**2023년 하계 저학년 예비직무 전문가 과정**

**직무 로드맵 과제 보고서**

폰트, 로고, 그래픽, 디자인이(가) 표시된 사진

자동 생성된 설명

|  |
| --- |
| 기업명 : 카카오  직무명 : 데이터 엔지니어 |

**2023년 08월 11일**

|  |  |
| --- | --- |
| **조 명: 7조** | **담임멘토: 박희주 멘토** |
| **팀 명: 너두주코나두좋조** | **팀 장 명: 박준하** |
| **팀원명: 권민지, 신지후, 심재민, 안병민, 이호영, 정용재, 홍예진** | |
| 부팀장 권민지 AI빅데이터융합경영학 010-2939-9425  정용재 AI빅데이터융합경영학 010-2427-9953 | |
| 서 기 홍예진AI빅데이터융합경영학 010-8283-7502 | |
| 총 무 신지후 AI빅데이터융합경영학 010-8510-0534  이호영 AI빅데이터융합경영학 010-2541-8099 | |
| 홍 보 심재민 AI빅데이터융합경영학 010-3057-5653  안병민 행정학과 010-3178-3638 | |

**[ 목차 ]**

Ⅰ. 개요

Ⅱ. 서론

        1. 공모전 참여 목적과 목표

                   A. 목적

                   B. 목표

        2. 직무와 산업/기업 선정이유

                   A. 직무 선정이유 – 데이터 엔지니어

                   B. 산업 선정이유 – 포털 및 기타 인터넷 정보 매개 서비스업

                   C. 기업 선정이유 – 카카오

        3. 주요 활동

                   A. 설문조사

                   B. 현직자 인터뷰

                   C. 자료조사

        4. 기대 효과

Ⅲ. 본론

        1. 산업분석

                   A. 포털 및 기타 인터넷 정보 매개 서비스업

                   B. 카카오 주요 산업 분야

                   C. 포털 및 기타 인터넷 정보 매개 서비스업 전망

        2. 기업분석

                   A. 기업명

                   B. 기업소개

                   C. 기업개요

                   D. 주력 사업 및 신규 사업

                   E. 채용 정보

        3. 직무 기본 분석 – 데이터 엔지니어

                   A. 직무 소개

                   B. 직무의 업무

                   C. 직무의 고객

                   D. 직무의 근무환경

                   E. 직무수행 시 만족감 및 성취감

                   F. 직무수행 시 힘든 점

                   G. 직무의 전망

        4. 직무 요구 역량 분석 – 데이터 엔지니어

                A. 핵심역량

B. 대학 교육

                   C. 대학원 과정

                   D. 교내활동

                   E. 자격증

                   F. 태도 및 성향

                   G. 필요경험

        5. 표준 커리어 플래너

Ⅳ. 결론

1. 결론 및 요약

        2. 보고서를 읽는 대학생들에게

       3. 개인 별 소감

**Ⅰ. 개요**

주니어 코렙 11기 너두주코나두좋조팀은 데이터 직무에 관심을 가진 다수의 AI빅데이터융합경영학 전공자로 구성되어 데이터 관련 취업을 준비하는 학생들에게 산업 직무의 정보 및 지식을 제공해 주고자 보고서를 작성한다.

본 보고서에는 데이터 관련 직무 중 데이터 엔지니어를 중점적으로 다룬다. 본 팀은 데이터 엔지니어의 업무, 근무 환경, 전망 및 필요 역량에 대해 알아보기 위해 카카오, 컬리, 우아한 형제들 총 3개 기업의 현직자 인터뷰를 진행했다. 관련 인터뷰를 위해 설문조사를 실시하였으며, 추가 자료조사를 통해 데이터 산업 분야의 데이터 직무에 대한 탐구를 진행하였다.

설문조사 결과, 데이터 직무에 관심이 있어도 자세한 취업 방안은 모르는 사람이 많다는 것을 알고 특정 기업을 선정하여 구체적인 지식 및 취업 정보를 제공해야 할 필요성을 느꼈다. 따라서 인터뷰 한 3가지 기업 중 카카오를 선정하여, 카카오의 데이터 엔지니어 직무를 중점으로 보고서를 작성하였다. 해당 직무에 대한 탐구를 위해, 크게 산업분석, 기업분석, 직무분석, 직무 요구 역량 분석 및 커리어 플래너를 나누었다.

산업과 기업분석의 세부적인 사항으로, 카카오의 주요 산업 분야 9가지, 복지, 주력 사업, 채용 기회를 알아보았다. 카카오의 데이터 엔지니어라는 직무분석을 위해 업무 프로세스, 근무 환경, 전망, 역량을 조사하였다. 추가로 구체적인 취업 플랜을 위해 학년별 커리어 플래너를 작성하여 전공, 교내 활동, 개인 활동을 정리하였다.

이 활동들을 통해 해당 기업 또는 직무를 목표로 하는 국민대학교 학생들에게 관련 이해도를 높여주어 명확한 진로 설정과 취업에 필요한 준비를 할 수 있도록 도움을 주고자 한다.

**Ⅱ. 서론**

**1. 공모전 참여 목적과 목표**

**A. 목적: 실무, 프로세스, 취업 역량 파악**

본 팀은 데이터 직무에 관심 있는 AI빅데이터융합경영학 전공자 및 행정학 전공자로 구성되어, 실무 및 역량 등을 알아보고자 주니어 코렙 공모전에 참여했다.

최근 몇 년간 전 세계적으로 비즈니스 프로세스의 디지털화가 급속히 진행되는 추세로, 데이터의 양과 속도 또한 향상되고 있다. 나날이 증가하는 데이터는 데이터 시장을 성장시키고 있으며, 관련 데이터 직무의 수요 또한 증가하고 있다. 따라서 데이터 산업으로의 진출을 위해서는 관련 지식을 넓혀야 한다고 판단하여, 실무에서의 세부적인 프로세스, 취업을 위한 역량 등을 파악하고자 했다.

주니어 코렙 공모전을 통해 데이터 직무를 이해하고 진로 방향성을 잡아서 취업을 위한 토대를 마련한다는 목적을 세워 실천하고자 한다.

**B. 목표**

**가. 현직자와 관계망 형성**

현직자와 관계망을 형성하는 것은 데이터 직무를 이해하고 진로 방향성을 잡는 데에 중요한 영향을 끼친다. 실제 업무를 경험하는 현직자에게 직무 이해와 실제 업무 경험을 공유 받을 수 있으며, 산업 및 회사 정보를 습득할 수 있다. 다양한 산업에서 데이터가 어떻게 활용되고 있는지, 어떤 회사들이 선두 주자인지 등의 정보를 얻으면서, 관심 분야에 대한 적합한 회사를 찾을 수 있게 될 것이라 기대한다. 또, 멘토링과 조언을 얻어 자신의 역량을 향상시키는 데 큰 도움이 될 것이라고 판단했다.

따라서, 해당 목표를 진행하기 위해 데이터 실무를 담당하고 있는 현직자 3명과 인터뷰 진행하고자 한다.

**나. 실무 환경과 근무 과정 파악하기**

이론적인 지식과 실무에서의 경험은 차이가 있기 때문에, 현직자를 통해 실무 환경에 대해 파악함으로써, 책으로 배우는 것 외에 실제 업무의 복잡성과 실무자의 역량을 이해할 수 있을 것이라 기대하여 다음과 같은 목표를 설정하였다. 이를 위해 기업별 프로젝트 과정을 조사하고 시각화하고 실무와 전공 공부와의 차이점을 자세하게 알아보는 세부적인 목표를 세웠다.

**다. 커리어 플래너 세우기**

본 팀의 목표 직무인 데이터 직무에 적합한 커리어 플래너를 작성해보고자 한다. 커리어 플래너는 명확한 목표를 설정하고 계획적인 학습과 성장을 이룰 수 있게 도울 뿐만 아니라 진행 상황을 파악하고 동기를 부여할 수 있는 성취감을 느낄 것을 기대해본다.

**2. 산업/기업과 직무 선정이유**

**A. 산업 선정이유 – 포털 및 기타 인터넷 정보 매개 서비스업**

지금은 웹 3.0 시대와 5차 산업혁명 시대이다. 웹 3.0은 ‘탈중앙화’, ‘개인의 컨텐츠 소유’가 특징인 차세대 인터넷을 부르는 용어이다. 쉽게 말해 데이터들이 개개인에 맞추어 커스터마이징이 되면서 이용자가 필요한 정보와 지식만을 추출하여 보여준다. 5차산업혁명은 4차산업혁명과 결합되어 있는데 4차산업은 자동차에 핵심을 두었다면 5차산업혁명은 사람과 로봇이 함께 협력하는 시대라고 한다. 이렇듯 점점 데이터의 가공과 활용, 인공지능의 중요도가 상승하는 시대에서 해당 산업은 주목받는 산업 중 하나이다. 데이터가 가장 많이 생성되고 활용되는 곳이자 인공지능이 결합되어 뿌리가 되는 곳이 인터넷 포털이다. 그러므로 이 산업은 점차 성장하고 더욱 확장되어 미래 핵심 주요 산업 중 하나라고 판단하여 목표 산업으로 선정하였다.

