

PYTHON  
FINAL PROJECT  
ANAK AGUNG AYU DIVA SHANTY DARMAWAN

# CASE 1

**Dear Data Analyst,**

Akhir tahun ini, perusahaan akan memberikan hadiah bagi pelanggan yang memenangkan kompetisi **Festival Akhir Tahun**. Tim Marketing membutuhkan bantuan untuk menentukan perkiraan hadiah yang akan diberikan pada pemenang kompetisi nantinya. Hadiah tersebut akan diambil dari **TOP 5 Produk** dari Kategori **Mobiles & Tablets** selama tahun 2022, dengan jumlah kuantitas penjualan (valid = 1) paling tinggi.

Mohon bantuan untuk mengirimkan data tersebut sebelum akhir bulan ini ke Tim Marketing. Atas bantuan yang diberikan, kami mengucapkan terima kasih.

Regards,  
**Tim Marketing**

## SYNTAX

```
# Tulis kode Anda di bawah ini. Dapat menggunakan lebih dari 1 blok kode
condition1 = df['is_valid'] == 1
condition2 = df['category'] == 'Mobiles & Tablets'
condition3 = df['year'] == 2022
```

```
df_1 = df[(condition1) & (condition2) & (condition3)]
df_1
```

```
df1 = df_1.groupby(by = ['sku_name'], as_index = False)['after_discount'].sum()
df1_final = df1.sort_values(by = 'after_discount', ascending = False).head(5)
df1_final
```

## RESULT

	sku_name	after_discount
1	IDROID_BALRX7-Gold	518752000
37	iphone_7-128GB-wof-Matt Black	34127490
36	iphone_7-128GB-wof-Jet Black	28535710
6	Samsung-Galaxy-S8-G955-Plus-Black	25853210
2	IDROID_BALRX7-Jet black	17079202

Dari 79 baris hasil dari penyisihan pada data df, berikut adalah nama 5 produk teratas dari kategori Mobiles & Tablets pada tahun 2022. Dengan produk bernama **IDROID\_BALRX7-Gold** yang memiliki hasil penjualan paling tinggi sebesar **518752000**.

## CASE 2

Menindaklanjuti meeting gabungan Tim Warehouse dan Tim Marketing, kami menemukan bahwa ketersediaan stock produk dengan Kategori Others pada akhir 2022 kemarin masih banyak.

1. Kami mohon bantuan untuk melakukan pengecekan data penjualan kategori tersebut dengan tahun 2021 secara kuantitas penjualan. Dugaan sementara kami, telah terjadi penurunan kuantitas penjualan pada 2022 dibandingkan 2021. (Mohon juga menampilkan data ke-15 kategori)
2. Apabila memang terjadi penurunan kuantitas penjualan pada kategori Others, kami mohon bantuan untuk menyediakan data TOP 20 nama produk yang mengalami penurunan paling tinggi pada 2022 jika dibanding dengan 2021. Hal ini kami gunakan sebagai bahan diskusi pada meeting selanjutnya.

Mohon bantuan untuk mengirimkan data tersebut paling lambat 4 hari dari hari ini. Atas bantuan yang diberikan, kami mengucapkan terima kasih.

Regards,  
**Tim Warehouse**

# SYNTAX



```
# Tulis kode Anda di bawah ini. Dapat menggunakan lebih dari
condition1 = df['is_valid'] == 1
condition2 = df['year'] == 2021
condition3 = df['category'] == 'Others'

df_211 = df[(condition1) & (condition2) & (condition3)]
df_211
```



```
▶ condition1 = df['is_valid'] == 1
condition2 = df['year'] == 2022
condition3 = df['category'] == 'Others'

df_212 = df[(condition1) & (condition2) & (condition3)]
df_212
```

[illegible]

