Proyek Analisis Data: Nama dataset

• Nama: Muhammad Amien Ramdhani

Email: dhanimacbull@gmail.com

• Id Dicoding: amienramdhani

Menentukan Pertanyaan Bisnis

- Berapa Kecepatan Rata rata angin pada tahun 2011
- Berapa jumlah penyewa sepeda yang telah terdaftar dan yang belum terdaftar di tiap tahun

Menyiapkan semua library yang dibuthkan

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Data Wrangling

Gathering Data

Ga	merin	y L	Jac	a									
	_df = _df.he			ad_csv('da	ay.cs	v')							
	instan	t		dteday	seas	on	yr	mnth	holi	day	weekda	ay w	orkingday
0		1	20	11-01-01		1	0	1		0		6	0
1		2	20	11-01-02		1	0	1		0		0	Θ
2		3	20	11-01-03		1	0	1		0		1	1
3		4	20	11-01-04		1	0	1		0		2	1
4		5	20	11-01-05		1	0	1		0		3	1
													-
	weathe istere		ıt \	temp		ate	mp	h	um w	vinds	peed	casua	ı L
0		u	2	0.344167	0.3	636	25	0.8058	33	0.160	9446	33	31
654 1 670			2	0.363478	0.3	537	39	0.6960	87	0.248	8539	13	31
2 122			1	0.196364	0.1	894	05	0.4372	73	0.248	8309	12	20
3	,		1	0.200000	0.2	121	22	0.5904	35	0.160	9296	16	08

```
1454
            1 0.226957 0.229270 0.436957
                                              0.186900
                                                            82
4
1518
    cnt
0
    985
1
   801
2
  1349
3
  1562
4
  1600
```

Assessing Data

```
Menilai Tabel day_df
day df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 731 entries, 0 to 730
Data columns (total 16 columns):
     Column
                 Non-Null Count
                                  Dtype
0
                 731 non-null
     instant
                                  int64
1
     dteday
                 731 non-null
                                  object
 2
                 731 non-null
     season
                                  int64
 3
                 731 non-null
                                  int64
     yr
 4
     mnth
                 731 non-null
                                  int64
 5
                 731 non-null
     holiday
                                  int64
 6
     weekday
                 731 non-null
                                  int64
 7
     workingday 731 non-null
                                  int64
 8
     weathersit 731 non-null
                                  int64
 9
                 731 non-null
                                  float64
     temp
 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
                731 non-null
    registered
                                  int64
15
                 731 non-null
                                  int64
     cnt
dtypes: float64(4), int64(11), object(1)
memory usage: 91.5+ KB
day_df.isna().sum()
              0
instant
              0
dteday
              0
season
yr
              0
              0
mnth
              0
holiday
              0
weekday
```

```
workingday
              0
weathersit
              0
              0
temp
atemp
              0
              0
hum
windspeed
              0
casual
              0
registered
              0
cnt
              0
dtype: int64
print("Jumlah duplikasi: ", day_df.duplicated().sum())
Jumlah duplikasi: 0
day df.describe()
          instant
                       season
                                       yr
                                                 mnth
                                                          holiday
weekday \
count 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
                     1.110807
std
       211.165812
                                 0.500342
                                             3.451913
                                                         0.167155
2.004787
                     1.000000
                                 0.000000
                                             1.000000
                                                         0.000000
min
         1.000000
0.000000
                                                         0.000000
25%
       183.500000
                     2.000000
                                 0.000000
                                             4.000000
1.000000
       366.000000
                     3.000000
                                             7.000000
                                                         0.000000
50%
                                 1.000000
3.000000
75%
       548.500000
                     3.000000
                                 1.000000
                                            10.000000
                                                         0.000000
5.000000
max
      731.000000
                     4.000000
                                 1.000000
                                            12.000000
                                                         1.000000
6.000000
       workingday
                   weathersit
                                     temp
                                                atemp
                                                              hum
windspeed \
                   731.000000
                               731.000000
                                           731.000000 731.000000
count 731.000000
731.000000
                     1.395349
                                             0.474354
mean
         0.683995
                                 0.495385
                                                         0.627894
0.190486
std
         0.465233
                     0.544894
                                 0.183051
                                             0.162961
                                                         0.142429
0.077498
                     1.000000
                                 0.059130
                                             0.079070
                                                         0.000000
         0.000000
min
0.022392
25%
         0.000000
                     1.000000
                                 0.337083
                                             0.337842
                                                         0.520000
0.134950
50%
         1.000000
                     1.000000
                                 0.498333
                                             0.486733
                                                         0.626667
0.180975
```

```
75%
                     2.000000
                                               0.608602
                                                           0.730209
         1.000000
                                  0.655417
0.233214
         1.000000
                     3.000000
                                  0.861667
                                               0.840896
                                                           0.972500
max
0.507463
            casual
                      registered
                                          cnt
        731.000000
                      731.000000
                                   731.000000
count
mean
        848.176471
                     3656.172367
                                  4504.348837
        686.622488
                     1560.256377
                                  1937.211452
std
min
          2.000000
                       20.000000
                                    22.000000
25%
        315.500000
                    2497.000000
                                  3152.000000
                    3662,000000
50%
        713.000000
                                  4548.000000
75%
       1096.000000
                    4776.500000
                                  5956.000000
       3410.000000
                    6946.000000
                                  8714.000000
max
```

