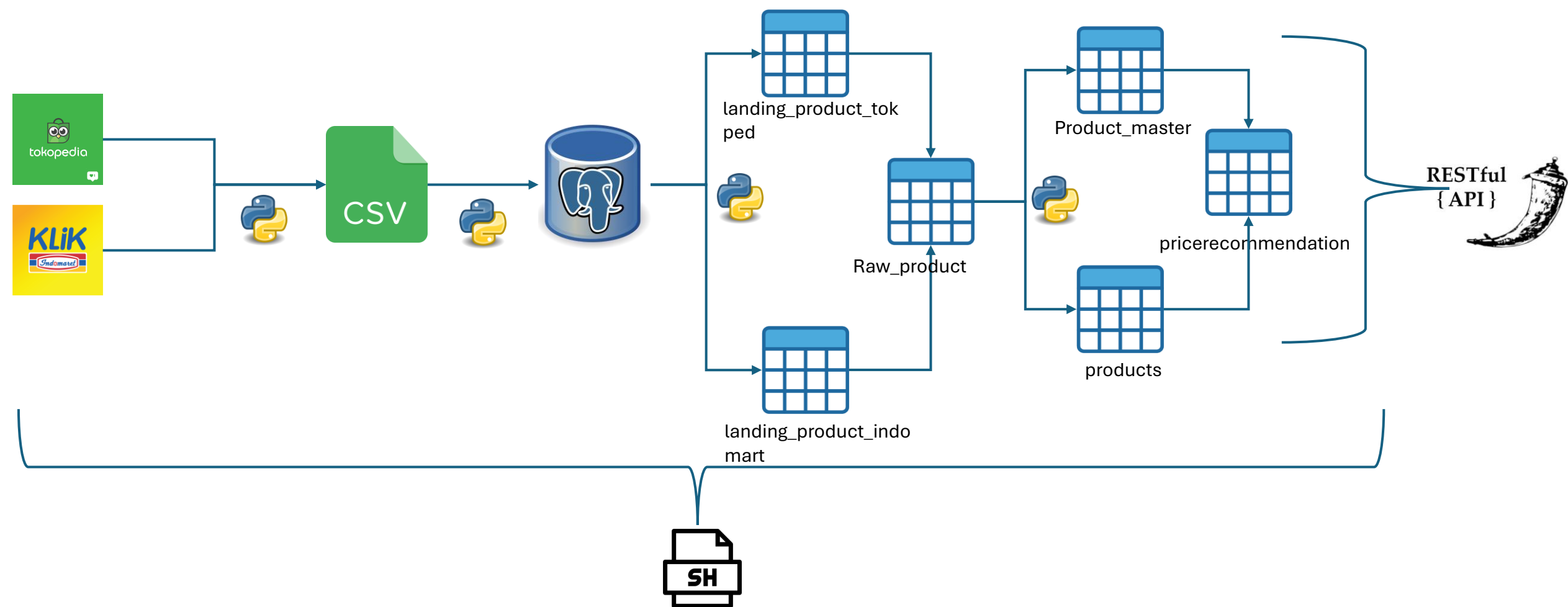


Dokumentasi

Flow Aplikasi sederhana

<https://github.com/yaafiasyari/tes.git>



Scraping Klikindomart



Pertama saya scraping perlink category product pada link yang di tugaskan
'https://www.klikindomaret.com/page/unilever-officialstore' agar mendapatkan data product dari
semua category



klikindomart_link.py

Setelah mendapatkan link setiap category product lalu saya scraping semua link berdasarkan product
agar mendapatkan detail data dari setiap product



klikindomart_link2.py

Lalu saya lanjut scraping perlink product dengan memanggil dari fungsi sebelumnya untuk di scraping
Sesuai data yang di butuhkan, kemudian saya export data menjaadi csv



klikindomart.py



products_indomart

Scraping Klikindomart



Pertama saya scraping perlink product pada link yang di tugaskan
<https://www.tokopedia.com/unilever/product> agar mendapatkan detail dari setiap product



tokpedlink.py

Lalu saya lanjut scraping perlink product dengan memanggil dari fungsi sebelumnya untuk di scraping
Sesuai data yang di butuhkan, kemudian saya export data menjaadi csv

