High Price Model 1.0

Content

Introduction

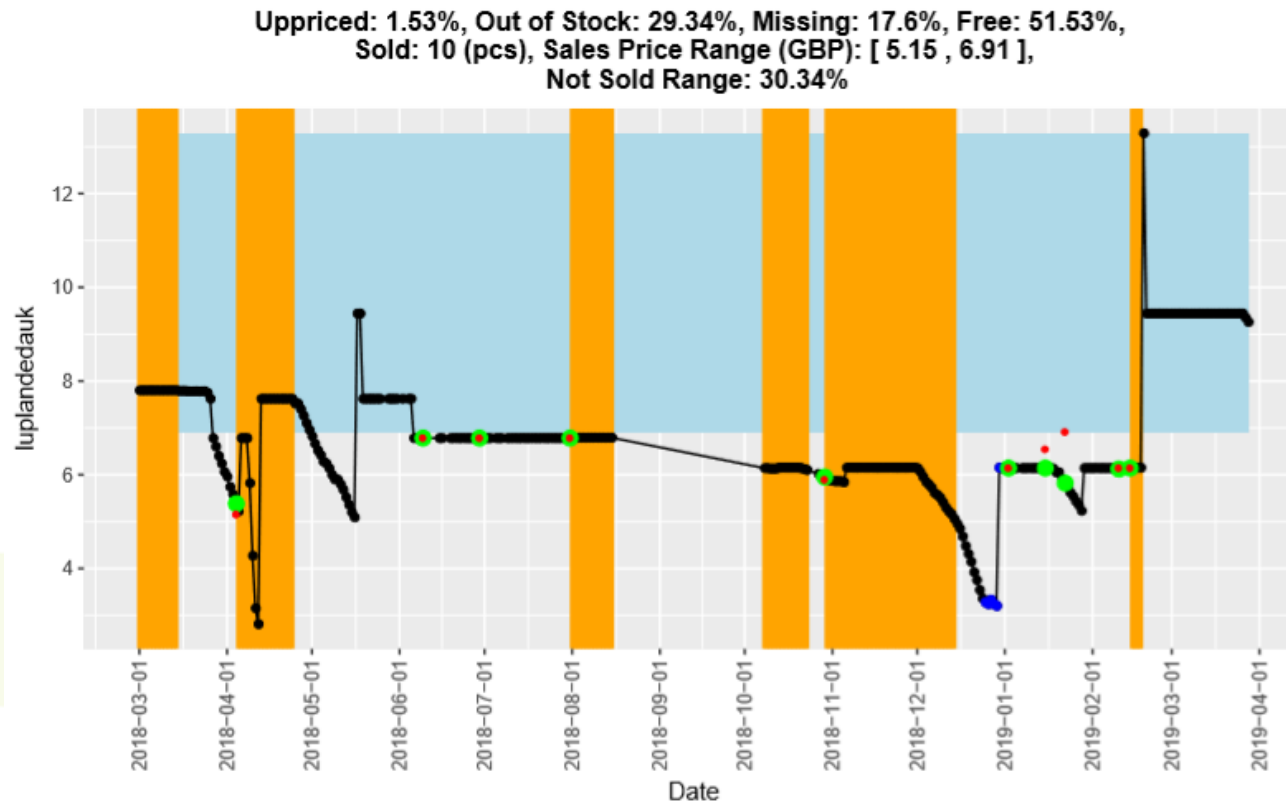❏ The aim of the model is to **handle overpriced situations**.

What is High Price?

❏ A price on an item is significantly harder to sell.

❏ There is a price competition and we unfortunately can not see other seller's sales prices. So we can **make decision based on our own price and selling history**.

Technical restriction:

❏ The company doesn't store an item forever. There is a limited keep time.

→ We don't want to list items on prices where we won't sell them with high probability and finally we will cull them. The **goal is to determine marketable prices**.