Cleaning Data

Membersihkan day_df

Mengubah tipe data dteday menjadi tipe data datetime

```
day df['dteday'] = pd.to datetime(day df['dteday'])
day df['temp'] = day df['temp']*41
day df['atemp'] = day df['atemp']*50
day df['hum'] = day df['hum']*100
day df['windspeed'] = day df['windspeed']*67
col = ['season', 'holiday', 'weekday',
       'workingday', 'weathersit']
for i in col:
  if i in day df.columns.to list():
    day df[i] = day df[i].astype('category')
day df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 731 entries, 0 to 730
Data columns (total 16 columns):
     Column
                 Non-Null Count
                                  Dtype
     -----
- - -
                                  - - - - -
 0
                 731 non-null
                                  int64
     instant
 1
                 731 non-null
                                  datetime64[ns]
     dteday
2
                 731 non-null
     season
                                  category
 3
     yr
                 731 non-null
                                  int64
 4
     mnth
                 731 non-null
                                  int64
 5
     holiday
                 731 non-null
                                  category
 6
     weekday
                 731 non-null
                                  category
 7
     workingday 731 non-null
                                  category
```

```
8
                 731 non-null
     weathersit
                                  category
 9
                                  float64
     temp
                 731 non-null
 10
     atemp
                 731 non-null
                                  float64
 11
     hum
                 731 non-null
                                  float64
 12
     windspeed
                 731 non-null
                                  float64
 13
     casual
                 731 non-null
                                  int64
                                  int64
14
     registered
                 731 non-null
 15
                 731 non-null
                                  int64
     cnt
dtypes: category(5), datetime64[ns](1), float64(4), int64(6)
memory usage: 67.4 KB
day df.head(5)
               dteday season yr mnth holiday weekday workingday
   instant
weathersit
         1 2011-01-01
                            1
                                0
                                      1
                                                       6
2
1
                                                                  0
         2 2011-01-02
                            1
                                0
                                      1
                                                       0
2
2
                                              0
                                                                  1
         3 2011-01-03
                            1
                                0
                                      1
                                                       1
1
3
         4 2011-01-04
                                0
                                      1
                                              0
                                                       2
                                                                  1
                            1
1
4
         5 2011-01-05
                            1
                                0
                                      1
                                                       3
                                                                  1
1
                             hum
                                  windspeed
                                             casual
                                                      registered
        temp
                 atemp
                                                                   cnt
0
   14.110847
              18.18125
                        80.5833
                                  10.749882
                                                331
                                                             654
                                                                   985
1
   14.902598
              17.68695
                        69.6087
                                  16.652113
                                                131
                                                             670
                                                                   801
2
               9.47025
                        43.7273
                                  16.636703
                                                            1229
                                                                  1349
    8.050924
                                                120
3
    8.200000
              10.60610
                        59.0435
                                  10.739832
                                                108
                                                            1454
                                                                  1562
4
    9.305237
              11.46350
                        43.6957
                                  12.522300
                                                 82
                                                            1518
                                                                  1600
day df.describe()
          instant
                            yr
                                      mnth
                                                   temp
                                                              atemp
hum \
count 731.000000
                   731.000000 731.000000 731.000000 731.000000
731.000000
       366.000000
                     0.500684
                                  6.519836
                                             20.310776
                                                          23.717699
mean
62.789406
std
       211.165812
                     0.500342
                                  3.451913
                                              7.505091
                                                           8.148059
14.242910
         1.000000
                     0.000000
                                  1.000000
                                              2.424346
                                                           3.953480
min
0.000000
25%
       183.500000
                     0.000000
                                  4.000000
                                             13.820424
                                                          16.892125
52.000000
       366.000000
50%
                     1.000000
                                  7.000000
                                             20.431653
                                                          24.336650
62.666700
       548.500000
                     1.000000
                                 10.000000
                                             26.872076
                                                          30.430100
75%
```

```
73.020850
                     1.000000 12.000000 35.328347 42.044800
       731.000000
max
97.250000
        windspeed
                         casual
                                  registered
                                                       cnt
       731.000000
                    731.000000
                                  731.000000
                                               731.000000
count
        12.762576
                    848.176471
mean
                                 3656.172367
                                              4504.348837
std
         5.192357
                    686,622488
                                 1560.256377
                                              1937.211452
         1.500244
min
                       2.000000
                                   20.000000
                                                 22.000000
25%
         9.041650
                    315.500000
                                 2497.000000
                                              3152.000000
50%
        12.125325
                    713.000000
                                 3662.000000
                                              4548.000000
75%
        15.625371
                   1096.000000
                                 4776.500000
                                              5956.000000
max
        34.000021
                   3410.000000
                                 6946.000000 8714.000000
for i in col:
    print("Name of {} col".format(i)) #Name of Col
    print("No. of NUnique", day_df[i].nunique()) #Total Nunique Values
print("Unique Values", day_df[i].unique())# All unique vales
    print('*'*30) # to make differnce i each col
    print()
    print()
Name of season col
No. of NUnique 4
Unique Values [1, 2, 3, 4]
Categories (4, int64): [1, 2, 3, 4]
**********
Name of holiday col
No. of NUnique 2
Unique Values [0, 1]
Categories (2, int64): [0, 1]
*********
Name of weekday col
No. of NUnique 7
Unique Values [6, 0, 1, 2, 3, 4, 5]
Categories (7, int64): [0, 1, 2, 3, 4, 5, 6]
**********
Name of workingday col
No. of NUnique 2
Unique Values [0, 1]
Categories (2, int64): [0, 1] *************
Name of weathersit col
```

Exploratory Data Analysis (EDA)

Explore day_df

```
day_df.describe(include='all')
```

<ipython-input-40-a3f85d0fb009>:1: FutureWarning: Treating datetime
data as categorical rather than numeric in `.describe` is deprecated
and will be removed in a future version of pandas. Specify
`datetime_is_numeric=True` to silence this warning and adopt the
future behavior now.

day_df.describe(include='all')

	instant		dteday	/ season	yr
mnth \ count 731.0006	731.000000		731	731.0	731.000000
unique	NaN		731	4.0	NaN
NaN .					
top	NaN	2011-01-01	00:00:00	3.0	NaN
NaN					
freq	NaN		1	188.0	NaN
NaN					
first	NaN	2011-01-01	00:00:00) NaN	NaN
NaN					
last	NaN	2012-12-31	00:00:00) NaN	NaN
NaN					
mean	366.000000		NaN	l NaN	0.500684
6.519836			NI - N	. N N.	0 500242
std	211.165812		NaN	l NaN	0.500342
3.451913			NI N	I NaN	0 000000
min	1.000000		NaN	l NaN	0.000000
1.000000 25%			NaN	l NaN	0.000000
4.000000	183.500000		Ivar	n Nain	0.00000
50%	366.000000		NaN	l NaN	1.000000
7.00000			Ivai	n ivaiv	1.000000
75%	548.500000		NaN	l NaN	1.000000
10.00000			ivai	· Nan	1.000000
max	731.000000		NaN	l NaN	1.000000
12.00000			itai		11000000
	holiday we	eekday work:	ingday v	veathersit	temp

