

**Tugas 6**  
**METODE PENGUMPULAN DATA DIGITAL - API**



**Disusun oleh Kelompok 11:**  
Christi Novelina Panggabean (2106725021)  
Jovina Nareswari Hubballah (2106708223)  
Nadia Amelia Kirana (2106653224)  
Nadia Sukesi Sianipar (2106700776)

**Mata Kuliah Metode Pengumpulan Data Digital**

**DEPARTEMEN MATEMATIKA**  
**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**  
**UNIVERSITAS INDONESIA**  
**TAHUN AJARAN 2023/2024**

## Website 1 : OLX

Website OLX adalah platform online yang memungkinkan pengguna untuk membeli, menjual, dan berinteraksi dengan barang-barang bekas atau baru secara lokal.

Tujuan :

1. Analisis Pasar : Dengan mengumpulkan data tentang handphone yang dijual di daerah Depok pada website OLX, pengguna dapat melakukan analisis pasar untuk memahami tren harga, permintaan, dan penawaran. Hal ini dapat membantu pengguna dalam mengambil keputusan sebelum membeli handphone.
2. Pemantauan Persaingan : Dengan menggunakan data yang sudah dikumpulkan, pengguna dapat memantau aktivitas pesaing dan strategi harga mereka.

Akan dilakukan web scraping menggunakan API pada website [www.olx.com](http://www.olx.com) untuk mengambil informasi mengenai handphone yang dijual di daerah Depok.

Berikut adalah tahapan nya :

- Inspect elemen, dan mencari URL API

```
https://www.olx.co.id/api/relevance/v4/search?category=208&facet\_limit=100&location=4000024&location\_facet\_limit=20&platform=web-desktop&query=handphone&relaxedFilters=true&spellcheck=true&user=18e980c4fd9x6843dcdd
```

- Menampilkan data di python

```
url =
"https://www.olx.co.id/api/relevance/v4/search?category=208&facet_limit=100&location=4000024&location_facet_limit=20&platform=web-desktop&query=handphone&relaxedFilters=true&spellcheck=true&user=18e980c4fd9x6843dcdd"
payload = {}
headers = {"User-Agent" : "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/123.0.0.0 Safari/537.36"}

response = requests.request("GET", url, headers=headers,
data=payload)
print(response.text)
```

- Ganti Tipe data ke JSON

```
# GANTI TIPE KE JSON
import json
```

```
json_data = json.loads(response.text) ['data']
```

## - Tabulasi Data

```
#Tabulasi
import pandas as pd
df = pd.json_normalize(json_data)
df
```

result :

id	score	category_id	ad_id	has_phone_param	description	created_at	inspection_info	title	multi_location	...	locations_resolved.COUNTRY_name	locations_resolved.ADMIN_LEVEL_1_id	locations_resolved.ADMIN
918643890	1.000000	208	918643890	False	Samsung a54 5gBaru\nBm buka dus.. InRam 8/2... 30T12:58:08+07:00	2024-03-30T12:58:08+07:00	None	Samsung a54 5g ram 8/128gb	None	...	Indonesia	2000009	
915426172	0.998221	208	915426172	False	Galaxy Gadget kini tersedia dengan In-RAM 8/2... 26T13:30:01+07:00	2024-03-26T13:30:01+07:00	None	Samsung Galaxy S24 Ultra	None	...	Indonesia	2000009	
918592543	0.998441	208	918592543	False	forsale\niphone 14 pro max 256gb ibox fullset...	2024-03-30T13:59:10+07:00	None	iPhone 14 Pro Max 256GB ibox	None	...	Indonesia	2000009	
917931714	0.994962	208	917931714	False	Cukup dengan EKTP saja bisa bawa pulang handph... 15T14:03:15+07:00	2024-03-15T14:03:15+07:00	None	Cukup bunga Ringan iPhone 15 Series bunga mu...	None	...	Indonesia	2000009	
917001156	0.992283	208	917001156	False	90Hz punch hold screen with magic ring\n18watt... 01T14:05:07+07:00	2024-04-01T14:05:07+07:00	None	INFINIX HOT 40i NFC 8+8256GB	None	...	Indonesia	2000009	
918759707	0.991103	208	918759707	False	No minus\ninRam 3/32GB - Hp dan charger	2024-04-01T14:02:24+07:00	None	Vivo Y12 3/32Gb	None	...	Indonesia	2000009	
903314125	0.989324	208	903314125	False	Jual Hp Oppo F15 Gold Ram 3/16 Iod baru & bat...	2024-04-01T13:49:20+07:00	None	Oppo F15 Gold Ram 3/16	None	...	Indonesia	2000009	
918758245	0.987544	208	918758245	False	Batangan dan dus only/inminus pasah di tombol v...	2024-04-01T13:32:15+07:00	None	Poco X3 GT 8/256	None	...	Indonesia	2000009	
918758003	0.985765	208	918758003	False	JUAL AJA VIVO Y21nRAM 4/64nKELENGKAPAN FILLS...	2024-04-01T13:27:05+07:00	None	VIVO Y21 RAM 4/65	None	...	Indonesia	2000009	
918757990	0.983988	208	918757990	False	Vivo y12 ram 3/32 nominus hn dan res sis maha...	2024-04-01T13:26:44+07:00	None	Vivo y12 ram 3/32 nominus	None	...	Indonesia	2000009	

