

01_Introduction to cloud computing

Cloud Computing과 데이터의 중요성

- 클라우드 컴퓨팅의 영향: 클라우드 컴퓨팅은 사람과 데이터를 빠르고 쉽게 연결하며, 우리 삶의 모든 측면을 변화시키고 있음
- 비즈니스에서의 데이터: 오늘날 기업에게 데이터는 핵심적인 요소 → 기업들은 판매, 고객 피드백, 재고, 시장 조사 등 다양한 데이터에 의존하며, 데이터에 대한 끊임없는 접근이 필수
- **Cloud Data Analytics** 전문가의 역할: 기업이 고객을 이해하고, 미래를 위한 전략을 세우며, 위험을 완화하고, 더욱 유연해지도록 돕기 위해 **Cloud Data Analytics** 전문가에 대한 수요가 계속해서 증가 중

Google Cloud Data Analytics 프로그램 소개

- 프로그램 목표: 이 프로그램은 Cloud Data Analytics 분야의 초급 직무에 필요한 지식과 기술을 제공
- 강사진: Joey, Eric, Alyx, C.J., Christine과 같은 Google 직원들이 강사로 참여하여 업계 내부의 관점과 개인적인 경험을 공유
- 학습 자료 및 활동:
 - 콘텐츠: Cloud Data Analytics 개념과 기술을 가르치는 비디오 및 읽기 자료.
 - 실습: 개념과 기술을 연습할 수 있는 대화형 활동 및 랩. 랩은 여러 번 반복 가능
 - 평가: 이해도를 확인하는 퀴즈 및 용어집.
 - 경력 준비: 이력서 및 면접 준비를 돕는 자료.

코스 구성

- 코스 1: 클라우드 컴퓨팅 및 데이터 분석 개론 (Joey)
- 코스 2: 데이터 구조 및 구성. BigQuery, Google Cloud Storage, DataProc과 같은 클라우드 구성 요소를 활용한 대규모 데이터 세트 저장, 분석 및 처리 경험 습득 (Eric)
- 코스 3: 데이터의 여정(수집에서 통찰까지). 비즈니스 문제를 해결하기 위한 데이터 변환 전략 연습 (Alyx)
- 코스 4: 클라우드에서의 데이터 시각화. Looker를 사용하여 데이터 시각화 및 대시보드 구축 (C.J.)
- 코스 5: 종합 프로젝트(Capstone). 코스 1~4에서 배운 모든 것을 적용하여 프로젝트를 완성하고, 고용주에게 보여줄 수 있는 포트폴리오를 제작 (Christine)

Google Cloud Data Analytics Certificate Overview

Google Cloud Data Analytics Certificate 개요

- 목표: 이 프로그램은 클라우드 데이터 분석의 성장 분야를 탐색하고, 클라우드 환경에서 데이터가 어떻게 저장, 구성, 관리, 처리되는지 학습하는 과정. 데이터 시각화 및 프리젠테이션을 통해 인사이트를 공유하는 방법을 익히게 됨.
- 사전 경험: 클라우드 데이터 분석에 대한 사전 경험은 필요하지 않음.
- 전망:
 - 데이터 분석 분야는 지속적으로 성장하고 있으며, 데이터의 복잡성이 커지면서 숙련된 전문가의 필요성이 증가함.
 - 점점 더 많은 기업들이 클라우드 기반 서비스와 솔루션을 활용함에 따라, 클라우드 데이터 분석 전문가에 대한 수요가 높아지고 있음.

Google Cloud Data Analytics Certificate 코스 구성

- 총 5개의 코스로 구성되며, 클라우드 환경에서의 데이터 작업 기초, 데이터 저장, 구성, 처리, 분석, 인사이트 전달, 그리고 구직 리소스 등 핵심 개념과 기술을 다룸.
- **Introduction to Data Analytics in Google Cloud:** Google Cloud에서의 데이터 분석 개론.
- **Data Management and Storage in the Cloud:** 클라우드에서의 데이터 관리 및 저장.
- **Data Transformation in the Cloud:** 클라우드에서의 데이터 변환.
- **The Power of Storytelling: How to Visualize Data in the Cloud:** 스토리텔링의 힘: 클라우드에서 데이터를 시각화하는 방법.
- **Put It All Together: Prepare for a Cloud Data Analyst Job:** 모든 것을 종합하여 클라우드 데이터 분석가 직업을 준비하는 과정.
- **Capstone Project:** 마지막 코스에서는 배운 내용을 적용하는 종합 프로젝트를 진행하며, 잠재적 고용주에게 보여줄 수 있는 포트폴리오를 제작함.

수료생을 위한 혜택

- 구직 리소스: 5개 코스를 모두 이수하면 Google이 제공하는 구직 리소스에 접근 가능.
- **Google Cloud Affiliate Employers:** 미국 내 Google Cloud 제휴 고용주의 채용 공고를 검색하고 지원할 기회 제공.
- **Interview Warmup:** Google이 개발한 인터뷰 연습 도구를 통해 기술 향상 및 자신감 증진.
- 자격증 배지: Google Cloud Data Analytics Certificate 배지를 획득하여 LinkedIn®에 공유함으로써 다른 지원자들 사이에서 돋보일 수 있음.

Introduction to Course 1

Cloud Computing과 데이터 분석

- 클라우드 컴퓨팅의 탄생:

- 과거 기업은 거대하고 시끄러운 컴퓨터들로 가득 찬 방을 필요로 했지만, 1960년대부터 컴퓨팅 파워를 공유하는 아이디어가 제기됨.
- 오늘날 우리는 원격 데이터 센터를 통해 데이터 저장, 앱 실행, 분석 등을 할 수 있게 됨.
- **Cloud Computing**의 중요성: 더 많은 조직들이 클라우드 기반 솔루션을 채택함에 따라, 이러한 전환을 돕기 위한 숙련된 클라우드 전문가에 대한 수요가 증가하고 있음.
- 강사의 개인적인 경험:
 - 철학을 전공하고 전문 경험이 없던 강사가 **Google**에서 다양한 역할을 경험한 후, 데이터 분석 및 엔지니어링에 대한 열정을 발견함.
 - **HR** 분석가로 일하며 **SQL**을 배우고, 비즈니스 요청을 분석 문제로 전환하고 **SQL**을 사용해 데이터를 정리하여 비전문가도 이해하기 쉬운 데이터셋을 제공하는 과정을 즐김.

이 과정에서 배우는 내용

- **Cloud Computing**의 이해: 클라우드 컴퓨팅을 정의하고, 그 구성 요소를 식별하며, 클라우드와 기존 컴퓨팅의 차이를 구별하는 방법을 학습.
- 데이터 분석: 온프레미스(on-premises) 물리적 데이터 분석과 클라우드 데이터 분석을 비교하고, **Google Cloud Architecture Framework**를 중심으로 클라우드 데이터 분석이 비즈니스에 미치는 영향에 대해 배움.
- 데이터 관리 및 수명 주기: 데이터 관리의 내부 작동 방식과 데이터 수명 주기, 그리고 이 두 가지가 원활하게 작동하도록 하는 클라우드 데이터 분석가의 역할에 대해 탐구함.
- 협업 및 도구:
 - 클라우드 전문가들이 함께 멋진 비즈니스 프로젝트를 만들기 위해 어떻게 협업하는지 살펴봄.
 - 또한 클라우드 데이터 분석가의 주요 도구들을 발견하고, 클라우드 컴퓨팅에서 프로세스 관리의 중요성을 학습.
- 주요 도구: **BigQuery**와 **Dataproc**을 포함한 **Google** 클라우드 기반 도구들을 소개함.
- 결과: 이 과정을 완료하면 클라우드 데이터 도구에 대해 알게 되고, 클라우드의 이점을 이해하고 전달하며, 시기적절한 인사이트를 공유하는 방법을 익히게 됨.

Google Cloud Data Analytics Certificate의 첫 번째 코스 개요

- 코스 1 목표:
 - 이 코스(Introduction to Data Analytics in Google Cloud)는 **Google Cloud Data Analytics Certificate** 프로그램의 첫 번째 과정임.
 - 이 코스에서는 프로그램에 대한 이해를 높이고, 앞으로의 과정에 대한 기대치를 설정하며, 성공적인 학습 방법을 익히는 데 중점을 둠.
- 주요 학습 내용:
 - 클라우드 컴퓨팅의 기초 및 구성 요소를 탐색하고, 클라우드 컴퓨팅과 기존 컴퓨팅의 차이점을 학습함.

