

TUGAS ANALISIS DATA MENGENAI DAFTAR PENERBANGAN OLEH BERBAGAI MASKAPAI



Mata Kuliah Pengenalan Komputasi KU1102

Kelas 37

Disusun Oleh:

Kelompok 4

1. Joseph Hansel 16720155
2. Caleb Effendi 16720317
3. Khairfirza Swandi 16720407

INSTITUT TEKNOLOGI BANDUNG

BANDUNG

Desember 2020

DAFTAR ISI

SAMPUL.....	i
DAFTAR ISI.....	ii
DESKRIPSI DATA DAN FILE.....	1
KARAKTERISTIK DATA	2
STATISTIK.....	4
VISUALISASI.....	7
KORELASI	12
PEMBAGIAN TUGAS.....	14

DESKRIPSI DATA DAN FILE

- **Jelaskan deskripsi data tersebut !**

Data yang kami pilih adalah data mengenai harga penerbangan berbagai maskapai dan ditambah dengan data jumlah penumpang dalam penerbangan tersebut. Di dalam data ini terdapat 6 kolom yang terdiri dari maskapai, asal, tujuan, tahun, penumpang, dan harga.

- **Jelaskan format data tersebut, data tersebut diambil dari mana, dan bagaimana dimensinya, serta ukuran file data !**

Format data yang kami ambil adalah csv, data ini diambil dari kumpulan dataset yang ada di internet, melalui [link https://drive.google.com/drive/folders/1dtbcOhJbCpaGaTS3h5h5gFm0tA0vbqGf](https://drive.google.com/drive/folders/1dtbcOhJbCpaGaTS3h5h5gFm0tA0vbqGf). Kami memilih dataset ini karena dunia industri dan transportasi saling berhubungan satu sama lain, dengan memahami harga transportasi, kami mengetahui cara paling efisien (secara finansial) untuk melakukan perjalanan.

Dimensi dari dataset ini adalah 9962 baris x 6 kolom ukuran filenya 342 KB

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9962 entries, 0 to 9961
Data columns (total 6 columns):
 #   Column      Non-Null Count  Dtype  
---  -
 0   maskapai    9962 non-null   object 
 1   asal        9962 non-null   object 
 2   tujuan      9962 non-null   object 
 3   tahun       9962 non-null   int64   
 4   penumpang  9962 non-null   int64   
 5   harga       9962 non-null   int64   
dtypes: int64(3), object(3)
memory usage: 350.3+ KB
```

KARAKTERISTIK DATA

- Jelaskan atribut/kolom apa saja yang ada di dalamnya, apa makna tiap atribut dan jelaskan jenisnya !

```
In [3]: # Apa saja kolom yang terdapat pada dataset
```

```
df.columns
```

```
Out[3]: Index(['maskapai', 'asal', 'tujuan', 'tahun', 'penumpang', 'harga'], dtype='object')
```

Atribut/kolom yang terdapat dalam dataset yaitu “maskapai”, “asal”, “tujuan”, “tahun”, “penumpang”, dan “harga”.

Makna tiap atribut

maskapai : nama maskapai yang digunakan
asal : asal penerbangan maskapai
tujuan : tujuan penerbangan maskapai
tahun : tahun penerbangan maskapai
penumpang : banyak penumpang yang menggunakan maskapai
harga : harga tiket maskapai untuk satu penumpang

```
In [10]: # Apa makna tiap atribut dan jelaskan jenisnya
```

```
df["maskapai"].value_counts()
```

```
Out[10]: mataram      2039  
country link    2007  
water asia      2004  
macan           1980  
rajawali        1932  
Name: maskapai, dtype: int64
```

```
In [11]: df["asal"].value_counts()
```

```
Out[11]: JOG      1717  
BDO      1698  
CGK      1661  
KNO      1631  
PKU      1628  
BTO      1626  
SRG         1  
Name: asal, dtype: int64
```

```
In [12]: df["tujuan"].value_counts()
```

```
Out[12]: DPS      1703  
BTH      1668  
UPG      1665  
SRG      1659  
PDG      1642  
HLP      1624  
BDO         1  
Name: tujuan, dtype: int64
```

```
In [13]: df["tahun"].value_counts()
```

```
Out[13]: 2014      2518  
2016      2488  
2017      2486  
2015      2470  
Name: tahun, dtype: int64
```

```
In [14]: df["penumpang"].value_counts()
```

```
Out[14]: 93      95  
76      90  
105     88  
180     86  
65      85  
..  
80      56  
97      54  
88      54  
135     53  
127     49  
Name: penumpang, Length: 141, dtype: int64
```

```
In [15]: df["harga"].value_counts()
```

```
Out[15]: 3779165      2  
3509357      2  
3545368      2  
1949158      2  
2285405      2  
..  
3215983      1  
1573448      1  
2783850      1  
1663593      1  
1556480      1  
Name: harga, Length: 9943, dtype: int64
```

Jenis tiap atribut

maskapai : kategorikal (nominal)
asal : kategorikal (nominal)
tujuan : kategorikal (nominal)
tahun : kuantitatif (diskrit)
penumpang : kuantitatif (diskrit)
harga : kuantitatif (diskrit)

- **Jelaskan karakteristik data untuk tiap atribut, misalnya terdiri atas nilai apa saja (pada atribut kategorikal) atau berapa *range* nilai datanya (pada atribut kuantitatif), berapa persen data kosong !**