## - Melihat semua isi kolom

```
df.columns
```

result :

```
Index(['id', 'score', 'category_id', 'ad_id', 'has_phone_param', 'description', 'created_at', 'inspection_info', 'title', 'multi_location', 'partner_id', 'user_type', 'partner_code', 'monetizationInfo', 'images', 'certified_car', 'is_kyc_verified_user', 'deal_price', 'business_platform', 'main_info', 'artificial_boost', 'display_date', 'isSpinViewAvailable', 'user_id', 'installment', 'created_at_first', 'locations', 'parameters', 'spell.id', 'spell.key', 'spell.version', 'spell.main', 'spell.facet_disabled', 'spell.default_sorting', 'status.status', 'status.allow_edit', 'status.ban_reason_id', 'status.display', 'status.translated_display', 'status.link', 'status.flags.new', 'status.flags.hot', 'status.message', 'status.allow_deactivate', 'favorites.count', 'favorites.count_label_next', 'price.value.raw', 'price.value.currency.iso_4217', 'price.value.currency.pre', 'price.value.display', 'price.key_name', 'price.key', 'package.id',
```

```
'locations_resolved.COUNTRY_id', 'locations_resolved.COUNTRY_name',
'locations_resolved.ADMIN_LEVEL_1_id',
'locations_resolved.ADMIN_LEVEL_1_name',
'locations_resolved.ADMIN_LEVEL_3_id',
'locations_resolved.ADMIN_LEVEL_3_name',
'locations_resolved.SUBLOCALITY_LEVEL_1_id',
'locations_resolved.SUBLOCALITY_LEVEL_1_name', 'favorites.count_label',
'package_id', 'favorites.count_label_prev'],
dtype='object')
```

- Melihat Keseluruhan data pada baris pertama

```
print(df.iloc[0].to_string())
```

result :

id	918643890
score	1.0
category_id	208
ad_id	918643890
has_phone_param	False
description	Samsung a54 5g\nBaru\nBlm buka dus.. \nRam 8/1...
created_at	2024-03-30T12:58:09+07:00
inspection_info	None
title	Samsung a54 5g ram 8/128gb
multi_location	None
partner_id	None
user_type	Regular
partner_code	None
monetizationInfo	None
images	[{'id': '2411503583', 'external_id': '6606b6c7...']}
certified_car	False
is_kyc_verified_user	False
deal_price	None
business_platform	None
main_info	None
artificial_boost	False
display_date	2024-03-29T12:40:49+0000
isSpinViewAvailable	False
user_id	7039993
installment	None
created_at_first	2024-03-29T19:40:49+07:00
locations	[{'lat': -6.375, 'lon': 106.823, 'region_id': ...}]
parameters	[{'type': 'single', 'key': 'make', 'value': 'e...'}]
spell.id	112
spell.key	DISCERN_LOCATION_ES
spell.version	1
spell.main	False
spell.facet_disabled	False
spell.default_sorting	DEFAULT
status.status	active
status.allow_edit	True
status.ban_reason_id	None
status.display	active
status.translated_display	Aktif
status.link	None
status.flags.new	False
status.flags.hot	False
status.message	None
status.allow_deactivate	True
favorites.count	0
favorites.count_label_next	1
price.value.raw	4800000.0
price.value.currency.iso_4217	IDR
price.value.currency.pre	Rp

```

price.value.display          Rp 4.800.000
price.key_name               Harga
price.key                   price
package.id                  2000057
locations_resolved.COUNTRY_id 1000001
locations_resolved.COUNTRY_name Indonesia
locations_resolved.ADMIN_LEVEL_1_id 2000009
locations_resolved.ADMIN_LEVEL_1_name Jawa Barat
locations_resolved.ADMIN_LEVEL_3_id 4000024
locations_resolved.ADMIN_LEVEL_3_name Depok Kota
locations_resolved.SUBLOCALITY_LEVEL_1_id 5001321
locations_resolved.SUBLOCALITY_LEVEL_1_name Beji
favorites.count_label        NaN
package_id                   NaN
favorites.count_label_prev   NaN

```

- Setelah itu, akan dipilih kolom kolom yang dibutuhkan

```

df2 =
df[['id','title','description','price.value.display','locations_
resolved.SUBLOCALITY_LEVEL_1_name','locations_resolved.ADMIN_LEVEL_3_name']]
df2