atemp \	731.0	731.0	731.0	731.0	731.000	200
count 731.0000		731.0	731.0	/31.0	/31.0000	900
unique	2.0	7.0	2.0	3.0		NaN
NaN top	0.0	0.0	1.0	1.0		NaN
NaN						
freq	710.0	105.0	500.0	463.0		NaN
NaN first	NaN	NaN	NaN	NaN	1	NaN
NaN	i i i i i i i i i i i i i i i i i i i	Han	Han	itait	•	lan
last	NaN	NaN	NaN	NaN		NaN
NaN mean	NaN	NaN	NaN	NaN	20.310	776
23.71769		IVAIV	IVAIN	IVAIV	20.310	770
std 8.148059	NaN	NaN	NaN	NaN	7.5050	991
0.140035 min	NaN	NaN	NaN	NaN	2.4243	346
3.953486	9					
25%	NaN	NaN	NaN	NaN	13.820	424
16.89212 50%	25 NaN	NaN	NaN	NaN	20.4310	653
24.33665						
75%	NaN	NaN	NaN	NaN	26.8720	976
30.43016 max	งง NaN	NaN	NaN	NaN	35.3283	347
42.04486			.10.1	110.11	33.323.	
	hum	windspeed	ca	sual reg	jistered	cnt
count	731.000000	731.000000	731.00	0000 731	.000000	731.000000
unique	NaN	NaN		NaN	NaN	NaN
·						
top	NaN	NaN		NaN	NaN	NaN
freq	NaN	NaN		NaN	NaN	NaN
first	NaN	NaN		NaN	NaN	NaN
last	NaN	NaN		NaN	NaN	NaN
mean	62.789406	12.762576	848.17	6471 3656	5.172367	4504.348837
std	14.242910	5.192357	686.62	2488 1560	.256377	1937.211452
min	0.000000	1.500244	2.00	0000 20	0.00000	22.000000
25%	52.000000		315.50	0000 2497	7.000000	3152.000000

50%	62.666700	12.125325	713.000000	3662.000000	4548.000000
75%	73.020850	15.625371	1096.000000	4776.500000	5956.000000
max	97.250000	34.000021	3410.000000	6946.000000	8714.000000

Mengurutkan Jumlah Peminjam Sepeda berdasarkan tahunnya

```
day_df.groupby(by="yr").cnt.nunique().sort_values(ascending=False)
yr
1     356
0     352
Name: cnt, dtype: int64
```

Mengurutkan Jumlah Peminjam Sepeda berdasarkan bulan

```
day_df.groupby(by="mnth").cnt.nunique().sort_values(ascending=False)
mnth
1
      62
5
      62
7
      62
8
      62
10
      62
12
      62
3
      60
4
      60
6
      60
9
      60
11
      60
      57
Name: cnt, dtype: int64
```

Mengurutkan Jumlah Peminjam Sepeda berdasarkan Musim

```
day_df.groupby(by="season").cnt.nunique().sort_values(ascending=False)
season
3    188
2    182
1    179
4    176
Name: cnt, dtype: int64
```

Mencari Kecepatan maksimal,minimal,dan rata-rata dari angin berdasarkan tahun

```
day_df.groupby(by="yr").agg({
    "windspeed": ["max", "min", "mean"]
})

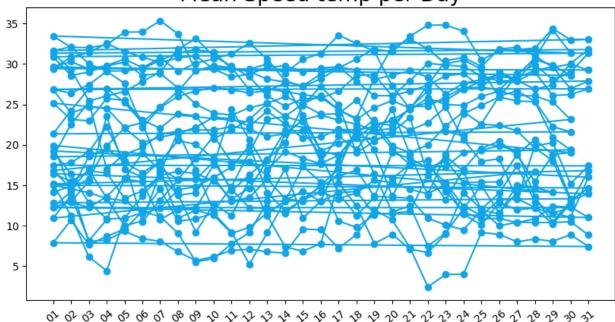
windspeed
    max    min    mean
yr
0     34.000021    1.500244    12.823977
1     29.584721    3.125550    12.701344
```

Mencari maksimal, minimal, dan rata-rata dari temperature berdasarkan tahun

```
day df.groupby(by="yr").agg({
    "temp": ["max", "min", "mean"]
})
         temp
                    min
          max
                              mean
yr
    34.815847
               2.424346 19.953263
0
    35.328347 4.407500 20.667313
1
day df.groupby(by="season").agg({
    "temp": ["max", "min", "mean"]
})
             temp
              max
                         min
                                    mean
season
        23.472500
1
                    2.424346 12.207650
2
        33.141653
                   10.374763 22.320611
3
        35.328347
                   19.235847
                              28.958682
4
        26.957500
                    9.054153
                              17.339148
day_df.groupby(by="season").agg({
    "windspeed": ["max", "min", "mean"]
})
        windspeed
              max
                        min
                                   mean
season
        34.000021
                   3.042356
                             14.373984
1
        26.000489
2
                   4.417256
                             13.634978
3
        25.166339
                   4.292744
                             11.530366
4
        28.292425
                   1.500244
                             11.523637
day temp = day df.resample(rule='D', on='dteday').agg({
    "temp": "mean"
})
day temp.index = day temp.index.strftime('%d')
```

```
plt.figure(figsize=(10, 5))
plt.plot(day_temp.index, day_temp['temp'], marker='o',
color='#0ea5e9')
plt.xticks(rotation=40)
plt.title(label='Mean Speed temp per Day', loc='center', fontsize=20)
plt.show()
```

