

알고 타자! 믿고 타자! 같이 타자!
오늘의 항공사 예측 분석 서비스

On : Air

A203 파이썬의 연어들
팀장 박지우/ 팀원 김채은 김현욱 나승호 안수빈



TABLE OF CONTENTS

- 01. 기획 배경
- 02. 타겟 사용자
- 03. 기존 서비스와의 차별점
- 04. 주요 기능
- 05. 개발 및 분석 계획
- 06. 디자인 컨셉
- 07. 기대 효과 및 고도화 전략



01 기획 배경

01 기획 배경



“

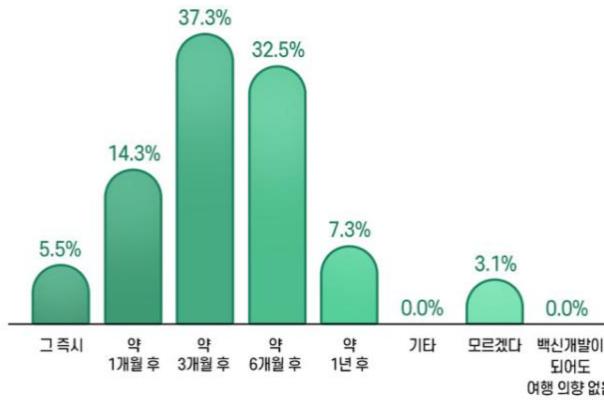
"빨리 백신 맞고 해외여행
가자"

...

여행 준비하는 MZ세대

”

코로나19 종식 후 해외여행 의향 및 시기



”

유럽행 특가 항공권
폭풍 품절

...

‘보복 여행’ 조짐 논란

”

“해외여행 목말라”

...

‘무착륙관광비행’
시행 9개월, 2만명 돌파

”

02 타겟 사용자

02 - 1 문제 상황과 니즈



항공사 이용객

12월 마카오행 비행기 지연된 날 봤더니…

*제주항공은 하루 2회 운항

항공	날짜(총 지연 횟수)
티웨이항공	5일, 15일, 16일, 18일, 21일, 22일, 23일, 25일(8회)
진에어	5일, 9일, 21일(3회)
제주항공	1일, 2일, 3일, 5일, 7일, 17일, 20일, 21일, 22일, 23일, 24일, 25일, 26일, 30일, 31일(15회)
에어서울	3일, 8일, 18일, 23일, 29일, 30일(6회)

〈자료: 항공정보포털시스템〉

국적항공사 국제선 지연율(단위: %)



〈자료: 한국공항공사·인천국제공항공사〉



다양한 항공사 지연의 원인을 분석해서 미리 예측할 수 없을까?

