Production Of Contaminated Water

Begin	Agenda	There's a problem	Plan of action	What causes the problem?	What contributed the most to contamination?	To explain further....	Separate the water	Top 3 models to do the separation and prediction	Threshold Setting

Production Of Contaminated Water

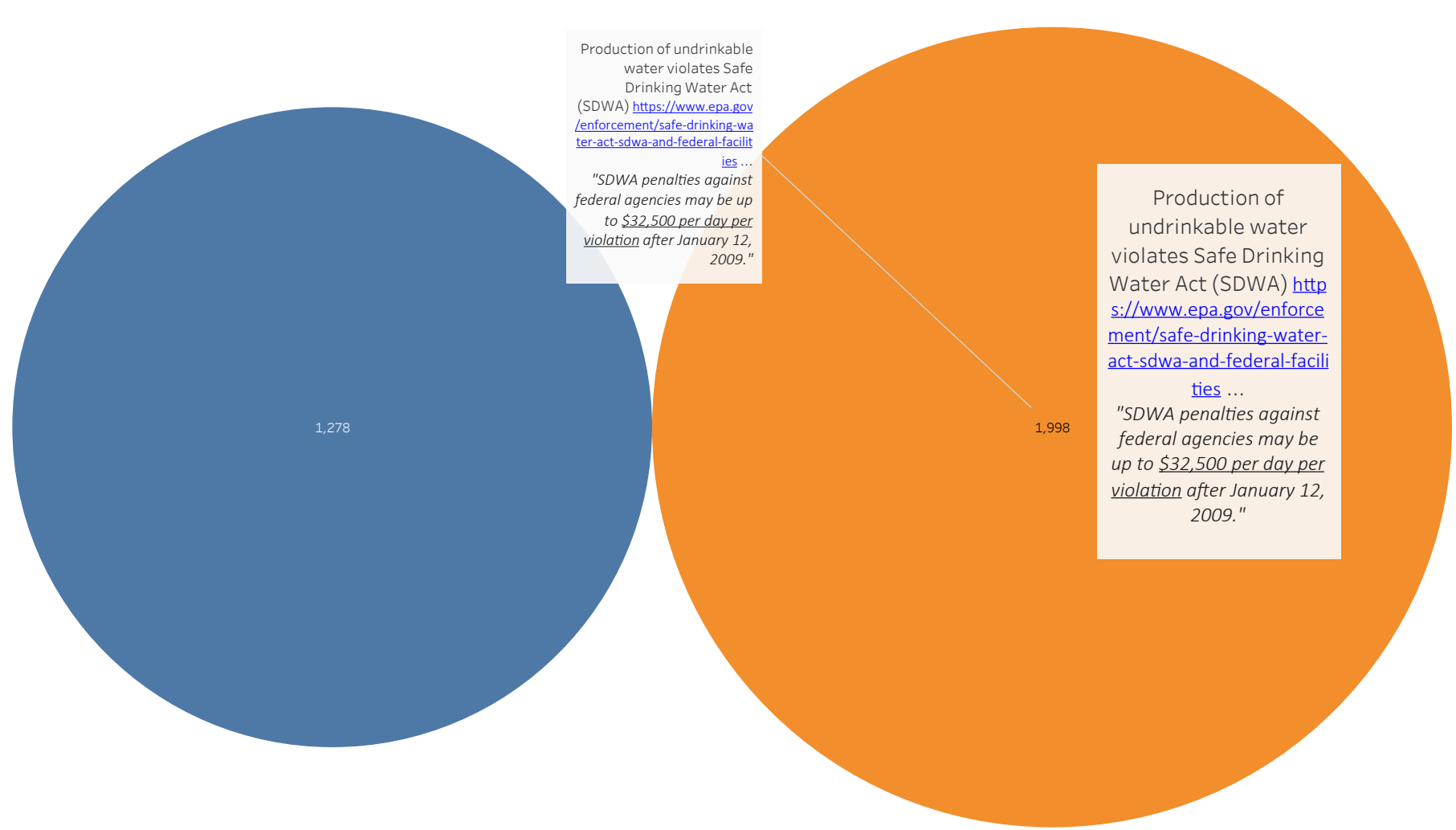
AGENDA

A friend of mine found himself in deep trouble when he took on a new job as a supervisor of a water treatment plant. A warning came from the federal authorities threatening to impose a big fine due to the violation of the Safe Water Drinking Act. He later discovered that most of the water produced in the plant were contaminated....

