

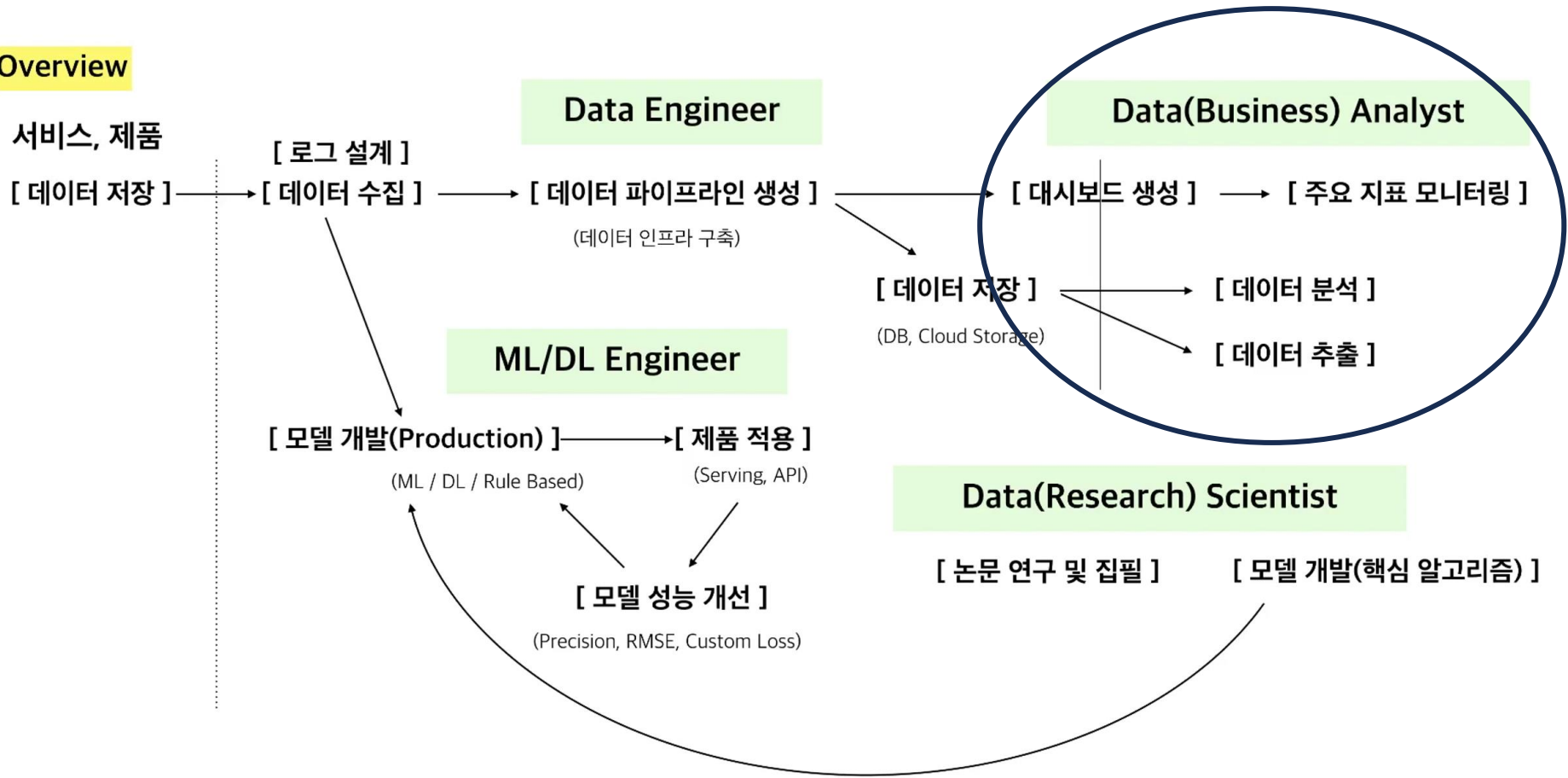
# Deba Lab

무역학부 전공 소모임

# Study Purpose

# Study Purpose

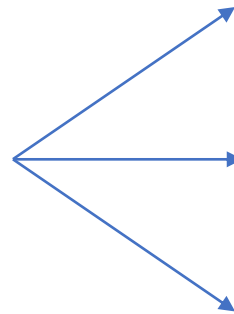
## Overview



Study Purpose

Data(Business) Analyst 란?