- 데이터 관리, 데이터 라이프사이클, 그리고 이 주제들이 클라우드 데이터 분석가의 역할과 어떻게 관련되는지 심층적으로 학습함.
- 클라우드 데이터 전문가들이 프로젝트에서 협업하여 주요 비즈니스 솔루션을 생성하는 방법을 탐색함.
- 클라우드 데이터 분석가에게 필요한 주요 도구들을 소개받고, 클라우드 컴퓨팅에서 프로세스 관리의 중요성을 학습함.

Google Cloud Data Analytics Certificate 프로그램 과정

- 프로그램은 총 5개의 코스로 구성되어 있음.
- 코스 1: Introduction to Data Analytics in Google Cloud (현재 코스)
 - 클라우드 데이터 분석 실무의 기초를 탐색
 - 데이터 수집, 저장, 처리, 시각화를 포함한 클라우드 데이터 분석가의 역할과 책임에 대해 학습함.
- 코스 2: Data Management and Storage in the Cloud
 - 데이터가 어떻게 구조화되고 구성되는지 배움
 - BigQuery, Google Cloud Storage, DataProc과 같은 클라우드 구성 요소를 활용하여 대규모 데이터 세트를 효율적으로 저장, 분석, 처리하는 실습 경험을 얻음.
- 코스 3: Data Transformation in the Cloud
 - 데이터 품질 및 유효성을 검사하고, 다양한 데이터 소스 간의 관계를 설정하며, 데이터 변환 방법을 탐색함.
 - PII(개인 식별 정보)를 안전하게 보호하기 위한 데이터 거버넌스 및 보안 모범 사례를 적용하는 방법을 배움.
- 코스 4: The Power of Storytelling: How to Visualize Data in the Cloud
 - 클라우드에서 데이터를 시각화하는 5가지 주요 단계(스토리텔링, 계획, 데이터 탐색, 시각화 구축, 데이터 협업)를 학습함.
 - UI/UX 기술을 활용하여 상호작용적이고 영향력 있는 클라우드 기반 시각화를 와이어프레임하고, 클라우드 네이티브 데이터 시각화 도구로 보고서와 대시보드를 만들.
- 코스 5: Put It All Together: Prepare for a Cloud Data Analyst Job -
 - 이전 코스에서 배운 기초 지식과 기술을 종합하여 전체 데이터 라이프사이클에 초점을 맞춘 종합 프로젝트(capstone project)를 제작함.
 - 이력서를 업데이트하고 면접 기술을 연습하여 취업을 준비함.

학습 자료 및 성공을 위한 팁

- 학습 자료: 구글 강사가 진행하는 비디오, 관련 개념을 소개하는 읽기 자료, 실습을 위한 활동 및 랩, 주요 용어 목록인 용어집, 이해도를 확인하는 퀴즈 등이 제공됨.
- 평가:
 - 각 과정의 주요 개념에 대한 이해를 증명하기 위해 채점되는 퀴즈가 있으며, 수료를 위해서는 80% 이상의 점수를 받아야 함.
 - 통과 점수를 받을 때까지 여러 번 응시할 수 있음.
- 성공을 위한 팁:
 - 각 레슨의 항목은 순서대로 진행하는 것이 좋음.

- 모든 학습 기회에 참여하여 최대한 많은 지식과 경험을 얻는 것이 권장됨.
- 이해가 안 되는 부분이 있다면 비디오를 다시 보거나, 읽기 자료를 검토하거나, 활동을 반복하는 것을 망설이지 말아야 함.
- 유용한 링크를 발견하면 북마크하여 나중에 참고할 수 있도록 함.
- 완료 예상 기간: 매주 1~2개 모듈을 완료할 경우, 전체 과정을 약 3~6개월 내에 마칠 수 있음.

Google Cloud Data Analytics Certificate 취득 조건

- 자격증 취득 조건: Google Cloud Data Analytics Certificate를 취득하려면 프로그램의 모든 코스를 완료해야 함.
- 평가 기준:
 - 자격증을 받기 위해서는 프로그램 내 모든 채점 과제에서 최소 **80%** 이상의 점수를 받아야 함.
 - 각 코스의 채점 퀴즈는 해당 코스의 누적 점수에 포함됨.

성공적인 학습 습관 및 팁

- 학습 계획:
 - 매주 규칙적인 학습 시간을 정하고 따르는 것이 좋음.
 - 달력을 사용하여 일정을 만들고 달성 가능한 목표를 설정할 수 있음.
- 보충 콘텐츠 활용:
 - 이 자격증은 **SQL**, 데이터 정리, 분석 및 시각화, 비즈니스 인텔리전스 등 기본적인 데이터 분석 개념에 대한 사전 지식이 있으면 더 효과적임.
 - 사전 지식이 없는 학습자를 위해 선택적 보충 콘텐츠가 제공됨.
- 탐구심: 흥미로운 아이디어가 생기면 질문하고, 온라인으로 더 많은 정보를 검색하며, 관심 있는 링크를 탐색하고, 발견한 내용을 메모하는 것이 좋음.
- 노트 필기: 중요한 정보를 기억하고, 주제들 간의 연관성을 파악하는 데 효과적임.
- 모범 답안 (**Exemplars**) 참고: 많은 활동에 모범 답안이 제공되므로, 이를 통해 자신의 작업이 기준을 충족하는지 확인하거나 오류를 점검할 수 있음.
- 링크 새 탭에서 열기: **Google Cloud Skills Boost** 플랫폼을 벗어나지 않도록 모든 하이퍼링크(활동 템플릿, 모범 답안, 외부 웹사이트 등)는 새 탭에서 여는 것이 권장됨.
- 면접 준비:
 - **Interview Warmup**은 면접 질문을 연습하고 실시간 피드백을 받아 면접에 대한 자신감과 편안함을 높이는 데 도움이 되는 도구임.
 - 자격증 과정 전반에 걸쳐 이 도구를 사용해 볼 기회가 주어짐.
- 활동에 필요한 소프트웨어: 프로그램의 일부 활동을 완료하기 위해 **Google Docs**, **Google Sheets**, **Google Slides**와 같은 디지털 문서 및 스프레드시트, 프리젠테이션 도구를 사용해야 함.
- 용어집 활용: 모듈, 코스, 그리고 전체 자격증 프로그램에 대한 용어집이 제공되므로, 채점 퀴즈를 준비하거나 향후 참고 자료로 활용할 수 있음.

Lab 실습을 위한 기술 팁

- **Lab 개요:** 각 Lab에는 활동 개요, 시나리오, 시작 지침, 그리고 각 과제를 완료하기 위한 지침이 포함된 Qwiklabs 랜딩 페이지가 있음.
- **시간 제한:** Lab 시작 버튼을 누르면 **90분** 타이머가 시작되며, 타이머는 일시 정지할 수 없으므로 이 시간 안에 모든 과제를 완료해야 함.
- **실습 횟수:**
 - **Lab**은 사전 구성된 가상 머신이나 프로비저닝된 리소스에 의존하기 때문에 총 **5번**만 완료할 수 있음.
 - 하지만 랜딩 페이지는 횟수 제한 없이 참고 가능함.
- **진행 상황 점검:**
 - 각 **Lab**에는 진행 상황 점검 및 객관식 질문이 포함되어 있으며, 할당된 시간 내에 무제한으로 시도할 수 있음.
 - 처음 시도에 모든 문제를 맞히지 못하는 것은 흔한 일이며, 이는 학습 과정의 일부임.
- **실습 환경:** Lab은 모바일 기기가 아닌 데스크톱 또는 노트북 컴퓨터에서 완료하도록 설계됨.