The plant produces water for consumption as well as for other manufacturing purposes.

Based on statistical results, we are to identify undrinkable water and separate them from drinkable ones. And also automate the whole process through Artificial Intelligence.

Production Of Contaminated Water



Production Of Contaminated Water

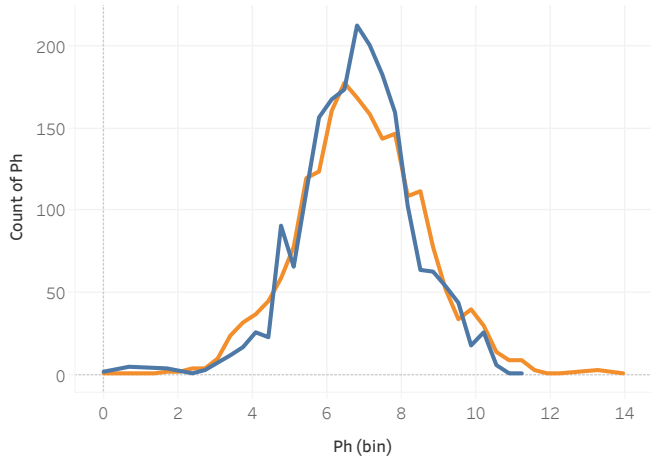
Plan Of Action:

1. Find out the reason behind the problem
2. Separate the waters based on these reasons
3. Use many models to separate the waters
4. Determine the best models
5. Determine the Threshold to separate the waters
6. Remove the contaminated water for consumption

