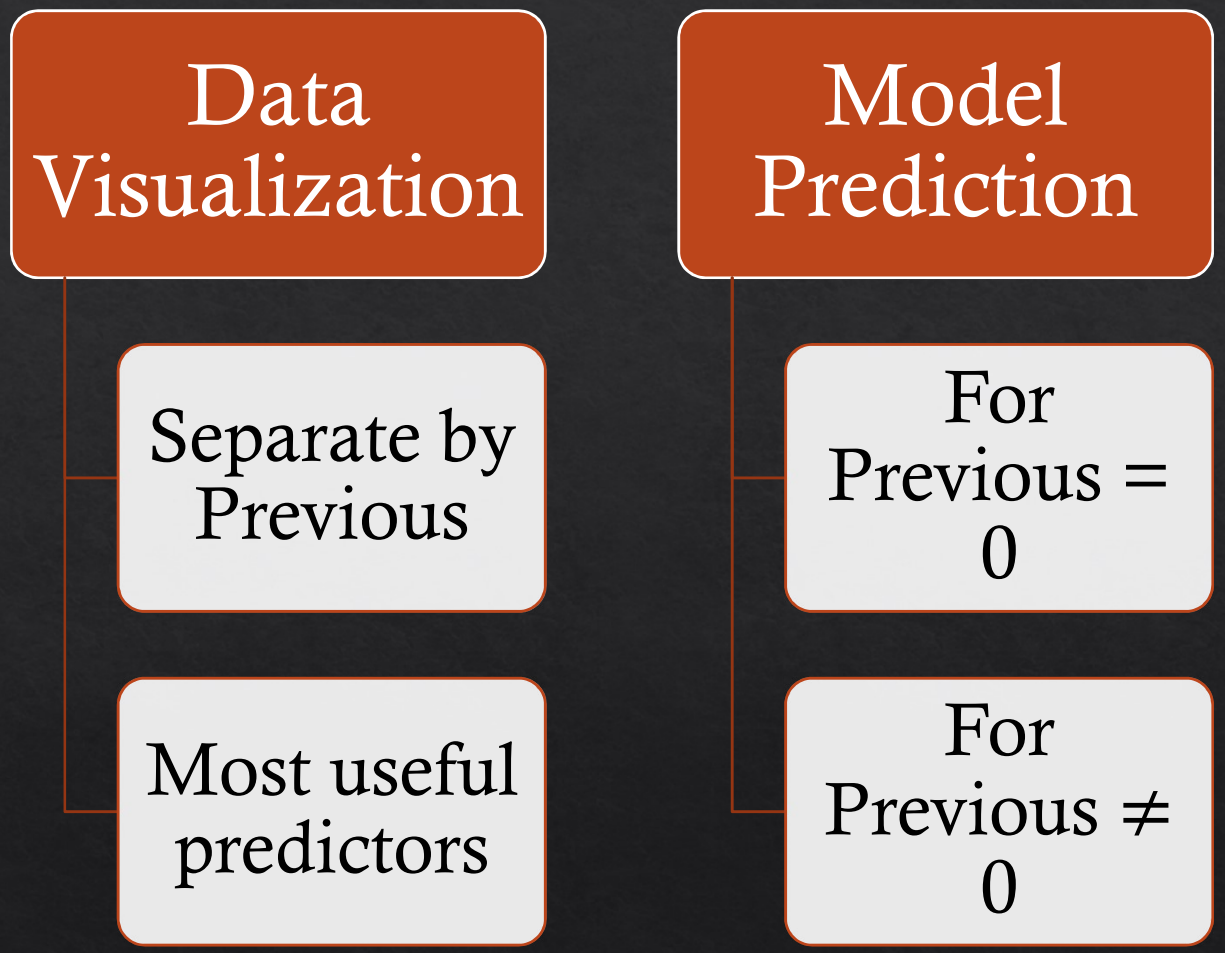
R-noobs for Hackathon

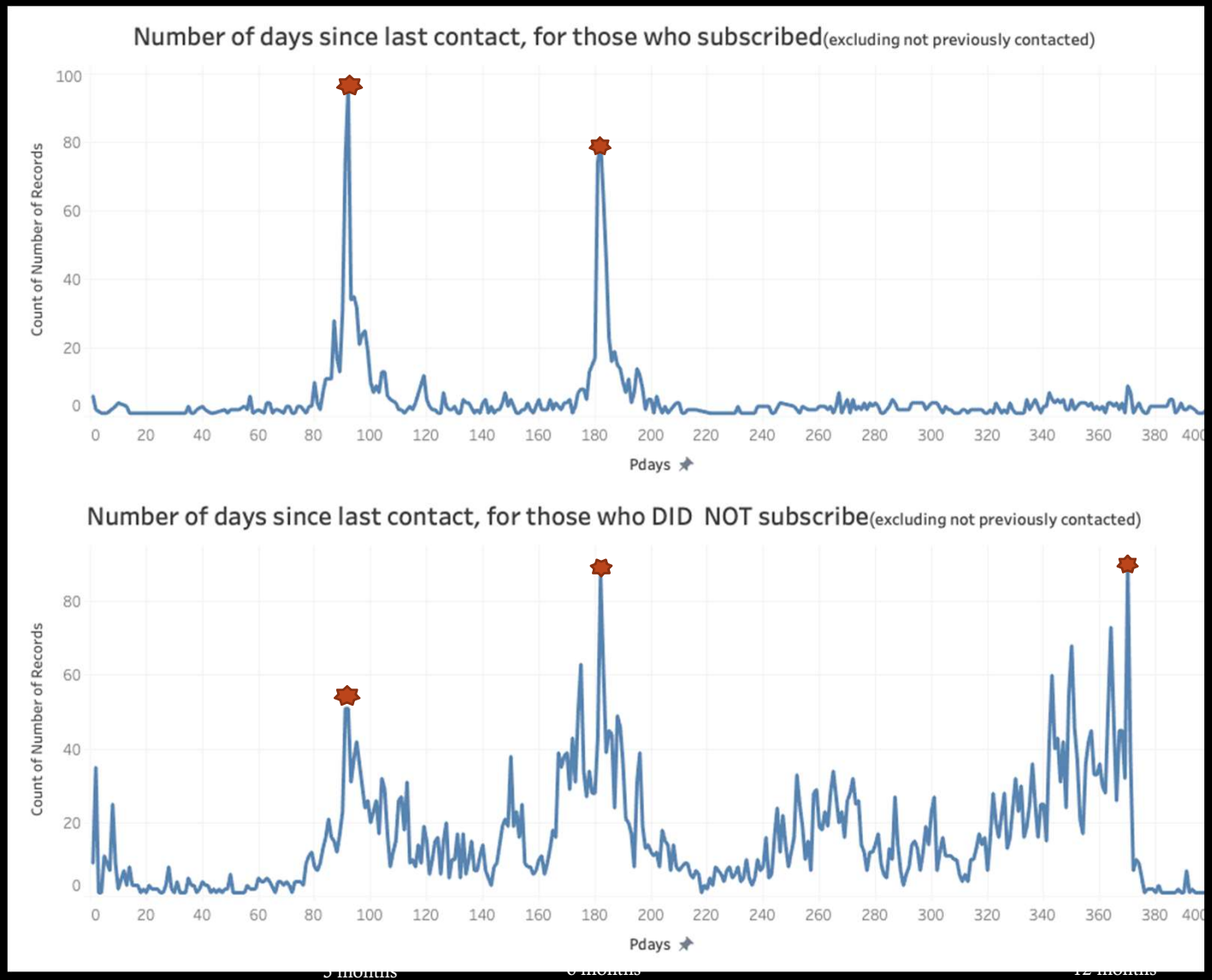
- Members:
- JT Sarmento
- Zhoulin Wang
- Temuulen Enebish
- Tingyu Yang


Data Visualization
Separate by Previous
Most useful predictors
Model Prediction
For Previous = 0
For Previous ≠ 0

Previous Days Since Last Contact

For Previous ≠ 0, which means clients who have been contacted before.