**B. 기업 선정이유 – 카카오**

카카오는 ‘데이터 엔지니어’로서의 업무를 보고 커리어를 쌓기 적합한 기업이라 판단하여 목표 기업으로 선정했다. 해당 기업은 2023년 대학생 취업 선호 기업 2위에 이름을 올렸고 꾸준히 상위권에 위치해 있을 정도로 인지도와 취업 선호도가 높다. 카카오는 카카오엔터테인먼트, 카카오게임즈, 카카오뱅크 등 다양하고 탄탄한 계열사를 가지고 있어 안정성이 높고 발전가능성이 무궁무진하다. 또한, 카카오는 자체적으로 데이터를 수집하고 분석하여 기술을 보완하여 더 나은 서비스를 제공한다. 업데이트된 부분을 매년 카카오 홈페이지에 있는 하이라이트 트랙 및 세션에 업로드하여 빅데이터 트랙에서의 개선점을 공유한다. 이정도로 방대한 데이터 양을 가지고 있고 이를 활용하는 기업이라는 점에서 데이터 직무의 다방면적인 역량을 키울 수 있을 것이라 판단했다.  AI 분야에서 세계적인 대기업들 사이의 경쟁이 치열해지고 있는 상황에서, 카카오는 해당 기술에 대한 투자를 강화하고 있는 것으로 평가되는 정도로 인공지능 기술 발전을 위해 과감히 투자하여 미래를 이끌어 갈 기업들 중 대표 기업으로 성장할 수 있을 것이라 기대되어 카카오를 선정하였다.

**C. 직무 선정이유 – 데이터 엔지니어**

“데이터는 잠들지 않는다.” 라는 말에서 알 수 있듯이 매초마다 새로운 데이터들이 생겨나고 있다. 1분에 세계에서 발생하는 데이터의 양으로 구글에서는 5.7millions, 즉 570만개고, 페이스북에서는 240k, 즉 24만개의 사진이 공유되고 있다. 오늘날 기업들은 이렇게 발생한 빅데이터를 활용하여 데이터 분석, 머신러닝, 딥러닝 등 데이터를 통한 가치창출을 할 수 있어 데이터의 중요도가 높아져 데이터 관련 직무의 필요와 선호도 높아지고 있다.

데이터 관련 직무는 크게 데이터 엔지니어, 데이터 사이언티스트, 데이터 분석가로 나뉜다. 그중 본 팀은 목표 직무를 데이터 엔지니어로 선정하였다. 데이터 엔지니어는 데이터 관련 다른 직무인 데이터 사이언티스트, 데이터 분석가보다 개발자 직군에 가깝다. 본 팀원 모두가 데이터 관련 직무에 관심과 흥미가 생긴 이유가 프로그래밍을 하는 데에 있다. 인사이트를 도출하는 과정도 중요하지만, 정답이 있는 것이 아니라 관점에 차이기 때문에 효율적인 인사이트를 정의하기에 확신이 없어 어려움을 많이 느낀다. 기획 부분보다는 개발 부분에 더 재미를 느껴 데이터의 개발자, 데이터 엔지니어 직무로 결정하게 되었다. 또한, 가장 먼저 데이터 엔지니어가 일을 하여 데이터 플로우 성능을 최적화해야 그 뒤에 데이터 사이언티스트가 정제된 데이터를 가지고 인사이트를 도출하고, 데이터 분석가가 그 인사이트를 가지고 비즈니스 가치 창출을 기획할 수 있다.

이렇듯, 데이터 관련 업무의 첫 단계인 데이터 엔지니어에 대해 깊은 탐색을 하여 심도 있는 이해를 함과 동시에 국민대학교 재학생들의 데이터 엔지니어에 대한 직무 이해도를 향상시키고, 더 나아가 그들이 취업 준비를 하는 데 도움을 주고자 한다.

**3. 주요 활동**

**A. 자료조사**

산업과 기업, 직무에 대해 다양한 경로로 정보를 수집했다. ‘구글’, ‘네이버’ 등 검색 플랫폼과 유튜브에서 신뢰할 수 있는 정보를 선별하였으며, 특히 직무에 대해서는 해당 기업 채널의 자체 유튜브 컨텐츠와 블로그를 통해 세부적이고 정확한 정보를 얻을 수 있었다. 검색만으로 알기 힘든 세밀한 정보들은 데이터 관련 직종 현직자 인터뷰를 통해 조사했다.

**B. 현직자 인터뷰**

**가.** 카카오 데이터 플랫폼 팀 사원

**나.** 컬리 데이터 농장팀 소속 데이터분석가

**다.** 우아한형제들 데이터 서비스실 소속 데이터분석가

**4. 기대 효과**

위 보고서는 AI빅데이터융합경영학과 신입생과 재학생, 복수전공 및 부전공을 희망하는 재학생 그리고 데이터 직무에 관심있는 재학생을 대상으로 만들었다. 데이터 직무를 희망하지만 정보가 부족한 대학생, 취업 준비생에게 산업 및 직무 분석 정보를 제공하여 구체적인 직무 선정에 도움을 줄 수 있다. 실제 데이터 취업을 계획하는 재학생과 졸업생에게 실질적이고 정확한 취업 정보를 제공하여 취업 준비를 도울 수 있다. 공모전 활동을 통하여 개인의 명확한 진로 설정과 커리어 플랜을 세우며 취업 준비에 대한 역량을 키울 수 있다.

**Ⅲ. 본론**

**1. 산업분석**

**A. 포털 및 기타 인터넷 정보 매개 서비스업**

**가. 정의**

포털 및 기타 인터넷 정보 매개 서비스업은 인터넷을 통해 다양한 정보와 서비스를 중개하고, 제공하는 기업 또는 업체를 말한다. 이러한 기업들은 인터넷 플랫폼을 통해 사용자에게 다양한 정보를 제공하거나, 다른 서비스와 연결하여 편리한 이용 환경을 제공한다. 이러한 기업들은 사용자들의 검색 요청에 따라 웹 검색 결과를 제공하는 검색 엔진뿐만 아니라, 포털 사이트와 소셜 미디어 플랫폼, 콘텐츠 제공 플랫폼 등을 포함한다. 또한 사용자들의 편의를 증진시키고, 다양한 정보와 서비스에 손쉽게 접근할 수 있도록 도와주며, 인터넷 사용의 중요한 요소로 자리 잡고 있다.

**나. 특징**

ㄱ. 정보중개

기업들은 다양한 정보를 수집하고 이를 사용자에게 중개하여 제공한다. 이러한 정보에는 뉴스, 날씨, 주식 정보, 금융 정보, 여행 정보, 쇼핑 정보 등이 포함될 수 있다.

ㄴ. 검색엔진

대표적인 기능으로 웹 검색 엔진을 운영하여 사용자들이 원하는 정보를 검색하고 찾을 수 있도록 도와준다. 검색 결과는 사용자의 검색어와 관련된 웹페이지들이 우선적으로 나열된다.

ㄷ. 포털 사이트

포털 사이트는 사용자들에게 다양한 정보와 서비스를 제공하는 웹페이지를 말한다. 포털 사이트는 뉴스, 커뮤니티, 쇼핑, 게임 등 다양한 서비스를 한 곳에서 제공하는 형태가 일반적이다.

ㄹ. 콘텐츠 제공 플랫폼

웹툰, 웹소설, 음악, 동영상 등 다양한 콘텐츠를 제공하는 플랫폼도 포털 및 기타 인터넷 정보 매개 서비스업에 속한다.

ㅁ. 소셜 미디어 플랫폼

소셜 미디어 플랫폼은 사용자들이 소셜 네트워크를 통해 커뮤니케이션하고 콘텐츠를 공유하는 서비스를 제공한다.

**B. 카카오 주요 산업 분야**

카카오의 주요 산업 분야는 9개로 나눌 수 있다.

**가. 데이터 베이스업**

데이터 베이스란 데이터를 체계적‧유기적으로 배열 또는 구성한 데이터의 집합체로서 개별적 또는 공으로 필요한 데이터에 접근하거나 검색‧활용할 수 있도록 한 것을 말한다. 또한 ‘데이터 베이스 산업’이란 데이터 베이스의 제작‧관리‧활용‧유통 등과 이와 관련된 서비스를 제공하는 산업을 말한다. 데이터 베이스 산업은 데이터 베이스를 구축하는 영역을 비롯하여 전문 컨설팅 영역, 전문 솔루션 영역으로 구분되며 그 외에도 데이터 베이스화된 지식과 정보를 제공하는 데이터 베이스 서비스 영역이 포함된다.