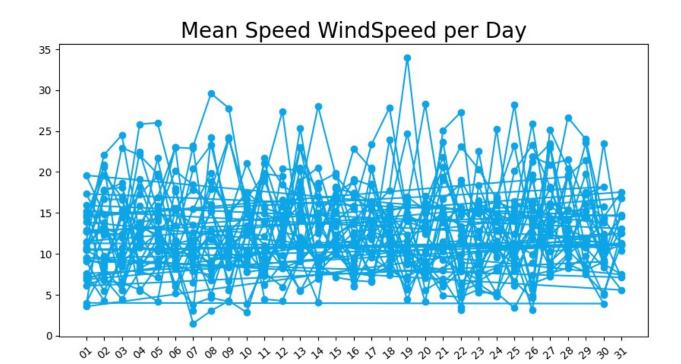
Mean Speed temp per Day



```
day_wind = day_df.resample(rule='D', on='dteday').agg({
    "windspeed": "mean"
})

day_wind.index = day_wind.index.strftime('%d')

plt.figure(figsize=(10, 5))
plt.plot(day_wind.index, day_wind['windspeed'], marker='o',
color='#0ea5e9')
plt.xticks(rotation=40)
plt.title(label='Mean Speed WindSpeed per Day', loc='center',
fontsize=20)
plt.show()
```



Visualization & Explanatory Analysis

Pertanyaan 1:

Berapa Rata Rata Kecepatan Angin Pada tahun 2011, Berapa Suhu Temperature pada tahun 2011

```
day df.groupby(by="yr").agg({
    "windspeed": ["max", "min", "mean"]
})
    windspeed
           max
                      min
                                 mean
yr
    34.000021
                1.500244
                           12.823977
0
    29.584721
               3.125550
                           12.701344
day df[day df.mnth == 4]
     instant
                  dteday season
                                   yr
                                       mnth holiday weekday workingday
90
           91 2011-04-01
                                2
                                    0
                                           4
                                2
                                                             6
91
           92 2011-04-02
                                    0
                                           4
                                                    0
                                                                         0
92
           93 2011-04-03
                                    0
                                                    0
                                                             0
                                                                         0
                                           4
                                2
                                                                         1
93
           94 2011-04-04
                                    0
                                           4
                                                    0
                                                             1
94
           95 2011-04-05
                                2
                                    0
                                           4
                                                    0
                                                             2
                                                                         1
95
           96 2011-04-06
                                2
                                           4
                                                             3
                                                                         1
                                    0
                                                    0
                                2
                                                    0
                                                             4
                                                                         1
96
           97 2011-04-07
                                    0
                                           4
97
                                2
                                                             5
           98 2011-04-08
                                           4
                                                    0
                                                                         1
                                    0
           99 2011-04-09
                                                    0
                                                             6
98
                                    0
                                           4
                                                                         0
```

00	100 2011 04 10	2	0	4	0	0	0
99	100 2011-04-10 101 2011-04-11	2	0	4	0	0	0
100		2	0	4	0	1	1
101	102 2011-04-12	2	0	4	0	2	1
102	103 2011-04-13	2	0	4	0	3	1
103	104 2011-04-14		0	4	0	4	1
104	105 2011-04-15	2 2	0	4	1	5	0
105 106	106 2011-04-16 107 2011-04-17	2	0 0	4 4	0 0	6	0
107	107 2011-04-17	2	0	4	0	0 1	0 1
107	109 2011-04-19	2	0	4	0	2	1
100	110 2011-04-19	2	0	4	0	3	1
110	111 2011-04-21	2	0	4	0	4	1
111	112 2011 04 21	2	0	4	0	5	1
112	113 2011-04-23	2	0	4	0	6	0
113	114 2011-04-24	2	0	4	0	Õ	0
114	115 2011-04-25	2	0	4	0	1	1
115	116 2011-04-26	2	0	4	0	2	1
116	117 2011-04-27	2	0	4	0	3	1
117	118 2011-04-28	2	0	4	0	4	1
118	119 2011-04-29	2	0	4	0	5	1
119	120 2011-04-30	2	0	4	0	6	0
456	457 2012-04-01	2	1	4	0	0	0
457	458 2012-04-02	2	1	4	0	1	1
458	459 2012-04-03	2	1	4	0	2	1
459	460 2012-04-04	2	1	4	0	3	1
460	461 2012-04-05	2	1	4	0	4	1
461	462 2012-04-06 463 2012-04-07	2 2	1 1	4	0	5	1
462 463	463 2012-04-07 464 2012-04-08	2	1	4 4	0 0	6 0	0 0
464	465 2012-04-09	2	1	4	0	1	1
465	466 2012-04-10	2	1	4	0	2	1
466	467 2012-04-11	2	ī	4	0	3	i
467	468 2012-04-12	2	1	4	0	4	1
468	469 2012-04-13	2	1	4	0	5	1
469	470 2012-04-14	2	1	4	0	6	0
470	471 2012-04-15	2	1	4	0	0	0
471	472 2012-04-16	2	1	4	1	1	0
472	473 2012-04-17	2	1	4	0	2	1
473	474 2012-04-18	2	1	4	0	3	1
474	475 2012-04-19	2	1	4	0	4	1
475	476 2012-04-20	2	1	4	0	5	1
476	477 2012-04-21	2	1	4	0	6	0
477	478 2012-04-22	2	1	4	0	0	0
478	479 2012-04-23	2	1	4	0	1	1
479	480 2012-04-24	2 2	1	4	0	2	1
480 481	481 2012-04-25 482 2012-04-26	2	1 1	4 4	0 0	3 4	1 1
482	483 2012-04-27	2	1	4	0	5	1
483	484 2012-04-28	2	1	4	0	6	0
1 05	707 Z01Z-04-Z0	_	_	7	J	9	U

