
Nginx로 ssl 인증서 인증하기
Nginx 설정 라이브 업데이트 하기
Nginx 로그의 이해와 활용
백엔드 엔지니어 되어보기 - API Gateway 사용하기
API gateway가 필요한 이유
Nginx API Gateway로 사용하기
AWS API Gateway 사용하기
백엔드 엔지니어 되어보기 - ELK로 서버 로그 수집하기
로그를 남기는 이유
로그를 남기는 법
파일비트로 로그 읽어서 보내기, 로그 레벨 필터링
파일비트가 로그를 읽는 방법
Logstash로 로그 수집, 변환, 보내기
Elasticsearch로 로그 저장하기
Kibana로 로그 시각화하기
데이터베이스(RDBMS) 구조와 동작방식 (MySQL을 기초로)
session과 transaction
RDBMS 내부 구조
RDBMS 동작 방식과, 튜닝 방법
RDBMS 가 indexing 하는 방식
MySQL의 HA
Data Modeling
데이터 모델링의 기초, 정규화 비정규화
파티셔닝
샤딩
엔지니어링 설계와 의사결정 series - (1) 내가 스타트업의 데이터 엔지니어라면? - 가장 빠르게 작은 규모의 분석환경 세팅하기
요구사항 분석의 기본
가장 빠르게 구성하는 작은 규모의 분석환경 아키텍처
아키텍처 세팅
연결관리와 모니터링

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.

해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

05. Observability

Observability 기초 개념
Observability vs Monitoring
Data Engineer가 모니터링해야할 대상
Observability를 위해 필요한 능력
Prometheus로 메트릭 수집하기
Prometheus 기초, push vs poll(scrap) 데이터 수집 방식의 변화
Prometheus 설치하기
Prometheus로 모니터링하기
Grafana 로 대시보드 구성하기
Grafana 설치하기
Grafana 로 대시보드 구성하기
host/network monitoring
모니터링을 위한 컴퓨터구조 기초
host/network monitoring 대상과 방법
Elastic Beats/PackageBeats 로 host/network 모니터링 하기
Prometheus node exporter 로 host/network 모니터링 하기
Kibana로 대시보드 구성하기
Grafana로 대시보드 구성하기
application/process monitoring
application monitoring의 대상과 방법, 모던 어플리케이션 모니터링 전략
application monitoring의 기본 - health check
Java application 모니터링
Kibana로 대시보드 구성하기
Grafana로 대시보드 구성하기
Application에서 내가 원하는 Metric 남기기 - Prometheus Client 사용하기
Prometheus Concept
Count로 traffic, throughput 모니터링하기
Histogram으로 latency 모니터링하기
메트릭 모니터링 응용해서 서비스 이상 가장 먼저 감지하기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

06.

분산 시스템 Fundamental

분산시스템의 이해

전통적인 서버 아키텍처의 역사

분산시스템이 필요한 이유

CAP theory, 분산시스템의 딜레마

분산시스템의 대표 use case

Zookeeper 이해하고 사용하기

Zookeeper 의 탄생 배경, 주요 기능

Zookeeper 실습 1 - Leader Election

Zookeeper 실습 2 - 분산 파일 상태 관리

Zookeeper의 Architecture 동작 원리

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.

해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

07.

ElasticSearch(OpenSearch)

Lucene이 indexing 하는 방법
inverted index 소개
Lucene search fundamental
OpenSearch의 구성과 동작 방식
ElasticSerach 설치하기
ElasticSearch의 CRUD
ElasticSerach의 구성요소
OpenSearch index 관리 방법
index 상태 살펴보기
index의 shard, replica
index lifecycle 관리하기
안전하게 index 변경하고 사용하기
OpenSearch cluster 운영
master, data node 분리하기, 모니터링
ingest node 운영하기, 모니터링
coordinator node 운영하기, 모니터링
Data aging. Hot-Warm-Cold 아키텍처
엔지니어링 설계와 의사결정 series - (2) 중견 기업 규모의 데이터 분석환경을 빠르게 구성하는 방법
요구사항 분석
OpenSearch로 분석환경을 구성한다면
대규모 분석환경을 위한 index, data 관리방법
OpenSource Software와 기업관계 이해하기
Elastic의 상장과 유료화, 그로 인한 생태계의 변화

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.

해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

08.