## SYNTAX 2.1

```
df21 = df211_final.merge(df212_final, on = 'sku_name', how = 'outer')
df21.fillna(0, inplace = True)
df21.rename(columns = {'qty_ordered_x' : 'Qty 2021',
                       'qty_ordered_y' : 'Qty 2022',
                       'sku_name' : 'Product Name'},
            inplace = True)
df21['Difference'] = df21['Qty 2022'] - df21['Qty 2021']
df21.head(15)
```

Gambar di kanan adalah hasil syntax yang menunjukkan perbandingan jumlah kuantitas penjualan kategori Others. Gambar menunjukkan 15 barang dari total 167 barang yang berada dalam kategori Others. Dari gambar juga dapat dilihat bahwa terdapat penurunan dalam penjualan tahun 2022.

## RESULT

	Product Name	Qty 2021	Qty 2022	Difference
0	Aladdin_Wrench_Snap N Grip_01	1.0	0.0	-1.0
1	Aladdin_bike_cover	3.0	2.0	-1.0
2	Am-PTV_ATS-004-M	5.0	0.0	-5.0
3	BFV_Black Friday voucher Rs 1500	1.0	0.0	-1.0
4	BFV_Black Friday voucher Rs 2000	1.0	0.0	-1.0
5	BOOSTER CABLE 500AMP	1.0	0.0	-1.0
6	BO_Hippo figured Child Seat	1.0	0.0	-1.0
7	BO_Rectangular Sea Pool	1.0	0.0	-1.0
8	Careem_25000 Credits	1.0	0.0	-1.0
9	Careem_50000 Credits	1.0	0.0	-1.0
10	Dawlance Split A/C LVS 30 Red 18000BTU LW	1.0	0.0	-1.0
11	Dawlance_MD 10 + DWB 600	23.0	0.0	-23.0
12	Dawlance_bundle_DWEK-8229_clear	1.0	0.0	-1.0
13	Entertainer Asia_Vouch 365 - 2017 Book Karachi	1.0	2.0	1.0
14	Entertainer Asia_Vouch 365-2017 Mobile App Lahore	2.0	1.0	-1.0



## SYNTAX 2.2

```
# Tulis kode Anda di bawah ini. Dapat menggunakan lebih dari 1 b  
df21.sort_values(by = 'Difference', ascending = True).head(20)
```

Berikut adalah top 20 produk yang mengalami penurunan paling tinggi pada tahun 2022. Penurunan terbesar dialami oleh produk **RB\_Dettol Germ Busting Kit-bf** dengan selisih perbedaannya 155 atau menurun sebesar 126.53% dari tahun 2021. Hal ini bisa terjadi dikarenakan kebutuhan dan keadaan yang sudah berbeda seiring dengan perubahan yang terjadi dari tahun sebelumnya.

## RESULT

	Product Name	Qty 2021	Qty 2022	Difference
39	RB_Dettol Germ Busting Kit-bf	200.0	45.0	-155.0
11	Dawlance_MD 10 + DWB 600	23.0	0.0	-23.0
43	Telemall_MM-DR-HB-L	23.0	2.0	-21.0
69	iu_Tickets General Enclosure-Islamabad	20.0	0.0	-20.0
40	RS_Rehmat-e-Shereen Mix Mithai	13.0	0.0	-13.0
73	kansai_NeverWet	10.0	1.0	-9.0
80	sindbad_Sindbad Gift Card-3	7.0	0.0	-7.0
66	emart_00-1	7.0	1.0	-6.0
2	Am-PTV_ATS-004-M	5.0	0.0	-5.0
50	Vouch 365 2016	5.0	0.0	-5.0
62	duma_4561253300294	4.0	0.0	-4.0
81	sockoye_QG in Quarter Grey	4.0	0.0	-4.0
44	The Vitamin Company Kojic Acid Whitening Cream...	4.0	0.0	-4.0
56	aw_Octane Booster-12oz./354ml	3.0	0.0	-3.0
26	MEGUIAR_G12711	4.0	1.0	-3.0
46	Trans2_LW 999	3.0	0.0	-3.0
17	JBS_IFAM-009	2.0	0.0	-2.0
47	Transcend_64GB JetFlash 810	2.0	0.0	-2.0
82	sstop_Universallensclipkit	2.0	0.0	-2.0
37	MEGUIAR_X1030EU	2.0	0.0	-2.0

