

# Smart HIR

**Find Talent More Efficiently**

Katie Huang  
February 10, 2021

# Mission

## Hire qualified Data Scientists

# Online Training Courses



# Online Training Courses



>20,000 enrollees



# Online Training Courses

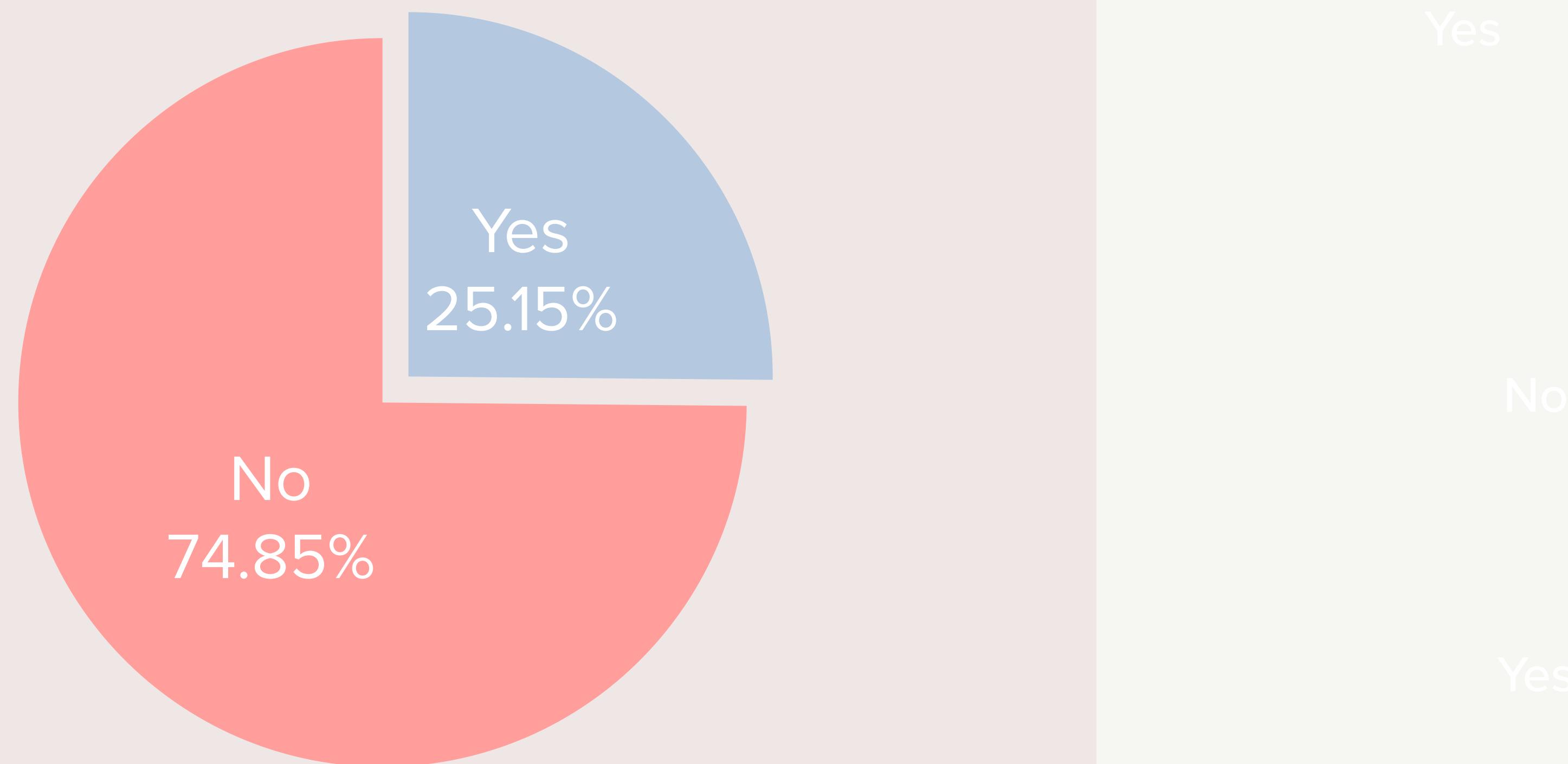


**>20,000 enrollees**

**Who would actually join  
your company?**

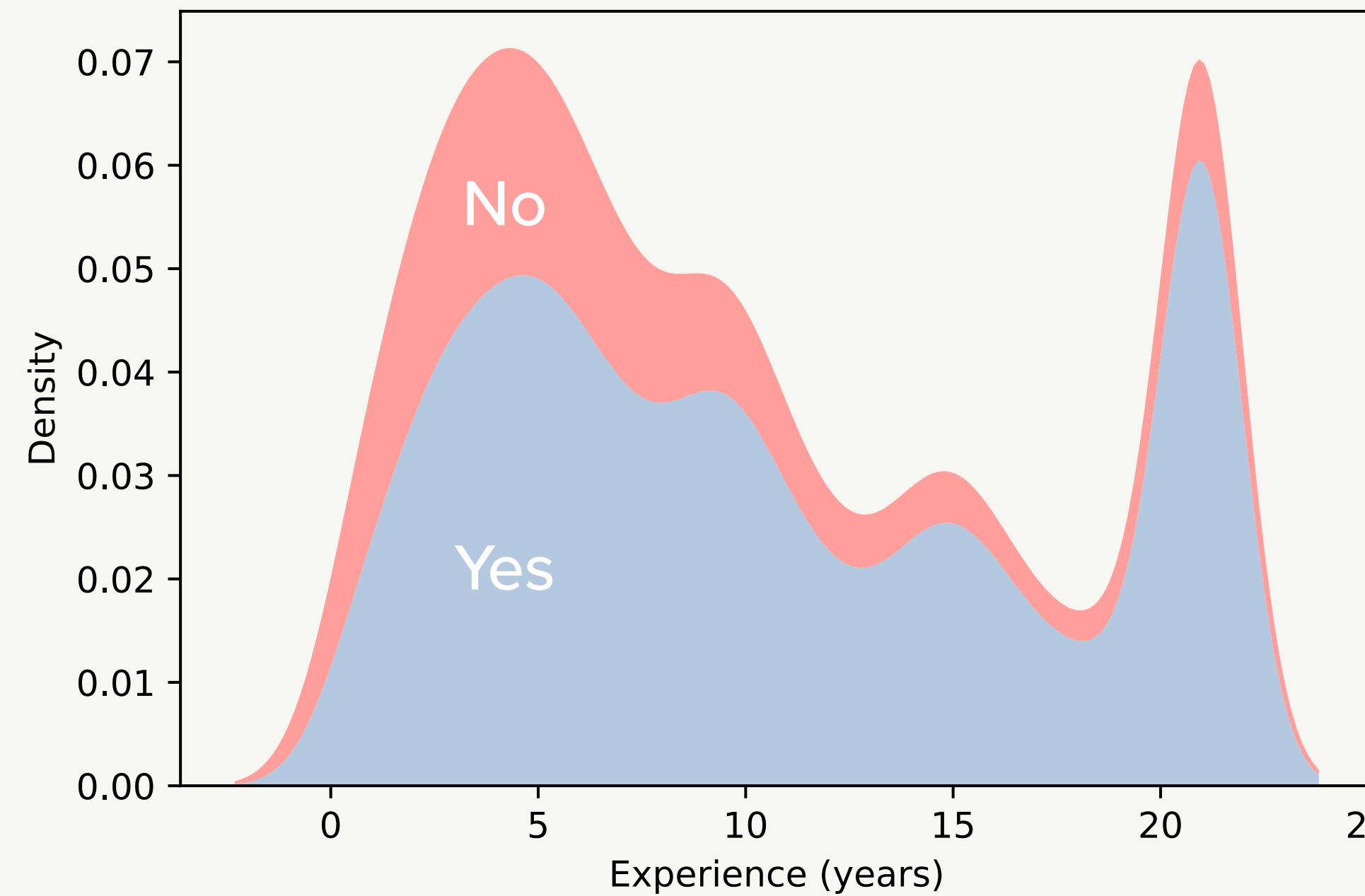
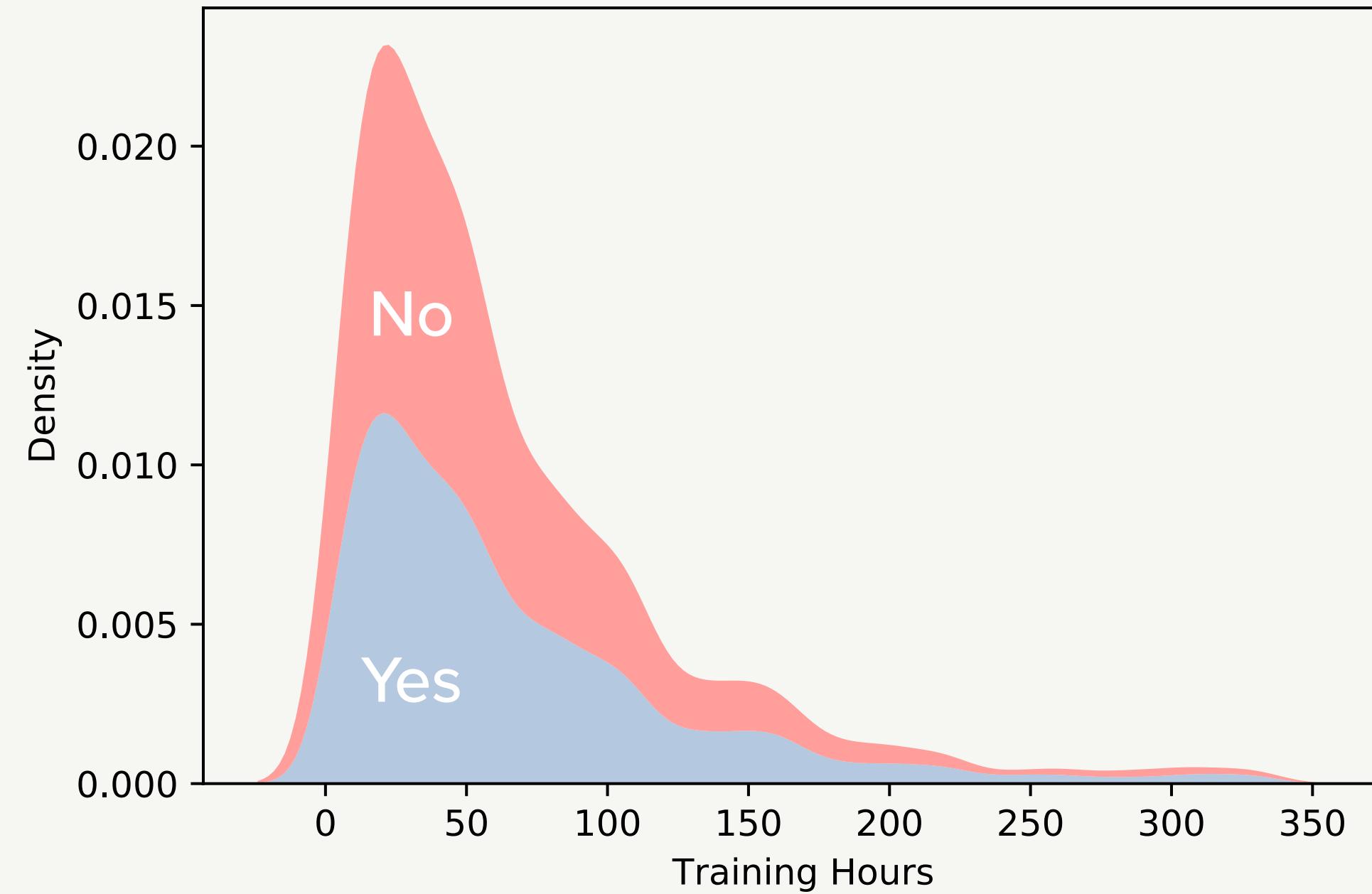
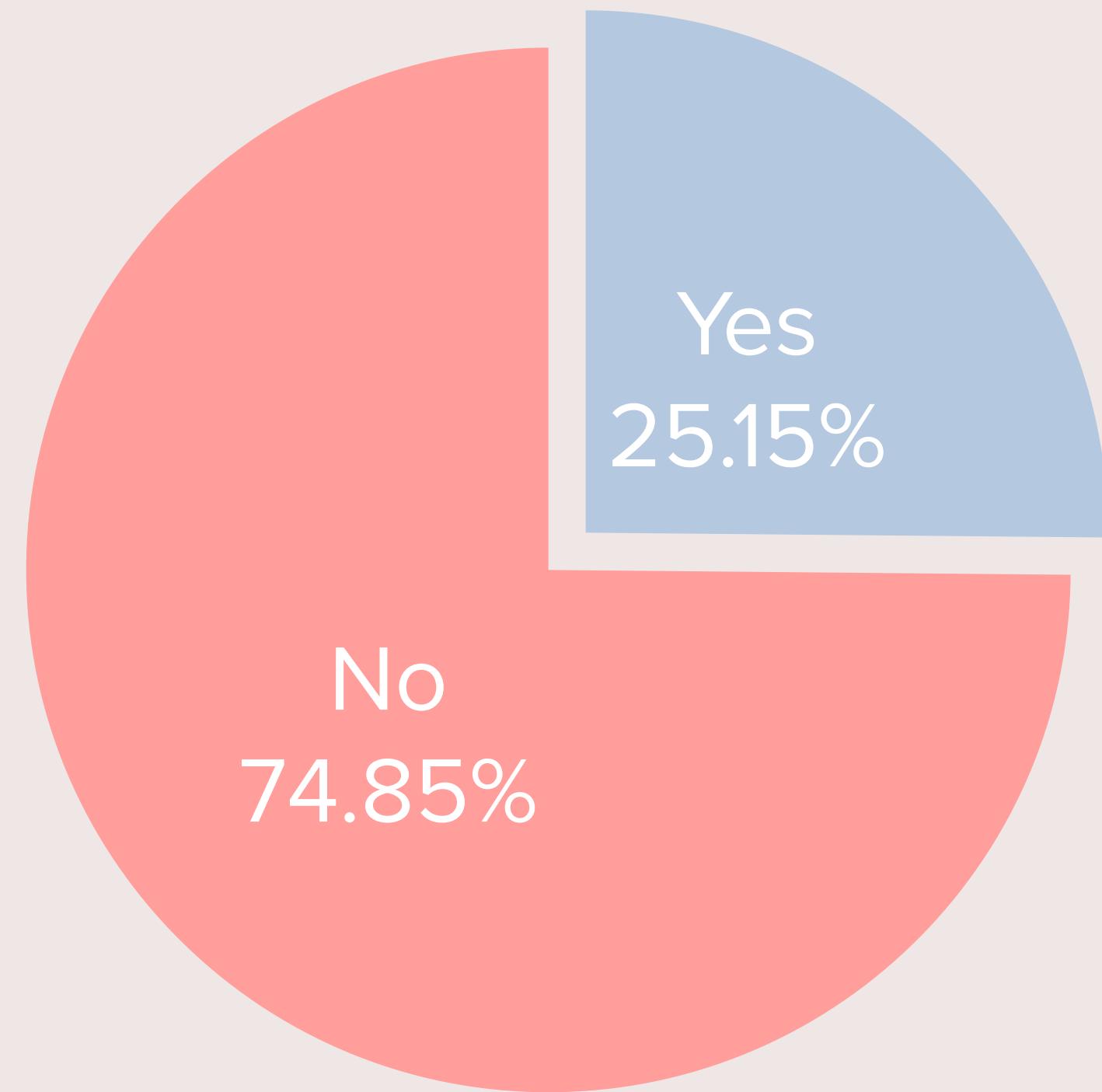
# Examine the enrollees