```

### result

id	title	description	price.value.display	locations_resolved.SUBLOCALITY_LEVEL_1_name	locations_resolved.ADMIN_LEVEL_3_name
0	Samsung a54 5g ram 8/128gb	Samsung a54 5g\nBaru\nBln buka dus.. InRam 8/1...	Rp 4.800.000	Beji	Depok Kota
1	Samsung Galaxy S24 Ultra	Galaxy Gadget kini tersedia dengan In-RAM 8/2...	Rp 21.999.000	Beji	Depok Kota
2	iPhone 14 Pro Max 256GB Ibox	forsale\niphone 14 pro max 256gb box fullset...	Rp 17.000.000	Cipayung	Depok Kota
3	Cicilan bunga Ringan Iphone 15 Series bunga mu...	Cukup dengan EKTP saja bisa bawa pulang handph...	Rp 13.999.000	Beji	Depok Kota
4	INFINIX HOT 40i NFC 8+8/256GB	90Hz punch hole screen with magic ring\n18watt...	Rp 1.699.000	Sawangan	Depok Kota
5	Vivo Y12 3/32Gb	No minus\nRam 3/32GB , Hp dan charger	Rp 850.000	Sawangan	Depok Kota
6	Oppo F1S Gold Ram 3 /16	Jual Hp Oppo F1S Gold Ram 3/16 lcd baru & bat...	Rp 550.000	Pancoran Mas	Depok Kota
7	Poco X3 GT 8/256	Batangan dan dus only\nminus patah di tombol v...	Rp 2.700.000	Beji	Depok Kota
8	VIVO Y21 RAM 4/65	JUAL AJA VIVO Y21\nRAM 4/64\nKELENGKAPAN FILLS...	Rp 1.200.000	Cipayung	Depok Kota
9	Vivo y12 ram 3/32 nominus hp dan cas aja	Vivo y12 ram 3/32 nominus hp dan cas aja mahar...	Rp 750.000	Pancoran Mas	Depok Kota
10	Iphone 11 - 64Gb	Iphone 11\n64 GB\npurple warna jarang\nfisik m...	Rp 4.200.000	Limo	Depok Kota
11	Samsung galaxy z fold 5 12 256gb	Samsung galaxy z fold 5 12 256gb Second ex Res...	Rp 18.600.000	Beji	Depok Kota
12	Samsung a23 6/128 like new fulset	Kelengkapan : \nfulset\nKondisi : \n* no m...	Rp 1.750.000	Beji	Depok Kota
13	Samsung a23 6/128 lengkap like new tanpa minus	Kelengkapan : \nfulset\nKondisi : \n* no m...	Rp 1.750.000	Pancoran Mas	Depok Kota
14	Samsung a23 6/128 lengkap like new	Kelengkapan : \nfulset\nKondisi : \n* no m...	Rp 1.750.000	Cilandong	Depok Kota
15	Samsung Z Flip3 Bunga 0%	Promo Ramadhan HOME CREDIT\nSamsung Bunga 0%	Rp 14.999.000	Sawangan	Depok Kota
16	Realme C31 second	Realme C31\nInRam 3/32	Rp 1.000.000	Cilandong	Depok Kota
17	Samsung A54 256Gb LIKE NEW KAK.PEMBELIAN BULAN FEB AKHIR MULUS B...	Rp 4.000.000	Sukmajaya	Depok Kota	
18	Iphone 13 bisa cicilan proses 3 menit	Cicilan mudah tanpa kartu kredit\nCUKUP KTP DA...	Rp 9.999.000	Beji	Depok Kota
19	Zenfone 9 8/256 hitam fulset	Flash sale edisi gajian nasus zenfone 9 8/256 ...	Rp 6.800.000	Pancoran Mas	Depok Kota

- Rename Kolom

```

df2.rename(columns={
    'title': 'Judul',
    'description': 'Deskripsi',
    'price.value.display': 'Harga',
    'locations_resolved.SUBLOCALITY_LEVEL_1_name': 'Kecamatan',
    'locations_resolved.ADMIN_LEVEL_3_name': 'Kota'
}, inplace=True)

```

```
df2.head(16)
```

## result

	<b>id</b>	<b>Judul</b>	<b>Deskripsi</b>	<b>Harga</b>	<b>Kecamatan</b>	<b>Kota</b>
0	918643890	Samsung a54 5g ram 8/128gb	Samsung a54 5g\nBaru\nBlm buka dus.. \nRam 8/1...	Rp 4.800.000	Beji	Depok Kota
1	915426172	Samsung Galaxy S24 Ultra	Galaxy Gadget kini tersedia dengan \n- RAM 8/2...	Rp 21.999.000	Beji	Depok Kota
2	918592543	iPhone 14 Pro Max 256GB Ibox	forsale\niphone 14 pro max 256gb ibox fullset\...	Rp 17.000.000	Cipayung	Depok Kota
3	917931714	Cicilan bunga Ringan Iphone 15 Series bunga mu...	Cukup dengan EKTP saja bisa bawa pulang handph...	Rp 13.999.000	Beji	Depok Kota
4	917001156	INFINIX HOT 40i NFC 8+8/256GB	90Hz punch hold screen with magic ring\n18watt...	Rp 1.699.000	Sawangan	Depok Kota
5	918759707	Vivo Y12 3/32Gb	No minus\nRam 3/32GB . Hp dan charger	Rp 850.000	Sawangan	Depok Kota
6	903314125	Oppo F1S Gold Ram 3 /16	Jual Hp Oppo F1S Gold Ram 3/16 Lcd baru & bat...	Rp 550.000	Pancoran Mas	Depok Kota
7	918758245	Poco X3 GT 8/256	Batangan dan dus only\nMinus patah di tombol v...	Rp 2.700.000	Beji	Depok Kota
8	918758003	VIVO Y21 RAM 4/65	JUAL AJA VIVO Y21\nRAM 4/64\nKELENGKAPAN FILLS...	Rp 1.200.000	Cipayung	Depok Kota
9	918757990	Vivo y12 ram 3/32 nominus hp dan cas aja	Vivo y12 ram 3/32 nominus hp dan cas aja mahar...	Rp 750.000	Pancoran Mas	Depok Kota
10	918756686	Iphone 11 - 64Gb	Iphone 11\n64 GB\npurple warna jarang\nnfisik m...	Rp 4.200.000	Limo	Depok Kota
11	918756318	Samsung galaxy z fold 5 12 256gb	Samsung galaxy z fold 5 12 256gb Second ex Res...	Rp 16.900.000	Beji	Depok Kota
12	918755947	Samsung a23 6/128 like new fulset	Kelengkapan : \n* fulset\n\nKondisi : \n* no m...	Rp 1.750.000	Beji	Depok Kota
13	918755910	Samsung a23 6/128 lengkap like new tanpa minus	Kelengkapan : \n* fulset\n\nKondisi : \n* no m...	Rp 1.750.000	Pancoran Mas	Depok Kota
14	918755851	Samsung a23 6/128 lengkap like new	Kelengkapan : \n* fulset\n\nKondisi : \n* no m...	Rp 1.750.000	Cilandak	Depok Kota
15	918755204	Samsung Z Flip5 Bunga 0%	Promo Ramadhan HOME CREDIT\n\nSamsung Bunga 0%...	Rp 14.999.000	Sawangan	Depok Kota