Example

Upriced: 1.53%, Out of Stock: 29.34%, Missing: 17.6%, Free: 51.53%,
Sold: 10 (pcs), Sales Price Range (GBP): [5.15 , 6.91],
Not Sold Range: 30.34%

❑ **Considering a Sales Event frequently the saved LULP differs from the realized Sales Price (red point)**
❑ **StockOut periods are kind of black spots**

Price Range without sales event
Out of Stock
Lowest Used Landed Price (LULP)
LULP at Sales Event
Sales Price
Uppriced

Model Idea
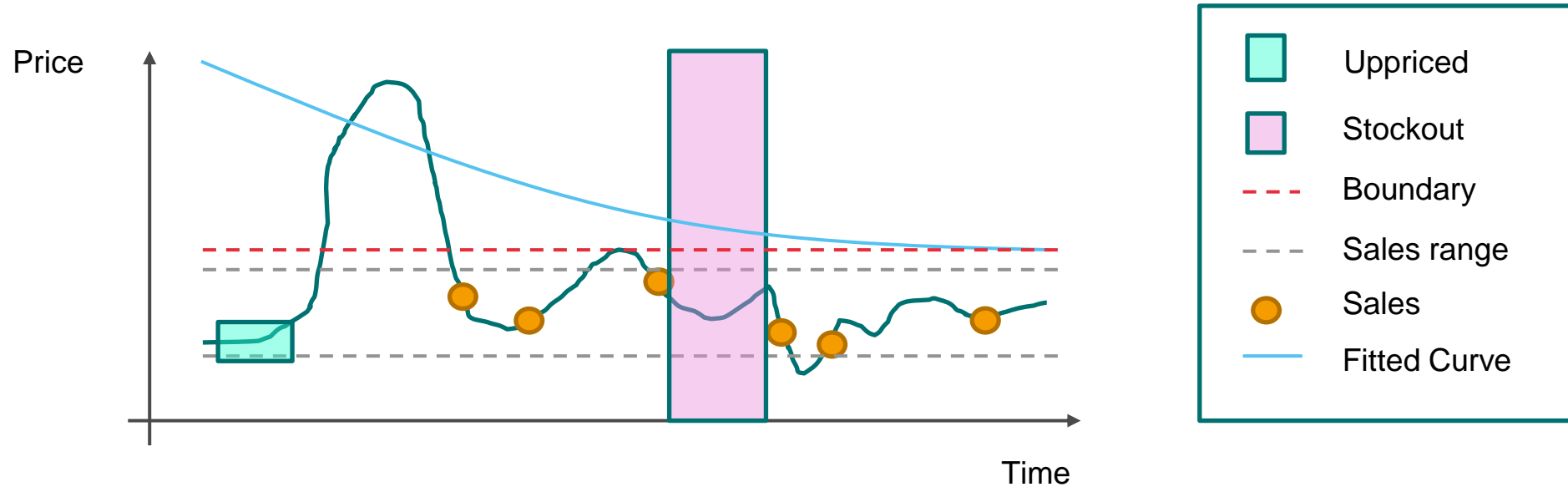
- ❏ The idea is to fit a curve representing a conservative upper limit on the price of an item based on it's price and sales history.

- ❏ The Curve depends on:

  - ❏ The highest Sales Price,

  - ❏ Observation period (the start is max(01.03.2018 [first saved data], first date on stock), StockOut days are not considered here)

  - ❏ Sold amount (during the observation period)

  - ❏ The proportion of days during the observation period where the price was above the highest Sales price
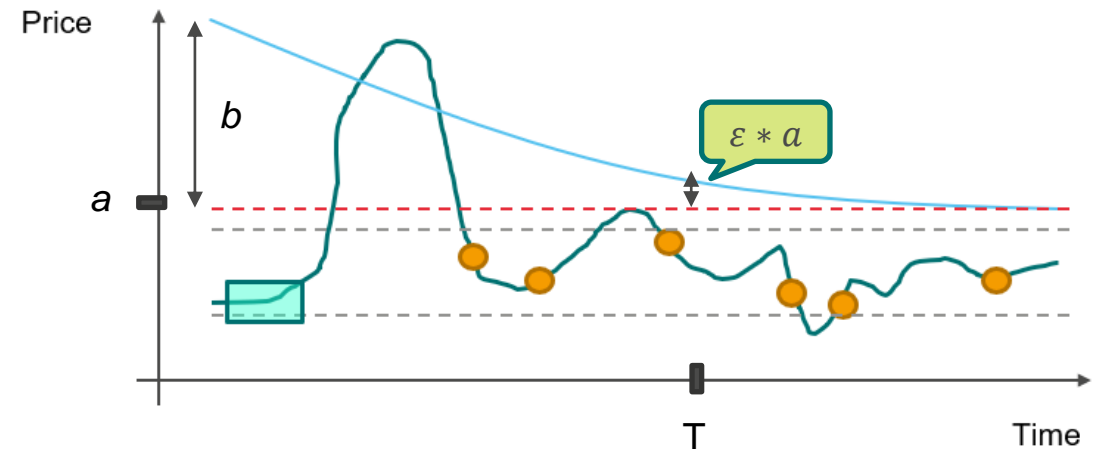
Assumptions and processes

❏ The model assumes at **least one sales event** during the observation period.