- Looking for a new job?



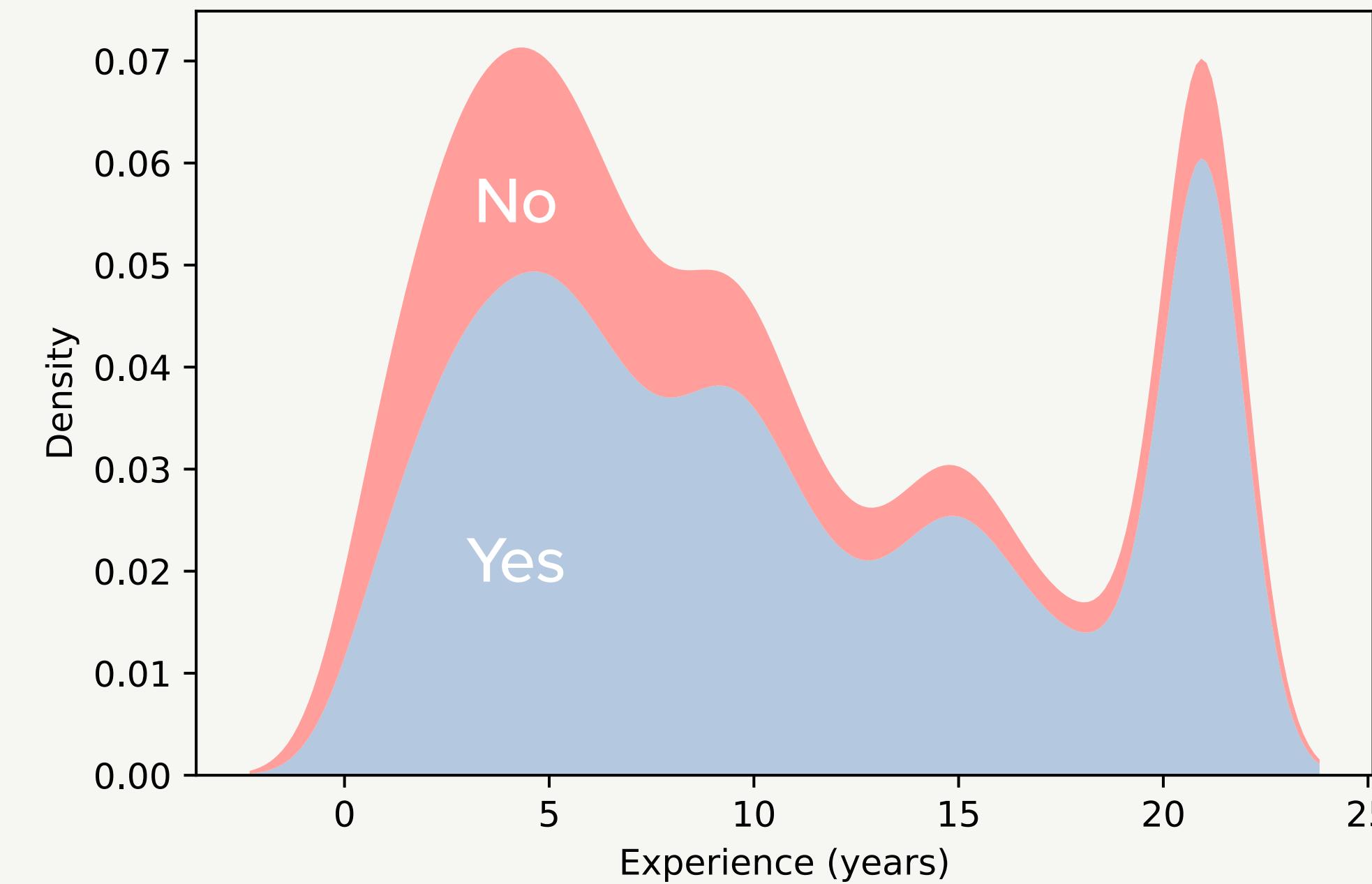
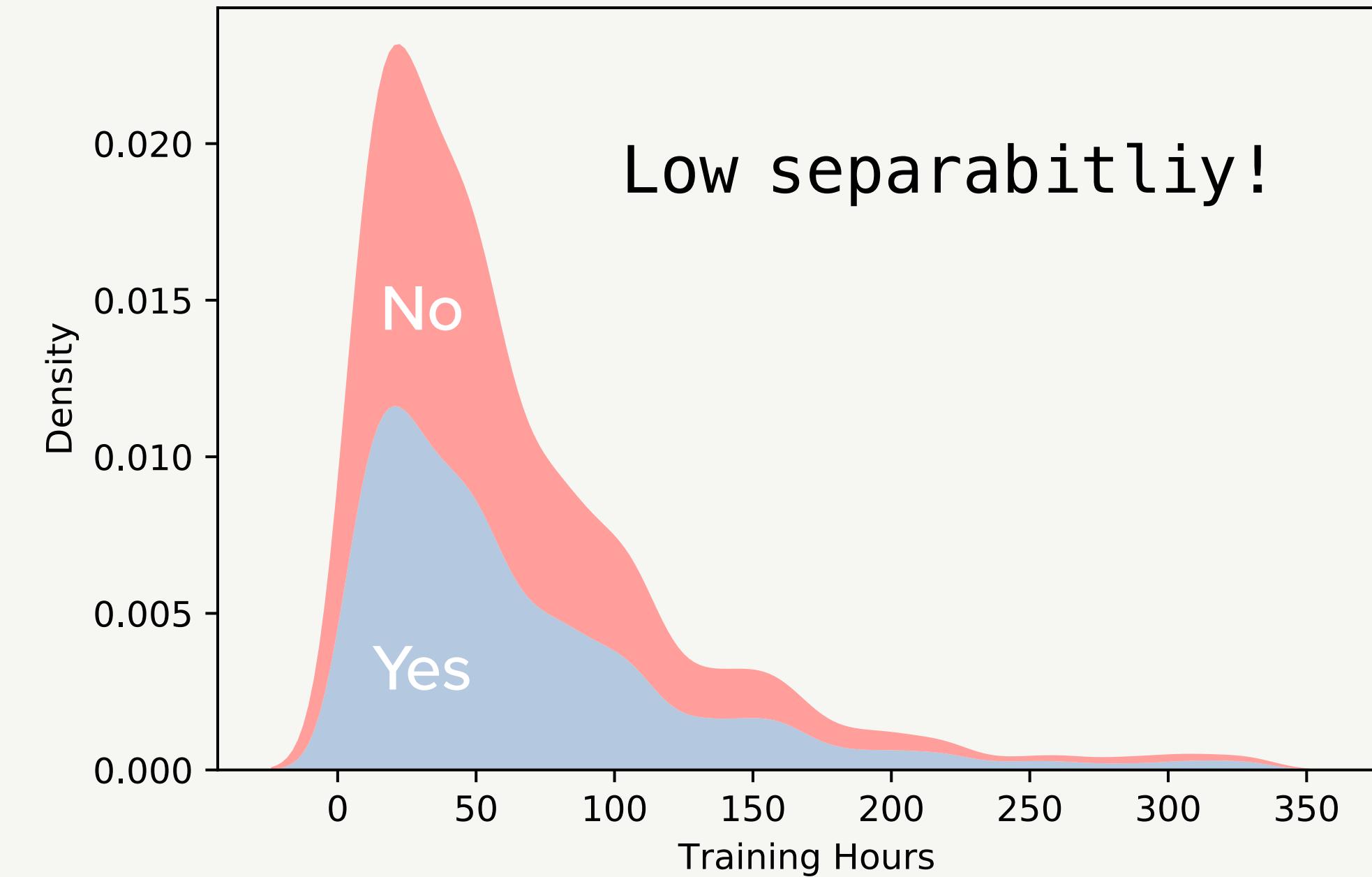
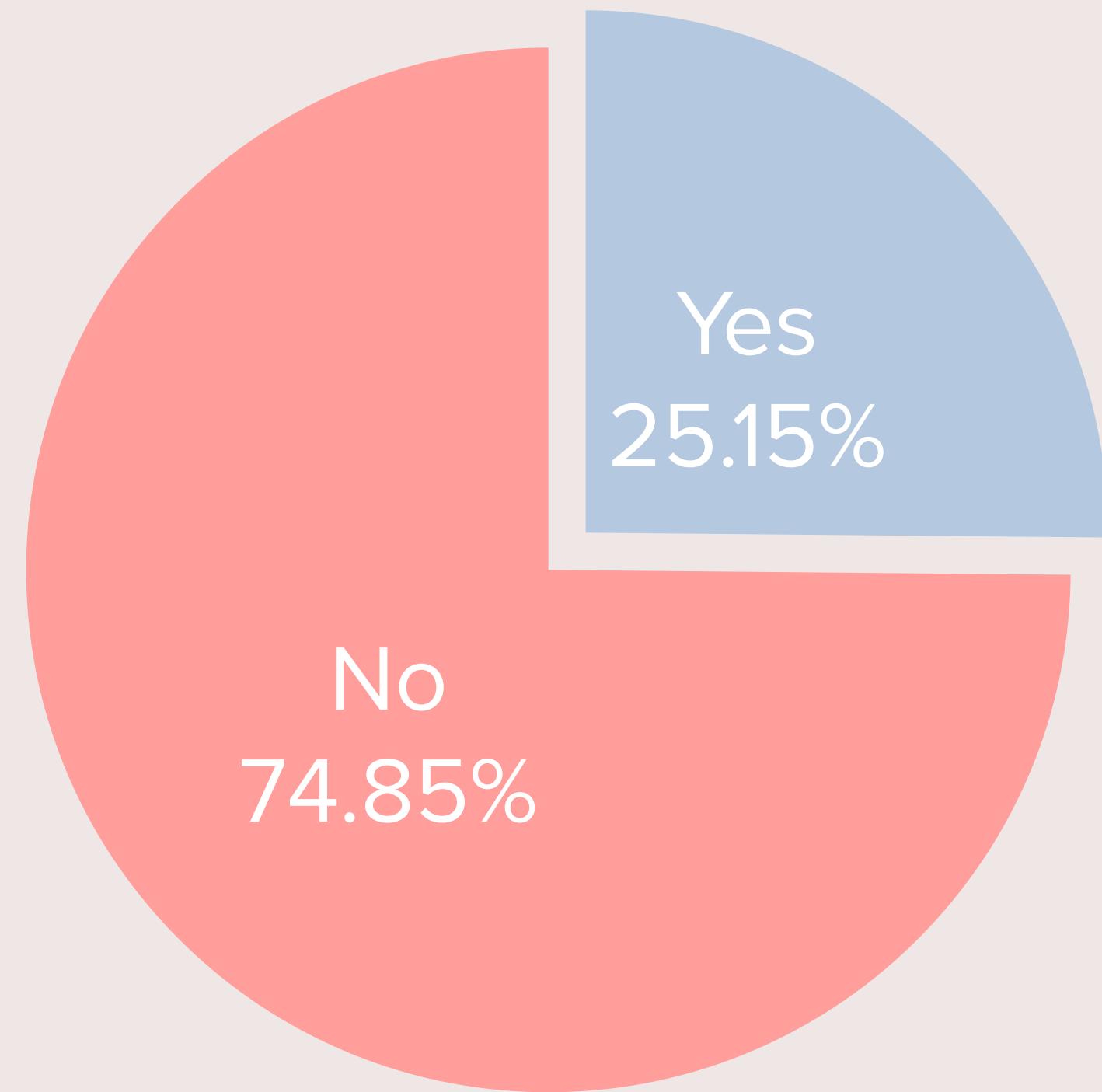
# Examine the enrollees

- Looking for a new job?



