

Model Prediction Report



Executive Summary

	ACTUAL CAMPAIGN RESULT	WYZOO OPTIMIZED CAMPAIGN		
Covered audience	100%	40%	60%	80%
Mailed Quantity	1,191,928	476,771	715,157	953,542
Response Rate (cumulative)	3.85%	5.46%	4.80%	4.37%
Responder Quantity	45,914	26,040	34,328	41,655
Average Cost per piece (incl. postage)	\$0.70	\$0.80	\$0.80	\$0.80
Total Cost	\$834,350	\$381,417	\$572,125	\$762,834
Average Purchase	\$48.83	\$43.42	\$45.57	\$46.30
Total Revenue	\$2,242,107	\$1,130,720	\$1,564,200	\$1,928,629
Net Revenue	\$1,407,757	\$749,303	\$992,075	\$1,165,795
Cost to acquire	\$18.2	\$14.6	\$16.7	\$18.3

Goal

Train the model to **increase the response rate** of direct mail campaigns, while simultaneously ensuring that the **monetary value** of each order remains consistent.

Performance Overview

The Xperra model has identified the **45%** of those in the responder file who have a higher than average likelihood of **responding positively** to a campaign. This model will then be used to score **new records**. Regardless of the quantity of those in a mailing, the resulting file should contain prospects with the highest probability of a response and include the **ideal mailing quantity** for this **highly targeted group**.

These records are then incorporated into the model which will **progressively identify** which respondents should be targeted in **future campaigns**.

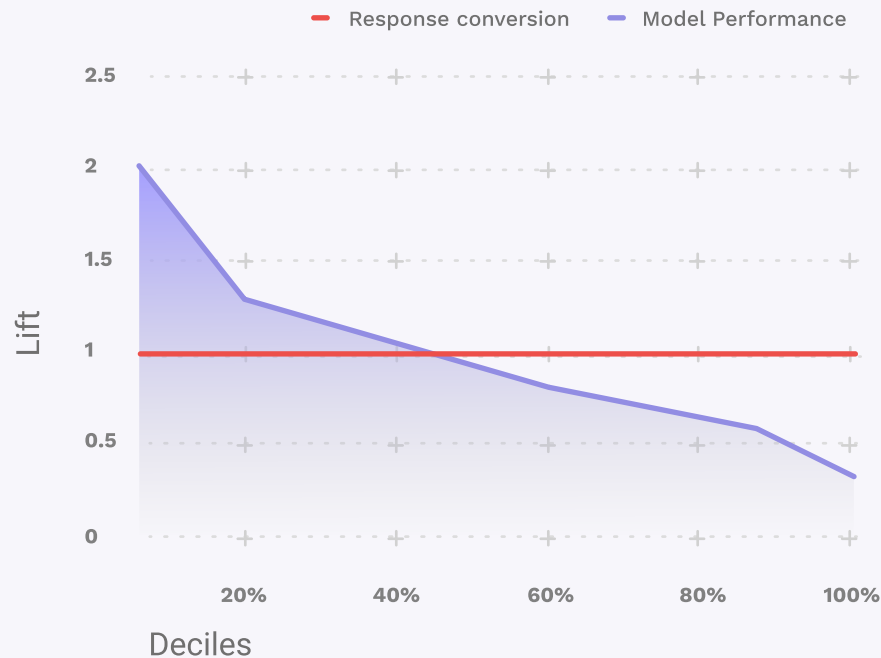
Recommendation

Here is a recommendation based on **Xperra's Artificial Intelligence Model**.

- We recommend an **initial test roll-out** comparing your **current marketing efforts** with a **test set of records** selected by the model. We will help you to determine the size of the test and **set up the experiment**.
- We will **analyze the responses** from the test campaign to identify where **improvements** are needed.
- We will import your latest mail and response files to **improve the calibration** of the model (not required, but recommended).
- Using the model to score and select people most likely to respond, we will confirm you are using the **best records** for each and **every campaign**.

Model Insights

- The below graph demonstrates how response rates using the model **increase over** those from the original file.



Power and performance through prediction

- Note that the response rates for each decile return a higher than average response until we reach the **45% marker**. After this point the you can see that that response rate of leads drops off considerably.

Model Insights

Optimizing your return on investment is a critical, **strategic activity**. We provide you with a snapshot of your data to help you identify your **ideal target markets**.

Actual Campaign

Xperra optimized campaign (increase size of mailing list based on data science model rank)

Performance of each decile

Audience covered	100%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Mailed Quantity	1,191,928	119,192	119,192	119,192	119,192	119,192	119,192	119,192	119,192	119,192	119,192
Response Rate (incremental)	3.85%	7.89%	5.19%	4.48%	4.04%	3.79%	3.57%	3.07%	2.70%	2.30%	1.48%
Responder Quantity	45,914	9,400	6,190	5,335	4,818	4,512	4,250	3,658	3,220	2,738	1,769
Average Cost per Piece (incl. postage)	\$0.7	\$0.80	\$0.80	\$0.80	\$0.80	\$0.80	\$0.80	\$0.80	\$0.80	\$0.80	\$0.80
Total Cost	\$834,350	\$95,354	\$95,354	\$95,354	\$95,354	\$95,354	\$95,354	\$95,354	\$95,354	\$95,354	\$95,354
Average Purchase	\$48.83	\$39.52	\$44.35	\$46.46	\$49.15	\$51.80	\$47.00	\$51.93	\$54.19	\$63.03	\$79.66
Total Revenue	\$2,242,107	\$371,520	\$274,503	\$247,901	\$236,796	\$233,706	\$199,774	\$189,946	\$174,484	\$172,588	\$140,890
Net Revenue	\$1,407,757	\$276,166	\$179,150	\$152,548	\$141,442	\$138,352	\$104,420	\$94,593	\$79,130	\$77,234	\$45,536
Cost to Acquire	\$18.2	\$10.1	\$15.4	\$17.9	\$19.8	\$21.1	\$22.4	\$26.1	\$29.6	\$34.8	\$53.9

Model Insights.