484 485			12-04-29 12-04-30	2		1 1	4 4		0 0		9 1		0 1
weath	nersi	+	temp		ate	emp		hum	wir	ndspeed	cas	แลโ	
register			cemp		u c ·	Silip		TTG.	W	laspeca	cus	uuc	
90		2	12.300000	14.	172	270	68.	6250	17.	333436		307	
1920													
91		2	12.915000	15.	783	185	65.	3750	13.	208782		898	
1354		_									_		
92		1	15.511653	18.	938	335	48.	0000	12.	208271	1	651	
1598		1	22 506652	27	1 4 /	245	42	6250	25	022257		724	
93 2381		1	23.506653	27.	140	045	42.	6250	25.	833257		734	
94		2	16.980847	19.	01:	750	64	2083	26	000489		167	
1628		2	10.900047	19.	91	/ 30	04.	2003	20.	000409		107	
95		1	16.024153	19.	380	140	47	0833	17	625221		413	
2395		_	10.024133	19.	500	J - U	7/.	0033	Ι/.	023221		713	
96		1	17.937500	21.	684	480	60.	2917	10.	874904		571	
2570		_											
97		2	13.769153	16.	223	395	83.	6250	15.	208464		172	
1299													
98		2	14.042500	17.	076	645	87.	7500	8.	916561		879	
1576													
99		2	17.493347	21.	336	585	85.	7500	9.	833389	1	188	
1707													
100		2	24.421732	28.	260	985	71.	6956	21.	739758		855	
2493		_	20 602500	2.4	.			0167	10	41.0000		.	
101		2	20.602500	24.	654	2/0	/3.	9167	18.	416893		257	
1777		2	16 012500	20	06	41E	0.1	0167	16	701220		200	
102 1953		2	16.912500	20.	004	+13	01.	9167	10.	791339		209	
1933		1	19.167500	23.	13.	710	5/	0417	7	416900		529	
2738		_	19.10/300	23.	I)	7 10	J 4 .	0417	, .	410900		J29	
104		1	18.313347	22.	091	565	67.	1250	15.	167125		642	
2484		_	101313317		03.	, ,	07.	1230		10,123		0 12	
105		3	17.664153	21.	274	460	88.	8333	22.	834136		121	
674													
106		1	18.723347	22.	284	480	47.	9583	20.	334232	1	558	
2186													
107		1	21.012500	25.	15	730	54.	2500	10.	958989		669	
2760													
108		2	20.739153	24.	462	290	66.	5833	10.	584057		409	
2795			24 205000	20	21		6.1	4167	1.0	222275		610	
109		1	24.395000	28.	219	960	bΙ.	4167	16.	208975		613	
3331		1	10 025047	22	60	160	40	7002	21	702206		715	
110 3444		1	18.825847	22.	094	+00	40.	7083	21.	792286		745	
111		2	13.803347	16.	ΘO.	770	72	9583	1/	707907		177	
1506		_	13.003347	10.	U J	, , 0	/	5505	17.	101301		± / /	
2300													

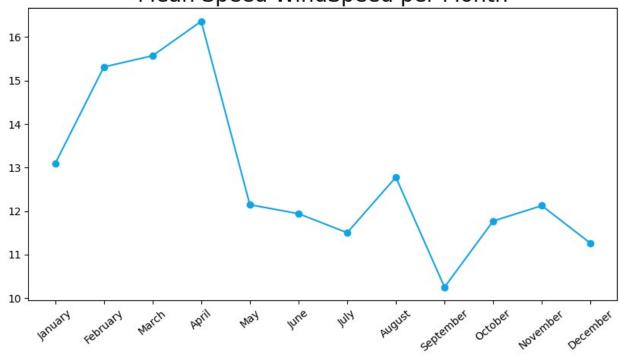
112	2	18.860000	22.50605	88.7917	15.458575	1462
2574 113	2	23.848347	27.58815	81.0833	12.875725	1710
2481		23:040347	27.30013	01.0055	12:075725	1710
114	1	24.873347	28.72500	77.6667	12.417311	773
3300	1	25 000247	20 70415	72 0167	21 075500	670
115 3722	1	25.898347	29.70415	72.9167	21.875500	678
116	2	25.420000	28.75710	83.5417	20.917400	547
3325	_					•
117	2	25.317500	28.94645	70.0833	21.500836	569
3489	_					
118	1	20.910000	24.87315	45.7083	16.084221	878
3717 119	1	19.372500	23.20105	50.3333	15.750025	1965
3347	1	19.372300	23.20103	30.3333	13.730023	1905
456	2	17.459153	20.86435	67.6250	11.541889	2347
3694						
457	1	17.790433	21.37565	50.4348	20.913313	1208
4728	_	10 100017	22 27415	20 6667	6 700011	10.40
458 5424	1	19.133347	23.07415	39.6667	6.708911	1348
459	1	22.208347	26.67250	46.9583	12.125325	1058
5378		22.200547	20.07250	40.3303	12.123323	1050
460	1	17.835000	21.55815	37.4167	14.708443	1192
5265						
461	1	16.536653	19.53835	37.7083	20.125996	1807
4653	1	17 027500	21 20045	25 4167	10 416257	2252
462 3605	1	17.937500	21.30645	25.4167	18.416357	3252
463	1	20.500000	24.62125	27.5833	15.583932	2230
2939	_	20130000	21102123	2713333	13.303331	2230
464	1	20.055847	23.83190	31.7500	23.999132	905
4680	_					
465	1	18.313347	21.81165	43.5000	16.708125	819
5099 466	1	14.296536	16.86370	46.9565	19.783358	482
4380		171230330	10.00570	+019303	131703330	702
467	1	16.297500	19.38020	46.6250	19.458743	663
4746						
468	1	18.142500	21.59040	40.8333	10.416557	1252
5146	1	20 205000	24 20000	E0 2017	12 701420	2705
469 4665	1	20.295000	24.39980	50.2917	12.791439	2795
470	1	24.873347	28.69375	50.7917	15.083643	2846
4286	_	,,		2017317		_0.0
471	1	27.230847	30.74625	56.1667	19.083543	1198
5172		04.045.55	20 02 125	20 6415	10 000115	6.00
472	1	24.941653	29.92435	39.0417	18.333143	989