## CASE 3

**Dear Data Analyst,**

Terkait ulang tahun perusahaan pada 2 bulan mendatang, Tim Digital Marketing akan memberikan informasi promo bagi pelanggan pada akhir bulan ini. Kriteria pelanggan yang akan kami butuhkan adalah mereka yang sudah melakukan check-out namun belum melakukan pembayaran (`is_gross = 1`) selama tahun 2022. Data yang kami butuhkan adalah ID Customer dan Registered Date.

Mohon bantuan, untuk mengirimkan data tersebut sebelum akhir bulan ini ke Tim Digital Marketing. Atas bantuan yang diberikan, kami mengucapkan terima kasih.

Regards,

**Tim Digital Marketing**



## SYNTAX

```
# Tulis kode Anda di bawah ini. Dapat menggunakan lebih dari 1 blok kode
condition1 = df['is_gross'] == 1
condition2 = df['is_valid'] == 0
condition3 = df['is_net'] == 0
condition4 = df['year'] == 2022
```

```
df3 = df[(condition1) & (condition2) & (condition3) & (condition4)]
df3
```

```
df3_final = df3[['id', 'customer_id', 'registered_date']]
df3_final
```

## RESULT

	id	customer_id	registered_date
9	ODR9699658949w	C246762L	2022-05-08
18	ODR1965502162e	C848774L	2021-11-07
19	ODR8450052777q	C693415L	2022-04-12
21	ODR7673587024b	C180595L	2022-04-22
22	ODR7333927150n	C587425L	2022-03-22

Berikut adalah hasil dari syntax yang menghasilkan daftar pelanggan yang memenuhi kriteria case pada tahun 2022. Syntax menghasilkan sebanyak **1052 data** yang memenuhi kriteria dan gambar di atas adalah beberapa data hasilnya.

## SYNTAX

```
[ ] from google.colab import files  
    df3_final.to_csv('audience_list.csv')  
    files.download('audience_list.csv')
```

## RESULT

	id	customer_id	registered_date
9	ODR96996	C246762L	08/05/2022
18	ODR19655	C848774L	07/11/2021
19	ODR84500	C693415L	12/04/2022
21	ODR76735	C180595L	22/04/2022
22	ODR73339	C587425L	22/03/2022
23	ODR55770	C612225L	24/04/2022
28	ODR85715	C315164L	05/11/2021
29	ODR69318	C904928L	18/05/2022
30	ODR36383	C317777L	06/12/2021
31	ODR90822	C353808L	09/05/2022

Syntax berikut berguna untuk mengunduh data dari sebuah dataframe menjadi csv sehingga dapat dengan mudah diterima dan dibaca oleh tim digital marketing yang akan menerimanya. Hasil csv yang telah dirapikan dapat dilihat pada gambar di atas.

# CASE 4

**Dear Data Analyst,**

Pada bulan October hingga Desember 2022, kami melakukan campaign setiap hari Sabtu dan Minggu. Kami hendak menilai, apakah campaign tersebut cukup berdampak pada kenaikan penjualan (before\_discount). Mohon bantuan untuk menampilkan data:

1. Rata-rata harian penjualan weekends (Sabtu dan Minggu) vs rata-rata harian penjualan weekdays (Senin-Jumat) per bulan tersebut. Apakah ada peningkatan penjualan pada masing-masing bulan tersebut.
2. Rata-rata harian penjualan weekends (Sabtu dan Minggu) vs rata-rata harian penjualan weekdays (Senin-Jumat) keseluruhan 3 bulan tersebut.

