

데이터 사이언티스트의 일상

(Focus on 하나카드)

하나카드 데이터 본부
최병정 상무

하나카드 트래블로그

혁신금융서비스: 해외여행 특화서비스



M 하나머니

트래블로그

400만

무료환전 통화 41종으로 확대

4월 중 예정

430만

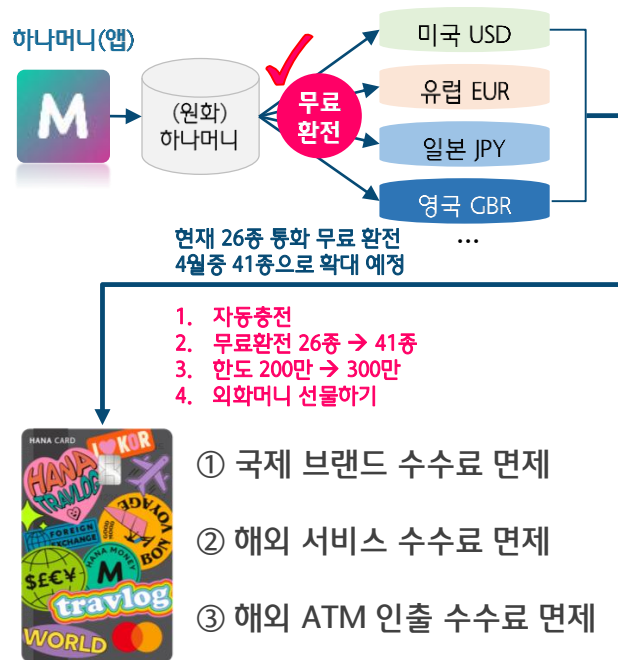
가입자

1조5천억

환전액

1조5천억

이용액



하나카드 트래블로그

혁신금융서비스: 해외여행 특화서비스



[해외여행 일타강사 EP.01 🗺️] 슬기로운 여행 생활,
여행 머니 편! → [허준](#)

[해외여행 일타강사 EP.02 🗺️] 슬기로운 여행 생활,
환율과 여행 편!

[해외여행 일타강사 EP.03 🗺️] 슬기로운 여행 생활,
추천 여행지 편!

데이터 사이언티스트의 일상

데이터 사이언티스트로 살아가기

요즘은 이른바 빅데이터와 데이터 사이언스, 인공지능의 시대입니다.

이 시대에 데이터 사이언티스트가 가져야 할 **세가지 Core Competency**는 ①**Analytics(AI 알고리즘)**, ②**IT Skill(DB, SQL)**, ③**Domain(Business) Knowledge**입니다. 세 가지 모두 필수적인 소양이지만, 특히 **도메인 지식이 없는 데이터 사이언티스트를 논하기 어렵** 것입니다. 그러나 현실적으로 학교에서 비즈니스 교육을 제공받기 어렵기 때문에 학생들은 학교의 교육과정만으로 이 모두를 갖추는 것은 매우 어려운 일입니다. 더군다나 세가지 Competency를 각각 **Silo하게 배우는 것이 아니라 함께 통합적으로 학습**해야 하기 때문에 더욱 더 어렵다고 생각합니다.

지금까지 25년 이상 **SAS**(고객사 포함)와 **하나카드**에서 경험한 수많은 프로젝트와 데이터 사이언티스트를 보며 느낀 여러 가지 **경험을 공유**하고자 합니다.

하나카드

기본 현황



하나카드 핵심역량

하나카드는 1,250만의 신용/체크카드 회원을 보유하고 있으며, 금융과 디지털 기술을 겸비한 융합형 내부 인재를 육성하여 수십억 건의 거래 데이터 등을 다양한 분야에서 활용하고 있습니다.

데이터 보유



카드 회원
1,250만 명



연간 카드 거래
18억 건



가맹점
400만 개

하나카드 기본 현황



데이터 역량

하나카드

ORACLE

python™

sas R

- 내/외부 연수를 통하여 데이터 분석 전문가 육성
- 누적 수료 : 219명

하나금융융합기술원

상품 추천
예측 모형

금융 이용
예측 모형

통화 최적 시간
예측 모형

신규 발급
효율 모형

고객센터
인입 예측 호형

상담사
스케줄 관리 모형

- 그룹 공동단위 연구 진행
- 총원 58명(박사 24명 / 석사 26명 등)

하나카드

기본 현황



데이터 활용



리스크 관리



한도 관리



상품 개발



영업 마케팅



고객 관리



소비자 보호



온라인 서비스

What is the Data Science?