5702 473		2	18.996653	22.85190	56.9167	11.250104	347
4020		2	10.990003	22.03190	20.910/	11.250104	347
474		1	20.431653	24.65230	61.2500	4.417256	846
5719		1	21 502247	2F 7007F	60 4502	10 041257	1240
475 5950		1	21.593347	25.78875	69.4583	10.041357	1340
476		1	23.370000	27.14605	68.2917	19.000329	2541
4083		_					
477		3	16.263347	19.47520	83.5417	23.084582	120
907 478		2	13.188347	15.05625	76.6667	20.334232	195
3019		_	131100317	13.03023	, 0.000,	20.55.252	100
479		1	16.946653	20.26415	45.4167	16.708661	518
5115 480		1	19.543347	23.51585	42.7917	7.959064	655
5541			19.040047	73,31303	74./31/	7.333004	000
481		2	20.431653	24.17915	75.6667	11.833875	475
4551		-	10 757500	22 62105	40 0022	22 201411	1014
482 5219		1	18.757500	22.63185	40.0833	23.291411	1014
483		2	15.443347	18.87520	48.9583	8.708325	1120
3100							
484		1	18.791653	22.50605	58.7083	7.832836	2229
4075 485		2	19.030847	22.88480	57.0000	11.499746	665
4907							
	cnt						
	227						
	252						
	249						
	115						
	795						
	808 141						
	471						
	455						
	895						
	348						
	034						
	162						
	267 126						
	126 795						
	795 744						
	429						
	204						
	944						

```
110 4189
111 1683
112 4036
113 4191
114 4073
115 4400
116 3872
117 4058
118 4595
119 5312
456 6041
457 5936
458 6772
459 6436
460 6457
461 6460
462 6857
463 5169
464 5585
465 5918
466 4862
467 5409
468 6398
469 7460
470 7132
471 6370
472 6691
473 4367
474 6565
475
    7290
476 6624
477 1027
478 3214
479 5633
480 6196
481 5026
482 6233
483 4220
484 6304
485 5572
year windspeed = day df[day df.dteday.dt.year == 2011]
year_windspeed = year_windspeed.resample(rule='M', on='dteday').agg({
    "windspeed": "mean"
})
year_windspeed.index = year_windspeed.index.strftime('%B')
plt.figure(figsize=(10, 5))
plt.plot(year windspeed.index, year windspeed['windspeed'],
```

```
marker='o', color='#0ea5e9')
plt.xticks(rotation=40)
plt.title(label='Mean Speed WindSpeed per Month', loc='center',
fontsize=20)
plt.show()
```

Mean Speed WindSpeed per Month

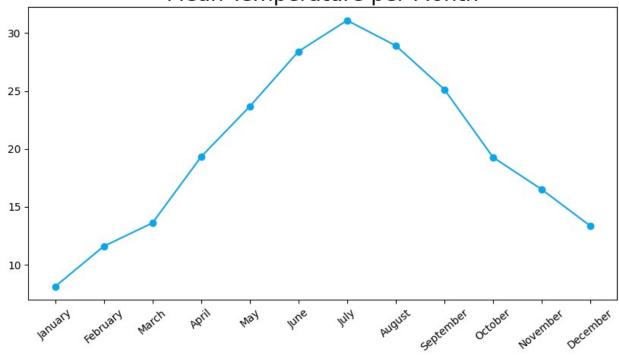


```
year_temp = day_df[day_df.dteday.dt.year == 2011]
year_temp = year_temp.resample(rule='M', on='dteday').agg({
    "temp": "mean"
})

year_temp.index = year_temp.index.strftime('%B')

plt.figure(figsize=(10, 5))
plt.plot(year_temp.index, year_temp['temp'], marker='o',
color='#0ea5e9')
plt.xticks(rotation=40)
plt.title(label='Mean Temperature per Month', loc='center',
fontsize=20)
plt.show()
```

Mean Temperature per Month

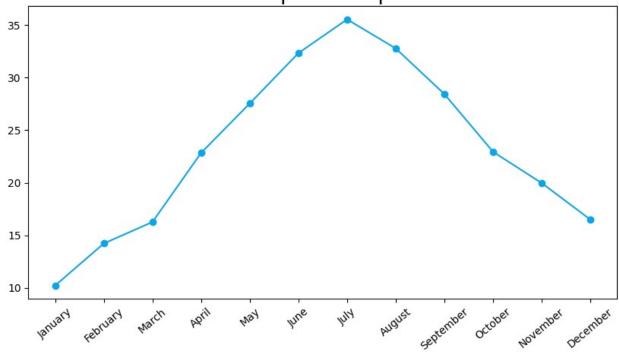


```
year_atemp = day_df[day_df.dteday.dt.year == 2011]
year_atemp = year_atemp.resample(rule='M', on='dteday').agg({
    "atemp": "mean"
})

year_atemp.index = year_atemp.index.strftime('%B')

plt.figure(figsize=(10, 5))
plt.plot(year_atemp.index, year_atemp['atemp'], marker='o',
color='#0ea5e9')
plt.xticks(rotation=40)
plt.title(label='Mean Atemperature per Month', loc='center',
fontsize=20)
plt.show()
```



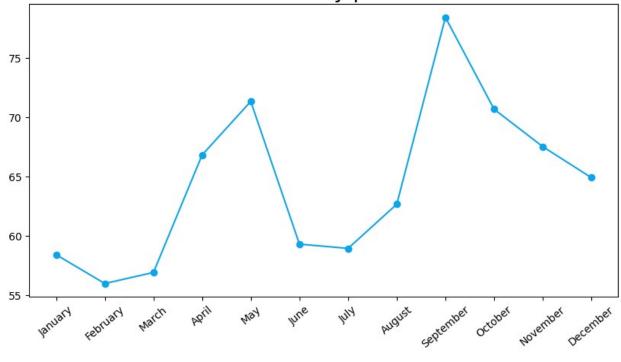


```
year_hum = day_df[day_df.dteday.dt.year == 2011]
year_hum = year_hum.resample(rule='M', on='dteday').agg({
    "hum": "mean"
})

year_hum.index = year_hum.index.strftime('%B')

plt.figure(figsize=(10, 5))
plt.plot(year_hum.index, year_hum['hum'], marker='o', color='#0ea5e9')
plt.xticks(rotation=40)
plt.title(label='Mean Humidity per Month', loc='center', fontsize=20)
plt.show()
```

