

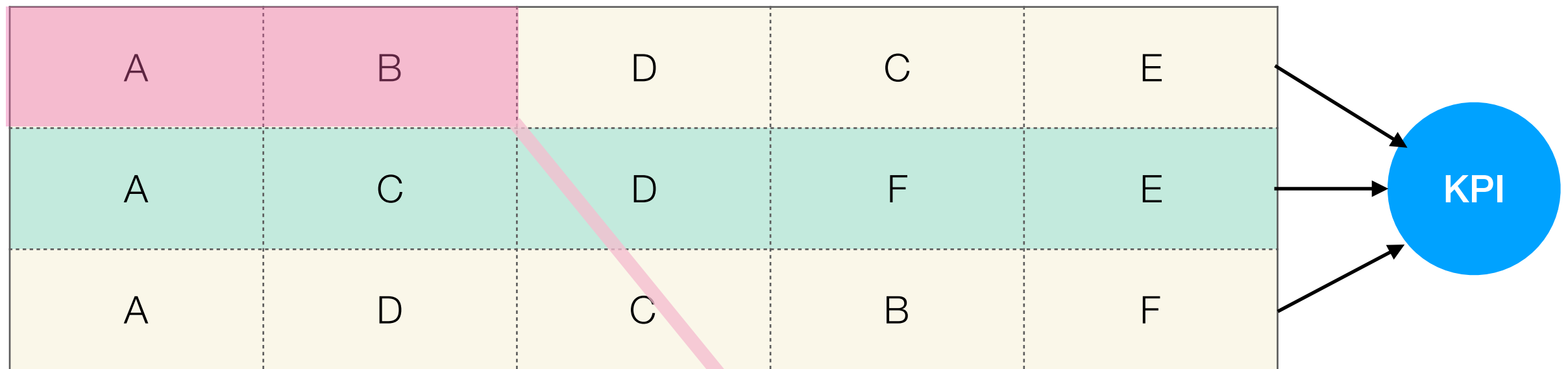
# CDJ 분석 방향

tan.d

# 분석 목표

- KPI 에 도달하는 경로 중 가장 빈번하게 나타나는 패턴을 제공
  - 패턴 마이닝 (sequential pattern mining) 기법을 활용
- 사용자가 KPI에 도달할 확률 제공
  - 이탈분석에 사용했던 로지스틱 회귀모형 또는 다른 여러 모형을 시도해 볼 예정
- 고객사가 위의 두 가지 서비스를 동시에 활용할 수 있는 기능을 구현 (뒤에서 자세히 설명)

## KPI 에 도달하는 경로 중 가장 빈번하게 등장하는 패턴



고객사는 패턴에 대한 정보 뿐만 아니라,  
특정 경로를 거친 고객이 KPI에 도달할 확률 또한 제공받는다

ex) A -> B 를 거쳐간 고객 중 KPI에 도달할 확률이 높은 segment (상, 중, 하)  
KPI에 도달할 확률이 낮은 segment(상, 중, 하) 제공



해당 기간에 특정 시퀀스를 경험한 유저 중에서 이 기간 안에 conversion을 할 유저와 그렇지 않을 유저

n일의 기간 동안 

A	B	D
---	---	---

 을 거친 고객 중

- KPI에 도달할 확률이 높은 그룹
- KPI에 도달할 확률이 낮은 그룹

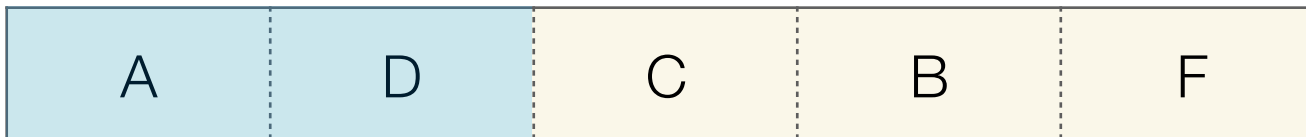
# 해당 기능이 제공하는 이점

- KPI에 도달하는 과정을 직관적으로 이해할 수 있음
  - 모바일 앱 사용자의 이동 경로는 매우 복잡하나 패턴 마이닝을 통해 핵심적인 경로 파악이 가능
- 경로 파악에 그치지 않고 실제 액션을 취할 수 있도록 고객 세그먼트 정보 제공
  - 고객마다 어떤 채널에 액션을 취해야 할 지 판단이 가능, 예측 확률에 따라 나뉜 세그먼트 별로 상이한 액션을 취할 수 있음

# 논의점

## 1. 패턴 마이닝

- 1) 패턴 마이닝 인풋에 들어갈 시퀀스의 단위를 세션 단위로?
- 2) 중복되는 채널 어떻게 처리? Unknown 채널 추가?
- 3) 제공되는 패턴의 길이는 어떻게 설정?
- 4) 단계별로 골에 도달하는 확률까지 제공?



해당 채널을 거친 경우 conversion 할 확률 30%

## 2. 예측 모델링

- 1) 패턴 별로 모델을 학습? 하나의 모델로 통일?
- 2) 다양한 종류의 모델 시도 가능
- 3) 성능에 관한 부분은 모델링 단계에 들어가야 논의 가능할 듯