## Website 2 : IndoGold

Akan dilakukan web scraping menggunakan API pada website [www.indogold.id](http://www.indogold.id) untuk mengambil informasi mengenai harga beli dan jual emas 7 tahun terakhir dan harga beli dan jual perak 7 tahun terakhir

Tujuannya adalah untuk mendapatkan harga dan jual beli emas serta harga jual dan beli perak untuk memperoleh akses terkini mengenai harga pasar dari logam mulia tersebut. Informasi tersebut sangat penting untuk para investor, pedagang, individu yang tertarik dalam investasi atau perdagangan logam mulia

Berikut tahapannya :

### 1. Harga Beli Emas

- Inspect elemen dan cari URL API nya

```
https://www.indogold.id/ajax/chart_range/GOLD/2017-03-30/2024-03-30
```

- Menampilkan data di Python

```
url =  
"https://www.indogold.id/ajax/chart_range/GOLD/2017-03-30/
```

```

2024-03-30"
payload={}
headers = {}

response = requests.request("GET", url, headers=headers,
data=payload)

print(response.text)

```

- Ganti tipe data ke JSON

```

import json
json_data = json.loads(response.text)
json_data

```

- Tabulasi data

```

data = json_data[0]['data']
df = pd.DataFrame(data, columns=['timestamp', 'open',
'high', 'low', 'close'])
# Menampilkan DataFrame
print(df)

```

Output :

	timestamp	open	high	low	close
0	1490832000000	542500	544500	541000	544000
1	1490918400000	545000	545000	539000	542500
2	1491004800000	545500	546000	545000	546000
3	1491091200000	546000	546000	545500	545500
4	1491177600000	547000	547000	543500	546000
...	...	...	...	...	...
2553	1711411200000	1122723	1130675	1115765	1120735
2554	1711497600000	1132166	1134154	1120735	1122723
2555	1711584000000	1144591	1147076	1129184	1132166
2556	1711670400000	1151549	1154531	1142106	1144094
2557	1711756800000	1152046	1153537	1151549	1151549

[2558 rows x 5 columns]

- tambahkan kolom “tanggal” ke dalam data diatas, untuk memudahkan kita melihat daftar harga belinya. Harga beli ada di kolom “open”

	timestamp	open	high	low	close	tanggal
0	1490832000000	542500	544500	541000	544000	2017-03-30
1	1490918400000	545000	545000	539000	542500	2017-03-31
2	1491004800000	545500	546000	545000	546000	2017-04-01
3	1491091200000	546000	546000	545500	545500	2017-04-02
4	1491177600000	547000	547000	543500	546000	2017-04-03
...	...	...	...	...	...	...
2553	1711411200000	1122723	1130675	1115765	1120735	2024-03-26
2554	1711497600000	1132166	1134154	1120735	1122723	2024-03-27
2555	1711584000000	1144591	1147076	1129184	1132166	2024-03-28
2556	1711670400000	1151549	1154531	1142106	1144094	2024-03-29
2557	1711756800000	1152046	1153537	1151549	1151549	2024-03-30

## 2. Harga Jual Emas

API :

```
https://www.indogold.id/ajax/chart_range_jual/GOLD/2017-03-30/2024-03-30
```

Harga Jual :

```
data = json_data[0]['data']

# Mengubah data menjadi DataFrame
df = pd.DataFrame(data, columns=['timestamp', 'open',
                                  'high', 'low', 'close'])

# Menampilkan DataFrame
print(df)

tanggal_range = pd.date_range(start='2017-03-30',
                               end='2024-03-30')

# Konversi rentang tanggal menjadi format yang diinginkan
tanggal_formatted = tanggal_range.strftime('%Y-%m-%d')

# Menambahkan kolom "tanggal" ke DataFrame
df['tanggal'] = tanggal_formatted

# Menampilkan DataFrame
print(df)
```

Output :

	timestamp	open	high	low	close	tanggal
0	1490832000000	531500	533000	533500	530000	2017-03-30
1	1490918400000	534000	531000	534000	528000	2017-03-31
2	1491004800000	534500	534500	535000	533500	2017-04-01
3	1491091200000	534500	534500	535000	534500	2017-04-02
4	1491177600000	535500	534500	536000	532500	2017-04-03
...	...	...	...	...	...	...
2553	1711411200000	1095000	1093000	1102500	1088000	2024-03-26
2554	1711497600000	1104000	1094500	1106000	1093000	2024-03-27
2555	1711584000000	1116000	1104000	1118500	1101000	2024-03-28
2556	1711670400000	1123000	1115500	1125500	1113500	2024-03-29
2557	1711756800000	1123500	1123000	1124500	1123000	2024-03-30

### 3. Harga Beli Perak

API :

```
https://www.indogold.id/ajax/chart_range/PERAK/2017-03-30/
2024-03-30
```

Harga Beli :

```
data = json_data[0]['data']
df = pd.DataFrame(data, columns=['timestamp', 'open',
'high', 'low', 'close'])
# Menampilkan DataFrame
print(df)

tanggal_range = pd.date_range(start='2017-03-30',
end='2024-03-30')

# Konversi rentang tanggal menjadi format yang diinginkan
tanggal_formatted = tanggal_range.strftime('%Y-%m-%d')

# Menambahkan kolom "tanggal" ke DataFrame
df['tanggal'] = tanggal_formatted

# Menampilkan DataFrame
print(df)
```