Mean Humidity per Month



day_	df[day_d	f.mnth == 4]]						
	instant		season	yr		holiday	weekday	workingday	\
90		2011-04-01	2	0	4	0	5	1	
91		2011-04-02	2	0	4	0	6	0	
92		2011-04-03	2	0	4	0	0	0	
93		2011-04-04	2	0	4	0	1	1	
94	95	2011-04-05	2	0	4	0	2	1	
95	96	2011-04-06	2	0	4	0	3	1	
96	97	2011-04-07	2	0	4	0	4	1	
97	98	2011-04-08	2	0	4	0	5	1	
98	99	2011-04-09	2	0	4	Θ	6	Θ	
99	100	2011-04-10	2	0	4	0	0	0	
100	101	2011-04-11	2	0	4	0	1	1	
101	102	2011-04-12	2	0	4	0	2	1	
102	103	2011-04-13	2	0	4	0	3	1	
103	104	2011-04-14	2	0	4	0	4	1	
104	105	2011-04-15	2	0	4	1	5	0	
105	106	2011-04-16	2	0	4	0	6	0	
106	107	2011-04-17	2	0	4	Θ	Θ	Θ	
107	108	2011-04-18	2	0	4	Θ	1	1	
108	109	2011-04-19	2	0	4	0	2	1	
109	110	2011-04-20	2	0	4	Ō	3	$\bar{1}$	
110	111	2011-04-21	2	0	4	0	4	$\bar{1}$	
111		2011-04-22	2	0	4	0	5	1	
112		2011-04-23	2	0	4	0	6	0	
			_	•	•				

113	114	2011-04-24	2	0	4	0	0		0
114	115	2011-04-25	2	0	4	0	1		1
115 116	_	2011-04-26 2011-04-27	2 2	0	4 4	0 0	2		1 1
117		2011-04-27	2	0	4	0	4		1
118	_	2011-04-29	2	0	4	0	5		1
119 456		2011-04-30 2012-04-01	2 2	0 1	4 4	0 0	6 0		0 0
457	458	2012-04-02	2	1	4	0	1		1
458 459		2012-04-03 2012-04-04	2 2	1 1	4 4	0 0	2 3		1 1
460		2012-04-04	2	1	4	0	4		1
461		2012-04-06	2	1	4	0	5		1
462 463		2012-04-07 2012-04-08	2 2	1 1	4 4	0 0	6 0		0 0
464	465	2012 - 04 - 09	2	1	4	Ö	1		1
465		2012-04-10 2012-04-11	2	1	4	0	2		1
466 467		2012-04-11	2 2	1 1	4 4	0 0	3 4		1 1
468		2012-04-13	2	1	4	0	5		1
469 470	_	2012-04-14 2012-04-15	2 2	1 1	4 4	0 0	6 0		0 0
471		2012-04-15	2	1	4	1	1		0
472		2012 - 04 - 17	2	1	4	0	2		1
473 474		2012-04-18 2012-04-19	2 2	1 1	4 4	0 0	3 4		1 1
475	476	2012-04-20	2	1	4	0	5		1
476 477		2012-04-21 2012-04-22	2 2	1 1	4 4	0 0	6 0		0 0
478	_	2012-04-22	2	1	4	0	1		1
479		2012 - 04 - 24	2	1	4	0	2		1
480 481		2012-04-25 2012-04-26	2 2	1 1	4 4	0 0	3 4		1 1
482		2012-04-27	2	1	4	0	5		1
483 484		2012-04-28 2012-04-29	2 2	1 1	4 4	0 0	6 0		0 0
485		2012 - 04 - 30	2	1	4	Ö	1		1
weat	hersi	t temp	at	emp	hum	winds	oeed (casual	
register	-			·	60 6250			207	
90 1920		2 12.300000	14.17	/2/0	68.6250	17.333	3430	307	
91 1354		2 12.915000	15.78	3185	65.3750	13.208	3782	898	
92 1598		1 15.511653	18.93	3835	48.0000	12.208	3271	1651	
93 2381		1 23.506653	27.14	1645	42.6250	25.833	3257	734	
94		2 16.980847	19.91	L750	64.2083	26.000	9489	167	

1628						
95	1	16.024153	19.38040	47.0833	17.625221	413
2395 96	1	17.937500	21.68480	60.2917	10.874904	571
2570		17.937500	21.00400	00.2917	10.0/4904	5/1
97	2	13.769153	16.22395	83.6250	15.208464	172
1299	_	131703133	10.22333	03.0230	131200101	172
98	2	14.042500	17.07645	87.7500	8.916561	879
1576						
99	2	17.493347	21.33685	85.7500	9.833389	1188
1707						
100	2	24.421732	28.26085	71.6956	21.739758	855
2493						
101	2	20.602500	24.65270	73.9167	18.416893	257
1777	_	16 012500	20 06415	01 0167	16 701220	200
102	2	16.912500	20.86415	81.9167	16.791339	209
1953 103	1	19.167500	23.13710	54.0417	7.416900	529
2738	Т	19.10/500	23.13/10	J4.841/	7.410900	329
104	1	18.313347	22.09565	67.1250	15.167125	642
2484	_	10.515547	22.03303	07.1230	13.107123	042
105	3	17.664153	21,27460	88.8333	22.834136	121
674						
106	1	18.723347	22.28480	47.9583	20.334232	1558
2186						
107	1	21.012500	25.15730	54.2500	10.958989	669
2760	_					
108	2	20.739153	24.46290	66.5833	10.584057	409
2795	1	24 205000	20 21060	61 4167	16 200075	612
109 3331	1	24.395000	28.21960	61.4167	16.208975	613
110	1	18.825847	22.69460	40.7083	21.792286	745
3444	_	10.023047	22.03400	40.7005	21.732200	743
111	2	13.803347	16.09770	72.9583	14.707907	177
1506	_					= • •
112	2	18.860000	22.50605	88.7917	15.458575	1462
2574						
113	2	23.848347	27.58815	81.0833	12.875725	1710
2481	-	24 072247	20 72500	77 6667	10 417011	770
114	1	24.873347	28.72500	77.6667	12.417311	773
3300 115	1	25.898347	29.70415	72.9167	21.875500	678
3722	1	23.09034/	29.70413	12.9107	21.0/3300	0/0
116	2	25.420000	28.75710	83.5417	20.917400	547
3325	_	231120000	20173710	3313417	201317400	547
117	2	25.317500	28.94645	70.0833	21.500836	569
3489						
118	1	20.910000	24.87315	45.7083	16.084221	878
3717						

