



# NAVER SHOPPING

BIGDATA 빅데이터분석

이상언

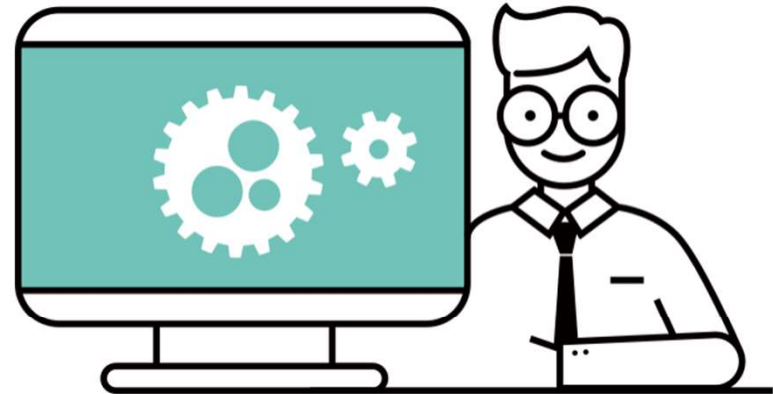
# 개요

1. 빅데이터 서비스 환경 분석서

2. 빅데이터 환경 분석서

3. 빅데이터 기술 환경 분석서

4. 인력 직무 기술서



# 목차(서비스 환경분석)

서비스 모델 개요

3C 기법 활용 환경분석

SWOT 분석 결과

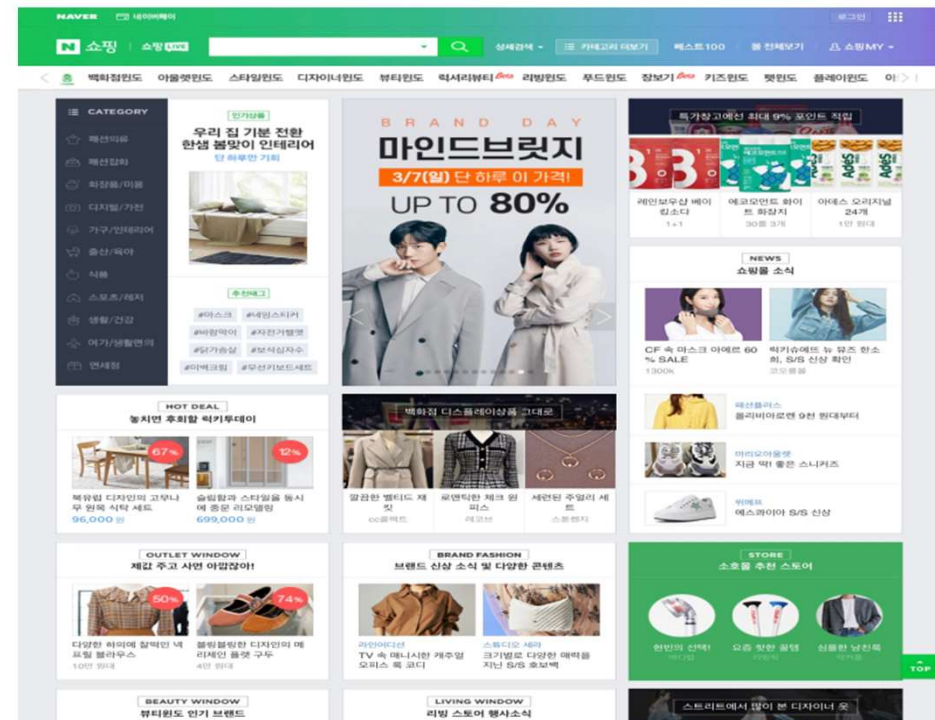


네이버쇼핑

# 서비스 소개

## < 네이버 쇼핑 >

NAVER 쇼핑 데이터를 분석한 후 판매율  
과 시장 점유율 상승을 이끌어낼 만한 인  
사이트와 홈페이지 화면 구성 제시



## 3C 기법을 활용한 환경 분석

구성 요소	세부 항목	서비스 모델 분석을 위한 세부 내용
NAVER 쇼핑	시장 규모	계속 되는 치킨 게임에도 매년 시장 전체 규모는 성장하고 있으며 신규 수요 창출 가능성도 큼
	성장 가능성	국내 이커머스 점유율 42%로 1위 네이버쇼핑은 가격정책과 직관적인 UI를 통해 향후 발전 가능
	고객 니즈	매출 증대를 위한 홈페이지 화면 구성 제시
본사	조직 내 외부 환경	조직 내 빅데이터 분석 인력은 충원 예정이며 프로젝트에 할당할 본사 여력은 충분
	기술 역량	데이터 서버 관리 기술이 부족하지만 이는 네이버 본사 서버를 이용하기로 함
	서비스 능력	데이터 분석 후 그에 맞는 적절한 인사이트 도출 가능
경쟁사	현존 경쟁사	빅데이터 분석에 관하여 본사 보다 큰 역량을 갖고 있는 회사가 많음 하지만 그에 대한 인사이트를 통해 홈페이지 UI 구성 방식 제안에 관해서는 경쟁력이 있음