Lab 성공을 위한 팁

- **사전 준비:** Lab 타이머를 시작하기 전에 관련 가이드, 비디오 및 코스 자료를 미리 검토하여 시간을 절약해야 함. 실습을 방해받지 않도록 준비해야 함.
- **계정 사용:**
 - 개인 계정과 학생 계정 간의 충돌을 방지하기 위해 시크릿 모드 또는 개인 브라우저 창을 사용하여 Lab을 완료해야 함.
 - 개인 Google Cloud 계정이 있더라도 Lab을 위해 사용하면 추가 요금이 발생할 수 있으므로 사용하지 않아야 함.
- **화면 구성:** Lab 지침과 Lab 환경을 별도의 브라우저 창에 나란히 배치하면 쉽게 전환할 수 있음.
- **로딩 시간:** Lab에 필요한 리소스가 로드되거나 프로비저닝되는 데 시간이 걸릴 수 있으며, 이 경우 브라우저에서 벗어나지 않아야 함.

Lab의 핵심 의의

- **실습 환경:** 이 프로그램의 Lab은 학습한 도구와 개념을 적용하는 안내식 실습 환경을 제공함.
 - **목표:** 실제 현장 경험을 시뮬레이션하도록 설계되어, 과제를 완수하고 자신의 능력에 대한 자신감을 키울 수 있음.
 - **기술 문제 해결:** 기술적인 문제가 발생하면 이 문서를 검토하여 해결하는 데 도움을 받을 수 있음.
-

Explore your course 1 scenario: TheLook eCommerce

- 목표:
 - 클라우드 데이터 분석의 성장 분야를 탐색하고, 클라우드 환경에서 데이터가 저장, 구성, 관리, 처리되는 방법을 학습함.
 - 또한, 산업 전문가들이 사용하는 도구를 배우고, 데이터 시각화 및 프리젠테이션을 통해 인사이트를 공유하는 기술을 개발함.
 - 이 프로그램은 사전 경험이 없어도 시작할 수 있음.
- 프로그램 구성:
 - **Introduction to Data Analytics in Google Cloud:** 클라우드에서의 데이터 분석 기초를 탐색함.
 - **Data Management and Storage in the Cloud:** 데이터가 어떻게 구조화되고 구성되는지 배우고, BigQuery, Google Cloud Storage, DataProc 등 클라우드 구성 요소를 활용하는 실습 경험을 얻음.
 - **Data Transformation in the Cloud:** 데이터 품질을 검사하고, 변환 전략을 적용하며, 데이터 거버넌스 및 보안 모범 사례를 학습함.
 - **The Power of Storytelling: How to Visualize Data in the Cloud:** 클라우드에서 데이터를 시각화하는 주요 단계(스토리텔링, 계획, 탐색, 시각화 구축, 협업)를 배우고, 클라우드 네이티브 시각화 도구를 활용함.
 - **Put It All Together: Prepare for a Cloud Data Analyst Job:** 이전 코스에서 배운 모든 지식과 기술을 적용하는 종합 프로젝트(capstone project)를 제작하고, 취업 준비를 위한 리소스를 활용함.
- 수료생을 위한 혜택:
 - 모든 코스를 완료하면 Google이 제공하는 구직 리소스에 접근할 수 있음.
 - Google Cloud 제휴 고용주의 채용 공고에 지원할 수 있고, Google이 개발한 Interview Warmup 도구로 면접 기술을 연습할 수 있음.
 - LinkedIn®에 공유하여 잠재적 고용주에게 자신의 역량을 보여줄 수 있는 자격증 배지를 받을 수 있음.

코스 1 시나리오: The Look - eCommerce

- 역할: The Look - eCommerce라는 가상의 회사에서 클라우드 데이터 분석가로 일하는 역할임.
- 주요 목표: 클라우드 기반 데이터 분석 솔루션을 구축하고 유지하며, 조직의 여러 팀과 협업하여 데이터 기반 의사 결정을 지원함.
- 주요 과제:
 - 데이터 라이프사이클 계획 설계: 이해관계자의 피드백을 바탕으로 효과적인 데이터 라이프사이클을 만들고, 데이터 소스를 추천하며, 데이터 변환을 제안하는 활동을 함.
 - **BigQuery** 탐색: The Look - eCommerce의 마케팅 부서를 위해 판매 데이터를 활용하여 보고서를 만들고, 테이블을 탐색하며 데이터를 필터링하여 정보를 추출하는 Lab을 진행함.
- 협업 팀원: 마케팅 매니저 Martina, 상품 책임자 Meredith, 데이터 아키텍트 Artem, 데이터 과학자 Huan, 재무 부사장 Kiran, 재무 책임자 Trevor와 같은 다양한 팀원들과 협업함.

- **핵심 의의:** 이 활동과 **Lab**은 클라우드 데이터 분석가의 역할과 업무를 시뮬레이션하는 환경에서 학습한 기술을 연습하고 적용하는 것을 목표로 함.
-

클라우드 데이터 분석 소개

- **클라우드 데이터 분석의 기본:** 이 모듈에서는 클라우드 데이터 분석의 기본을 살펴보고, 클라우드 컴퓨팅의 역사와 다양한 이점을 탐구함.
- **클라우드 vs. 기존 컴퓨팅:** 클라우드 컴퓨팅과 기존 컴퓨팅의 차이점, 즉 네트워크 인프라, 특징, 장점 및 한계를 비교하여 검토함.
- **성공적인 학습 계획:** 프로그램 리소스를 활용하여 성공적인 학습 계획을 세우고, 경력 준비를 할 수 있도록 도움.
- **주요 용어 학습:** 학습 내용을 마무리하면서 주요 용어와 정의가 포함된 용어집을 숙지함.

Google 분석 매니저 **Joey**의 이야기

- **직무 및 역할:** Google의 분석 매니저로서 비즈니스 분석가 팀을 이끌며, 데이터 기반 인사이트를 제공하여 주요 비즈니스 의사결정을 지원함.
- **프로젝트 경험:** 미래 직장 설계 프로젝트를 통해 건축가, 프로그램 매니저, UX 팀과 협력하여 **Googler**를 위한 새로운 사무실 경험을 설계함. 데이터를 활용하여 약 20만 명의 사용자가 필요로 하는 사무실 공간과 서비스를 구현함.
- **배경 및 경력:** 로스앤젤레스에서 멕시코계 미국인으로서 한부모 가정에서 성장함. 다양하고 복잡한 도시 생활은 다양한 그룹과 소통하는 열정을 심어주었고, 대인 관계 기술을 키우는 데 도움이 되었음.
- **비전형적 경력 경로:** 철학을 전공하고 피자 가게에서 3년간 일했지만, Google에서의 비기술적 인턴십을 통해 데이터 분석 및 엔지니어링에 대한 열정을 발견함. HR 운영, K-12 교육 봉사, 비즈니스 인텔리전스 등 다양한 분야를 경험하며 기술적인 역할로 전환함.
- **개인의 목표:** 소외된 집단을 위한 롤 모델이 되는 것이 목표임. 사람들이 새로운 열정을 발견하고 불가능하다고 생각했던 것을 성취하도록 영감을 주고 싶음.
- **데이터에 대한 관점:** 분석가로서 데이터는 모두를 위한 것이며, 기술적인 작업을 덜 위협적이고 더 접근하기 쉽게 만들고 싶다는 목표를 가지고 있음.

Google의 **Senior Vice President Ben**의 이야기

- **학습의 중요성:**
 - Google의 **Learning and Sustainability** 부문 수석 부사장인 Ben은 배움이 사람들을 다른 지점으로 이끌어 주는 중요한 수단이라고 믿음.
 - 교육 덕분에 꿈도 꾸지 못했던 곳에 갈 수 있었고, 모두가 그러한 기회에 접근할 수 있어야 한다고 생각함.
- **성장 배경:** 인도에서 자랐으며, 부유한 가정은 아니었지만 좋은 학교에 다닐 수 있었고, 이것이 자신의 삶에 큰 변화를 가져왔음을 깨달음.
- **기술의 역할:** 기술의 도움으로 더 많은 사람에게 더 많은 기회를 제공할 수 있다고 믿음.

