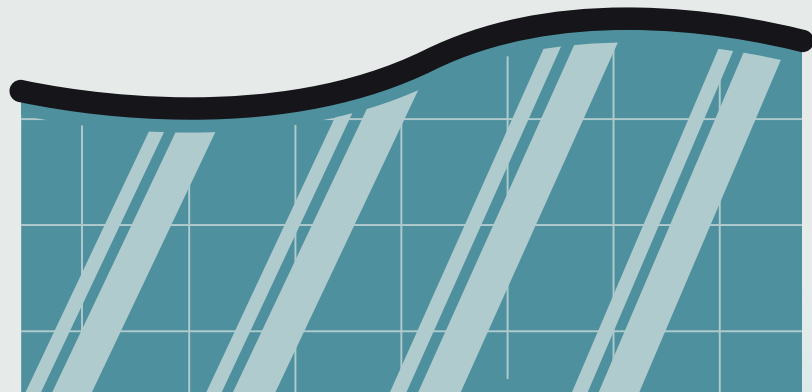


# Improving American Airlines Passanger Satisfaction

AIB 16 김재이



# 프로젝트 필요성

- 항공 산업은 코로나19 대유행 이후 큰 후퇴를 겪었습니다.
  - ICAO(국제 민간 항공 기구)에 따르면 2020년에 항공 산업은 3710억 달러의 손실을 입었다.
- 따라서 이러한 불황을 감안하여 업계를 회생시키기 위해서 고객의 애로사항을 파악하고 제공되는 서비스에 대한 만족도를 높이는 것은 매우 중요하다.
- 여러 연구에 따르면 항공사의 서비스 품질을 높이는 것은 고객의 재이용률과 충성도에 긍정적인 영향을 끼치는 것으로 확인되었습니다.
- 따라서 고객의 **만족도에 크게 영향을 미치는 서비스는 무엇인지 파악하고 부족한 점을 개선**하는 일은 몹시 중요합니다.

# 데이터셋 설명

## US Airlines Passanger Satisfaction

- <https://www.kaggle.com/datasets/teejmahal20/airline-passenger-satisfaction/code>

총 25개의 column

Train data size : 103,904

- Train / Validation data로 나누어서 교차검증을 진행

Test data size : 25,976

# 데이터셋 설명

Gender - 고객 성별

Customer Type - 충성 고객 여부 (Disloyal 또는 Loyal)

Age - 고객 나이

Type of Travel - 비행의 목적 (Business 또는 Personal Travel)

Class - 좌석 클래스 (비즈니스, 이코노미, 기타)

Flight Distance - 비행거리

Seat comfort - 좌석 편안도

Departure/Arrival time convenient - 출발/도착 시간의 만족도

Food and drink - 식음료 제공 만족도

Gate location - 게이트 위치의 만족도

Inflight wifi service - 기내 와이파이 서비스 만족도

Inflight entertainment - 기내 오락시설 만족도

Ease of Online booking - 온라인 예약의 과정이 얼마나 간편하고 만족스러웠는가

Leg room service - 좌석의 다리 공간 만족도

Baggage handling - 수하물 처리 만족도

Checkin service - 체크인 서비스 만족도

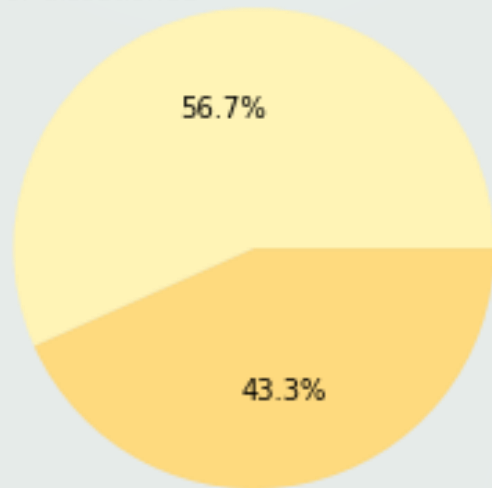
Cleanliness - 청결에 대한 만족도

Departure Delay in Minutes - 출발 지연 시간

Arrival Delay in Minutes - 도착 지연 시간

**satisfaction** - 고객 만족 여부

Neutral or dissatisfied

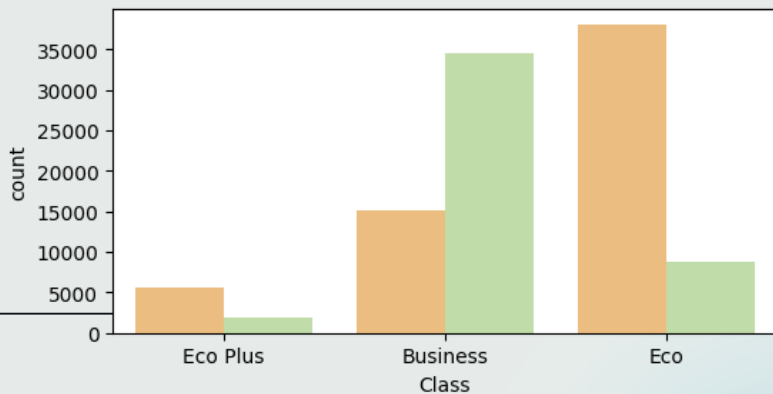
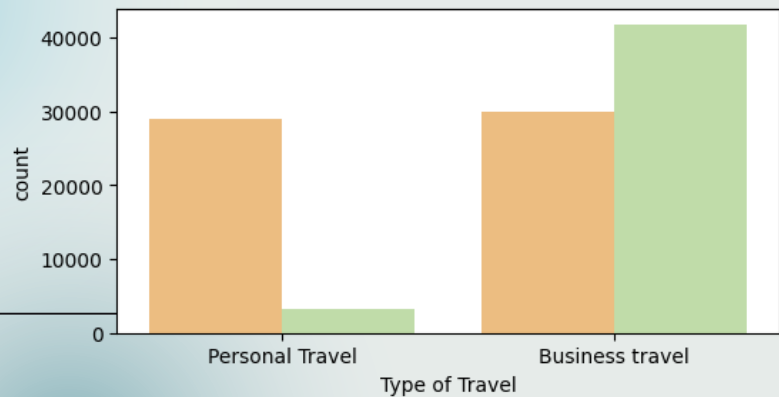
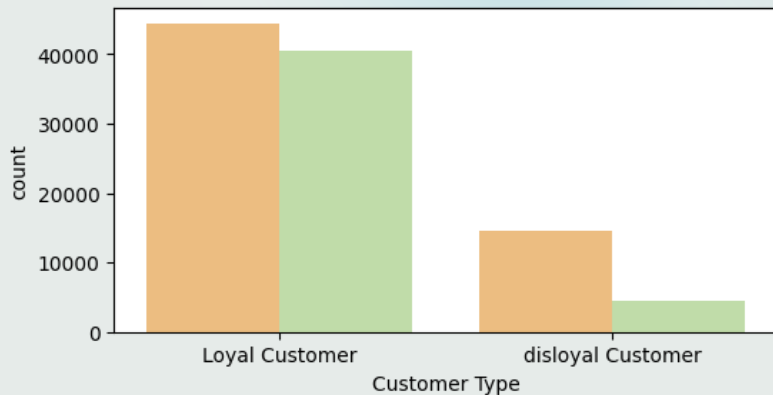
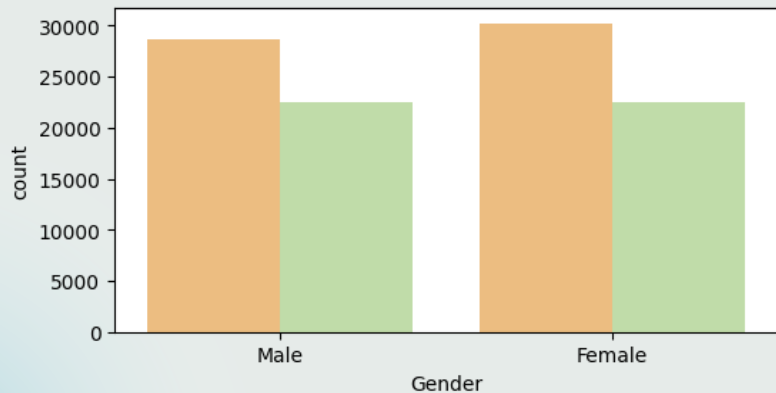
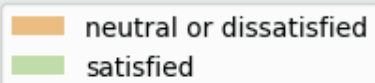