텍스트, 폰트, 원, 스크린샷이(가) 표시된 사진

자동 생성된 설명

(그림 1-B-가) DB산업 주요 시장 영역

**나. 인터넷관련 소프트웨어**

인터넷 관련 소프트웨어 산업은 인터넷과 관련하여 소프트웨어를 개발, 제조, 판매, 운영 등과 관련된 활동을 수행하는 산업을 말한다. 이 산업은 인터넷을 기반으로 하는 다양한 소프트웨어와 서비스를 개발하고 제공함으로써 사용자들에게 정보, 커뮤니케이션, 엔터테인먼트 등 다양한 기능을 제공한다.

**다. 시스템통합**

시스템 통합 산업은 다양한 정보 기술 (IT) 시스템과 서비스를 하나로 통합하여 효율적으로 운영하고 제공하는 산업을 말한다. 이 산업은 기업이나 기관의 IT 시스템과 솔루션을 통합하고, 다양한 기술과 애플리케이션들을 유기적으로 연결하여 시스템을 효과적으로 관리하고 최적화하는 것을 목표로 한다.

**라. 인터넷광고사업**

인터넷 광고 산업은 인터넷을 통해 상품, 서비스, 브랜드 등을 홍보하고 마케팅하는 데에 사용되는 광고 산업을 말한다. 이 산업은 기업이나 기관이 인터넷을 활용하여 대중들에게 광고 메시지를 전달하고, 제품이나 서비스를 홍보하는 데에 매체로서의 역할을 수행한다.

**마. 무료 이메일서비스**

무료 이메일 서비스 산업은 사용자들에게 무료로 이메일 계정을 제공하고, 이메일을 주고받을 수 있는 서비스를 제공하는 산업을 말한다. 이 산업은 인터넷 사용자들에게 편리하고 안정적인 이메일 서비스를 제공하여, 개인적인 또는 비즈니스적인 목적으로 이메일을 사용하는 데에 도움을 준다.

**바. 전자상거래**

전자상거래 산업은 전자적인 방법을 통해 상품과 서비스를 인터넷을 비롯한 디지털 매체를 통해 구매, 판매, 교환하는 비즈니스 활동을 말한다. 이 산업은 전통적인 오프라인 상거래를 대체하거나 보완하며, 소비자와 판매자 사이의 거래를 디지털적으로 연결하는 기술과 시스템을 기반으로 한다.

**사. 인터넷접속서비스(PPP서비스)**

인터넷접속서비스(PPP 서비스) 산업은 개인이나 기업들이 인터넷에 접속하여 온라인으로 정보를 주고받을 수 있도록 인터넷 접속 서비스를 제공하는 산업을 말한다. PPP는 Point-to-Point Protocol의 약어로, 인터넷에 접속하기 위한 규약 중 하나이다. PPP를 이용하여 사용자들은 인터넷 서비스 제공자(ISP)의 네트워크에 접속하여 인터넷을 이용할 수 있게 된다.

**아. 호스팅서비스**

호스팅 서비스 산업은 웹사이트나 애플리케이션, 데이터베이스 등과 같은 인터넷 서비스를 인터넷을 통해 제공하기 위해 서버와 관련된 인프라 및 기술을 제공하는 산업을 말한다. 이 산업은 기업이나 개인들이 인터넷 상에서 자신들의 웹사이트, 애플리케이션, 블로그, 이메일 등을 운영하기 위해 필요한 서버를 호스팅하는 서비스를 제공한다.

**자. 캐릭터디자인**

캐릭터 디자인 산업은 다양한 분야에서 사용되는 캐릭터들을 창조하고 디자인하는 산업을 말한한다. 이 산업은 애니메이션, 만화, 영화, 비디오 게임, 광고, 상품 디자인 등 다양한 영역에서 사용되는 캐릭터들을 개발하고 디자인하는데 중점을 둔다.

**C. 포털 및 기타 인터넷 정보 매개 서비스업의 전망**

포털 및 기타 인터넷 정보 매개 서비스업의 전망은 디지털 기술의 발전과 소비자 행태의 변화에 따라 크게 영향을 받는다. 포털 및 기타 인터넷 정보 매개 서비스업은 디지털 시대의 핵심 산업으로서 계속해서 성장하고 혁신을 이루어 나갈 것으로 예상된다. 사용자들의 다양한 요구와 행태를 반영하여 편리하고 유용한 서비스를 제공하는 기업들이 성공할 수 있을 것이다. 또한, 기술의 발전과 변화에 능동적으로 대응하는 기업들이 시장에서 선두를 유지할 수 있을 것으로 기대된다. 아래는 포털 및 기타 인터넷 정보 매개 서비스업의 전망에 대한 주요 측면에는 모바일 접속의 증가, 개인화와 AI 기술 활용의 확대, 빅데이터와 데이터, 신기술과 경험의 혁신, 글로벌 시장 확대와 경쟁, 데이터 보안과 개인정보 보호 강화 등이 있다.

**2. 기업분석**

**A. 기업명: ㈜ 카카오**

**B. 기업소개**

카카오는 국내 1위 메신저인 카카오톡을 중심으로 서비스를 제공하는 모바일 라이프 플랫폼 기업이다. 인터넷 포털 사이트 다음(Daum)을 비롯해 모바일 플랫폼, 디지털 콘텐츠, 온라인 금융, 인공지능 등 다양한 분야에서 사업을 전개하며 혁신적인 기술과 사용자 중심의 서비스를 제공하고 있다. 현재 글로벌 시장에서도 큰 성과를 거두고 있으며, 글로벌 확장과 제휴를 통해 글로벌 플랫폼 기업으로서의 입지를 강화하고 있다. 더불어, 인공지능 분야에서의 연구 개발과 혁신적인 비즈니스 모델을 추구하여 산업의 트렌드를 이끌어가고 있다.

**C. 기업개요**

**가.** 업종: 포털 및 기타 인터넷 정보매개 서비스업

**나.** 대표자명: 홍은택

**다.** 기업주소: 제주 제주시 첨단로 242

**라.** 사업내용: 데이터베이스업, 인터넷 관련 소프트웨어, 시스템 통합, 인터넷 광고사업, 무료 이메일 서비스, 전자상거래, 인터넷 접속 서비스(PPP서비스), 호스팅 서비스, 캐릭터 디자인

**마.** 브랜드: 카카오톡, 다음, kakao, Daum

**바.** 계열사: 카카오(135)

**사.** 기업비전: 카카오는 새로운 연결을 통해 더 편리하고 즐거운 세상을 꿈꿉니다. 사람과 사람, 사람과 기술을 한층 가깝게 연결함으로써 세상을 어제보다 더 나은 곳으로 만들기 위해 노력하고 있습니다.

**D. 기업 복지**

**가. 생활 속 지원**

개인별로 연 360만원의 지원을 받아 자기개발과 여가/취미 활동에 활용할 수 있다. 또한, 설이나 추석과 같은 명절에는 현금 30만원을 선물로 받게 되며, 경조사나 보험 사항에도 지원이 제공된다. 주택 구입을 고려하는 경우, 대출금 및 이자에 대한 지원도 받을 수 있다. 이러한 혜택 외에도, 사내에 어린이집이 있어 아이를 동반하여 출퇴근이 가능하다.

**나. 업무 몰입 지원**

사내 및 사외에서 MOOC, 사외교육, 그리고 해외 컨퍼런스 등의 다양한 강연 및 교육 프로그램을 운영한다. 서울 및 경기 지역 출퇴근 시에는 셔틀 서비스 및 교통비를 지원하며, 야근 시에는 지역 상관없이 택시비가 제공된다. 또한, 편의를 위해 음료와 간식을 제공하며, 야근 시에는 저녁 식대도 지원된다.

**다. 휴식 지원**

다양한 휴식 공간을 내부에 마련하여 직원들이 편하게 쉴 수 있도록 한다. 사내 동호회 활동 또한 장려하며 지원한다. 개인의 워크 라이프 밸런스를 고려해, 매월 마지막 주 금요일은 자율 활동 날로 정하고 직원들이 원하는 활동에 자유롭게 참여할 수 있도록 한다. 특히, 매 3년마다 1개월의 휴가와 휴가비 200만원을 지원해주어 여가를 즐기고 휴식을 취할 수 있도록 돕는다. 더불어, 다양한 호텔과 리조트와의 제휴를 맺어 할인된 가격으로 숙박할 수 있도록 하여 직원들의 휴가 숙소를 지원한다.

**라. 건강 증진 지원**

월 2회 안마, 지압, 수기 치료를 제공하여 직원들의 건강을 관리한다. 뿐만 아니라 명상 및 상담 프로그램도 운영하여 정신적인 안정과 균형을 도모하도록 지원한다. 직원들의 안녕과 안전을 생각하여 기업 내 양호실을 설치하고, 응급의약품을 항상 구비하여 긴급 상황에도 대비한다. 더불어 매년 1회 종합 건강검진 비용을 회사에서 전액 부담하며 가족 보험료 및 의료비 또한 지원한다.

**E. 주력 사업 및 신규 사업**

 카카오의 주력 사업은 크게 세 가지로 구분할 수 있다.