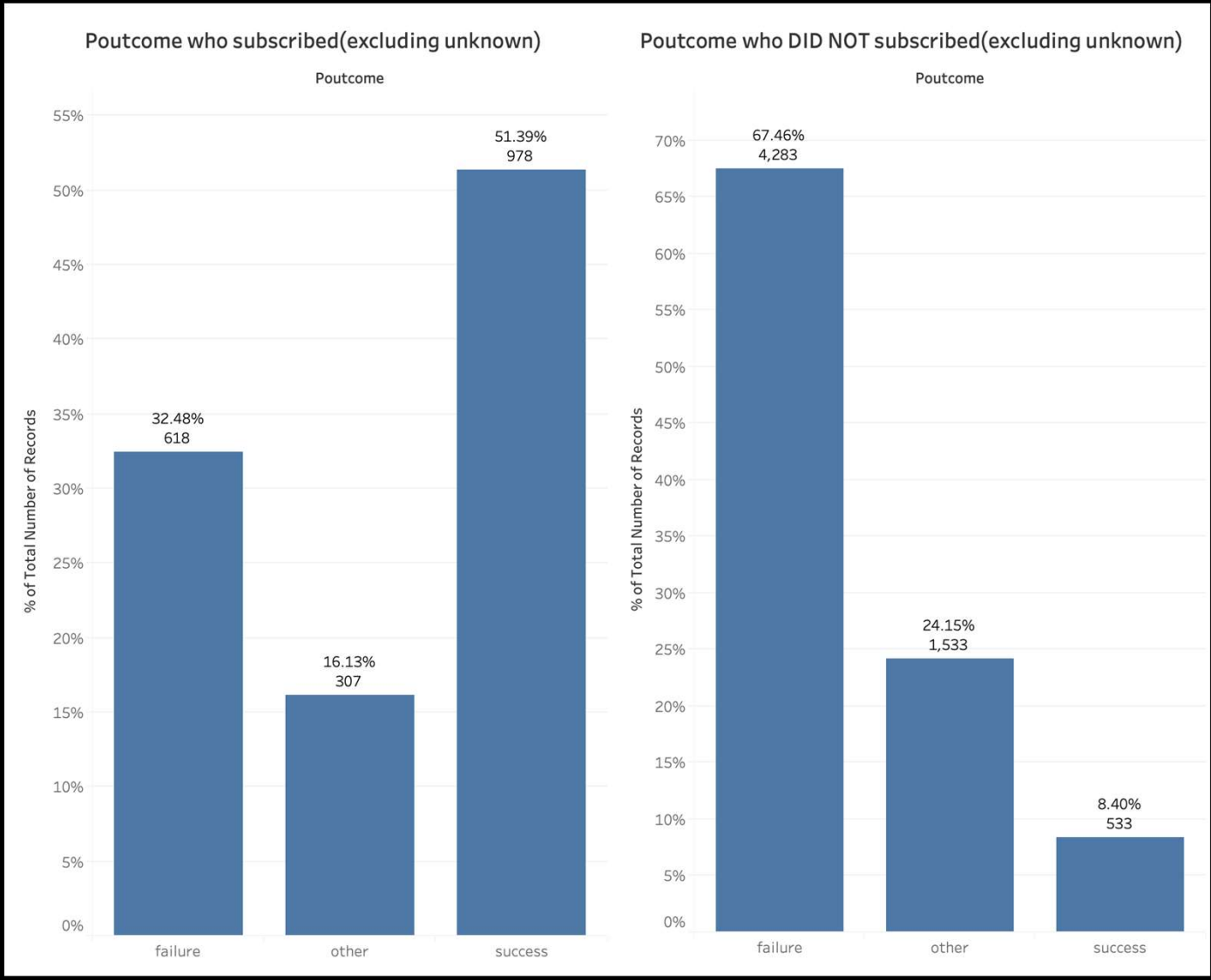
We can imply **a sales strategy of reconnecting with customer every 3rd 6th and 12th month**.



Outcome of Last Marketing Campaign

For Previous ≠ 0, which means clients who have been contacted before.

If you had been marketed to in the past, you have a high probability of subscribing again.