# SWOT 분석을 활용한 환경 분석

1) SWOT 분석		
	우호적 환경	비 우호적 환경
외부 환경	<b>기회 (Opportunity)</b> <ul style="list-style-type: none"> <li>- 오랜 시간 데이터 분석을 통해 업무 교류를 한 고객사</li> <li>- 데이터 분석을 통해 추천할 수 있는 UI 상세를 요구</li> <li>- NAVER Cor. 측에서 쇼핑 사업의 지속적 확장을 원함</li> </ul>	<b>위협 (Threat)</b> <ul style="list-style-type: none"> <li>- 고객사의 잔여 계약 기간과 잦은 니즈 변화</li> <li>- 데이터 3법에서 파생된 법안 변동에 따른 변화</li> <li>- 경쟁사 측에 높은 성장률</li> <li>- 사회적 거리두기로 인한 정량적인 미팅 부족</li> </ul>
내부 환경	<b>강점 (Strength)</b> <ul style="list-style-type: none"> <li>- 고객사에서 요청한 UI 제안서를 직접 제작 가능</li> <li>- 고객사의 데이터웨어하우스에 접근 권한 소지</li> <li>- 경쟁사의 쇼핑몰에 대한 데이터 소유</li> </ul>	<b>약점 (Weakness)</b> <ul style="list-style-type: none"> <li>- 비주얼라이저 대비 애널리스트의 양적 부족</li> <li>- 본사에서 활용할 데이터 한도가 부족</li> <li>- 분석결과 도출 후 작성 될 API 개발 경험 부족</li> <li>- 사회적 거리두기로 인한 인력 누수</li> </ul>