DATA



DEPLOYMENT



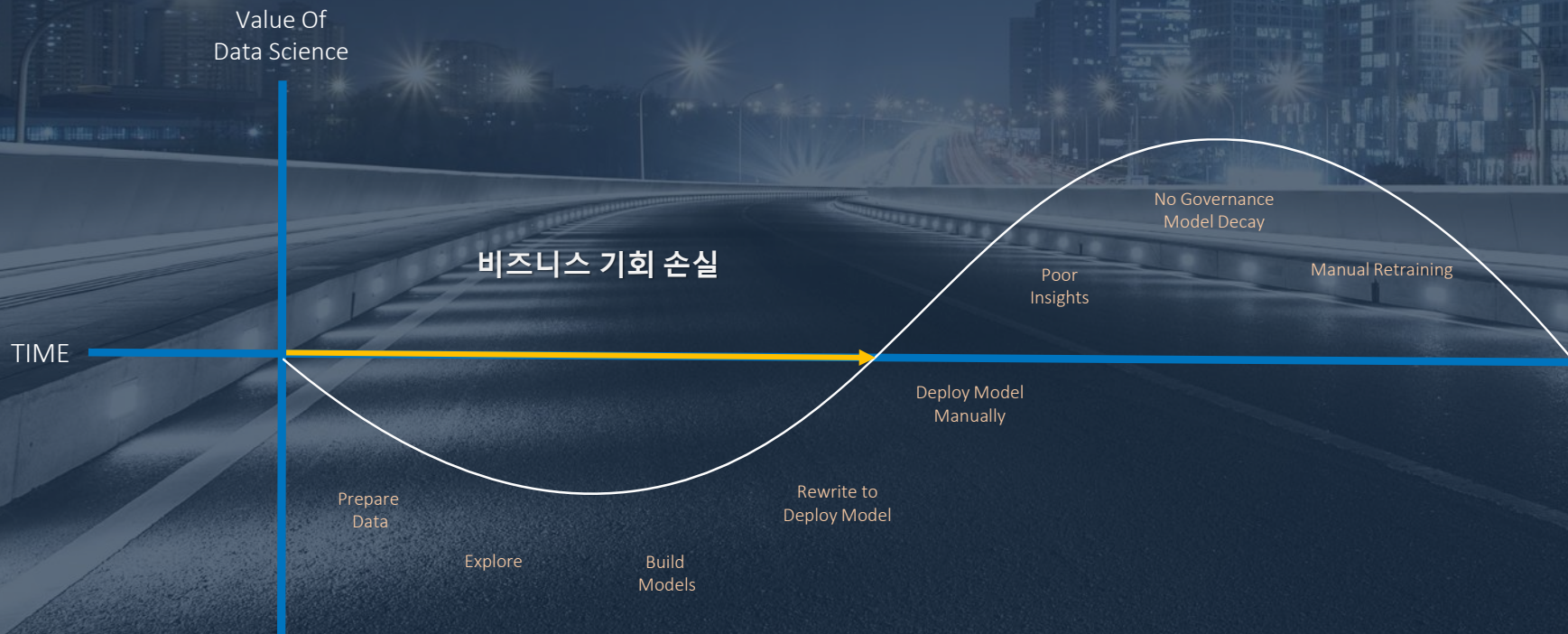
DISCOVERY

ACCELERATE

The **Data Science** Life Cycle

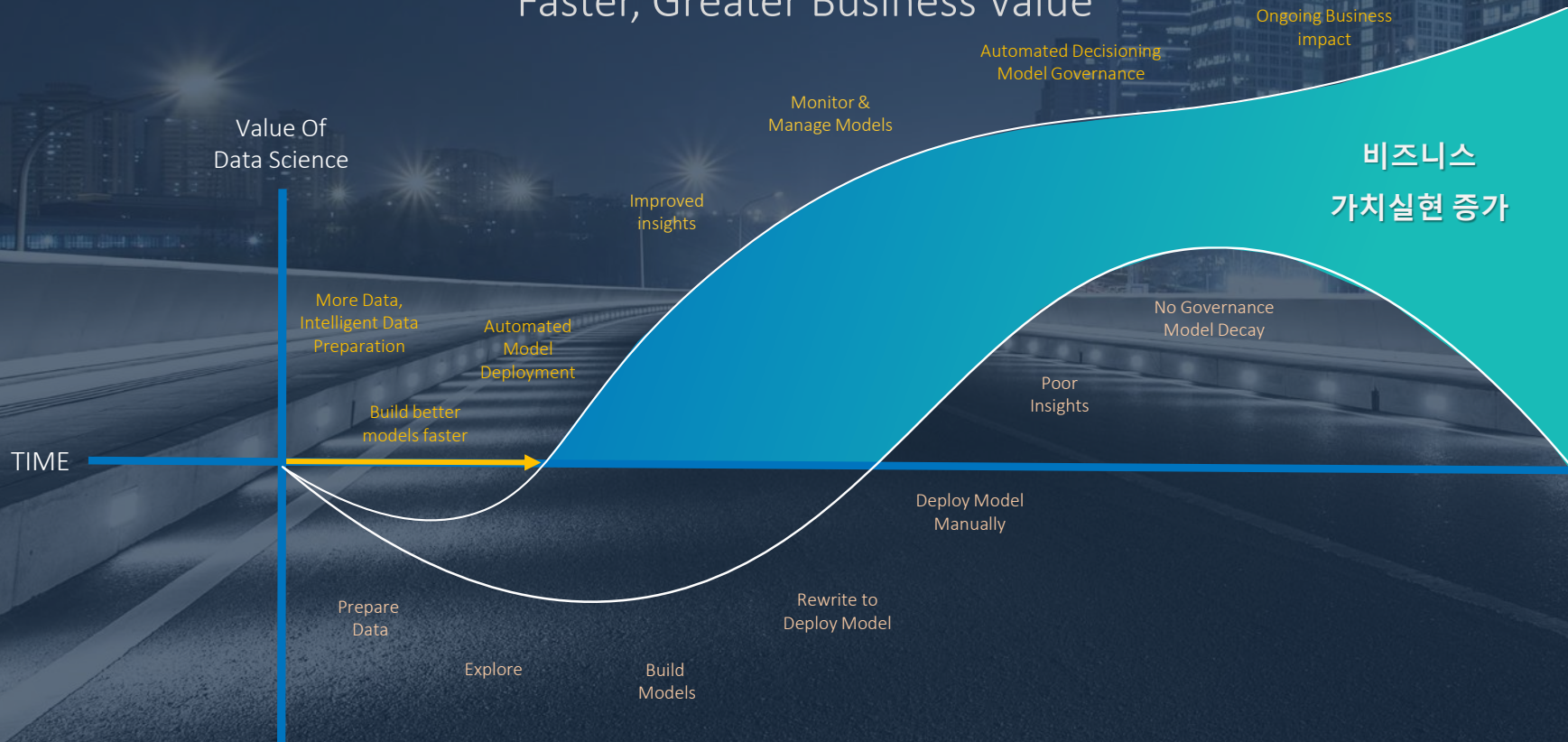
(분석의 가속화, 운영의 가속화)

Value of Data Science



분석의 가속화, 운영의 가속화 (AIOps)

Faster, Greater Business Value



AI로 경영하라, 이준기 교수

“기초적 판단은 인공지능이, 최종 선택은 인간이...”



프로로그: 인공지능 시대, 비즈니스는 어떤 모습인가

1장 인공지능은 당신의 비즈니스를 어떻게 바꾸는가

2장 인간을 위협하는 인공지능의 정체

3장 인공지능의 의사결정, 어디까지 믿을 수 있나

4장 AI로 경영하는 사람들

5장 인간의 머리와 인공지능의 다리를 합치는 법

6장 그래서 우리는 인공지능을 어떻게 사용할 것인가

AI로 경영하라, 이준기 교수

5장. 인간의 머리와 인공지능의 다리를 합치는 법

퍼플 피플 = 인공지능 시대를 이끌어 갈 데이터 통역가

지금은 단순히 성능이 뛰어난 인공지능을 만드는 데 총력을 기울이고 있지만, 사실 더 고민해야 할 것은 **인간과 인공지능의 하이브리드 모델**이다. 단순히 두 능력을 합치는 것을 넘어 **효율적인 조합으로 시너지를 내게 하는 방법**을 고민할 필요가 있다.

케이론 모델을 보다 섬세하게 디자인하려면 **조직, 업무 프로세스, 개인의 의사결정**이라는 세 측면에서 조망해볼 필요가 있다. 가장 집중해야 할 것은 **머리와 힘의 역할 분배와 상호작용**이다.

케이론 모델의 활용과 쟁점

적용 영역	머리와 힘의 역할 조정	쟁점
조직	머리: 기업의 전략, 현업 부서 힘: 인공지능 부서	프로젝트 선정, 목적 설정, 인공지능팀 구성, 두 조직 간의 커뮤니케이션
업무 프로세스	머리: 프로세스에서 인간이 할 일 힘: 프로세스에서 인공지능이 할 일	프로세스 분할 방식, 역할 전환, 상호 모니터링
개인의 의사결정	머리: 전문가의 판단 힘: 인공지능의 판단	언제 인간의 판단에 따르면 언제 인공지능의 판단에 따를지 결정 기준

퍼플 피플이란 **데이터 분석과 공학기술**을 비즈니스의 생태계에 맞춰 설명할 수 있는 사람을 의미한다.

퍼플 피플의 주요 업무는 **자신의 업무 지식을 바탕으로 빅데이터나 인공지능을 활용**해 기업에 **가장 큰 가치**를 안겨줄 수 있는 프로젝트를 선별하고... (중략) **프로젝트의 중재자 역할**을 한다.

* 케이론: 반인반마인 켄타우로스 중 매우 현명하고 뛰어난 현자

수학을 읽어드립니다

세바시_문과생이 수학과 코딩의 귀재가 된 이유

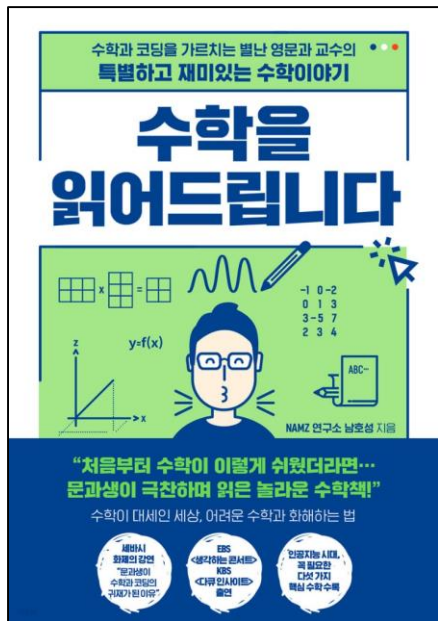


문과생이 수학과 코딩의 귀재가 된 이유 | 남호성 고려대학교 영어영문학과 교수 | 인공지능 음성인식 음성변환 NAMZ | 세바시 1135회

https://youtu.be/oGHZ_0wT_PE?si=WLOCdjpSQGnoRv8I

수학을 읽어드립니다

“함수, 미분, 행렬, 벡터, 확률”



프로로그_인공지능을 만드는 영문과 교수의 수학이야기

PART 1. 하마터면 수학을 포기할 뻔했다

PART 2. 수학 포기자에서 수학 예찬자가 되다

PART 3. 수학이 대세인 세상이 온다

PART 4. 우리가 미처 몰랐던 수학과 화해하기

PART 5. 미래에 꼭 필요한 수학

1. 인공지능 시대에 꼭 필요한 다섯 가지 수학



과정소개

Undergrad >

Graduate >

경영학과

비즈니스 인포메틱스

MBA >

AMP 및 기업교육 >

비즈니스 인포메틱스

[Home](#) > [과정소개](#) > [Graduate](#) > [비즈니스 인포메틱스](#)

소개

교육과정

교과목

일반대학원 비즈니스 인포메틱스 학과소개

기존 데이터 분석의 틀을 넘어선
종합적인 데이터 운용 및 분석 능력부터
혁신적 비즈니스 모델을 창출할 수 있는
능력을 갖춘 융합적이고 창조적인 전문인력 양성

인공지능(70%) & 비즈니스 적용(30%)

기업 경쟁력의 핵심이 단순한 품질이나 가격 중심에서 어떤 기업이 먼저 기회를 발굴할 것인가로 옮겨가고 있습니다. 어디에 언제 기회가 있을가를 발견하기 위해 기업은 내외부의 모든 데이터를 분석 및 활용하여 새로운 시장의 부상이나 창출을 누구보다 일찍 알아내야만 합니다.

이러한 경쟁력 패러다임의 전환에도 불구하고 국내에서는 데이터를 창조적으로 분석 및 활용할 수 있는 능력을 갖춘 인력이 매우 부족할 뿐만 아니라 이러한 인력을 양성하는 전문 기관 역시 찾아보기 힘듭니다. 이러한 문제를 해결하고자 한양대학교에서는 데이터 수집부터 분석까지 그리고 이를 바탕으로 혁신적인 경영 전략을 수립하고 행동에 옮길 수 있는 능력을 갖춘 전문인력을 양성하고자 국내 최초의 대학원 과정인 비즈니스 인포메틱스 (Business Informatics) 학과를 설립하였습니다.

비즈니스 인포메틱스 학과는 거대한 시스템을 마음대로 다루며 데이터를 모으고 이를 새로운 형태로 결합, 변환할 수 있는 능력, **기존 데이터 분석의 틀을 넘어선 종합적인 데이터 분석 능력, 분석 결과를 기업의 혁신적 비즈니스 모델 창출로 연결하여 행동으로 옮길 수 있는 능력**을 갖춘 종합적이며 융합적인 인력을 양성할 것을 목표로 하고 있습니다. 이를 위하여 비즈니스 인포메틱스 학과는 한양대학교 경영대학과 공과대학 컴퓨터공학부 교수들이 의기투합하여 설립되었습니다. 경영학 분야에서는 김중우 교수, 백승익 교수, 신민수 교수, 임규건 교수, 조남재 교수께서 참여하고 계십니다. 컴퓨터 공학 분야에서는 강수용 교수, 김상욱 교수, 박희진 교수, 백은옥 교수, 유민수 교수, 이춘화 교수, 임윤규 교수, 차재혁 교수, 최용석 교수, 차경진 교수님께서 참여 하고 계십니다.