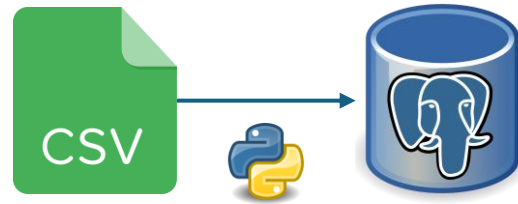


tokped_data.py



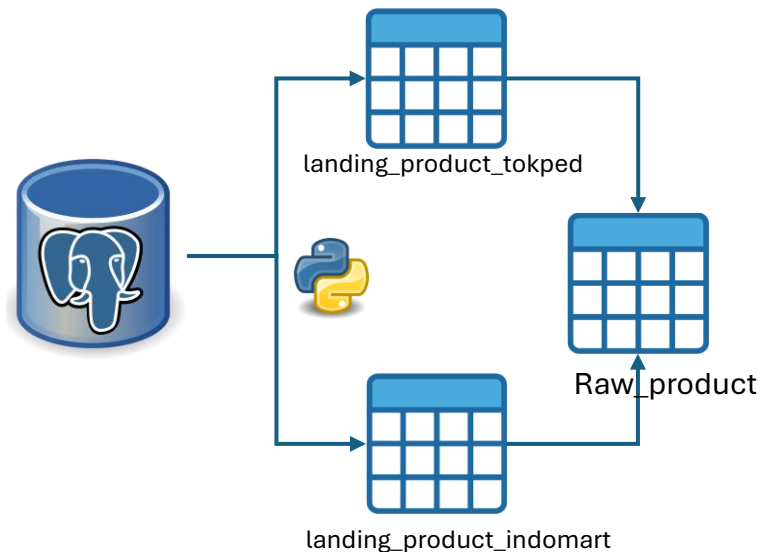
products_tokopedia

Normalisasi dan Ingest to database



Pertama saya export file csv menjadi dataframe pandas lalu saya normalisasikan kolom category agar sama, lalu saya cleansing value yang tidak di butuhkan seperti penulisan 'RP' pada harga dan angka yang tidak di butuhkan. Lanjut setelah melakukan normalisasi dan cleansing data langsung saya ingest pertable dari csv ada landing_product_indomart dan landing_product_tokped

Data processing raw_product



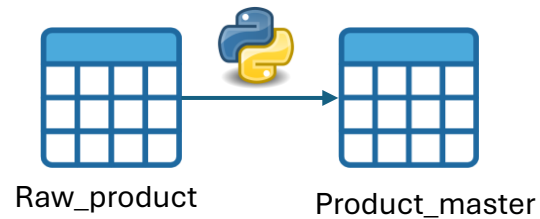
Setelah ingest data dari csv lanjut menggabungkan kedua data tersebut menjadi 1 table yaitu `raw_product` untuk di oleh ke tahap selanjutnya pada query ini juga terdapat beberapa cleansing untuk memisahkan detail dengan name