Used for Audience Selection – RR, AOV, \$/BK



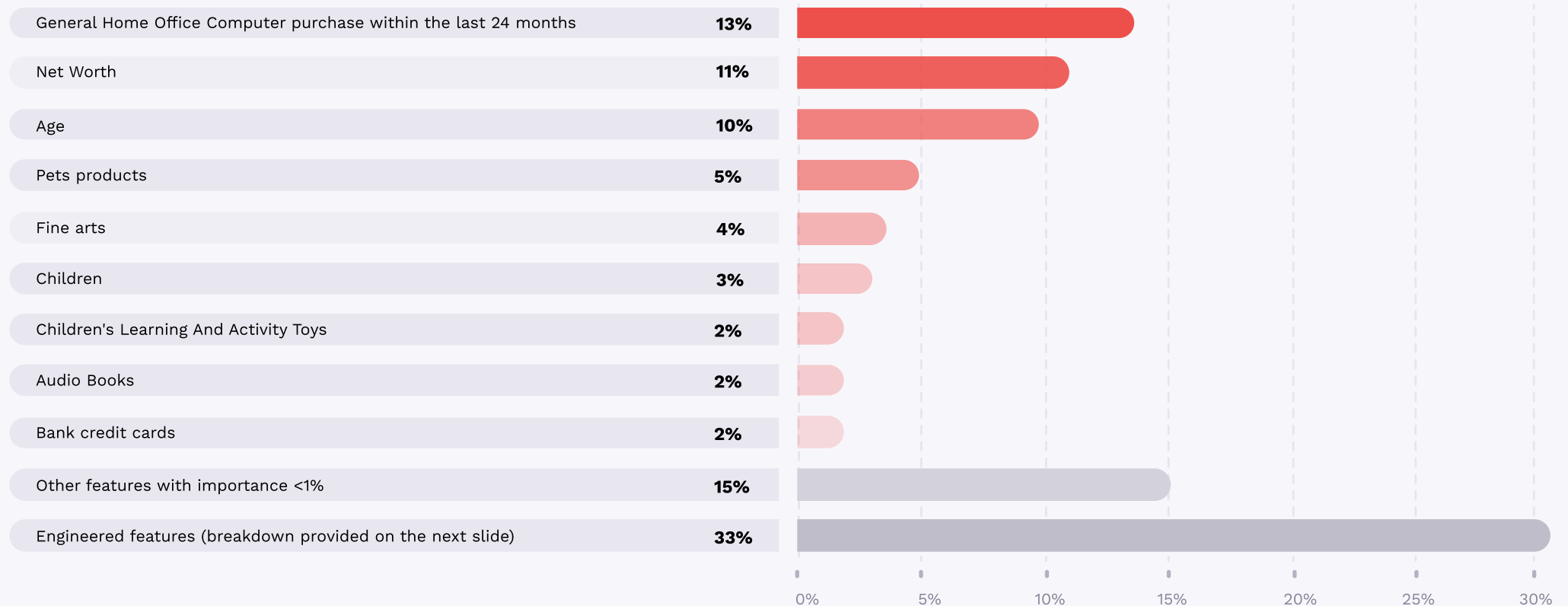
Model Insights

Most influential features

Below is a list of the most important features we have identified for predicting customer response. Based on the Xperra model, the graph below highlights the relative importance of those features.

Most Important Features

Features importance



Model Insights

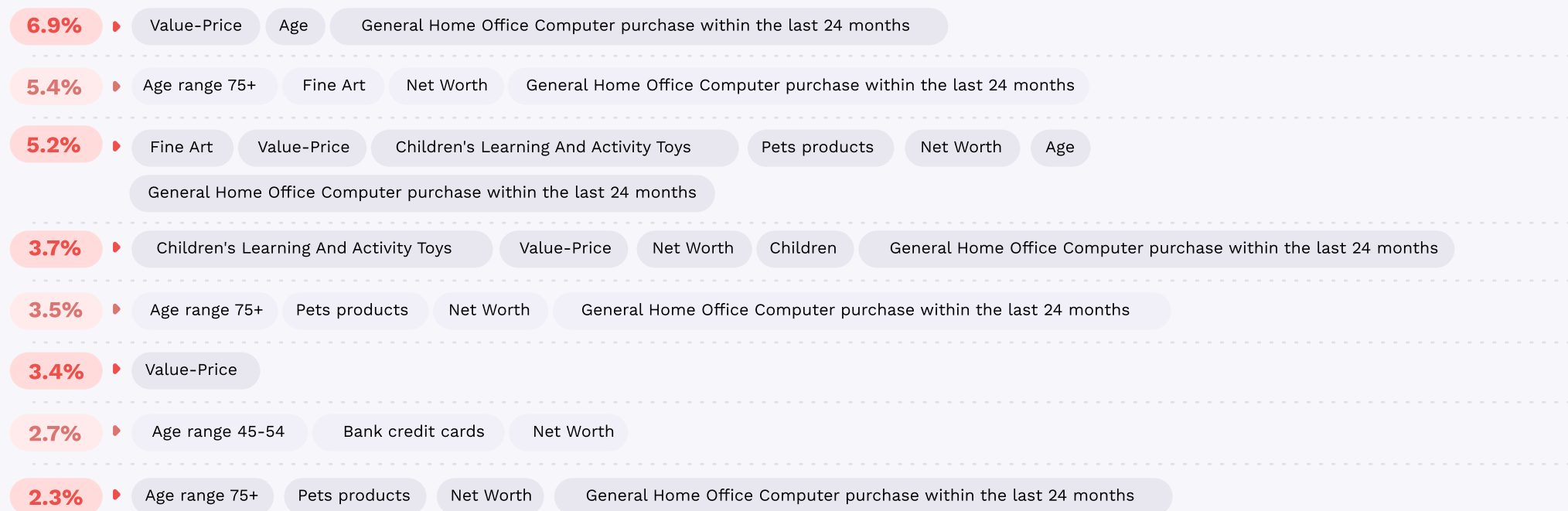
Using internal algorithms, the Xperra model builds a number of “engineered features.” These are designed to identify interactions between existing categories. These interactions are measured for impact, and then evaluated to determine whether to keep within the final model. Many features are created during the modeling process, however only those determined to have an actual impact on the predictive power of the model are kept.