# SWOT 분석을 활용한 환경 분석

## 2) SWOT 분석을 활용한 전략 도출

	강점(Strength)	약점(Weakness)
<b>기회</b> (Opportunity)	<b>SO 전략</b> <ul style="list-style-type: none"> <li>- 데이터 분석 이후 전문 비주얼라이저를 통해 높은 완성도의 UI 제안서 작성</li> <li>- 데이터 전 처리 과정 없이 정제된 데이터를 고객사를 통해 직접 수신 가능</li> <li>- 이번 프로젝트 이후 향후 고객사와 지속적 교류 가능</li> </ul>	<b>WO 전략</b> <ul style="list-style-type: none"> <li>- 고객사의 서버를 임시적으로 활용하는 방안 가능</li> <li>- 고객사의 협조 아래 API를 상세 UI제안으로 변경 가능</li> <li>- 향후 지속적인 고객사와의 교류를 바탕으로 데이터 전문 인력 충원에 부담을 경감 시킬 수 있음</li> </ul>
<b>위협</b> (Threat)	<b>ST전략</b> <ul style="list-style-type: none"> <li>- 데이터 분석 뿐 아니라 직접 UI 제공으로 고객사의 니즈 충족 가능</li> <li>- CRISP_DM 프로세스 활용으로 보다 빠른 피드백 가능</li> <li>- 경쟁사의 데이터 분석 결과와 비교하여 고객사에게 본사의 역량을 제시 할 수 있음</li> </ul>	<b>WT 전략</b> <ul style="list-style-type: none"> <li>- 부족한 데이터전문 인력의 충원을 위해 현재 인력기술서를 바탕으로 한 기획 진행 중</li> <li>- 원활한 비대면 업무를 위해 고객사와 협의중</li> <li>- 변화된 데이터 법에 대한 전문 법률 자문이 필수적임</li> </ul>

## 목차(데이터 환경분석)

데이터 후보 식별

데이터 선정 평가

최종 데이터 선정

정보 보안, 개인정보 보호 문제

제약 사항 및 대안





네이버쇼핑

## 데이터 후보 식별

후보 데이터명	데이터 유형	보관 형태	생산 주체	데이터 크기
분야	비정형, 명사형	NoSQL	NAVER	최대 1.0MB 평균 0.5MB
연령	정형, 이산형	DB	NAVER DATALAB	최대 0.2MB 평균 0.1MB
성별	정형, 이산형	NoSQL	NAVER DATALAB	최대 0.3MB 평균 0.2MB
지역	정형, 이산형	DB	공공데이터포털	최대 1.5MB 평균 0.7MB
기기	비정형, 명사형	NoSQL	B사 데이터마트	최대 0.4MB 평균 0.3MB
경쟁사 UI	비정형, 명사형	NoSQL	본사 데이터	최대 2.5MB 평균 2.0MB
시간대별 판매율	정형, 이산형	DB	NAVER DATALAB	최대 0.3MB 평균 0.2MB
쇼핑 관련 검색어	비정형, 이산형	NoSQL	NAVER DATALAB	
날씨	정형, 이산형	DB	NAVER DATALAB	최대 0.4MB 평균 0.3MB
결제원	정형, 이산형	DB	NAVER DATALAB	최대 1.5MB 평균 0.7MB

네이버쇼핑

## 데이터 선정 평가

후보 데이터명	평가 세부 항목	평가 내용	평가 점수
분야	서비스 목표 부합 여부	개인에게 맞춤 분야를 추천 하기 위한 UI 개선을 위해 매우 적합	5
	데이터 활용 수준	네이버 내부에서 제공해주는 데이터 이용으로 매우 적합	5
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	5
	데이터 처리 분석 능력	데이터랩으로 유지 보수 가능	5
연령	서비스 목표 부합 여부	연령별 쇼핑품목을 추천 하기 위한 UI 개선을 위해 매우 적합	5
	데이터 활용 수준	네이버 내부에서 제공해주는 데이터 이용으로 매우 적합	5
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	5
	데이터 처리 분석 능력	데이터랩으로 유지 보수 가능	5
성별	서비스 목표 부합 여부	성별에 따라 추천 하기 위한 UI 개선을 위해 매우 적합	5
	데이터 활용 수준	네이버 내부에서 제공해주는 데이터 이용으로 매우 적합	4
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	5
	데이터 처리 분석 능력	데이터랩으로 유지 보수 가능	5