```
In [10]: # Apa makna tiap atribut dan jelaskan jenisnya
df["maskapai"].value_counts()

Out[10]: mataram      2039
country link    2007
water asia      2004
macan           1980
rajawali        1932
Name: maskapai, dtype: int64

In [11]: df["asal"].value_counts()

Out[11]: JOG      1717
BDO      1698
CGK      1661
KNO      1631
PKU      1628
BTO      1626
SRG         1
Name: asal, dtype: int64

In [12]: df["tujuan"].value_counts()

Out[12]: DPS      1703
BTH      1668
UPG      1665
SRG      1659
PDG      1642
HLP      1624
BDO         1
Name: tujuan, dtype: int64
```

Karakteristik data untuk atribut kategorikal

maskapai : mataram, country link, water asia, macan, rajawali
 asal : JOG, BDO, CGK, KNO, PKU, BTO, SRG
 tujuan : DPS, BTH, UPG, SRG, PDG, HLP, BDO

```
In [16]: # Karakteristik tiap atribut
minimum = df.min()
print(minimum)

maskapai      country link
asal          BDO
tujuan        BDO
tahun         2014
penumpang      60
harga        1001488
dtype: object
```

```
In [17]: maksimum = df.max()
print(maksimum)

maskapai      water asia
asal          SRG
tujuan        UPG
tahun         2017
penumpang      200
harga        3999482
dtype: object
```

Karakteristik data untuk atribut kuantitatif

tahun : rangenya dari 2014 hingga 2017
 penumpang : rangenya dari 60 hingga 200
 harga : rangenya dari 1001488 hingga 3999482

```
In [19]: # Persen Data yang Kosong
df.isnull().sum()

Out[19]: maskapai      0
asal      0
tujuan     0
tahun      0
penumpang  0
harga      0
dtype: int64
```

Persen data yang kosong adalah 0% dikarenakan tidak terdapat data yang kosong pada dataset.

STATISTIK

- Tunjukkan sampel data dan berikan penjelasan pada sampel data tersebut
Data dengan penumpang lebih besar dari 100

```
In [20]: # Sampel data
# Data dengan penumpang besar dari 100
df.loc[df["penumpang"] > 100]
```

Out [20]:

	maskapai	asal	tujuan	tahun	penumpang	harga
0	macan	JOG	SRG	2014	126	2414238
1	country link	BTO	BTH	2016	141	2707440
2	macan	BDO	SRG	2016	189	3751886
3	rajawali	BDO	PDG	2017	117	2224218
5	macan	KNO	PDG	2016	172	3841582
...
9956	rajawali	BDO	PDG	2015	191	3780245
9957	macan	BTO	BTH	2016	152	2955327
9958	rajawali	KNO	HLP	2014	197	3922810
9959	country link	KNO	HLP	2016	193	3986309
9961	country link	CGK	SRG	2015	141	2711080

7079 rows x 6 columns

10 Data dengan harga tertinggi

```
In [21]: # 10 data dengan harga tertinggi
df.sort_values(["harga"], ascending = [0])[:10]
```

Out [21]:

	maskapai	asal	tujuan	tahun	penumpang	harga
1367	rajawali	BTO	BTH	2017	200	3999482
3608	macan	BTO	SRG	2014	200	3998965
2692	rajawali	BTO	HLP	2015	200	3998458
9153	country link	CGK	SRG	2015	200	3997547
734	water asia	BDO	BTH	2015	200	3997398
3187	water asia	BTO	BTH	2014	200	3996937
6572	mataram	BTO	BTH	2017	200	3996845
4651	rajawali	PKU	UPG	2015	200	3996837
6510	rajawali	BTO	HLP	2015	200	3996561
6004	country link	CGK	SRG	2016	200	3996457

Data dengan penumpang terbanyak pada tahun 2016, jika lebih dari satu data, maka tampilkan semuanya

```
In [22]: # Data dengan penumpang terbanyak pada tahun 2016, jika lebih dari satu data tampilkan semuanya
df1 = df.loc[df["tahun"] == 2016]
maks = df1.max()
df1.loc[df1["penumpang"] == maks["penumpang"]]
```

Out [22]:

	maskapai	asal	tujuan	tahun	penumpang	harga
331	country link	BTO	DPS	2016	200	3974585
1216	rajawali	CGK	SRG	2016	200	3971021
2192	rajawali	BDO	SRG	2016	200	3405316
2666	macan	BTO	BTH	2016	200	3982678
3130	water asia	BTO	BTH	2016	200	3971364
4153	mataram	PKU	SRG	2016	200	3974727
4774	water asia	CGK	DPS	2016	200	3980113
4874	country link	KNO	HLP	2016	200	3995749
5275	country link	KNO	BTH	2016	200	3992169
5705	macan	KNO	SRG	2016	200	3974594
6004	country link	CGK	SRG	2016	200	3996457
6999	country link	BDO	BTH	2016	200	3980337
7139	mataram	JOG	BTH	2016	200	3990589
7175	rajawali	BDO	UPG	2016	200	3977544
8152	water asia	PKU	PDG	2016	200	3987509
9447	rajawali	BDO	PDG	2016	200	3982286
9885	mataram	PKU	BTH	2016	200	3985918

Data dengan harga terkecil pada maskapai rajawali, jika ditemukan lebih dari satu, maka tampilkan pada baris pertama kali ditemukan