02 – 2 타겟 사용자



여행 계획을 세울 때
자세한 항공사 정보를
얻고 싶은 여행객

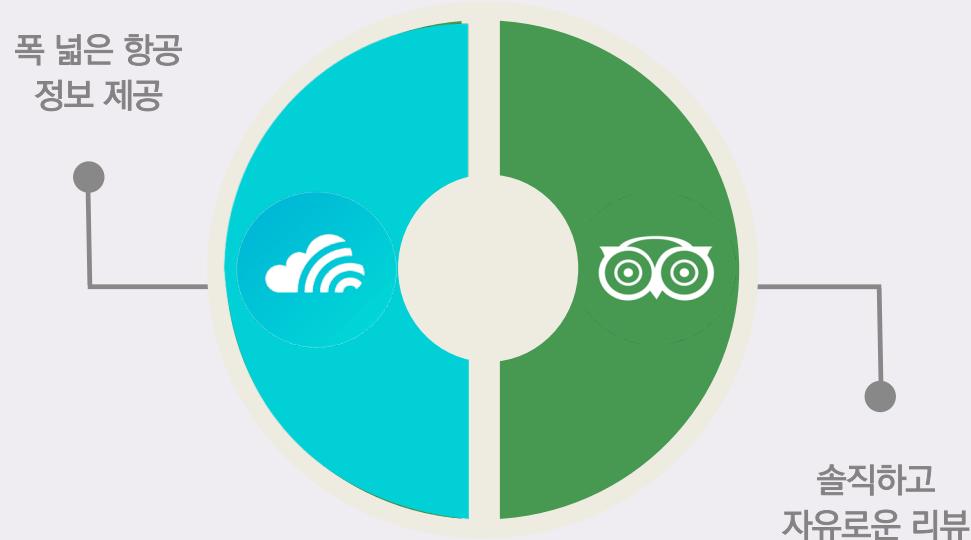


항공사에 대한
솔직한 리뷰를
나누고 싶은 여행객



자사에 대한
솔직한 평가를
알고 싶은 항공사

02 - 3 요구사항



항공사 리포팅 서비스와
사용자의 리뷰
모두를 한 번에
확인할 수 있는 서비스

03 기존 서비스와의 차별점

03 기존 서비스와의 차별점

On : Air

트립 어드바이저

구글 플라이트



항공사 기본 정보 제공	0	0	0
항공사 분석 리포트 제공	0	X	△
리뷰	0	0	X
지연 정보 및 사유	0	X	△

오늘의 항공사 예측 분석 서비스

On : Air 온에어



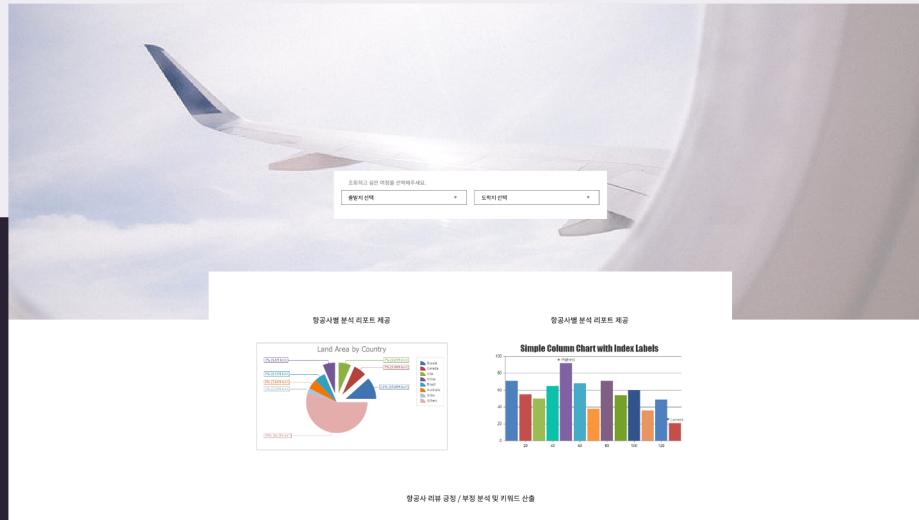
04 주요 기능

04 – 1 주요 기능 Lookup

항공편 조회

날짜와 도착지를 입력하면 해당 노선을 운항하는 항공사의 정보 제공

- ✓ 통합 지연율
- ✓ 해당 노선의 출발 지연율
- ✓ 평균 지연시간
- ✓ 항공사 특징: 리뷰 기반 키워드, 기내식, 서비스, 게이트 및 터미널 등



04 – 2 주요 기능 Report

항공사별 분석 상세 조회

항공사 선택 시 분석 리포트 확인 가능

- ✓ 리뷰 기반의 대표 키워드
- ✓ 10분 이내 출발률 등 세부적인 출도착 정보 제공
- ✓ 지연율: 평균 및 예측 지연율, 지연 사유 제공
- ✓ 여객 인원수 추이 등 기본 정보 제공