Hadoop, Hive, HBase

Hadoop의 이해
Hadoop의 필요성
Hadoop의 주요기능
Hadoop Architecture와 특징
Hadoop MapReduce
MapReduce 개념
MapReduce 실습 - Batch Data Aggregation
Hadoop을 사용할 때 주의할 점
Small File problem
replica 관리
commision, decommision
Hadoop, MapReduce 의 단점과 한계
Hadoop모니터링
기본 System, Java 모니터링
Hadoop File System 모니터링
Yarn의 이해
Resource Manager의 필요성
Yarn 컴포넌트와 Architecture 이해
Yarn에서 모니터링 해야할 것
Hive의 이해
metastore의 필요성
Hive 의 구성요소와 아키텍처
Hive 설치하기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

08.

Hadoop, Hive, HBase

Hive의 활용
Hive JDBC로 사용하기
Hive thrift 로 사용하기
Hive의 HA구성하기
HBase의 이해
Big Table 개념
HBase의 Architecture와 동작 방식 1 - 파일 관리
HBase의 Architecture와 동작 방식 2 - CRUD
HBase의 활용
테이블과 row key 디자인
row key design antipattern
HBase 에서 MapReduce, Column Family의 활용
HBase 모니터링
Hadoop으로서의 모니터링
Compaction 이해하고 세팅하기
HBase Use Case 1: SNS 서비스의 DB로 HBase 사용하기
요구사항 분석
테이블, 로우키 설계
HBase 데이터 적재, 조회 어플리케이션 구현
배치 작업 설계
HBase 배치작업과 데이터 활용 구현
HBase Use Case 2: 쇼핑몰 서비스의 DB로 HBase 사용하기
요구사항 분석
테이블, 로우키 설계
HBase 데이터 적재, 조회 어플리케이션 구현
배치 작업 설계
HBase 배치작업과 데이터 활용 구현
엔지니어링 설계와 의사결정 series - (3) 언제 Hadoop 쓰는게 좋을까?
Hadoop + eco system 요약
Hadoop의 장단점과 도입 전에 현실적으로 고려해야할 내용

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.

해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

09.

Batch, ETL, Workflow Scheduler

Cron Schedule
Cron Schedule 표현식
Linux에서 Cron Schedule 하기
Batch와 ETL의 개념과 사례
Batch와 ETL의 개념
java 프레임워크로 보는 Batch와 ETL 사례
Airflow 이해하기
소프트웨어의 발전의 흐름으로 이해하는 Workflow Scheduler 의 필요성
Airflow 아키텍처
Airflow 장단점
Airflow 설치하기
database 세팅하기
salary 세팅하기
job queue와 worker node 세팅하기
web admin 운영하기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

09.

Batch, ETL, Workflow Scheduler

DAG 개발하고 운영하기
DAG의 기본 구성요소
DAG 개발하기
Task 재사용하기
Operator로 공통화 하기
Admin을 통한 변수관리, 모니터링, 상태관리
Error handling
DAG 실습 1: HDFS File Merge 하기
요구사항
하드코딩으로 만들기
HDFS Operator 만들기
DAG 실습 2: Workflow 구성, Data Aggregation, Analysis, Report/Alarm
Sensor로 DAG간의 관계를 연결하기
요구사항
Airflow로 데이터 통합, 분석, 리포트 연결하기
Airflow 운영하기
Airflow 모니터링 대상과 메트릭 수집
Grafana Dashboard 구성하기
로그 보관기한 늘리기
Dedicated queue/worker 로 resource 보장하기
Provider로 빠르게 통합하기
Docker로 Task 환경 관리하기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

10.

Distributed Streaming Processing

Streaming 데이터 처리의 필요성
스트리밍 데이터 처리의 필요성
Kafka의 이해
분산 메세지 큐가 필요한 이유, 분산메세지 큐의 치명적 단점
Kafka Architecture
Kafka Producer
Kafka Producer API 사용법
Kafka Producer의 구조와 동작방식
Kafka Producer 모니터링하기
Kafka Producer의 성능을 높이는 법
대용량 데이터 처리시 Producer 에서 주의할 사항
Kafka Consumer
Kafka Consumer API 사용법 기초
Kafka Consumer의 구조와 동작방식
Kafka Consumer 모니터링하기
Kafka Consumer를 안정성 있게 운영하는 법
Kafka Consumer 에러 핸들링 방법
Kafka Streams
Kafka Streams API 사용해보기
Kafka Streams 사용할 때 주의할 점
단일 어플리케이션으로 streaming 데이터 처리하는 것의 한계
스트리밍에서 주요 장애 상황과 개념들
Kafka의 한계 + 단일 어플리케이션의 한계
Storm을 이용한 분산 Streaming 이론, 아키텍처
Storm Architecture와 Programming model
Storm Protocol, Consistent model 이해
Storm을 이용한 분산 Streaming 구현과 운영
Storm Application 만들기
분산된 환경에 Storm Application 배포하기
Storm 모니터링하기
Prometheus client를 이용해서 Storm 커스텀 모니터링 하기
Window Operation
Backpressure 처리하기
Storm의 상태 관리
Storm의 한계

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
 해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

10.