Most influential engineered features

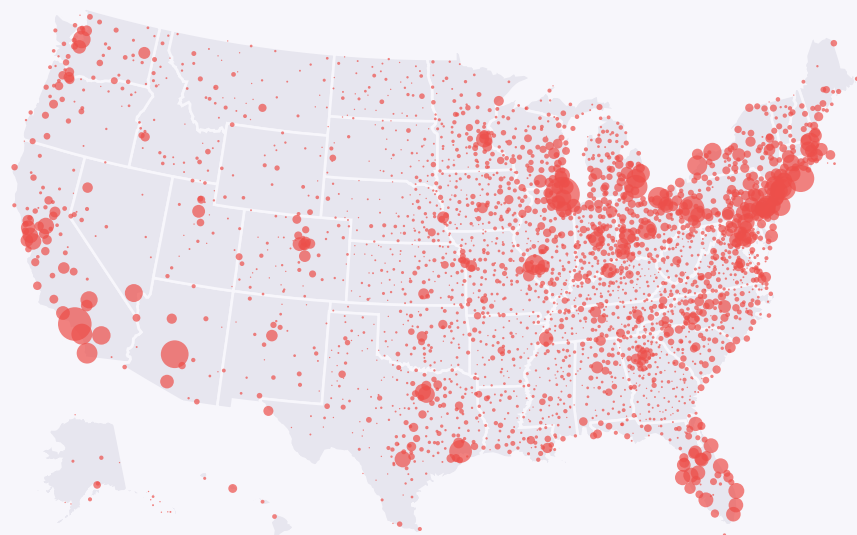
We have listed the most important engineered features identified by the Xperra model. The graph below highlights their relative importance.

Feature Importance

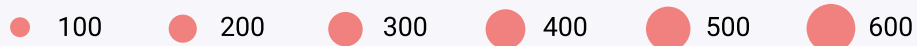
Most Important Features



Model Insights .

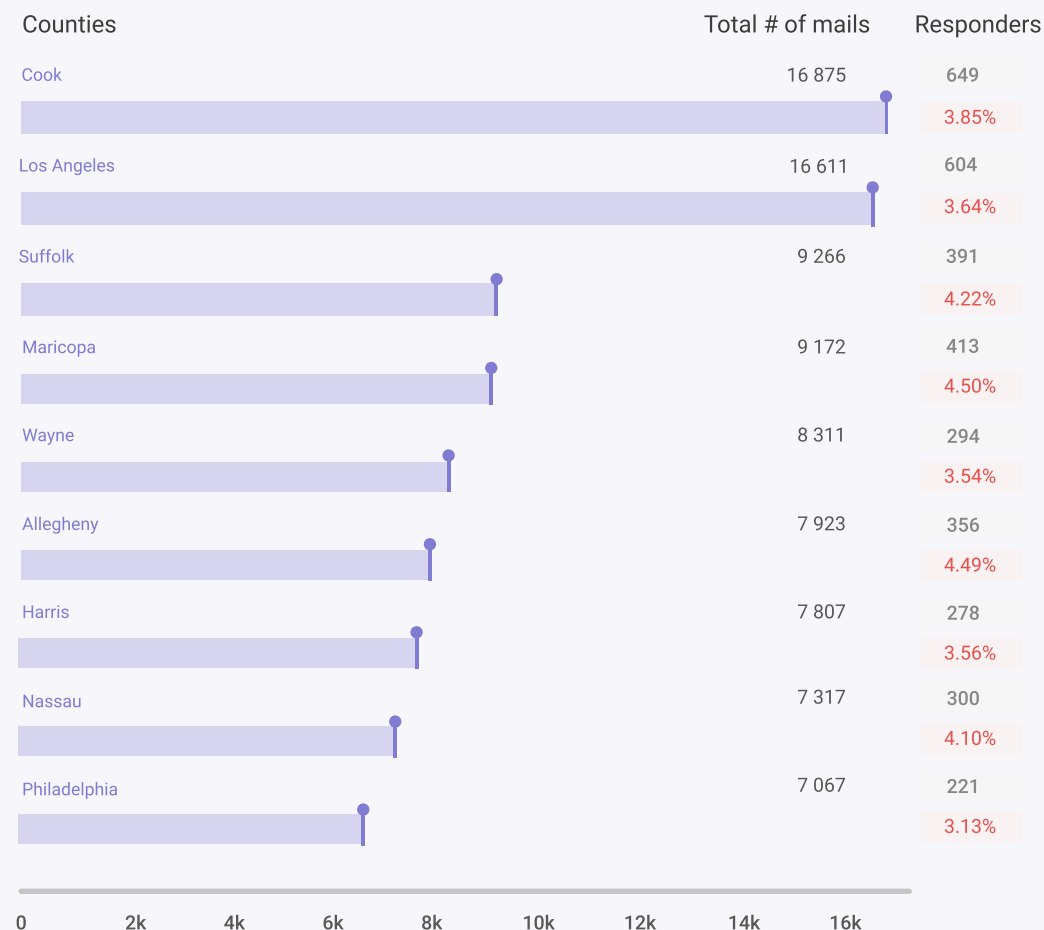


Responders number



Counties

We group your customer by county and find where you have the largest number of responders and the highest conversion rate.



Model Insights

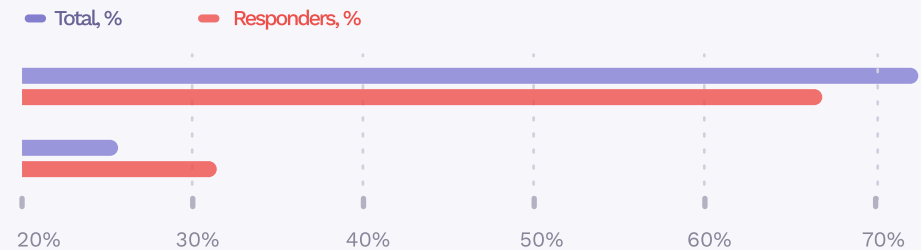
Statistics

Importance of feature to response rate: **13%**

General Home Office Computer Purchase

Your responders recently (within the last 24 months) purchased Home Office Computers.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total, % *	Impact on response rate
No	871,821	30,419	73.1%	66.3%	0.90	↓
Yes	320,107	15,495	26.8%	33.7%	1.25	↑