# Examine the enrollees

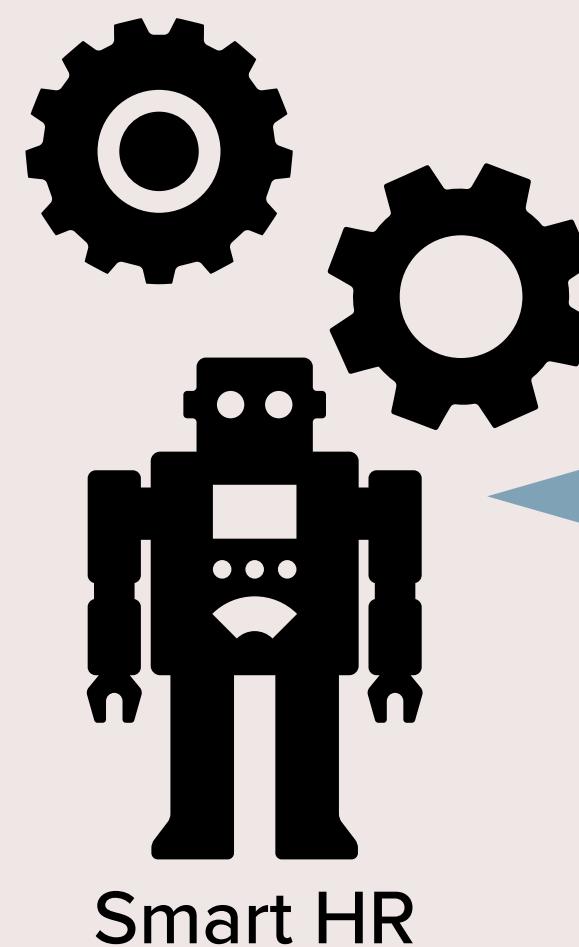
- Looking for a new job?



# Collect Data, Learn & Predict

## Candidate Info

- ◆ Gender
- ◆ Major
- ◆ Education level
- ◆ Current company
- ◆ Current city
- ◆ Last new job
- ◆ Relevant experience
- ◆ Experience
- ◆ Enrolled courses
- ◆ Training hours



**Looking for a new job?**

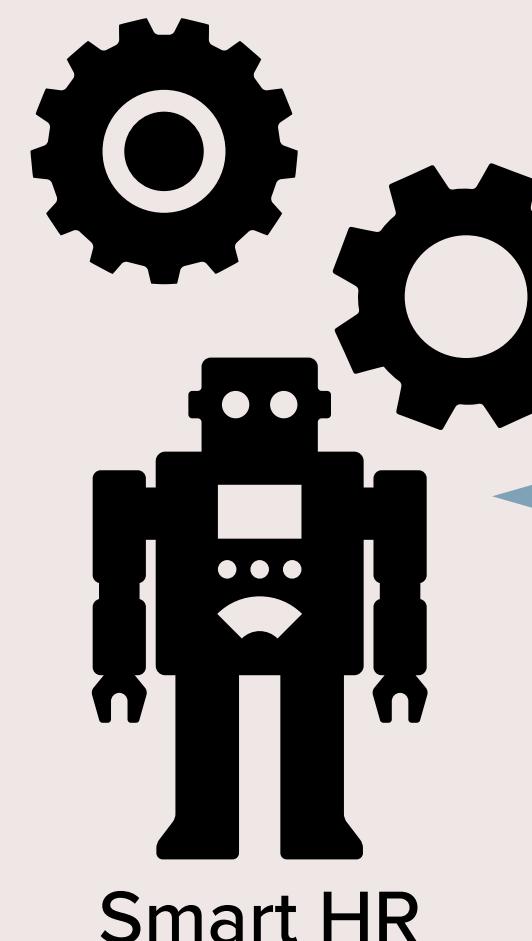
**No**

**Yes!**  
("switcher")

# Collect Data, Learn & Predict

## Candidate Info

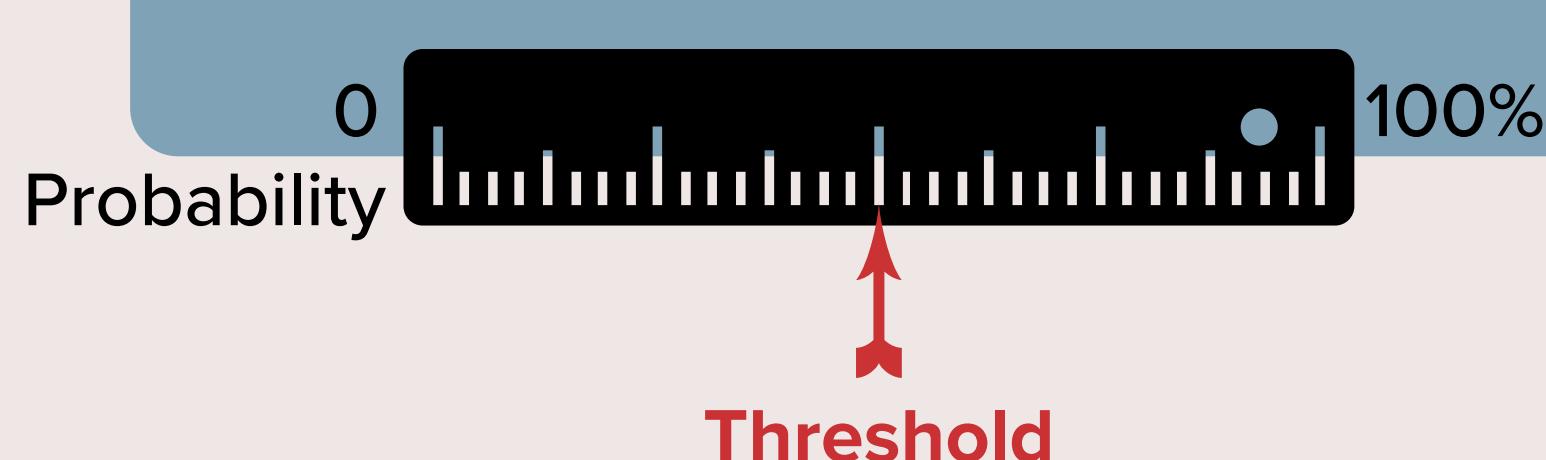
- ◆ Gender
- ◆ Major
- ◆ Education level
- ◆ Current company
- ◆ Current city
- ◆ Last new job
- ◆ Relevant experience
- ◆ Experience
- ◆ Enrolled courses
- ◆ Training hours



**Looking for a new job?**

No

**Yes!**  
("switcher")

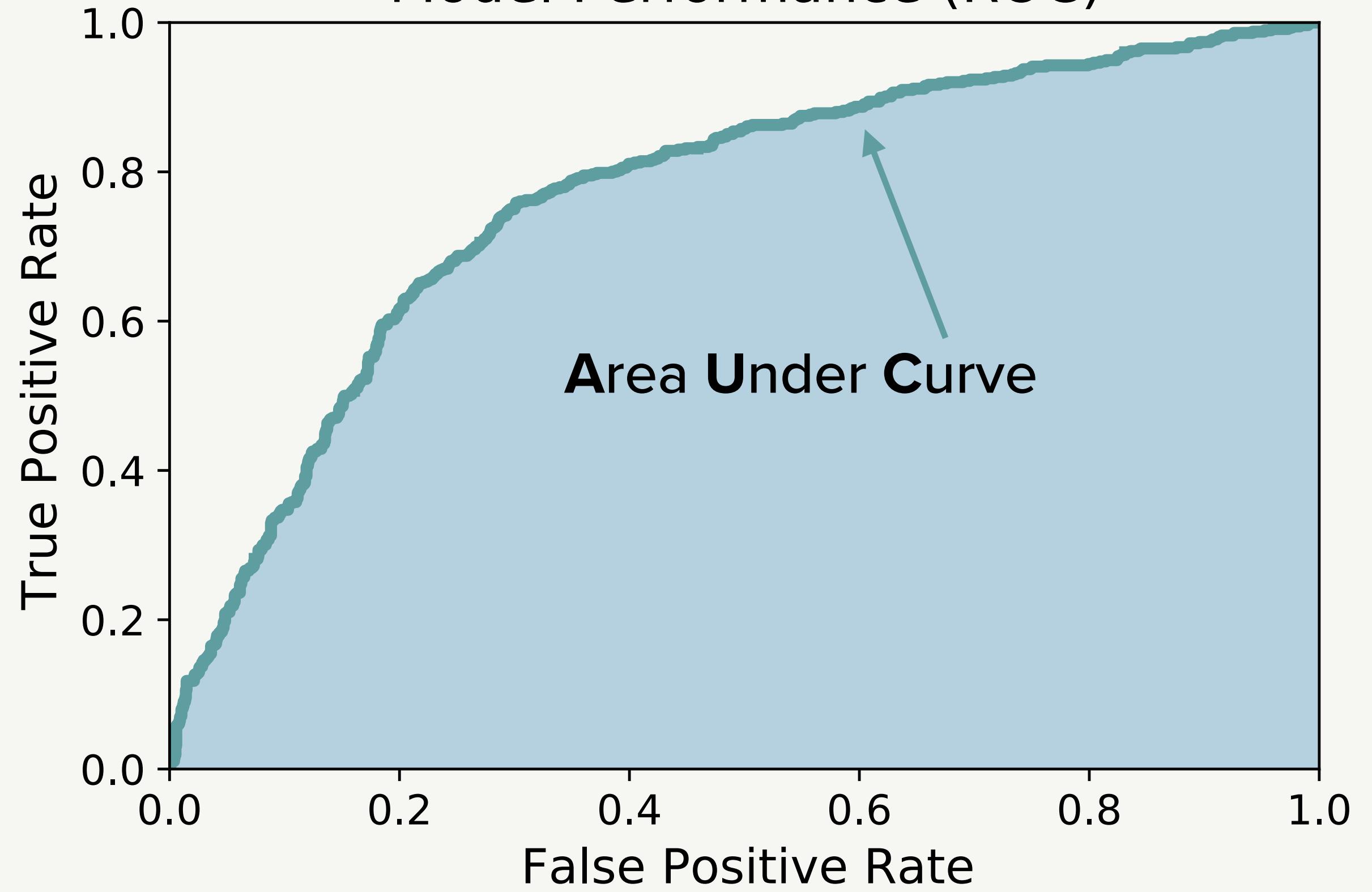


# Model

Logistic Regression

Overall performance: AUC score

Model Performance (ROC)