데이터를 분석하여



분석 및 모니터링

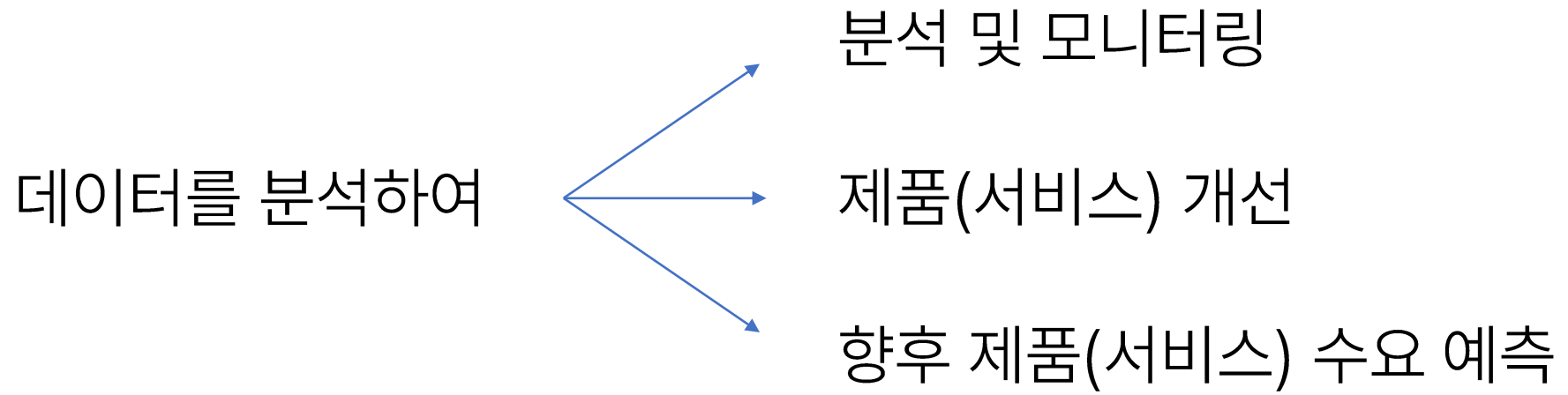
제품(서비스) 개선

향후 제품(서비스) 수요 예측

Study Purpose

Data(Business) Analyst 란?

기획가, 분석가적 성향



# Study Purpose



# Study Purpose

모집부문		모집인원	주요 업무
UB	기업금융	000명	기업고객 대상 금융상담 및 상품판매, 기업금융 서비스 등
	자산관리		개인 및 PB고객 대상 금융상담 및 상품판매, 고객자산관리서비스 등
IB		00명	M&A 인수금융, 인프라금융, 부동산금융 등 투자금융(IB) 업무
글로벌		00명	국외점포 등 설립 및 운영, 현지 영업 추진, 해외 지분 투자 전략기획 등
자본시장		00명	외환, 이자율 상품 및 파생상품 트레이딩, 자본시장 상품 세일즈 등
ICT	IT	00명	은행 내 IT 서비스 개발 및 운영, 보안 및 개인정보보호 관련 업무 등
	IT_플랫폼개발	00명	금융/비금융 플랫폼 개발, 신기술 기반 디지털 신사업 추진 등
	데이터·AI	0명	정형/비정형 데이터분석, 인공지능 및 빅데이터 기반 서비스 개발 등
	ICT_장애인	0명	은행 내 IT서비스, 금융/비금융 플랫폼 개발 및 운영 등
전문자격		00명	변호사(韓), 공인회계사(KICPA)
직무전문가		00명	리스크관리, 전략기획, 재무관리, 리크루팅
보훈 및 ESG동반성장		00명	은행 일반 업무(입출금, 예금 신규 및 해지, 제신고, 카드, 여신, 외환 등)

# Study Contents



# Study Contents

## 1. Basic Study

- 데이터 분석 툴(파이썬, R, 엑셀)
- 김경원 교수님 강좌 및 외부 소스 활용

## 2. Outcome

- 공모전 > 논문 > 창업 > 프로젝트
- 자격증

## Study Contents

### 증명할 수 있는 Outcome

- 공모전 > 논문 > 창업 > 프로젝트
- 자격증

# Study Contents

## 지난 기수 활동

### Achievements of the Previous Term

DATA 100

03 (2) 회귀분석

선형 회귀분석

1. 독립변수 값에 따라 종속변수 값이 영향을 받는 관계성을 분석

2. R<sup>2</sup>값으로 모형이 데이터에 잘 적합한 정도를 파악

3. T-검정으로 t분포기반 독립변수와 종속변수 간의 영향력정도 의사결정

4. DW값으로 데이터의 패턴이 모델링에 잘 반영되었는지 확인

```
# 시도 삭제
df=df.drop('시도',axis=1)

# 종속변수를 총이용자수로 설정
Y_colname = ['총이용자수']
X_colname = [x for x in df.columns if x not in Y_colname]
Y_colname, X_colname

# 데이터프레임 만들기
Y= pd.DataFrame(df['총이용자수'])
X = pd.DataFrame(df[X_colname])

# Min-Max스케일링 실행
scaler = preprocessing.MinMaxScaler()
X_p = scaler.fit_transform(X)
X_p = pd.DataFrame(X_p, columns=X_colname)

#결과
import statsmodels.api as sm
X_p = sm.add_constant(X_p)
model_lr = sm.OLS(Y, X_p).fit()
model_lr.summary()
```

○

데이터 전체의 왜곡을 줄이기 위해 Y절편 또는 상수항을 출력력을 위한 상수 변수를 추가

DATA 100

03 (1) 군집분석

Silhouette Plot of K-Medoids Clustering for 16 Samples in 4 Centers

K-Medoids Clustering with k=4

Silhouette Plot of K-Medoids Clustering for 16 Samples in 5 Centers

K-Medoids Clustering with k=5

군집의 실루엣 계수와 분포상 데이터가 가장 적절히 구분되는 K=4로 선정

유출이 유입보다 많은 Cluster2의 광역시도를 선정  
-> 인천, 세종, 울산, 충북