텍스트, 스크린샷, 폰트, 원이(가) 표시된 사진

자동 생성된 설명

**가. IP-IT 결합 통한 글로벌 문화 생태계**

ㄱ. 카카오엔터테인먼트

카카오를 구성하는 계열사 상당 부분은 ‘IP-IT 결합 통한 글로벌 문화 생태계’ 부문에 해당한다. 그 중에서도, 이 부문의 69개사 중 44개사가 카카오엔터테인먼트와 그 산하의 소속회사이다. 카카오는 K-콘텐츠의 IP와 플랫폼의 시너지를 바탕으로 글로벌 시장 진출 및 경쟁력 강화를 추구하고 있다. 스타쉽엔터테인먼트와 이담엔터테인먼트를 보유하고 있음에도 아직 글로벌 시장에서 영향력이 작다는 평가를 받는 카카오는 최근 SM 지분 39.87%를 확보했다. SM 인수를 통해 K팝 시장에서 약 50억 달러로 평가되는 시장 규모를 갖게 됨으로써, 글로벌 엔터 사업을 확대한다는 전략이다.

- 주요 계열사: 카카오엔터테인먼트, 카카오게임즈

**나. AI·헬스케어 중심 미래 성장동력**

카카오는 미래 성장동력을 발굴하고 육성하기 위해, AI와 헬스케어 분야에서 기술 및 서비스를 개발하고 있다. 이때 블록체인 기술 및 서비스의 고도화를 진행하고 있으며, 미래의 비즈니스 파트너 발굴과 스타트업 생태계 기여를 위한 투자도 진행한다.

ㄱ. 카카오브레인

2017년 설립된 카카오브레인은 인간의 생산성 및 삶의 질을 향상시킬 수 있는 다양한 AI 모델 및 서비스를 개발하고 있다. 특히 데이터, 인프라, 모델 개발, 서비스 출시 전체를 아우르고 있는 기술인 ‘초거대 AI’를 이용하여 AI 연구와 활용 측면에서 장기적인 경쟁력을 만들어 내고 있다. 현재 카카오브레인은 p Large-scale AI Studio(초거대 AI 스튜디오), pGenerative Model Studio(생성 모델 스튜디오), pHealthcare AI Studio(헬스케어 AI 스튜디오)와 같은 3개 사업에 주력하고 있다.

ㄴ. 카카오엔터프라이즈

카카오엔터프라이즈는 클라우드를 중심으로 다양한 인공지능 영역을 아우르는 IT 전문성을 갖춘 기술 기업이다. 초거대 AI 모델, 고성능컴퓨팅(HPC) 등 고성능⬝고집적 IT 환경이 필요한 기업 및 기관에게 안정적인 클라우드 서비스를 제공하며 클라우드 네이티브 기반의 디지털 혁신을 지원한다. 다양한 파트너사와의 협력을 기반으로 동반 성장 및 건강한 디지털 생태계 구축을 위해 힘쓰며, 기술 고도화를 위한 R&D 투자를 강화하고 있다.

- 주요 계열사: 카카오브레인, 카카오엔터프라이즈, 카카오헬스케어, 그라운드엑스, 카카오벤처스, 카카오인베스트먼트

**다. 일상의 혁신을 위한 디지털 전환**

카카오는 플랫폼을 통해 이용자 및 파트너가 보다 편리하고 효율적으로 디지털 기술의 혜택을 누릴 수 있도록 디지털 전환을 돕고 있다. 현재 카카오는 핀테크, 모빌리티, 이커머스 분야에서 디지털 전환 비즈니스를 강화해 나가고 있다

ㄱ. 카카오페이

카카오페이 주요 사업부문은 결제서비스, 금융서비스 및 기타서비스로 나누어진다. 주력 부문인 결제서비스 사업은 카드정보를 입력하는 과정 없이 사전에 등록한 카드를 선택할 수 있고, 복잡한 절차 대신 안면인식, 지문인증, 비밀번호 입력만으로 결제할 수 있는 편의성을 제공한다. 금융서비스 사업은 카카오페이 모바일 플랫폼을 통해 대출상품, 투자상품, 보험상품에 대한 소비자 및 공급자의 니즈를 충족시키고, 효익을 증진하기 위한 서비스를 개발 및 제공한다. 신규사업으로는 ‘내 주변’ 서비스 고도화를 통해 가맹점주와 손님을 직접 연결시켜주는 개인형 맞춤 혜택 강화가 있다. 더불어, 금융 거래 데이터와 코GPT 2.0을 연계한 AI 기술을 토대로 금융 서비스를 제공해주는 ‘금융비서 AI’를 개발하고 있다.

ㄴ. 카카오모빌리티

카카오모빌리티는 정교한 위치 추정 기술을 활용하여 실내에서도 완벽한 길 안내를 제공하는 세계 최초의 FIN 기술을 개발하였으며, 중형택시 최초 앱미터기를 도입하는 등 모빌리티 분야의 혁신을 이끌어내고 있다. 현재 주요 산업은 AI 엔진 제작을 통해 새로운 모빌리티 시대를 구축하기 위한 자율주행 모빌리티 개발이며, 대리운전 수요 예측 시스템 또한 제작하고 있다. 더불어 자율주행 이동체들의 관제 플랫폼을 구축하고, 주행 로직과 성능을 관리하며, 모델 성능 개선에 기여하고 있다. 최근에는 플랫폼(카카오 T)를 활용하여 UAM 관제를 준비하고 있다. 영국의 UAM 기체 제조사 '버티컬 에어로스페이스'와 파트너십을 체결하고, UAM 서비스의 국내 상용화 및 글로벌 공동 사업을 위해 협력하고 있다.

- 주요 계열사: 카카오페이, 카카오뱅크, 카카오모빌리티, 카카오스타일

**F. 채용 기회**

**가. 채용 절차**

직무마다 상이하지만, 카카오 데이터 엔지니어의 경우 서류 전형 -> 코딩 테스트 -> 원격 인터뷰 -> 1차 인터뷰 -> 2차 인터뷰 -> 입사 확정 순으로 이루어진다. 이때, 불합격 시 1년간 재지원이 불가능하다.

ㄱ. 서류전형

2022년 하반기 인턴 기준, 테크 직무 공통 자기소개서 항목은 다음과 같다.

- 졸업(예정) 시기를 기재해주세요.

- 지원하는 분야와 관련하여 학습한 전공과목 혹은 외부강의를 3개 이내로 기재해주세요.

- 자신의 열정과 기술적인 전문성을 나타낼 수 있는 경험/이력/생각을 자유롭게 기술해주세요. (대회, 해커톤, 프로젝트, 오픈소스 코드 기여, 논문 등)

- 자신이 다뤄본 경험이 있는 분야에 대해 1~4단계로 표현해주세요. 4단계가 가장 높은 단계입니다. (예: DB-4 / Python-3 / Java, Spring-3)

ㄴ. 코딩 테스트

Hackerrank로 링크를 받아 문제를 푸는 방식으로 진행된다. 약 7문제 정도가 출제되며, 각 문제 당 점수 배점이 다르다. 문제는 쉬운 난이도부터 어려운 난이도 순으로 배치된다.

ㄷ. 원격 인터뷰

지원자마다 면접 방식, 면접 시간 등이 다르다. 원격 인터뷰 단계는 주로 코딩테스트 결과에 대해 추가 검증을 하고, 경력직의 경우 간단한 경력 관련 질문/답변이 진행된다. 다만, 부서에 따라 경력 관련 질문/답변은 추후 1차 인터뷰 단계에서 진행하기도 한다.

ㄹ. 1차 인터뷰

면접관 n명, 지원자 1명의 다대일 면접이며, 대략 60~90분가량 진행된다. 1분 자기소개, 자기소개서 및 포트폴리오에 적혀 있는 프로젝트에 대한 질문, 그리고 관련 분야의 이론적 지식에 관한 질문 등의 방식으로 이루어진다.

질문 예시는 다음과 같다.  
 - 진행했던 프로젝트에 대한 소개

- 딥러닝 지식 (relu가 무엇이고, 왜 사용하는가? / local minimum 해결방법이 무엇인가?)

- computer science 질문 (정렬 알고리즘 설명 / 멀티 스레드와 멀티 프로세스 차이 설명)

ㅁ. 2차 인터뷰

마찬가지로 대략 한시간가량 진행되며, 주로 1차 인터뷰에서 질문받지 않았던 기존 프로젝트와, github와 같은 개인 포트폴리오, 직무 관련 기술적 지식에 대한 질문을 받는다.