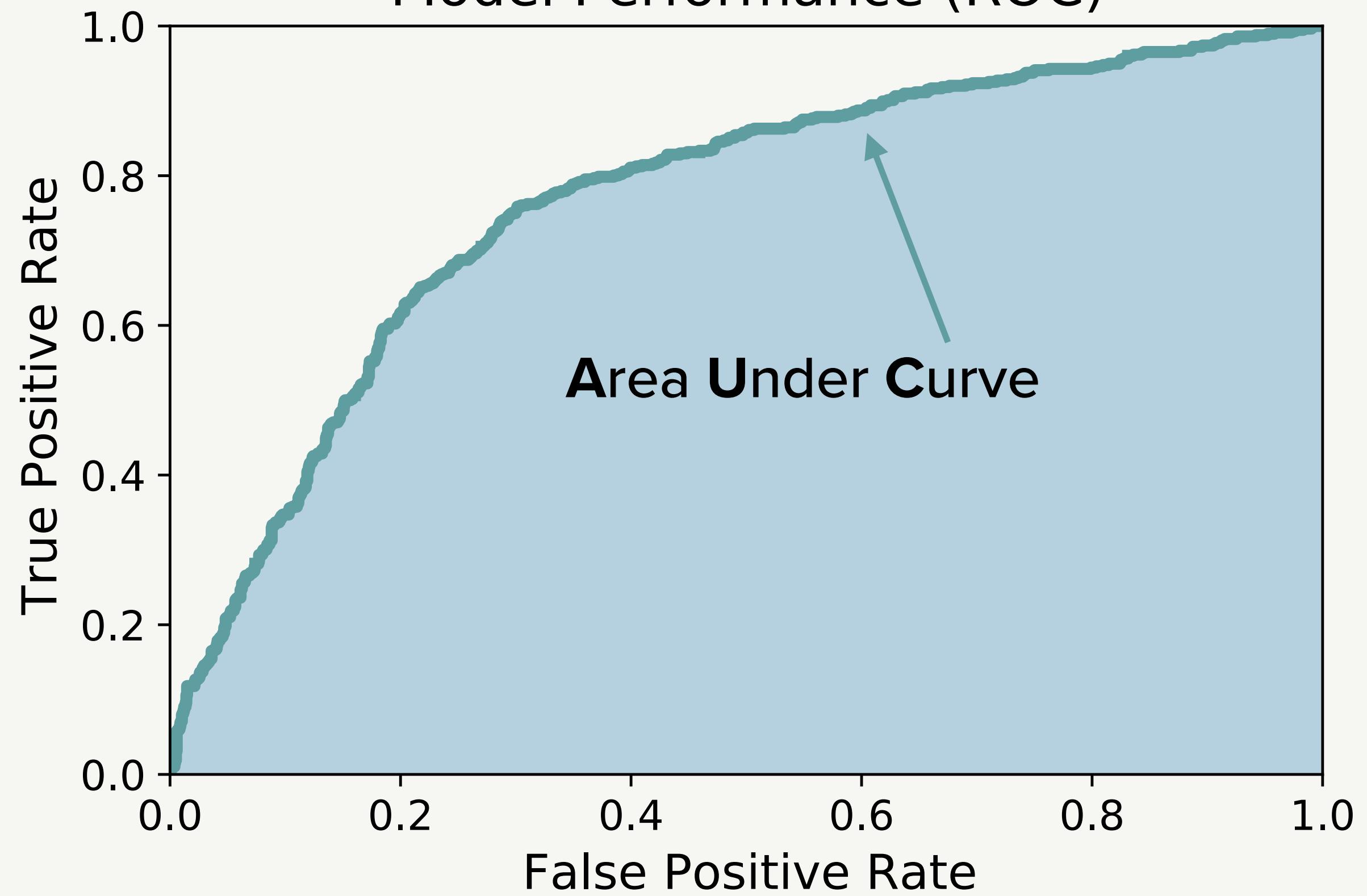
# Model

## Logistic Regression

**Overall performance: AUC score**

- Perfect model: 100%
- Random guess: 50%

Model Performance (ROC)



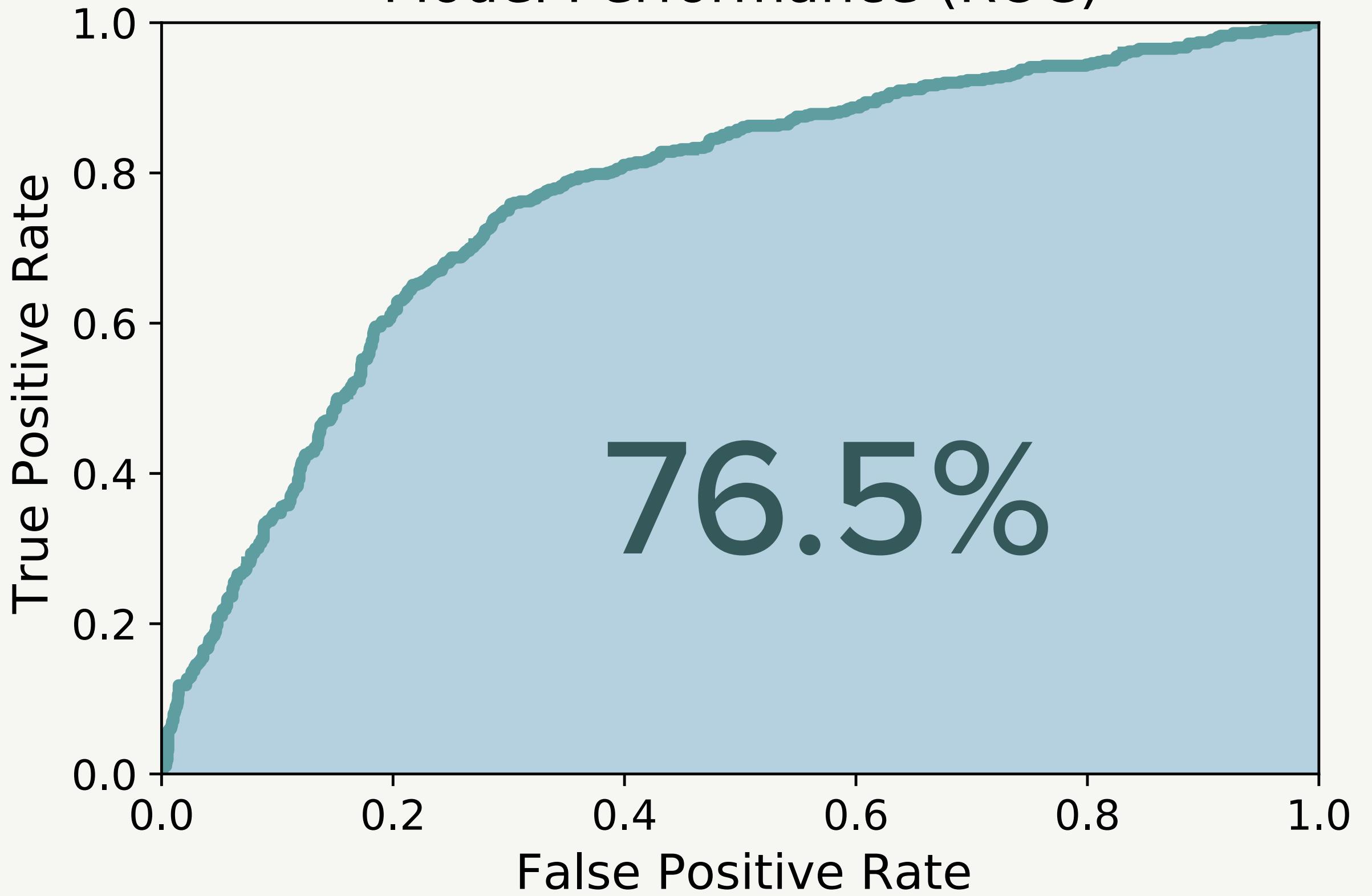
# Model

## Logistic Regression

**Overall performance: AUC score**

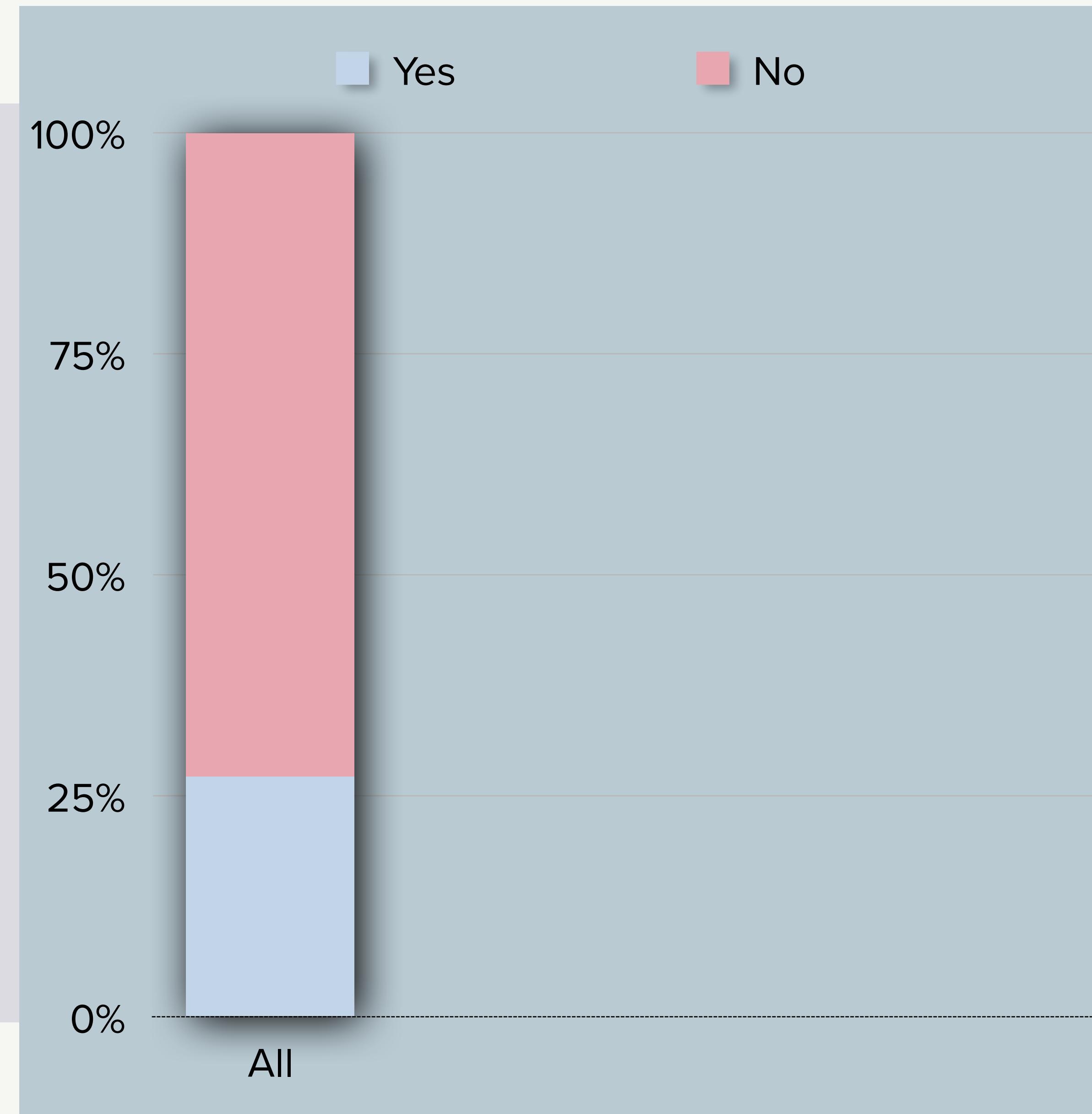
- Perfect model: 100%
- Random guess: 50%

Model Performance (ROC)