네이버쇼핑

## 데이터 선정 평가

후보 데이터명	평가 세부 항목	평가 내용	평가 점수
지역	서비스 목표 부합 여부	지역별 쇼핑품목을 추천 하기 위한 UI 개선을 위해 매우 적합	3
	데이터 활용 수준	외부에서 제공해주는 데이터 이용으로 활용 가능함	4
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	4
	데이터 처리 분석 능력	네이버 데이터랩부서에서 유지 보수 가능	5
기기	서비스 목표 부합 여부	기기별 UI 개선을 위해 매우 적합	5
	데이터 활용 수준	외부에서 제공해주는 데이터 이용으로 활용 가능함	4
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	4
	데이터 처리 분석 능력	네이버 데이터랩부서에서 유지 보수 가능	5
경쟁사 UI	서비스 목표 부합 여부	인사이트 도출 후에 UI 디자인시 아이디어 반영에 필수	5
	데이터 활용 수준	기존 본사 데이터로 충분히 활용 가능	5
	데이터 처리 가능 여부	디자인팀과 마케팅팀에 자문을 구하여 처리 가능	3
	데이터 처리 분석 능력	디자인팀과 마케팅팀에 자문을 구하여 처리 가능	3

네이버쇼핑

## 데이터 선정 평가

후보 데이터명	평가 세부 항목	평가 내용	평가 점수
시간대별 판매율	서비스 목표 부합 여부	시간대별로 판매율을 통계내어 결과가 반영된 UI를 자동으로 배치	5
	데이터 활용 수준	네이버 내부에서 제공해주는 데이터 이용으로 매우 적합	4
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	5
	데이터 처리 분석 능력	데이터랩으로 유지 보수 가능	5
쇼핑 관련 검색어	서비스 목표 부합 여부	쇼핑에 관련된 검색어들을 종합 분석	2
	데이터 활용 수준	네이버 내부에서 제공해주는 데이터 이용으로 매우 적합	5
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	5
	데이터 처리 분석 능력	쇼핑 UI 제작과는 맞지 않고, 지나치게 분석량이 증가됨	1
날씨	서비스 목표 부합 여부	날씨별 판매율이 높은 상품을 자동으로 우선 배치해주는 UI 개발	5
	데이터 활용 수준	네이버 내부에서 제공받지만 일부는 기상청에서 직접 API를 통해 발굴	3
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	3
	데이터 처리 분석 능력	데이터랩으로 유지 보수 가능	5

네이버쇼핑

## 데이터 선정 평가

후보 데이터명	평가 세부 항목	평가 내용	평가 점수
결제원	서비스 목표 부합 여부	결제원에 따른 판매 추이 분석	1
	데이터 활용 수준	카드 결제시스템과 네이버페이 시스템 모두 분석해야함으로 복잡성 증가	2
	데이터 처리 가능 여부	네이버 내부에서 데이터 활용 가능	3
	데이터 처리 분석 능력	추후 충원될 인력에 따라 달라짐	2
상기 점수 기준 : 1~5 점 까지 (예시) 상 : 5 / 중 : 3 / 하 : 1			

네이버쇼핑

## 최종 데이터 선정

데이터 명	선정 사유
분야	분야별 데이터를 활용하여 상위 분야 품목 우선 표시 및 추천으로 수익률 증가 기대
연령	연령별 데이터를 활용하여 연령대에 맞는 품목 우선 표시 및 추천으로 수익률 증가 기대
성별	성별에 따른 품목 우선 표시 및 추천으로 수익률 증가 기대
지역	지역별 관심도 높은 품목 우선 표시 및 추천으로 수익률 증가 기대
기기	기기 별 UI 개선으로 수익률 증가 기대
시간대별 판매율	시간대별로 판매율을 통계 내어 결과가 반영된 UI를 자동으로 배치
날씨	날씨 별 판매율이 높은 상품을 자동으로 우선 배치해주는 UI 개발
경쟁사 UI	인사이트 도출 후에 UI 디자인시 아이디어 반영에 필수