<https://youtu.be/kvKA0mfZCWM>

카드사의 데이터 사이언스 업무

하나카드

데이터 본부 = 데이터 사이언스 + 데이터 전략 + AICC

1

데이터 기획

데이터 전략 기획

- 중장기 데이터 전략/기획
- 데이터 전략 비전 목표 및 추진과제 수립
- 데이터 기반 생태계 확장(데이터전문기관 준비) 및 사회기여 전략 수립 등

내/외부 환경 모니터링

- 시장/트렌드/수요 서비스 조사 및 이슈 분석
- 잠재고객/트렌드 분석
- 리포트 주제 발굴

데이터 사업 기획

- 데이터 기반 신사업 기획/사업 발굴
- 데이터 분석/결합 결과 유통 전략 수립
- 데이터 홍보

데이터 컨설팅

- 데이터 유통 및 데이터 전문기관 활성화를 위한 영업
- 데이터 결합 컨설팅

데이터 비즈니스 플레이어로의
도약을 위한
“데이터 전략 수립 및 신사업 기획”

2

데이터 사이언스

비즈니스 분석과제 도출

- 분석 트렌드 모니터링
- 데이터 정의
- 분석 신기술 검토 (HFG&SKG 가명결합 등)

데이터 분석

- 정형/비정형 데이터 기반 추천/개인화 모형
- 예측 모형 등 분석 모형 개발
- 알고리즘 및 인덱스 개발
- 분석 모형 서비스 적용
- 고객 세그먼트 및 타겟 분석

분석지원

- 분석 교육 기획
- 성과 공유 정례화
- 분석 솔루션 및 시간화 도구 활용 교육

데이터 플랫폼 운영

- 데이터 Pipeline관리
- 데이터 생성/수집/저장/가공 관리

카드 본업 강화를 위한 데이터 기반의
“비즈니스 인사이트 도출 및
초개인화 마케팅 지원”

3

AICC

챗봇 (Chatbot) > 큐미(QMe)

- 1:1 대화를 통해 휴먼상담과 유사한 경험 제공
- 오타, 초성, 영한 자동변환 등 학습을 통한 편의성 제공
- 개인화를 통한 손님 맞춤형 챗봇 서비스 제공

콜봇 (Callbot)

- 인바운드: 한도조회, 결제예정금액 등 32종
- 아웃바운드: 한도부족 승인거절, 연체안내
- 스마트 라우팅을 통한 손님 콜봇 자동연결

불완전 판매 모니터링 (STT)

- AI가 대화 내용을 분석하여 불완전 판매 여부를 판단
- AI가 1차 판단 > 담당자 최종 판단 > 민원조기대응
- 표준 스크립트와 STT를 비교하여 유사성 측정

상담 어시스턴트

- 초급상담 직원의 상담 지원
- 실시간 AI 기술을 활용한 추천 상담 제공
- 실시간 대화(STT), KMS 연동(검색), 관리자 코칭

AI기술로 단순/반복적인 업무를 처리하여
“손님에게 빠르고 편리한 상담을,
상담 직원에게는 업무효율성을 제공”



하나카드

데이터 전략부(데이터 유통/판매)



상품 소개

하나카드는 지역, 상권, 외국인, 지역화폐 거래 정보 등 8가지의 판매 데이터 및 통계 모델링, 정보 가공 등 4가지의 가공서비스를 제공하고 있습니다.

판매 데이터

1. 지역별 카드이상에 관한 정보

· 지역 거주민이 사용하는 카드 소비정보로, 소비인구 거주지역, 소비업종, 소비자 프로파일, 카드 이용정보를 제공

2. 지역별, 업종 유입고객 및 상권 특성 정보

· 특정 지역의 지역 거주민과 비거주민을 구분한 소비정보로, 상권지역, 소비업종, 소비인구 거주지역, 소비자 프로파일, 카드 이용정보를 제공

3. 지역별 외국인 소비정보

· 지역별로 외국인이 국내 가맹점에서 사용한 소비정보로, 상권지역, 소비업종, 외국인정보(국적), 거래 정보를 제공

4. GIS Block 코드별 업종 매출 현황

· 가맹점의 주소정보를 GIS 정재연진을 활용하여 Block 단위별 매출 정보로, 상권지역(Cell id), 상권지역, 소비업종, 카드 이용정보를 제공

5. 지역화폐 발급 및 이용정보

· 하나카드가 발행한 지역화폐 카드 발급 회원의 소비 정보로, 카드발급시점, 발급회원 거주지역 및 데모그래픽, 소비업종, 지역화폐 이용 정보를 제공
(하나카드 지역화폐 카드 발행 지역: 김포, 울산, 부산, 철곡, 익산, 세종, 광주, 대전)

6. 학생증 체크카드 이용현황

· 하나카드가 발행한 학생증 체크카드 발급회원의 소비정보로, 학교, 발급회원 거주지역 및 데모그래픽, 소비 업종, 학생증 체크카드 이용 정보를 제공

7. 개인사업자 평가 정보

· 개인사업자의 사업자 평가 정보로, 개인사업자별 평가항목별 사업자 정보, 시장 추정매출 정보, 평가 점수 및 등급 정보를 제공

8. 좌표기준 거리별 외국인 소비 정보

· 기존 지역 인근 가맹점의 외국인 소비 정보로, 국질별 외국인 관광 동선 분석, 국적별 선호 상권 및 업종 분석 정보를 제공

상품 소개

하나카드는 지역, 상권, 외국인, 지역화폐 거래 정보 등 8가지의 판매 데이터 및 통계 모델링, 정보 가공 등 4가지의 가공서비스를 제공하고 있습니다.

가공 서비스

통계자료 가공

· 개요: 구매 데이터에 대하여 유의미한 통계분석을 위한 데이터 생성 · 가공
· 가공기술: 원천데이터에 내재하는 오류값/특이값 보정 분석을 위한 데이터 카테고리화, 그래프, 차트, 분석 대시보드 등 시각화 등

통계모델링 분석

· 개요: 통계기법을 이용한 예측, 유형화, 연관성분석 등 분석결과 도출
· 가공기술: 추정/예측모델, 유형화(군집화), 연관성 분석, 교차판매분석, 요인분석, 추천 모델, 미스매치분석 등

주소정보 가공

· 개요: 주소지 주변의 다양한 정보를 활용하기 위해 주소 정제, 및 코드화
· 가공기술: 주소정제, 주소코드화, Apt/건물코드 부여

공간정보 가공

· 개요: 위치정보를 다양한 공간정보와 결합하여 공간 특성정보 생성
· 가공기술: 좌표부여, 좌표변환, 주소추출(역지오코딩), 지역 프로파일링(상권정보보강 등)

활용 방법

소비 트렌드 분석

· 마케팅/홍보 효율화 검증
· 소비자 프로파일 분석을 통한 타겟마케팅

상권 분석

· 가맹점 신규입점 가능 상권 추출
· 경쟁사 분석을 통한 마케팅 전략 수립

하나카드
판매데이터
가공서비스

지역별 업종별 소비규모 예측

· 지방자치단체의 정책 입안 또는 실행효과 분석
· 지역 이벤트, 축제, 박람회 유입자 분석

외국인 국적별 소비 트렌드 분석

· 외국인 대상 관광 물건을 개발
· 외국인 유치관련 관광 정책 마련

인공지능 학습을 위한 데이터

1. 합성데이터 (NIA 사업)
2. 탈회, FDS, 상품 가입 등

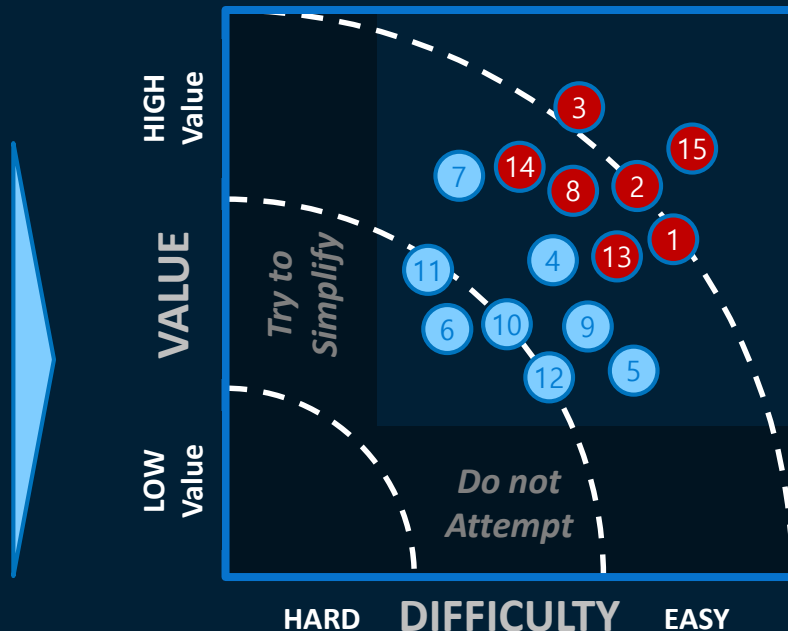
학습을 위한 데이터 판매

1. 데이터 처리
2. Classification 문제
3. Regression 문제
4. Forecasting
5. NLP (STT) 등

#카드사에서 하는 분석 주제들

국내외의 손님 Single-view 기반이 되는 “데이터 관리”부터 “상품추천 및 마케팅”, “손님행동 및 성향분석” 및 “리스크 분석”의 4가지 관점의 주제에 대한 Use Case