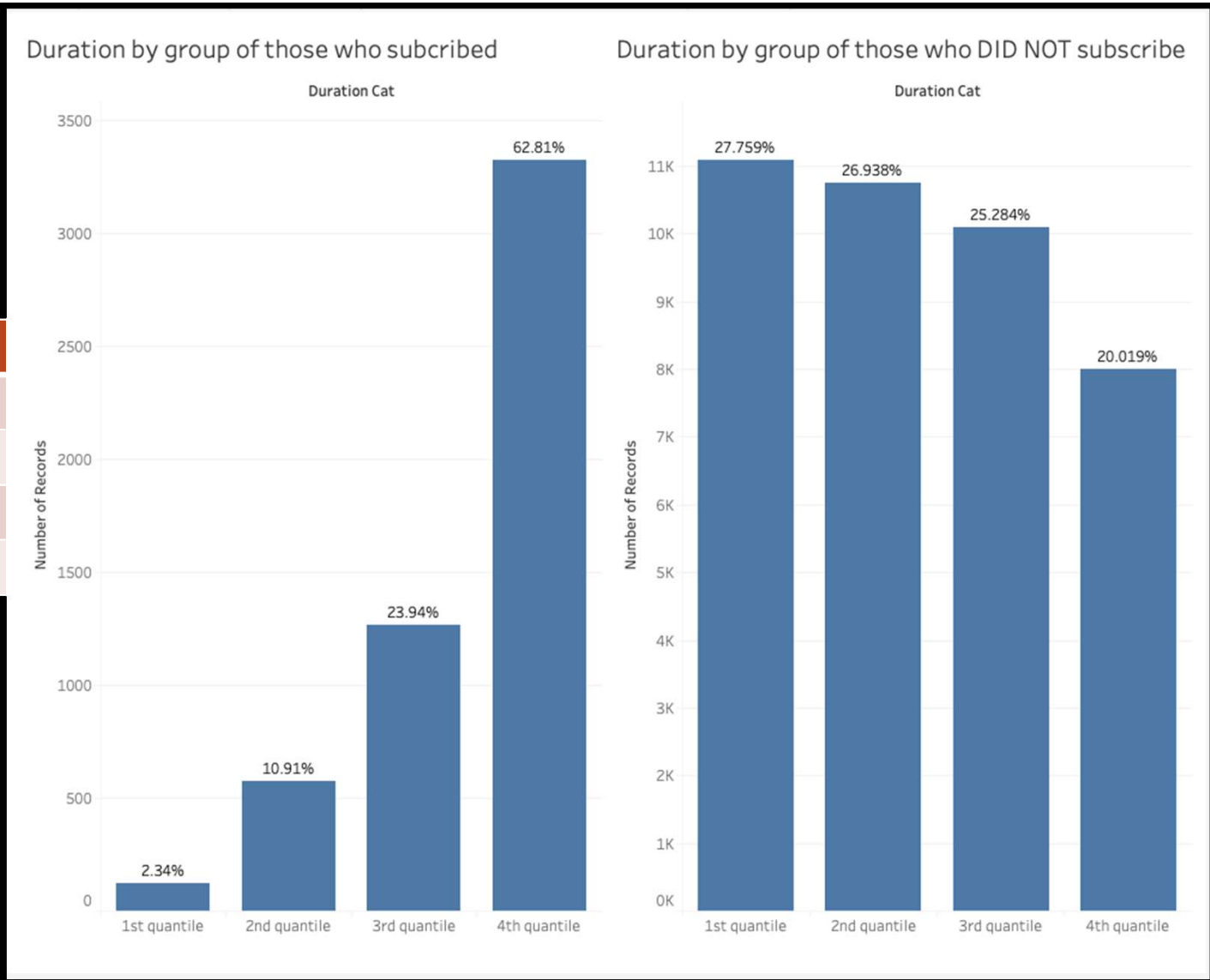


Duration

Category	Quantile
1st quantile	[0, 103)
2nd quantile	[103, 180)
3rd quantile	[180, 319)
4th quantile	[319, 4918]