- 클라우드 **Computing**의 중요성:
 - 클라우드는 컴퓨팅이 나아가야 할 방향이므로 매우 중요함.
 - 우리가 사용하는 거의 모든 주요 제품이 이제 온라인 클라우드 데이터 센터를 기반으로 하고 있으며, 이를 통해 사용자에게 놀라운 서비스를 제공함.
 - 클라우드 기술 학습의 가치: 클라우드 기술을 배우는 것은 전 세계 사람들이 사용하는 강력한 시설을 클라우드에 구축하는 데 필요한 일자리와 기회가 있는 전체 경제에 참여할 수 있게 해줌.
 - 기술 교육의 진화:
 - 교육은 사람들이 다양한 스킬링 코스를 통해 개별적인 기술을 배우도록 진화하고 있음.
 - 모든 사람이 모든 교육에 접근할 수 있는 것은 아니지만, 스킬링을 통해 필요에 따라 교육의 부분들을 쌓아 나갈 수 있음.
 - 이는 시간과 자원에 맞춰 교육을 받을 수 있는 방법임.
 - 기본의 중요성: 무엇이든 배우는 초기 단계는 기초와 기본을 배우는 것임. 인내심을 가지고 배우면 흥미로운 점을 발견하게 됨.
 - 컴퓨터 경력: 30년 넘게 컴퓨터와 함께 일했지만, 여전히 매일매일이 흥미롭고 매혹적이라고 언급함.
-

Cloud Computing의 이해

- **Cloud Computing**이란:
 - 인터넷을 통해 서비스로 호스팅되는 온디맨드(on-demand) 컴퓨팅 리소스를 사용하는 것.
 - 기업들이 직접 리소스를 찾고, 설정하고, 관리할 필요가 없으며, 사용한 만큼만 비용을 지불함.
- 작동 방식:
 - **Cloud Computing**은 네트워크를 사용하여 사용자를 클라우드 플랫폼(가상 공간)에 연결함.
 - 주 컴퓨터가 장치와 서버 간의 모든 통신을 처리하여 정보를 공유하고, 개인 정보 보호 및 보안 조치를 통해 모든 것을 안전하게 보호함.
- 비유:
 - 클라우드 컴퓨팅은 공동 주방에 비유할 수 있음. 즉 임차인이 개인 주방용품을 살 필요 없이 잘 갖춰진 주방을 사용하는 것과 비슷함
 - 기업은 자체 스토리지, 하드웨어, 소프트웨어를 구매하고 유지보수하는 데 시간과 비용을 들이지 않고도 온디맨드 방식으로 컴퓨팅 리소스에 접근 가능함

Cloud Computing 인프라의 주요 구성 요소

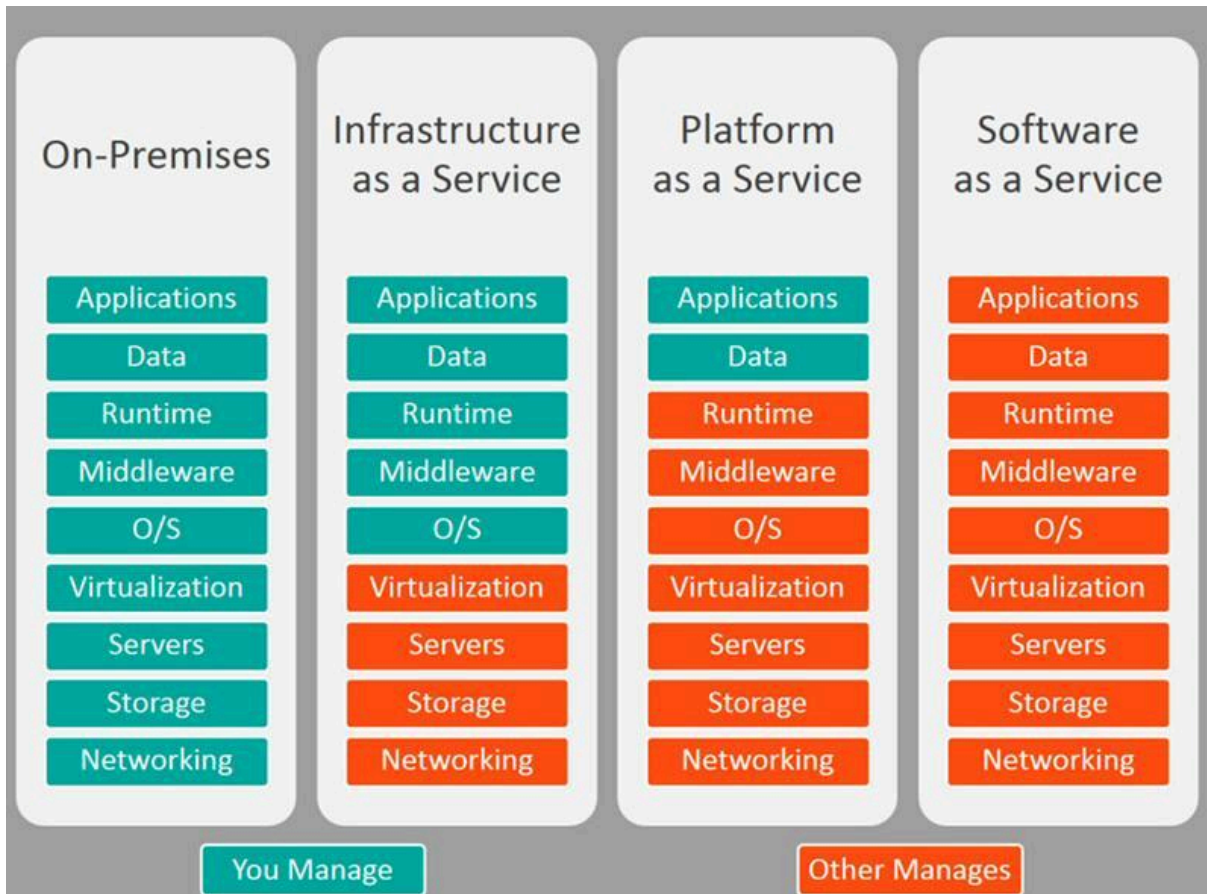
- 하드웨어(**Hardware**): 서버, 프로세서, 메모리, 네트워크 스위치, 라우터, 케이블, 방화벽, 로드 밸런서, 냉각 시스템, 전원 공급 장치 등 물리적 항목을 포함함.
- 스토리지(**Storage**):
 - 파일(**File**) 스토리지:

- 파일들을 폴더 계층 구조로 정리하여 한 곳에 저장함.
 - 가장 오래되고 널리 사용되지만, 기능이 다소 제한적임.
 - 오브젝트(**Object**) 스토리지:
 - 비정형 데이터와 그 메타데이터(데이터에 대한 데이터)를 저장함.
 - 예를 들어, 스마트폰 사진에 위치, 날짜 등 정보를 담는 것과 같음.
 - 블록(**Block**) 스토리지:
 - 대용량 데이터를 최적으로 정리된 작은 조각으로 나누고 고유한 레이블을 부여함.
 - 데이터 접근이 용이하지만, 비용이 비싸고 메타데이터 처리 기능이 제한적임.
 - 네트워크(**Network**):
 - 클라우드 인프라가 백엔드 리소스에 연결되는 수단.
 - 물리적 네트워크 설정이 가상 네트워크를 작동시키는 기반이 됨.
 - 가상화(**Virtualization**):
 - 서버, 스토리지, 네트워크 등 물리적 인프라의 가상 버전을 생성하는 기술.
 - 이 기술 덕분에 연결 없이도 서비스가 작동할 수 있음.
-

Google Cloud Data Center와 클라우드 분석가

- 데이터 센터:
 - 서버, 컴퓨터 시스템 및 관련 구성 요소를 포함하는 물리적 건물.
 - 방대한 양의 데이터를 위한 중앙 집중식 위치를 제공함.
- 클라우드 분석가:
 - 클라우드 분석가들은 클라우드를 통해 이 데이터를 활용하여 다양한 비즈니스 작업과 프로젝트를 수행함. 관련 데이터를 선택하고 추출한 다음, 처리 및 조사를 위해 준비함.
 - 또한 데이터를 분석하고, 시각화하며, 공유하여 가치 있는 인사이트를 발견하고 현명한 비즈니스 의사 결정을 내림.

Cloud Service Model 3가지



- **IaaS (Infrastructure as a Service)**

- 개념:
 - 하드웨어, 스토리지, 네트워크, 가상화 도구 등 IT 인프라 서비스를 온디맨드 방식으로 제공하는 모델.
 - 서비스 제공자가 인프라를 호스팅하고 유지보수하며, 사용자는 운영 체제, 데이터, 애플리케이션 등 나머지를 모두 관리함.
- 특징: IT 리소스에 대한 가장 높은 수준의 제어 권한을 제공하며, 기존 온프레미스 IT와 유사함.
- 예시: 온라인 폴더에 이메일을 정렬하는 클라우드 스토리지. 리스한 자동차처럼 사용자는 인프라를 사용하지만 소유하지는 않음.