Output :

	timestamp	open	high	low	close	tanggal
0	1490832000000	11050	11100	11000	11050	2017-03-30
1	1490918400000	11050	11050	10950	11050	2017-03-31
2	1491004800000	11100	11100	11050	11100	2017-04-01
3	1491091200000	11100	11100	11100	11100	2017-04-02
4	1491177600000	11100	11100	11000	11100	2017-04-03
...	...	...	...	...	...	...
2553	1711411200000	18315	18535	18205	18425	2024-03-26
2554	1711497600000	18425	18480	18205	18315	2024-03-27
2555	1711584000000	18645	18645	18260	18425	2024-03-28
2556	1711670400000	18700	18755	18590	18645	2024-03-29
2557	1711756800000	18700	18755	18700	18700	2024-03-30

#### 4. Harga Jual Perak

##### API

[https://www.indogold.id/ajax/chart\\_range\\_jual/PERAK/2017-03-30/2024-03-30](https://www.indogold.id/ajax/chart_range_jual/PERAK/2017-03-30/2024-03-30)

##### Harga Jual

```
data = json_data[0]['data']
df = pd.DataFrame(data, columns=['timestamp', 'open',
                                'high', 'low', 'close'])

# Menampilkan DataFrame
print(df)

tanggal_range = pd.date_range(start='2017-03-30',
                               end='2024-03-30')

# Konversi rentang tanggal menjadi format yang diinginkan
tanggal_formatted = tanggal_range.strftime('%Y-%m-%d')

# Menambahkan kolom "tanggal" ke DataFrame
df['tanggal'] = tanggal_formatted

# Menampilkan DataFrame
print(df)
```

Output :

	timestamp	open	high	low	close	tanggal
0	1490832000000	8950	8950	9000	8900	2017-03-30
1	1490918400000	9000	8950	9000	8900	2017-03-31
2	1491004800000	9000	9000	9000	9000	2017-04-01
3	1491091200000	9000	9000	9000	9000	2017-04-02
4	1491177600000	9000	9000	9000	8950	2017-04-03
...	...	...	...	...	...	...
2553	1711411200000	13350	13400	13500	13300	2024-03-26
2554	1711497600000	13450	13350	13500	13250	2024-03-27
2555	1711584000000	13600	13450	13600	13300	2024-03-28
2556	1711670400000	13650	13600	13650	13550	2024-03-29
2557	1711756800000	13650	13650	13650	13650	2024-03-30

### Website 3: Agoda

Agoda adalah sebuah platform pemesanan perjalanan online yang memungkinkan penggunaanya untuk mencari dan memesan hotel dan akomodasi di berbagai destinasi. Akan dilakukan web scrapping menggunakan API pada website [www.agoda.com](http://www.agoda.com) untuk mengambil beberapa informasi mengenai hotel di Yogyakarta.

Berikut adalah tahapan nya :

1. Inspect elemen, dan mencari URL API

```
https://www.agoda.com/api/personalization/PersonalizeRecommendedProperties/v1?recommendationType=2&hotelId=0&finalPriceView=1&hasSearchCriteria=true&cityId=14018&lengthOfStay=1&checkIn=2024-04-10T00%3A00%3A00&adults=1&children=0&rooms=1&\_ts=1711957536477
```

2. Menampilkan data pada python

```
agoda=
"https://www.agoda.com/api/personalization/PersonalizeRecommendedProperties/v1?recommendationType=2&hotelId=0&finalPriceView=1&hasSearchCriteria=true&cityId=14018&lengthOfStay=1&checkIn=2024-04-10T00%3A00%3A00&adults=1&children=0&rooms=1&_ts=1711957536477"

payload={}
headers = {}

response1 = requests.request("GET", agoda, headers=headers,
data=payload)
```

```
response1.text
```

```
'{"PersonalizedRecommendedProperties":[], "PersonalizedRecommendationTitle": "Based on properties you\u0027ve viewed, you might like:", "PersonalizedRecommendationPriceText": "Prices from", "RecommendationType": 2}'
```

```
agodayogyakarta = !gdown 1Ku2nPDkkNNH0g503kXGRccSRAagcjPXn
```

### 3. Mengganti tipe data menjadi JSON

```
agodayogyakarta_file=open('agoda_yogyakarta.json')
yogyakarta=json.load(agodayogyakarta_file)
```

### 4. Tabulasi data

```
yogyakartal = yogyakarta['PersonalizedRecommendedProperties']

property_list = []

for property_data in yogyakartal:
    property_dict = {
        'PropertyId': property_data['PropertyId'],
        'HotelName': property_data['HotelName'],
        'HotelUrl': property_data['HotelUrl'],
        'StarRating': property_data['StarRating'],
        'ReviewCount': property_data['ReviewCount'],
        'ReviewScoreText': property_data['ReviewScoreText'],
        'ReviewScore': property_data['ReviewScore'],
        'Price': property_data['Price'],
        'DiscountMessage': property_data['DiscountMessage']
    }
    property_list.append(property_dict)

df = pd.DataFrame(property_list)

df
```

	PropertyId	HotelName	HotelUrl	StarRating	ReviewCount	ReviewScoreText	ReviewScore	Price	DiscountMessage
0	409746	Hotel Tentrem Yogyakarta	/hotel-tentrem-yogyakarta/hotel/yogyakarta-id....	star-5	4587	Exceptional	9.2	151.37	25% DISCOUNT
1	746720	Greenhost Boutique Hotel Prawirotaman	/greenhost-boutique-hotel-prawirotaman/hotel/y...	star-4	4653	Excellent	8.9	49.49	41% DISCOUNT
2	772296	Adhisthana Hotel Yogyakarta	/adhisthana-hotel-yogyakarta/hotel/yogyakarta-...	star-3	2715	Excellent	8.8	53.37	35% DISCOUNT
3	1199795	Lafayette Boutique Hotel	/lafayette-boutique-hotel/hotel/yogyakarta-id....	star-5	2776	Exceptional	9.1	104.43	28% DISCOUNT
4	2989957	Yogyakarta Marriott Hotel	/yogyakarta-marriott-hotel_2/hotel/yogyakarta-...	star-5	2328	Exceptional	9.2	134.34	30% DISCOUNT

## Website 4: tiket.com

tiket.com adalah sebuah platform yang menyediakan jasa pemesanan tiket perjalanan secara online. Akan dilakukan web scrapping menggunakan API pada website [www.tiket.com](https://www.tiket.com/ms-gateway/tix-train-search-v2/journeys?orig=BD&otype=STATION&dest=GMR&dtype=STATION&ttype=ONE WAY&ddate=20240401&rdate=&acount=1&icount=0) untuk mengambil beberapa informasi mengenai jadwal keberangkatan dan kedatangan kereta dari Bandung (BD) ke Gambir (GMR).