## 정보보안 및 개인정보 보호 검토 사항

항목	데이터 취득 주체	문제 사항	보완 사항
개인 정보 보호법	본사, NAVER	개인정보보호에 관한 문제	고객사에서 제공해주는 데이터 사항의 경우 익명화 처리 요구
통계법	NAVER	비밀 보장 위배 여부	고객사에서 제공해주는 통계에 대해 보안 사항 제출 요구
공공데이터의 제공 및 이용 활성화에 관한 법률	본사	공공데이터의 영리적 이용 및 권리침해에 관한 문제	공공기관에서 제공해주는 데이터의 익명화 및 불법영리행위 금지 요구
정보통신망법	NAVER	직무상 비밀 누설 및 목적 외 사용	고객사에서 제공해주는 통계에 대해 보안 사항 제출 요구

## 데이터 활용에 대한 제약사항 파악 및 대안

구분	상세 내용	대안 및 필요 사항
활용제한	산업기술의 유출방지 및 보호에 관한 법률	네이버 법률팀과 조율 후 사측 내용 유출에 관한 계약서 작성
법	개인정보 보호법 위반 사항 확인	법률 회사와 고객사 법률팀과 협의 후 보고서 작성
가이드라인	개인정보 비 식별 조치 가이드라인	고객사 데이터에 개인 정보 보호 가이드라인을 준수했는 확인 및 확인서 송부 요구
데이터 관리	데이터 분석을 위한 클라우드 서버 제공	본사 서버 활용 능력의 부족으로 고객사 측 에게 제공 받기로 함



## 목차(기술 환경 분석)

데이터 수집 방식

데이터 저장 방식

데이터 처리 방식

데이터 분석 방식

분석 결과 사용자 컴퓨터 환경에  
제공하는 방식 정의



네이버쇼핑

## 데이터 수집 방식

후보 데이터	데이터 유형 및 크기	수집 경로	제약 사항	수집 방식
분야	비정형 / 5GB	데이터 레이크	비정형, 공개 데이터, API제공	API
연령	정형 / 2GB	데이터 레이크	파일 데이터, 실시간 정보 수집	API
성별	정형 / 1.5GB	데이터 레이크	파일 데이터, 실시간 정보 수집	API
지역	정형 / 3GB	NAVER 데이터베이스	조회 전용, 대용량 데이터	EII
기기	비정형 / 1GB	NAVER 데이터베이스	실시간 정보 수집, 데이터 수시 변경	ETL
시간대별 판매율	정형 / 3GB	NAVER 데이터베이스	조회 전용, 대용량 데이터	EII
날씨	반정형 / 15GB	기상청	실시간 정보 수집, 데이터 수시 변경	ETL
경쟁사 UI	비정형 / 7GB	데이터 레이크	비정형, 공개 데이터	본사 데이터

네이버쇼핑

## 데이터 저장 방식

후보 데이터	데이터 특성	저장 위치 유형
분야	비정형 / 5GB / 실시간 활용 / 온라인 활용	NoSQL
연령	정형 / 2GB / 실시간 활용 / 온라인 활용	HADOOP
성별	정형 / 1.5GB / 실시간 활용 / 온라인 활용	NoSQL
지역	정형 / 3GB / 대용량 데이터 / 분산데이터	Big Data Appliance
기기	비정형 / 1GB / 실시간 활용 / 중앙 집중 데이터	NoSQL
시간대별 판매율	정형 / 3GB / 실시간 활용 / 온라인 활용	HADOOP
날씨	반정형 / 15GB / 대용량 데이터 / 분산데이터	NoSQL
경쟁사 UI	비정형 / 7GB / 본사 데이터 소지	HDD

네이버쇼핑

## 주요 데이터 처리 방식

데이터	데이터 특성	처리 방식
분야별 판매추이	웹 데이터 생성	중앙 집중, 병렬처리
연령별 쇼핑 경향 분석	웹 데이터 생성	중앙 집중, 병렬처리, HADOOP
성별 쇼핑 품목 분석	웹 데이터 생성	중앙 집중, 병렬처리
지역별 판매 추이	데이터 생성 지역 분산	지역 분산, 병렬 처리
기기 별 UI 환경	수집 경로를 통해 데이터 센터에 저장	중앙 집중, 일렬 처리, 실시간 데이터 처리