Model Insights

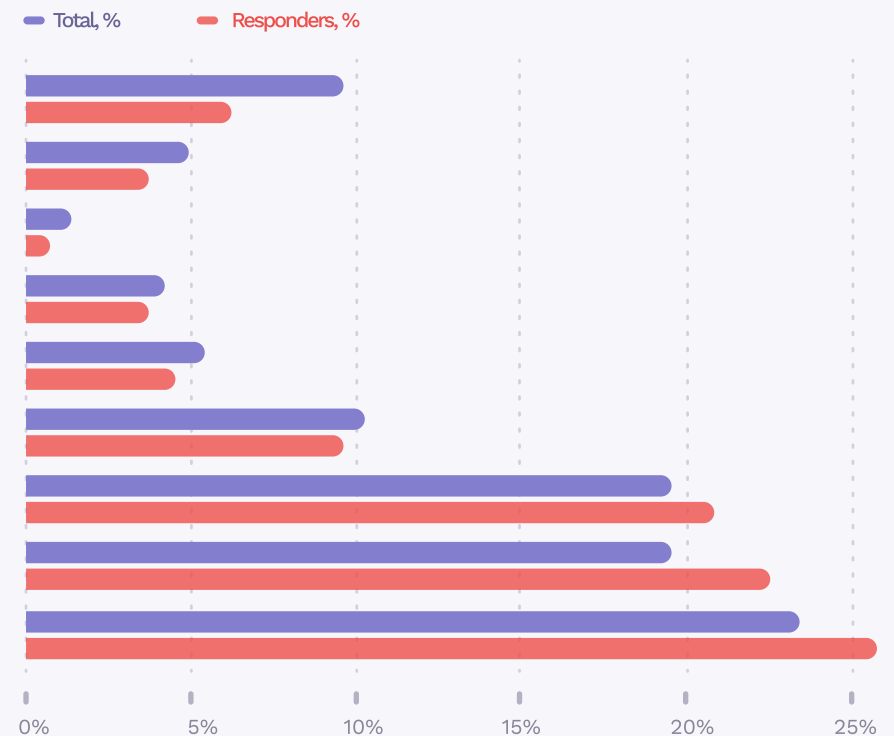
Statistics

Importance of feature to response rate: **11%**

Net Worth

Predicts the range in which a household's NetWorth is likely to fall.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total, % *	Impact on response rate
\$1<	108619	2717	9.9%	6.3%	0.6	↓
\$1 - \$4,9K	55321	1543	5.0%	3.6%	0.7	↓
\$5K - \$9,9K	16695	524	1.5%	1.2%	0.8	↓
\$10K - \$24,9K	47260	1604	4.3%	3.7%	0.8	↓
\$25K - \$49,9K	57811	2015	5.2%	4.7%	0.8	↓
\$50K - \$99,9K	111001	4208	10.1%	9.8%	0.9	↓
\$100K - \$249,9K	213386	8724	19.5%	20.5%	1.0	↑
\$250K - \$499,9K	213929	9411	19.5%	22.1%	1.1	↑
<\$499,9K	268498	11795	24.5%	27.7%	1.2	↑



Model Insights

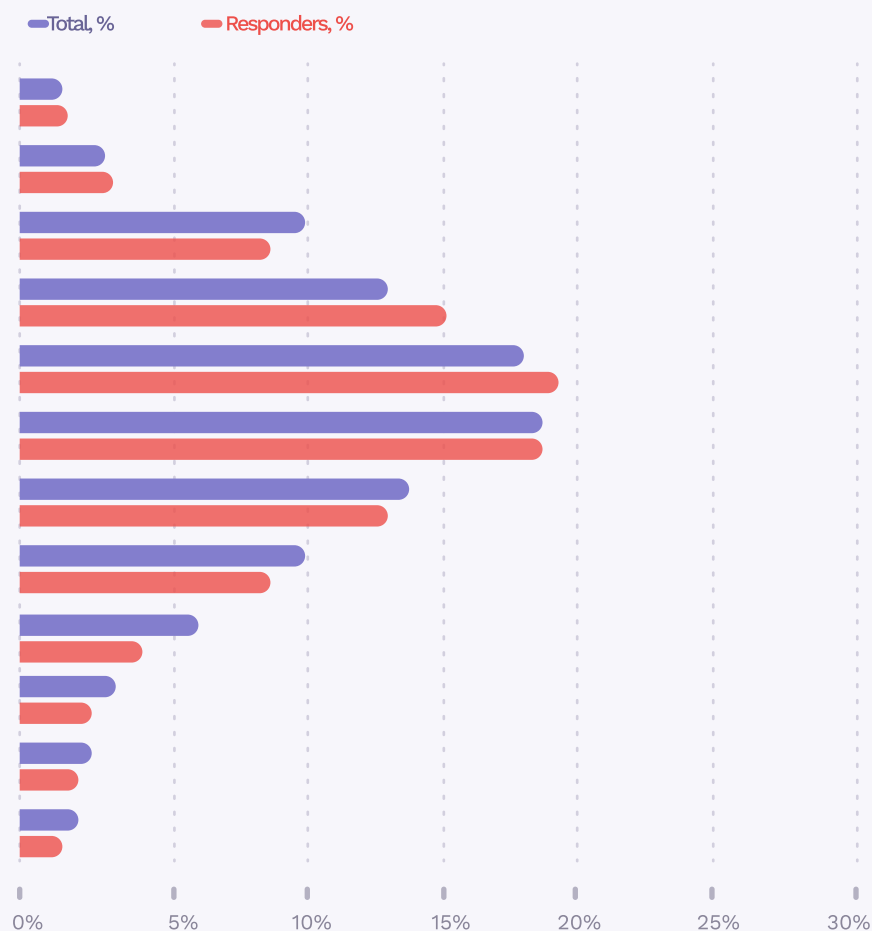
Statistics

Importance of feature to response rate: **10%**

Age

Age ranges of any adults present in the household.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total, % *	Impact on response rate
95 - 100	12,964	446	1.1%	1.0%	0.9	↓
90 - 94	44,099	1,847	3.9%	4.2%	1.1	↑
85 - 89	99,430	4,670	8.8%	10.6%	1.2	↑
80 - 84	150,369	6,667	13.3%	15.0%	1.1	↑
75 - 79	190,862	8,478	16.8%	19.1%	1.1	↑
70 - 74	204,316	8,364	18.0%	18.9%	1.0	↑
65 - 69	165,196	6,116	14.6%	13.8%	0.9	↓
60 - 64	114,110	3,676	10.0%	8.3%	0.8	↓
55 - 59	67,478	1,854	5.9%	4.1%	0.7	↓
50 - 54	35,047	853	3.1%	2.0%	0.6	↓
45 - 49	19,340	426	1.7%	1.0%	0.6	↓
40 - 44	11,743	300	1.0%	0.7%	0.7	↓