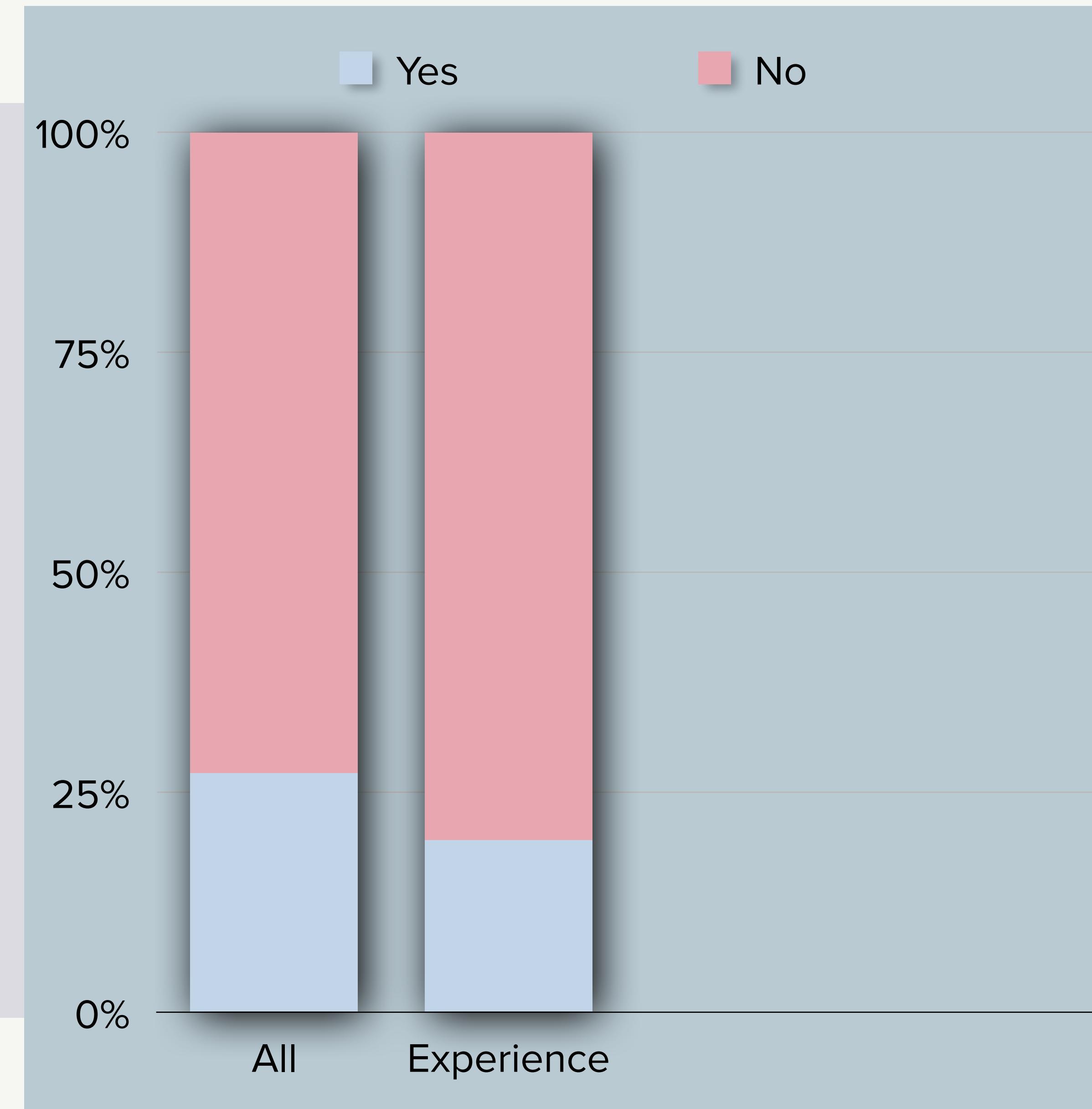
## Select candidates to interview

- All candidates



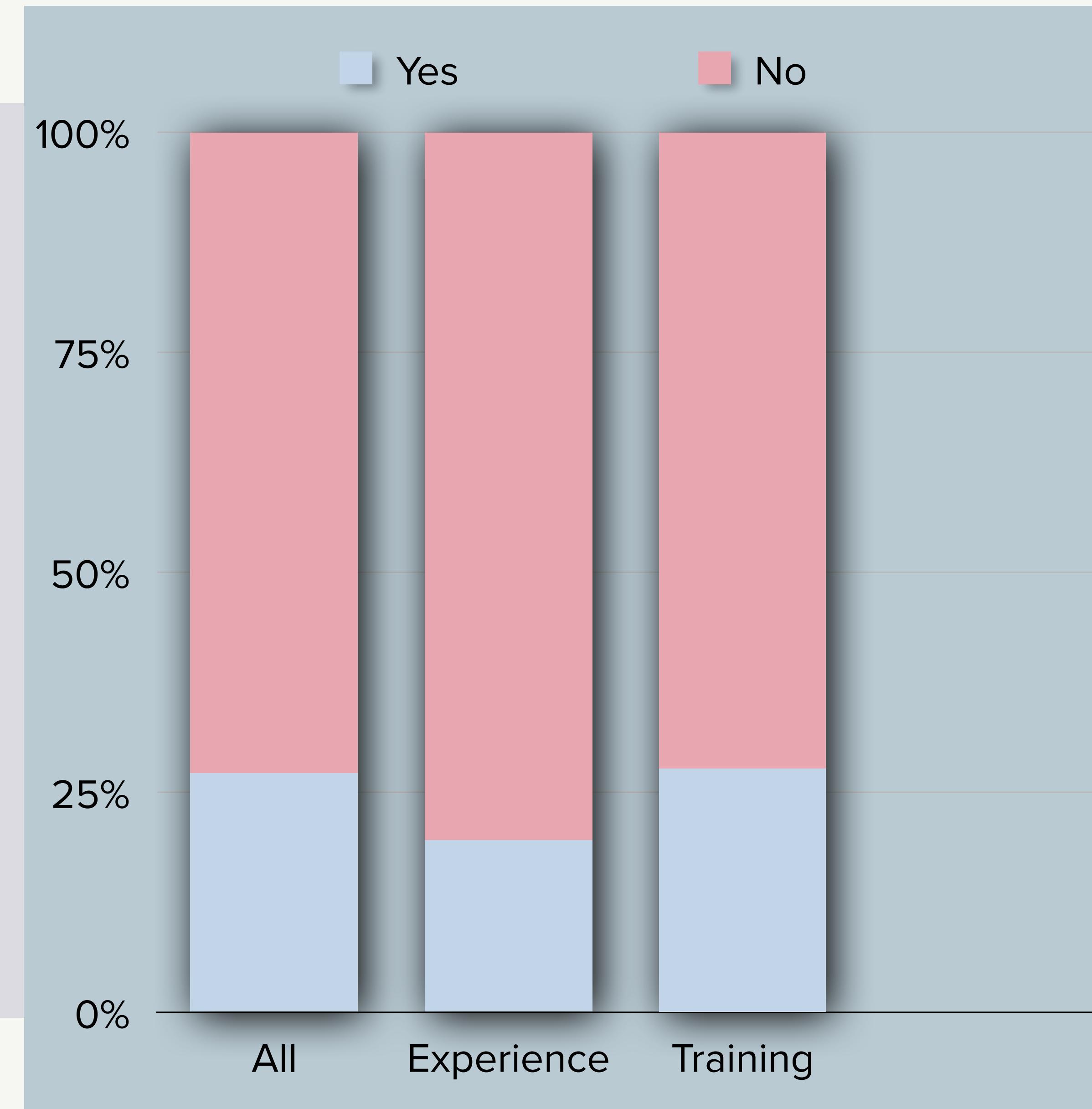
## Select candidates to interview

- All candidates
- By experience > 5 years



## Select candidates to interview

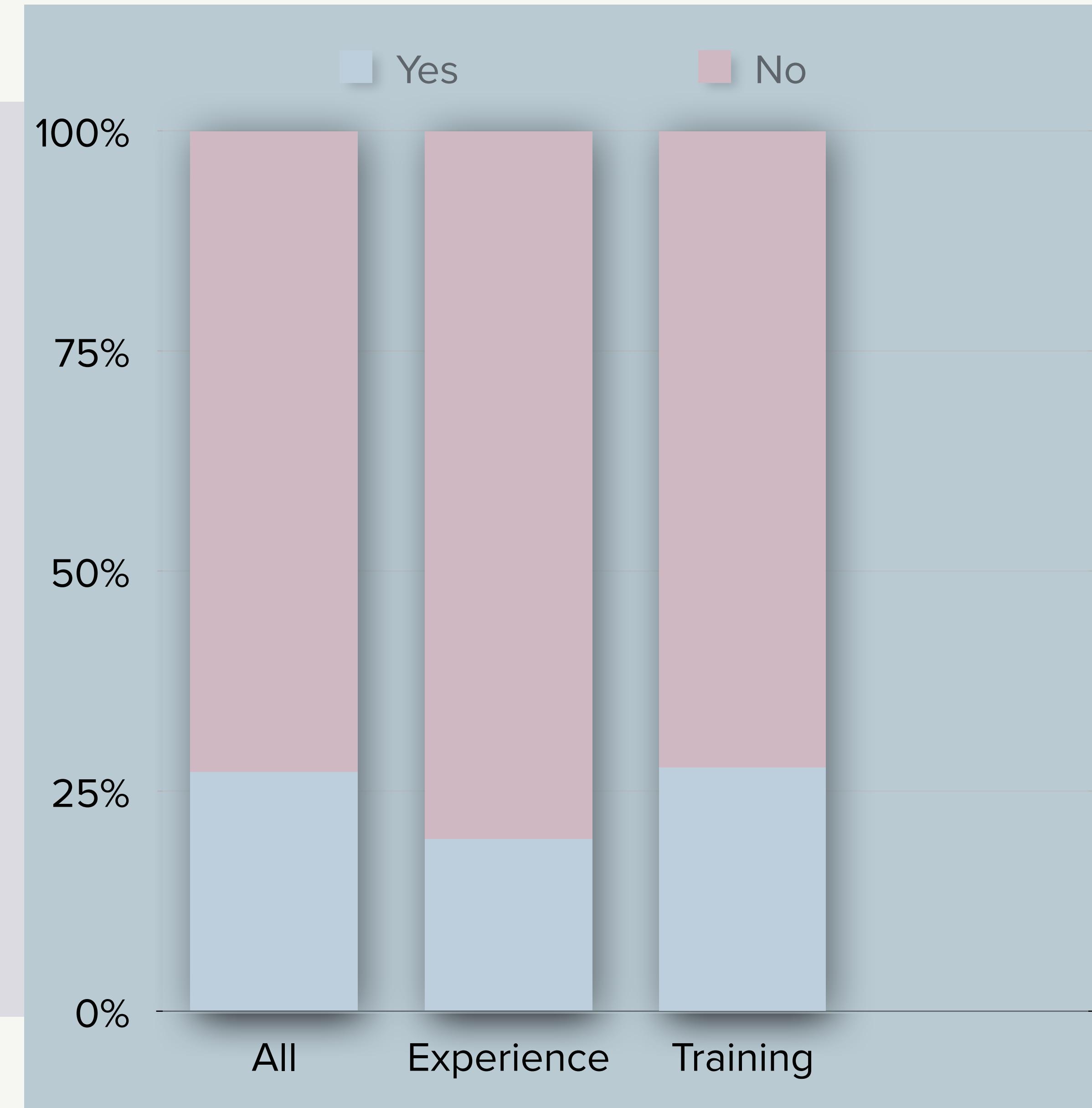
- All candidates
- By experience > 5 years
- By training hours > 100 hrs



## Select candidates to interview

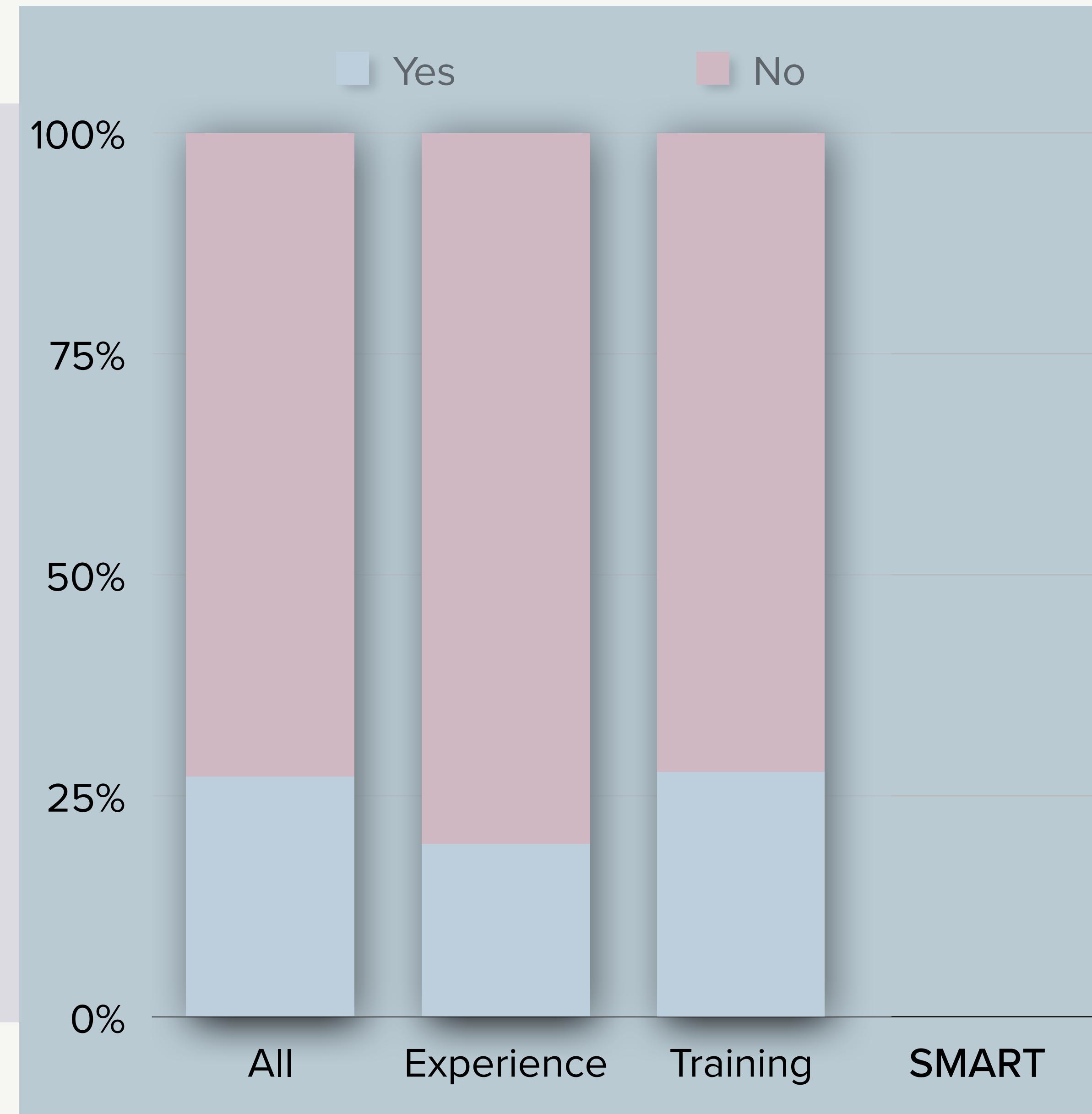
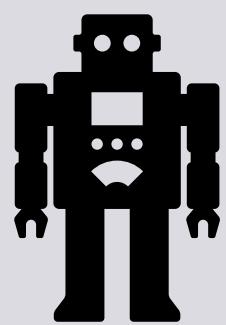
- All candidates
- By experience > 5 years
- By training hours > 100 hrs

Low yield!