Satisfied

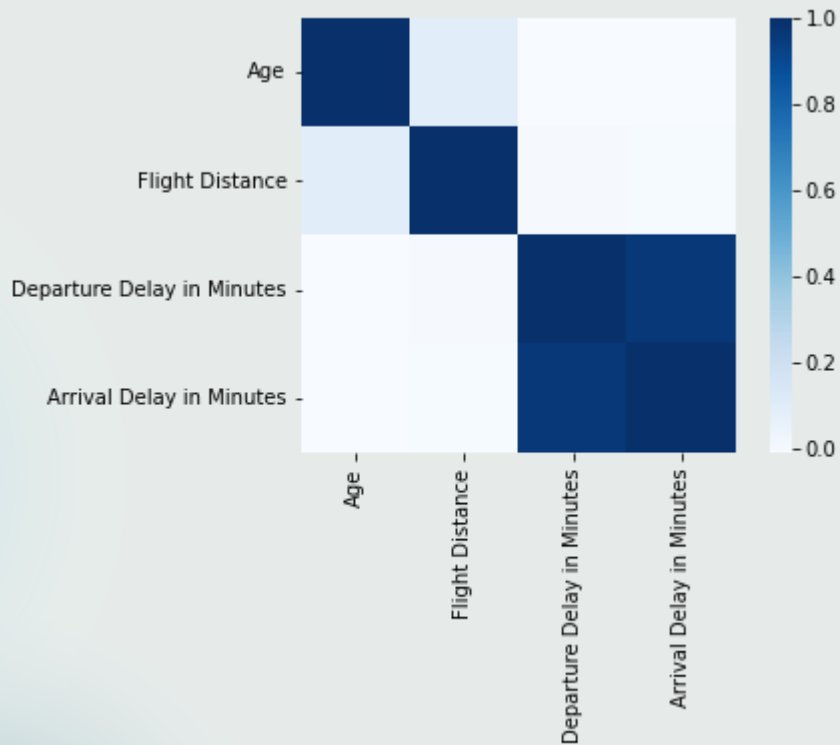
# 가설 제시

- 여러 요인들 중 고객의 만족도 향상에 가장 중요한 것은 기내 서비스이다.
- [기내 서비스] - Inflight wifi service, Inflight entertainment, Food and drink, Inflight service
- [공항여객 서비스] - Gate location, Baggage handling, Checkin service
- [지연] - Departure Delay in Minutes, Arrival Delay in Minutes