## 분석 유형에 따른 서비스 모델 특성 파악

구분	분석 방식 요건	관련 기술 사례
배치(온라인 분석)	시간 단위, 일 단위 등의 비교적 긴 시간의 데이터 분석 수행 필요	HADOOP
감성 분석	텍스트에 표현된 사람들의 의견, 성향 및 태도 등 분석	신규 핸드폰에 대한 선호도 조사(SNS 감성분석)
통계 분석	회귀 분석, 시계열 분석, 분산 분석 등의 통계 알고리즘	시간별 쇼핑패턴 분석(회귀 분석, 시계열 분석)

네이버쇼핑

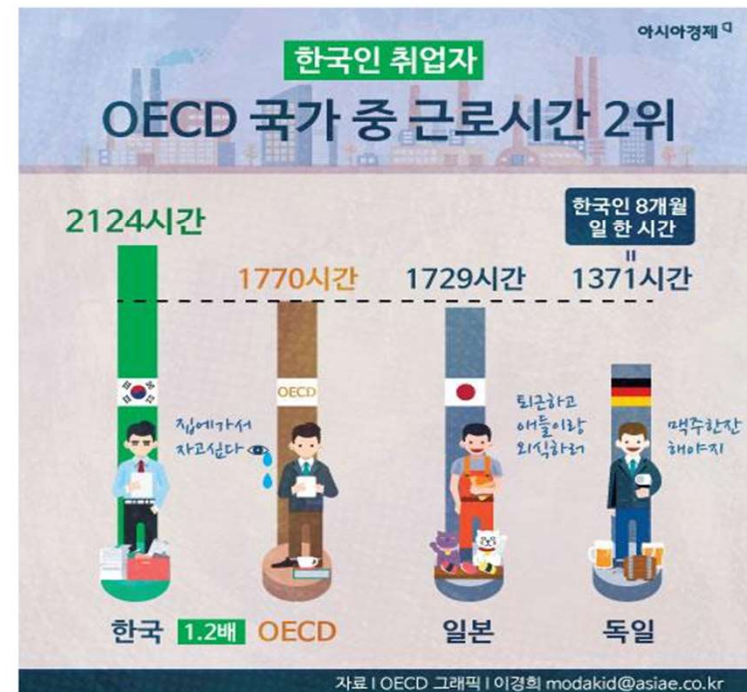
## 서비스 결과 전달

구분	상세	데이터 내용	비고
시각화	인포그래픽	연령대별, 성별 판매 추이를 그래프로 표현	
	막대 그래프	분야별 판매 추이를 그래프로 표현	
	히트맵	지역별 판매 추이를 히트 맵 표현	
	파이차트	기기 별 접속현황 표현	
실시간 전송	분석된 통계	본사가 분석한 내용을 실시간으로 확인 가능, 클라우드 사용(고객사 서버)	
	웹 스크래핑	웹 스크래핑 알고리즘을 통해 실시간 업데이트 사항을 공용 클라우드로 전송	
기타	UI 디자인	구체적인 UI 개선 방향을 제시하기 위해 주 1회 새로운 디자인 공개	본사 디자인팀과 협의 조율 중
	파일 전송	분석된 데이터를 일괄 정리하여 주 2회 고객사에 전체 파일 제공	