구분	Use Case 주제
A 데이터 관리	1. 데이터 품질 개선을 통한 분석 데이터 확보 2. 예측 모델 고도화를 위한 통합 마트 구축
B 상품추천 및 마케팅	3. 가맹점 추천 (초개인화; CLO 서비스) 4. 신용/체크 카드 사용 활성화 5. 카드 손님 이탈(무실적) 6. 손님 Journey 분석 7. 손님 Life Stage 기반 캠페인 8. 옴니채널 손님경험 최적화 및 마케팅 자동화 9. 손님 성향 평가
C 손님 분석	10. 카드 소유자 행동분석 11. 서비스/제품 순서 (가치 창출 경로)
D 리스크 분석	12. 브랜드와 서비스에 대한 평판 모니터링 13. 손님불만 사전 예측 14. 카드 부정 사용 적발(FDS) 15. 신용평가(CSS) 모델



[참고] Use Case 별 분석 기법 및 관련 과제 (1/2)

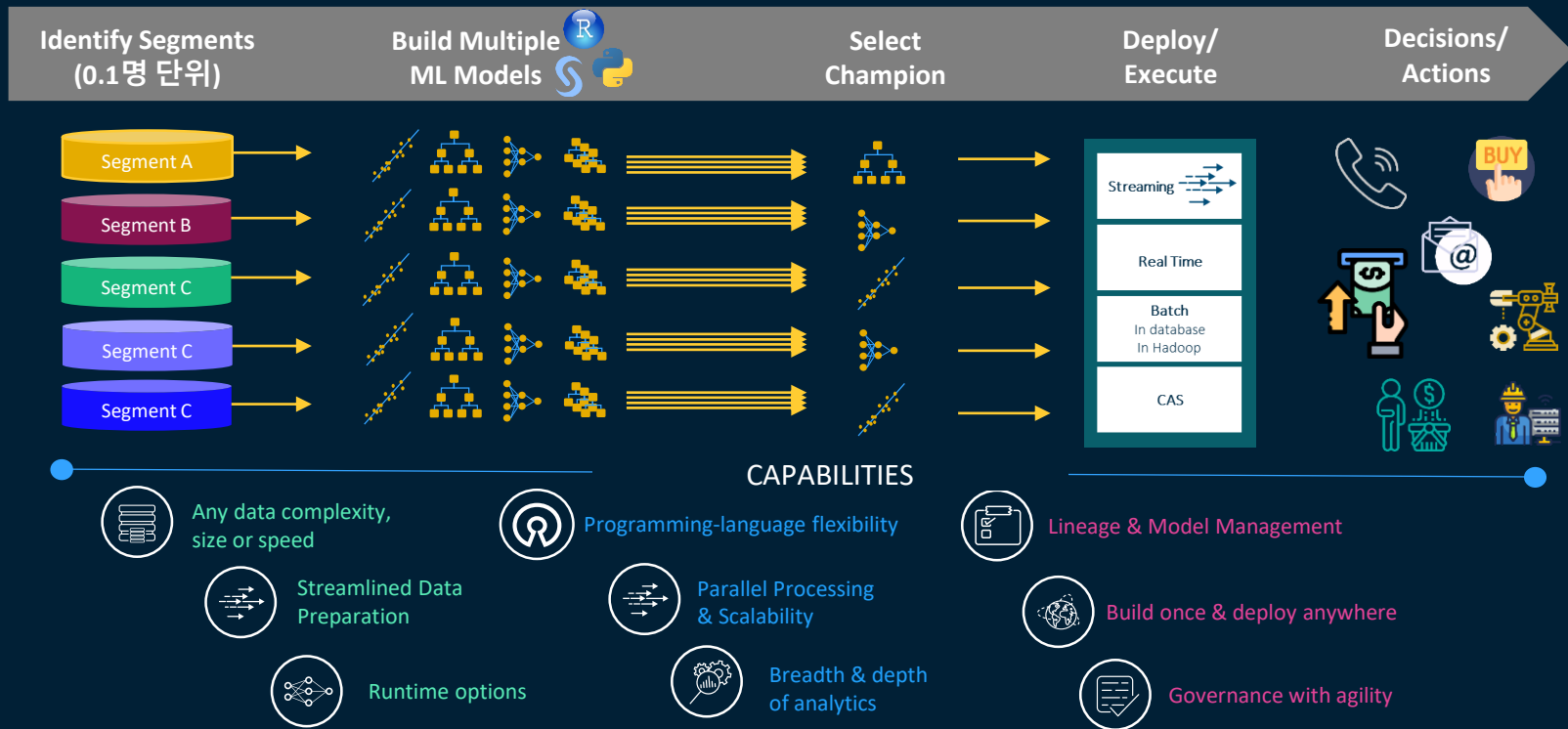
Use Case 주제		주요 내용	분석 기법	관련 과제
A 데이터 관리	1	데이터 품질 개선을 통한 분석 데이터 확보	• 데이터 프로파일링 • 데이터 정제, 확장, 통합	• 이기종 데이터 Access • 데이터의 정제 및 프로파일링 • 데이터 관리 및 손님에 대한 단일 뷰 생성 • 손님행동 데이터 분석 및 시각화
	2	예측 모델 고도화를 위한 통합 마트 구축	• Data Lake 관점의 통합마트 구축을 통하여 손님의 Insight 도출 및 다양한 업무에 활용	• 손님, 계약, 설계사 기준 단일 뷰 생성 • 손님 및 설계사 세분화(군집화) • Data Mining & M/L을 적용한 예측 모델 (이탈, X-Sell)
B 상품 추천 및 마케팅	3	가맹점 추천 (초개인화; CLO 서비스)	• 신용카드 이용에 따라 개인에게 적합한 가맹점 추천	• 추천 알고리즘(CF-협업필터링, Factorization machine 등) • AutoML을 이용한 상황별 예측분석 • 카드 사용 패턴에 따라 위치기반의 가맹점간 연계 분석 • 유형별 이용목적 파악 및 손님 선호 기반 신제품 개발 • 개별화된 이용업종별 추천 지수 개발
	4	신용/체크 카드 사용 활성화	• 무/저실적 손님 중 잠재력이 높은 손님 대상 사용 활성화	• 추천 알고리즘 • 손님 세분화 • 머신러닝 & 딥러닝 • 손님 세그먼트별 카드 이용 분석 • 손님 Seg별 카드 사용을 유도하는 분류 및 예측 모델 • 가맹점-손님 매칭
	5	카드 손님 이탈(무실적)	• 손님 세그먼트별 이탈 이벤트를 식별하여 카드 무실적 및 이탈을 최소화	• Path & 네트워크 분석 • 머신러닝 & 딥러닝 • 이탈(무실적 포함)로 이어지는 모든 여정을 식별 • 식별된 여정을 바탕으로 한 동인 식별 • 머신러닝 적용을 통한 잠재 이탈자 식별
	6	손님 Journey 분석	• 손님 Journey와 카드 사용 순서 분석을 통한 적절한 시기에 캠페인 실행	• 추천 알고리즘 • Path & 네트워크 분석 • Path 분석과 CF 알고리즘을 통한 손님 Journey와 카드사용 순서를 시각화 • 특정 트리거 후 손님 행동 패턴 분석 (예: 해외여행 전후 일반적인 행동은?) 및 캠페인 반응 결과 분석
	7	손님 Life Stage 기반 캠페인	• 졸업, 주택 구입, 결혼 등 손님의 중요한 생활 이벤트를 식별하여 캠페인 실행	• Path & 네트워크 분석 • 선호도 분석 • 머신러닝 & 딥러닝 • 다양한 손님의 Insight 확보 • 손님 세그먼트별 선호도 분석 • 손님 세그먼트별 캠페인 반응도 분석
	8	옴니채널 손님경험 최적화 및 마케팅 자동화	• 옴니채널 서비스 프로세스 최적화를 통한 손님에게 일관된 서비스 제공	• 손님 세분화/군집화/최적화 • Path & 네트워크 분석 • 머신러닝 & 딥러닝 • 채널 및 손님 가치 세분화/군집화 • 실시간 스코어링 및 추천을 위한 Auto ML 모델 개발 • 캠페인 관리 및 평가 • App에서 토탈 서비스 제공(셀프 서비스)

[참고] Use Case 별 분석 기법 및 관련 과제 (2/2)