# 시각화



# 시각화



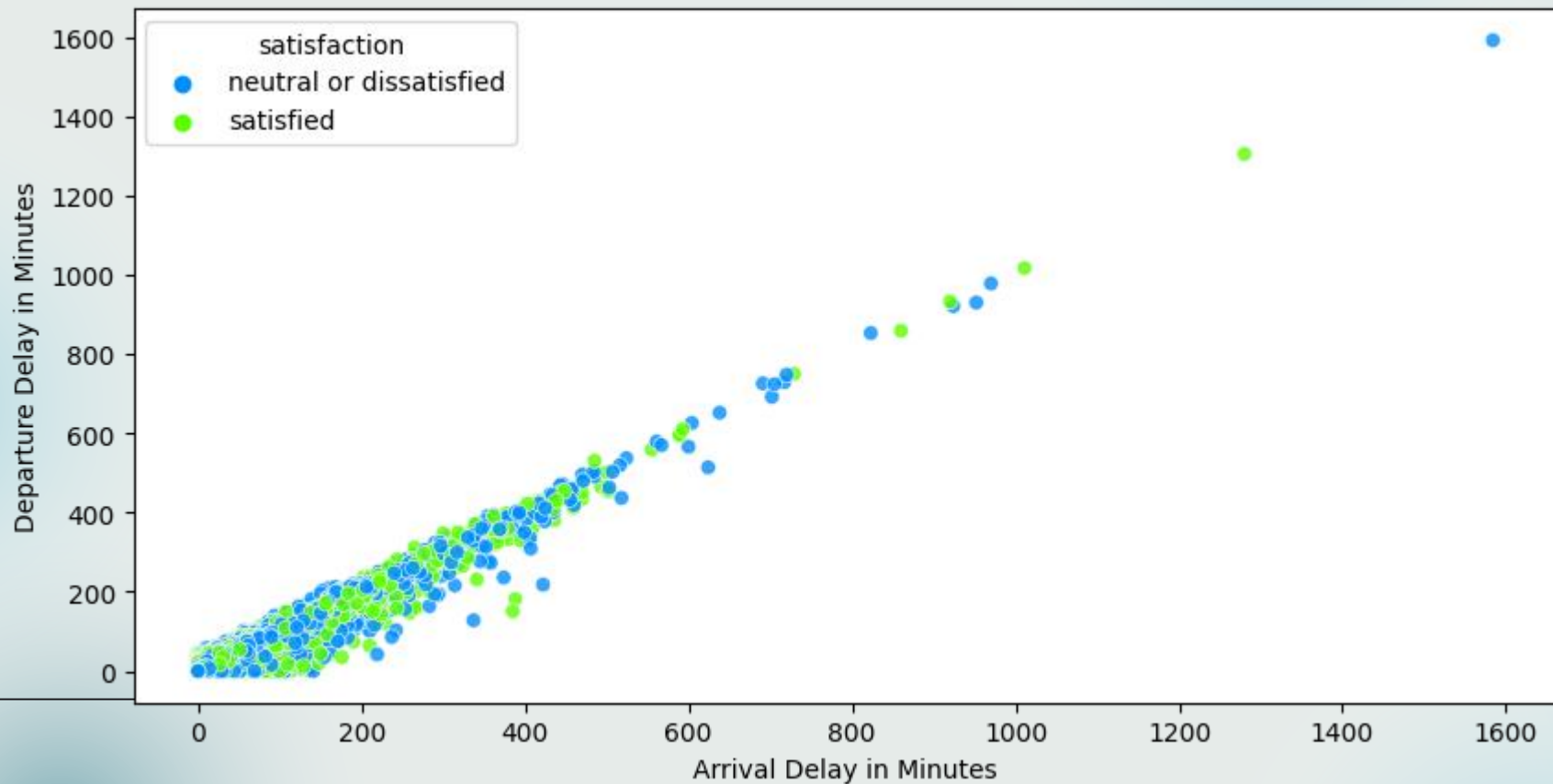
## <수치형 컬럼끼리의 상관관계 분석>

Departure Delay in Minutes  
Arrival Delay in Minutes

두 특성은 높은 상관관계에 있다

# 시각화

두 특성은 선형관계를 이루는 것을 확인





# EDA

```
0  Unnamed: 0  
1  id
```

분석에 필요없는 컬럼 삭제

```
Checkin service      0  
Inflight service     0  
Cleanliness          0  
Departure Delay in Minutes  0  
Arrival Delay in Minutes 310  
satisfaction         0
```

Arrival Delay in Minutes 에 310개의 결측치 발견  
→ Departure Delay in Minutes 값으로 채워주었음

```
def transform_satisfaction(x):  
    if x == 'satisfied':  
        return 1  
    elif x == 'neutral or dissatisfied':  
        return 0  
    else:  
        return -1
```

Satisfaction 컬럼의 값을 0과 1로 바꿔주었음

# 모델링

```
ROC_AUC: 0.9590923832232073
      precision    recall  f1-score   support

     0       0.95      0.98      0.97     11713
     1       0.97      0.94      0.96      9068

 accuracy          0.96      20781
 macro avg       0.96      0.96      0.96      20781
 weighted avg    0.96      0.96      0.96      20781
```

RandomForest

```
ROC_AUC: 0.9603196756072893
      precision    recall  f1-score   support

     0       0.96      0.98      0.97     11713
     1       0.97      0.94      0.96      9068

 accuracy          0.96      20781
 macro avg       0.96      0.96      0.96      20781
 weighted avg    0.96      0.96      0.96      20781
```