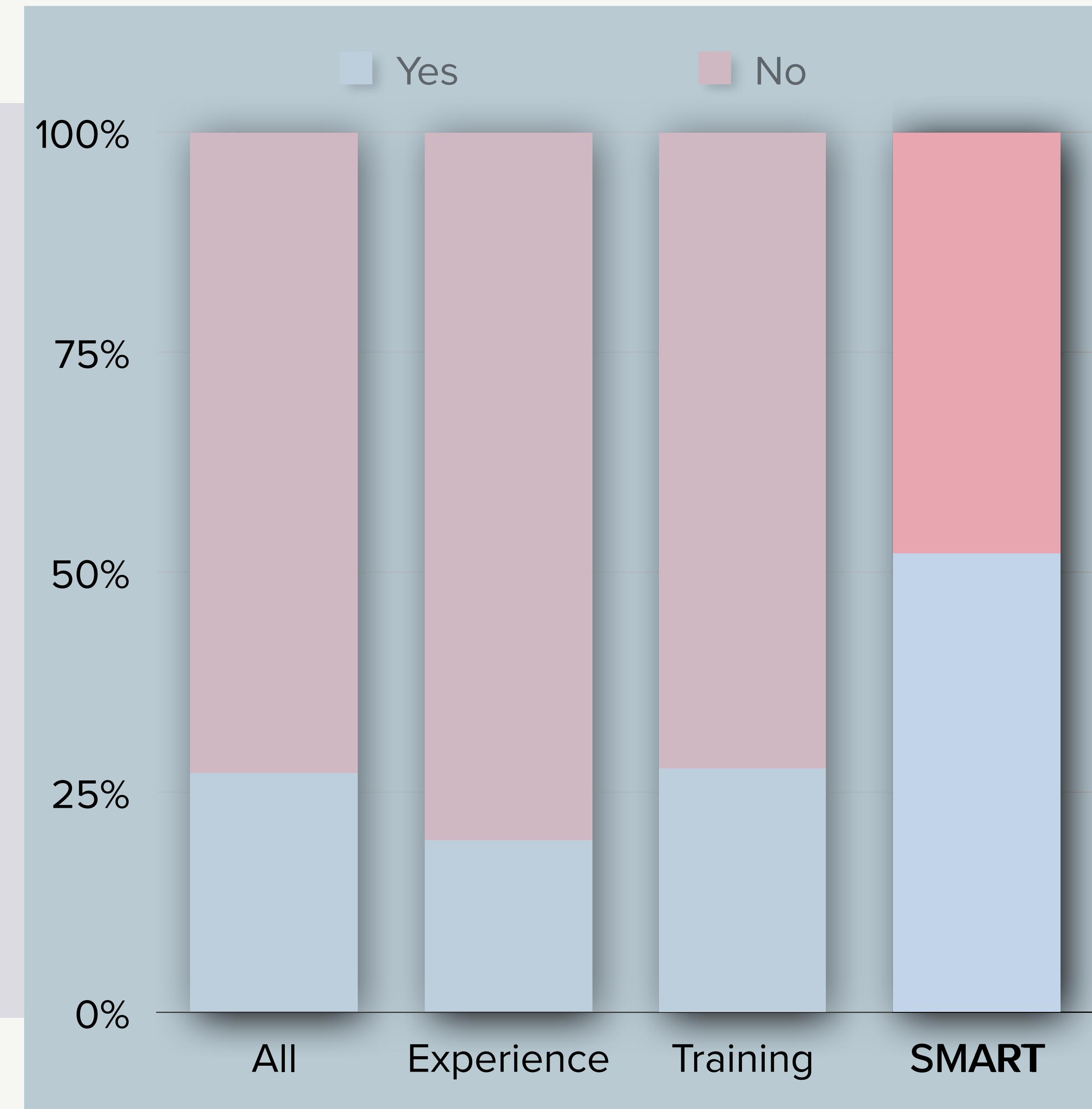
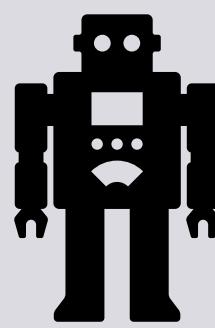
## Select candidates to interview

- All candidates
- By experience > 5 years
- By training hours > 100 hrs
- By smart prediction



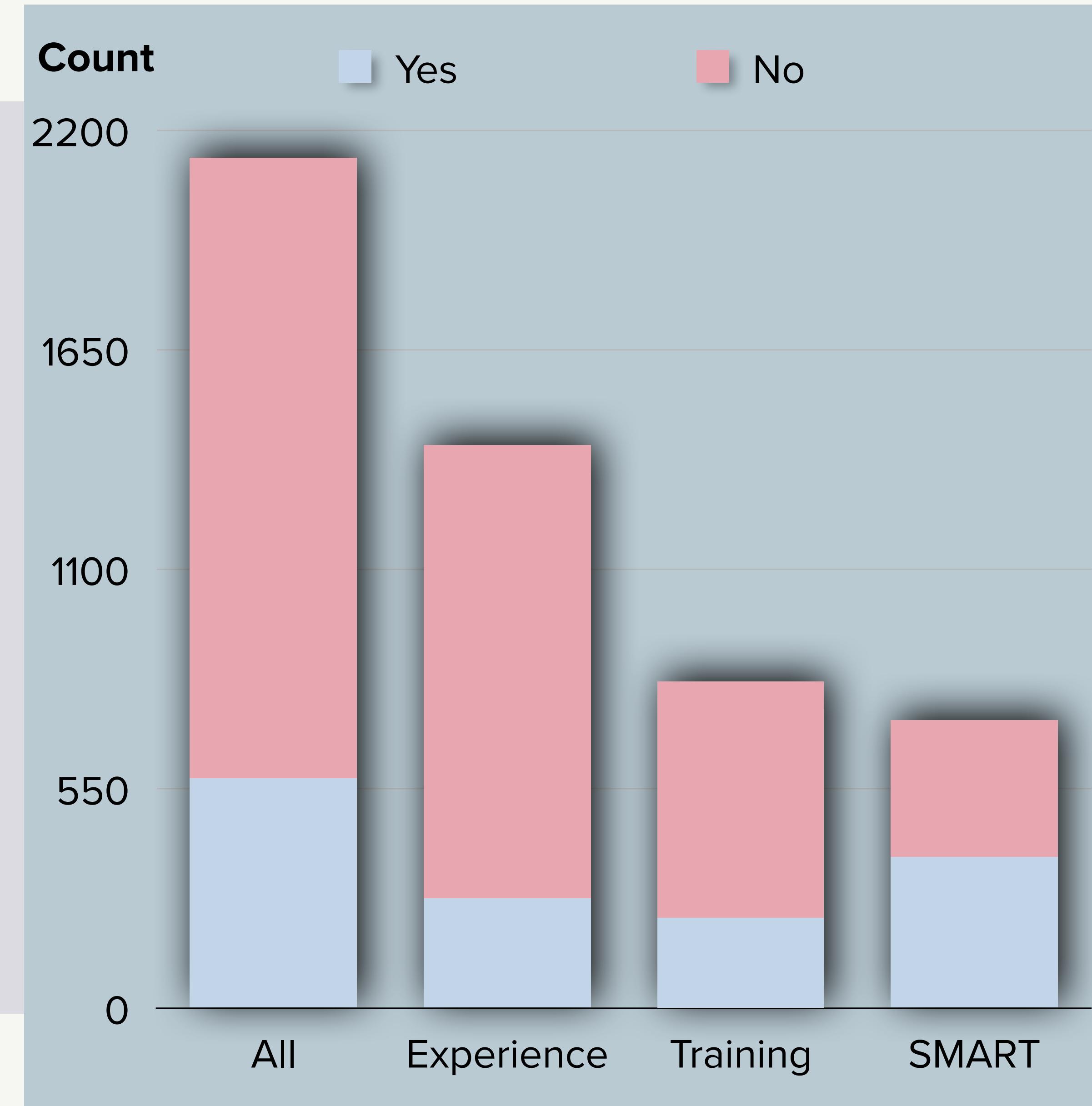
## Select candidates to interview

- All candidates
- By experience > 5 years
- By training hours > 100 hrs
- By smart prediction



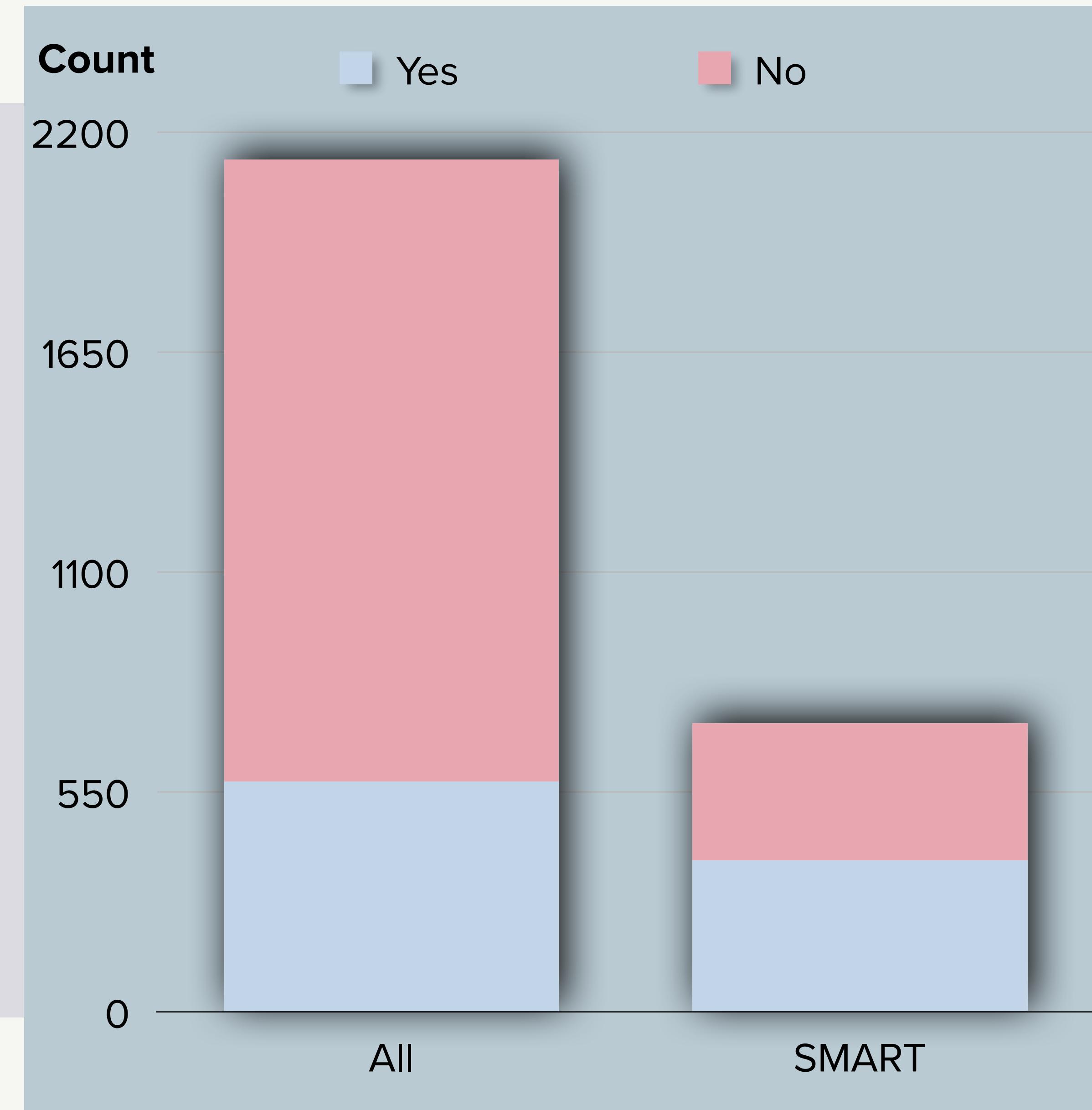
# ⌚ Time Saved

- All candidates
- By experience > 5 years
- By training hours > 100 hrs
- By smart prediction