Use Case 주제			주요 내용	분석 기법	관련 과제
C	손님 분석	9 손님 성향 평가	카드 사용실적, 캠페인 반응에 따른 손님 성향 및 제품 선호도 평가	- 성향 스코어링 - 제품 선호도 스코어링 - 선형 및 로지스틱 회귀분석	- 손님 동인 및 행동에 대한 Insight 도출 - 손님 세분화 및 핵심 속성 식별 - 성향 및 제품 선호도 측정을 위한 스코어링 모델 개발
		10 카드 소유자 행동분석	평생가치 기준 손님을 분류하고, 행동 경로를 식별하여 고가치 손님 세그먼트로 전이 유도	- Path & 네트워크 분석 - 머신러닝 & 딥러닝	- 카드발급에서 이탈까지 경로 탐색 - 손님 고가치 세그먼트로 전이시킬 수 있는 주요 이벤트/시나리오 식별 - 각 세그먼트에 대한 시각화된 행동 경로 예측
		11 서비스/제품 순서 (가치 창출 경로)	손님의 가치 상승 경로를 식별하고 이에 따른 손님을 사전관리 방법 모색	- 추천 알고리즘 - Path & 네트워크 분석	- 고유한 손님 여정을 식별하기 위한 데이터 세션화 및 이벤트 식별 - 가치에 이르는 경로 식별(예: 수익성을 높이기 위해 1년차 손님에게 어떤 일이 발생해야 하는가?)
D	리스크 분석	12 브랜드와 서비스에 대한 평판 모니터링	내부 및 소셜 미디어의 비정형 데이터 분석을 통한 브랜드 및 서비스 평판 평가	- 텍스트 분석 및 텍스트 마이닝 - 머신러닝 & 딥러닝	- 웹 채팅 등 텍스트 분석(형태소 분석, 사용자 정의사전, 텍스트 분류 등) - 불만유형 분류를 위한 예측모델 개발 - 분석결과에 대한 데이터 탐색 및 시각화
		13 손님불만 사전 예측	I/B Call 접수 시 사전 손님불만 예측을 통한 Call 배분 및 관련 정보 제공	- 텍스트 분석 및 텍스트 마이닝 - 머신러닝 & 딥러닝	- VOC 텍스트(STT, 상담 메모) 분석(형태소 분석, 사용자 정의사전, 텍스트 분류 등) - 머신러닝 & 딥러닝을 적용한 불만유형별 사전 예측모델 개발
		14 카드 부정 사용 적발(FDS)	국내/국외 온라인 부정사용 적발을 위한 모델 개발 및 실시간 스코어링 적용 및 운영	- 데이터 마이닝 & 머신러닝 - 실시간 스코어링	- 머신러닝 & 딥러닝을 적용한 국내/국외 온라인 부정사용 예측모델 개발 - 개발된 예측모델에 대한 실시간 스코어링 적용
		15 신용평가(CSS) 모델	개인 및 기업 대상 위험관리를 위한 신용 평가 모델 개발	- 신용평가모형(CSS) - Path & 네트워크 분석 - 머신러닝 & 딥러닝	- 위험 평가를 위한 신용평가 모델링 - 경로 및 네트워크 분석을 통한 이상패턴 감지 - 홀드아웃 테스트를 통한 정확도 검증

초개인화 성공을 위한 Platform

“What you want to achieve? Automate Modeling”



가명결합

금융, 소비, 생활트렌드통합분석

- SK그룹과 하나금융그룹 교차고객의 금융, 소비, 생활을 아우르는 통합 트렌드 발굴
- 나이스지니데이터와 하나카드 결제데이터 기반 외국인관광객의 소비특성 등을 분석

SK그룹 & 하나금융그룹 : 금융, 비금융 특성과 소비 트렌드 통합 분석

하나은행 금융 활동

- 선호 금융상품
 - 예적금, 간접투자 등
- 대출 금액/비중
 - 주택담, 신용대출 등
- 상품가입 경로
- 예적금 만기 후 전환상품
- 2,3 금융권 대출/연체 정보

하나카드 소비 활동

- 주요 소비 업종
- 선호 생활/소비 습관
 - 선호 교통수단, 음식점, 커피 브랜드 등
- 쇼핑 채널 및 행태
 - 대형마트/쇼핑몰, 온라인, 동네 소매점, 숍세권 선호

하나증권 투자 정보

- 투자 규모
 - 약정금액, 금융자산 비중,
- 투자성향
 - 대형/소형주 투자비중, 업종별 투자 비중, 수익률
- 투자 채널
 - HTS, MTS별 이용횟수

SKT 모바일 이용 특성

- 주요 활용 앱 유형
 - 여행, 골프, 쇼핑, 게임, 반려동물 등 관심 분야
- 은행, 증권 등 금융 활동
- 이동 정보
 - 평균 이동거리
 - 방문지 (관광지, 학교 등)

11st 온라인 쇼핑

- 온라인 쇼핑 선호 카테고리
 - 반려동물, 스포츠, 도서, 음반, 디지털 기기 등

SKB 영상/미디어 소비

- IPTV, 미디어 소비 특성
 - 주중/주말 시청 비율
- 드라마/스포츠/예능/교양/홈쇼핑 등 분야별 시청 시간대

나이스지니데이터 & 하나카드 : 외국인 관광객 선호 메뉴 분석

외국인 관광객 몰리는 'K디저트' 성지

- 계동(점유율 94%)
주요 메뉴 빵도르·크루아상
- 재동(82.7%)
스콘·초코 케이크
- 이화동(45.9%)
티라미수·밀크티

서빙고동
97.7%
와플·
케이크



-성북동1가(42.6%)
도나트·마들렌

-석관동(31.9%)
샌드위치·페이스트리

관련 기사 일부

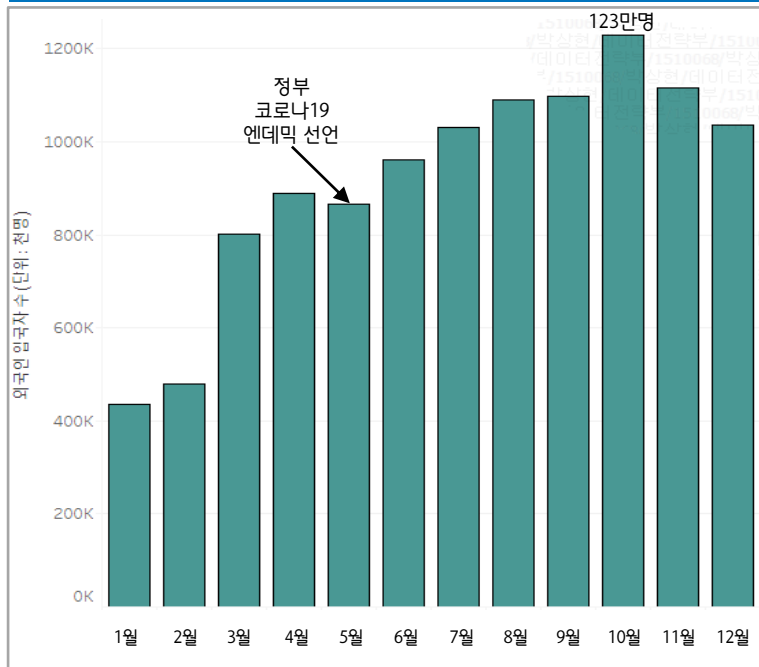
서울관광재단이 하나카드와 빅데이터 업체 나이스지니데이터를 통해 작년 7월부터 올 6월까지 서울을 방문한 외국인 관광객들의 신용카드 결제 데이터를 분석한 결과, 가장 많이 먹은 음식은 국적 불문하고 갈비와 삼겹살 등 '고기류'였다. 하지만 일본과 러시아, 대만 등 아시아권 관광객들은 고기류 다음으로 와플과 파이 등 '디저트'를 선택한 것으로 조사됐다.

나라별 디저트 구매 비율은 러시아가 전체 소비한 음식의 83.6%로 가장 높았다. 이어 자메이카 35.3%, 일본 11.6% 등이었다. 전체 외국인 관광객 중 가장 많은(20.7%) 일본인 관광객은 고기류 다음으로 디저트를 선호했고, 주로 빵도르와 크루아상을 구매했다