`raw_products.py`

raw_products												
Enter a SQL expression to filter results (use Ctrl+Space)												
	id	full_name_product	type	name	detail	original_price	discounted_price	discount_percentage	platform	created_date		
1	7,390	Pond's Sunserum SPF50 Bottle 30g FREE Ponds	Serum Wajah	Pond's Sunserum SPF50 Bottle	30g FREE Ponds Bright Serum Uclar 14g	210,900	121,300		42	www.tokopedia.com	2024-06-23 02:43:16.022	
2	7,391	Rinso Molto Deterjen Bubuk Japanese Peach 7	Pewangi Pelembut Pakaian	Rinso Molto Deterjen Bubuk Japanese Peach	770g FREE Cover Baju	57,900	16,600		71	www.tokopedia.com	2024-06-23 02:43:16.022	
3	7,392	Rinso Matic Deterjen Cair Bukanan Depan 1.45L	Pewangi Pelembut Pakaian	Rinso Matic Deterjen Cair Bukanan Depan	45L FREE Wireless Vacuum 2 in 1	130,900	63,300		52	www.tokopedia.com	2024-06-23 02:43:16.022	
4	7,393	Dove Deodorant Serum Roll On 3% Niacinamic	Deodorant	Dove Deodorant Serum Roll On 3% Niacinamic	45ml	39,900	25,000		37	www.tokopedia.com	2024-06-23 02:43:16.022	
5	7,394	Cif Pembersih Serbaguna Original Botol 660g F	Pembersih Lantai & Toilet	Cif Pembersih Serbaguna Original Botol	660g FREE Handuk Microfiber	59,900	36,100		40	www.tokopedia.com	2024-06-23 02:43:16.022	
6	7,395	Cif Pembersih Serbaguna Original Botol 660g F	Pembersih Lantai & Toilet	Cif Pembersih Serbaguna Original Botol	660g FREE Sponge 3M	89,900	36,100		60	www.tokopedia.com	2024-06-23 02:43:16.022	
7	7,396	Cif Pembersih Serbaguna Lemon Botol 660g FF	Pembersih Lantai & Toilet	Cif Pembersih Serbaguna Lemon Botol	660g FREE Handuk Microfiber	59,900	36,100		40	www.tokopedia.com	2024-06-23 02:43:16.022	
8	7,397	Cif Pembersih Serbaguna Lemon Botol 660g FF	Pembersih Lantai & Toilet	Cif Pembersih Serbaguna Lemon Botol	660g FREE Sikat Nagata	54,900	36,100		34	www.tokopedia.com	2024-06-23 02:43:16.022	
9	7,398	Cif Pembersih Serbaguna Original Botol 660g F	Pembersih Lantai & Toilet	Cif Pembersih Serbaguna Original Botol	660g FREE Sikat Nagata	54,900	36,100		34	www.tokopedia.com	2024-06-23 02:43:16.022	
10	7,399	CIF Cream Pembersih Serbaguna Lemon Botol	Pembersih Lantai & Toilet	CIF Cream Pembersih Serbaguna Lemon Botol	CIF Cream Pembersih Serbaguna Lemon Botol	54,900	37,400		32	www.tokopedia.com	2024-06-23 02:43:16.022	
11	7,400	CIF Cream Pembersih Serbaguna Original Boto	Pembersih Lantai & Toilet	CIF Cream Pembersih Serbaguna Original Boto	CIF Cream Pembersih Serbaguna Original Boto	54,900	37,400		32	www.tokopedia.com	2024-06-23 02:43:16.022	
12	7,401	Love Beauty&Planet Body Mist Lavender&Vani	Sabun Mandi	Love Beauty&Planet Body Mist Lavender&Vani	100ml Free Food Storage	119,900	80,500		33	www.tokopedia.com	2024-06-23 02:43:16.022	
13	7,402	Royco Kaldu Sapi 220gr dan Bango Manis 735n	Bumbu Masak Instan	Royco Kaldu Sapi	220gr dan Bango Manis 735ml Free Voucher VII	81,800	40,200		51	www.tokopedia.com	2024-06-23 02:43:16.022	
14	7,403	Royco Kaldu Ayam 220gr 3pcs Free Voucher VII	Bumbu Masak Instan	Royco Kaldu Ayam	220gr 3pcs Free Voucher VII	64,700	28,600		56	www.tokopedia.com	2024-06-23 02:43:16.022	
15	7,404	Vixal Kuat Segar Pouch 600ml	Pembersih Lantai & Toilet	Vixal Kuat Segar Pouch	600ml	20,900	14,400		31	www.tokopedia.com	2024-06-23 02:43:16.022	
16	7,405	Vixal Kuat Segar Pouch 180ml	Pembersih Lantai & Toilet	Vixal Kuat Segar Pouch	180ml	6,900	4,600		33	www.tokopedia.com	2024-06-23 02:43:16.022	
17	7,406	Vixal Kuat Segar Pouch 600ml & Wipol Karbol	Pembersih Lantai & Toilet	Vixal Kuat Segar Pouch	600ml & Wipol Karbol Cemara 750ml FREE Bot	53,800	31,300		42	www.tokopedia.com	2024-06-23 02:43:16.022	
18	7,407	Ponds Men UV Protect Sunscreen 30g	Krim Wajah	Ponds Men UV Protect Sunscreen	30g	61,900	42,400		32	www.tokopedia.com	2024-06-23 02:43:16.022	