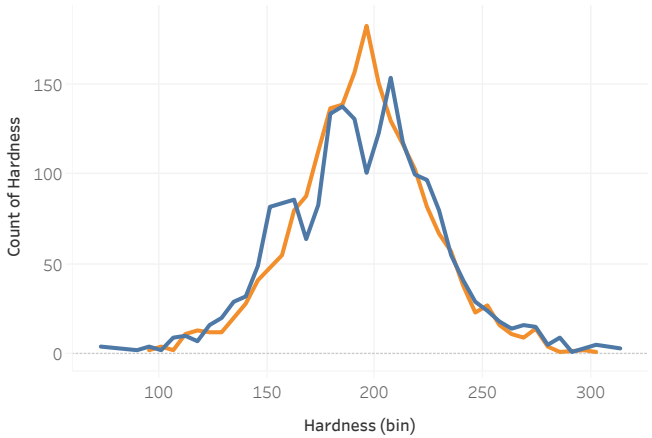
Production Of Contaminated Water

Drinkability drinkable undrinkable

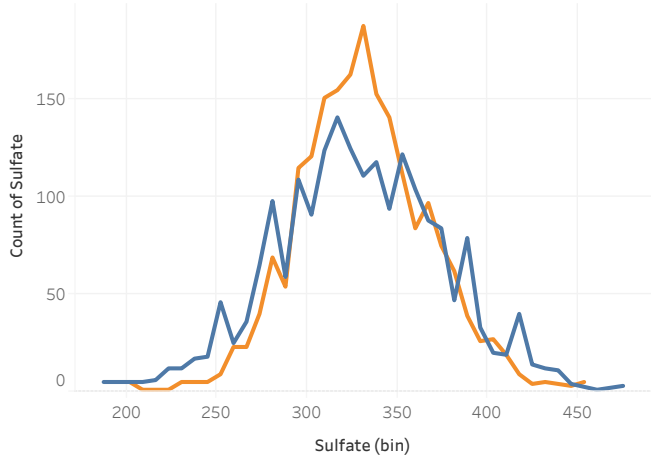
Ph



Hardness

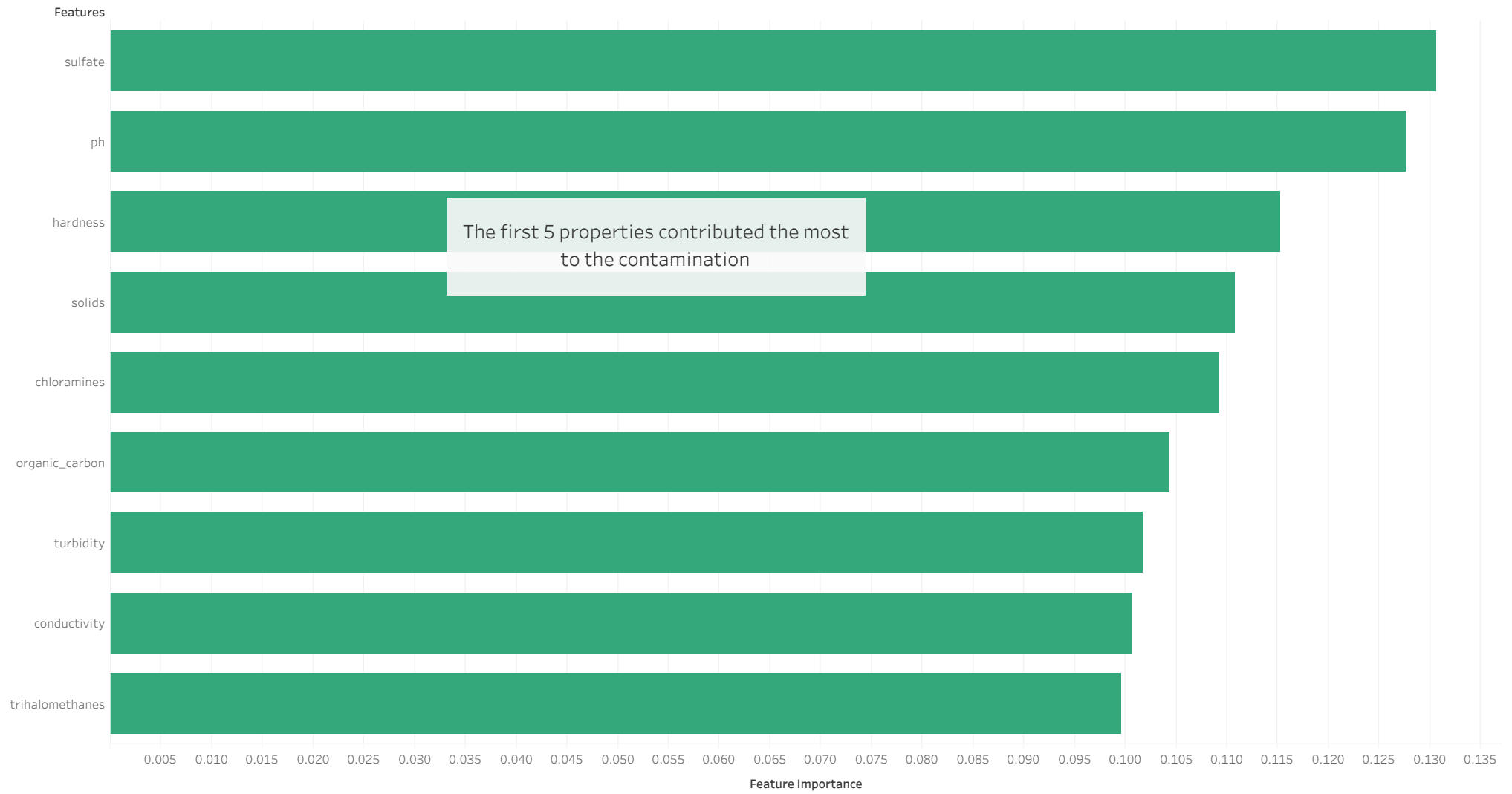


Sulfate

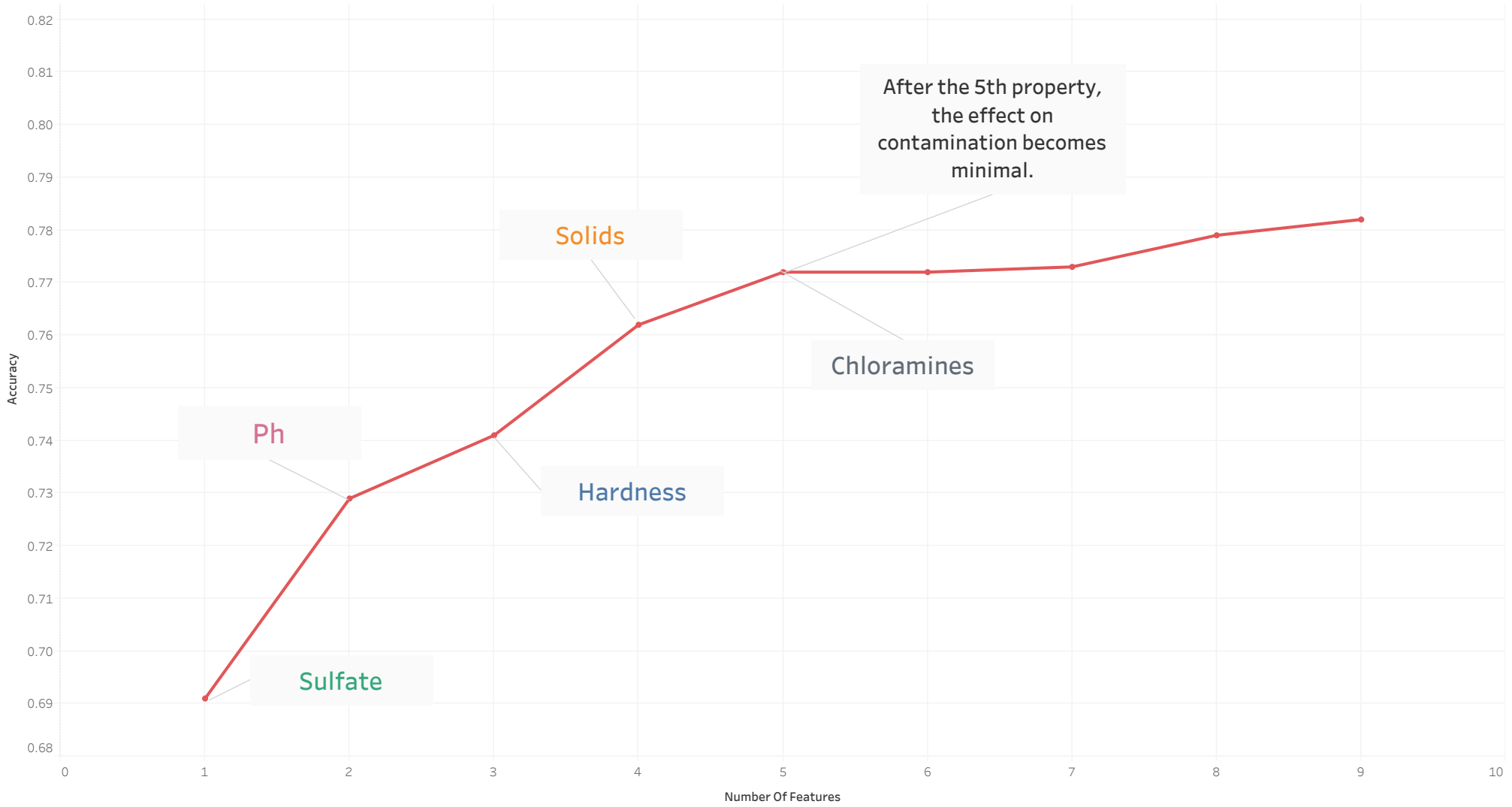


Solids

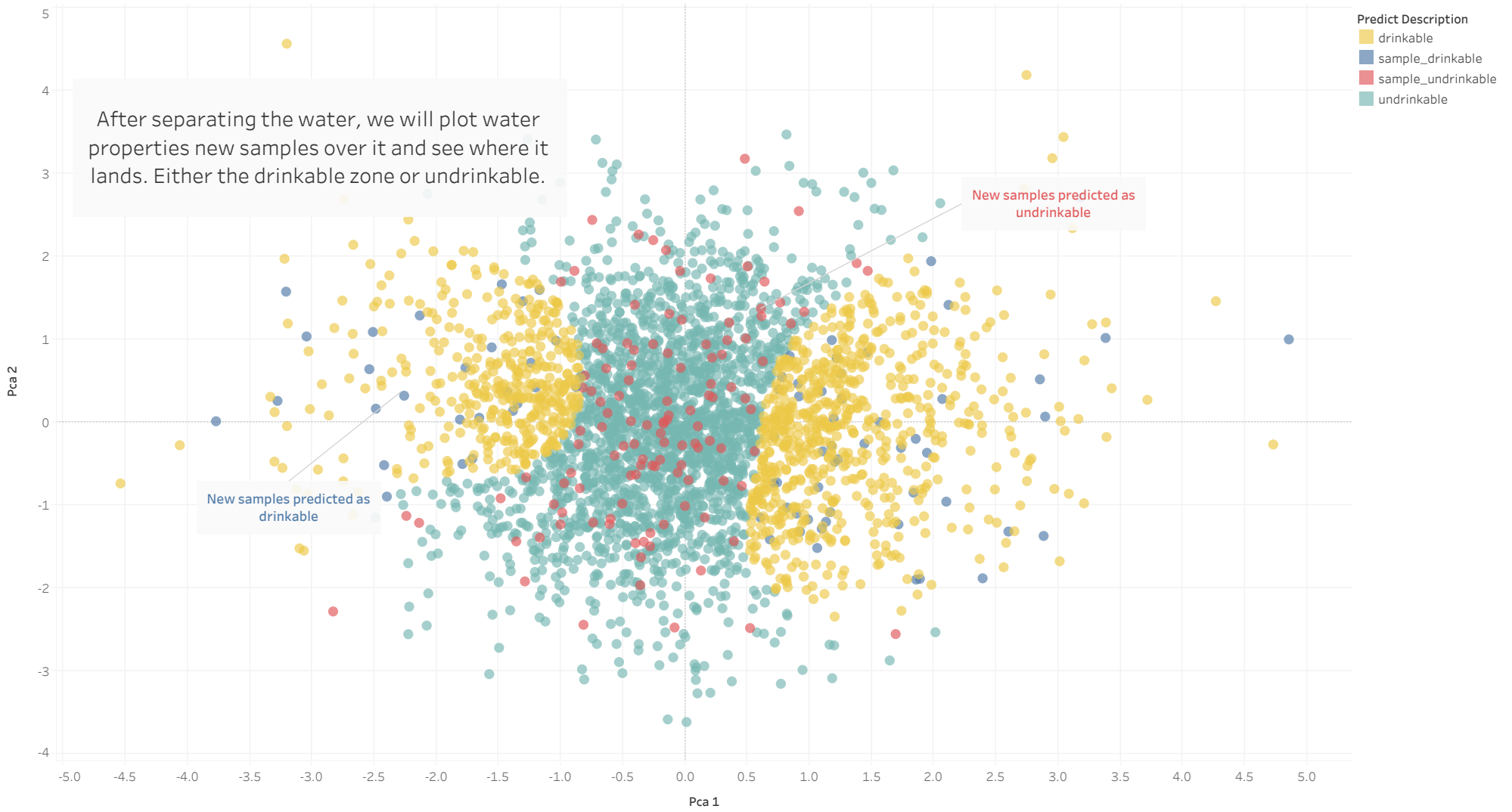
Production Of Contaminated Water



Production Of Contaminated Water



Production Of Contaminated Water



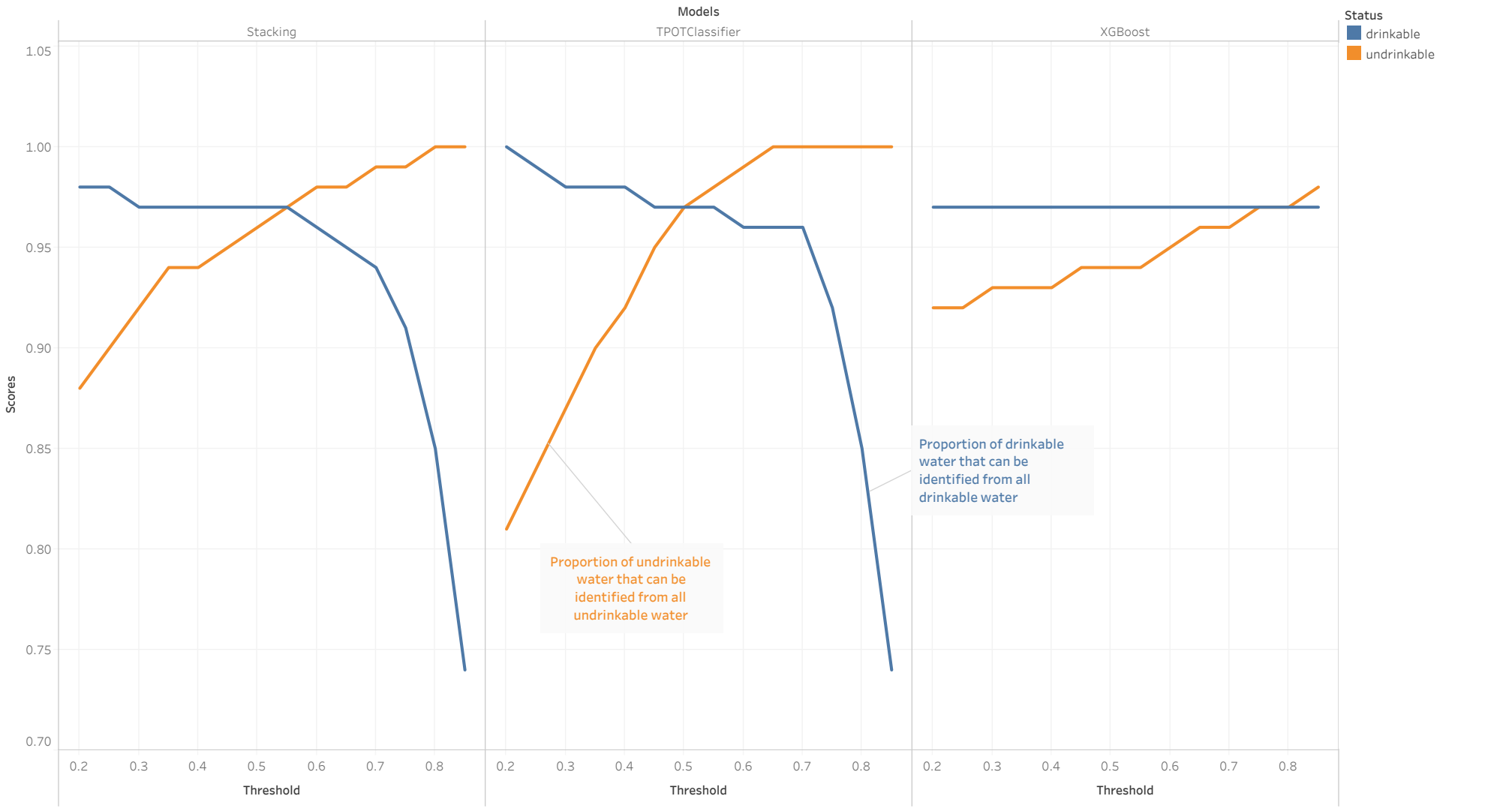
Production Of Contaminated Water

Production Of Contaminated Water

Threshold setting is a management decision