Model Insights

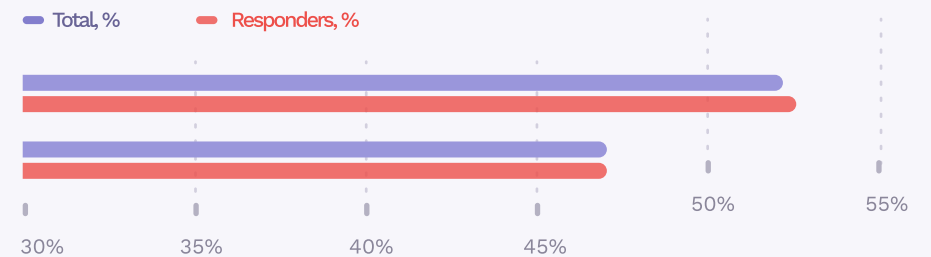
Statistics

Importance of feature to response rate: **5%**

Pet Products

Someone in the household purchased Pet Products in the past 24 months.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total, % *	Impact on response rate
No	620,631	24,223	52.06%	52.7%	1.01	↑
Yes	571,297	21,691	47.9%	47.2%	0.98	↓



Model Insights

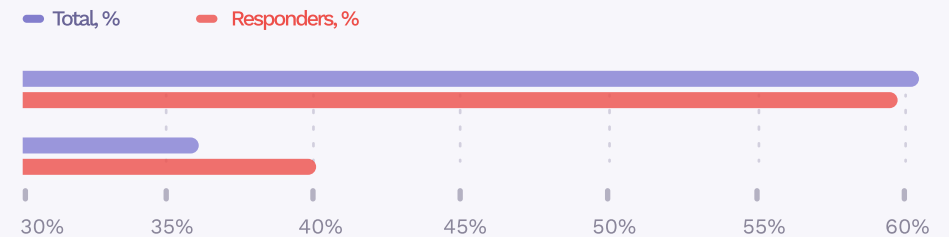
Statistics

Importance of feature to response rate: **4%**

Fine Art

Someone in the household has an interest in fine arts, such as painting, sculpting, filming, architecture, literature, or textiles.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total, % *	Impact on response rate
No	762,073	27,485	63.9%	59.8%	0.93	↓
Yes	429,855	18,429	36.06%	40.13%	1.11	↑



Model Insights

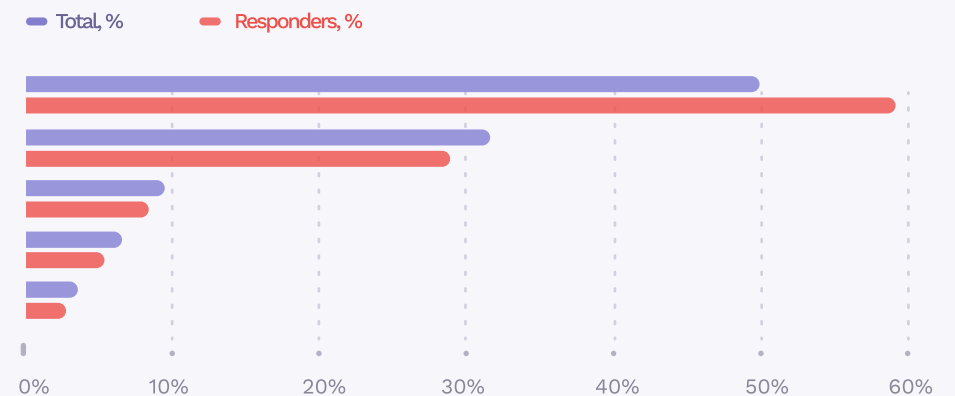
Statistics

Importance of feature to response rate: **3%**

Children

Presence/absence of children age 0-17 in the household.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total %	Impact on response rate
0 kids	304,211	1,312	49.9%	56.9%	1.13	↑
1 child	194,094	6,694	31.9%	28.8%	0.9	↓
2 kids	60,132	1,861	9.8%	8.02%	0.8	↓
3 kids	27,402	796	4.5%	3.4%	0.8	↓
4 kids	15,956	441	2.6%	1.9%	0.7	↓



Model Insights

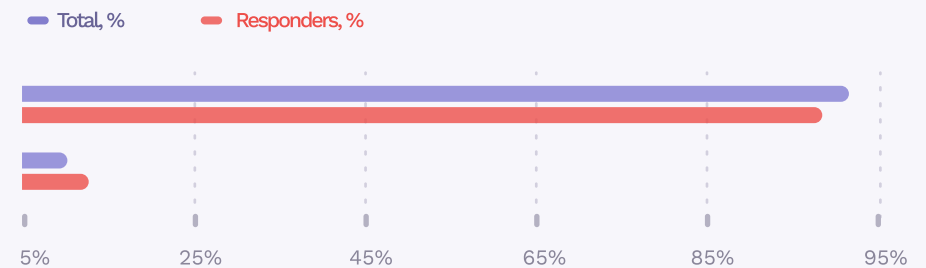
Statistics

Importance of feature to response rate: **2%**

Children's Learning And Activity Toys

Someone in the household purchased Learning or Activity Toys for children within the last 24 months.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total, % *	Impact on response rate
No	1,086,532	41,120	91.2%	89.5%	0.98	↓
Yes	105,396	4,794	8.8%	10.4%	1.18	↑



Model Insights

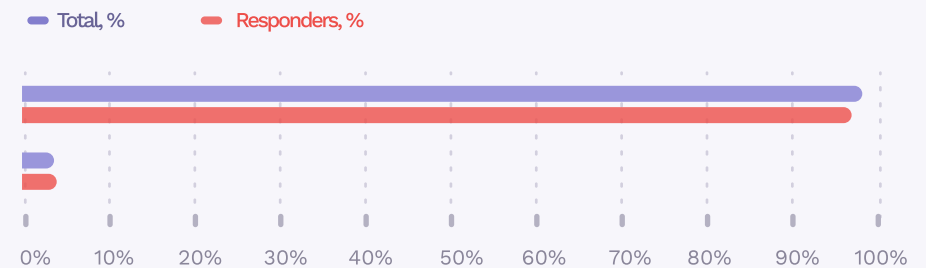
Statistics

Importance of feature to response rate: **2%**

Audio Books

Someone in the household has purchased Audio Books within the last 24 months.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total, % *	Impact on response rate
No	1,179,399	45,350	98.9%	98.7%	0.99	↓
Yes	12,529	564	1.0%	1.2%	1.16	↑