Distributed Streaming Processing

Flink의 아키텍처, 동작 원리, 기본 API
Flink Feature & Architecture 1 - 프로그래밍 모델
Flink Feature & Architecture 2 - Task 모델
Flink Dataset API 소개
Flink Dataset API 실습
Flink DataStream API 소개
Flink DataStream API 실습
Flink의 상태관리
Flink의 상태관리 실습
Flink를 이용한 분산 Streaming 구현과 운영
Flink Streaming Application 구현하기
Flink Batch Application 구현하기
Flink 모니터링하기
Prometheus client를 이용해서 Flink 커스텀 모니터링 하기
Flink로 실시간 aggregation하기
Flink failover 전략
Apache Beam으로 streaming처리 추상화하고 쉽게 구현하기
Beam 프로그래밍 방법
Beam 으로 Streaming Application 구현하고 배포하기
Lambda Architecture - 배치와 실시간의 통합
Lambda Architecture 개념과 사례
Lambda Architecture로 실시간 랭킹 서빙하기 실습 1 - 요구사항, 아키텍처
Lambda Architecture로 실시간 랭킹 서빙하기 실습 1 - 구현
KAPPA Architecture
Kafka Architecture 개념과 사례
KAPPA Architecture 실습 1 - 요구사항, 아키텍처
KAPPA Architecture 실습 1 - 구현
Streaming을 위한 File Format
Streaming에서 사용하는 FileFormat 종류와 장단점
protobuf 사용하기
avro를 위한 schema registry 환경 구축
avro 사용하기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

10.

Distributed Streaming Processing

Use Case 1: 데이터엔지니어링으로 시스템 고도화하기 - 대용량 log, event pipeline
Use Case 1 - 요구사항 분석
Use Case 1 - 아키텍처
Use Case 1 - 로그 수집기 구현
Use Case 1 - 로그 index lifecycle 구현
Use Case 1 - event validator 구현
Use Case 1 - event joiner 구현
Use Case 1 - 모니터링
Use Case 2: 데이터엔지니어링으로 서비스 수준 높이기 - 대용량 실시간 랭킹
Use Case 2 - 요구사항 분석
Use Case 2 - 아키텍처
Use Case 2 - 구현 1
Use Case 2 - 구현 2
Use Case 2 - 구현 3
Use Case 2 - 구현 4
Use Case 2 - 모니터링 1
Use Case 2 - 모니터링 2
CDC
CDC의 개념, CDC가 필요한 이유
Debezium과 Kafka를 이용한 CDC환경 구축

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

11.

Spark

Spark 언제쓰나요?
Spark 언제쓰나요?
Spark 기초
Spark 개념 이해하기 - Distributed Dataset
Spark의 구조, Architecture
RDD 기초
RDD 실습, 디버깅
Spark SQL, DataFrame, Dataset
Spark SQL의 동작원리
Spark SQL 실습
DataFrame의 특징, 장단점
DataFrame 실습
DataSet의 특징, 장단점
DataSet 실습
Spark로 배치 처리하기
Spark Driver의 이해
Client 모드로 배치잡 띄우기
Cluster 모드로 배치잡 띄우기
Spark Streaming으로 실시간 처리하기
Spark Structured Streaming 기초 - 개념
Spark Structured Streaming 기초 - 실습
Spark Structured Streaming 고급 - 개념
Spark Structured Streaming 고급 - 실습
Jupyter Notebook 설치하기
Jupyter Notebook 소개
Jupyter Notebook 설치하기
Spark로 데이터 분석하기
평균, 분산, 표준편차 구하기
데이터 오류 찾고 보정하기
Spark 어플리케이션 운영방법
Spark Job 모니터링
Suffle 튜닝하기
리소스 관리 방법
안정적인 Driver 운영을 위한 전략

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

12. BigData OLAP

OLAP vs OLTP
OLAP, OLTP 개념과 사례
Presto
Hive의 한계와 Presto의 대두
Presto Architecture
Presto의 Processing
Presto 설치하기
Presto 운영, 모니터링
OLAP Cube - Druid
OLAP Cube 개념
Start Schema
Druid Architecture
Druid의 Indexing, Processing
Druid 설치하기
Druid에 큐브 빌딩하기
Lambda Architecture 로 Druid로 실시간 큐브 서빙하기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

13.