본사: Korean Air Operations Center, 1370 Gonghang-Dong, 서울 157712

1588-2001

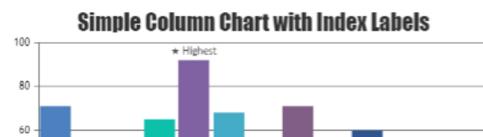
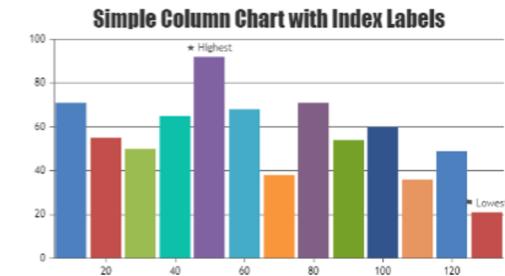
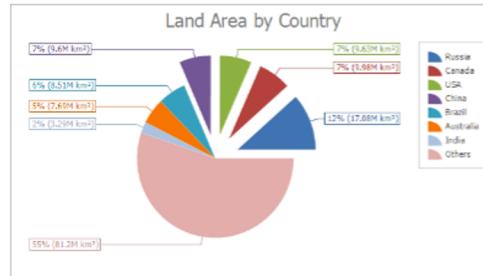
웹사이트 ↗

코로나19 업데이트 ↗

#국적기 #기내식 #서비스 #승무원

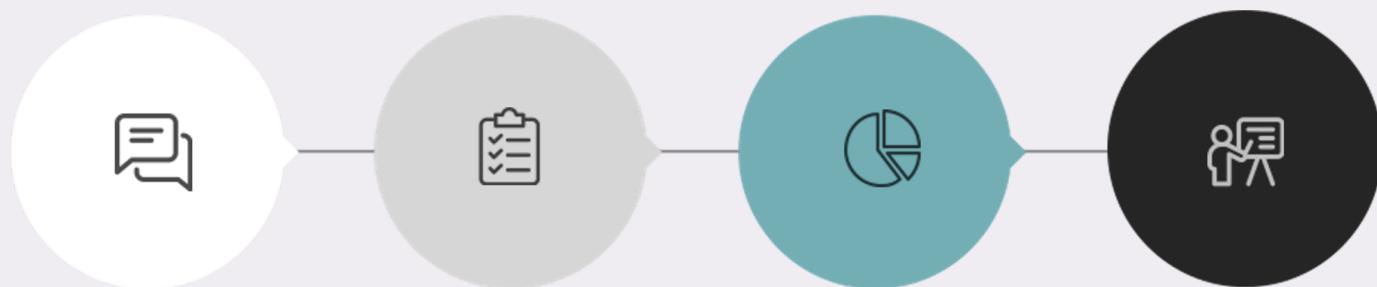
통합 지연률 12% | 파리(CDG)로 출발 항공기 지연률 20% | 평균 지연시간 23분