Model Insights

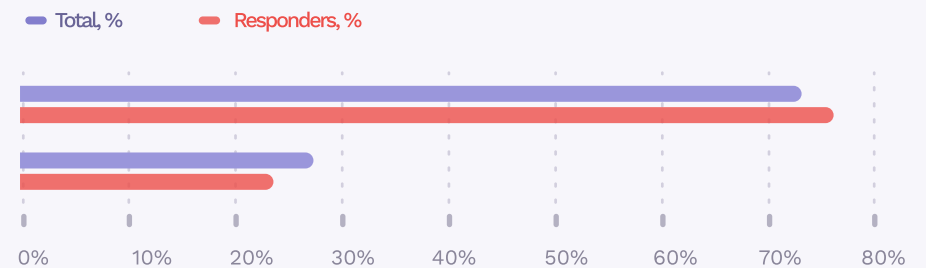
Statistics

Importance of feature to response rate: **2%**

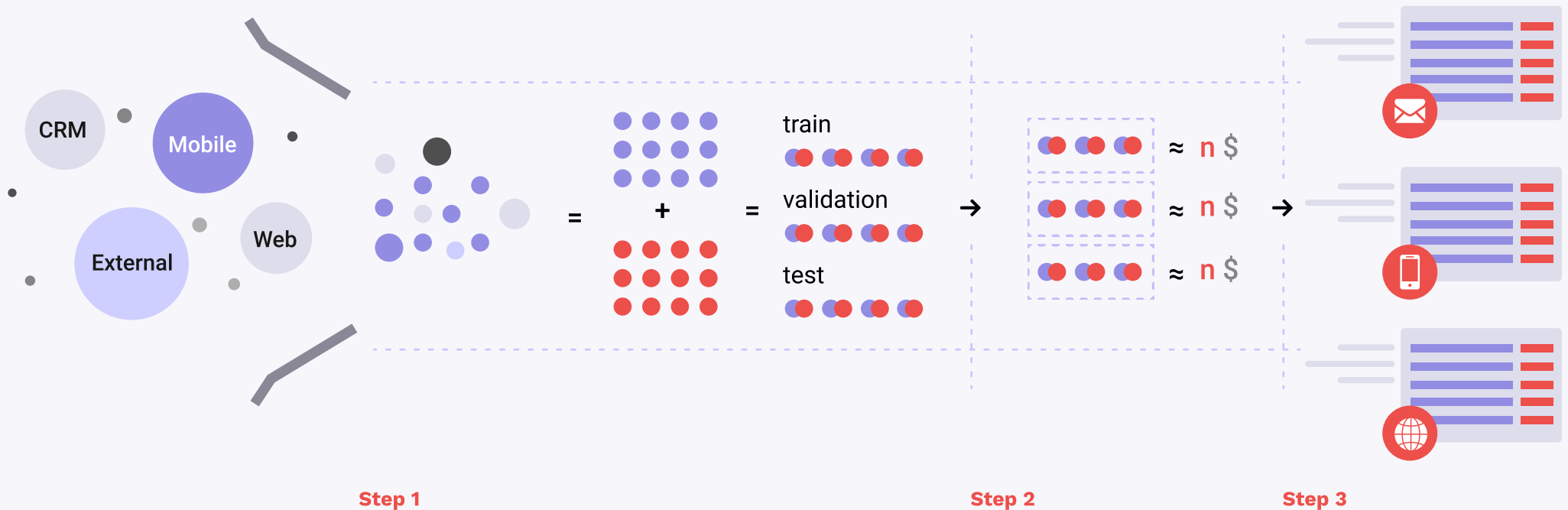
Bank Credit Cards

Indicates possession of a bank credit cards in a household.

	Total # of records	# of Responders	Total # of records, %	Responders, %	Responders, % / Total, % *	Impact on response rate
No	861,402	34,886	72.2%	75.9%	1.05	↓
Yes	330,526	11,028	27.7%	24.0%	0.86	↑



Model Process .



• Step 1

We take data that you have provided from different sources, then condition them for modeling, then combine them into a clear structure.

• Step 2

In this step, we split this information into 3 files for the purpose of training the data, validating, and testing it. This step is critical for establishing its usefulness outside of the modeling environment.

• Step 3

The resulting records are scored, and selected for delivery into the appropriate marketing channels.

Model Training Data .

Data ingestion – files recieved

2017, Miles Kimball catalog drops, both responders and non responders



17MKFL10_BESTRF-2122019-122107.csv

Campaign **Fall 2017** Responders list

77,280



17MKFL10_Mail_File_parsedNames.csv

Campaign **Fall 2017** Mailing list

2,092,150

Model Training Data .

Data Waterfall Report

In order to create the **Machine Learning model**, we prepare data by removing any outliers, and then drop or fix any missing values.

- **Step 1**

We first align any incoming data to condition it for modeling, and then confirm the target channel for modeling.

- **Step 3**

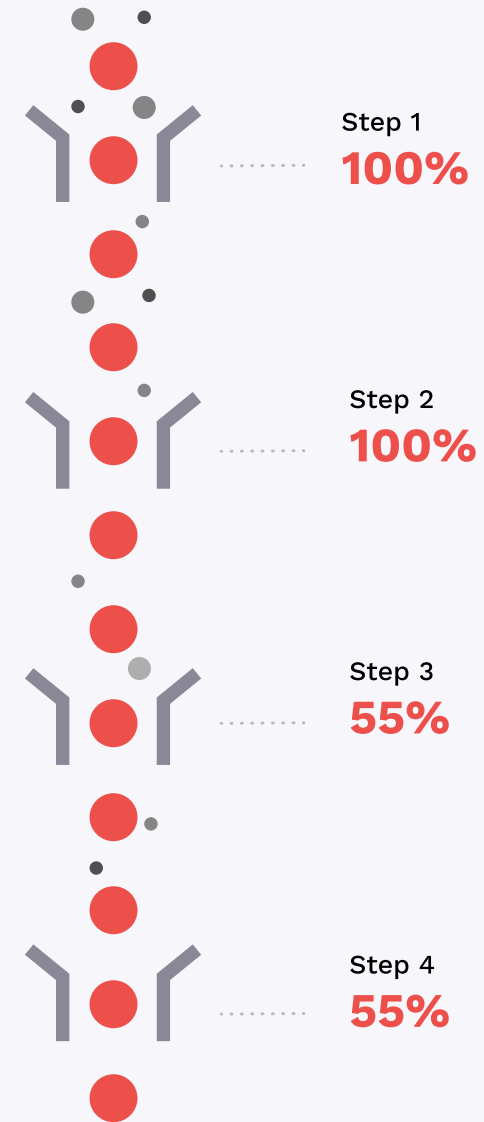
We then append 3rd party data sources for modeling. Using Human Intelligence we add these to the model, and assist the AI to create new engineered features.