```
In [23]: # Data dengan harga terkecil pada maskapai rajawali
# jika ditemukan lebih dari satu tampilkan pada baris pertama kali ditemukan

df2 = df.loc[df["maskapai"] == "rajawali"]
imin = df2["harga"].idxmin()
df[imin:imin+1]
```

Out[23]:

	maskapai	asal	tujuan	tahun	penumpang	harga
3135	rajawali	CGK	BTH	2017	170	1002623

Mengakses data pada indeks baris ke-20 sampai ke-25 pada kolom “asal”

```
In [26]: # Mengakses data pada indeks baris ke-20 sampai ke-25 pada kolom asal

df.loc[20:25, "asal"]
```

Out[26]:

20	CGK
21	PKU
22	CGK
23	BD0
24	CGK
25	CGK

Name: asal, dtype: object

Mendapatkan rata-rata penumpang per kelompok maskapai

```
In [29]: # Mendapatkan rata-rata penumpang per kelompok maskapai

df.groupby(["maskapai"])["penumpang"].mean()
```

Out[29]:

maskapai	
country link	131.031390
macan	130.116667
mataram	130.601275
rajawali	129.780021
water asia	129.295908

Name: penumpang, dtype: float64

- Buatlah statistik untuk setiap atribut data, minimum terdiri dari rata-rata, standar deviasi, persentil, nilai maksimum, nilai minimum, distribusi frekuensi nilai pada data

Rata-rata, standar deviasi, persentil, nilai maksimum, dan nilai minimum

```
In [30]: # Statistik untuk setiap atribut data

df.describe()
```

Out[30]:

	tahun	penumpang	harga
count	9962.000000	9962.000000	9.962000e+03
mean	2015.496085	130.169745	2.500100e+06
std	1.120146	40.815328	8.625361e+05
min	2014.000000	60.000000	1.001488e+06
25%	2014.000000	95.000000	1.765279e+06
50%	2015.000000	130.000000	2.490682e+06
75%	2016.000000	166.000000	3.251025e+06
max	2017.000000	200.000000	3.999482e+06

```
In [43]: # Percentile 10%

df.quantile(0.1)
```

Out[43]:

tahun	2014.0
penumpang	74.0
harga	1304164.8

Name: 0.1, dtype: float64

```
In [44]: # Percentile 90%

df.quantile(0.9)
```

Out[44]:

tahun	2017.0
penumpang	187.0
harga	3704837.6

Name: 0.9, dtype: float64

Distribusi frekuensi nilai pada data

```
In [10]: # Apa makna tiap atribut dan jelaskan jenisnya
df["maskapai"].value_counts()

Out[10]: mataram      2039
country link    2007
water asia     2004
macan          1980
rajawali       1932
Name: maskapai, dtype: int64

In [11]: df["asal"].value_counts()

Out[11]: JOG      1717
BD0      1698
CGK      1661
KNO      1631
PKU      1628
BT0      1626
SRG         1
Name: asal, dtype: int64

In [12]: df["tujuan"].value_counts()

Out[12]: DPS      1703
BTH      1668
UPG      1665
SRG      1659
PDG      1642
HLP      1624
BD0         1
Name: tujuan, dtype: int64
```

```
In [13]: df["tahun"].value_counts()

Out[13]: 2014      2518
2016      2488
2017      2486
2015      2470
Name: tahun, dtype: int64

In [14]: df["penumpang"].value_counts()

Out[14]: 93      95
76      90
105     88
180     86
65      85
..
80      56
97      54
88      54
135     53
127     49
Name: penumpang, Length: 141, dtype: int64

In [15]: df["harga"].value_counts()

Out[15]: 3779165      2
3509357      2
3545368      2
1949158      2
2285405      2
..
3215983      1
1573448      1
2783850      1
1663593      1
1556480      1
Name: harga, Length: 9943, dtype: int64
```

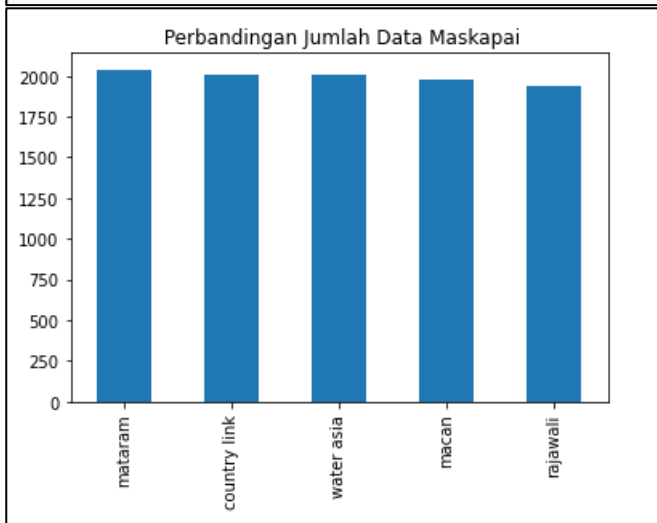

VISUALISASI

- Buatlah visualisasi dalam bentuk grafik/chart masing-masing minimum 2 buah untuk setiap kategori berikut :
 1. Perbandingan Kategori
 2. Penampilan Perubahan terhadap Waktu
 3. Penampilan hierarki dan hubungan keseluruhan bagian
 4. *Plotting relationships*

1. Perbandingan Kategori

- Perbandingan kategori jumlah data maskapai