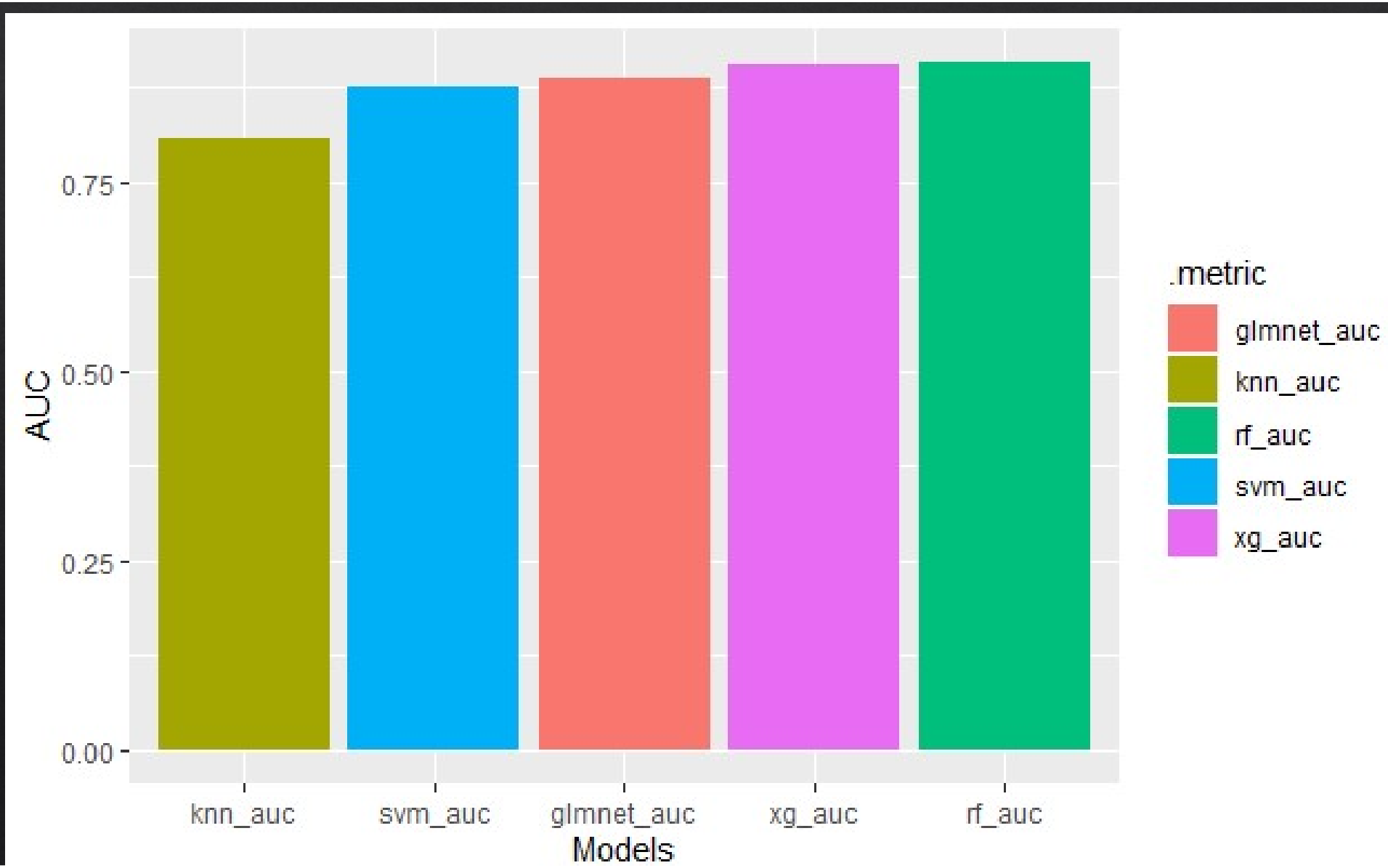
Of people who subscribed, majority of them spend more time during the phone call.