대용량 데이터 에코 시스템 구축하기

Data Warehouse
Data Warehouse의 개념과 기술스택
Data Lake
Data Lake의 개념과 기술스택
내가 Data Lake를 만든다면?
Data Mart
Data Mart의 개념과 기술스택
내가 Data Mart를 만든다면?
Data Lakehouse
Data Lakehouse의 개념
Apache Hudi 소개
Apache Delta 소개
Apache Iceberg 소개
Hudi vs Delta vs Iceberg
Lakehouse 구현을 위해 필요한 것

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.
해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

CURRICULUM

14.

Engineer Way

완강 선물
완강 후 강의의 활용법
기술스택보다 중요한 것
스스로 성장하는 엔지니어되기

본 과정은 현재 촬영 및 편집이 진행되고 있는 **사전 판매 중인 강의**입니다.

해당 교육과정 소개서는 변경되거나 추가될 수 있습니다.

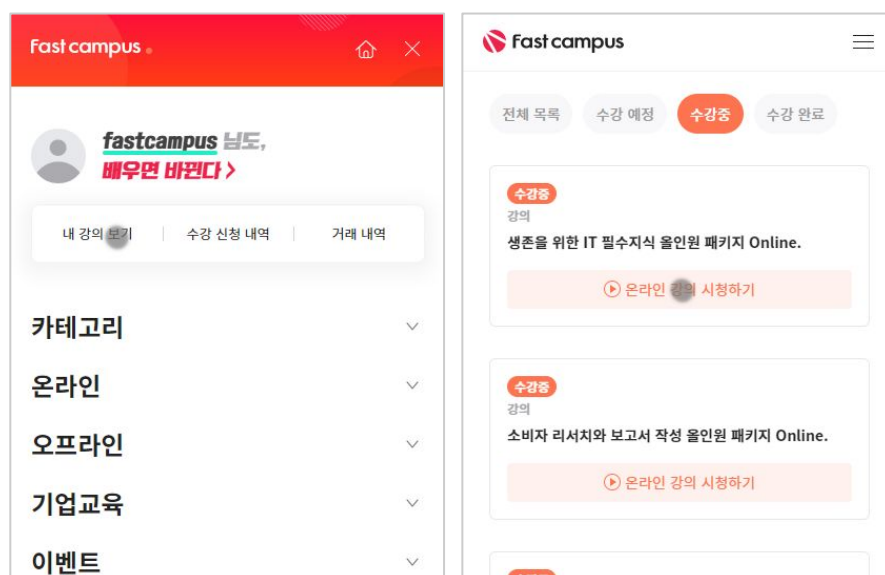


주의 사항

- 상황에 따라 사전 공지 없이 할인이 조기 마감되거나 연장될 수 있습니다.
- 패스트캠퍼스의 모든 온라인 강의는 **아이디 공유를 금지**하고 있으며 1개의 아이디로 여러 명이 수강하실 수 없습니다.
- 별도의 주의사항은 각 강의 상세페이지에서 확인하실 수 있습니다.

수강 방법

- 패스트캠퍼스는 크롬 브라우저에 최적화 되어있습니다.
- 사전 예약 판매 중인 강의의 경우 1차 공개일정에 맞춰 '온라인 강의 시청하기'가 활성화됩니다.



환불 규정

- 온라인 강의는 각 과정 별 '정상 수강기간(유료수강기간)'과 정상 수강기간 이후의 '복습 수강기간(무료수강기간)'으로 구성됩니다.
- 환불금액은 실제 결제금액을 기준으로 계산됩니다.

수강 시작 후 7일 이내	100% 환불 가능 (단, 수강하셨다면 수강 분량만큼 차감)
수강 시작 후 7일 경과	정상(유료) 수강기간 대비 잔여일에 대해 환불규정에 따라 환불 가능

※ 강의별 환불규정이 상이할 수 있으므로 각 강의 상세페이지를 확인해 주세요.