Mohon bantuan untuk mengirimkan data tersebut paling lambat minggu depan. Atas bantuan yang diberikan, kami mengucapkan terima kasih.

Regards,  
**Tim Campaign**

# SYNTAX

```
[46] # Tulis kode Anda di bawah ini. Dapat menggunakan lebih dari  
condition1 = df['is_valid'] == 1  
condition2 = df['day'] == 'Saturday'  
condition3 = df['day'] == 'Sunday'
```

```
df_weekend = df[(condition1) & (condition2) | (condition3)]  
df_weekend
```

```
[47] condition1 = df_weekend['year'] == 2022  
condition2 = df_weekend['quarter'] == 4  
  
dfweekend = df_weekend[(condition1) & (condition2)]  
dfweekend
```

```
[54] condition1 = df['is_valid'] == 1  
condition2 = df['day'] == 'Monday'  
condition3 = df['day'] == 'Tuesday'  
condition4 = df['day'] == 'Wednesday'  
condition5 = df['day'] == 'Thursday'  
condition6 = df['day'] == 'Friday'
```

```
df_weekdays = df[(condition1) & (condition2) |  
                  (condition3) | (condition4) |  
                  (condition5) | (condition6)]  
  
df_weekdays
```

```
[49] condition1 = df_weekdays['year'] == 2022  
condition2 = df_weekdays['quarter'] == 4  
  
dfweekdays = df_weekdays[(condition1) & (condition2)]  
dfweekdays
```

Sebelum mengawali, pertama akan dibagi terlebih dahulu sesuai dengan kondisi yang telah diberikan. Penjualan pada hari Sabtu dan Minggu akan dibagi dan dimasukkan ke dalam variabel **dfweekend**. Penjualan untuk hari-hari sisanya akan dimasukkan ke dalam variabel **dfweekdays**. Kemudian kedua variabel ini akan difilter lagi sesuai dengan tahun dan bulan yang terpilih. Kita menggunakan quarter 4 karena yang bulan yang terpilih adalah bulan Oktober hingga bulan Desember.

## SYNTAX 4.1

```
[52] dfweekend_final = dfweekend.groupby(by = ['month'],
                                         as_index = False)['after_discount'].mean()
      dfweekend_final.rename(columns= {'after_discount' : 'Weekend Sales'}, inplace = True)
      dfweekend_final

[51] dfweekdays_final = dfweekdays.groupby(by = ['month'],
                                              as_index = False)['after_discount'].mean()
      dfweekdays_final.rename(columns= {'after_discount' : 'Weekdays Sales'}, inplace = True)
      dfweekdays_final

df_4 = dfweekend_final.merge(dfweekdays_final,
                           left_on = 'month', right_on = 'month')
df_4['Difference'] = df_4['Weekdays Sales'] - df_4['Weekend Sales']
df_4
```

## RESULT

	month	Weekend Sales	Weekdays Sales	Difference
0	December	866444.346154	8.515144e+05	-14929.958654
1	November	819607.227273	1.239608e+06	420000.284632
2	October	830026.774194	1.105340e+06	275313.419836

Syntax pada gambar di kiri menghasilkan hasil pada gambar di atas. Dapat dilihat bahwa dari ketiga bulan yang diminta, peningkatan hanya terjadi pada bulan October dan November, sedangkan bulan Desember mengalami penurunan pada nilai rata-rata penjualan.