```
df["maskapai"].value_counts().plot(kind = "bar", title = " Perbandingan Jumlah Data Maskapai")  
plt.show
```

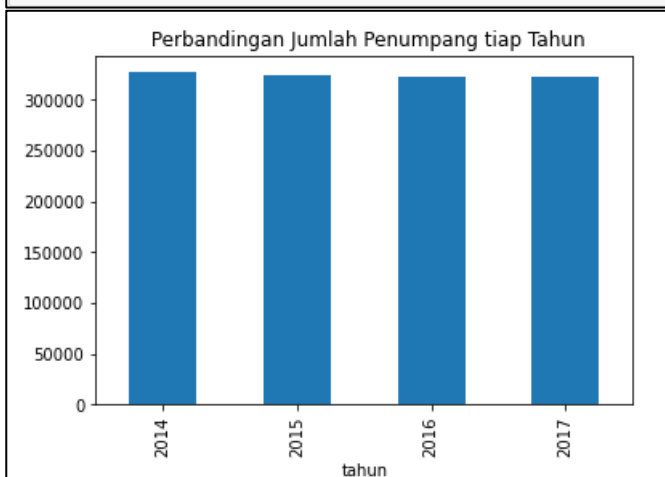


Keterangan : sumbu- x merupakan jenis- jenis maskapai dan sumbu-y merupakan banyaknya data

Dari grafik *bar chart* di atas, dapat dilihat perbandingan jumlah data tiap maskapai. Dari grafik tersebut diketahui bahwa maskapai Mataram memiliki jumlah data terbanyak dan maskapai Rajawali memiliki jumlah data yang paling sedikit. Jumlah data tiap maskapai tidak memiliki perbedaan yang terlalu signifikan antara satu sama yang lain.

- Perbandingan kategori jumlah penumpang tiap tahun

```
df.groupby("tahun")["penumpang"].sum().plot(kind = "bar", title = " Perbandingan Jumlah Penumpang tiap Tahun")  
plt.show
```



Keterangan : sumbu- x merupakan tahun dan sumbu-y merupakan banyaknya penumpang

Dari grafik *bar chart* di atas, dapat dilihat perbandingan jumlah total penumpang setiap tahun. Dari grafik tersebut diketahui bahwa pada tahun 2014 memiliki jumlah penumpang terbanyak dan pada tahun 2017 memiliki jumlah penumpang yang paling sedikit. Jumlah total penumpang tiap tahunnya tidak memiliki kenaikan ataupun pengurangan yang signifikan.

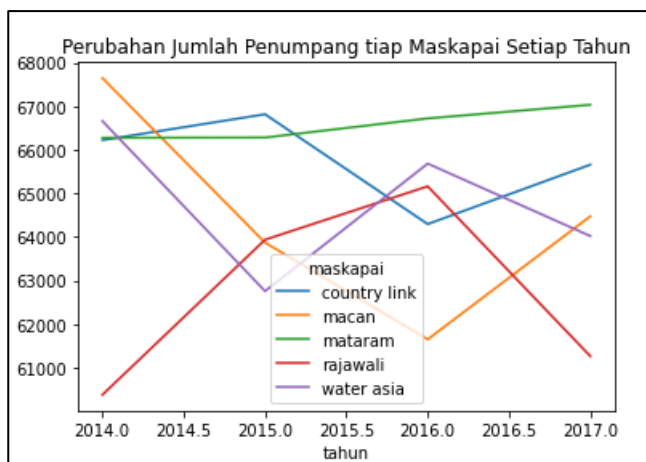
2. Penampilan Perubahan Terhadap Waktu

- Penampilan perubahan jumlah penumpang tiap maskapai terhadap waktu

```
df.groupby(["tahun", "maskapai"])["penumpang"].sum().unstack().plot(kind = "line", title = "Perubahan Jumlah Penumpang tiap Maskapai Setiap Tahun")
plt.show
```

Kode lengkap :

```
df.groupby(["tahun", "maskapai"])["penumpang"].sum().unstack().plot(kind = "line", title = "Perubahan Jumlah Penumpang tiap Maskapai Setiap Tahun")
plt.show
```



Keterangan : sumbu- x merupakan tahun dan sumbu-y merupakan banyaknya penumpang

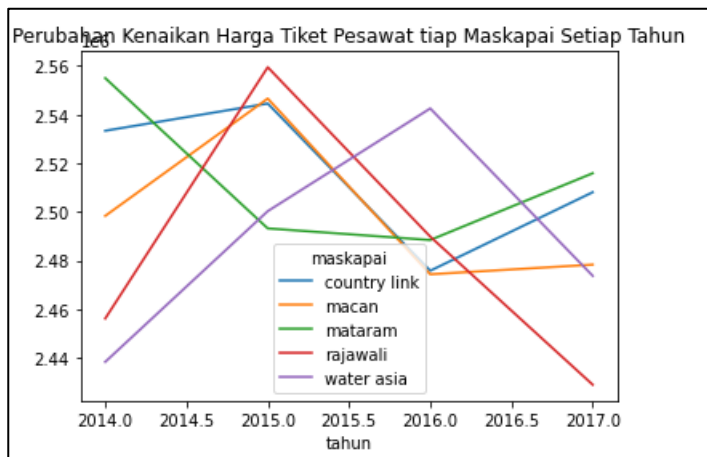
Dari grafik *line chart* di atas, dapat dilihat perbandingan perubahan jumlah total penumpang tiap maskapai setiap tahunnya. Dari grafik tersebut diketahui bahwa maskapai Mataram tidak mengalami penurunan jumlah penumpang dan selalu mengalami kenaikan jumlah penumpang walaupun kenaikannya tidak terlalu signifikan, maskapai Country Link memiliki perubahan jumlah penumpang yang tidak konsisten setiap tahunnya, maskapai Water Asia juga memiliki perubahan jumlah penumpang yang tidak konsisten setiap tahunnya, maskapai Macan mengalami penurunan jumlah penumpang hingga tahun 2016 namun mengalami kenaikan jumlah penumpang pada tahun 2017, maskapai Rajawali mengalami kenaikan jumlah penumpang hingga tahun 2016 namun mengalami penurunan jumlah penumpang pada tahun 2017. Setiap maskapai memiliki perbedaan perubahan jumlah penumpang yang cukup signifikan setiap tahunnya. Pada grafik ini diberi warna yang berbeda- beda bertujuan untuk membandingkan jumlah penumpang berbagai macam maskapai tiap tahunnya.

- Penampilan perubahan kenaikan harga tiket pesawat tiap maskapai terhadap waktu

```
df.groupby(["tahun", "maskapai"])["harga"].mean().unstack().plot(kind = "line", title = "Perubahan Kenaikan Harga Tiket Pesawat tiap Maskapai Setiap Tahun")
plt.show
```

Kode lengkap :

```
df.groupby(["tahun", "maskapai"])["harga"].mean().unstack().plot(kind = "line", title = "Perubahan Kenaikan Harga Tiket Pesawat tiap Maskapai Setiap Tahun")
plt.show
```

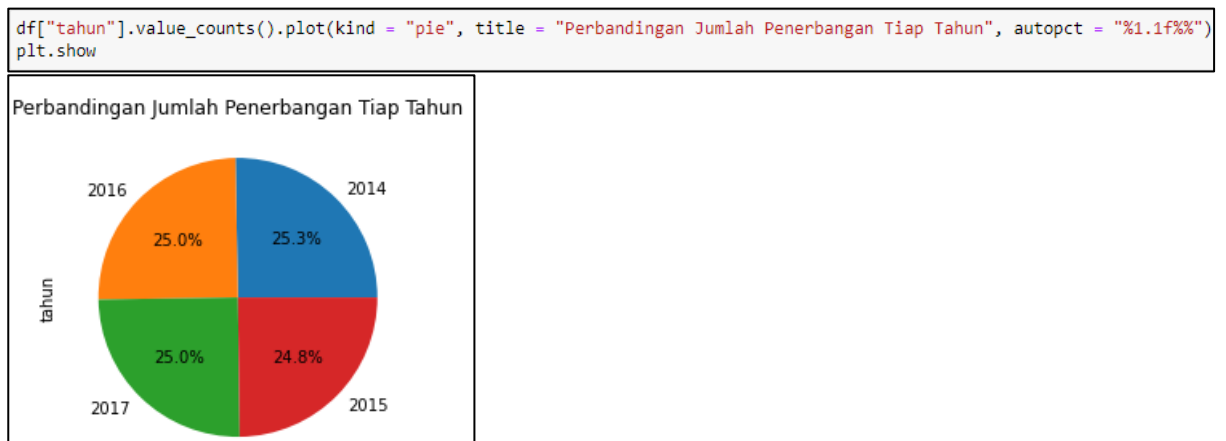


Keterangan : sumbu- x merupakan tahun dan sumbu-y merupakan harga tiket pesawat

Dari grafik *line chart* di atas, dapat dilihat perbandingan perubahan harga tiket pesawat rata-rata tiap maskapai setiap tahunnya. Dari grafik tersebut diketahui bahwa maskapai Mataram mengalami penurunan harga tiket pesawat hingga tahun 2016 namun mengalami kenaikan harga tiket pesawat pada tahun 2017, maskapai Country Link mengalami perubahan harga tiket pesawat yang tidak konsisten setiap tahunnya, maskapai Water Asia mengalami kenaikan harga tiket pesawat hingga tahun 2016 namun mengalami penurunan harga tiket pesawat pada tahun 2017, maskapai Macan mengalami perubahan harga tiket pesawat yang tidak konsisten, maskapai Rajawali mengalami kenaikan harga tiket pesawat hingga tahun 2015 namun mengalami penurunan harga tiket pesawat pada tahun 2016 hingga tahun 2017. Setiap maskapai memiliki perbedaan perubahan harga tiket pesawat yang cukup signifikan setiap tahunnya. Pada grafik ini diberi warna yang berbeda- beda bertujuan untuk membandingkan harga tiket pesawat berbagai maskapai tiap tahunnya.

3. Penampilan Hierarki dan Hubungan Keseluruhan-Bagian

- Perbandingan hierarki jumlah penerbangan tiap tahun



Keterangan : warna biru menunjukkan tahun 2014, warna merah menunjukkan tahun 2015, warna jingga menunjukkan tahun 2016, warna hijau menunjukkan tahun 2017.

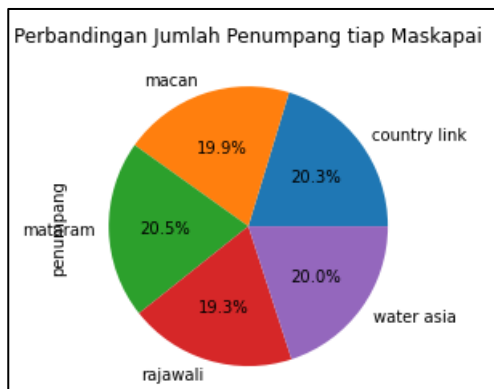
Dari grafik *pie chart* di atas, dapat dilihat perbandingan jumlah penerbangan setiap tahun. Dari grafik tersebut diketahui bahwa pada tahun 2014 memiliki jumlah penerbangan terbanyak dan pada tahun 2015 memiliki jumlah penerbangan yang paling sedikit, dan pada tahun 2016 dan 2017 memiliki jumlah penerbangan yang sama. Perbandingan jumlah penerbangan tiap tahunnya tidak memiliki perbedaan yang signifikan. Grafik tersebut diberi warna yang berbeda- beda bertujuan untuk membandingkan persentase perbandingan jumlah penerbangan tiap tahunnya.

- Perbandingan hierarki jumlah penumpang tiap maskapai

```
df.groupby("maskapai")["penumpang"].sum().plot(kind = "pie", title = "Perbandingan Jumlah Penumpang tiap Maskapai", autopct = "%1
plt.show
```

Kode lengkap :

```
df.groupby("maskapai")["penumpang"].sum().plot(kind = "pie", title = "Perbandingan Jumlah
Penumpang tiap Maskapai", autopct = "%1.1f%%")
plt.show
```



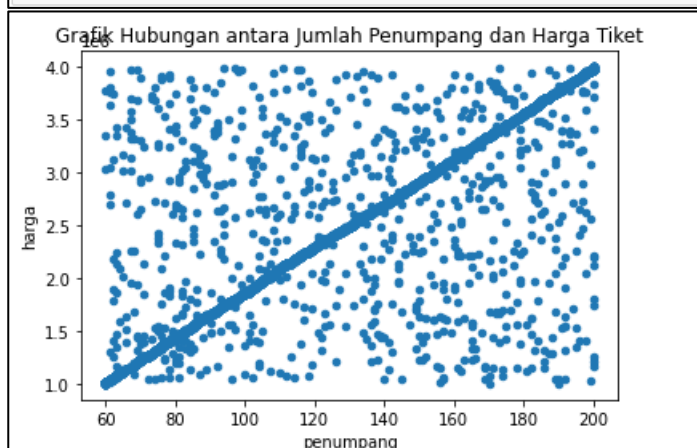
Keterangan : warna biru menunjukkan maskapai Country Link, warna ungu menunjukkan maskapai Water Asia, warna merah menunjukkan maskapai Rajawali, warna hijau menunjukkan maskapai Mataram, dan warna jingga menunjukkan maskapai Macan.

Dari grafik *pie chart* di atas, dapat dilihat perbandingan jumlah penumpang tiap maskapai. Dari grafik tersebut diketahui bahwa maskapai Mataram memiliki jumlah total penumpang terbanyak dan maskapai Rajawali memiliki jumlah penumpang yang paling sedikit. Perbandingan jumlah total penumpang tiap maskapai tidak memiliki perbedaan yang signifikan. Grafik tersebut diberi warna yang berbeda-beda bertujuan untuk membandingkan persentase perbandingan jumlah total penumpang tiap maskapai.

4. Plotting-Relationship

- Hubungan korelasi antara jumlah penumpang dan harga tiket

```
df.plot(kind="scatter", x = "penumpang", y ="harga", title = "Grafik Hubungan antara Jumlah Penumpang dan Harga Tiket")
plt.show()
```



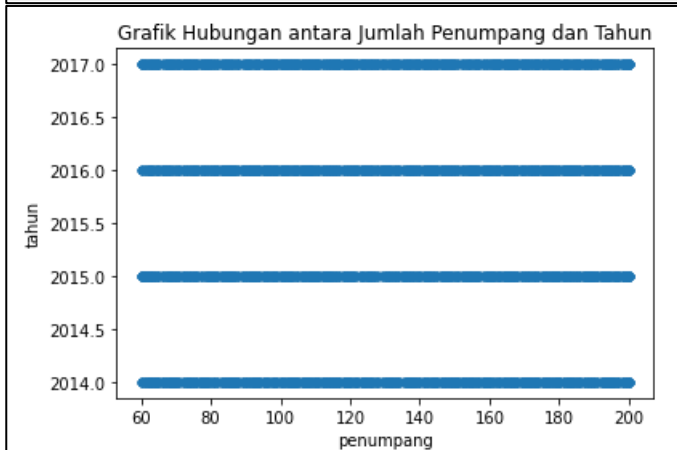
Keterangan : sumbu- x merupakan jumlah penumpang dan sumbu-y merupakan harga tiket pesawat

Dari grafik *scatter plot* di atas, dapat dilihat perbandingan hubungan antara jumlah penumpang dengan harga tiket pesawat, grafik tersebut menyatakan keeratan hubungan antara dua variabel. Karena grafik tersebut memiliki kemiringan atau gradien yang positif, maka variabel harga tiket pesawat dan

jumlah penumpang memiliki hubungan linier yang positif, hal ini juga dapat dibuktikan dengan mencari nilai korelasinya, dan didapatkan bahwa kedua variabel memiliki nilai korelasi yang mendekati 1. Oleh karena itu, dapat disimpulkan bahwa jika jumlah penumpang meningkat maka harga tiket pesawat juga akan meningkat

- Hubungan korelasi antara jumlah penumpang dan tahun

```
df.plot(kind="scatter", x = "penumpang", y ="tahun", title = "Grafik Hubungan antara Jumlah Penumpang dan Tahun")  
plt.show()
```



Keterangan : sumbu- x merupakan jumlah penumpang dan sumbu-y merupakan tahun

Dari grafik *scatter plot* di atas, dapat dilihat perbandingan hubungan antara jumlah penumpang dengan tahun, grafik tersebut menyatakan keeratan hubungan antara dua variabel. Karena grafik tersebut bentuknya menyebar, maka tahun dan jumlah penumpang tidak memiliki hubungan yang erat, hal ini juga dapat dibuktikan dengan mencari nilai korelasinya, dan didapatkan bahwa kedua variabel memiliki nilai korelasi yang mendekati 0. Oleh karena itu, dapat disimpulkan bahwa variabel tahun dan jumlah penumpang memiliki hubungan yang sangat kecil atau tidak memiliki hubungan sama sekali.

KORELASI

- Tunjukkan dan jelaskan korelasi antar semua atribut kuantitatif yang ada dalam tabel !
- Jelaskan dalam bentuk angka dan perjas dengan menggunakan visualisasi dalam bentuk grafik yang terkait !

```
data.corr()
```

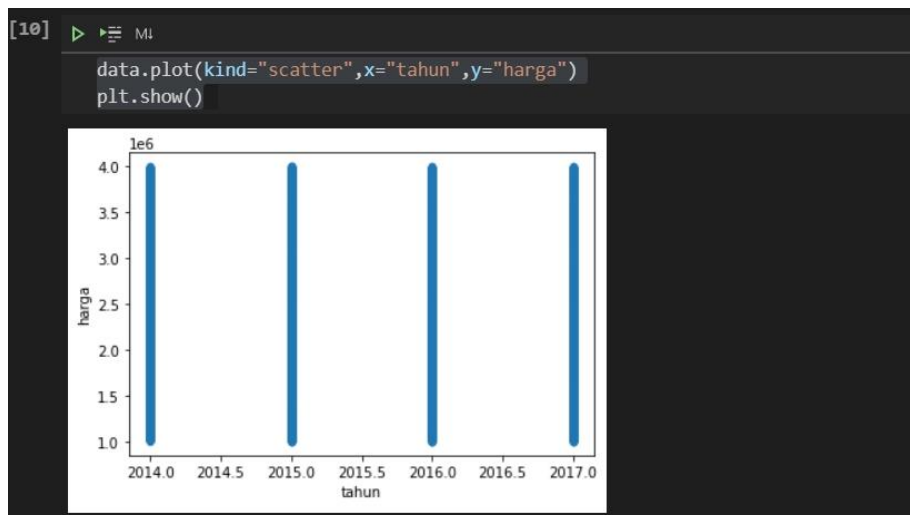
	tahun	penumpang	harga
tahun	1.000000	-0.004587	-0.010025
penumpang	-0.004587	1.000000	0.917927
harga	-0.010025	0.917927	1.000000

Dari matriks korelasi di atas kita tahu bahwa nilai koefisien korelasi antara:

- “tahun” dengan “penumpang” adalah -0.004587 (mendekati 0, tidak berkorelasi)
- “tahun” dengan “harga” adalah -0.010025 (mendekati 0, tidak berkorelasi)
- “penumpang” dengan “tahun” adalah -0.004587 (mendekati 0, tidak berkorelasi)
- “penumpang” dengan “harga” adalah 0.917927 (mendekati 1, berkorelasi positif)
- “harga” dengan “tahun” adalah -0.010025 (mendekati 0, tidak berkorelasi)
- “harga” dengan “penumpang” adalah 0.917927 (mendekati 1, berkorelasi positif)

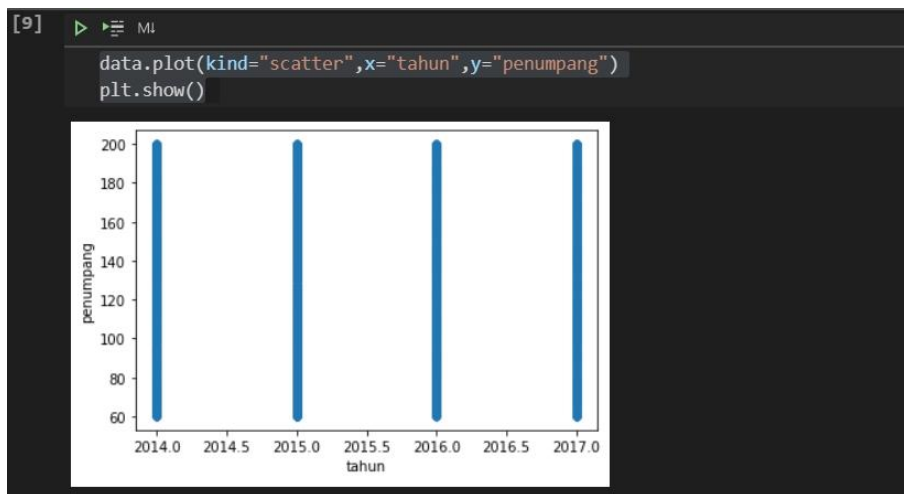
Visualisasi korelasi

1. Korelasi “tahun” dengan “penumpang”



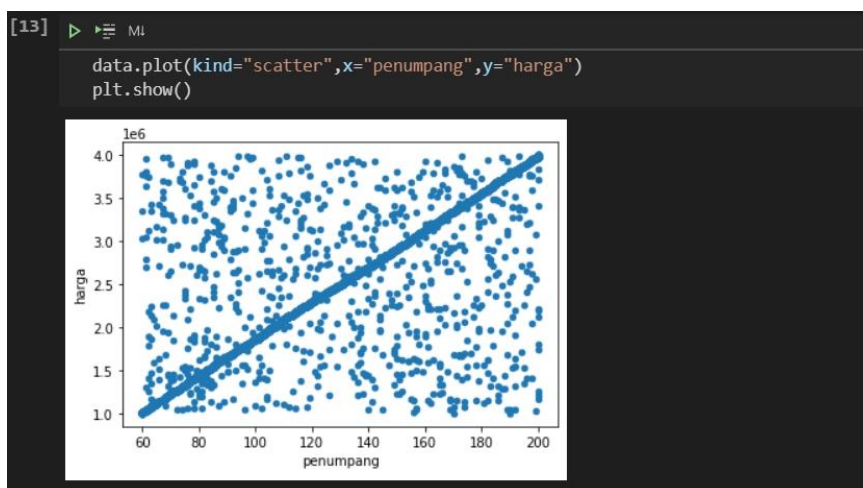
Dari visualisasi di atas dapat dilihat bahwa tahun dan penumpang tidak memiliki korelasi, karena ketika perubahan tahun, jumlah penumpang cenderung sama dan tidak ada dampak signifikan pada perubahan jumlah penumpang. Untuk bahasa pemrograman yang digunakan dalam membuat visualisasi ada di gambar di atas.

2. Korelasi “tahun” dengan “harga”



Sama dengan hubungan tahun dan penumpang, tahun dengan harga juga tidak memiliki korelasi ditunjukkan pada visualisasi, yaitu ketika tahun berbeda, harga cenderung sama. Untuk bahasa pemrograman yang digunakan dalam membuat visualisasi ada di gambar di atas.

3. Korelasi “penumpang” dengan “harga”



Hubungan antara penumpang dengan harga adalah berbanding lurus, ditunjukkan dengan bertambahnya penumpang, harga pun cenderung naik, walaupun ada beberapa kasus yang menunjukkan penumpang yang sedikit namun harga tinggi. Karena jumlah kasusnya sedikit, kasus-kasus tersebut dapat diabaikan, dan tetap dapat disimpulkan bahwa harga dan penumpang berbanding lurus. Untuk bahasa pemrograman yang digunakan dalam membuat visualisasi ada di gambar di atas.

PEMBAGIAN TUGAS

No	Nama	Pembagian Tugas
1	Joseph Hansel	Mengerjakan deskripsi data dan file serta korelasi.
2	Caleb Effendi	Mengerjakan visualisasi data.
3	Khairfirza Swandi	Mengerjakan karakteristik data dan statistik.