**나. 지원자격 (경력직)**

ㄱ. 필수사항

a. 데이터 마이닝/분석 인프라 또는 데이터 엔지니어링 2년 이상 업무경험 보유자

b. 하둡 에코시스템의 전반적 이해를 바탕으로 ETL 가능하신 분

c. python(필수), java, go, bash 등 언어 중 하나 이상 활용 가능하신 분

d. 새로운 기술에 대한 관심이 높고 거부감이 없는 분

ㄴ. 우대사항

 a. 프로덕션 레벨에서의 데이터 처리 인프라/파이프라인 구축 및 운영 경험이 있으신 분

 b. 데이터 분석가, 머신러닝 전문가와 협업 경험이 있으신 분

 c. 클라우드 서비스 기반 빅데이터 인프라 구축 및 운영 경험이 있으신 분

 d. Kubernetes 환경에서의 분석 및 추론 자동화에 관심 있으신 분

 e. 원천 데이터를 가치있는 결과로 만들어내는 데이터 플로우에 관심 있으신 분

**3. 직무 기본 분석 - 데이터 엔지니어**

**A. 직무 소개**

데이터 엔지니어는 원시 데이터를 수집 및 관리하는 데이터 처리 시스템을 개발하여 데이터 과학자나 분석가가 인사이트를 도출하는 것을 돕는 직무이다.

**가. 데이터 과학 분야 대표 직무**

|  |  |  |
| --- | --- | --- |
| 직무 | 업무 | 필요 능력 |
| 데이터  엔지니어 | 1. 최종 데이터 사용자의 요구사항 분석  2. 파이프라인 구축, 관리 및 유지 보수  3. 데이터 응용 솔루션 및 앱 개발 | 1. 고급 파이썬, SQL 코딩 능력  2. 데이터베이스에 대한 이해  3. IT 및 컴퓨터 시스템에 대한 이해 |
| 데이터  분석가 | 1. 데이터 전처리  2. 분석을 통한 비즈니스 인사이트 도출  3. 데이터 시각화 | 1. SQL, 엑셀, 기초 코딩 능력  2. 시각화 감각  3. 프레젠테이션, 의사소통 능력 |
| 데이터  과학자 | 1. 데이터 전처리  2. 모델링 및 검증  3. 데이터 시각화 | 1. 파이썬, R, SQL 코딩 능력  2. 머신러닝, 통계에 대한 이해  3. 컴퓨터 공학적 지식 |
| 데이터  아키텍처 | 1. 데이터베이스 설계와 이용  2. 데이터베이스 모델링  3. 데이터 표준화 | 1. 비즈니스적 능력  2. 파이프라인 설계 능력  3. 파이썬, 자바, C++ 코딩 능력[[1]](#footnote-1) |

**나. 데이터 엔지니어 선택 이유**

많은 AI빅데이터융합경영학과생들이 소프트웨어융합대학의 학과를 복수/융합 전공으로 선택하는 것을 고려하면, 데이터 엔지니어나 아키텍처를 목표 직무로 삼은 학생들이 많을 것으로 예상된다. 이 중에서 데이터 아키텍처는 신입을 뽑는 경우가 적으므로 조사 직무로 데이터 엔지니어를 선택하였다.

**B. 직무의 업무**

**가. 개발하는 시스템 목록**

ㄱ. 데이터 수집 시스템

a. Data Collection SDK

조직의 여러 프로그램 및 시스템에서 데이터 수집을 가능하게 해주는 소프트웨어 개발 도구이다.

b. Data Collection Server

조직의 여러 프로그램 및 시스템에서 수집된 데이터를 저장하는 서버이다.

ㄴ. 데이터 가공 및 저장 시스템

a. Data Pipeline

조직의 데이터를 비즈니스 요구 사항에 적합하게 필터링, 병합 등의 과정을 거쳐서 가공하는 데이터 처리 시스템이다. Data Pipeline은 데이터 처리 시간 간격에 따라 두 가지 유형으로 나누어진다.

* batch Process Pipeline

특정한 시간 간격 동안 저장된 데이터를 묶어서 불러와서 가공을 하는 형태의 데이터 파이프라인이다

* Streaming Data Pipeline

시스템이 돌아가는 와중에도 데이터를 지속적으로 업데이트 해야 할 때 사용하는 데이터 파이프라인이다.

b. Data Lake

조직에서 수집되는 모든 정형, 비정형, 반정형 데이터를 원시 형태로 저장하는 데이터 저장소이다.[[2]](#footnote-2)

ㄷ. 데이터 처리 서비스

a. Report & Analytics Service

사내의 여러 시스템에서 수집된 데이터를 변환하여 비즈니스 인텔리전스 및 분석을 실행하는 시스템이다.[[3]](#footnote-3)

b. Self Service BI Service

직원 누구나 접속하여 간단한 데이터 분석 및 시각화를 할 수 있는 서비스이다.

c. Analytics Notebook Service

데이터 처리 시스템이 돌아가는 와중에도 코드를 수정할 수 있는 ‘라이브 코딩’ 과 시각화를 돕는 플랫폼이다.[[4]](#footnote-4)

d. AB TEST Service

웹 사이트나 웹 콘텐츠의 여러가지 버전을 비교하여 최고의 버전을 가려내는 AB 테스트를 실행하는 서비스다.

ㄹ. 데이터 관리 서비스

a. Data Discovery Service

서로 다른 데이터베이스에서 사용자가 필요로 하는 데이터를 검색하여 제공받는 시스템이다.[[5]](#footnote-5)

b. Data Access Control Service

직원들에게 사내 데이터에 액세스할 수 있는 권한을 부여하는 시스템이다.[[6]](#footnote-6)

c. Data Quality Monitoring Service

조직에서 생성되는 모든 데이터들의 정밀도, 일관성, 신뢰도과 같은 품질적 요소를 평가하고 관리하는 시스템이다.

d. Data Lineage Service

현재 쓰고 있는 데이터의 생성 위치, 전처리 과정, 활용 역사 등의 계보를 보여주는 시스템이다.[[7]](#footnote-7)

ㅁ. 데이터 가공 서비스

a. Self Service Data Preparation Service, Automated EDA Report Service

데이터 처리 과정에서 80% 이상의 시간을 차지하는 데이터 요청, 데이터 변환 및 정제, 프로파일링을 분석가 대신 수행하는 서비스이다.[[8]](#footnote-8)

b. Real Time Event Trigger Service

특정 이벤트나 조건의 발생을 기반으로 데이터를 실시간으로 처리하는 시스템이다.

ㅂ. 데이터 활용 서비스

a. Data Query API Service

사내 데이터 이용자들의 컴퓨터와 데이터 베이스를 연결하는 API를 만들어서, 이용자들이 데이터를 찾을 수 있게 하는 서비스이다.[[9]](#footnote-9)

**C. 직무의 고객**

**가. 내부 고객**

ㄱ. 사내 데이터 과학자 및 모든 직원들

카카오의 데이터 과학자가 쉽게 전처리를 할 수 있도록 도와주는 자동 데이터 전처리 서비스나 모든 직원들이 쉽게 데이터 시각화를 할 수 있는 Self Service BI Service를 개발한다.

ㄴ. 사내 웹 개발자, 마케팅 팀

웹 사이트나 웹 콘텐츠의 여러가지 버전을 비교하여 최고의 버전을 가려내는 AB 테스트 시스템을 설계해준다.

**나. 외부 고객**

ㄱ. 광고를 원하는 기업

이용자의 방문 기록, 활동 로그, 검색 기록 등을 이용하여, 다른 기업의 광고를 가장 적합한 이용자에게 보여주는 맞춤형 광고 시스템을 설계한다.

**D. 직무의 근무환경**

**가. 재택근무**

조직 별로 재택근무 비율을 자율적으로 결정할 수 있어서 특정 요일에는 사무실로 출근하고 나머지 요일은 재택근무를 하는 것이 가능하다.

**나. 자율 출퇴근제**

총근무시간인 8시간만 준수하면 출퇴근 시간은 개인이 자유롭게 선택할 수 있다.

**E. 직무수행 시 만족감 또는 성취감**

**가. 프로젝트 마무리의 성취감**

데이터 엔지니어는 업무 특성상 성취감을 느끼기는 어렵지만 오랜 기간 준비했던 프로젝트가 끝나고 서비스를 시작했을 때, 그리고 서비스에 대한 사용자의 액션이 나타날 때 성취감을 느낀다.

**나. 서비스가 성장하는 것을 느낄 때**

자신이 엔지니어링 한 서비스가 발전해 나갈 때 성취감을 느낀다.

**다. 문제를 해결할 때**

다양한 관점에서 고민하여 문제를 해결하거나 새로운 기능을 개발할 때 성취감을 느낀다.

**라. 자신이 발전하는 것을 느낄 때**

새로운 기술을 배워 나가고, 다양한 분야의 전문가와 협업하면서 이전에 알지 못했던 도메인과 비즈니스 영역을 배우며 성취감을 느낀다[[10]](#footnote-10)

**F. 직무수행 시 힘든 점**

**가. 신체적 피로**

업무 중 오래 앉아 있어야 하는 것이 힘들다. 또한 장시간 컴퓨터를 사용하므로 눈이 아프고 일자목이 될 가능성이 있다.