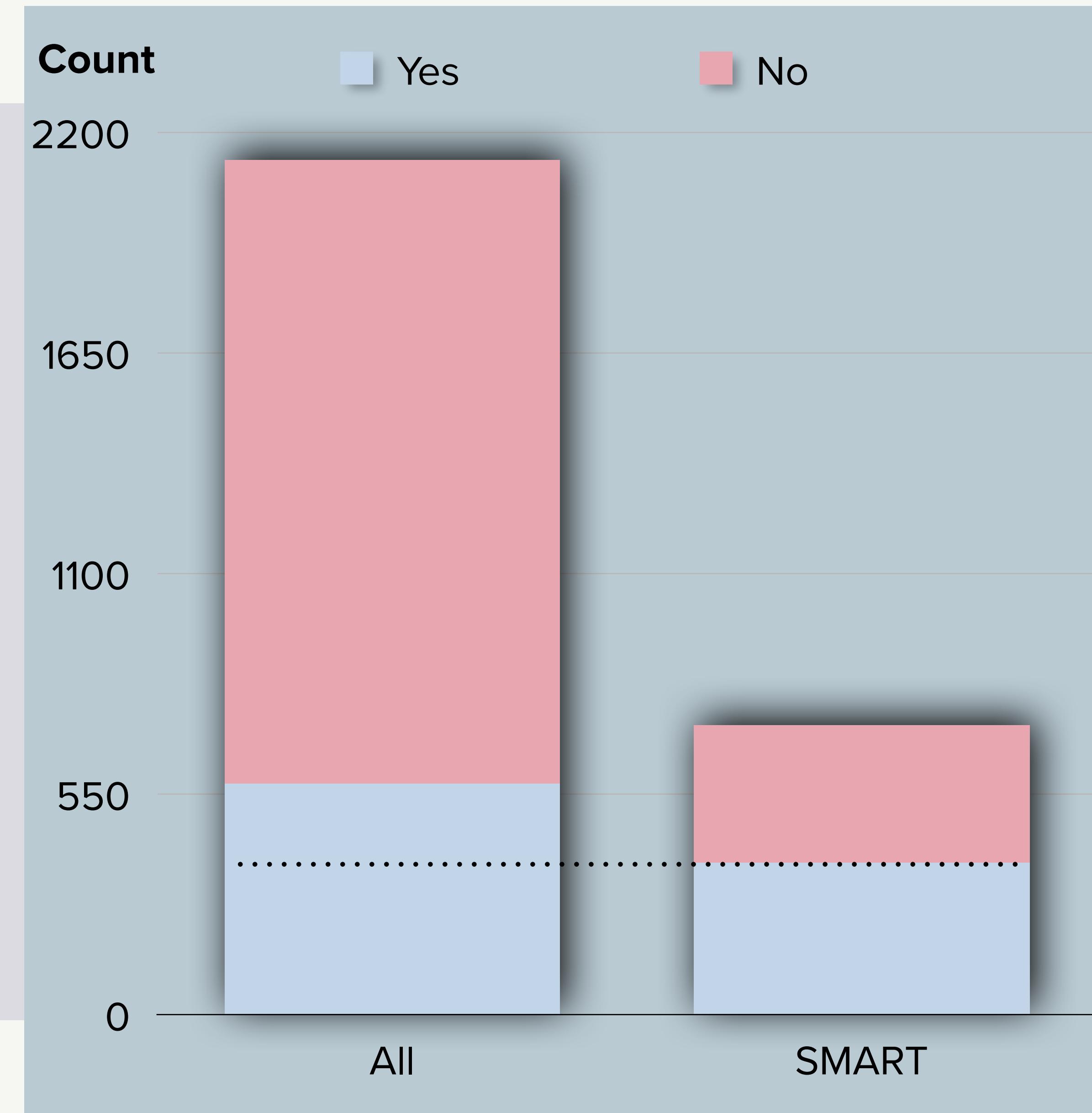
# Trade-off?

- Saved 67% time



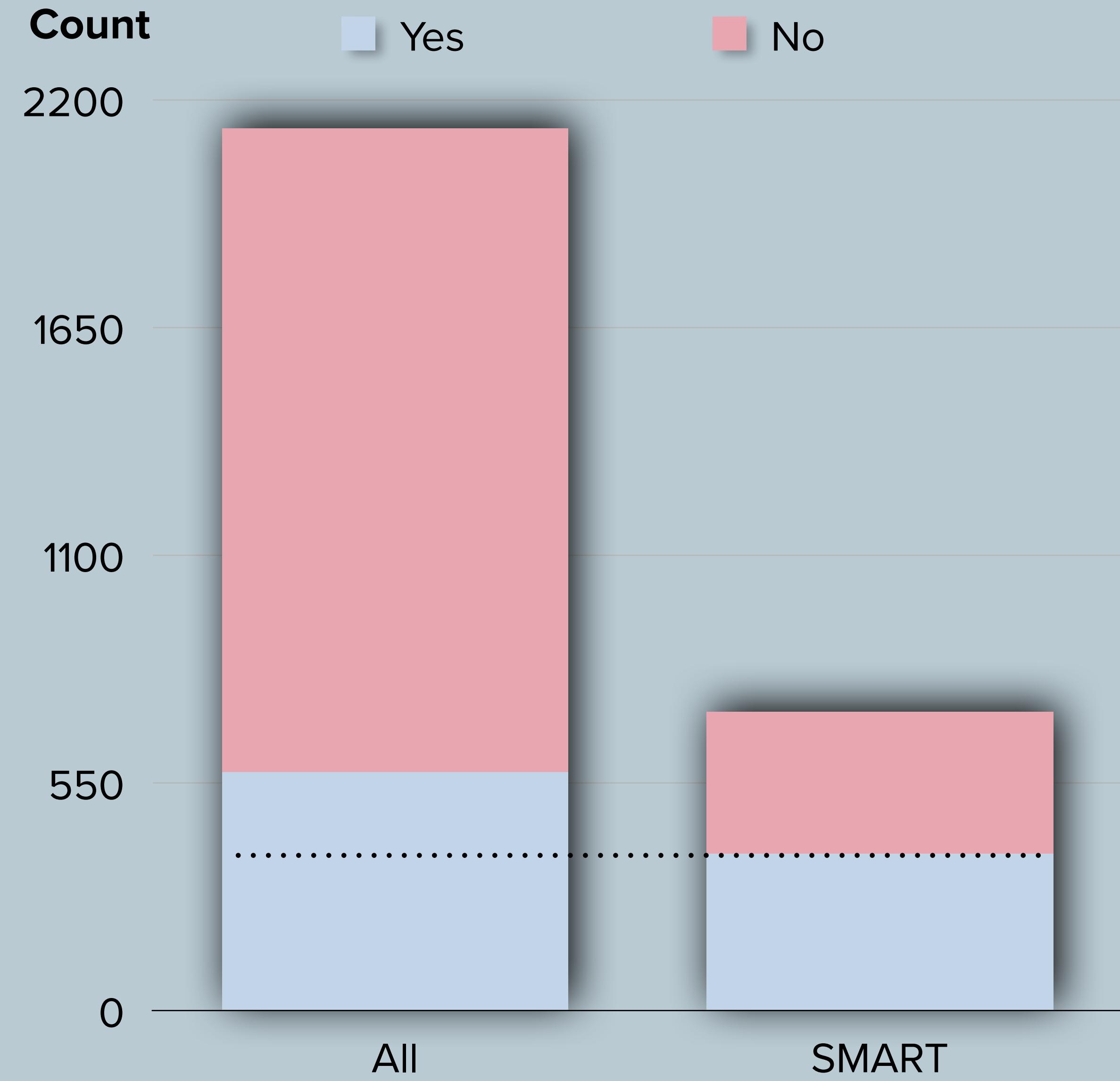
# Trade-off?

- Saved 67% time
- Recall: caught 65% switchers



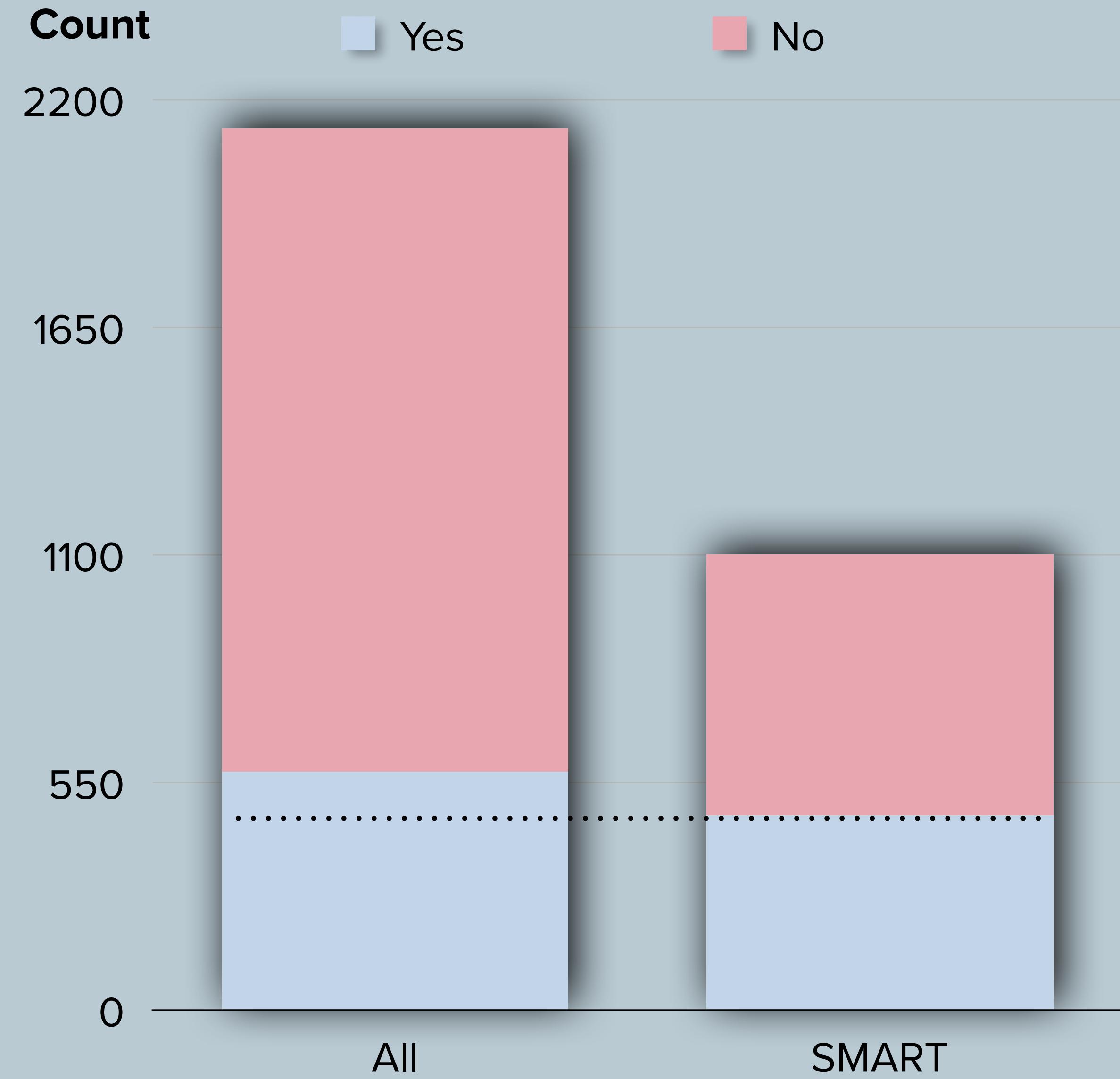
# 65%

Increase recall;



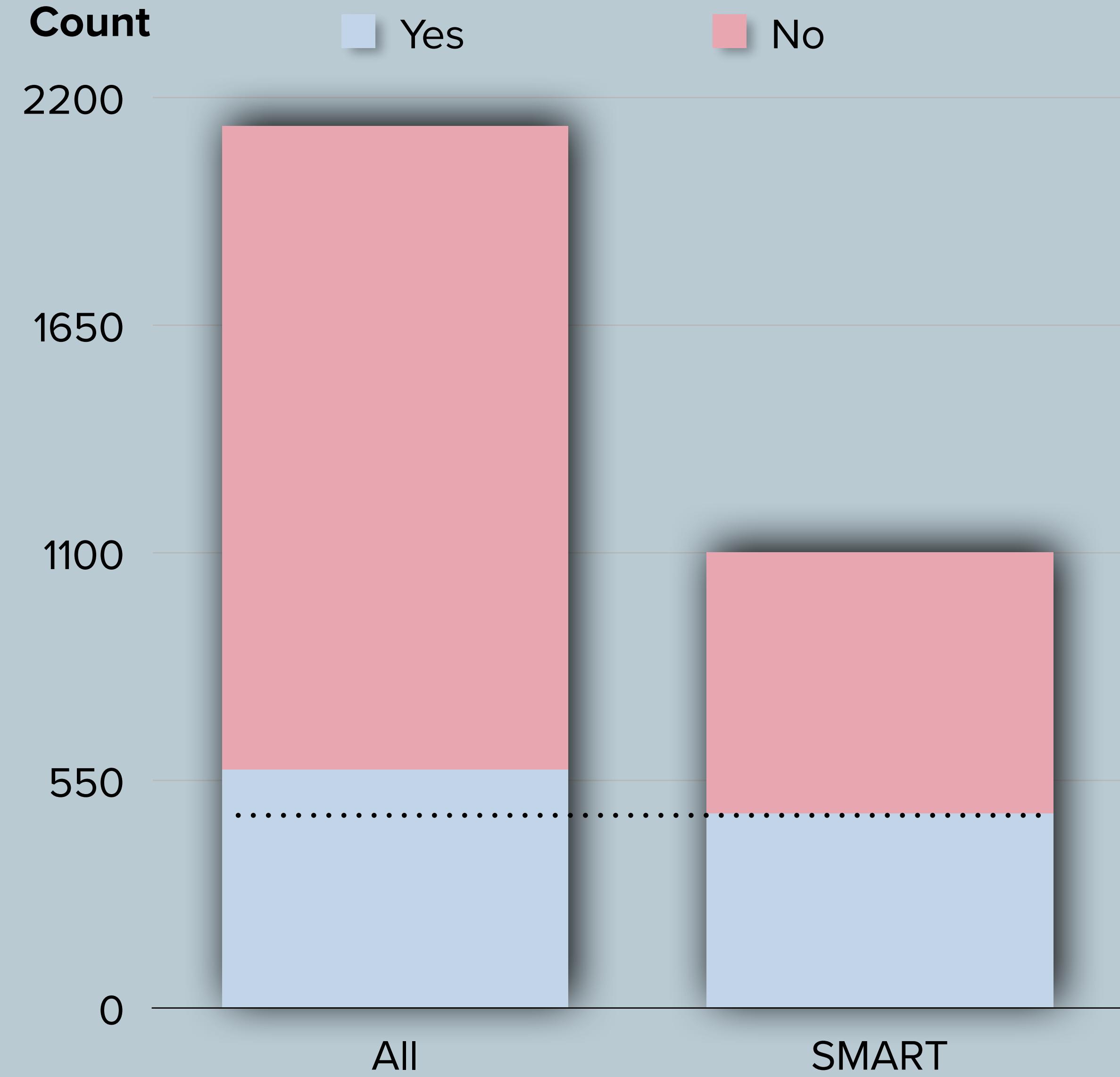
# 83%

Increase recall;



# 83%

**Increase recall; still save 46% time!**



# Customize



Web app for easy access

**Welcome to Smart HR!**

Tell me about the candidates and I will let you know if they are actually looking for a new job! This will save you considerable amount of time reaching out or interviewing the candidates!