관광 빅데이터 포럼

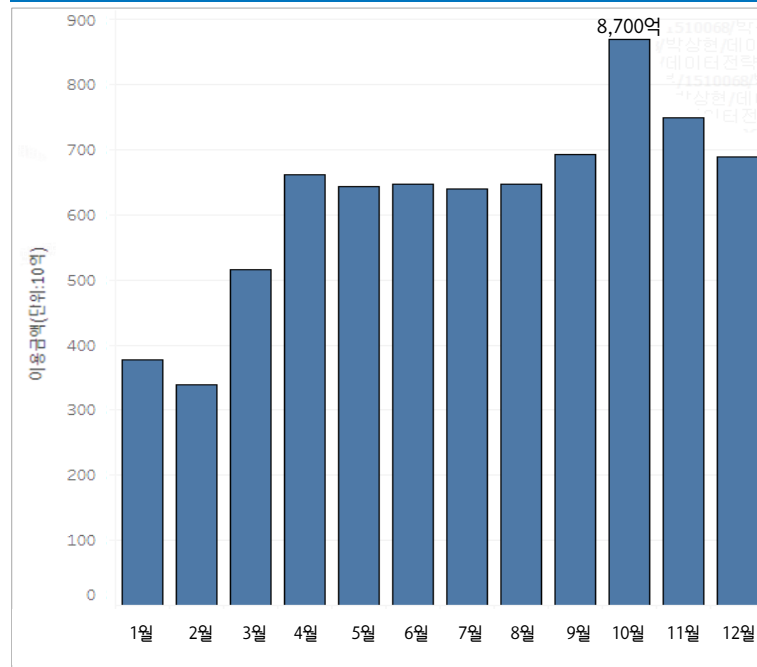
23년도 월별 외국인 입국자 수와 카드 이용 금액 추이

월별 외국인 입국자 수



자료 : 한국관광 데이터랩

월별 외국인 카드 이용 금액



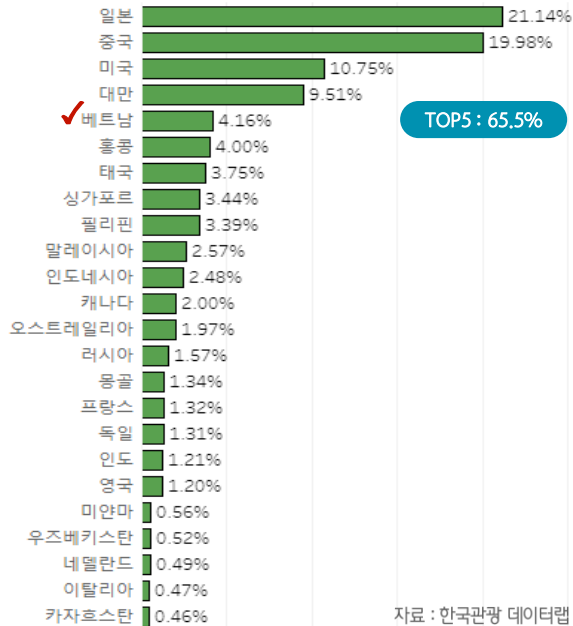
자료 : 하나카드

주. (1) 당사 기준 국내 외국인의 카드 이용 금액 월별 집계, 장기 체류 외국인 제외 (장기 체류 외국인은 1년중 3개월 이상 카드 이용 내역이 존재하는 경우)
당사의 외국인 매입 데이터를 기반으로 전체 외국인 이용 금액을 산정한 것으로 실제 외국인의 카드 이용 금액과는 다름

관광 빅데이터 포럼

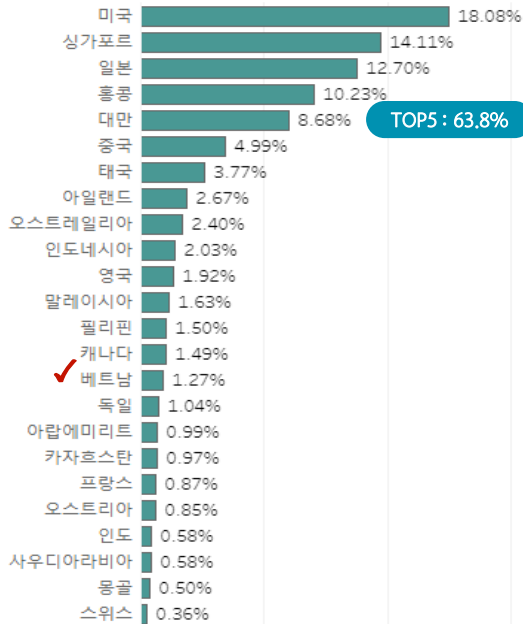
23년도 주요 국가별, 시도별 외국인 카드 이용 금액과 비율

국가별 외국인 입국자 비율

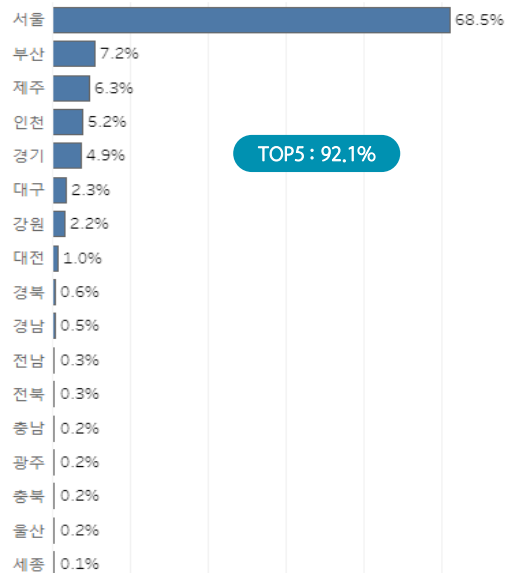


자료 : 한국관광 데이터랩

국가별 카드 이용 금액 비율



17개 시도별 카드 이용 금액 비율

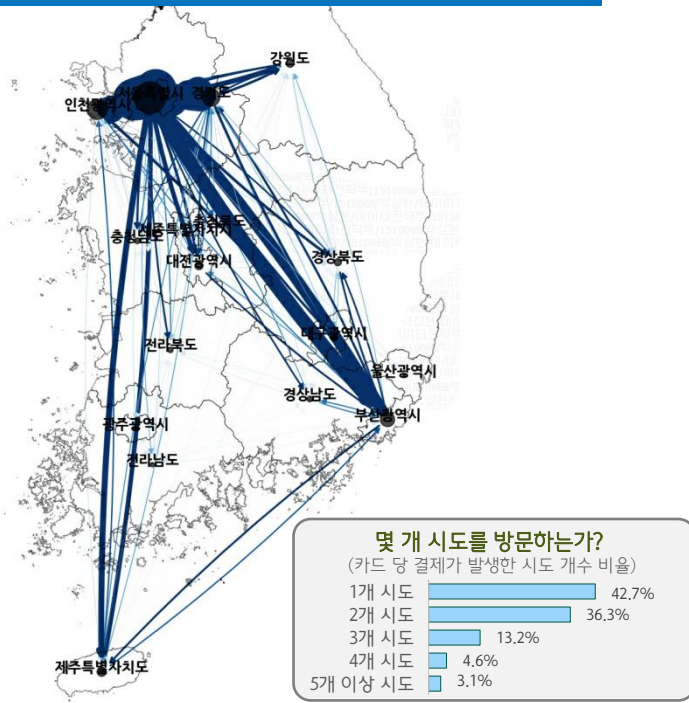


주. (1) 이용 금액 관련 국적은 비자, 마스터 기준 카드 식별 번호로 구분, 해당 국가 금융기관을 이용하는 타 국적 외국인도 포함될 수 있음
 - 예. 카드 식별 번호는 싱가포르 국적이지만 실제로는 중국 국적인 경우 등

관광 빅데이터 포럼

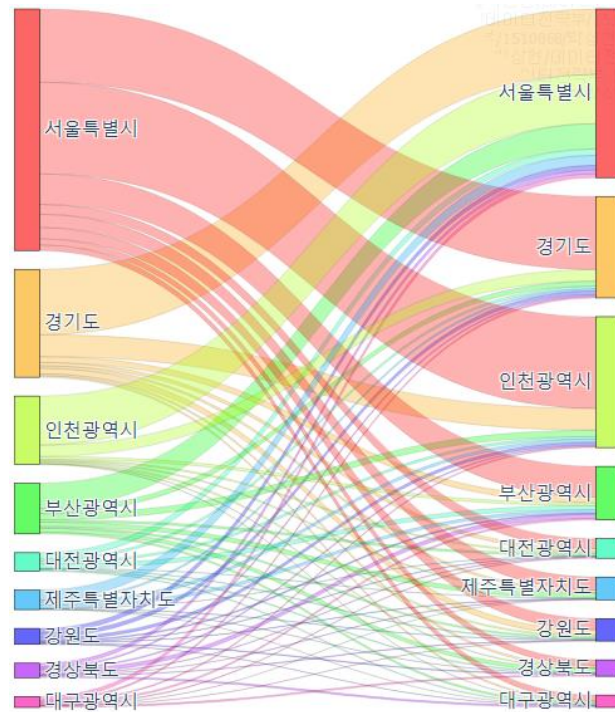
외국인 카드 소비 동선

카드 소비 지역과 순서에 따른 동선 맵



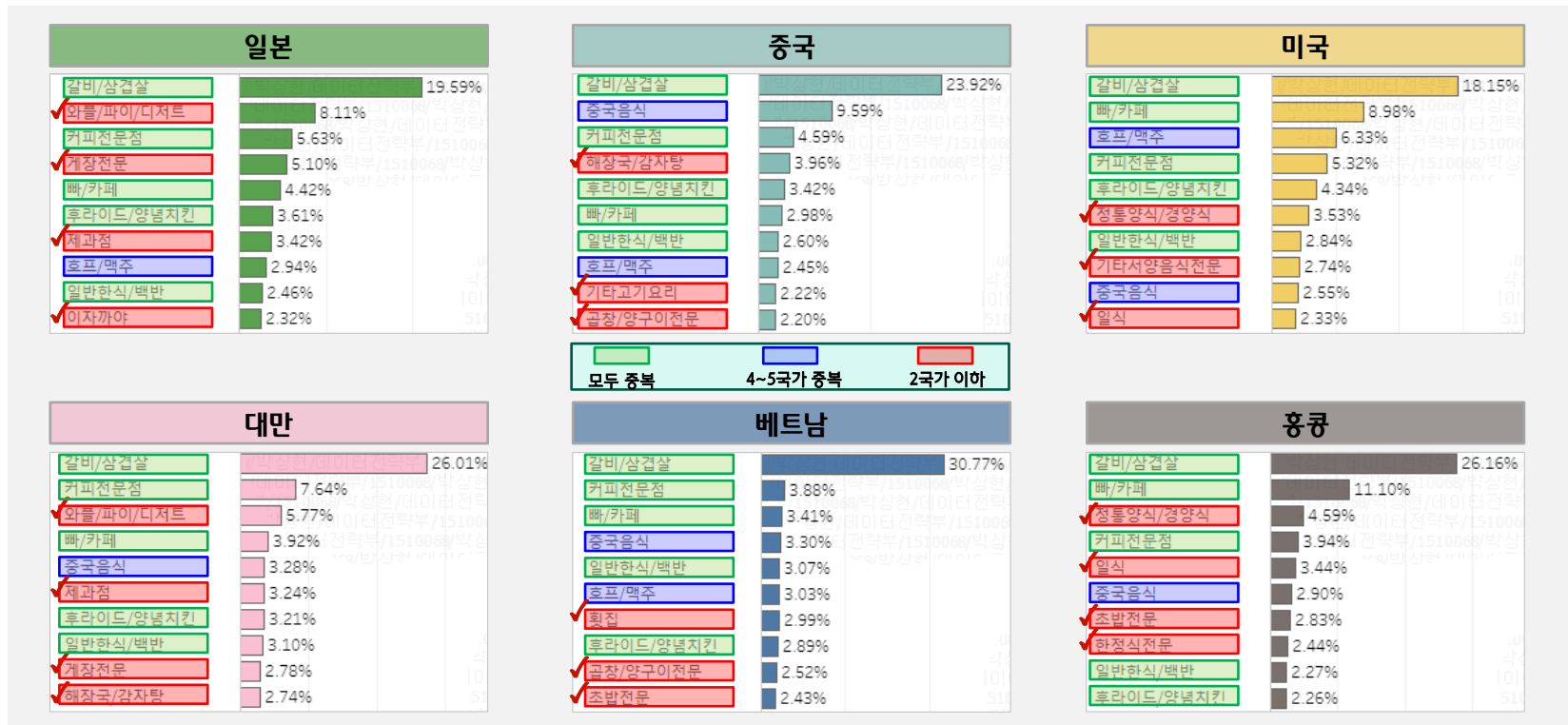
주. 23년도 12개월 중 외국인 카드 이용 금액이 가장 많은 10월을 기준으로 결제 동선을 시각화

