

# Eksplorasi Penyewaan Sepeda : Tren Musiman dan Suhu

oleh Dewi Maharani



Eksplorasi data mengenai tren penyewaan sepeda berdasarkan musim dan suhu.



# Latar Belakang

Sistem berbagi sepeda di kota besar menunjukkan penggunaan yang dipengaruhi musim dan cuaca, seperti suhu, kelembapan, dan kecepatan angin. Analisis ini bertujuan mengidentifikasi pola penggunaan dan mengoptimalkan layanan dengan data penyewaan sepeda 2011-2012.

Data Penyewaan Sepeda



# Alat yang Digunakan



# Pertanyaan Bisnis



1

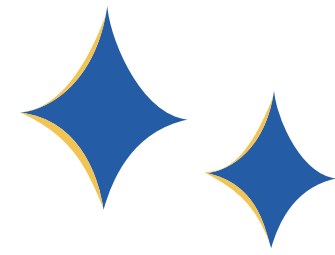
Bagaimana distribusi total penyewaan sepeda berdasarkan musim?

2

Apa hubungan antara suhu dan total penyewaan sepeda?



# Pengolahan Data



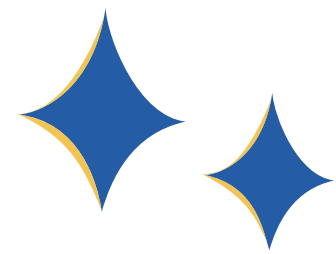
# Akuisisi Data

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
0	1	2011-01-01	1	0	1	0	6	0	2	0.344167	0.363625	0.805833	0.160446	331	654	985
1	2	2011-01-02	1	0	1	0	0	0	2	0.363478	0.353739	0.696087	0.248539	131	670	801
2	3	2011-01-03	1	0	1	0	1	1	1	0.196364	0.189405	0.437273	0.248309	120	1229	1349
3	4	2011-01-04	1	0	1	0	2	1	1	0.200000	0.212122	0.590435	0.160296	108	1454	1562
4	5	2011-01-05	1	0	1	0	3	1	1	0.226957	0.229270	0.436957	0.186900	82	1518	1600

	instant	dteday	season	yr	mnth	hr	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
0	1	2011-01-01	1	0	1	0	0	6	0	1	0.24	0.2879	0.81	0.0	3	13	16
1	2	2011-01-01	1	0	1	1	0	6	0	1	0.22	0.2727	0.80	0.0	8	32	40
2	3	2011-01-01	1	0	1	2	0	6	0	1	0.22	0.2727	0.80	0.0	5	27	32
3	4	2011-01-01	1	0	1	3	0	6	0	1	0.24	0.2879	0.75	0.0	3	10	13
4	5	2011-01-01	1	0	1	4	0	6	0	1	0.24	0.2879	0.75	0.0	0	1	1

Kode ini mengimpor dua dataset (day.csv dan hour.csv) dari GitHub ke Python, dengan data harian di day.csv dan data per jam di hour.csv.

Google Colab



# Pengukuran Data

Google Colab

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 731 entries, 0 to 730
Data columns (total 16 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   instant     731 non-null    int64
1   dteday      731 non-null    object
2   season      731 non-null    int64
3   yr          731 non-null    int64
4   mnth        731 non-null    int64
5   holiday     731 non-null    int64
6   weekday     731 non-null    int64
7   workingday  731 non-null    int64
8   weathersit   731 non-null    int64
9   temp        731 non-null    float64
10  atemp       731 non-null    float64
11  hum         731 non-null    float64
12  windspeed   731 non-null    float64
13  casual      731 non-null    int64
14  registered  731 non-null    int64
15  cnt         731 non-null    int64
```

Jumlah Duplikasi: 0

	instant	season	yr	mnth	holiday	weekday	workingday	weathersit
count	731.000000	731.000000	731.000000	731.000000	731.000000	731.000000	731.000000	731.000000
mean	366.000000	2.496580	0.500684	6.519836	0.028728	2.997264	0.683995	1.395349
std	211.165812	1.110807	0.500342	3.451913	0.167155	2.004787	0.465233	0.544894
min	1.000000	1.000000	0.000000	1.000000	0.000000	0.000000	0.000000	1.000000
25%	183.500000	2.000000	0.000000	4.000000	0.000000	1.000000	0.000000	1.000000
50%	366.000000	3.000000	1.000000	7.000000	0.000000	3.000000	1.000000	1.000000
75%	548.500000	3.000000	1.000000	10.000000	0.000000	5.000000	1.000000	2.000000
max	731.000000	4.000000	1.000000	12.000000	1.000000	6.000000	1.000000	3.000000

Ditemukan kolom dteday yang memiliki tipe data objek seharusnya bertipe data tanggal, dan kolom holiday yang bertipe int64 seharusnya bertipe boolean. Selain itu, data tersebut tidak mengandung duplikasi atau nilai yang hilang.



Google Colab

# Pengukuran Data



