

교육 과정 소개서.

7개 프로젝트로 완벽하게 끝내는 AWS 데이터 파이프라인
구축



강의정보

| | |
|-------|---|
| 강의장 | 온라인 강의 데스크탑, 노트북, 모바일 등 |
| 수강 기간 | 평생 소장 |
| 상세페이지 | https://fastcampus.co.kr/data_online_awspipeline |
| 강의시간 | 20시간 |
| 문의 | 고객센터 |

강의특징

나만의 속도로 **낮이나 새벽이나 내가 원하는 시간대**에 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

- 클라우드 환경에서 데이터 파이프라인을 구축하면서 경로 별로 발생하는 장애/특수 상황에 바로 대응할 수 있는 데이터 수집부터 서버리스 파이프라인까지 7개의 파이프라인을 직접 구축합니다.
- 클라우드 환경에서 서비스의 품질을 계속 유지할 수 있는 파이프라인 최적화 노하우 학습합니다.
- 현업에서 데이터 파이프라인을 구축하면서 발생할 수 있는 7개의 실제 사례 기반 프로젝트를 진행합니다.

강의요약

- 데이터의 수집 경로 별로 실 서비스에서 발생할 수 있는 다양한 파이프라인 구조를 직접 구축해볼 수 있는 실습을 진행합니다.
- Web/APP Log 데이터를 처리하면서 실시간으로 발생한 데이터를 Batch/Stream 방식으로 수집하고 저장하는 방법을 학습합니다.
- AWS 환경에서 데이터 분석을 간소화하기 위한 EMR 클러스터를 학습하고 컴퓨팅 양을 조절하며 분석 과정을 최적화하는 방법을 학습합니다.
- 분산 환경에서 데이터를 처리할 수 있는 Presto를 학습하고 Presto와 Elastic Search를 연동하여 데이터 시각화를 학습해볼 수 있습니다.
- 클라우드 데이터 파이프라인을 구축하고 나서 데이터 분석을 더 용이하게 할 수 있는 방법들을 학습합니다.



강사

권낙주

과목

- 7개 프로젝트로 완벽하게 끝내는 AWS 데이터 파이프라인 구축

약력

- 현) 금융권 SI플랫폼 리더, 클라우드 빅데이터 아키텍처 전문가
- 전) SK C&C Bigdata 플랫폼, Public Cloud 전환 기획 및 아키텍트 책임자
- 전) 직방 데이터분석팀 데이터 아키텍트 책임자
금융권 데이터표준화 컨설팅
- 데이터 파이프라인 구축 오프라인 캠프 1-5기 강사
- 인덕대학교 겸임교수
- 고려사이버대 특강



CURRICULUM

01.

데이터 수집을
위한 파이프라인
구성하기

파트별 수강시간 07:57:21

| |
|---|
| CH01. 데이터 파이프라인의 이해 |
| 01. 강의의 구성 |
| 02. 데이터 아키텍처의 변화 |
| 03. 데이터 파이프라인 흐름의 이해 - 데이터 파이프라인을 위한 AWS 서비스 이해 |
| CH02. 실습 준비 및 Kafka 실습 |
| 01. 실습 준비를 위한 설명 및 EC2 생성 - EC2 서비스(1) |
| 02. 실습 준비를 위한 준비 - EC2 서비스 (2) |
| 03. Kafka Consumer 실습 |
| 04. Kafka 서버 설치 및 Producer 실행 |
| 05. Kafka 실습을 위한 EC2 실행 및 웹 서버 테스트 |
| CH03. AWS 서비스를 활용한 수집 파이프라인 구축 |
| 01. AWS 서비스를 활용한 데이터 수집에 대한 설명 |
| 02. API Gateway 구성하기 |
| 03. Kinesis Stream 구성하기 (1) |
| 04. Kinesis Stream 구성하기 (2) |
| 05. 수집 파이프라인 테스트 |
| CH04. 수집 미니 프로젝트 |
| 01. 프로젝트 개요 설명 |
| 02. Applog 수집 설명 및 구성도 |
| 03. Applog 수집을 위한 API-gateway 구성하기 |
| 04. Applog 수집을 위한 Kinesis 구성하기 |
| 05. Applog 수집 테스트 및 쉬운 로그 수집 방안 |
| 06. 외부 SaaS 서비스 수집 설명 및 구성도 |
| 07. 외부 SaaS 서비스 로그 수집 실습 및 테스트 |
| 08. 외부 SaaS 서비스 - Lambda를 이용한 Filtering 설명 |
| 09. 외부 SaaS 서비스 로그 수집 - Mariadb 생성 및 준비 |
| 10. 외부 SaaS 서비스 로그 수집 - Lambda 프로그램 배포 |
| 11. OpenAPI 공공 데이터 수집 설명 및 구성도 |
| 12. OpenAPI 수집 준비 |
| 13. OpenAPI 수집 Python 프로그램 테스트 |
| 14. OpenAPI 법정동 코드 FileToDB |
| 15. OpenAPI 저장 테이블 생성 |
| 16. OpenAPI Lambda 프로그램 DB에 데이터 입력 |
| 17. OpenAPI Lambda 배포 |
| 18. Web Scrapping 설명과 실습 |



CURRICULUM

02.