10분 이내 출발율 80% 30분 초과 지연 12% 총 운항횟수 899회



05 개발 및 분석 계획

05 - 1 개발 계획



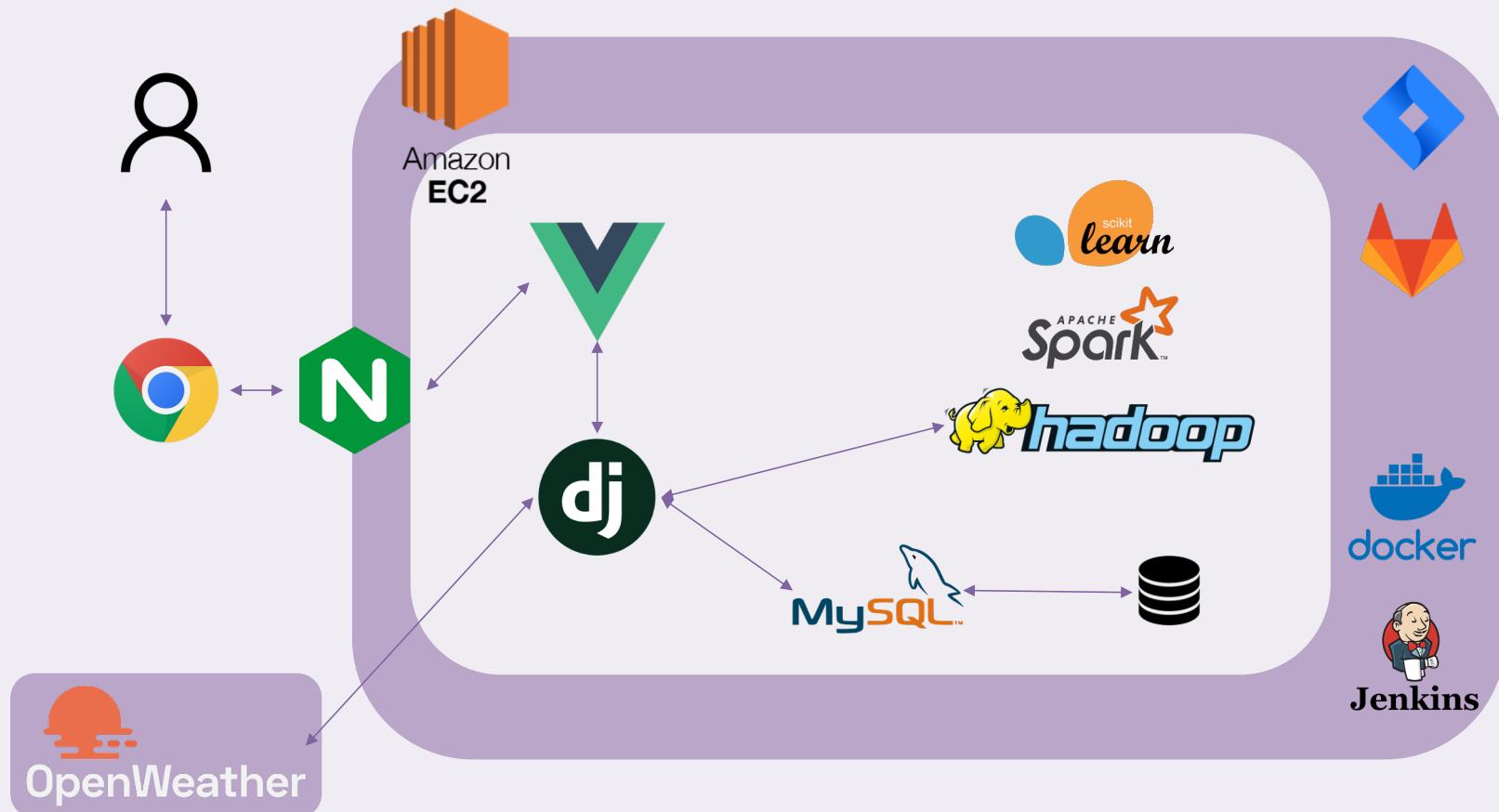
1 – 2주차
아이디어 회의
주제 선정

3주차
기획
데이터 계획

4 – 5주차
빅데이터 분석
개발

6주차
최종 점검
발표

05 – 2 프로젝트 아키텍처



05 - 3 분석 계획

항공사 리뷰기반 감성분석 계획



목표

항공사 별 리뷰 긍부정 비율 계산

항공사 별 높은 빈도의 키워드 추출

항공사 별 분석 리포트에서 참고 기준으로 제공



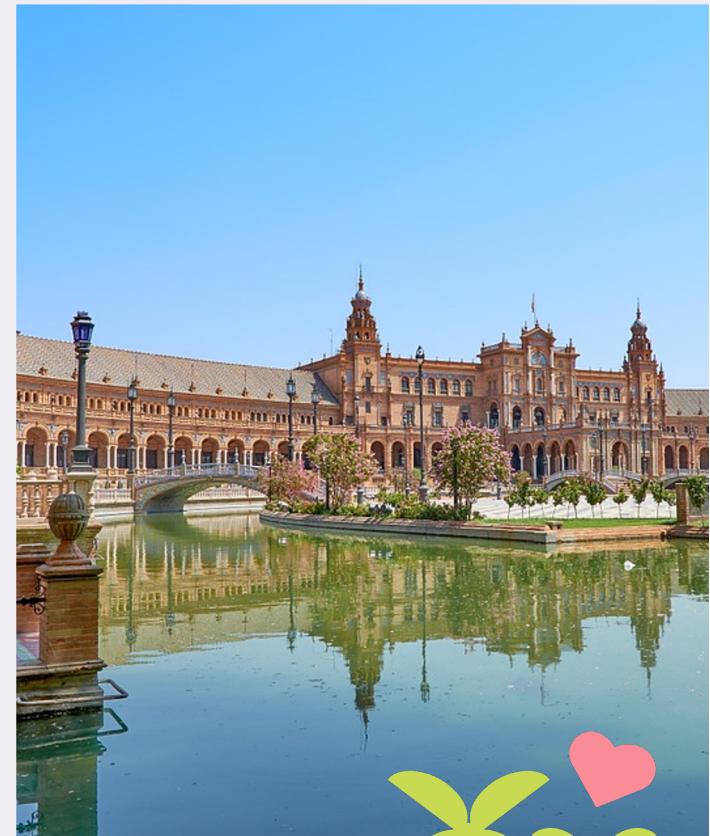
데이터

트립어드바이저

<https://www.tripadvisor.co.kr/>

주요 항공사의 한국어 리뷰 크롤링

대한항공, 아시아나 등 13개 항공사





항공사 리뷰 기반 감성분석 계획

○ 분석계획

1. 리뷰 데이터 크롤링 트립어드바이저 API
2. 리뷰 데이터 하둡 저장 및 분산 처리
3. 리뷰 데이터 전처리
4. 머신러닝 모델 학습 및 구축
 - 모델 학습: 네이버 영화 리뷰 데이터 활용
 - 분류 정확도 확인 후 비교하여 모델 선택