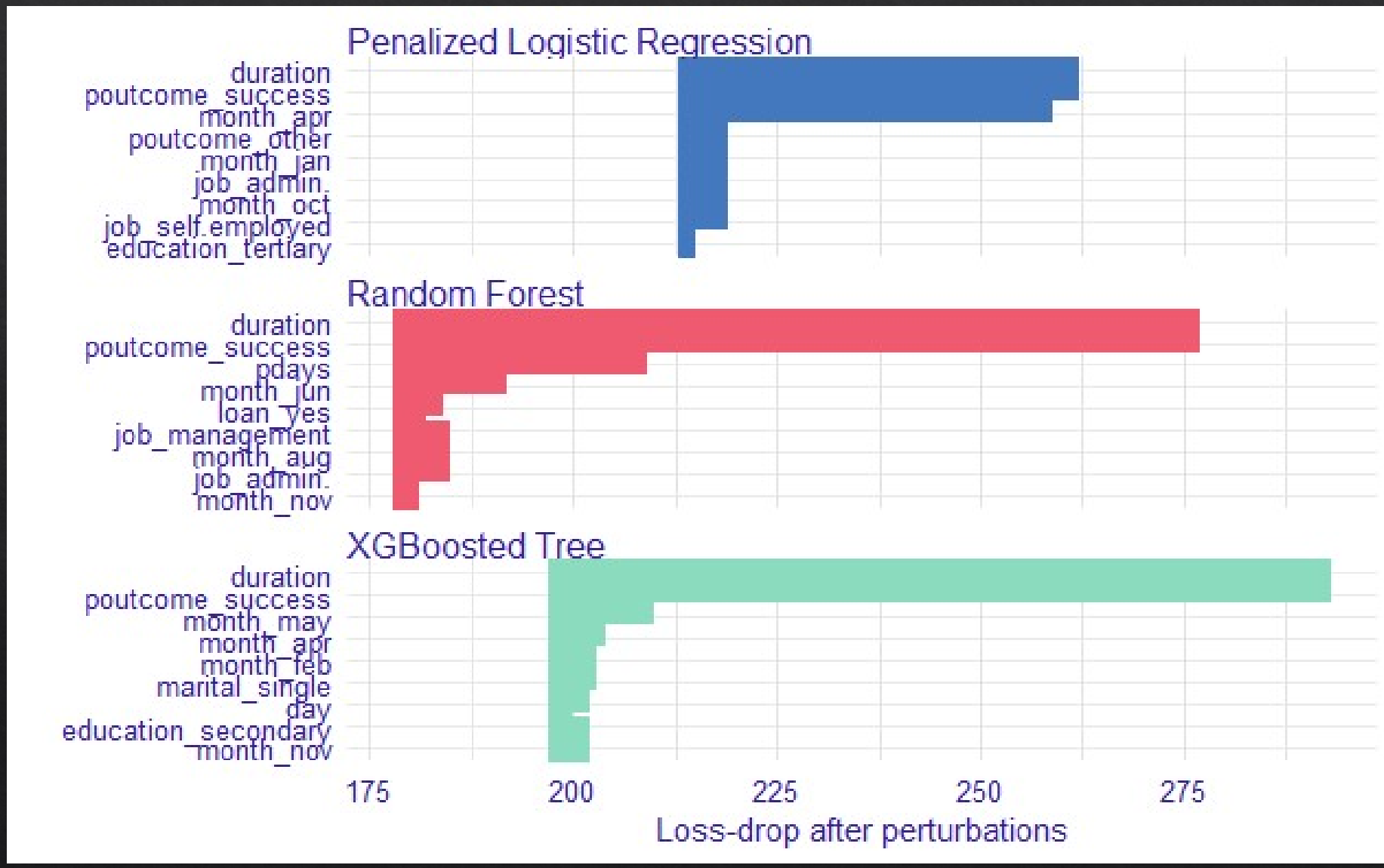
Of people who did not subscribe, the majority of them spend less time.