The higher the threshold the more undrinkable water we can identify (aka 'Specificity')

Production Of Contaminated Water



Production Of Contaminated Water

In [53]:
```
_ = sample.loc[:,'sulfate':'turbidity']
model3 = stc
threshold = 0.5

predict_water(_, model3, threshold)
```

```
There are 120 batches of water that are undrinkable

Undrinkable:
0        batch_1
1        batch_2
2        batch_3
3        batch_4
4        batch_5
         ...
182    batch_183
186    batch_187
188    batch_189
195    batch_196
196    batch_197
Name: batch, Length: 120, dtype: object

Drinkable:
5        batch_6
6        batch_7
7        batch_8
8        batch_9
9       batch_10
         ...
193    batch_194
194    batch_195
197    batch_198
198    batch_199
199    batch_200
Name: batch, Length: 80, dtype: object
```

In [54]:
```
_ = sample.loc[:,'sulfate':'turbidity']
model3 = stc
threshold = 0.6

predict_water(_, model3, threshold)
```

```
There are 124 batches of water that are undrinkable, a

Undrinkable:
0        batch_1
1        batch_2
2        batch_3
3        batch_4
4        batch_5
         ...
186    batch_187
187    batch_188
188    batch_189
195    batch_196
196    batch_197
Name: batch, Length: 124, dtype: object

Drinkable:
5        batch_6
6        batch_7
7        batch_8
8        batch_9
9       batch_10
         ...
193    batch_194
194    batch_195
197    batch_198
198    batch_199
199    batch_200
Name: batch, Length: 76, dtype: object
```

StackingClassifier:

Threshold :0.5
Identify 120 batches of undrinkable

Threshold :0.6
Identify 124 batches of undrinkable

Production Of Contaminated Water

Plan of action	What causes the problem?	What contributed the most to contamination?	To explain further....	Separate the water	Top 3 models to do the separation and prediction	Threshold Setting	Choose the threshold	Predict the sample	Thank you!

THANK YOU!