 - 대표적 모델: 단어를 벡터로 바꾸는 단어 임베딩 방법론

Word2Vec : CBOW, Skip-Gram 단어 예측, 단어 문맥 행렬 빈도수

Glove : 단어 동시 등장 여부 보존, 단어 벡터의 내적은 코사인 유사도

5. 리뷰 데이터 분석
6. 워드 클라우드 시각화

05 – 3 분석 계획

항공사 지연시간 예측분석 계획



목표

항공사 별 지연 관련 통계 제공

- ✓ 기상별 지연 시간 예측
- ✓ 월별(이용객이 예측된) 지연 시간 예측
- ✓ 오늘의 기상 & 이달의 이용객수에 따른 지연시간 예측



데이터

- ✓ 항공기 출/도착 운행 데이터
- ✓ 항공사 월별 이용객수 데이터





항공사 지연 시간 예측분석 계획

○ 분석계획

1. 항공기 출/도착 운행 데이터 크롤링, 항공사 월별 이용객수 데이터 파일 수집
2. 항공기 운행 데이터 하둡 분산 처리 및 전처리
3. 기본 통계 그래프 도출
4. 지수평활법 활용 시계열 예측 모델, 다중선형회귀 머신러닝 모델 학습 및 구축
 - 지수평활법을 통해 여러 모델 구축
 - RMSE, MAPE, MASE 지표를 확인 후 모델 선택
 - 경사하강법을 통해 다중선형회귀 모델 구축
5. 선택된 모델로 데이터 분석