- **Step 2**

Next we standardize all addresses, and prepare the structure to allow appending of external data.

- **Step 4**

Depending on the quality of the data, we may use algorithms to estimate missing values for some features. Not all features will qualify for this process.



Model Training Data .

Dataset preparation

Records passed to modeling from Step 4 - 1,191,928

Test split

- Training dataset – 715,157
- Validation dataset – 238,385
- Blind test dataset – 238,386

Combining Features

Variable Processing

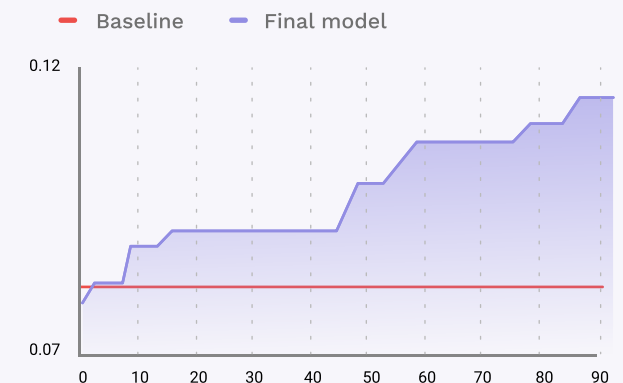
- Original number of appended features – 96
- Number of engineered feature – 417

Feature evolution

Final model parameters

Final Model

- Number of Model Improvement Stages – 91



Model Training Data

Sampling for Modeling

Total data records – 1,191,928

Training Data

715,157

3.85%
(27,576)

Mail Qty	Response Qty	Response Rate	AVG Purchase amount
715,157	27,576	3.85%	49.06\$

Training data is used to **estimate the model**. We allow Xperra's algorithms to learn and train the model with this information.

Validation Data

238,385

3.83%
(9,134)

Mail Qty	Response Qty	Response Rate	AVG Purchase amount
238,385	9,134	3.83%	48.74\$

Validation data is used to confirm the **model performance** after a draft model is created from the training data. This may happen repeatedly until a final model is confirmed.

Test Data

238,386

3.86%
(9,204)

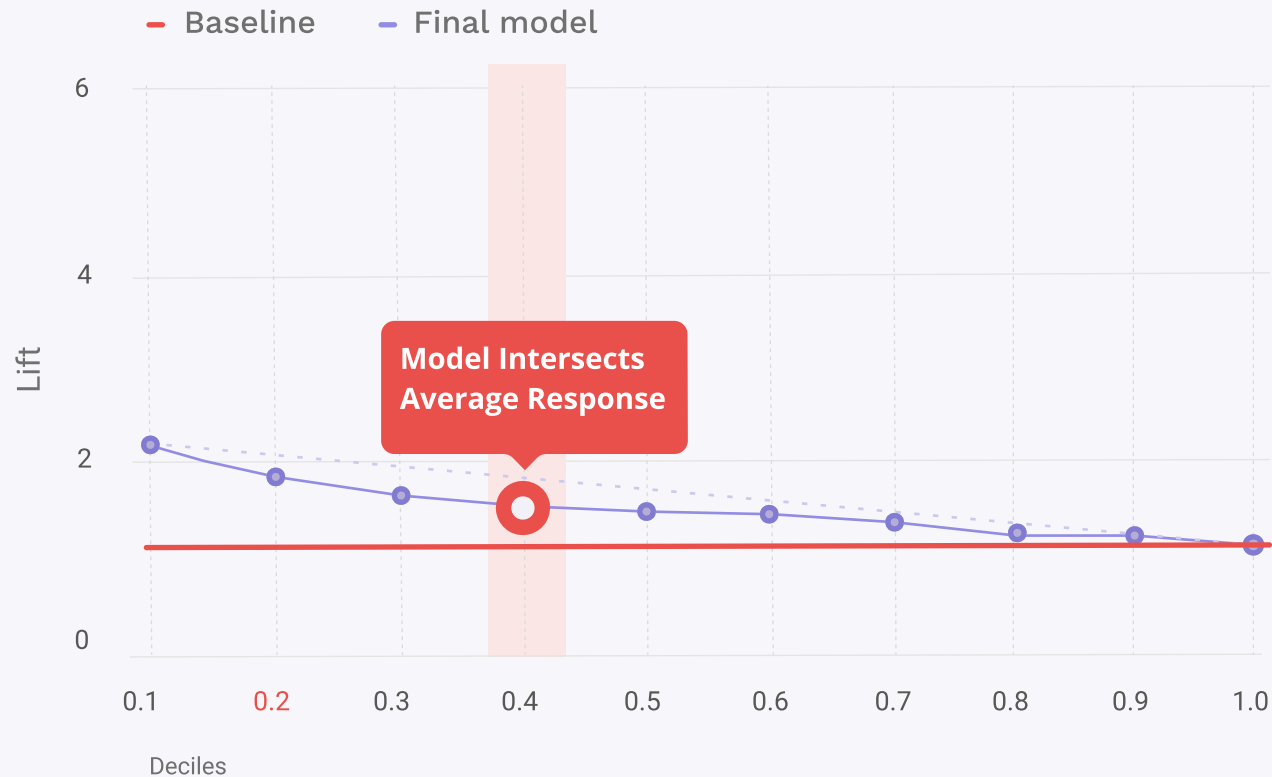
Mail Qty	Response Qty	Response Rate	AVG Purchase amount
238,386	9,204	3.86%	48.41\$

Blind testing occurs when previously unused data is scored against the model and results are evaluated for **final confirmation** that the model is ready for use.