```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17379 entries, 0 to 17378
Data columns (total 17 columns):
#   Column      Non-Null Count  Dtype
---  -
0   instant     17379 non-null  int64
1   dteday      17379 non-null  object
2   season      17379 non-null  int64
3   yr          17379 non-null  int64
4   mnth        17379 non-null  int64
5   hr          17379 non-null  int64
6   holiday     17379 non-null  int64
7   weekday     17379 non-null  int64
8   workingday  17379 non-null  int64
9   weathersit   17379 non-null  int64
10  temp        17379 non-null  float64
11  atemp       17379 non-null  float64
12  hum         17379 non-null  float64
13  windspeed   17379 non-null  float64
14  casual      17379 non-null  int64
15  registered  17379 non-null  int64
16  cnt         17379 non-null  int64
```

Jumlah Duplikasi: 0

	instant	season	yr	mnth	hr	holiday
count	17379.0000	17379.000000	17379.000000	17379.000000	17379.000000	17379.00000
mean	8690.0000	2.501640	0.502561	6.537775	11.546752	0.02877
std	5017.0295	1.106918	0.500008	3.438776	6.914405	0.16716
min	1.0000	1.000000	0.000000	1.000000	0.000000	0.00000
25%	4345.5000	2.000000	0.000000	4.000000	6.000000	0.00000
50%	8690.0000	3.000000	1.000000	7.000000	12.000000	0.00000
75%	13034.5000	3.000000	1.000000	10.000000	18.000000	0.00000
max	17379.0000	4.000000	1.000000	12.000000	23.000000	1.00000

Pada data Hour\_df, ditemukan kolom dteday yang memiliki tipe data objek seharusnya bertipe tanggal, dan kolom holiday yang bertipe int seharusnya bertipe boolean, dengan tidak ada duplikasi atau nilai yang hilang dalam data tersebut.



# Pembersihan Data

Google Colab

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 731 entries, 0 to 730
Data columns (total 16 columns):
#   Column      Non-Null Count  Dtype
---  -
0   instant     731 non-null    int64
1   dteday      731 non-null    datetime64[ns]
2   season      731 non-null    int64
3   yr          731 non-null    int64
4   mnth        731 non-null    int64
5   holiday     731 non-null    bool
6   weekday     731 non-null    int64
7   workingday  731 non-null    int64
8   weathersit   731 non-null    int64
9   temp        731 non-null    float64
10  atemp       731 non-null    float64
11  hum         731 non-null    float64
12  windspeed   731 non-null    float64
13  casual      731 non-null    int64
14  registered  731 non-null    int64
15  cnt         731 non-null    int64
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17379 entries, 0 to 17378
Data columns (total 17 columns):
#   Column      Non-Null Count  Dtype
---  -
0   instant     17379 non-null  int64
1   dteday      17379 non-null  datetime64[ns]
2   season      17379 non-null  int64
3   yr          17379 non-null  int64
4   mnth        17379 non-null  int64
5   hr          17379 non-null  int64
6   holiday     17379 non-null  bool
7   weekday     17379 non-null  int64
8   workingday  17379 non-null  int64
9   weathersit   17379 non-null  int64
10  temp        17379 non-null  float64
11  atemp       17379 non-null  float64
12  hum         17379 non-null  float64
13  windspeed   17379 non-null  float64
14  casual      17379 non-null  int64
15  registered  17379 non-null  int64
16  cnt         17379 non-null  int64
```

Tipe data pada kolom dteday dan holiday telah diubah di kedua dataset, yaitu Day\_df dan Hour\_df.



