Promo Depth Calculation Rulebook

Specification for Computing Final Promo Depth in Shelftrak Price & Promotion Data

0. Input Format

Each row contains a single string in the Excel column “Price & Promo”, generally in this form:

```
"{BASE PRICE} - {PROMO_TEXT}"
```

The left side is the captured base unit price (if available). The right side is the raw promotional text from the shelf label.

Examples:

```
68 - Save $80, now $68
33.33 - SGD 60 for any 2 bottles
0 - UP TO 25% OFF
```

The goal is to convert each row into a promo depth expressed as a number from 0 to 100. If a discount cannot be reliably quantified, the depth is set to 0.

1. Pre-processing

1.1 Split and Extract

Split the input string on the first “-” into BASE_PART and PROMO_TEXT_RAW. Extract the first numeric token from BASE_PART:

• If > 0, this is the base unit price P.

• If 0 or missing, set P = None (unknown base price).

1.2 Number Normalisation in Promo Text

In PROMO_TEXT, normalise numbers as follows:

Remove thousands separators: 2,199 → 2199.

Convert decimal commas to dots: 12,50 → 12.50.

1.3 Text Cleaning

Create two versions of the promo text:

promo: cleaned text.

promo_lower: lowercase version for keyword detection.

Fix common OCR issues, e.g.:

"but 2 get 1" → "Buy 2 get 1".

"grt" → "Get".

1.4 Empty Promo Text

If PROMO_TEXT is empty or whitespace only, promo depth = 0.

2. Hard Exclusions (Always 0%)

The following patterns always yield a promo depth of 0%:

2.1 Device + Sticks/Cartons Bundles

If promo_lower includes “device” AND any of: "stick", "sticks", "carton", "pod", "pods", "capsule", "capsules".

2.2 Gift-with-Purchase (e.g. Absolut Free Citron)

If promo contains “receive a free bottle” (e.g. Absolut + free Citron 37.5cl).

2.3 Pure Rating / Award Messages

If promo contains rating keywords ("points", "decanter", "world wine awards", "iwsc", "james suckling", "wine spectator", "robert parker", "medal")

AND does not contain promo keywords ("save", "off", "buy", "get", "free", "was", "now", "discount", "offer", "deal", "pay", " for ", "%", "x", "@").

2.4 Combos Without Explicit Discount

If promo contains “combo” plus “+” or “&” or “ and ” but not “get”, “free”, “glass”, “backpack”, “trolley”, “gift”, “save”, or “off”.

2.5 Free Merchandise Only

Promos that only give free glasses/trolleys/backpacks/etc. with no price reduction.

2.6 Solo / Only Price Without Promotion

Text like “Solo X€” or “Only X€” without “save/off/was/now/for”.

3. Special Global Rules

3.1 “UP TO X% OFF”

If promo contains “UP TO X% OFF”, promo depth = X, even if base price P is unknown.

3.2 SGD 60 Any 2 Bottles

If base price P is known and promo matches “SGD 60 for any 2 bottles” or similar:

```
discount = 1 - 60 / (2 * P)
```

3.3 “60 SGD OFF 2 Bottles”

If base price P is known and promo says “60 SGD OFF 2 Bottles” (or similar):

```
discount = 60 / (2 * P)
```

4. Candidate Discounts

Every rule that yields a valid discount adds a candidate to a list:

```
discounts = [d1, d2, ...]
```

Final promo depth will be the smallest discount in this list (see Section 10).

4.1 Direct Percent Discounts

Any “X%” is a direct candidate: “30% off” → discount = 30.

4.2 Chinese “■” Discounts

For “z■” notation (pay z/10 of original price):

```
discount = (1 - z / 10) * 100
```

5. Price Reduction & Save / Off Patterns

5.1 “Was A, now B”

When promo explicitly states old and new price:

```
discount = 1 - (B / A)
```

5.2 “Y (was X)” with Currency

With exactly two numbers and a currency symbol, old = X, new = Y (X > Y > 0):

```
discount = 1 - (Y / X)
```

5.3 “Save S, now N”

When both saving and new price are given:

```
old = N + S
discount = S / (N + S)
```

5.4 “Save S” with Base Price P

If promo states “Save S” and base price P is known but “now” is not:

```
old = P + S
discount = S / (P + S)
```

5.5 “S off” with Base Price P

For “S off” (money, not %), with base price P:

```
old = P + S
discount = S / (P + S)
```

5.6 “S■■■”

Japanese discount notation with base price P:

```
old = P + S
discount = S / (P + S)
```

6. Multi-buy Rules

6.1 Multi-Buy Without Prices

Applies when there is no currency symbol in promo.

6.1.1 “A for B” / “AxB” / “A x B”

Treat larger number as total units, smaller as paid units:

```
total_units = max(A, B)
paid_units  = min(A, B)
discount    = 1 - paid_units / total_units
```

6.1.2 “2 for get 1”

Two units for the price of one: discount = 1 - 1/2 = 50%.

6.1.3 Special “3 for 5” Approximation

Approximate discount as 1 - 3/5 = 40%.

6.2 Multi-Buy With Prices (Currency Present)

6.2.1 “N for T” with Base Price P

```
promo_unit = T / N
if promo_unit < P:
    discount = 1 - (promo_unit / P)
```

6.2.2 “BUY N PAY T”

Same formula as above.

6.2.3 “X each” / “Unit Price X”

```
promo_unit = X
if promo_unit < P:
    discount = 1 - (promo_unit / P)
```

6.2.4 “Buy 1 for A, 2 for B”

Base unit = A, promo unit = B / 2:

```
promo_unit = B / 2
if promo_unit < A:
    discount = 1 - (promo_unit / A)
```

6.2.5 “BUY N FOR P SAVE S”

```
old_total = P + S
discount  = S / old_total
```

7. Buy X Get Y / Free Promotions

7.1 General Rule

“Buy X Get Y (Free)” → receive X + Y units, pay for X:

```
discount = 1 - X / (X + Y)
```

7.2 “X + Y Free”

Same structure as above.

7.3 “Leve X Pague Y”

```
discount = 1 - Y / X
```

7.4 Business Overrides

Override generic formulas to match business rules:

Buy any 2 get 1 free → ~33.33%.

Buy any 3 get 1 free → 25%.

Buy any 4 get 1 free → 25%.

Buy 2 Get 3 → ~33.33%.

8. Buy-1-Get-2nd and Multi-tier Offers

8.1 “Buy 1 Get 2nd for X” with Base Price P

Two units cost P + X, effective unit price = (P + X)/2:

```
promo_unit = (P + X) / 2
if promo_unit < P:
    discount = 1 - (promo_unit / P)
```

8.2 “Buy 1 For A on purchase of 2” with Base Price P

Interpreted as both bottles at price A, compare with base price P:

```
discount = 1 - P / (2 * A)
```

8.3 Multi-tier Percent Offers

For offers like “Buy 1 save 10%, buy 2 save 20%”: collect all percentages (10, 20) and later choose the smallest.

9. Dataset-Specific Clarifications

9.1 Device + Sticks/Cartons

Always 0% promo depth (bundle).

9.2 Absolut Free Citron Bottle

Gift-with-purchase only; 0% depth.

9.3 Award-only Lines

Scores/medals without promo keywords: 0% depth.

9.4 Complex Combos Without Reference Pricing

When combos have no “save/off/was/now” references, they can be treated as 0% by design.

9.5 Solo / Only Price Statements

“Solo only X€” or “Only 10€” without other promo signals → 0%.

10. Final Aggregation & Output

10.1 Candidate Set

```
discounts = [d1, d2, d3, ...]
```

10.2 No Candidates

If discounts is empty → promo depth = 0%.

10.3 Multiple Candidates – Use the Lowest Discount

Business rule: when multiple candidate discounts exist, always use the smallest (lowest) discount value.

```
if not discounts:
    depth = 0
else:
    depth = min(discounts)
    depth = max(0, min(depth, 100))  # clamp to [0, 100]
```

10.4 Rounding & Presentation

Clamp depth to [0, 100] and typically round to two decimals (e.g. 33.33).

10.5 QA Highlighting (Optional)

Highlight promo depths > 85% in red or flag for manual review as potentially erroneous.

> This rulebook is the single source of truth for promo depth logic and should match the implementation in code and any LLM-based app prompts.