- **PaaS (Platform as a Service)**

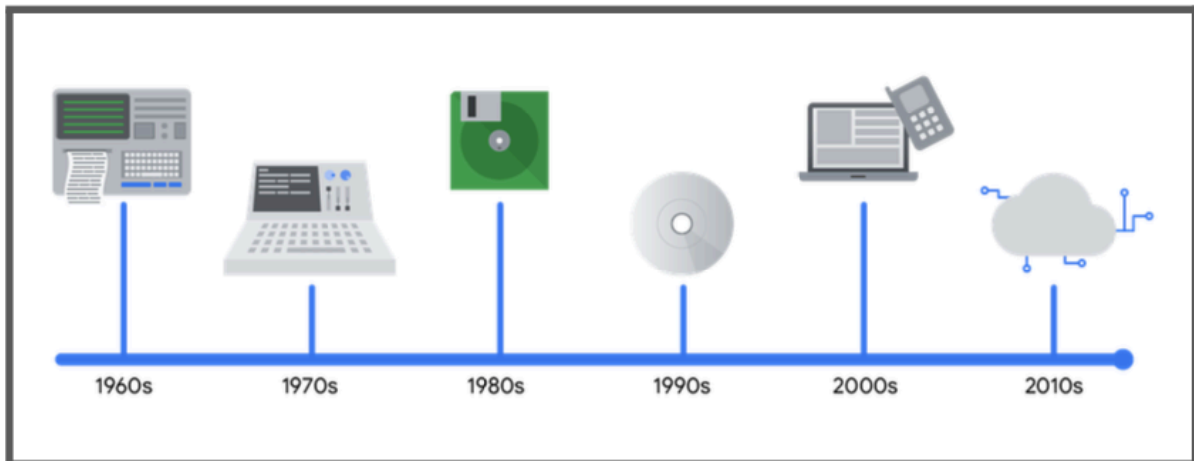
- 개념: 클라우드 애플리케이션 개발을 위한 하드웨어 및 소프트웨어 도구를 제공하여 개발 과정을 단순화함.
- 특징: 사용자는 기본 인프라 관리 부담 없이 앱 개발에만 집중할 수 있음.
- 예시:
 - 택시를 타고 목적지를 말하는 것과 유사함.
 - 사용자가 직접 운전하지 않고(인프라 관리), 운전자가(플랫폼) 사용자를 목적지(앱 개발)까지 데려다주는 것을 신뢰함.

- **SaaS (Software as a Service)**

- 개념: 인프라, 유지보수, 업데이트 및 애플리케이션 자체를 포함한 완벽한 소프트웨어 패키지를 구독 방식으로 제공함.
- 특징: 사용자는 인터넷을 통해 앱에 연결하기만 하면 되며, 다른 사용자와 같은 서비스를 공유함.

- 예시: 버스를 타는 것과 유사함. 정해진 노선(설정된 서비스)에서 정류장을 선택하고, 다른 사람들과 버스를 함께 사용함.
- 비유적 표현: **IaaS, PaaS, SaaS**에 대한 비유는 각 모델이 제공하는 개별적인 맞춤화 수준을 설명하기 위한 것이며, 실제 하드웨어 또는 소프트웨어의 세부 사항을 의미하지는 않음.

클라우드 컴퓨팅의 역사



- **1960년대:** 컴퓨터는 방 크기였고, 여러 사람이 사용할 수 있는 컴퓨터 개발이 시작됨.
- **1970년대:**
 - 가상 머신(**virtual machines**)의 개념이 정립
 - **ARPANET**에서 네트워크를 나타내는 상징으로 '클라우드' 기호가 처음 사용됨.
- **1980년대:** 네트워크를 통해 파일에 접근할 수 있게 해주는 네트워크 파일 시스템 프로토콜(**network file system protocol**) 개발이 시작됨.
- **1990년대:**
 - 통신 산업 등에서 공유 컴퓨팅(**shared computing**)이 대중화됨.
 - 이 시기 말에는 많은 기업이 공유 컴퓨팅을 사용하게 됨.
- **2000년대:** 더 많은 기업이 클라우드 컴퓨팅 도구를 구축함.
 - **Google Docs**가 2006년에 출시
 - 대학 같은 기관들이 공유 시스템의 처리 능력을 활용하여 대용량 데이터 세트를 더 빠르게 분석하기 시작함.
- **2010년대:**
 - 안전하고 보안이 강화된 공용 클라우드(**public cloud**) 도구의 필요성이 커짐.
 - 2010년대 말에는 클라우드 컴퓨팅이 많은 온라인 비즈니스의 표준 시스템이 됨.
- 현재와 미래:
 - 다국적 기업, 원격 및 오프쇼어(**offshore**) 근무 환경이 확산되면서 클라우드 시스템에 대한 수요가 그 어느 때보다 높아짐.

- 데이터 저장, 분석, 머신러닝(machine learning)의 발전과 함께 클라우드 컴퓨팅은 계속해서 성장할 것으로 예상함.

Cloud Computing의 장점

- 접근성 (**Accessibility**):
 - 조직은 인터넷을 통해 언제 어디서나 데이터, 소프트웨어, 스토리지 및 클라우드 인프라에 접근하고 관리함.
 - 하드웨어와 소프트웨어가 설치된 물리적 위치에 있을 필요가 없음.
- 확장성 (**Scalability**): 변화하는 요구 사항에 맞춰 컴퓨팅 리소스를 쉽게 확장하거나 업그레이드 할 수 있음. 물리적 컴퓨팅의 한계를 없앴.
- 비용 절감 (**Cost Savings**):
 - 사용한 컴퓨팅 리소스에 대해서만 비용을 지불하는 '측정된 서비스(measured service)' 모델을 사용함.
 - 전기나 수도와 같이 사용량(트랜잭션 수, 스토리지 용량, 데이터 전송량 등)에 따라 요금이 부과되어, 비즈니스 이니셔티브를 더 수익성 있고 지속 가능하게 만들.
- 보안 (**Security**):
 - 조직의 시스템, 데이터, 컴퓨팅 리소스가 도난, 손상, 손실, 무단 사용으로부터 보호됨.
 - 데이터 센터에 저장된 정보는 암호화되어 있어 전통적인 네트워크 인프라보다 보안이 더 강력하다고 인정됨.
- 효율성 (**Efficiency**):
 - 새롭거나 업그레이드된 애플리케이션에 즉시 접근할 수 있음.
 - 인프라 상태를 걱정하거나 시간과 비용이 많이 드는 구현 과정을 거치지 않아도 됨.
- 리소스 확보 (**Freeing Resources**):
 - '관리형 서비스(managed services)'를 통해 제3자 제공업체가 클라우드 인프라의 유지보수, 관리, 지원을 담당
 - 사용자는 더 부가가치가 높은 작업에 집중할 수 있음.