# EDA

## Analisis Data Eksploratif

## Google Colab

	instant	dteday	season	yr	mnth	holiday
count	731.000000	731	731.000000	731.000000	731.000000	731
unique	NaN	NaN	NaN	NaN	NaN	2
top	NaN	NaN	NaN	NaN	NaN	False
freq	NaN	NaN	NaN	NaN	NaN	710
mean	366.000000	2012-01-01 00:00:00	2.496580	0.500684	6.519836	NaN
min	1.000000	2011-01-01 00:00:00	1.000000	0.000000	1.000000	NaN
25%	183.500000	2011-07-02 12:00:00	2.000000	0.000000	4.000000	NaN
50%	366.000000	2012-01-01 00:00:00	3.000000	1.000000	7.000000	NaN
75%	548.500000	2012-07-01 12:00:00	3.000000	1.000000	10.000000	NaN
max	731.000000	2012-12-31 00:00:00	4.000000	1.000000	12.000000	NaN
std	211.165812	NaN	1.110807	0.500342	3.451913	NaN

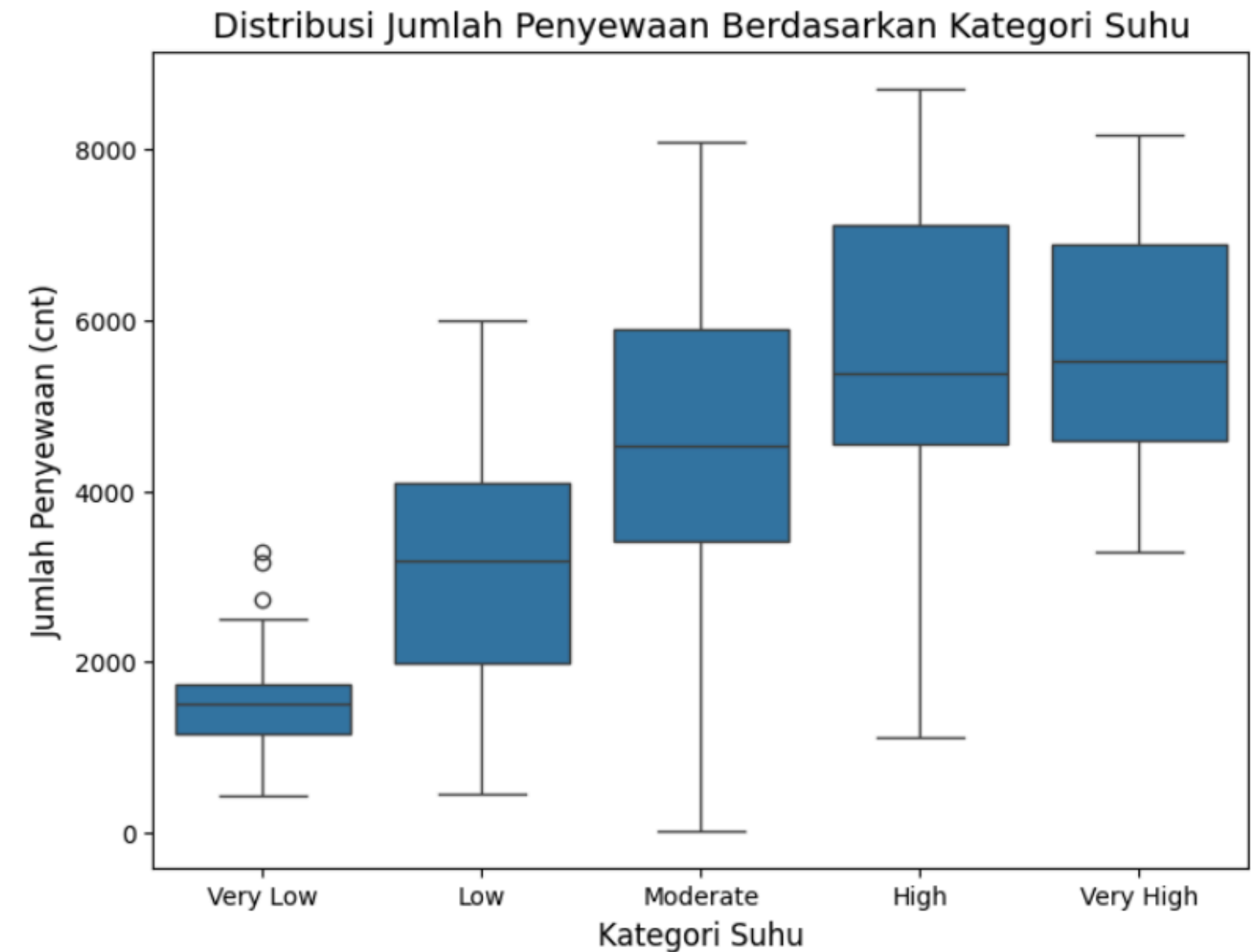
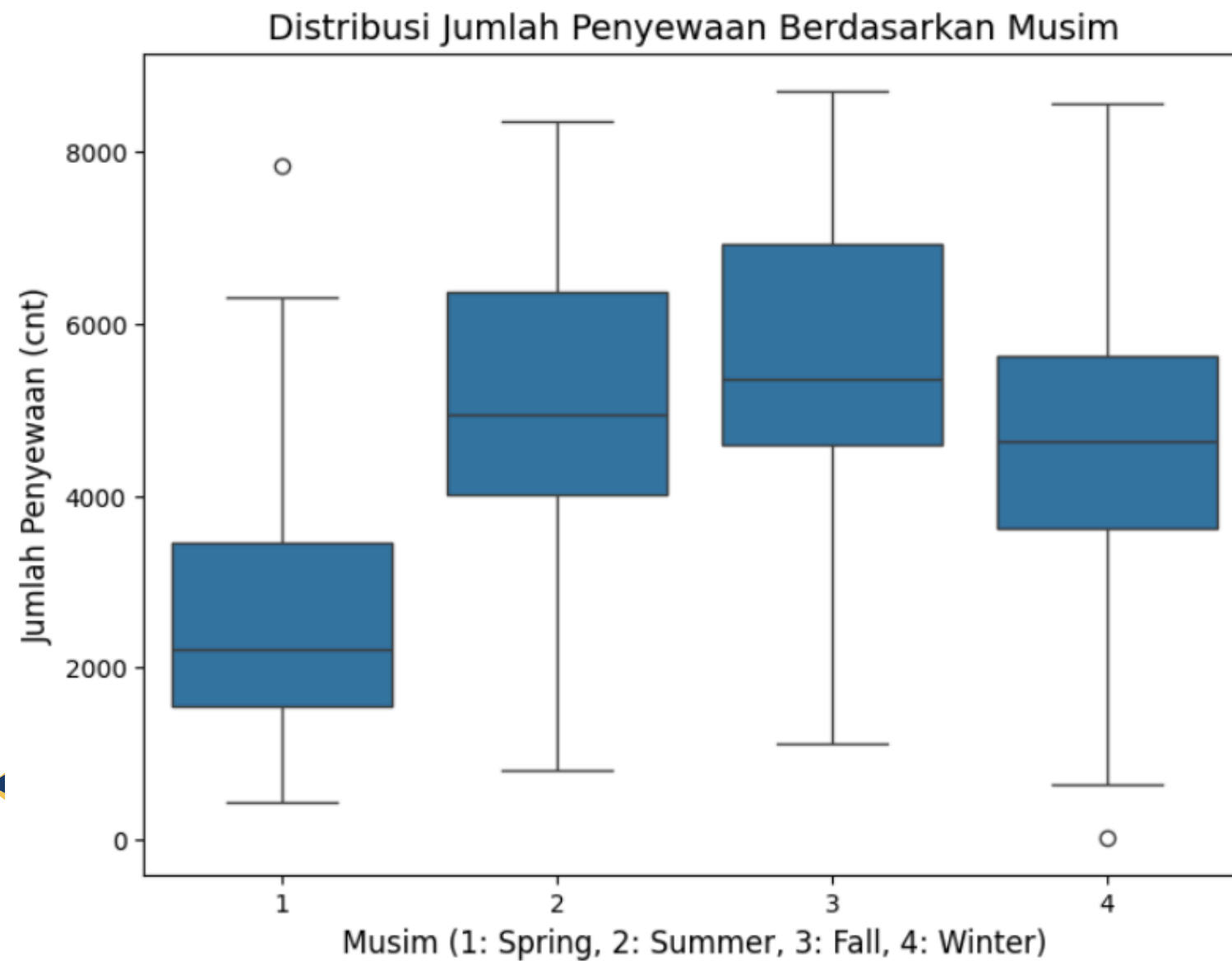
	mean cnt	median cnt	max cnt	min cnt
season				
1	2604.132597	2209.0	7836	431
2	4992.331522	4941.5	8362	795
3	5644.303191	5353.5	8714	1115
4	4728.162921	4634.5	8555	22