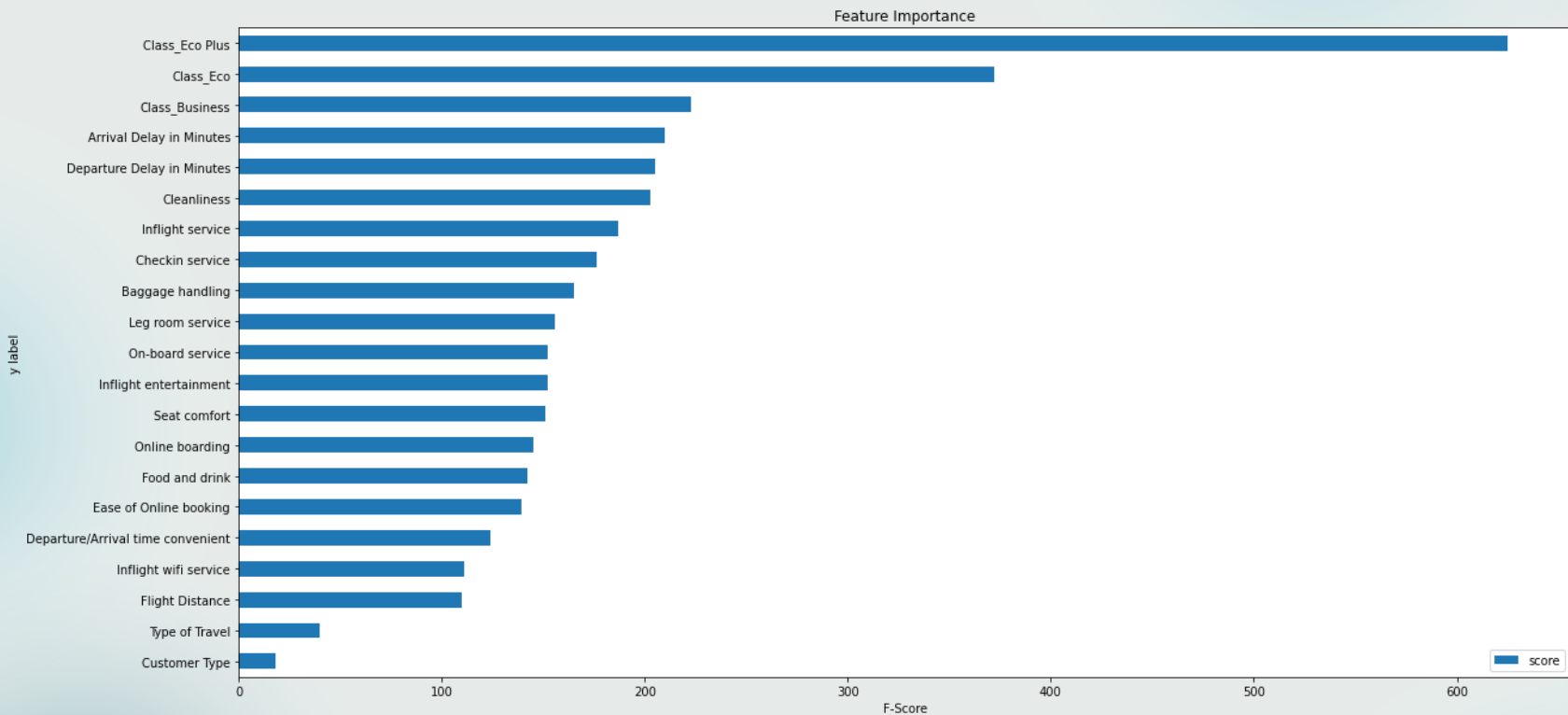
XGBoost

# 모델링

ROC\_AUC: 0.960452889085572

	precision	recall	f1-score	support
0	0.96	0.98	0.97	14573
1	0.97	0.94	0.96	11403
accuracy			0.96	25976
macro avg	0.96	0.96	0.96	25976
weighted avg	0.96	0.96	0.96	25976

# 특성중요도



# 결론

고객의 만족도에 큰 영향을 끼치는 특성은 다음과 같다

- ✓ Class of travel
- ✓ Arrival / Departure delays
- ✓ Services provided in the flight

비행 지연시간과 좌석클래스도 기내서비스 못지 않게 중요하다는 결론이 나왔다

따라서 항공사 측은 비행 지연시간 관리에 더 힘써야 하며, 낮은 클래스의 승객을 대상으로도 양질의 서비스를 제공해야 한다.