Berikut adalah tahapan nya :

1. Inspect elemen, dan mencari URL API

```
https://www.tiket.com/ms-gateway/tix-train-search-v2/journeys?orig=BD&otype=STATION&dest=GMR&dtype=STATION&ttype=ONE WAY&ddate=20240401&rdate=&acount=1&icount=0
```

2. Menampilkan data pada python

```
tiket=
"https://www.tiket.com/ms-gateway/tix-train-search-v2/journeys?
orig=BD&otype=STATION&dest=GMR&dtype=STATION&ttype=ONE WAY&ddat
e=20240401&rdate=&acount=1&icount=0"

payload={}
headers = {}

response3 = requests.request("GET", tiket, headers=headers,
data=payload)

response3.text
```

```
'<!DOCTYPE html><html lang="en-US"><head><title>Just a moment...</title><meta http-equiv="Content-Type" content="text/html; charset=UTF-8"><meta http-equiv="X-UA-Compatible" content="IE=Edge"><meta name="robots" content="noindex, nofollow"><meta name="viewport" content="width=device-width,initial-scale=1"><style>*{box-sizing:border-box;margin:0;padding:0}html{line-height:1.15;-webkit-text-size-adjust:100%;color:#313131}button,html{font-family:system-ui,-apple-system,BlinkMacSystemFont,Segoe UI,Roboto,Helvetica Neue,Arial,Noto Sans,sans-serif,Apple Color Emoji,Segoe UI Emoji,Segoe UI Symbol,Noto Color Emoji}@media (prefers-color-scheme:dark){body{background-color:#222;color:#d9d9d9}body a{color:#fff}body a:hover{color:#ee730a;text-decoration:underline}body .lds-ring div{border-color:#999 transparent transparent}body .font-red{color:#b20f03}body .big-button,body .pow-button{background-color:#4693ff;color:#1d1d1d}body #challenge-success-text{background-image:url(...'
```

```
tiketfile = !gdown 1D217EKq83Ko3QcxXsGlz31T6SOfMAJm0
```

3. Mengganti tipe data menjadi JSON

```

bdgjkt_file=open('bdg_jkt.json')
bdgjkt=json.load(bdgjkt_file)

bdgjkt_code = bdgjkt['code']
bdgjkt_message = bdgjkt['message']
bdgjkt_errors = bdgjkt['errors']
bdgjkt_data = bdgjkt['data']

bdgjkt_departjourneys = bdgjkt_data['departJourneys']

bdgjkt_journeys = bdgjkt_departjourneys['journeys']
bdgjkt_journeys

```

#### 4. Tabulasi data

```

data_bdgjkt = pd.json_normalize(bdgjkt_journeys)
data_bdgjkt.info()

```

#	Column	Non-Null Count	Dtype
0	id	10 non-null	object
1	segmentSchedules	10 non-null	object
2	labelType	5 non-null	object
3	nudges	10 non-null	object
4	segmentCount	10 non-null	int64
5	departDate	10 non-null	object
6	departTime	10 non-null	object
7	arriveDate	10 non-null	object
8	arriveTime	10 non-null	object
9	totalDuration	10 non-null	int64
10	transitTimings	10 non-null	object
11	nextDays	10 non-null	int64
12	originStation.id	10 non-null	int64
13	originStation.code	10 non-null	object
14	originStation.name	10 non-null	object
15	originStation.locator	0 non-null	object
16	originStation.city	0 non-null	object
17	destinationStation.id	10 non-null	int64
18	destinationStation.code	10 non-null	object
19	destinationStation.name	10 non-null	object
20	destinationStation.locator	0 non-null	object
21	destinationStation.city	0 non-null	object

```

dropkolom = ['segmentSchedules', 'labelType', 'nudges',
'transitTimings',
'destinationStation.locator','destinationStation.city',