Cloud Computing의 일반적인 용도

- 재해 복구 (**Disaster Recovery**): 비상 상황 시 데이터와 정보의 안전을 보장하기 위해 더 많은 데이터 센터에 접근할 수 있게 함.
 - 데이터 스토리지 (**Data Storage**): 대용량 데이터를 저장하여 데이터 센터를 간소화하고, 데이터 접근, 분석, 백업을 용이하게 함.
 - 대규모 데이터 분석 (**Large-scale Data Analysis**): 여러 데이터 소스에 쉽고 빠르게 접근할 수 있게 하고, 직관적인 사용자 인터페이스를 통해 데이터를 쿼리하고 탐색할 수 있어, 데이터 기반 인사이트를 발견하는 과정을 가속화함.
-



인공지능, 머신러닝 및 딥러닝 관계

AI (Artificial Intelligence) 개요

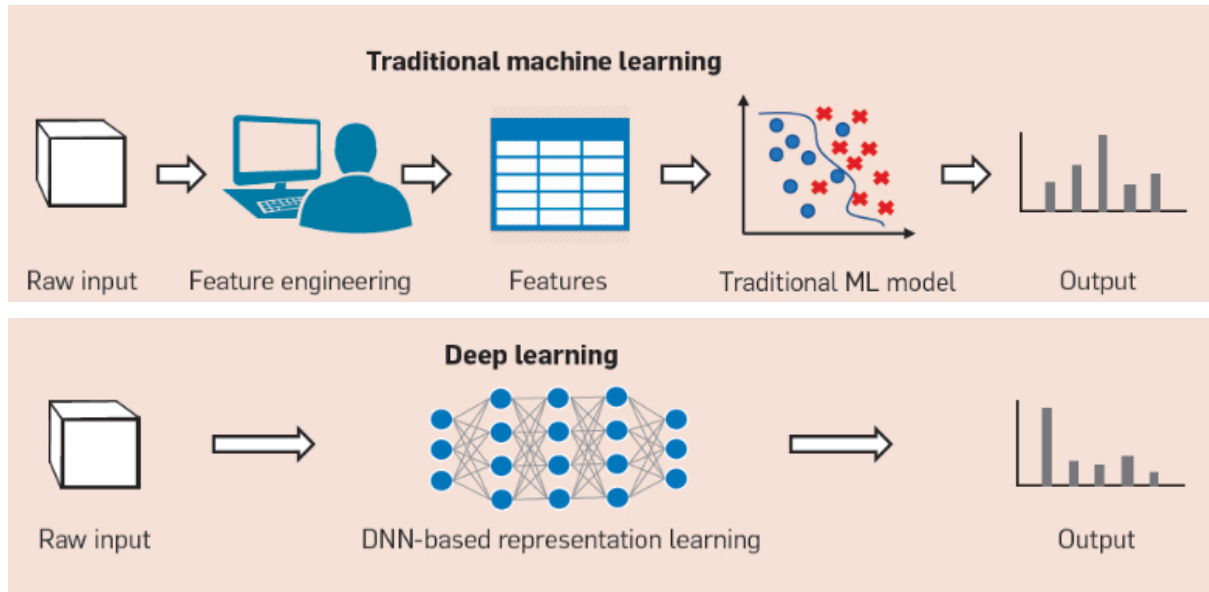
- **AI의 정의:**
 - 일반적으로 인간의 지능이 필요한 작업을 수행할 수 있는 컴퓨터 시스템의 이론 및 개발을 의미함.
 - AI는 음성-텍스트 변환이나 언어 번역 같은 다양한 작업에 보조 기술로 사용됨.
- **데이터 기반 의사결정의 중요성:**
 - AI는 데이터 분류 및 분석을 자동화
 - 예측 분석에서 중요한 역할을 수행하여 데이터 기반 의사결정에 필수적
- **AI의 하위 분야:**
 - AI는 여러 하위 분야로 구성
 - ML, Deep learning, LLM, GenAI 등이 포함

ML (Machine Learning)

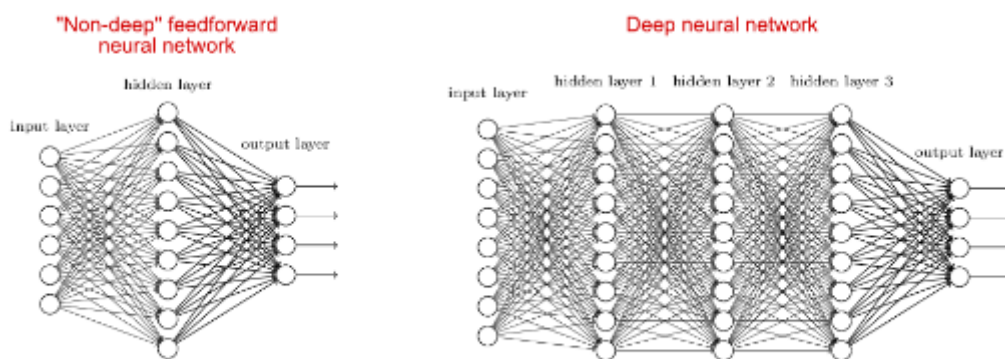
- **ML의 정의:** 명시적인 프로그래밍 없이 데이터 학습을 통해 유용한 예측을 할 수 있도록 알고리즘을 사용하는 AI의 하위 분야
- **모델(Model)과 신경망(Neural network):**
 - ML은 인공 신경망을 사용하여 데이터를 분류하며, 신경망은 인간의 뇌처럼 인공 뉴런을 통해 데이터를 연결하고 분석함.
 - 알고리즘에 더 많은 정보가 제공될수록 예측 정확도가 높아짐.
- **학습 방법:**
 - 지도 학습(Supervised learning): '
 - 레이블이 지정된 데이터(labeled data)를 사용하여 모델이 작업의 기본 개념을 학습하도록 도움.
 - 예를 들어, 개 사진에 '개'라는 레이블을 지정하여 모델을 훈련함.
 - 비지도 학습(Unsupervised learning):

- 레이블이 없는 데이터 (**unlabeled data**)를 사용하며, 모델은 태그가 없는 데이터에서 스스로 학습하여 데이터를 분류함.
- 새로운 사례에 대해 일반화하여 예측할 수 있게 됨.

Deep learning



- **Deep learning**의 정의:
 - ML의 하위 집합
 - 여러 겹의 인공 신경망을 사용하여 기존 ML보다 더 복잡한 문제를 처리할 수 있게 함.
- 특징: 인공 신경망의 수가 증가하여 모델이 더 많은 양의 데이터를 분석하고 더 정확한 예측을 할 수 있음.



인공신경망 vs 딥러닝

- 반지도 학습 (**Semisupervised learning**): Deep learning은 소량의 레이블이 지정된 데이터와 다량의 레이블이 없는 데이터를 사용하여 모델을 훈련하는 반지도 학습 방법을 사용하기도 함.

LLM (Large Language Model)

- **LLM의 정의:**
 - Deep learning의 하위 집합
 - 질문에 답변하고, 문서를 요약하며, 텍스트를 분류하거나 생성하는 등 언어 관련 문제를 해결하는 데 도움을 줌.
 - **훈련 데이터:** 방대한 양의 훈련 데이터를 포함하며, 대량의 데이터로 사전 훈련 (pretrained) 하거나 특정 도메인에 맞춰 미세 조정 (fine-tuned)할 수 있음.
 - 사전 훈련된 모델: 이전에 확립된 데이터를 사용하여 작업을 완료함.
 - 미세 조정된 모델: 새로운 데이터로 LLM을 훈련하여 새로운 도메인이나 맞춤형 사용 사례에 맞춰 조정함.
 - **LLM의 이점:**
 - 전문 지식이 없거나 모델 훈련 방법을 몰라도 기술을 사용할 수 있음.
 - 작업을 수행하기 위한 명확한 프롬프트만 있으면 됨.
 - LLM은 훈련 데이터를 기반으로 수학적으로 정확하다고 예측되는 콘텐츠를 생성함.
-

전통적인 컴퓨팅 (Traditional Computing)

- **정의:**
 - 물리적 하드웨어와 소프트웨어를 사용하여 데이터 저장, 접근, 관리를 가능하게 하는 컴퓨팅 모델.
 - 일반적으로 온프레미스(on-premises) 네트워크 인프라 내에 위치함.
- **작동 방식:**
 - IT 전문가가 전용 공간에 하드웨어를 설치하고, 필요한 소프트웨어를 구매 및 설치함.
 - 이후 IT 담당자가 전체 시스템을 유지보수하고 관리함.
- **비유:**
 - 소중한 사진을 책장에 보관된 앨범에 넣어두는 것에 비유할 수 있음.
 - 어느 정도의 통제권, 편리함, 보안을 제공하지만, 리소스(돈, 재료, 시간)에 제약을 받고, 물리적 손상 위험이 있으며, 공간의 한계가 있음.

전통적인 컴퓨팅의 장점

- **제어 (Control):**
 - 조직이 하드웨어, 소프트웨어, 데이터를 완벽하게 제어할 수 있음.
 - 특정 요구 사항에 맞게 로컬 네트워크 인프라를 맞춤 설정하는 것이 가능함.
- **보안 (Security):** 적절히 유지보수될 경우, 시스템과 민감한 정보에 대한 단독 접근 권한을 가지므로 보안에 대한 신뢰도가 높음.
- **규정 준수 (Compliance):** 특정 규제, 규칙, 법률(예: 데이터 보안)을 준수해야 하는 산업의 경우, 데이터를 온프레미스에 저장하는 전통적인 컴퓨팅이 유일한 선택일 수 있음.