	mean cnt	median cnt	count cnt
temp_bins			
Very Low	1571.162791	1501.0	43
Low	3152.015464	3192.0	194
Moderate	4585.797688	4521.0	173
High	5674.958763	5384.0	194
Very High	5664.118110	5531.0	127

Musim Gugur dan Musim Panas menunjukkan angka penyewaan sepeda yang lebih tinggi dibandingkan musim lainnya, sementara penyewaan sepeda cenderung lebih tinggi pada suhu yang lebih tinggi.



# Visualisasi Data & Analisis Deskriptif



Berdasarkan boxplot, sewa sepeda lebih tinggi pada musim Gugur dan suhu Tinggi/Sangat Tinggi, serta lebih rendah pada musim Dingin dan suhu Sangat Rendah. Dan rentang IQR pada musim Gugur dan suhu Tinggi menunjukkan variasi yang lebih besar.

# Dashboard

Dashboard



Dashboard Streamlit ini memungkinkan pengguna untuk mengeksplorasi data sewa sepeda berdasarkan musim dan suhu dengan filter interaktif. Menampilkan statistik deskriptif, tabel pivot, dan boxplot untuk memvisualisasikan distribusi sewa berdasarkan musim dan kategori suhu.

## Dashboard Bike Sharing

Dashboard ini menyajikan eksplorasi data penyewaan sepeda berdasarkan musim dan suhu.

Pilih mode filter:

- ☐ Musim
- ☐ Suhu
- ☒ Musim dan Suhu

Pilih Musim

Summer x

Pilih Kategori Suhu

Low x

Very Low x

Moderate x

High x

Very High x

## Eksplorasi Data

Statistik deskriptif data yang telah difilter:

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp
unique	None	184	None	None	None	2	None	None	None	None
top	None	2011-03	None	None	None	False	None	None	None	None
freq	None	25	None	None	None	4493	None	None	None	None
mean	7,008.1328	None	2	0.5003	4.654	None	2.9915	0.6954	1.442	0.5447
std	4,543.9996	None	0	0.5001	0.9485	None	1.9949	0.4603	0.6518	0.1386



## ✦✦ **Kesimpulan** ✦✦

Berdasarkan analisis di atas, dapat disimpulkan bahwa Musim Gugur dan Musim Panas menunjukkan sewa sepeda yang lebih tinggi, ideal untuk promosi dan insentif. Sewa sepeda juga meningkat seiring suhu yang lebih hangat, sehingga manajer perlu menambah ketersediaan sepeda pada periode suhu puncak. Penempatan stasiun penyewaan di area dengan lalu lintas tinggi akan meningkatkan aksesibilitas bagi pengendara.

TERIMA KASIH

My\_Portofolio

