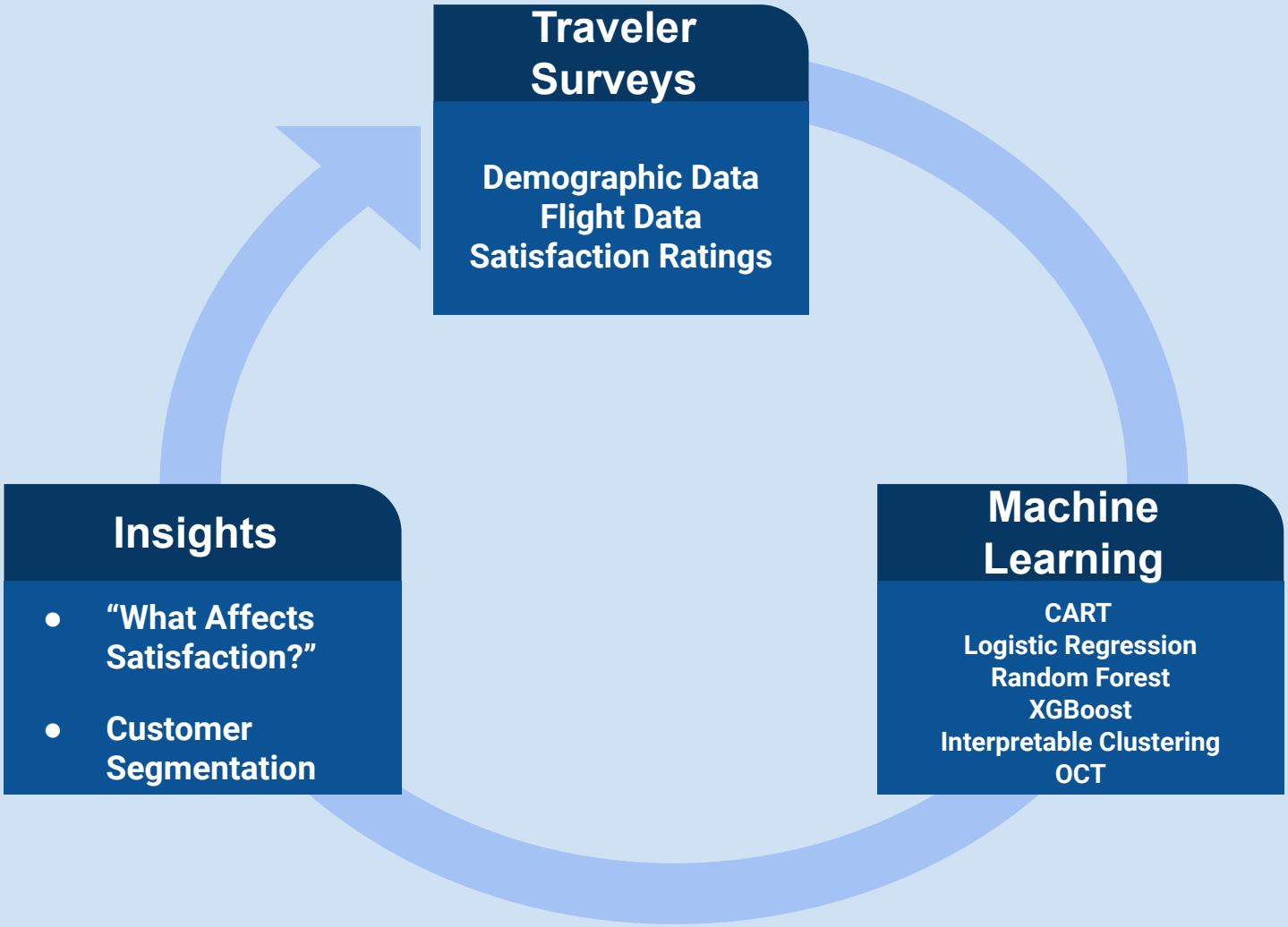
What Makes Flying Enjoyable?

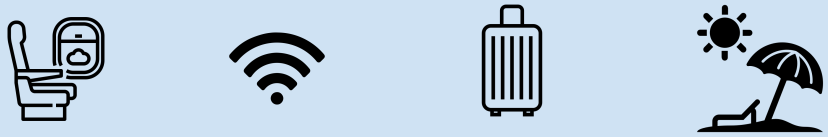
Using ML to Predict Passenger Satisfaction

Estella Dentinger and Brittany Nguyen
15.095 Fall 2022

Problem: “Analytics in the Air”



Data: Passenger Satisfaction Surveys



	F_1	F_2	F_3	...	F_m
P_1	1	2	5		1
P_2	5	1	3		0
P_3	4	2	4		1
...					
P_n	4	4	2		1

m = 23

Overall Satisfaction

Overall Satisfaction
0
1
0
1

n = 129,487

Features: Satisfaction Scores (0-5), Flight Info, Passenger Demographics

Our Approach: Binary Classification Methods

	In-Sample AUC	Out-of-Sample AUC	In-Sample Accuracy	Out-of-Sample Accuracy
CART	0.968	0.968	0.905	0.905
Logistic Regression	0.926	0.929	0.874	0.878
Random Forest	0.973	0.974	0.915	0.916
XGBoost	**0.997**	**0.997**	**0.970**	**0.971**
OCT	0.976	0.977	0.926	0.927

Key Insights: Top 5 Most Important Variables

1

Online Boarding
(0-5 satisfaction)

2

Inflight Wifi Service
(0-5 satisfaction)

3

Personal Traveler
(1=personal, 0=business)

4

Loyal Customer
(1=loyal, 0=disloyal)

5

Business Class
(1=business, 0=other)

More Insights: Some Passenger Archetypes

Business traveler who is overall satisfied when they can do their work and have entertainment on their flight

Despite good inflight entertainment, departure/arrival time was inconvenient and had poor overall ratings

Good ease of online booking and baggage handling, but flight itself was uncomfortable and unenjoyable

Why We Care

Cost Effectiveness

Customer Segmentation & Focus Groups

- **Interpretable Clusters + OCTs**

Interpretability

Best Performing: XGBoost, Random Forest & OCT

- **Lower performance at cost of higher understanding**

Scale of Impact

~$785 billion dollar industry

~6 million flyers/day

> 5K airlines globally