**Current company**

- Early Stage Startup
- Funded Startup
- NGO
- Other
- Public Sector
- Pvt Ltd
- unknown

**Probability**

- 20
- 40
- 60
- 80

### Candidate Statistics

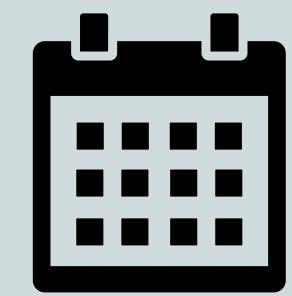
Qualified candidate count: 19158

Show dataframe

	Enrollee ID	Gender	Major	Education Level	Current City	Current Company
0	8949	Male	STEM	Graduate	0.9200	Public Sector
1	29725	Male	STEM	Graduate	0.7760	Other
2	11561	unknown	STEM	Graduate	0.6240	Public Sector
3	33241	unknown	Business	Graduate	0.7890	Other
4	666	Male	STEM	Masters	0.7670	Public Sector
5	21651	unknown	STEM	Graduate	0.7640	Public Sector
6	28806	Male	unknown	High School	0.9200	Other
7	402	Male	STEM	Graduate	0.7620	Public Sector
8	27107	Male	STEM	Graduate	0.9200	Public Sector
9	699	unknown	STEM	Graduate	0.9200	Public Sector
10	29452	unknown	unknown	High School	0.6240	Other

16,000  
14,000  
12,000  
26

# Prioritize



## Final list of candidates

Enrollee ID	Probability of Yes!
22527	93.6%
32795	92.8%
26694	92.5%
31758	92.4%
7626	92.1%
5391	92.1%
7266	92%
4377	92%
3432	91.7%
462	91.7%
...	...

# Outlook

## Smart HR

### Find talent

- Reduce cost and time for HR

# Outlook

## Smart HR

### Find talent

- Reduce cost and time for HR

### Retain talent

- Categorization of candidates
- Future planning of courses
- Factors of employee decision

# Important Features

Current Company type: Undisclosed

Current Company type: Public Sector

Major: Undisclosed

Current Company Type: Other

No Relevant Experience

Last New Job



# Important Features

Current Company type: Undisclosed

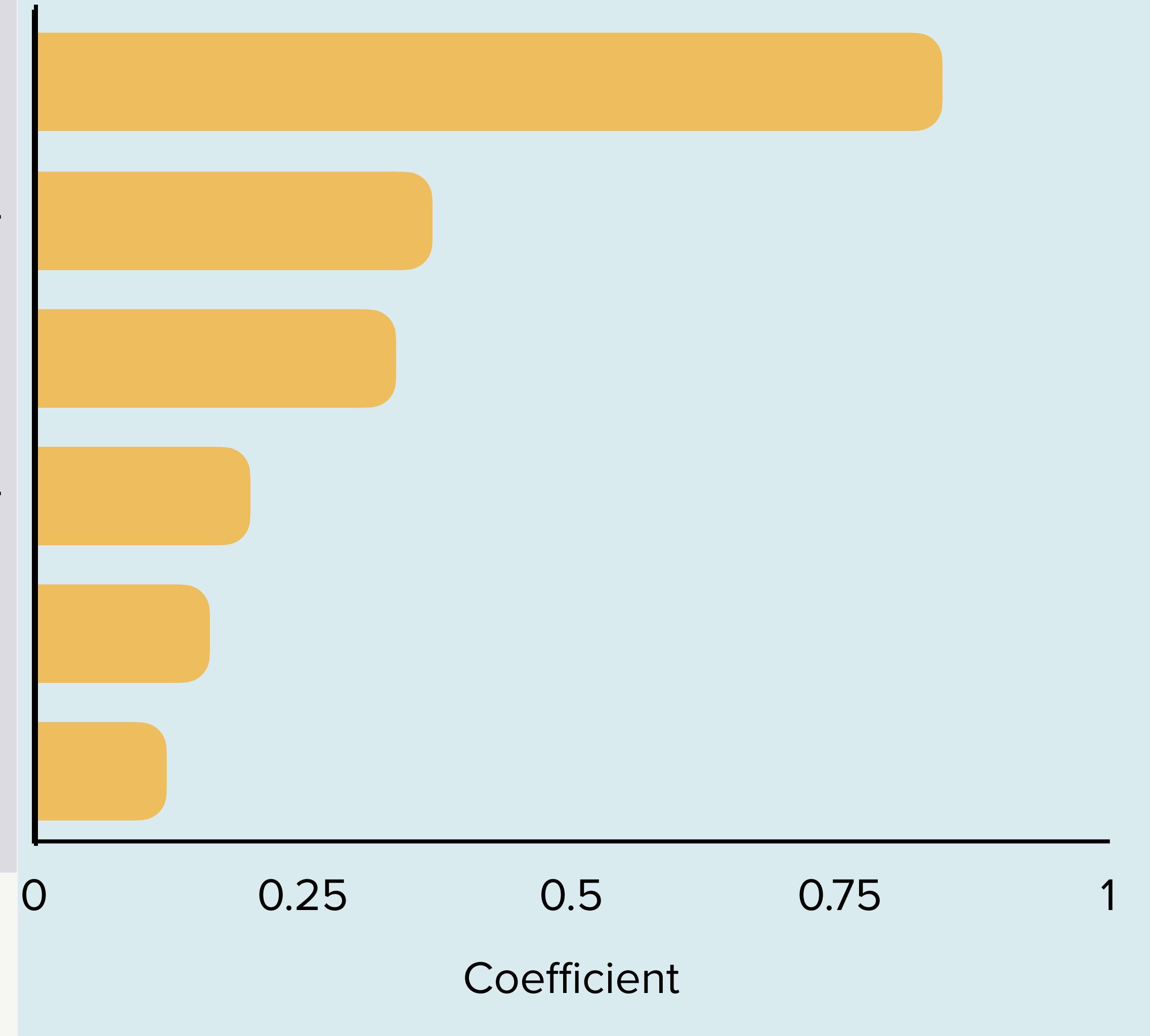
Current Company type: Public Sector

Major: Undisclosed

Current Company Type: Other

No Relevant Experience

Last New Job



# Important Features

Current Company type: Undisclosed

Current Company type: Public Sector

Major: Undisclosed

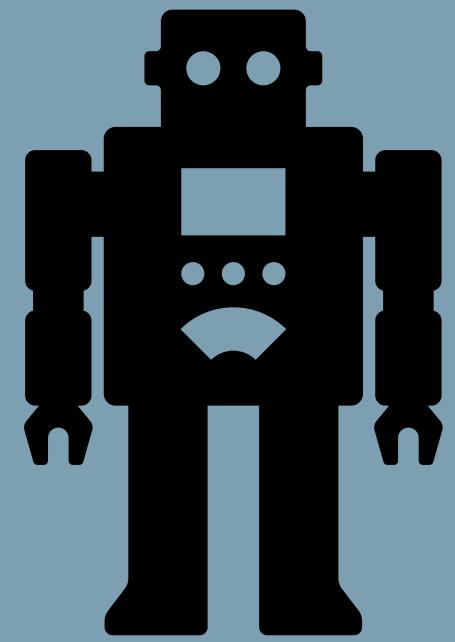
Current Company Type: Other

No Relevant Experience

Last New Job



# Thank you!



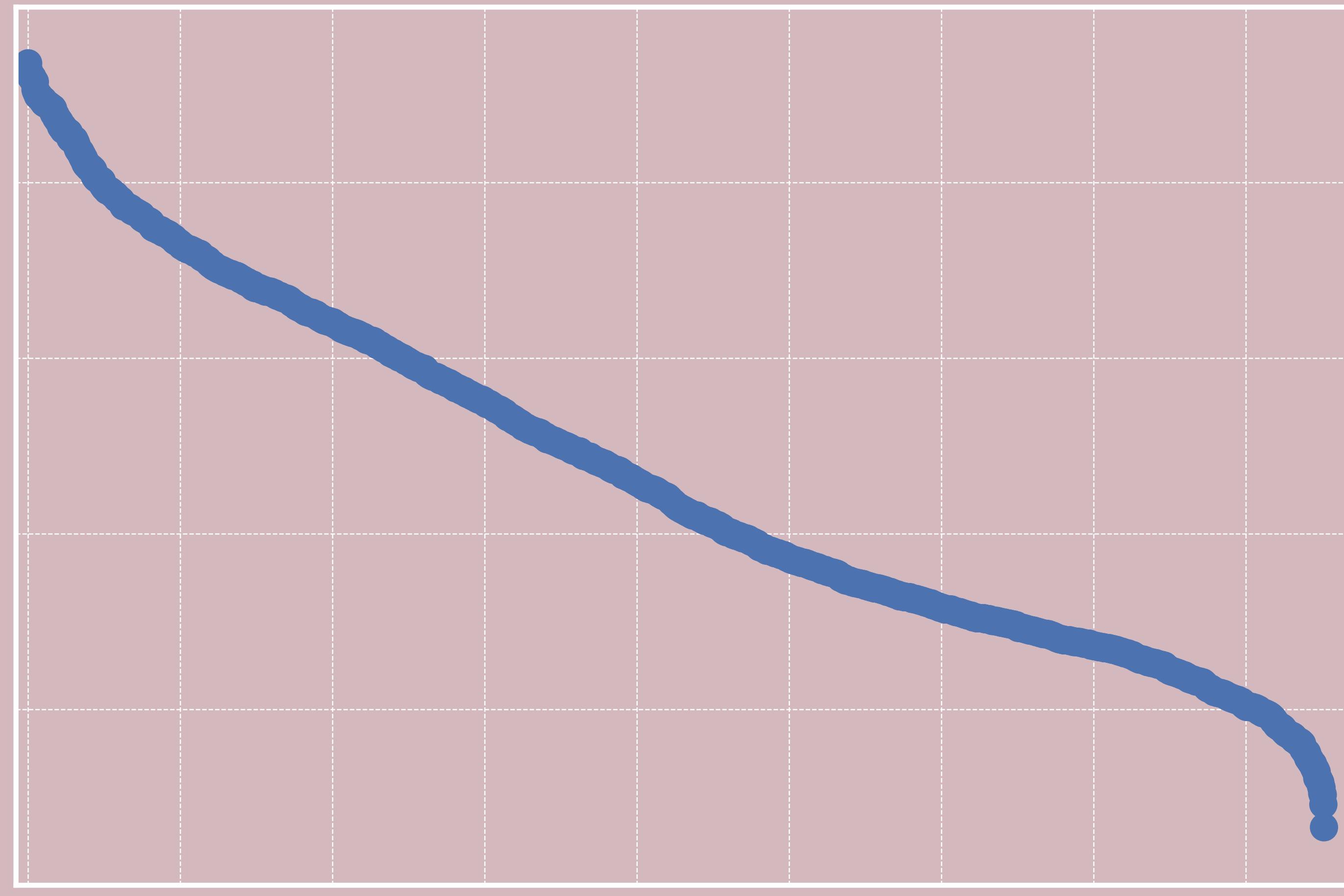
# Appendix

# Models Comparison

	Logistic Regression	kNN	Naive Bayes	Decision Tree	Random Forests	XGBoost
ROC AUC	0.765	0.690	0.687	0.775	0.750	0.715
Precision	0.41	0.41	0.35	0.49	0.47	0.29
Recall	0.83	0.67	0.76	0.76	0.78	0.98
Note	threshold = 0.4	n_neighbors = 5		(Grid Search results) max_depth = 6 max_features = 0.2 min_samples_leaf = 0.08	(Grid Search results) max_depth = 4 max_features = 'sqrt' min_samples_leaf = 0.1 n_estimators = 300	eval_metric = 'error' early_stopping_rounds = 50

Probability of Looking for a New Job (%)

## Switcher probability vs Number of Candidate



Model: Logistic Regression