Model Performance

LIFT

Lift is a **practical measure** of the expected model performance once we move into real world use. Lift is calculated on the blind test file and is reported below. Lift demonstrates the **degree** to which the model is **performing better** than an average unweighted model, working from the left side of the chart with the **highest performing groups** toward groups with poorer response rates.

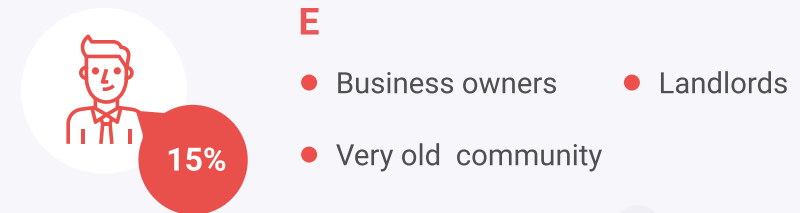
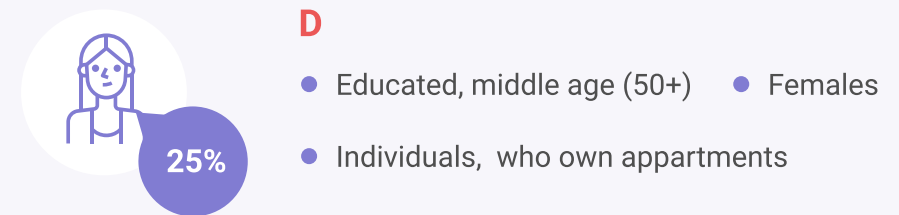
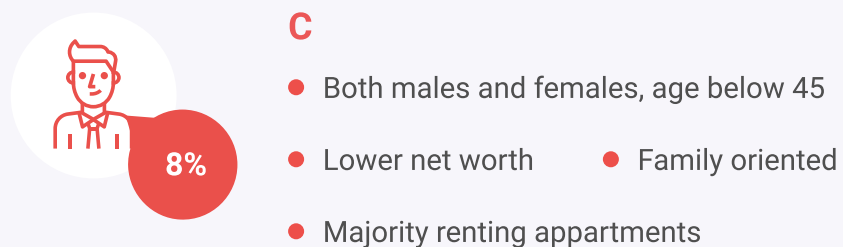
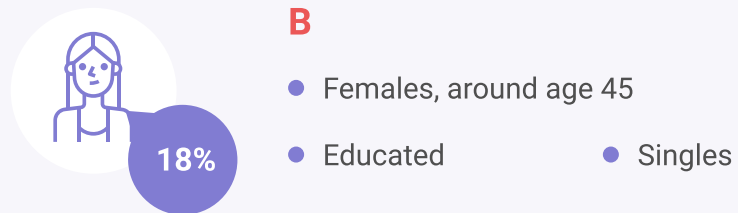
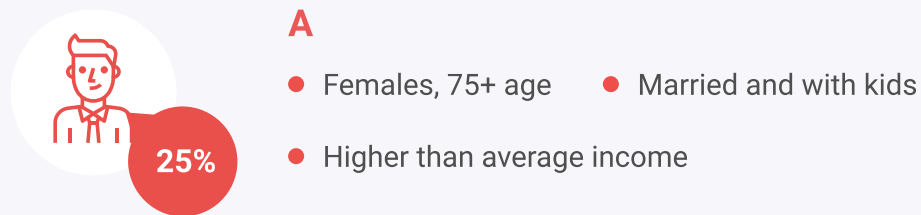


10% Deciles	2.04 Lift
20% Deciles	1.60 Lift
40% Deciles	1.40 Lift
50% Deciles	1.31 Lift
80% Deciles	1.12 Lift

Model Insights .

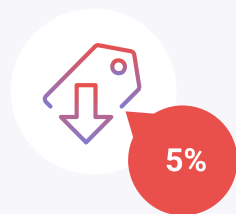
By using demographic data, the AI was able to determine that **91% of respondents were able to be narrowed to these five groups.**

These groups can then be used to select **ideal target audiences** for future marketing campaigns. Results from these campaigns can be applied back into a new model for each target group, and can result in **increasingly better results** for any subsequent campaigns.



Model Insights .

Based on demographic and customer behavior data, the AI was able to determine that **70% of respondents were found in these 5 groups.**



Price-sensitive

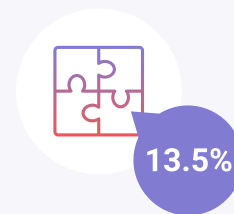
This group tends to **rent an apartment** rather than own their own home (3 times more frequently than average). They purchase only necessary items, and **buy less** than those in all other categories. They are highly unlikely to purchase memberships, art, or antiques.



Selective

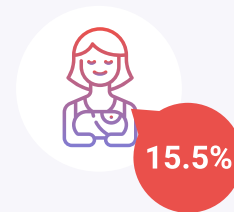
This is the oldest and **wealthiest group**. They are predominantly **business owners** who are single and do not have many children. They **enjoy spending**, but only in certain categories: antiques, art, collectibles, etc. As a result, their purchasing power is somewhat low. They are the only category that uses premium credit cards. They are **readers** (at three times the average rate).

Consistent within all clusters was a tendency toward a higher response rate for both **older** and **female respondents**. All purchasers were considerably more likely to be **homeowners**. In 4 out of 5 categories (other than the group labeled “**Selective**”), most did not own a premium American Express or Visa Credit card and fewer “**value-price**” items were purchased within the last 24 months. For all five groups, **hunting** purchases were relatively high within the Sports & Leisure category, as were “Crafts” products.



Leisure lovers

This group tends to make purchases in **hobby-driven categories**, such as gardening, books, and collectibles.



Family comfort

Predominantly **female**, this is the youngest group. They are **family-oriented** and have more children than average. They tend to purchase items related to children.




Spenders


This group enjoys shopping in **every category**; they typically purchase at a **30% higher** than average rate. They tend to have a relatively large income, while their level of savings is average.



Thank you .

 375 Park Avenue, Suite 2607
New York, NY 10152

 www.xperra.com

 (212) 660 6600

 explore@xperra.com