**나. 야간 작업**

다른 개발자 직무처럼 야간에 큰 시스템 장애가 발생하면 긴급하게 대응해야 하고, 사고 기간에는 단기간에 일을 빠르게 처리해야 하므로 야간 작업을 해야 하는 경우가 생긴다.

**G. 직무의 전망**

**가. 마이데이터 시대**

데이터 3법의 시행 이후 여러 기업이 보유한 개인정보 데이터를 하나의 플랫폼에 모을 수 있는 마이데이터 산업이 발달하면서 큰 규모의 데이터를 다루는 사람이 많이 필요해졌다. 그래서 데이터 엔지니어를 채용하는 공고는 계속 늘고 있다.

**나. 많은 수요**

데이터 엔지니어는 다양한 데이터 업무를 종합적으로 수행하는 직업이기 때문에, 조금만 데이터를 다뤄도 채용 공고에 데이터 엔지니어를 구한다고 올리는 경향이 있으므로 수요가 많다. 또한 앞으로 데이터를 많이 가지고, 잘 다룰 줄 아는 기업이 시장에서 성공하는 구조이기 때문에, 그러한 업무를 담당할 데이터 엔지니어에 대한 수요는 많을 수밖에 없다.

**다. 적은 공급**

분석과 모델링 업무는 하고 싶어하는 사람이 많지만, 이 업무를 돕는 데이터 엔지니어는 하고 싶어하는 사람이 적으므로 다른 데이터 직무에 비해 취업에 유리하다.

**4. 직무 요구 역량 분석 – 데이터 엔지니어**

**A. 핵심 역량**

**가. 파이썬 사용 역량**

파이썬 개발환경인 파이참(PyCharm)을 사용해서 개발이나 코딩이 주로 이루어진다. 파이썬과 그와 관련된 개발환경에 대해 익숙해지는 것이 좋다.

- 프로그래머스, 백준, 캐글 등의 코딩 사이트를 통해서 파이썬 및 코딩 역량을 키울 수 있다.

- AI빅데이터융합경영학과에서 진행하는 방학 파이썬 스터디에서 코딩의 기초를 다질 수 있다.

**나. 하둡(Hadoop), 스파크(Spark) 사용 역량**

프로젝트를 할 때 주로 사용되는 프레임워크. 데이터 관리 및 처리에 필요한 기술이다.

- 하둡(Hadoop): 프로그래밍 모델을 통해, 여러 대의 컴퓨터를 하나로 묶어서 대용량의 데이터를 처리할 수 있게 해주는 데이터 프레임워크. 페이스북, SK, 아마존 등 대용량 데이터를 다루는 기업들이 사용한다.

- 스파크(Spark): 위의 하둡과 같은 데이터 처리 시스템이다. 반복적인 데이터 처리가 빠르다는 장점이 있고, 최근 데이터 처리에 하둡과 스파크를 같이 쓰는 경우가 많아지고 있다.

**B. 대학 교육**

**가**. AI빅데이터융합경영학과와 같이 데이터 관련 학과라면 다음과 같은 수업의 내용이 직, 간접적으로 도움이 된다. 또한 데이터 직무에 관심이 있다는 것도 나타낼 수 있다.

- ‘데이터사이언스개론’

- ‘AI개발환경의이해’

- ‘머신러닝’

- ‘데이터베이스와 SQL’

**나**. 다음의 수업과 학회 활동으로 프로그래밍에 익숙해지는 것이 추후 코딩 실력 향상에 도움이 될 것이다.

- ‘AI빅데이터프로그래밍I’

- ‘객체지향프로그래밍’

- ‘D&A 학회 basic session’

**C. 대학원 과정**

**가**. 데이터 분석, 데이터 모델링을 연구하는 분야에서는 대학원 과정 유무 등, 해당 분야를 깊이 연구해본 경험이 있는 사람들을 선호한다. 다음과 같은 대학원에서 관련 분야를 연구해볼 수 있다.

- 국민대 경영대학원 AI빅데이터전공: 비전AI, 클라우드머신러닝

-성균관대학교 데이터사이언스융합학과: 자연어처리, 서버시스템이해, 데이터 마이닝

**D. 교내 활동**

**가. D&A 학회**

ㄱ. D&A 학회 Basic session

ㄴ. D&A 학회 ML(Machine Learning)session

ㄷ. D&A 학회 Deep session

ㄹ. D&A 학회 Conference session

**나. 실무형 핵심직무 전문가 양성과정(CoREP)**

**다. 파이썬 프로그래밍 스터디**

**라. 취업연계 인턴십 프로그램**

**E. 자격증**

**가. 컴퓨터 활용 관련 자격증**

|  |  |
| --- | --- |
| 자격증 | 내용 |
| 컴퓨터 활용 능력 1급 | 업무 능률을 위해 사무/자동화의 필수 프로그램인 스프레드시트   데이터베이스 활용능력을 갖춘다. |
| 마이크로소프트 오피스    활용 능력 자격증(MOS) | 사무현장에서 프로그램에 대한 이해도와 활용력을 기르기 위해 엑셀, 엑세스,   워드, 파워포인트, 아웃룩 등 다양한 오피스 프로그램을 활용한 필수기능 및   실무 능력을 갖춘다. |
| I ITQ 정보기술자격증 | 실무중심의 작업 문제로 현장 활용도를 높이고, 정보기술 활용 능력을   갖춘다. |
| 디지털 정보 활용    능력 자격증 (DIAT) | 디지털정보시대에 기본적인 정보통신기술, 정보처리기술의 활용분야에 대해   학습이나 사무업무를 수행할 수 있도록 능력을 갖춘다. |

**나. 데이터 업무 관련 자격증**

|  |  |
| --- | --- |
| 자격증 | 내용 |
| 데이터분석   준전문가(ADsP) | 데이터 분석 기획 및 데이터분석 등의 직무를 수행하기 위해 데이터 이해의   기본지식을 쌓는다. |
| 데이터분석   전문가(ADP) | 프로세스 혁신 및 마케팅 전략 결정 들의 과학적 의사결정을 도출하기 위해   데이터 처리 기술에 대한 지식을 쌓는다. |
| 데이터 과학자  ]능력인증자격(DSAC) | 데이터 과학자로서 유능한 실무진이 되기 위한 구체적 역량을 습득한다. |
| SQL 전문가 (SQLP) | 데이터 모델링에 기본 지식을 바탕으로 SQL 작성, 성능 최적화 등   데이터베이스 개체 설계 및 구현 등에 대한 전문지식 및 실무적 수행 능력을   습득한다. |
| 데이터 아키텍처  준전문가(ADsP)자격증 | 전사아키텍처에 대한 지식을 바탕으로 데이터 요건분석, 데이터 표준화, 데이터   모델링에 대한 기본지식 및 실무적 수행 능력을 필요로 한다. |
| 데이터 아키텍처   전문가(DAP)자격증 | 전사아키텍처 및 데이터품질관리에 대한 지식을 바탕으로 데이터 요건분석,   데이터 표준화, 데이터 모델링, 데이터베이스 설계와 이용에 대한 전문지식 및   실무적 수행 능력을 그 필수로 한다. |
| 빅데이터 분석 기사   자격증 | 대용량의 데이터 집합으로부터 유용한 정보를 찾고 결과를 예측하기 위해 목적에 따라 분석기술과 방법론을 기반으로 정형/비정형 대용량 데이터를 구축, 탐색, 분석하고 시각화를 수행하는 업무를 수행한다. |

**다. 외국어 관련 자격증**

|  |  |
| --- | --- |
| 자격증 | 내용 |
| TOEIC | 일상 커뮤니케이션 또는 비즈니스 등에 필요한 실용 영어 능력을 갖춘다. |
| TOEFL | 영어권 대학이나 대학원 입학 필수 요건, 국제적인 커뮤니케이션 역량을 확보하는데 있어 TOEIC보다 전문적인 영어 능력을 갖춘다. |

**F. 태도 및 성향**

데이터 직무 특성상 중간에 전처리를 잘못하면 데이터 과학자들에게 쓸모없는 데이터를 줄 수 있기 때문에, 정확도를 높이는 세심함이 필요하다. 또한 명확한 의사소통 능력이 회의에서 중요하다. 그리고 직무를 하면서 꼼꼼한 성향을 가지면 발생할 수 있는 사고들을 방지하는 것에 도움이 된다.

- 의사소통 능력

- 꼼꼼함

- 세심함

**G. 필요경험**

데이터와 관련한 다양한 소식들을 꾸준히 접해보는 것이 중요하다. 프로젝트를 하더라도 최대한 데이터 엔지니어의 일과 연관되는 활동(데이터 수집, 처리, 보관, 분석 과정)을 찾아서 경험하는 것이 중요하다. 위에서 언급한 다양한 교육 프로그램(부트캠프)에 참여해보는 것도 큰 도움이 된다.

- 구글 개발자 머신러닝 부트캠프

- 부스트 캠프 AI Tech 부분

**5. 표준 커리어플래너 – 데이터 엔지니어**

|  |  |
| --- | --- |
| 1학년 | 목표: 직무 기본지식 습득 및 목표 진로 설정 |
| 1학기 | **[수강 계획]**  1) 본 전공 이수  - 데이터사이언스개론: 데이터 사이언스에 대한 원론적 내용 수강  - 현대경영과기업가정신: 경영학에 대한 원론적 내용 수강  - AI빅데이터프로그래밍1: 입문 수준의 python 프로그래밍 기술  - 경영영어실습: 비즈니스에서 필요한 영어 의사소통능력 강화  - 인공지능수학: 빅데이터 분석을 통해 경영현상을 이해하기 위한 수학적 기초 배양  **[교내 활동]**  1) 중앙 동아리  - 사회적인 활동능력 함양  2) 학생회  - communication 능력 함양 |
| 하계방학 | **[교내 활동]**  1) 주니어 코렙  - 직무 로드맵 설정  2) 여름 python 스터디  - 기초 프로그래밍 능력 함양 및 AI프로그래밍1 복습  **[개인 활동]**  1) 토익문제 꾸준히 풀기  - 영어독해 능력 및 토익 자격증 공부 베이스  2) ADsP(데이터분석준전문가) 자격증 취득  - 데이터 분석 업무에 대해 대부분의 기업에서 기본적으로 요구하는 자격증으로,  저학년에 취득하는 것이 좋다.  3) 독서  - Data science 및 engineering에 관한 책 읽기  - 폭 넓은 영역 독서 |
| 2학기 | **[수강 계획]**  1) 본 전공 이수  - 경영통계: 현대 경영학을 공부하는데 필요한 수학 공부  - 경영정보학원론: 경영정보학에 대한 원론적 내용 수강  - AI빅데이터프로그래밍2: AI 알고리즘의 구현을 위한 객체지향 프로그래밍의 개념  - 회계학원론: 회계와 관련된 기본적인 개념 학습  **[교내 활동]**  1) Global Peers  - 본교 초청 외국인 학생(교환학생)과의 언어교환, 문화교류 등을 통해 외국어 역량을 증진할 수 있는 프로그램  2) 경영대학 Global Buddy Program  - 경영대학에 재학 중인 외국인 학생들과 교류하며 외국어 역량을 함양하는 프로그램  - 다양한 source에서 data를 수집하고 가공하는 Data Scientist는  영어 및 외국어에 대한 스킬을 갖추어야 한다.  3) 경력개발지원단 추천 프로그램 개인상담  - 개인상담은 월 2회, 학기당 4회까지  - 직업 흥미검사, 진로상담, 취업상담 등 학교생활 전반에 대한 상담 가능 |
| 동계방학 | **[개인 활동]**  1) 컴퓨터활용능력 1급 필기  2) 겨울 python 스터디  3) 독서  - Data science 및 engineering에 관한 책 읽기  - 폭 넓은 산업영역 독서 |

|  |  |
| --- | --- |
| 2학년 | 목표: 데이터 엔지니어 직무 구체화 및 교외 활동  \* 2~3학년 전반에 걸쳐 진행되는 다양한 빅데이터 분석 공모전에 참여해 엔지니어링 파트에 대한 경험을 쌓는 것이 매우 중요하다. |
| 1학기 | **[수강 계획]**  1) 본 전공 이수  - 빅데이터와 시각화: 빅데이터 가공처리법과 시각화 기술 습득  - 수리통계: 통계학의 기초개념 학습  - 회귀분석: 여러 변수들 간의 선형적 함수관계를 탐색과 R 이해  - AI개발환경의 이해(빅데이터엔지니어링)  : 빅데이터 분석을 위한 개발환경을 구축하는 방법 학습  - 공간빅데이터분석: 기업 경쟁력 향상을 위한 공간 빅데이터 분석능력 함량  2) 프로그래밍에 대한 이해도 높이기 위해 소프트웨어융합전공 전공과목 수강  **[교내 활동]**  1) D&A Basic session: 파이썬 프로그래밍과 시각화, EDA Competition 수행  **[교외 활동]**  1) 프로그래밍 플랫폼을 통한 프로그래밍 공부: 코드잇, 인프런 등  2) 데이터 분석 공모전 출품 |
| 하계방학 | **[개인 활동]**  1) SQLD (SQL 개발자) 자격증 취득  - 데이터모델링에 기본 지식을 바탕으로 SQL 작성, 성능 최적화 등  데이터베이스 개체 설계 및 구현 등에 대한 전문지식 및 실무적 수행 능력  - 2학기 때 듣게 될 데이터베이스와SQL 과목을 미리 예습하는 겸 익혀둔다.  2) 머신러닝 예습 스터디 진행  - 2학기 때 듣게 될 머신러닝 과목을 미리 팀원을 구성해 스터디하여 예습한다. |
| 2학기 | **[수강 계획]**  1) 본 전공 이수  - 머신러닝: 심화된 파이썬 실습을 병행하여 머신러닝 프로젝트 수행  - 데이터베이스와SQL: SQL 작성 능력을 배양하여 데이터 조작 및 추출  - 경영데이터분석: 엑셀의 분석, 최적화, 프로그램 기능을 학습  - 선형대수: 행렬과 벡터공간에 대한 교육 및 파이썬 프로그래밍 실습  2) 소프트웨어융합 전공 과목 복수 전공  **[교내 활동]**  1) D&A ML session: 파이썬 프로그래밍과 머신러닝 기술, ML Competition 수행  **[교외 활동]**  1) 프로그래밍 플랫폼을 통한 프로그래밍 공부: 코드잇, 인프런 등  2) 데이터 분석 공모전 출품 |
| 동계방학 | **[개인 활동]**  1) 컴퓨터활용능력 1급 자격증 취득  2) 국민대 경영대학 국제교류 프로그램 – 교환학생 참여 |

|  |  |
| --- | --- |
| 3학년 | 목표: 데이터 엔지니어에 대한 직무 역량 심화  \* 2~3학년 전반에 걸쳐 진행되는 다양한 빅데이터 분석 공모전에 참여해 엔지니어링 파트에 대한 경험을 쌓는 것이 매우 중요하다. |
| 1학기 | **[수강 계획]**  1) 본 전공 심화 과목 이수  - 텍스트데이터분석, 통계소프트웨어와모델링, 공간빅데이터분석, 다변량통계분석 ,회귀분석, 등 심화 전공 과목 수강  2) 프로그래밍에 대한 이해를 높이기 위해 소프트웨어융합 전공과목 수강  - 소프트웨어융합대학 교과목: 컴퓨터구조, 데이터베이스 등  프로그래밍의 원론적 내용 수강  **[교내 활동]**  1) D&A Deep session: 딥러닝에 필요한 개념과 기술 심화학습 |
| 하계방학 | **[개인 활동]**  1) 토익 900점 이상 취득하기 위해 토익 공부  2) GitHub를 통한 개인 포트폴리오 제작  3) 데이터 마이닝, ML, 딥러닝 등 빅데이터/인공지능 응용 분야 논문 공부 |
| 2학기 | **[수강 계획]**  1) 본 전공 이수  2) 소프트웨어융합 전공 과목 복수 전공  **[교내 활동]**  1) D&A 학회의 Conference session을 이수 |
| 동계방학 | **[개인 활동]**  1) ADP, SQLP, IBM Certified Data Engineer 등 자격증 취득  2) 경영대학 국제교류 프로그램 – 해외인턴십  3) 학부 연구생 |

|  |  |
| --- | --- |
| 4학년 | 목표: 실무 경험을 쌓고 취업을 위한 준비 시작 |
| 1학기 | **[수강 계획]**  1) 본 전공 심화 과목 이수: ‘캡스톤디자인’  2) 복수 전공(소프트웨어융합학과) 심화 과목 이수  **[교내 활동]**  1) 경영대의 국제교류 프로그램 또는 해외 인턴십 참여  2) 실제 기업의 데이터를 활용하는 공모전 ‘빅콘테스트’ 참여 |
| 하계방학 | **[개인 활동]**  1) 토익 900점 이상 취득 목표 공부  2) 깃허브를 통한 포트폴리오 제작  3) 다양한 논문을 찾아보며 데이터 엔지니어링에 대한 지식 습득 |
| 2학기 | **[수강 계획]**  1) 본 전공 심화 과목 이수: ‘캡스톤디자인’  2) 복수 전공(소프트웨어융합학과) 심화 과목 이수 |
| 동계방학 | **[개인 활동]**  1) 카카오(목표기업)에 맞는 자소서, 이력서와 면접 준비 시작  2) 심화 공부를 원하고 필요하다면 대학원 석사과정 준비  3) 빅콘테스트, 빅매칭캠프, kaggle 등의 큰 공모전 수상으로 취업연계 |