- 인터넷 의존성 없음 (**No reliance on the internet**): 네트워크나 데이터 접근 시 인터넷 연결에 의존하지 않으므로, 인터넷 서비스가 중단되어도 중요한 정보에 접근할 수 있음.

전통적인 컴퓨팅의 단점

- 접근성 제한: 데이터 접근이 하드웨어와 소프트웨어가 설치된 장치와 위치로 제한됨.
- 확장성 부족: 소프트웨어 제한, 하드웨어 구매 및 설정에 필요한 시간, 물리적 공간 제약 등으로 인해 확장하기가 어렵고 비용이 많이 듦.
- 높은 비용: 하드웨어와 소프트웨어 구매 비용 외에도, 네트워크 인프라의 지속적인 유지보수가 필요하여 비용이 많이 듦.
- 비효율성: 각 사용자마다 소프트웨어를 구매해야 하며, 자동 업데이트가 되지 않아 비효율적일 수 있음.

전통적인 컴퓨팅과 클라우드 컴퓨팅의 비교

- 클라우드 컴퓨팅의 우위:
 - 클라우드는 더 높은 접근성, 확장성, 비용 절감 효과를 제공함.
 - 보안 및 효율성이 뛰어나고, 직원들이 더 부가가치 있는 업무에 집중할 수 있도록 리소스를 확보해 줌.
- 선택 기준:
 - 전통적인 컴퓨팅과 클라우드 컴퓨팅은 각각 장단점이 있으므로, 비즈니스의 우선순위에 따라 적절한 모델을 선택하는 것이 중요함.
 - 클라우드 데이터 분석가로서 이러한 차이점을 아는 것은 매우 중요함.

Cloud Data Warehouse의 필요성

- 전통적 데이터베이스의 한계: 전통적인 데이터베이스는 진화하는 데이터 분석의 요구를 따라잡는 데 어려움을 겪음.
- 해결책: 클라우드 데이터 웨어하우스(**cloud data warehouse**)가 이러한 필요를 충족하기 위해 등장함.

Cloud Data Warehouse의 특징

- 정의: 클라우드 서비스 제공자가 원격 서버에서 호스팅하는 대규모 데이터 스토리지 솔루션.
- 기능:
 - 다양한 곳에서 오는 대량의 컨테이너를 보관하는 거대한 창고에 비유
 - 데이터를 수집하고, 저장하며, 통합하고, 분석할 수 있음.
- 관리:
 - 클라우드 제공업체가 완전히 관리함.
 - 사용자는 인프라 관리에 신경 쓰지 않고 데이터 활용 및 인사이트 도출에 집중할 수 있어 시간, 비용, 리소스를 절약할 수 있음.

- 가동 시간(**Uptime**): 온프레미스 데이터 웨어하우스에 비해 가동 시간이 길어, 증가하는 데이터 수요에 맞춰 확장하고 지원하는 능력이 뛰어남.
- 통합(**Integration**):
 - 판매 시스템, 이메일 리스트, 웹사이트 등 조직 내 다양한 정형화된 소스에서 데이터를 수집하여 한 곳으로 통합됨.
 - 이렇게 통합된 데이터는 유용한 비즈니스 인사이트를 얻기 위해 분석될 수 있음.

Cloud Data Warehouse의 주요 이점

- 실시간 분석(**Real-time analytics**): 사용자가 최신 정보에 빠르게 접근할 수 있도록 보장하여, 경쟁에서 앞서나가는 데 도움이 됨.
- **AI/ML** 기능:
 - AI 및 ML 기능을 적용하여 데이터 분석의 가능성을 크게 높일 수 있음.
 - 예를 들어, Google에서 사무실 편의 시설에 대한 수요를 예측하는 모델을 구축하여 비용을 절감하고 낭비를 줄인 사례가 있음.
- 맞춤형 보고 및 분석(**Custom reporting and analysis**): 현재 비즈니스 거래 및 일상 업무 관련 데이터와 분리된 서버에 저장된 과거 데이터를 사용하여 사용자가 특정 보고서를 분석하고 생성할 수 있음.

결론

- 기업이 조직화해야 할 데이터의 종류와 양은 계속해서 증가하고 있음.
- 클라우드 데이터 웨어하우스는 이러한 과제를 해결하며, 관리 및 분석 기능의 추가적인 이점을 제공하여 데이터를 더욱 쉽게 활용할 수 있도록 함.

클라우드 데이터 웨어하우스의 장점

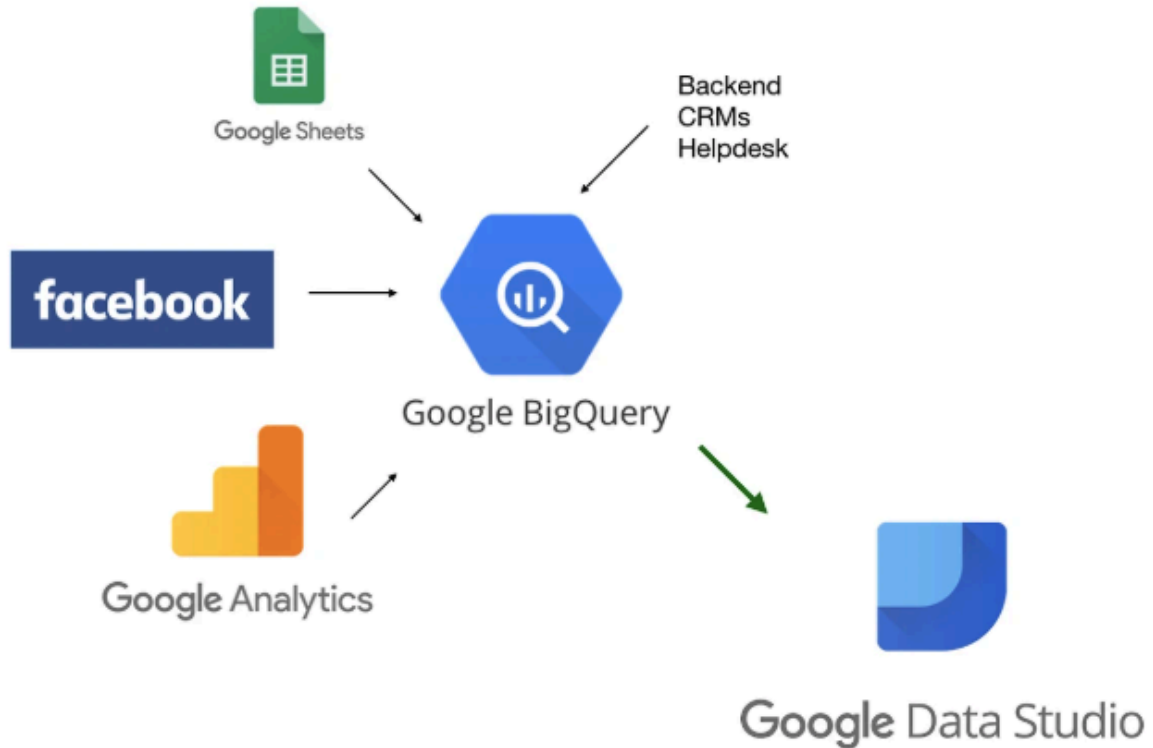
- 클라우드 **vs.** 전통적 데이터 웨어하우스:
 - 전통적 데이터 웨어하우스: 여러 소스 시스템의 데이터를 통합하여 데이터 일관성, 정확성 및 효율적인 접근을 가능하게 하는 훌륭한 스토리지 솔루션임.
 - 클라우드 데이터 웨어하우스: 이와 같은 이점을 제공하지만, 조직이 활용할 수 있는 추가적인 개선 사항이 있음.
- 주요 비교

	클라우드 데이터 웨어하우스	전통적 데이터 웨어하우스
비용	서버를 로컬에 호스팅할 필요가 없어 시간과 리소스를 절약함.	서버를 구현, 유지보수, 최적화하는 데 리소스가 필요함.
확장성 및 유연성	autoscaling을 통해 언제든지 변화하는 데이터 요구에 맞춰 조정 가능함.	변경 요청이 있을 때마다 추가적인 물리적 컴퓨터 리소스가 필요함.
가동 시간 및 접근성	전통적 서버 웨어하우스보다 가동 시간이 더 많음. 서로 다른 위치와 시간에 일하는 사용자에게 데이터가 더 쉽게 접근 가능함.	가동 시간이 더 적고, 서버가 실행 중일 때만 운영됨. 데이터 접근이 현장 보안 메커니즘으로 인해 제한될 수 있음.
데이터 관리	클라우드 제공업체가 관리하므로 내부 리소스가 덜 필요함.	스토리지 시스템을 관리하기 위해 현장 전문가와 리소스가 필요함.
실시간 분석 및 새로운 도구	AI 및 ML 도구를 사용하여 데이터를 즉시 수집하고 분석할 수 있어 실시간 분석 제공이 더 쉬움.	스트리밍 워크로드를 분석하고 AI 및 ML 도구를 추가하는 데 더 많은 리소스가 필요함.