네이버쇼핑

## 서비스 결과 전달

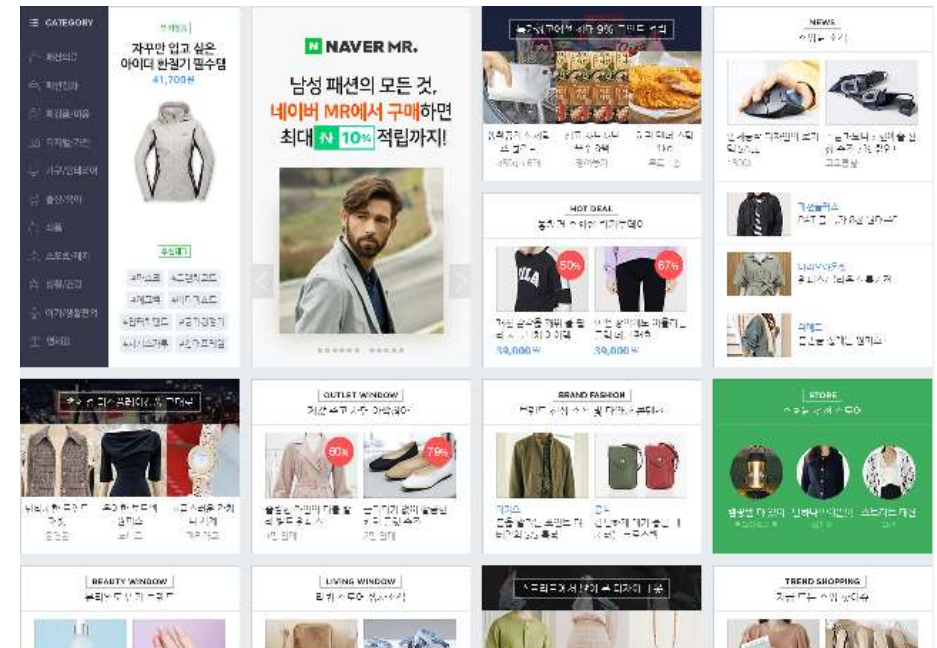
인포그래픽  
예시



인포그래픽

## 서비스 결과 전달

# 최종 비 디자인 예시





# 목차(인력 직무 기술서)

프로젝트 진행 인력 직무 기획

추가 인력 충원 직무 기술



## 프로젝트 진행을 위한 인력직무 기획

단위 그룹	관련 부서	직무	직무 내용	역량
업무 전문	기획	기획 인력	고객사와의 원활한 커뮤니케이션을 위해 지속적 교류 요구	
	마케팅	마케팅 담당 인력	타 쇼핑몰의 마케팅 내용에 대해 기획팀에 지속적 피드백	데이터 분석팀에 마케팅 기술 지원
	인사	인사 담당 인력	프로젝트 진행을 위해 필요한 전문 인력을 신속히 확충	
	구매	구매 담당 인력	본사 서버 확충을 위해 필요한 제반 시설 준비	
디자인 전문	디자인	UI 디자인	◎ 고객사에 제안할 UI 제작	◎ 고객사의 피드백 수용 능력 ◎ UI 제작 실무 경험 ◎ 쇼핑몰 프레임 워크 제작 경험