```

```

tabel_bdgjkt = data_bdgjkt.drop(columns=dropkolom)
tabel_bdgjkt

```

		<b>id</b>	<b>segmentCount</b>	<b>departDate</b>	<b>departTime</b>	<b>arriveDate</b>	<b>arriveTime</b>	<b>totalDuration</b>	<b>originStation.code</b>	<b>originStation.name</b>	<b>destinationStation.id</b>	<b>destinationStation.code</b>	<b>destinationStation.name</b>
0	BD-GMR-2024-04-01-7047-EKO-Q-910125968		1	2024-04-01	15:00:00	2024-04-01	17:45:00	9900	BD	Bandung	11	GMR	Gambir
1	BD-GMR-2024-04-01-7047-EKO-S-910126023		1	2024-04-01	15:00:00	2024-04-01	17:45:00	9900	BD	Bandung	11	GMR	Gambir
2	BD-GMR-2024-04-01-7047-EKS-J-910126532		1	2024-04-01	15:00:00	2024-04-01	17:45:00	9900	BD	Bandung	11	GMR	Gambir
3	BD-GMR-2024-04-01-7047-EKS-I-910125473		1	2024-04-01	15:00:00	2024-04-01	17:45:00	9900	BD	Bandung	11	GMR	Gambir
4	BD-GMR-2024-04-01-7047PC-EKS-A-939990631		1	2024-04-01	15:00:00	2024-04-01	17:45:00	9900	BD	Bandung	9	GMR	Gambir
5	BD-GMR-2024-04-01-7027A-EKO-S-910553898		1	2024-04-01	22:00:00	2024-04-02	00:49:00	10140	BD	Bandung	15	GMR	Gambir
6	BD-GMR-2024-04-01-7027A-EKO-Q-910553793		1	2024-04-01	22:00:00	2024-04-02	00:49:00	10140	BD	Bandung	15	GMR	Gambir
7	BD-GMR-2024-04-01-7027A-EKS-I-910553258		1	2024-04-01	22:00:00	2024-04-02	00:49:00	10140	BD	Bandung	15	GMR	Gambir
8	BD-GMR-2024-04-01-7027A-EKS-J-910553363		1	2024-04-01	22:00:00	2024-04-02	00:49:00	10140	BD	Bandung	15	GMR	Gambir
9	BD-GMR-2024-04-01-7027PC-EKS-A-939994985		1	2024-04-01	22:00:00	2024-04-02	00:49:00	10140	BD	Bandung	12	GMR	Gambir

## Website 5: FlightStats

FlightStats adalah sebuah platform online yang menyediakan informasi penerbangan. Akan dilakukan webscraping dengan API pada website flightstats.com untuk memperoleh infomasi kedatangan pesawat di bandara Soekarno-Hatta (CGK) pada tanggal 30 Maret 2023.

- URL API

```
"https://www.flighstats.com/v2/api-next/flight-tracker/arr/CGK/2024/3/30/6?carrierCode=&numHours=6&rqid=y3kzflvjw9"
```

- Menampilkan data pada python

```
url_fs =
"https://www.flighstats.com/v2/api-next/flight-tracker/arr/CGK/2024/3/30/6?carrierCode=&numHours=6&rqid=y3kzflvjw9"

payload_fs = { }
headers_fs = { }

response_fs = requests.request("GET", url_fs, headers=headers_fs,
data=payload_fs)
print(response_fs.text)
```

- Mengganti tipe data menjadi JSON

```
jdata_fs = json.loads(response_fs.text)
jdata_fs1 = jdata_fs['data']
jdata_fs2 = jdata_fs1['flights']
```

- Tabulasi Data

```
df_fs = pd.json_normalize(jdata_fs2)
df_fs
```

sortime	operatedby	url	departureTime.timeAMPM	departureTime.time24	arrivalTime.timeAMPM	arrivalTime.time24	carrier.fs	carrier.name	carrier.flightNumber	airport.fs	airport.city	isCodeshare
0	2024-03-29T13:55:00.000Z	None /flight-tracker/CTV/491?year=2024&month=3&date...	8:20PM	20:20	8:55PM	20:55	CTV	Citilink	491	BDJ	Baljarmasin	NaN
1	2024-03-29T14:55:00.000Z	None /flight-tracker/CTV/725?year=2024&month=3&date...	8:30PM	20:30	9:55PM	21:55	CTV	Citilink	725	SUB	Surabaya	NaN
2	2024-03-29T23:00:00.000Z	None /flight-tracker/NH/871?year=2024&month=3&date=...	12:05AM	00:05	6:00AM	06:00	NH	ANA	871	HND	Tokyo	NaN
3	2024-03-29T23:00:00.000Z	Operated by ANA 871 /flight-tracker/GA/9351?year=2024&month=3&date=...	12:05AM	00:05	6:00AM	06:00	GA	Garuda Indonesia	9351	HND	Tokyo	True
4	2024-03-29T23:00:00.000Z	Operated by ANA 871 /flight-tracker/UA/7902?year=2024&month=3&date=...	12:05AM	00:05	6:00AM	06:00	UA	United Airlines	7902	HND	Tokyo	True
...	...	...	...	...	...	...	...	...	...	...	...	...
227	2024-03-30T05:00:00.000Z	None /flight-tracker/JT/713?year=2024&month=3&date...	10:35AM	10:35	12:00PM	12:00	JT*	Lion Air	713	PNK	Pontianak	NaN
228	2024-03-30T05:00:00.000Z	None /flight-tracker/JT/171?year=2024&month=3&date...	11:10AM	11:10	12:00PM	12:00	JT*	Lion Air	171	TKG	Bandar Lampung	NaN
229	2024-03-30T05:05:00.000Z	None /flight-tracker/CTV/715?year=2024&month=3&date=...	10:40AM	10:40	12:05PM	12:05	CTV	Citilink	715	SUB	Surabaya	NaN
230	2024-03-30T05:15:00.000Z	None /flight-tracker/CTV/337?year=2024&month=3&date=...	10:50AM	10:50	12:15PM	12:15	CTV	Citilink	333	URG	Ujung Pandang	NaN
231	2024-03-30T05:20:00.000Z	None /flight-tracker/ID/7154?year=2024&month=3&date=...	11:25AM	11:25	12:20PM	12:20	ID*	Batik Air	7154	SIN	Singapore	NaN

232 rows x 13 columns

## Website 6: FlightRadar24

FlightRadar24 merupakan sebuah platform yang dapat melacak penerbangan dan mengetahui detailnya. Akan dilakukan webscraping untuk mengetahui detail penerbangan kedatangan pada bandara Soekarno-Hatta (CGK).