Data processing product_master & products

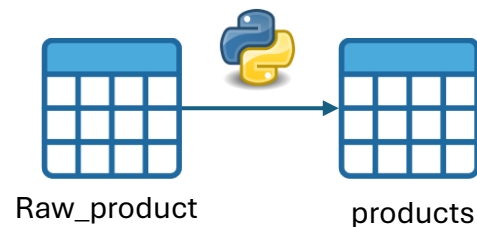


Setelah mendapatkan raw_product lalu insert ke table product_master dengan hanya mengambil kolom name,detail dan type dengan distinct berdasarkan name dan detail



product_master.py

	id	name	detail
1	1	Es Krim	Wall's Ice Cream Mini Cornetto Oreo 12X28ml
2	2	Saus & Dressing	Royco Saus Tiram
3	3	Shampoo	Sunsilk Shampoo Hijab Refresh
4	4	Deterjen Laundry	Rinso Kapsul Matic Konsentrat 4'S Sakura Blossom
5	5	Sabun Mandi	Lifebuoy Body Wash Sabun Mandi Cair Refill Total 1
6	6	Sabun Mandi	Lifebuoy Sabun Cair Lemon Fresh Refill
7	7	Shampoo	LIFEBUOY SHAMPOO ANTI DANDRUFF
8	8	Pembersih Lantai & Toilet	Wipol Karbol Wangi Lemon
9	9	Pembersih	Ponds Men Pembersih Wajah Ultra Bright Oil Fighte
10	10	Shampoo	Sunsilk Shampoo Soft & Smooth
11	11	Es Krim	Wall's Ice Cream Magnum White Almond
12	12	Sabun Mandi	Glow & Lovely Body Wash Multivitamin Niacinamic
13	13	Pewangi Pelembut Pakaian	Molto Pelembut dan Pewangi Parfum Boost Casabl
14	14	Sikat Gigi	Pepsodent Sikat Gigi Expert Clean 3'S Soft
15	15	Teh	Lipton Teh Celup Stl Enveloped 25's English Breakfa
16	16	Es Krim	Wall's Ice Cream Magnum Mini Mix 6X45mL
17	17	Pewangi Pelembut Pakaian	Rinso Molto Deterjen Bubuk Japanese Peach
18	18	Sikat Gigi	Pepsodent Sikat Gigi Double Care Sensitive Mp3 Sof
19	19	Vitamin Rambut	Sunsilk Vitamin Hair Mist 3 In 1
20	20	Treatment Liquid Soap	Lifebuoy Sabun Mandi Cair Lemon Fresh



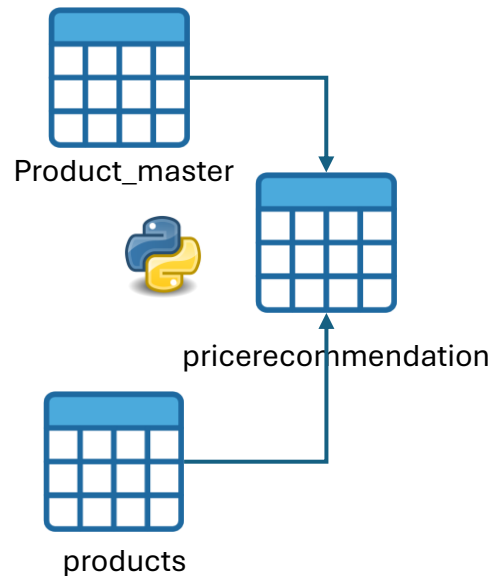
Lanjut setelah itu insert dari table raw_products di join dengan product_master untuk mengambil productmasterid



