

## Executive Summary for the Executive Team at ShopHQ

In order to help understand and minimize return rates, we recommend having consumers fill out surveys upon returns for the items with far higher than expected and far lower than expected return rates.

- These surveys would record additional information such as time owned prior to return, shipping time/any delays, item satisfaction, customer demographics, customer expectations, and reason for return.

### Above-Average Return Rates

- Item code-names Heliodor, Datolite, Clinohumite, and Amazonite all have return rates high above the expected average based on price, duration, and showings.

Item Name	Expected Return Rate	Actual Return Rate
Heliodor	15.1%	17.9%
Datolite	18.2%	20.9%
Clinohumite	22.5%	25.0%
Amazonite	22.9%	25.0%

- Example survey questions could include the following:
  - When was the item purchased, and what is the return date?
  - Were there any issues with shipping? Any delays?
  - What is the primary reason for returning?
  - Customer demographics: age, ethnicity, gender, purchase history, geographic location, etc.

### Below-Average Return Rates

- Item code-names Diopside, Amethyst, Andradite, Barite, and Goshenite all have return rates well below the expected based on price, duration, and showings.

Item Name	Expected Return Rate	Actual Return Rate
Diopside	10.7%	4.4%
Amethyst	10.6%	7.3%
Andradite	8.5%	5.6%
Barite	10.7%	8.0%
Goshenite	28.4%	26.3%

- Example survey questions could include the following:

- Would you purchase this item again or recommend it?
- Have you ever had an issue with this product? If so, what?
- What was your reason for choosing this product over alternatives?
- Customer demographics: age, ethnicity, gender, purchase history, geographic location, etc.

Of the important details (price, showings, and duration), price is by far the largest driver for return rate accounting for nearly 50% of the variation - but it does not explain everything.

**Figure 1**

Price in \$	Average Return Rate
1000+	42.7%
500-1000	35.8%
200-500	33.6%
100-200	26.9%
50-100	21.4%
20-50	15.1%
10-20	10.9%
0-10	6.6%

*(Figure 1 - The average return rate based on Price in \$ is 24%. Price Bins in green are above this marker; those in red are below.)*

- Duration and showings also impact return rates:
  - Duration
    - Accounts for 14% of the variation in return rates
    - Each additional unit of duration raises return rate by 3.4%
  - Showings
    - Accounts for 13.6% of the variation in return rates
    - Each additional unit of showing raises return rate by 5.8%

Our data was week-level transactional data from roughly 300,000 data points, each containing a particular Codename, Price Bin, and Week combination. Only the 30 most frequently appearing Codenames were examined with the rest being combined as “other”.

## Technical Discussion

I used a Random Forest model for my testing, fitted with Predicted ReturnRate against all other parameters. The model has a predicted error of ~0.064. It is sorted in ascending order based on Deviation.

Codename	Actual Average	Predicted Average	n	Deviation
<i>Diopside</i>	4.4	10.7	2457	-6.3
<i>Amethyst</i>	7.3	10.6	660	-3.3
<i>Andradite</i>	5.6	8.5	744	-2.9
<i>Barite</i>	8	10.7	724	-2.7
<i>Goshenite</i>	26.3	28.4	613	-2.1
<i>Ametrine</i>	21.1	22.4	826	-1.3
<i>Opal</i>	17.1	18	932	-0.9
<i>Tourmaline</i>	19.6	20	2385	-0.4
<i>Grossular</i>	20.2	20.3	1131	-0.1
<i>Beryl</i>	19	19	648	0
<i>Rhodochrosite</i>	16.2	16	1578	0.2
<i>Coral</i>	10.3	10.1	643	0.2
<i>Alexandrite</i>	21.7	21.4	791	0.3
<i>Garnet</i>	19.5	19.2	815	0.3
<i>Jasper</i>	19.1	18.8	770	0.3
<i>Axinite</i>	17.4	17	1636	0.4
<i>Zircon</i>	34.8	34.1	870	0.7
<i>Tigereye</i>	20.4	19.6	1062	0.8
<i>Apatite</i>	17.3	16.5	1418	0.8
<i>Andalusite</i>	20.3	19.4	1106	0.9
<i>Vandium</i>	19.5	18.4	2036	1.1
<i>Diaspore</i>	20.4	19.2	663	1.2
<i>Kornerupine</i>	14.9	13.6	1083	1.3
<i>Thulite</i>	21.2	19.7	1364	1.5
<i>Emerald</i>	25.2	23.6	2451	1.6
<i>Amazonite</i>	25	22.9	1240	2.1
<i>Clinohumite</i>	25	22.5	883	2.5
<i>Datolite</i>	20.9	18.2	626	2.7
<i>Heliodor</i>	17.9	15.1	736	2.8

Here is a scatter plot showcasing the same data; the actual average return rate vs. the predicted average return rate:

