

# Problem Statement



flyer_name	product_name	unit_promo_price	uom	least_unit_for_promo	save_per_unit	discount	organic
week_1_page_1	Ground Sirloin or Round	3.99	lb	1	4.00	0.50	0
week_1_page_1	Blueberries	2.5	1 Pint	2	3.49	0.58	0
week_1_page_2	Morning Rounds	2.00	6 Pack	1	0.50	0.11	0

# Data Processing: Feature Detection



Colors, and brushed or polished metal.  
See your Earth Fare bakery team to order!

Sales Valid Wed. September 6 – Tues. September 12

**HEALTHY REWARDS OFFER WITH CARD**

**\$3<sup>99</sup>**

SAVE \$3

New Crop

**Organic Gala Apples**

3 lb. Bag

Crisp, with a mildly sweet flavor and floral aroma.

Discount Taken at Register

**HEALTHY REWARDS OFFER WITH CARD**

**2/<sup>96</sup>**

SAVE \$2.98 on 2

**Blueberries**

Half Pint

A healthy snack—perfect for smoothies, added to yogurt, or make blueberry pancakes! High in Vitamin D and antioxidants.

Discount Taken at Register

**HEALTHY REWARDS OFFER WITH CARD**

**\$3<sup>99</sup>/lb.**

SAVE \$3.50/lb.

Ground Fresh 3X Daily

**USDA Choice or Better Ground Round**

Never Administered Antibiotics or Growth Hormones. 100% Vegetarian Diet. 3rd Party Certified Humanely Handled, Genetically Certified Angus. Traceable from Farm to Fork

Discount Taken at Register

**HEALTHY REWARDS OFFER WITH CARD**

**\$8<sup>99</sup>/lb.**

SAVE \$6/lb.

Fresh, Wild Caught

**Alaska Coho Salmon Fillets**

100% Traceable and Sustainably Sourced. All Earth Fare Salmon is non-GMO, Superior Flavor, Great Source of Omega-3 Fatty Acids. Product of the USA (Alaska)

Discount Taken at Register

**ORGANIC**

**HEALTHY REWARDS OFFER WITH CARD**

**\$2<sup>99</sup>/lb.**

SAVE \$1/lb.

**Organic Red & Green Seedless Grapes**

"Known to boost an important heart antioxidant known as glutathione, grapes can protect against heart failure and high blood pressure."—Angela Hind, M.D. and Wellness Expert

Discount Taken at Register

**HEALTHY REWARDS OFFER WITH CARD**

**2/<sup>4</sup>**

SAVE \$1 on 2

**Organic Green Kale**

A nutritional Superfood, kale is low in calories, high in fiber, iron and vitamins.

Discount Taken at Register

Excludes Red & Lacinato

**HEALTHY REWARDS OFFER WITH CARD**

**\$1<sup>29</sup>/lb.**

SAVE \$1.30/lb.

Fresh

**Whole Chickens**

Never Administered Antibiotics or Any Added Hormones. 100% Vegetarian Diet, 3rd Party Certified Humanely Raised

Discount Taken at Register

**HEALTHY REWARDS OFFER WITH CARD**

**\$9<sup>99</sup>/lb.**

SAVE \$5/lb.

**Cooked Tail-On Shrimp**

31-40 Count

No Water Additives or Preservatives. 100% Traceable and Sustainably Sourced, Caught with Turtle-Safe Nets

Discount Taken at Register

**HEALTHY REWARDS OFFER WITH CARD**

**\$7<sup>99</sup>/lb.**

SAVE \$2/lb.

**Napa Chicken Salad**

From the Deli

Signature item, made fresh in-house in small batches using American Humane Certified chicken that is never administered antibiotics or growth hormones.

Discount Taken at Register

SIGN UP FOR EMAIL DEALS & MORE AT [EARTHFARE.COM](http://EARTHFARE.COM)

Eroding and Dilating Images

**Cooked  
Tail-On Shrimp**

31-40 Count

No Water Additives or Preservatives.  
100% Traceable and Sustainably  
Sourced, Caught with Turtle-Safe Nets  
**Discount Taken at Register**

# Text Identification - Tesseract OCR



**2/\$6**  
**SAVE \$2.98 on 2**  
**Blueberries**

Half Pint

A healthy snack—perfect for smoothies, added to yogurt, or make blueberry pancakes! High in Vitamin C and antioxidants.

**Discount Taken at Register**



HEALTHY REWARDS  
OFFER WITH CARD

2/°6  
SAVE \$2.98 on 2

Blueberries

Half Pint

A healthy snack—perfect for smoothies, added to yogurt, or make blueberry pancakes! High in Vitamin C and antioxidants.

Discount Taken at Register

# Sentimental Analysis - Smart Heuristics

A few heuristics were integrated to improve the results:

## 1) Correct Translations:

a) E.g. lb = ["ib", "lb", "Lb", "ib.", "lb.", "Lb."]

## 2) Problem Dependent Ordering:

a) E.g. Line appearing before "SAVE" line includes useful numeric attribute information

## 3) Use Cases:

a) E.g. Numeric values appearing in "SAVE" line usually indicate one of the following scenarios:

i) Uom and save\_per unit

ii) save\_per\_unit and least\_unit\_for\_promo

iii) discount

# Sentimental Analysis - Fuzzy

```
process.extract("Cooked Tail-On Shrimp", product, limit=10, scorer=fuzz.token_sort_ratio)
```

```
[(['Cooked Tail-On Shrimp'], 100),  
(['Cooked Shrimp'], 76),  
(['Key West Pink Shrimp'], 63),  
(['Jumbo Key West Pink Shrimp'], 60),  
(['Wild American Shrimp'], 59),  
(['Bacon Wrapped Shrimp'], 59),  
(['Coconut & Tornado Shrimp Rings'], 57),  
(['New Zealand Sharp Cheddar'], 57),  
(['East Coast White Shrimp'], 55),  
(['Wild White American Shrimp'], 55)]
```

# Solution

	A	B	C	D	E	F	G	H	I
1	flyer_name	product_name	unit_promo	unit	least_unit	save_per_unit	discount	organic	
2	week_10_page_1	Cooked Tai	9.99 lb		1	5		0	
3	week_10_page_1	Organic Green Kale			2	0.5		1	
4	week_10_page_1	Whole Chic	0 lb		1	1.3		0	
5	week_10_page_1	NY Strip	lb		1	6		0	
6	week_10_page_1	Organic Red & Green S	lb		1	1		1	
7	week_10_page_1	Alaska Coho	6 lb		1			0	
8	week_10_page_1	Ground Ro	0 lb		1	3.5		0	
9	week_10_page_1	Blueberries	pint		2	1.49		0	
10	week_10_page_2	Brand Wide Sale			1		0.2	0	
11	week_10_page_2	Cleanses, Fibre Support, & Skinny C			1			1	
12	week_10_page_2	Clawhamm	949 pack		1	0.5		0	
13	week_10_page_2	Rustic Red	0 lb		1	5		0	
14	week_10_page_2	Pumpkin D	9.99 pack		1	0		0	
15	week_10_page_2	Maple Cara	0 inch		1	1		0	
16	week_10_page_2	Hummus	oz		2	1.99		1	