❏ Even if we saved LULP during stock out periods we can not decide about the suitability of the price, because we can not observe the outcome (sales). That is why **stock out periods are filtered out**.

❏ **Uppriced periods are taken into account**, though the selling rate during these periods is significantly lower, it is not zero. An item can go in and out multiple times in a day.

❏ Observation period is max(01.03.2018 [first saved data in Spark], first date on stock)
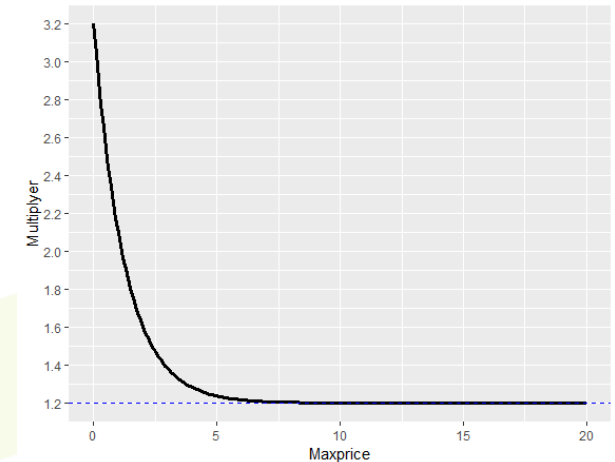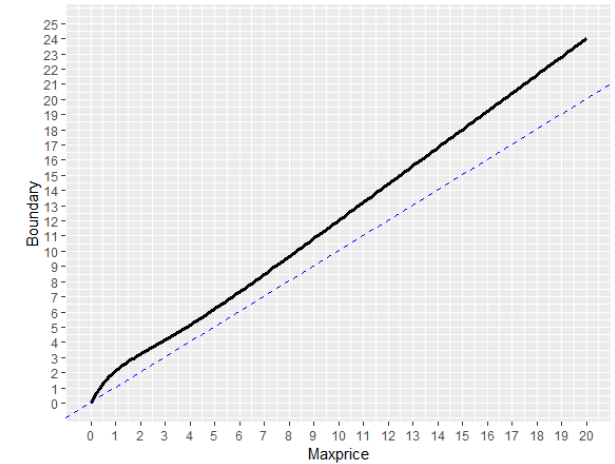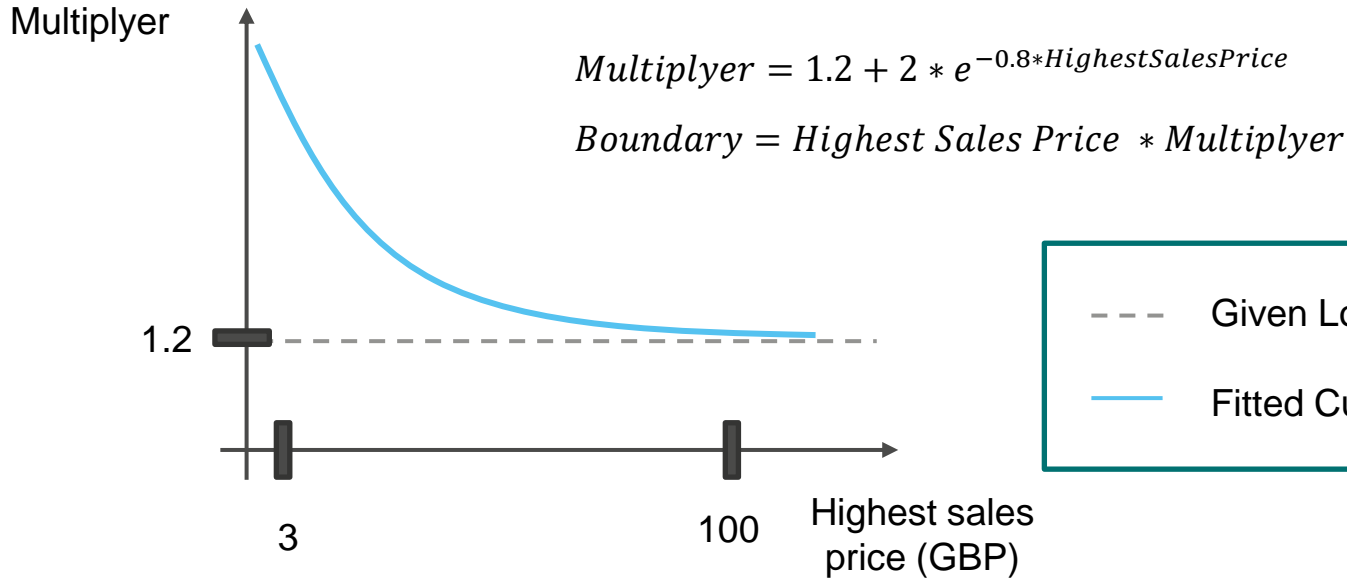
High Price Model