항공사 지연 시간 예측분석 계획

○ 분석계획

1. 원핫인코딩을 통해 범주형 변수인 도착지, 항공사, 기상을 수치값으로 변환
2. 해당 목적지로의 기상별 지연시간 예측
 - $H(x) = w1 * \text{도착지} + w2 * \text{항공사} + w3 * \text{기상} + b$
 - 기상 상태별로 머신러닝 모델을 통해 $H(x)$ 값을 구한 후 그래프로 표현
3. 해당 목적지로의 월별(이용 객별) 지연시간 예측
 - $H(x) = w1 * \text{도착지} + w2 * \text{항공사} + w3 * \text{월별 이용객수} + b$
 - 월별 이용 객 수를 통해 향후 12개월의 지연시간을 예측한 후 그래프로 표현
4. 해당 목적지로의 오늘 기상, 이 달의 이용 객수를 기반으로 한 지연시간 예측
 - $H(x) = w1 * \text{도착지} + w2 * \text{항공사} + w3 * \text{기상} + w4 * \text{월별 이용객수} + b$
 - 4개의 변수를 사용한 모델을 만든 후 각 값을 넣어 예측값을 도출

06 디자인 컨셉

오늘의 항공사 정보 분석 서비스

On : Air

온에어

알고 타자! 믿고 타자! 같이 타자!

위드 코로나(With-Corona) 시행에 대한 기대감과 함께 해외여행 계획을 짜고 있는 사람들이 늘고 있다. On:Air(온에어)는 이러한 기대감에 힘입어 실제 이용객이 항공사를 선택하고, 평가하는데 도움을 줄 수 있도록 개발되었다.

On-Air(온에어)는 빅데이터를 활용하여 소비자가 항공사 별 서비스 품질을 시각화한 통합 웹 서비스이다.

데이터를 가공, 분석하여 지역별, 결항률 통계를 시각화하여 제공하며 현재 운행중인 비행기의 지역 시간을 다각도로 예측하여 예비 이용객이 참고 할 수 있도록 하였다. 또한 항공사에 대한 리뷰를 남기고 리뷰를 기반으로 한 실사용자의 항공사 별 평가를 감정 분석하여 제공한다.

디자인 컨셉

도시적인 세련미와 구름 위에 오른 것 같은 몽환적인 풍경을 연상시키는 라벤더 / 베이지 계열로 전개

기획 배경

기존의 어렵고 복잡한 항공 정보를 실제 이용객이 보기 쉽게 데이터를 가공하고 분석하여 제공

팀장

박지우

팀원

김채은 김현욱 나승호 안수빈

글꼴

제목

넥슨 풋볼 고딕

본문

KoPubWorld돋움체 Bold

KoPubWorld돋움체 Medium

메인 색상



배경색 #EFEDF2

그라데이션 예시



디자인 예시



07 기대 효과 및 고도화 전략

07 - 1 기대 효과

“ 이상하게 제주항공 비행기를 탈 때마다
연착이 되었는데 방송만 지나치다 싶게 하고
정확한 이유는 들을 수가 없어 불안감은 증폭되었다.”
– 트립어드바이저 제주항공 리뷰 중



보다 공개적으로 항공사 이용 경험을
공유하여 정보의 풀(pool) 확대



제주	대한항공	KE 1259
울산	SK항공	KE 1607
진주	제주항공	4H 1347
제주	에어서울	RS 907
제주	제주항공	OZ 8015
제주	티웨이항공	TW 731
광주	진에어	LJ 373
부산	진에어	LJ 361
제주	제주항공	OZ 8981
제주	대한항공	KE 1263
부산	제주항공	OZ 8825
부산	티웨이항공	TW 9977
부산	제주항공	TC 274
제주	제주항공	OZ 8983
제주	제주항공	OZ 8983

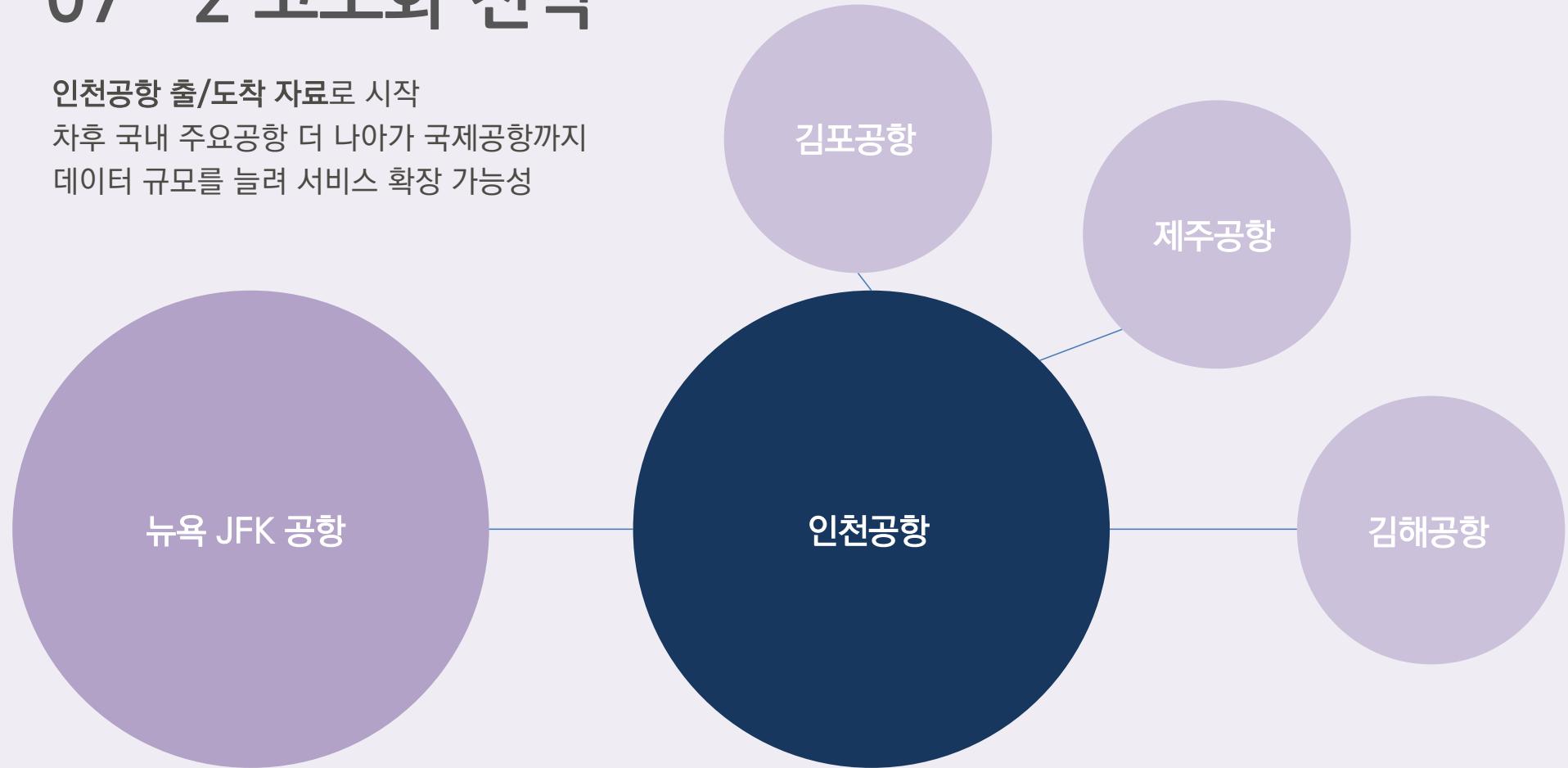
지연/결항 가능성을 예측하여
고객 스트레스 절감의 효과



항공사 서비스 개선에
실사용자의 리뷰 반영

07 - 2 고도화 전략

인천공항 출/도착 자료로 시작
차후 국내 주요공항 더 나아가 국제공항까지
데이터 규모를 늘려 서비스 확장 가능성



Q & A

감사합니다!