주요 지역간 카드 소비 동선



관광 빅데이터 포럼

23년도 입국자수 상위 6개 국가별 선호 음식 유형과 특징

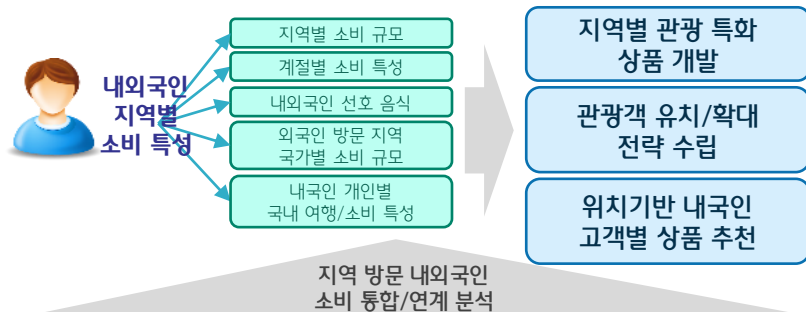


주. 나이스지니데이터 POS정보와 하나카드 외국인 결제 내역의 가명 결합 데이터 기반 집계 (전국 17개 시도, 22년3분기~23년2분기)

관광 빅데이터 포럼

하나카드 데이터 비즈니스 소개

내외국인 지역별 소비 세분화로 활용 분야 확대



지역별 내외국인 소비 현황



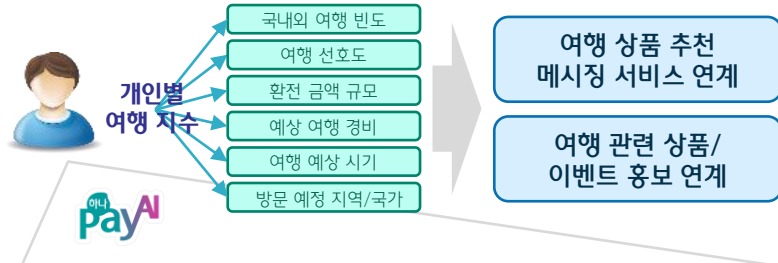
연령/성별/관심 분야 등 (내국인)
국적(외국인)
방문지역
소비업종
소비규모
소비동선
체류기간

데이터 연계
가명/익명
정보 결합 등

관련 기업/플랫폼 연계

- 쇼핑 품목 유형
- 소비 음식 유형
- 소비 가맹점 유형
- POS 결제 정보
- ...

내국인 여행 특성 기반 개인화 서비스 연계



하나머니 앱



- 해외머니/여행특화
- 트래블로그 서비스

여행 로그



- 여행 로그
- 여행 분석 리포트 (방문 장소 등)

여행 큐레이션

손님 여정에 맞는 맞춤 정보제공

여행 전

여행 중

여행 후



개인화된 여행 특화 서비스

생성형 AI

금융 및 카드 활용 분야

대 직원 서비스

자동 코드 생성

SQL, Python, SAS 등 코드 생성 및 코드 컨버팅으로 데이터 처리 업무 효율화

마케팅 메시지

캠페인 실행 시 대 손님 LMS/Push 마케팅 및 안내 메시지 자동 생성/디자인 자동 생성

기획안/기안문서 등 업무 자동화

다양한 부서의 업무 효율화를 위한 문서 자동 생성
이메일 자동 생성
회의록 자동 요약 및 주요 결정사항 도출

하나 금융 그룹 vLLM (금융특화)