❑ A curve is fitted with exponential decay (see the formula and the blue curve on the chart).

❑ „**a**" parameter represent a lower bound for the curve.

❑ „**b**" parameter means the dilation of the curve.

❑ „**c**" parameter determines the curvature.

❑ „**ε**" parameter is an addon applied on „a", which will determine for the curve where to land after a given time (T). ε is fixed in time T and depends on two variables:

    1. **Not Sold Rate (NSR)**: The proportion of days during the observation period where the price was above the highest Sales price.

    2. **Sales Count** (pcs)

❑ „**T**" parameter means an expert based, but statistically reliable time as of ε is calibrated. In this case 250 days was chosen.

$$f(t) = a + b * e^{-c*t}$$

„a" Parameter (lower bound)

Multiplyer

$$Multiplyer = 1.2 + 2 * e^{-0.8*HighestSalesPrice}$$

$$Boundary = Highest\ Sales\ Price\ * Multiplyer$$

1.2

- - - Given Lower bound

—— Fitted Curve

3   100   Highest sales
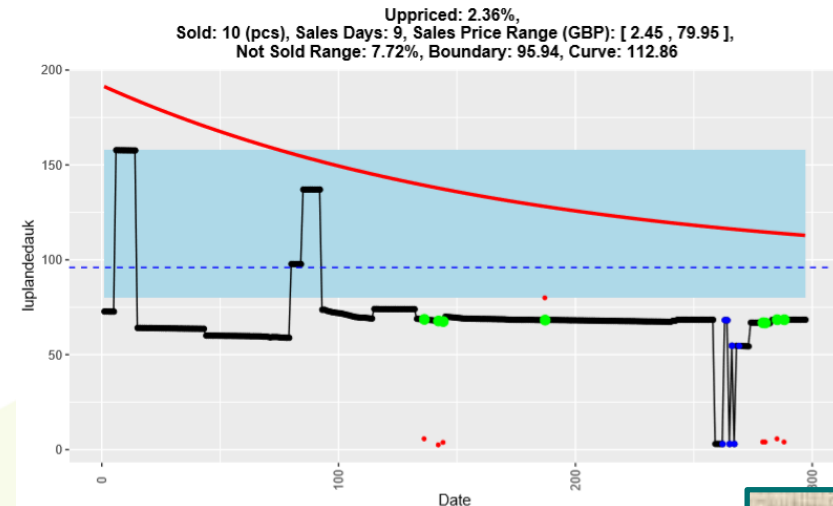   price (GBP)

❑ **Boundary** means the lower bound for the curve („**a**" parameter of the formula). The Boundary is always somewhat higher than the Highest Sales Price not to force the price strictly into a given range, let a little freedom to be able to move upward.

❑ Where the **Multiplyer** is a function of the **Highest Sales Price**, it means that for lower observed Highest Sales Price the Multiplyer is higher and in terms of higher observed Highest Sales Price it is lower but not less than a given number (1.2). So the upper limit for a price could not be less than this value multiplied by the Highest Sales Price (Boundary). The addon applied is at least 20% of the Highest Sales Price.

„b" Parameter (dilation)

❑ b means the dilation of the curve, normally the curve takes the value of $a + 1$ at time 0, so we need to delate the curve to start from the highest observed Price (LULP) or from some high enough price level to express better the uncertainty of the upper bound of the sales price range on the short run.

❑ $b = \max(\max(LULP), 2 * Boundary) - Boundary$

Uppriced: 0%,
Sold: 9 (pcs), Sales Days: 6, Sales Price Range (GBP): [19.61 , 55.86],
Not Sold Range: 96.41%, Boundary: 67.03, Curve: 68.03

Uppriced: 2.36%,
Sold: 10 (pcs), Sales Days: 9, Sales Price Range (GBP): [2.45 , 79.95],
Not Sold Range: 7.72%, Boundary: 95.94, Curve: 112.86
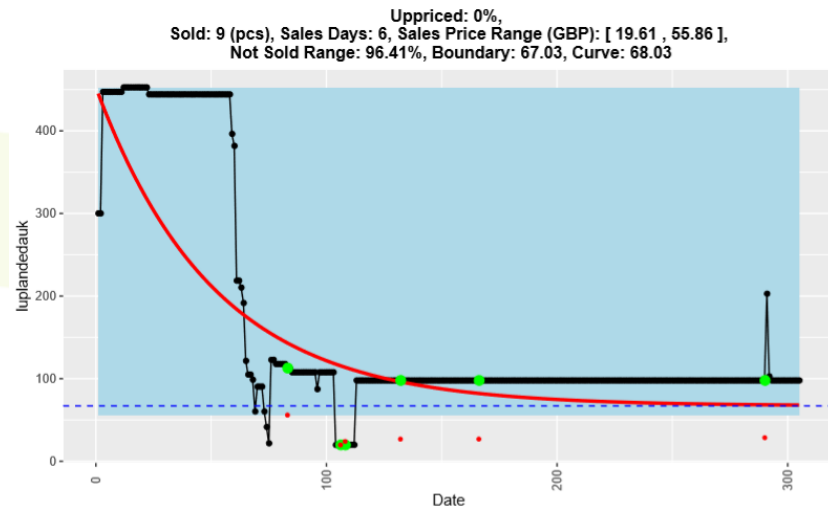
„ε" Parameter

- [ ] „ε" parameter is an addon applied on „a", which will determine for the curve where to land after a given time (T). ε is fixed in time T and depends on two variables:

  1. **Not Sold Rate (NSR)**: The proportion of days during the observation period where the price was above the highest Sales price.
  2. **Sales Count** (pcs): it was transformed by the following formula:

  SoldTrans $= 1 - 1/(1 + \exp(-5 * 0.55 * Soldqty))$

- [ ] Polynomial regression was fit based on expert based reference points:

- [ ] $\varepsilon = 0.1772 - 0.6903 * TransfSold + 0.2247 * TransfSold^2 - 0.1876 * TransfSold^3 + 0.1025 * TransfSold^4 - 0.5216 * NSR + 0.111 * NSR^2$
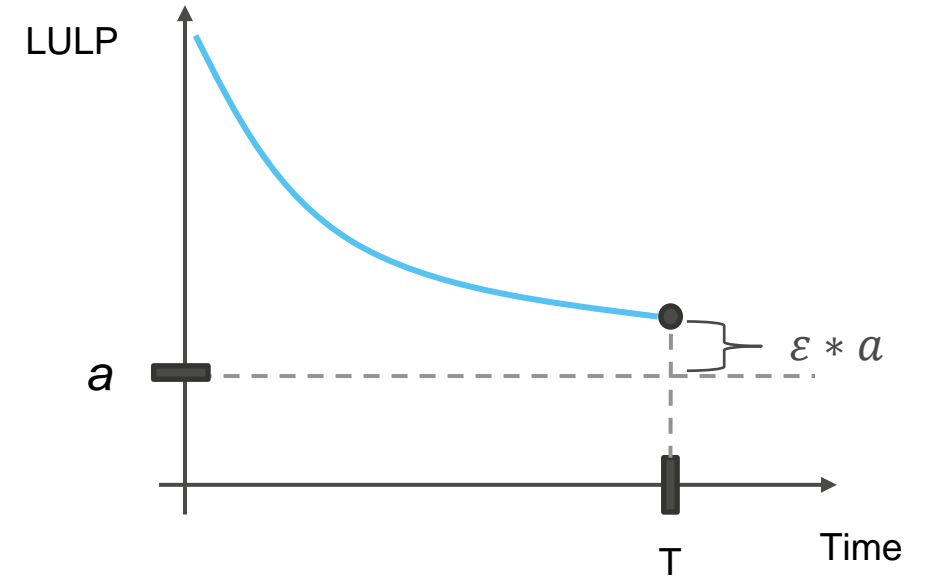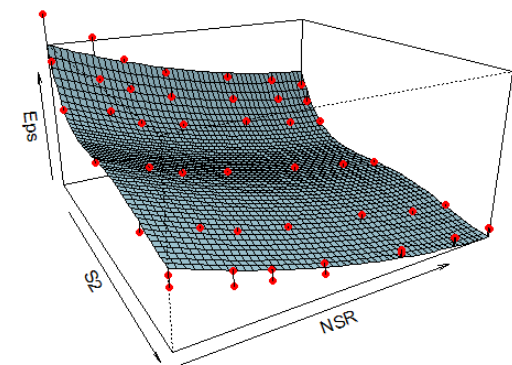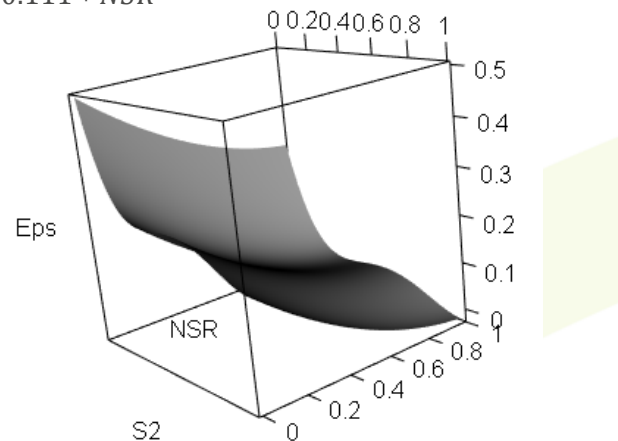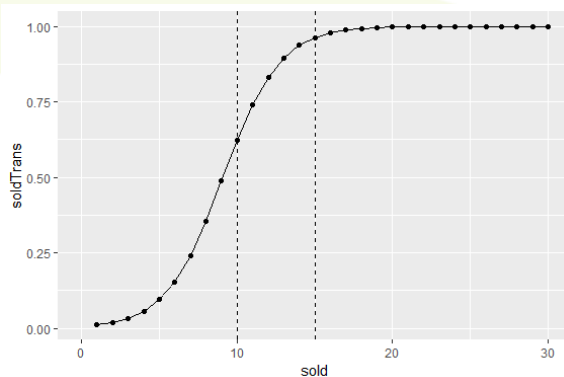
„c" Parameter (curvature)

❑ The „c" Parameter (curvature) can be determined by knowing the „a" and „b" Parameters, T is fixed, lets say it is 250 (days). The „ε" parameter is also needed.

$$a + b * e^{-c*T} = (1 + \varepsilon) * a$$

$$c = \frac{-\ln\left(\frac{\varepsilon * a}{b}\right)}{T}$$

$$c = \frac{-\ln\left(\frac{\varepsilon * a}{b}\right)}{max(250, observation\ period)}$$

Trial details

❏ Sample selection criteria:

- Item: book

- Yearly sold quantity ≥ 5

- There is at least 1 pcs on stock

- Actual Listed Price > 1.1 * [1 year highest sales price]

- Item is not „PENNY"

❏ Sample was created at 07.06.2019

- 13 308 Asins are in the total sample, it was split into 2 parts randomly in 50-50% proportion

- **Treatment group**: 6 654 Asins, **Control group**: 6 654 Asins

❏ Trial Started: 08.06.2019

Treatment and Control group similarity test

❏ Population Stability Index (PSI) was used to test similarity

❏ $PSI = \sum_{i=1}^{n}((Treatment_i\% - Control_i\%) * (\ln(\frac{Treatment_i\%}{Control_i\%})))$

  o Where n is the number of groups created.

❏ PSI < 0.1 is good, 0.1 ≤ PSI < 0.2 means warning level, PSI ≥ 0.2 is critical.

❏ Test results are good. There is high similarity between Treatment and Control groups.

Sold	Control count	Control %	Treatment count	Treatment %	PSI
[1, 6)	1 827	27,46%	1 834	27,56%	0,00
[6, 8)	2 012	30,24%	2 063	31,00%	0,00
[8, 11)	1 336	20,08%	1 293	19,43%	0,00
[11,Inf]	1 479	22,23%	1 464	22,00%	0,00
Total	6 654	100%	6 654	100%	0,00

Accepted	Control count	Control %	Treatment count	Treatment %	PSI
[0, 8)	1 915	28,78%	1 911	28,72%	0,00
[8, 12)	2 038	30,63%	2 018	30,33%	0,00
[12, 20)	1 687	25,35%	1 696	25,49%	0,00
[20,Inf]	1 014	15,24%	1 029	15,46%	0,00
Total	6 654	100%	6 654	100%	0,00

SalesRank	Control count	Control %	Treatment count	Treatment %	PSI
NA	3	0,05%	3	0,05%	0,00
[0, 50000)	39	0,59%	72	1,08%	0,00
[50000, 250000)	429	6,45%	410	6,16%	0,00
[250000, 500000)	848	12,74%	767	11,53%	0,00
[500000,1000000)	1 753	26,35%	1 745	26,22%	0,00
[1000000,2000000)	2 460	36,97%	2 556	38,41%	0,00
[2000000,6000000)	1 089	16,37%	1 064	15,99%	0,00
[6000000, Inf]	33	0,50%	37	0,56%	0,00
Total	6 654	100%	6 654	100%	0,01

Arrived	Control count	Control %	Treatment count	Treatment %	PSI
[0, 8)	1 574	23,65%	1 540	23,14%	0,00
[8, 12)	1 824	27,41%	1 868	28,07%	0,00
[12, 20)	1 673	25,14%	1 641	24,66%	0,00
[20,Inf]	1 583	23,79%	1 605	24,12%	0,00
Total	6 654	100%	6 654	100%	0,00

Trial Assessment

- ❑ Trial Started: 08.06.2019
- ❑ 10 weeks Statistics:

	Group	Before						After					
		-5th week	-4th week	-3rd week	-2nd week	-1st week	Total	1st week	2nd week	3rd week	4th week	5th week	Total
Sold quantity	Treatment	1 090	921	825	660	777	4 273	1 163	1 210	1 241	1 213	1 287	6 114
	Control	1 066	914	867	703	769	4 319	936	963	1 066	1 096	1 094	5 155
	Treat./Cont.	102,25%	100,77%	95,16%	93,88%	101,04%	98,93%	124,25%	125,65%	116,42%	110,68%	117,64%	118,60%
Sold GBP	Treatment	7 253,00	6 612,64	6 160,75	4 400,66	6 675,66	31 103	12 499,31	12 019,96	11 231,93	10 925,90	11 313,22	57 990
	Control	7 194,80	5 764,12	5 719,04	4 966,07	6 962,13	30 606	10 667,17	9 191,93	9 834,68	9 369,73	9 871,98	48 935
	Treat./Cont.	100,81%	114,72%	107,72%	88,61%	95,89%	101,62%	117,18%	130,77%	114,21%	116,61%	114,60%	118,50%
Arrived	Treatment	3 098	3 431	3 368	3 014	2 362	15 273	2 283	2 266	2 121	2 002	2 054	10 726
	Control	3 044	3 404	3 518	3 205	2 528	15 699	2 239	2 135	2 152	2 054	2 020	10 600
	Treat./Cont.	101,77%	100,79%	95,74%	94,04%	93,43%	97,29%	101,97%	106,14%	98,56%	97,47%	101,68%	101,19%
Accepted	Treatment	2 395,00	2 744,00	2 787,00	2 459,00	1 854,00	12 239	1 721,00	1 612,00	1 531,00	1 492,00	1 481,00	7 837
	Control	2 435,00	2 727,00	2 866,00	2 598,00	1 980,00	12 606	1 633,00	1 575,00	1 555,00	1 459,00	1 457,00	7 679
	Treat./Cont.	98,36%	100,62%	97,24%	94,65%	93,64%	97,09%	105,39%	102,35%	98,46%	102,26%	101,65%	102,06%

Sales improvement: +19%

❑ **Supply dropped, but not significantly different by groups**

Sold quantity

Accepted

Arrived