products.py

	id	name	price	originalprice	discountpercentage	detail	platform	productmasterid	created_date
1	1	Pond's Sunserum SPF50 Bottle	121,300	210,900		42 30g FREE Ponds Bright Serum Uclar 14g	www.tokopedia.com	205	2024-06-23 02:43:16.022
2	2	Rinso Molto Deterjen Bubuk Japanese Peach	16,600	57,900		71 770g FREE Cover Baju	www.tokopedia.com	17	2024-06-23 02:43:16.022
3	3	Rinso Matic Deterjen Cair Bukan Depan	63,300	130,900		52 45L FREE Wireless Vacuum 2 in 1	www.tokopedia.com	115	2024-06-23 02:43:16.022
4	4	Dove Deep Moist Serum Roll On 3% Niacinamide + 10x Collagen Pink	25,000	39,900		37 45ml	www.tokopedia.com	477	2024-06-23 02:43:16.022
5	5	CIF Pembersih Serbaguna Original Botol	36,100	59,900		40 660g FREE Handuk Microfiber	www.tokopedia.com	599	2024-06-23 02:43:16.022
6	6	CIF Pembersih Serbaguna Original Botol	36,100	89,900		60 660g FREE Sponge 3M	www.tokopedia.com	262	2024-06-23 02:43:16.022
7	7	CIF Pembersih Serbaguna Lemon Botol	36,100	59,900		40 660g FREE Handuk Microfiber	www.tokopedia.com	316	2024-06-23 02:43:16.022
8	8	CIF Pembersih Serbaguna Lemon Botol	36,100	54,900		34 660g FREE Sikat Nagata	www.tokopedia.com	372	2024-06-23 02:43:16.022
9	9	CIF Pembersih Serbaguna Original Botol	36,100	54,900		34 660g FREE Sikat Nagata	www.tokopedia.com	293	2024-06-23 02:43:16.022
10	10	CIF Cream Pembersih Serbaguna Lemon Botol	37,400	54,900		32 CIF Cream Pembersih Serbaguna Lemon Botol 660 g	www.tokopedia.com	46	2024-06-23 02:43:16.022
11	11	CIF Cream Pembersih Serbaguna Original Botol	37,400	54,900		32 CIF Cream Pembersih Serbaguna Original Botol 660 g	www.tokopedia.com	570	2024-06-23 02:43:16.022
12	12	Love Beauty&Planet Body Mist Lavender&Vanilla	80,500	119,900		33 100ml Free Food Storage	www.tokopedia.com	600	2024-06-23 02:43:16.022
13	13	Royco Kaldu Sapi	40,200	81,800		51 220gr dan Bango Manis 735ml Free Voucher VIU	www.tokopedia.com	408	2024-06-23 02:43:16.022
14	14	Royco Kaldu Ayam	28,600	64,700		56 220gr 3pcs Free Voucher VIU	www.tokopedia.com	556	2024-06-23 02:43:16.022
15	15	Vital Kuat Segar Pouch	14,400	20,900		31 600ml	www.tokopedia.com	587	2024-06-23 02:43:16.022
16	16	Vital Kuat Segar Pouch	4,600	6,900		33 180ml	www.tokopedia.com	345	2024-06-23 02:43:16.022

Machine learning Pricerecomendation



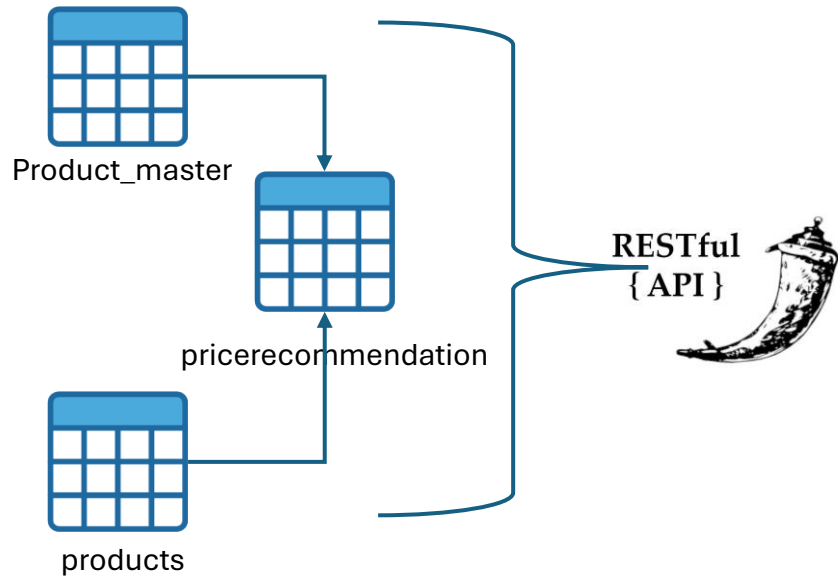
Disini saya export data products dan product master menjadi dataframe agar dapat di proses oleh python disini saya menggunakan library sklearn untuk mengambil rumus regresi linier untuk memprediksi 3 hari kedepan price rekomendasi



ml.py

pricerecommendation					
Enter a SQL expression to filter results (use Ctrl+Space)					
	id	productmasterid	predicted_price	date	
1	1	157	59,000	2024-06-25 00:00:15.358	
2	2	157	59,000	2024-06-26 00:00:15.358	
3	3	157	59,000	2024-06-27 00:00:15.358	
4	4	497	59,000	2024-06-25 00:00:15.367	
5	5	497	59,000	2024-06-26 00:00:15.367	
6	6	497	59,000	2024-06-27 00:00:15.367	
7	7	572	29,900	2024-06-25 00:00:15.374	
8	8	572	29,900	2024-06-26 00:00:15.374	
9	9	572	29,900	2024-06-27 00:00:15.374	
10	10	538	40,300	2024-06-25 00:00:15.382	
11	11	538	40,300	2024-06-26 00:00:15.382	
12	12	538	40,300	2024-06-27 00:00:15.382	
13	13	234	18,300	2024-06-25 00:00:15.390	
14	14	234	18,300	2024-06-26 00:00:15.390	