대 상담사(손님) 서비스

챗봇 / 콜봇

LLM+RAG+시나리오 개발을 통해 정확한 손님 의도 파악 및 자동 답변 제공

상담 도우미

게시판 답변 생성, 채팅 상담, 콜 상담 등에 상담 Assistant 활용, 손님 답변 자동 생성/수정

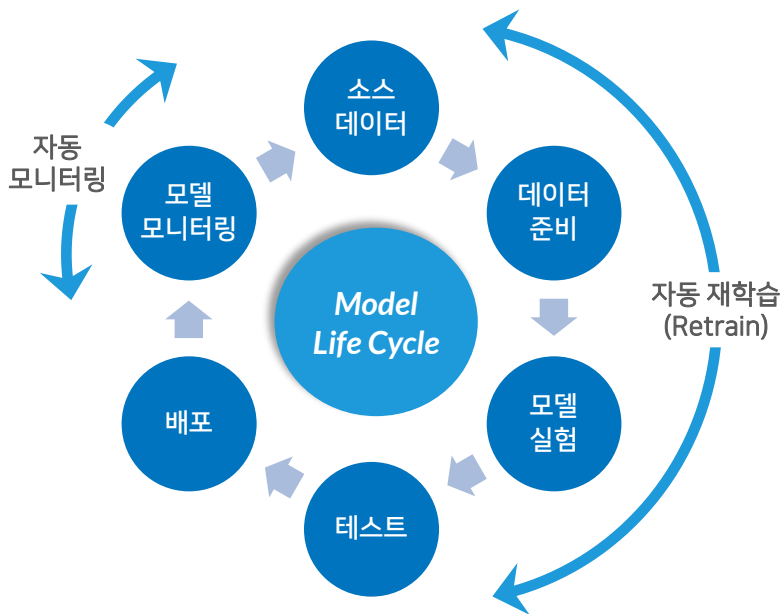
KMS 요약

손님 상담을 위한 다양한 내부 문서 요약 및 제공

AI Ops, MLOps, LLM Ops

금융 및 카드 활용 분야

모델 라이프사이클 전 과정 자동화



MLOps의 핵심 성능

1 중앙 집중화

- 모델의 자산화, 검색, 저장
- 모델과 적합한 메타데이터와 함께 중앙 장소에 저장 및 활용

2 거버넌스

- Responsible AI, 모델 라이프사이클 전과정에 대한 위험 평가, 법규 위반 예방, 편향 감지 등

3. 검증

- 전체 모델 인스턴스가 정확하고 유용한 예측을 도출하는지 확인

4 배포 및 모니터링

- 개발된 모델을 손쉽게 운영 환경에 적용
- 운영에 적용된 모델의 성능(Performance) 모니터링

5 자동화

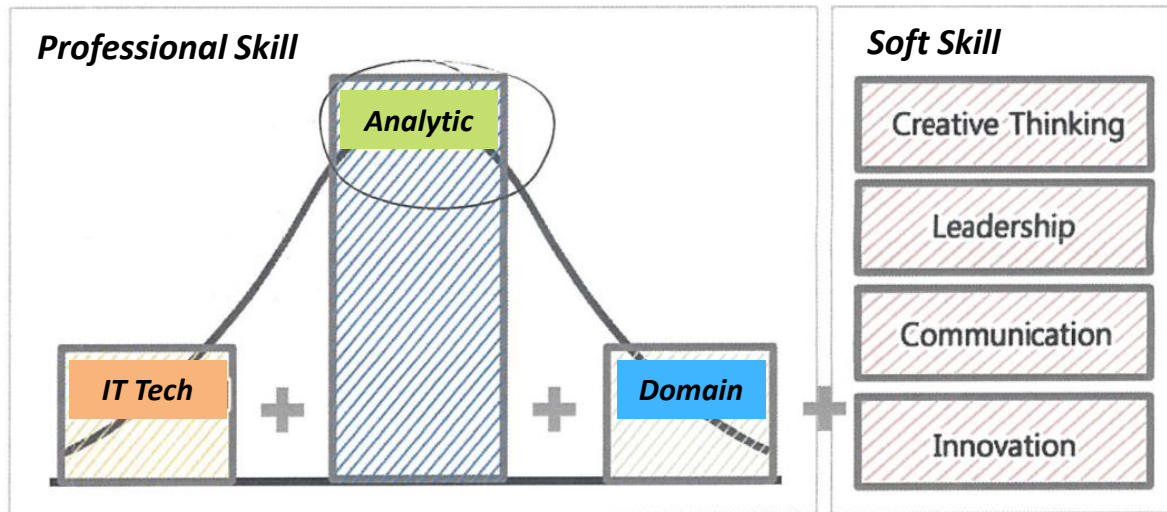
- 인적 감독과 시간 절약형 자동화의 균형

데이터 사이언티스트를 위한 준비

Data Scientist

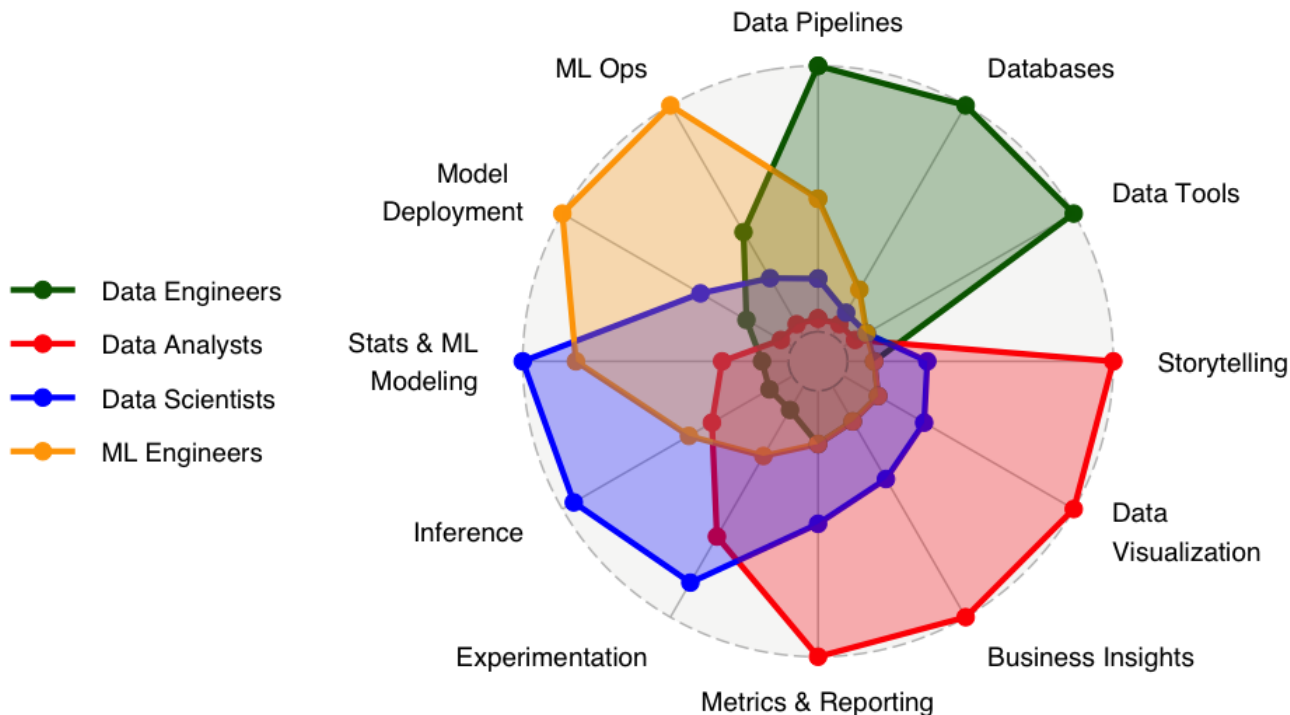
Core Competency

- *Data Scientist: The Sexiest Job of the 21st Century (Harvard Business Review)*



- 분석 경진대회 참여 전 vs 후
 - 데이터를 바라보는 시각, 데이터 처리 능력, 분석 능력, 창의력, 의사소통 능력
 - 분석 프로세스 (분석 라이프사이클)

Data Scientist vs Data Engineer vs ML Engineer vs Data Analyst



빅데이터 동아리

Invited Lecture

외부 강사 초청 강연 / 학회 선배와의 대화

Data-Mining Education

SAS / R / Python

빅데이터를 활용한 데이터 분석의 기초를 쌓기 위해 신입기수를 대상으로 Data-Mining 교육을 진행하고 있습니다.
교육은 SAS Enterprise Guide 및 Miner 프로그램을 사용하여 SEMMA(Sample/Explore/Modify/Model/Assessment)의 순서를 따라 진행됩니다.
또한 기본적인 분석 방법론을 제독한 후에는 명령어 기반의 R/Python 등을 추가로 활용하여 다양한 분석을 수행합니다.

 Sas Base/EG/EM

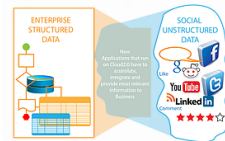
 R Program

 python Python Program

Project Session

Structured/Unstructured data analysis

Data-Mining 방법론에 대한 기본적인 교육 후, 실제적인 데이터 분석을 위한 Project Session을 진행합니다.
주로 Kaggle의 공개된 데이터를 활용하여 각 주제에 알맞은 분석을 팀 단위로 준비 및 발표하게 되며, 해당 결과물에 대한 구성원들 간의 건설적인 토론 및 지도교수의 피드백 하에 분석을 발전시켜 나갑니다. 신입기수 위주의 주니어팀의 경우 주로 정형화된 틀의 Structured data에 대한 분석을, 2학기 이상 활동한 시니어팀의 경우 그와 더불어 소셜 데이터 등을 활용하여 Unstructured data에 대한 분석에 임하게 됩니다.



Introduce B.a.f

B . a . f (비이더플)

B

a

f

Became

a

flower

Bigdata

analysis

for

빅데이터는 데이터 자체만으로는 아무 의미가 없다.

한낱 몸짓에 지나지 않았던 존재가 이름이 불리어진 후 꽃이 되었던 것처럼, 데이터 속의

의미를 찾는 과정을 통해 데이터에게 진정한 가치를 부여할수 있는 학회가 되고자한다.

꽃

-김준수-

내가 그의 이름을 불러주기 전에는
그는 다만 하나의 몸짓에 지나지 않았다.

내가 그의 이름을 불러주었을 때
그는 나에게로 와서
꽃이 되었다.

내가 그의 이름을 불러준 것처럼
나의 이 빛깔과 향기에 알맞은
누가 나의 이름을 불러다오.
그에게로 가서 나도
그의 꽃이 되고 싶다.

우리들은 모두
무엇이 되고 싶다.
너는 나에게 나는 너에게
맞혀지지 않는 하나의 눈짓이 되고 싶다.

<https://business.facebook.com/dgubaf/>

<https://www.dgubaf.com/>

데이터 사이언스 관련 자격증

설계, 개발, 분석 관련 자격증

데이터아키텍처 (Data Architecture) 자격시험 DAP Data Architecture Professional 국·가·공·인 국가공인 DA 전문가 (공인자격 제2019-03호) DAP Data Architecture Professional DAoP Data Architecture Semi Professional DA 준전문가 (공인자격 제2008-0307호) DAoP Data Architecture Semi-Professional	SQL (Structured Query Language) 자격시험 SQLP SQL Professional 국·가·공·인 국가공인 SQL 전문가 (공인자격 제2018-02호) SQLP SQL Professional SQL Developer SQL Developer 국·가·공·인 국가공인 SQL 개발자 (공인자격 제2018-02호) SQLD SQL Developer	데이터 분석 (Advanced Data Analytics) 자격시험 ADP Advanced Data Analytics Professional 국·가·공·인 국가공인 데이터 분석 전문가 (공인자격 제2018-01호) ADP Advanced Data Analytics Professional ADsP Advanced Data Analytics Semi-Professional 국·가·공·인 국가공인 데이터 분석 준전문가 (공인자격 제2018-01호) ADsP Advanced Data Analytics Semi-Professional
---	--	---



데이터 사이언스 관련 자격증

설계, 개발, 분석 관련 자격증

카테고리	등급	필기	실기	난이도
분석	국가기술 빅데이터 분석기사	- 빅데이터 분석 기획 - 빅데이터 탐색 - 빅데이터 모델링 - 빅데이터 결과해설	- 빅데이터 분석 실무 - R, Python에서 선택	★★★★
	데이터 분석 준전문가 (ADsP)	- 데이터 이해 - 데이터 분석기획 - 데이터 분석	X	★★★
	국가공인 데이터 분석 전문가 (ADP)	- 데이터 이해 - 데이터 처리 기술 이해 - 데이터분석 기획 - 데이터분석 - 데이터 시각화	- 데이터 분석 실무 - R, Python에서 선택	★★★★★

카테고리	등급	필기	실기	난이도
개발	국가공인 SQL 개발자	- 데이터 모델링의 이해 - SQL 기본 및 활용	X	★★★
	국가공인 SQL 전문가	- 데이터 모델링의 이해 - SQL 기본 및 활용 - SQL 고급 활용 및 튜닝	- 데이터 모델링의 이해 - SQL 기본 및 활용 - SQL 고급 활용 및 튜닝 (SQL)	★★★★

A person is sitting at a desk in a dimly lit office at night. They are looking at a large monitor displaying a bar chart. There is another monitor and a laptop on the desk. The background shows blurred city lights through a window.

Thank you!!

Byoungjeong.choi@gmail.com