클라우드 데이터 전문가들은 이러한 장점을 이해하고, 조직의 기존 스토리지 시스템을 어떻게 개선할 수 있을지 파악하여 가장 효과적인 데이터 스토리지를 설계하는 데 도움을 줄 수 있음.

BigQuery 소개

- **BigQuery란:**
 - Google Cloud에서 제공하는 데이터 웨어하우스로, 사용자가 데이터를 저장하고 분석할 수 있게 도와줌.
 - 방대한 양의 데이터를 빠르게 쿼리하고, 필터링하며, 결과를 집계하고, 복잡한 연산을 수행할 수 있음.
- **SQL 사용:**
 - 데이터베이스와 소통하는 컴퓨터 프로그래밍 언어인 SQL(structured query language)을 사용함.
 - Google 인프라를 활용하여 대량의 데이터를 매우 빠르게 검색할 수 있음.



- 데이터 분석가의 역할: 클라우드 데이터 분석가로서 **BigQuery**의 통합 SQL 인터페이스와 **ML** 기능을 활용하여 중요한 비즈니스 결정을 위한 데이터 도구를 발견, 구현 및 관리하는 데 도움을 줄 수 있음.
- 다른 도구와의 통합: **BigQuery** 작업의 결과는 일반적인 비즈니스 인텔리전스 도구나 스프레드시트와 통합될 수 있음.

BigQuery의 주요 기능 및 장점

- 데이터 마이그레이션: 다른 클라우드 서비스 제공업체의 기존 데이터 웨어하우스를 쉽게 마이그레이션할 수 있음.
- **Dry-run** 매개변수:
 - 쿼리를 실제로 실행하기 전에 쿼리 로직을 확인하고 계획할 수 있음.
 - 쿼리 실행에 필요한 바이트 수를 알려주어 예상 비용을 추정하는 데 도움을 줌.
- 데이터 쿼리 및 저장: 서버, 센서 및 기타 장치에서 수집된 데이터를 저장하고, 탐색하며, 쿼리를 실행하는 데 사용될 수 있음.
- 예정된 쿼리 (**Scheduled queries**):
 - 데이터를 자동으로 새로 고치고 테이블을 최신 상태로 유지하는 데 사용될 수 있음.
 - 시간 단위, 일 단위, 주 단위로 데이터를 업데이트하여 이해관계자에게 가장 역동적이고 시의적절한 메트릭을 제공함.
- 실제 활용 사례:
 - Google 팀은 **BigQuery**를 거의 매일 사용하여 데이터를 쿼리, 변환, 보고함. **SQL**을 사용해 데이터셋을 결합하고 데이터를 변환하여 테이블과 차트를 생성함.

- 미래에 유용할 수 있는 결과를 발견하면, 셀프 서비스 보고서와 대시보드를 구축하여 사용자가 같은 답변을 반복적으로 빠르게 얻을 수 있도록 확장함.

결론

- **BigQuery**는 존재하는 데이터와 사용자가 해결하려는 문제 또는 답변하려는 질문 사이의 다리 역할을 함.
 - 다른 도구와의 원활한 통합, 사용자 친화적인 인터페이스, 효율적인 프로그래밍을 위한 **SQL** 사용을 통해 복잡한 데이터셋 내에서 가치 있는 정보를 발견하는 것을 간단하고 생산적으로 만듦.
 - 클라우드 데이터 분석가의 도구 상자에서 필수적인 부분임.
-

Cloud Data Analytics 과정 요약

- 클라우드 컴퓨팅의 이해: 클라우드 컴퓨팅은 전통적인 컴퓨팅의 많은 한계를 해결하고, 전 세계 개인과 기업의 진화하는 데이터 컴퓨팅 요구를 충족시키는 진보된 강력한 컴퓨팅 모델임.
 - 주요 장점: 클라우드 컴퓨팅은 인터넷을 통해 온디맨드(**on-demand**) 방식으로 컴퓨팅 리소스를 서비스로 제공함. 이는 접근성, 확장성, 비용 절감, 보안 및 효율성이라는 장점을 제공함.
 - 리소스 확보: 클라우드 컴퓨팅은 직원들이 더 가치 있는 작업에 집중할 수 있도록 시간과 리소스를 확보해줌.
 - 학습 내용:
 - 코스 소개: 클라우드 컴퓨팅의 역사, 현재의 특징, 그리고 사용 모델의 장점에 대해 학습함.
 - 비교 분석: 클라우드 컴퓨팅과 전통적인 컴퓨팅의 차이점, 예를 들어 물리적 네트워크 인프라와 클라우드 네트워크 인프라의 차이를 이해함.
-

Module 1 용어 정리

- 클라우드 컴퓨팅 (**Cloud computing**): 인터넷을 통해 서비스로 제공되는 온디맨드 컴퓨팅 리소스를 사용하는 방식
- 클라우드 데이터 분석 (**Cloud data analytics**): 클라우드 기반 서비스 및 솔루션을 사용하여 대량의 데이터를 수집, 저장, 관리, 구성, 변환 및 분석하는 과정
- 클라우드 데이터 웨어하우스 (**Cloud data warehouse**): 클라우드 서비스 제공업체가 원격 서버에서 호스팅하는 대규모 데이터 스토리지 솔루션
- 데이터 센터 (**Data center**): 서버, 컴퓨터 시스템 및 관련 구성 요소를 포함하는 물리적 건물

- **Google BigQuery:** Google Cloud의 데이터 웨어하우스로, 데이터를 쿼리하고, 대규모 데이터셋을 필터링하며, 결과를 집계하고, 복잡한 작업을 수행하는 데 사용함
 - **IaaS (Infrastructure as a service):** 하드웨어, 스토리지, 네트워킹 및 가상화 도구를 포함한 IT 인프라 서비스를 온디맨드로 제공하는 클라우드 컴퓨팅 모델
 - **관리형 서비스 (Managed service):** 제3자 제공업체가 조직의 클라우드 인프라 및 애플리케이션에 대한 지속적인 유지보수, 관리 및 지원을 담당하는 행위
 - **측정된 서비스 (Measured service):** 트랜잭션 수, 스토리지 용량, 전송된 데이터 양을 기준으로 사용한 만큼만 요금을 부과하는 서비스
 - **온프레미스 (On-premises):** 조직 자체 데이터 센터 또는 사무실에 물리적으로 위치한 정보 기술 인프라
 - **PaaS (Platform as a service):** 클라우드 애플리케이션 개발 환경을 구축하기 위한 하드웨어 및 소프트웨어 도구를 제공하여 애플리케이션 개발 프로세스를 단순화하는 클라우드 컴퓨팅 모델
 - **SaaS (Software as a service):** 완전한 소프트웨어 패키지에 대한 라이선스 기반 구독을 사용자에게 제공하는 클라우드 컴퓨팅 모델
 - **전통적 컴퓨팅 (Traditional computing):** 일반적으로 온프레미스에 위치한 네트워크 인프라 내에서 물리적 하드웨어 및 소프트웨어를 사용하여 데이터 저장, 접근 및 관리를 가능하게 하는 컴퓨팅 모델
 - **가동 시간 (Uptime):** 기계가 작동 가능한 시간의 양
 - **가상화 (Virtualization):** 서버, 스토리지, 네트워크와 같은 물리적 인프라의 가상 버전을 생성하는 기술
-