- URL API

```
"https://api.flightradar24.com/common/v1/airport.json?code=cgk&plugin[]=&plugin-setting[schedule][mode]=arrivals&plugin-setting[schedule][timestamp]=1711965142&page=1&limit=100&fleet=&token="
```

- Menampilkan data pada python

```
flightradar =
"https://api.flightradar24.com/common/v1/airport.json?code=cgk&plugin[]=&plugin-setting[schedule][mode]=arrivals&plugin-setting[schedule][timestamp]=1711965142&page=1&limit=100&fleet=&token="

pl={ }
hd = { }

response = requests.request("GET", flightradar, headers=hd,
data=pl)
!gdown 1oK80AWVl1f9kiN7osF6Ja1zOSYhvJvLu
```

- Mengganti tipe data menjadi JSON

```
df_fr24 = open('jsonformatter_fightradar24_cgkArrivals.txt')
data_fr24 = json.load(df_fr24)
data_fr24
jsonfr = data_fr24['result']
jsonfr2 = jsonfr['response']
jsonfr3 = jsonfr2['airport']
jsonfr4 = jsonfr3['pluginData']
jsonfr5 = jsonfr4['schedule']
jsonfr6 = jsonfr5['arrivals']
flightradar_data = jsonfr6['data']
flightradar_data
```

- Tabulasi data dan kolom pada tabel

```
df_radar = pd.json_normalize(flightradarfinal)
df_radar
```

	flight.identification.id	flight.identification.row	flight.identification.number.default	flight.identification.number.alternative	flight.identification.callsign	flight.identification.codeshare	flight.status.live	flight.status.text	flight.status.icon	flight.status.estimated	...	flight.time.scheduled.dep
0	34960893	5488381886	QG967	None	CTV967	None	False	Landed 16:51	green	None	...	17116
1	3497011F	5488304507	GA113	None	GA113	[MH5785 NH5500]	False	Landed 16:50	green	None	...	17116
2	3496e690a	5488304854	Q4411	QG411	Q4411	[WY5411 TH3371 KJ4055 EV7114 VN3939 SV6367]	False	Landed 16:30	green	None	...	17116
3	3496e2c13	5488324053	ID6523	None	BTX6523	None	False	Landed 16:45	green	None	...	17116
4	None	5488332405	JT795	None	None	None	False	Scheduled*	None	None	...	17116
...	...	...	...	...	...	...	...	...	...	...	...	...
95	None	5488331737	QG425	None	None	None	False	Scheduled	None	None	...	17116
96	None	5488331893	QG969	None	None	None	False	Scheduled	None	None	...	17116
97	None	5488401600	TR278	None	None	[SQ8592]	False	Scheduled	None	None	...	17116
98	None	5488450140	BB5109	None	None	[A15109]	False	Scheduled	None	None	...	17116
99	None	548827679	BB5537	None	None	[A15120]	False	Scheduled	None	None	...	17116

100 rows x 75 columns

```
df_radar.columns
```

```
Index(['flight.identification.id', 'flight.identification.row',
       'flight.identification.number.default',
       'flight.identification.number.alternative',
       'flight.identification.callsign', 'flight.identification.codeshare',
       'flight.status.live', 'flight.status.text', 'flight.status.icon',
       'flight.status.estimated', 'flight.status.ambiguous',
       'flight.status.generic.status.text',
       'flight.status.generic.status.type',
       'flight.status.generic.status.color',
       'flight.status.generic.status.diverted',
       'flight.status.generic.eventTime.utc',
       'flight.status.generic.eventTime.local', 'flight.aircraft.model.code',
       'flight.aircraft.model.text', 'flight.aircraft.registration',
       'flight.aircraft.country.id', 'flight.aircraft.country.name',
       'flight.aircraft.country.alpha2', 'flight.aircraft.country.alpha3',
       'flight.aircraft.hex', 'flight.aircraft.restricted',
       'flight.aircraft.serialNo', 'flight.aircraft.age.availability',
       'flight.aircraft.availability.serialNo',
       'flight.aircraft.availability.age', 'flight.owner.name',
       'flight.owner.code.iata', 'flight.owner.code.icao', 'flight.owner.logo',
       'flight.airline.name', 'flight.airline.code.iata',
       'flight.airline.code.icao', 'flight.airline.short',
       'flight.airport.origin.code.iata', 'flight.airport.origin.code.icao',
       'flight.airport.origin.timezone.name',
       'flight.airport.origin.timezone.offset',
       'flight.airport.origin.timezone.abbr',
       'flight.airport.origin.timezone.abbrName',
       'flight.airport.origin.timezone.isDst',
       'flight.airport.origin.info.terminal',
       'flight.airport.origin.info.baggage', 'flight.airport.origin.info.gate',
       'flight.airport.origin.name', 'flight.airport.origin.position.latitude',
       'flight.airport.origin.position.longitude',
       'flight.airport.origin.position.country.name',
       'flight.airport.origin.position.country.code',
       'flight.airport.origin.position.country.id',
       'flight.airport.origin.position.region.city',
       'flight.airport.origin.visible',
       'flight.airport.destination.timezone.name',
       'flight.airport.destination.timezone.offset',
       'flight.airport.destination.timezone.abbr',
       'flight.airport.destination.timezone.abbrName',
       'flight.airport.destination.timezone.isDst',
       'flight.airport.destination.info.terminal',
       'flight.airport.destination.info.baggage',
       'flight.airport.destination.info.gate', 'flight.airport.real',
       'flight.time.scheduled.departure', 'flight.time.scheduled.arrival',
       'flight.time.real.departure', 'flight.time.real.arrival',
       'flight.time.estimated.departure', 'flight.time.estimated.arrival',
       'flight.time.other.eta', 'flight.time.other.duration',
       'flight.aircraft.images', 'flight.owner'],
      dtype='object')
```