데이터 분석을
위한 파이프라인
구성하기

파트별 수강시간 01:34:07

| |
|---|
| CH01. 데이터분석 파이프라인 구성하기 |
| 01. 데이터 분석 아키텍처 설명 |
| 02. EMR 구조에 대한 이해 |
| 03. EMR 실행 |
| 04. EMR 실행 환경 및 구성 변경 |
| 05. EMR 퍼블릭 액세스 차단 |
| 06. EMR 모니터링 방안 |
| CH02. 데이터 처리 방법 |
| 01. Zeppeline을 활용한 데이터 처리 |
| 02. Spark-Semidata 핸들링하기 |
| 03. Spark-Textfile 핸들링하기 |
| 04. Spark-Mariadb 핸들링하기 |
| CH03. Glue의 DataCatalog 이해 |
| 01. 4주차 강의 실습 설명 |
| 02. AWS Glue의 이해 |
| 03. AWS Glue의 Crawler 실습 |
| 04. AWS Glue를 활용한 분석 Table 생성-설명 |
| 05. AWS Glue를 활용한 분석 Table 생성 (Silver Data) |
| 06. Hiva table를 활용한 데이터 처리 (Gold Data) |
| CH04. Zeppelin Cron 실습 |
| 01. EMR에서 Zeppelin Cron 설정 |
| CH05. 분석 미니 프로젝트 (GA 분석 내재화) |
| 01. 분석 미니 프로젝트 설명 |
| 02. 분석 미니 프로젝트 - 실습 데이터 생성 |
| 03. 분석 미니 프로젝트 - 기본 마케팅 분석 |



CURRICULUM

03.

데이터 분석 및 시각화

파트별 수강시간 02:26:22

| |
|--|
| CH01. 데이터 분석 및 시각화 Architecture |
| 01. 데이터 분석 및 시각화 Architecture 설명 |
| CH02. Presto를 활용한 분석 |
| 01. Presto 설명 |
| 02. Presto-Cli 실습 |
| 03. Presto-MySQL을 연동한 분석 |
| 04. Presto-Zeppelin 연동 및 분석 |
| CH03. Tableau 시각화 방안 |
| 01. Presto와 Tableau를 연동, 시각화 실습 |
| CH04. 시각화 미니 프로젝트(신규 오픈 시스템 모니터링) |
| 01. 시각화 미니 프로젝트 설명 |
| 02. 시각화 미니 프로젝트 실습 |

CURRICULUM

04.

데이터 파이프라인 Extension

파트별 수강시간 03:15:44

| |
|--|
| CH01. 데이터 파이프라인 Extension |
| 01. 데이터 파이프라인 Extension 설명 |
| CH02. EMRPyspark |
| 01. EMR 환경에서 Pyspark 실습 |
| CH03. Redshift 구성 및 실습 |
| 01. Redshift 설명 |
| 02. Redshift 생성 |
| 03. Redshift와 EMR 연동 실습 |
| 04. Redshift-Spectrum 실습 |
| CH04. Cloud DB Snowflake 활용 |
| 01. Snowflake 서비스 설명 및 Architecture 설명 |
| 02. Snowflake 실습을 위한 가입 및 설명 |
| 03. Snowflake 실습 - SemiData 분석 |
| 04. Snowflake 실습 - EMR 활용한 textData_Load |
| CH05. Serverless 데이터 파이프라인 구성 실습 |
| 01. Extension 미니 프로젝트 - 설명 |
| 02. Spark 실행과 Steps 실습 |
| 03. Airflow 통한 Serverless 실습 |

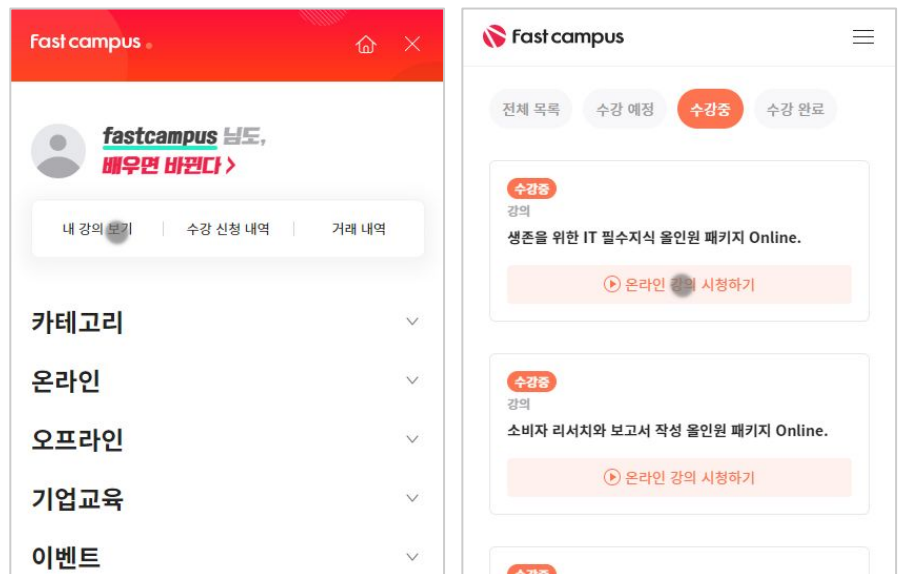


주의 사항

- 상황에 따라 사전 공지 없이 할인이 조기 마감되거나 연장될 수 있습니다.
- 패스트캠퍼스의 모든 온라인 강의는 아이디 공유를 금지하고 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.