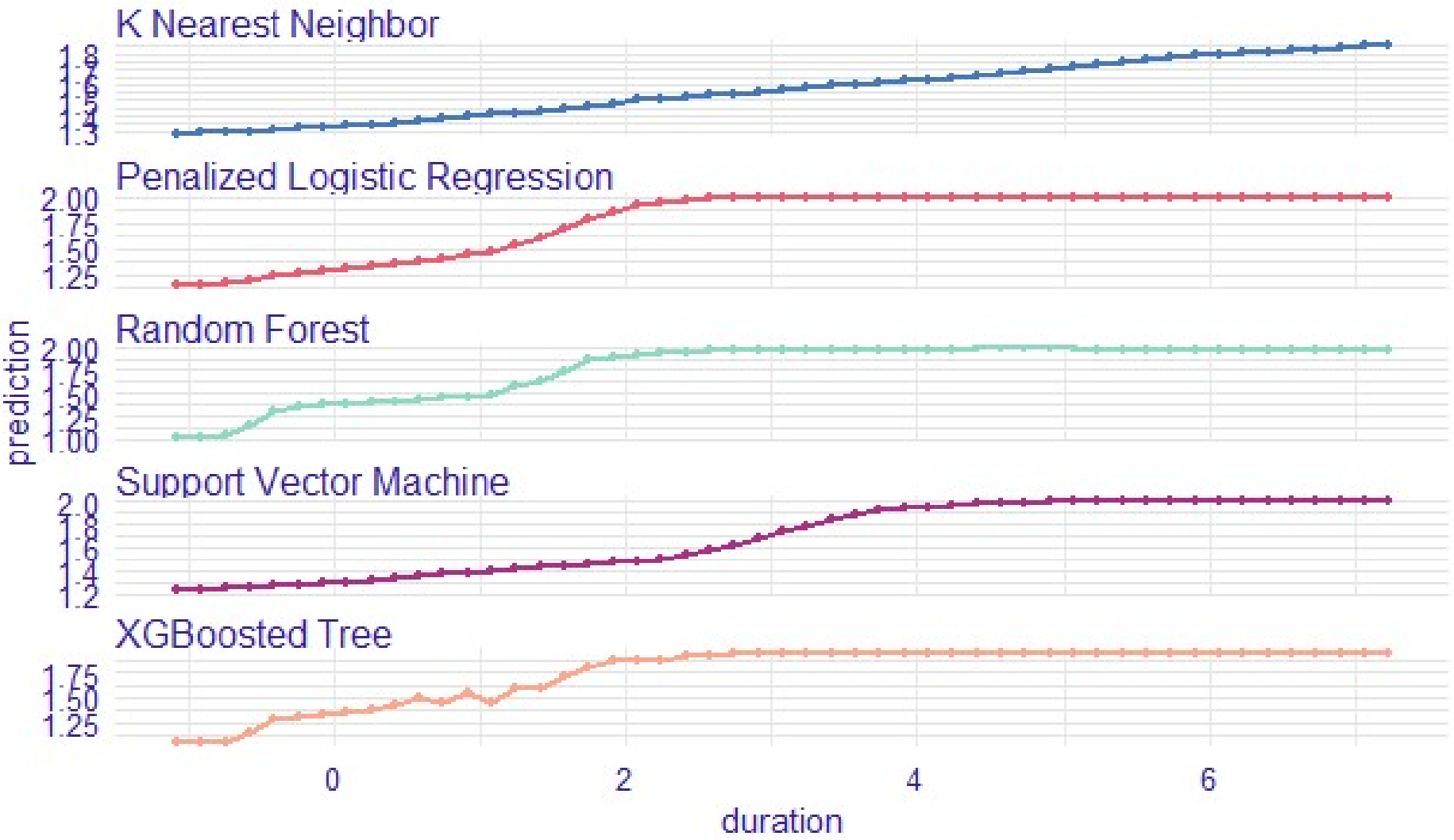


Predictive Modeling

- Two different models
- Hyperparameter tuning
- Time constraint
- Model explanation


AUC
0.75
0.50
0.25
0.00
knn_auc
svm_auc
glmnet_auc
xg_auc
rf_auc
Models
.metric
glmnet_auc
knn_auc
rf_auc
svm_auc
xg_auc


Penalized Logistic Regression
duration
poutcome_success
month_apr
poutcome_other
month_jan
job_admin.
month_oct
job_self.employed
education_tertiary
Random Forest
duration
poutcome_success
pdays
month_jun
loan_yes
job_management
month_aug
job_admin.
month_nov
XGBoosted Tree
duration
poutcome_success
month_may
month_apr
month_feb
marital_single
day
education_secondary
month_nov
175
200
225
250
275
Loss-drop after perturbations


K Nearest Neighbor
Penalized Logistic Regression
Random Forest
Support Vector Machine
XGBoosted Tree
prediction
2.00
1.75
1.50
1.25
1.00
0
2
4
6
duration