**Ⅳ. 결론**

**1. 결론 및 요약**

지금까지 포털 및 기타 인터넷 정보 매개 서비스 산업의 카카오, 그 중에서도 데이터 엔지니어(Data Engineer)라는 직무에 대해서 알아보았다. 기존 어느 시점의 비즈니스 결과라고 취급되었던 데이터는 현재 소비자의 행동 패턴, 니즈와 결합하여 기업에서 데이터를 더 가치있는 분야로 인식하며 넓게 사용되고 있다. 생산되는 데이터량은 현재 매우 빠르게 증가하고 있으며, 그래프와 같이 데이터량은 두배씩 증가하는 추세이다. 때문에 모든 산업, 모든 기업에서 데이터를 다룰 정도로 그 중요성은 계속 높아지고 있다.[[11]](#footnote-11)

텍스트, 스크린샷, 라인, 그래프이(가) 표시된 사진

자동 생성된 설명그 중 카카오는 국내 대표 IT기업에 걸맞게 다양한 데이터를 활용하여 서비스를 제공하고 있다. 거의 모든 국민이 사용하는 카카오톡부터 카카오페이, 카카오모빌리티와 같은 서비스를 통해 많은 양의 유저 데이터를 수집하고 이를 다시 활용해 서비스를 제공한다. 예시로 이러한 유저 데이터를 바탕으로 사용자의 결제, 관심, 소비성향 등을 분석해 카카오페이에서는 소비패턴에 맞는 금융상품 또는 카드를 추천 및 특정 가맹점의 할인 혜택 추천하는 서비스를 제공하고 카카오모빌리티에서는 택시 호출이 발생하는 장소와 시간을 예측하는 택시수요예측 서비스를 제공한다.

데이터 엔지니어는 이러한 데이터 관련 모든 서비스가 이뤄지기 위한 기반이라고 할 수 있다. 데이터를 수집, 가공, 저장한 후 제공하는 역할이기에 다른 데이터 관련 직무보다는 기술적인 성향을 더 필요로 하지만, 기업에서 데이터를 다루기 위해서는 반드시 필요한 직무이다. 또한 기업에서의 큰 데이터를 다룰 수 있는 데이터 엔지니어의 수요가 많고 공급은 적기에 데이터 관련 직무 중에서 고민하고 있는 학생에게는 고려해볼만한 충분히 매력적인 직무라고 생각한다.

**2. 보고서를 읽는 대학생들에게**

아마 이 보고서를 읽는 분들은 대부분 데이터 엔지니어를 꿈꾸고 있거나, 데이터 관련 직무에 관심이 있는 분들일 것이다. 그리고 자신의 목표 직무를 이루기 위해서 무엇을 해야 할지 고민하고 있을 수 있다. 우리 팀도 데이터 관련 직무를 꿈꾸는 학생들로 이루어졌지만, 모일 당시에는 직무에 필요한 역량, 준비과정, 정보 등을 잘 알지 못했고 아직 직무 자체를 정하지 못한 팀원도 있었다. 하지만 이 공모전의 다양한 활동을 통해 어떤 데이터를 어떻게 다루는지 등 데이터 엔지니어의 핵심 역량과 스펙 등을 알 수 있었고, 역량, 스펙을 채우기 위해 커리어 플래너를 작성하여 자신의 목표 직무를 구체화하였다. 이 보고서를 읽는 모든 분들도 본 내용을 참고한다면 진로 고민에 도움이 될 것이라고 생각한다. 또한 목표 직무를 탐색하는 활동을 해보고 싶다면, 주니어 코렙에 신청해 보는 것도 추천한다. 목표 직무를 설정하는 것은 물론, 산업/기업 분석, 인터뷰 등 모든 과정을 직접 경험해 볼 수 있다. 직무를 고민하는 모든 학생들을 응원한다.

**3. 개인 별 소감**

권민지: 주니어 코렙 하면서 뚜렷한 목표를 설정할 수 있었고 그에 맞는 커리어 플래너를 세우면서 내가 언제 무엇을 해야 하는지 파악할 수 있었습니다.

박준하: 주니어 코렙을 하기 전에 학과 전공을 아무 생각 없이 따라가다 보면 뭔가 데이터 관련 분야로 취업할 수 있을 거라는 근거 없는 자신감으로 대학 생활을 하고 있었습니다. 그런데 이번 방학에 주니어 코렙 활동을 통해서 데이터 직무에 대한 깊이 있는 이해를 할 수 있었습니다. 또한 저의 목표 직무를 확실하게 정하고 그에 대한 구체적인 커리어 플래너를 작성함을 통해 제가 앞으로 무엇을 해야 하는지 명확하게 알게 되었습니다.

신지후: 6주동안 주니어 코렙 활동을 하며 평소에 배우지 못하는 전화예절이나 이메일 예절, 인사예절 등 배울 수 있어서 새롭고 의미 있었습니다. 더불어, 단순히 꿈꿨던 미래를 구체적으로 계획해보고 조사해볼 수 있는 기회였습니다.

심재민: 주니어 코렙에 참가한 지난 6주는 막연했던 진로 계획을 되돌아보고, 구체화할 수 있었던 활동이었습니다. 또한, 현직자 컨택이나 팀 활동 동영상 제작 같은 하기 전에는 불가능해 보였던 일을 팀원들과 함께 헤쳐 나가면서 앞으로의 취업 준비에 대한 자신감이 생겼습니다.

이호영: 커리어 플래너 작성, 멘토링, 현직자 인터뷰 등 지금까지 주니어 코렙 활동을 하면서 꿈과 목표를 정할 수 있었던 의미 있는 경험이었습니다.

안병민: 주니어 코렙 프로그램의 체계적인 활동들을 통해 목표 직무에 대해 진지하게 고민해 볼 수 있는 시간이었습니다. 같은 목표 직무를 가진 팀원 들과의 활동을 통해 더 좋은 결과를 만들 수 있었고, 목표 직무를 찾으며 앞으로의 계획을 미리 구상해 보는 의미 있는 방학을 보낼 수 있었습니다.

정용재: 의미 없는 대학생활을 이어가던 중에 주니어 코렙에 참여하게 되었습니다. 덕분에 이 학과를 다니며 어떤 활동을 하여 어떤 길로 나아갈 수 있을지를 알 수 있게 되었고, 뜻깊은 시간을 보낸 것 같아 참여하길 잘 했다고 생각합니다.

홍예진: ‘어떻게든 되겠지’ 라고 생각하며 대학생활을 했던 것에 대해 많이 반성했습니다. 약 한 달이라는 시간 동안 혼자서는 하기 어려웠을 현직자 인터뷰, 선배와의 만남 등을 통해 많이 배웠고 성장한 거 같아 뿌듯합니다.

1. [[정보TALK] Data Architecture이 뭘까요? - DACON](https://dacon.io/forum/405922) [↑](#footnote-ref-1)
2. [빅데이터로 가치를 만드는 호수 ‘데이터 레이크’ 이야기 | 인사이트리포트 | 삼성SDS (samsungsds.com)](https://www.samsungsds.com/kr/insights/big_data_lake.html) [↑](#footnote-ref-2)
3. [보고 및 분석 전략 | 마이크로소프트 런 (microsoft.com)](https://learn.microsoft.com/en-us/dynamics365/guidance/implementation-guide/business-intelligence-reporting-analytics-reporting-analytics-strategy) [↑](#footnote-ref-3)
4. [Azure Synapse Analytics Notebook 개요 - Azure Synapse Analytics | Microsoft Learn](https://learn.microsoft.com/ko-kr/azure/synapse-analytics/spark/apache-spark-notebook-concept) [↑](#footnote-ref-4)
5. [Overview (discoverydataservice.org)](https://www.discoverydataservice.org/Content/Overview.htm) [↑](#footnote-ref-5)
6. [데이터 액세스 제어 - Satori (satoricyber.com)](https://satoricyber.com/data-protect-guide/data-access-control/) [↑](#footnote-ref-6)
7. [[기획특집] 보이는 데이터 거버넌스, ‘데이터 리니지’ < 기획특집 < 기획특집 < 기사본문 - 아이티데일리 (itdaily.kr)](http://www.itdaily.kr/news/articleView.html?idxno=83523) [↑](#footnote-ref-7)
8. [IDG\_Self\_Service\_Data\_Preparation.pdf (zalesia.com)](http://www.zalesia.com/down/IDG_Self_Service_Data_Preparation.pdf) [↑](#footnote-ref-8)
9. [API란 무엇인가요? - 애플리케이션 프로그래밍 인터페이스 설명 - AWS (amazon.com)](https://aws.amazon.com/ko/what-is/api/) [↑](#footnote-ref-9)
10. [2021 우아한형제들 경력 개발자 인터뷰 #1편 | 우아한형제들 기술블로그 (woowahan.com)](https://techblog.woowahan.com/2714/) [↑](#footnote-ref-10)
11. 데이터넷, “데이터 폭증으로 앞당겨진 디지털 수요 적극 대처해야” 데이터넷, 2021년 11월 12일, https://www.datanet.co.kr/news/articleView.html?idxno=166457 [↑](#footnote-ref-11)