네이버쇼핑

## 프로젝트 진행을 위한 인력직무 기획

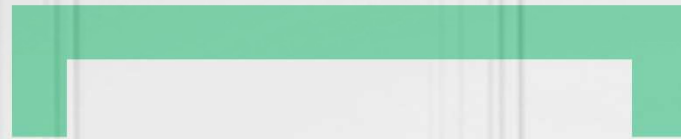
단위 그룹	관련 부서	직무	직무 내용	역량
시스템 전문	데이터 분석	데이터 사이언티스트	<ul style="list-style-type: none"> <li>○ 분석 기획, 데이터 준비, 데이터 분석, 시스템 구현, 피드백에 걸치는 전 과정에 책임</li> <li>○ 본사에 데이터 분석 능력에 대한 총괄적 평가</li> </ul>	<ul style="list-style-type: none"> <li>○ SOFT SKILL(인사이트 도출 능력, 고객사와의 커뮤니케이션 능력, 타 부서와의 협업 능력),</li> <li>○ HARD SKILL(데이터와 분석기술에 대한 이론적 지식과 숙련도) 모두 필요</li> </ul>
		데이터 분석가	<ul style="list-style-type: none"> <li>○ 제공 되는 데이터를 원천 데이터와 정제 데이터를 구분</li> <li>○ 비정형 데이터와 반정형 데이터를 분석하기 위해 논리적 데이터 스토어 개설</li> </ul>	<ul style="list-style-type: none"> <li>○ 분석적 능력, 데이터와 정보에 대한 기술을 활용하여 본 프로젝트에 맞는 데이터 셋업 능력</li> <li>○ 쇼핑 마케팅에 대한 기본적인 이해도가 요구</li> </ul>
		데이터 엔지니어	<ul style="list-style-type: none"> <li>○ 데이터 분석에 요구되는 사항과 데이터 식별</li> <li>○ 원천 데이터 1차 가공과 데이터 셋업</li> <li>○ 고객사에서 제공되는 기본 데이터 뿐 아니라 지속적으로 웹크롤링을 통해 경쟁사 모니터링</li> </ul>	<ul style="list-style-type: none"> <li>○ EAI, ETL의 소프트웨어 솔루션에 대한 이해</li> <li>○ 데이터 품질에 대한 이해도</li> <li>○ 데이터 식별을 위한 논리적 능력</li> </ul>
	시스템 관련	데이터베이스 관리자	<ul style="list-style-type: none"> <li>○ 본사에 부족한 서버 확충을 위한 기반 마련</li> <li>○ 안정적인 서버 운영</li> </ul>	<ul style="list-style-type: none"> <li>○ 데이터 정책 수립 능력</li> <li>○ 데이터베이스 운영 능력</li> <li>○ 리눅스(or 유닉스) 활용 가능</li> </ul>

# 추가 인력 충원을 위한 직무 기술서

직무명	직무 수행 내용	필요 역량
개발자	<ul style="list-style-type: none"> <li>○ 고객사 요구에 맞는 API 개발을 위해 사전 작업 준비</li> <li>○ 필요에 따라 자동 업데이트를 위한 머신 개발</li> </ul>	<ul style="list-style-type: none"> <li>○ 풀스택 개발자</li> <li>○ 기초적 데이터 처리 지식</li> <li>○ API 개발 실무 경험(5년 이상)</li> <li>○ PYTHON 활용 능력</li> </ul>
빅데이터 전문가	<ul style="list-style-type: none"> <li>○ 고객사에서 제공되는 대량의 정형, 비정형 데이터를 다양한 분석 기법으로 핸들링</li> <li>○ 머신러닝을 통한 예측 데이터 작성</li> </ul>	<ul style="list-style-type: none"> <li>○ 빅데이터 분석기사 자격증 필수</li> <li>○ 빅데이터 핸들링 실무 경험(3년 이상)</li> <li>○ 머신러닝 데이터 분석 능력 (TENSORFLOW 가능)</li> <li>○ 텍스트 마이닝 분석 경험</li> <li>○ API 개발 가능 우대</li> </ul>
데이터 분석가 (비주얼라이저)	<ul style="list-style-type: none"> <li>○ 데이터 분석 결과를 스토리텔링 및 시각화</li> <li>○ 데이터 분석 결과로부터 새로운 인사이트 도출</li> <li>○ 데이터 사이언티스트와 협업</li> </ul>	<ul style="list-style-type: none"> <li>○ R 또는 PYTHON(비주얼라이제이션 스킬) 활용 능숙</li> <li>○ 데이터 사이언티스트와의 협업 경험</li> <li>○ 데이터 사이언티스트에 견줄 현업에서의 경험</li> <li>○ 데이터분석의 전략적 시각 소지</li> <li>○ 분석 결과를 스토리라인으로 완성시킬 스킬 필수</li> <li>○ SQL 개발 응용, 활용 능력</li> </ul>



감사합니다



# NAVER SHOPPING

BIGDATA 빅데이터분석