API Development



Lalu saya membuat endpoint untuk melihat table product_master, products dan pricerecommendation menggunakan Flask API dengan Authorization Basic Auth username "Yaafi" password "password"

```
{
  "created_date": "Sun, 23 Jun 2024 02:43:16 GMT",
  "detail": "30g FREE Ponds Bright Serum Uclar 14g",
  "discountpercentage": 42,
  "id": 1,
  "name": "Pond's Sunserum SPF50 Bottle",
  "originalprice": 210900,
  "platform": "www.tokopedia.com",
  "price": 121300,
  "productmasterid": 205
},
{
  "created_date": "Sun, 23 Jun 2024 02:43:16 GMT",
  "detail": "770g FREE Cover Baju",
  "discountpercentage": 71,
  "id": 2,
  "name": "Rinso Molto Detergen Bubuk Japanese Peach",
  "originalprice": 57900,
  "platform": "www.tokopedia.com",
  "price": 16600,
  "productmasterid": 17
}
```

products

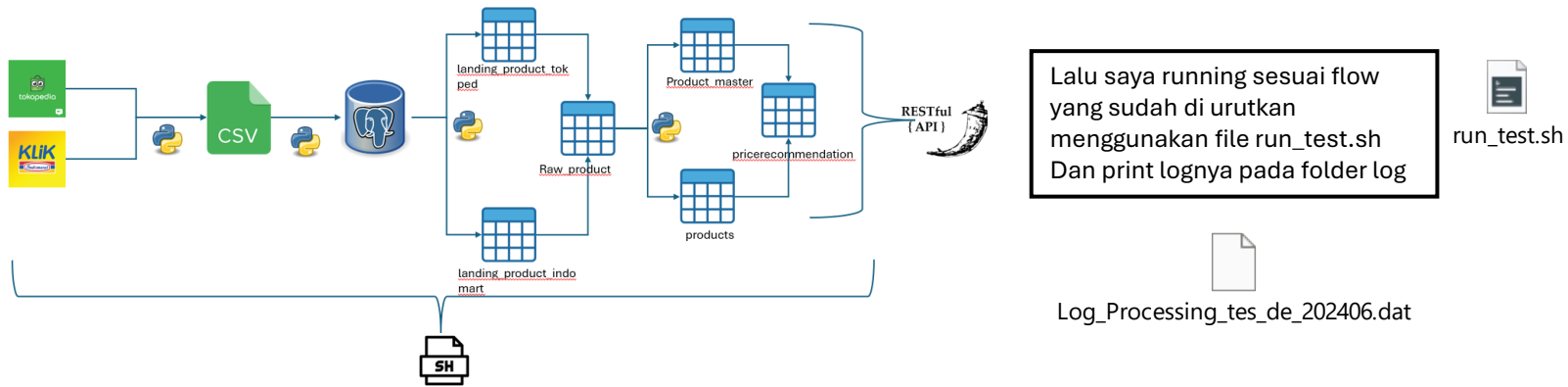
```
{
  "detail": "28ml",
  "id": 1,
  "name": "Wall's Ice Cream Mini Cornetto Oreo 12X28ml",
  "type": "Es Krim"
},
{
  "detail": "275ml Free Botol Minyak",
  "id": 2,
  "name": "Royco Saus Tiram",
  "type": "Saus & Dressing"
},
{
  "detail": "160mL",
  "id": 3,
  "name": "Sunsilk Shampoo Hijab Refresh",
  "type": "Shampoo"
}
```

Product_master

```
{
  "date": "Tue, 25 Jun 2024 00:00:15 GMT",
  "predicted_price": 59000.0,
  "productmasterid": 157
},
{
  "date": "Wed, 26 Jun 2024 00:00:15 GMT",
  "predicted_price": 59000.0,
  "productmasterid": 157
},
{
  "date": "Thu, 27 Jun 2024 00:00:15 GMT",
  "predicted_price": 59000.0,
  "productmasterid": 157
},
{
  "date": "Tue, 25 Jun 2024 00:00:15 GMT",
  "predicted_price": 59000.0,
  "productmasterid": 497
}
```

pricerecommendation

Running



Terimakasih