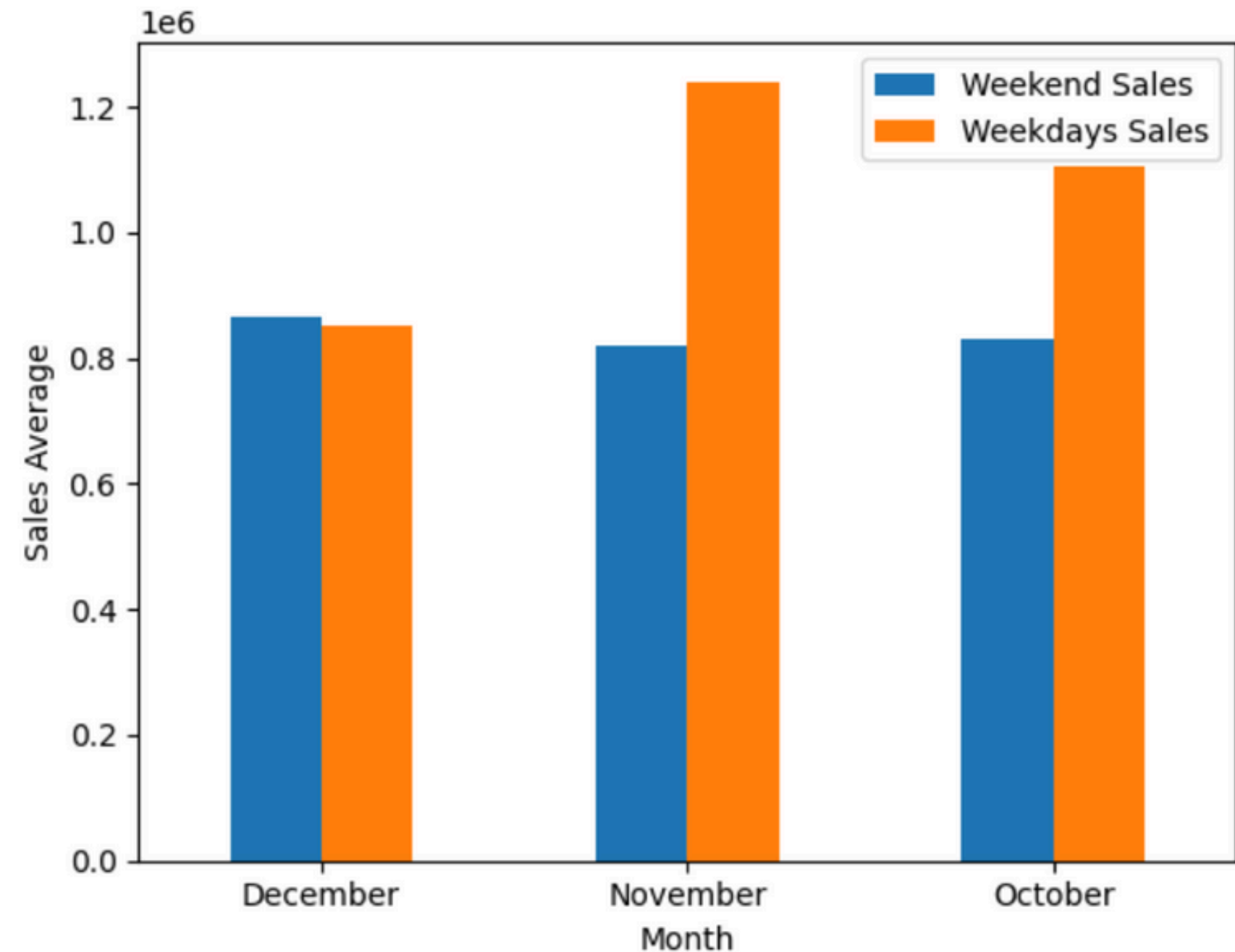
## SYNTAX 4.2

```
[56] df_4.plot(x = 'month',  
              kind = 'bar',  
              ylabel = 'Sales Average',  
              xlabel = 'Month',  
              rot = 0)
```

Selanjutnya akan dilihat perbandingannya secara lebih jelas menggunakan diagram batang. Dimana dapat dilihat bahwa memang yang mengalami peningkatan hanya pada bulan October dan November, sedangkan bulan December mengalami sedikit penurunan.

## RESULT

<Axes: xlabel='Month', ylabel='Sales Average'>





THANK YOU

ANAK AGUNG AYU DIVA SHANTY DARMAWAN