119	1	19.372500	23.20105	50.3333	15.750025	1965
3347 456	2	17.459153	20.86435	67.6250	11.541889	2347
3694	1	17 700422	21 27565	FO 4240	20 012212	1200
457 4728	1	17.790433	21.37565	50.4348	20.913313	1208
458	1	19.133347	23.07415	39.6667	6.708911	1348
5424 459	1	22.208347	26.67250	46.9583	12.125325	1058
5378	-	17 025000	21 55015	27 4167	14 700442	1100
460 5265	1	17.835000	21.55815	37.4167	14.708443	1192
461	1	16.536653	19.53835	37.7083	20.125996	1807
4653 462	1	17.937500	21.30645	25.4167	18.416357	3252
3605		17.937300	21.30043	23.4107	10.410337	3232
463	1	20.500000	24.62125	27.5833	15.583932	2230
2939 464	1	20.055847	23.83190	31.7500	23.999132	905
4680		20.033647	23.03190	31.7300	23.999132	903
465	1	18.313347	21.81165	43.5000	16.708125	819
5099 466	1	14.296536	16.86370	46.9565	19.783358	482
4380		14.290550	10.00570	40.9303	19.705550	402
467	1	16.297500	19.38020	46.6250	19.458743	663
4746 468	1	18.142500	21.59040	40.8333	10.416557	1252
5146		10.142500	21.39040	40.0333	10.410557	1232
469	1	20.295000	24.39980	50.2917	12.791439	2795
4665 470	1	24.873347	28.69375	50.7917	15.083643	2846
4286		24.073347	20.09373	30.7917	13.003043	2040
471	1	27.230847	30.74625	56.1667	19.083543	1198
5172 472	1	24.941653	29.92435	39 0417	18.333143	989
5702	_					
473	2	18.996653	22.85190	56.9167	11.250104	347
4020 474	1	20.431653	24.65230	61.2500	4.417256	846
5719	_					
475	1	21.593347	25.78875	69.4583	10.041357	1340
5950 476	1	23.370000	27.14605	68.2917	19.000329	2541
4083						
477	3	16.263347	19.47520	83.5417	23.084582	120
907 478	2	13.188347	15.05625	76.6667	20.334232	195
3019						
479	1	16.946653	20.26415	45.4167	16.708661	518

5115 480	1	19.543347	23.51585	42.7917	7.959064	655
5541		19.545547	23.31303	42.7917	7.959004	055
481	2	20.431653	24.17915	75.6667	11.833875	475
4551 482	1	18.757500	22.63185	40.0833	23.291411	1014
5219		18.757500	22.03103	40.0033	23.291411	1014
483	2	15.443347	18.87520	48.9583	8.708325	1120
3100	1	18.791653	22 50605	E0 7002	7 022026	2220
484 4075	1	18./91055	22.50605	58.7083	7.832836	2229
485	2	19.030847	22.88480	57.0000	11.499746	665
4907						
cnt 90 2227 91 2252 92 3249 93 3115 94 1795 95 2808 96 3141 97 1471 98 2455 99 2895 100 3348 101 2034 102 2162 103 3267 104 3126 105 795 106 3744 107 3429 108 3204 109 3944 110 4189 111 1683 112 4036 113 4191 114 4073 115 4400 116 3872 117 4058 118 4595 119 5312 456 6041 457 5936 458 6772 459 6436						

```
460 6457
    6460
461
462
    6857
463
    5169
464 5585
    5918
465
466 4862
467 5409
468 6398
469 7460
470 7132
471 6370
472 6691
473 4367
474 6565
475 7290
476 6624
477 1027
478 3214
479 5633
480 6196
481 5026
482 6233
483 4220
484 6304
485 5572
day df[day df.mnth == 7]
    181
        182 2011-07-01 3 0 7
                                           0
        183 2011-07-02 3 0
184 2011-07-03 3 0
185 2011-07-04 3 0
186 2011-07-05 3 0
182
183
                                    7
                                                   6
                                           0
                                                             0
                                   7
                                           0
                                                   0
                                                             0
184
185
                                    7
                                           1
                                                   1
                                    7
                                           0
                                                   2
                                                             1
        574 2012-07-27 3 1
575 2012-07-28 3 1
576 2012-07-29 3 1
573
574
                                           0
                                                  5
                                    7
                                                             1
                                    7
                                                  6
                                           0
                                                             0
575
                                    7
                                                             0
                                           0
                                                   0
        577 2012-07-30
                          3 1
                                    7
                                           0
                                                   1
                                                             1
576
                          3
                                    7
                                           0
                                                   2
577
        578 2012-07-31
                              1
weathersit temp atemp hum windspeed casual
registered \
            1 29.622500 32.60810 39.6250 6.874736
181
                                                      1246
4116
            1 30.271653 33.36540 44.4583 7.709154
                                                      2204
182
2915
183
            2 29.383347 33.42875 68.2500 15.333486
                                                      2282
2367
```

```
184
                29.793347 33.27085
                                      63.7917
                                                 5.459106
                                                             3065
2978
185
                30.613347
                            34.81690
                                      59.0417
                                                 8.459286
                                                             1031
3634
. . .
                32.048347 36.71085
                                      59.4583
                                                10.250464
                                                             1259
573
5645
574
                30.989153 34.88020
                                      61.3333
                                                10.542450
                                                             2234
4451
575
                29.588347 33.39665
                                      62.3750
                                                11.416532
                                                             2153
4444
576
                29.964153
                           34.24935
                                      66.8750
                                                10.292339
                                                             1040
6065
577
                29.246653 33.14480
                                      70.4167
                                               11.083475
                                                              968
6248
      cnt
181
     5362
182
     5119
183
     4649
184
     6043
185
     4665
573
     6904
574
     6685
575
     6597
576
     7105
577
    7216
[62 rows x 16 columns]
day df.groupby(by="season").cnt.nunique().sort values(ascending=False)
season
3
     188
2
     182
1
     179
     176
Name: cnt, dtype: int64
```

Pertanyaan 1:

Berapa Rata Rata Kecepatan Angin Pada tahun 2011, Berapa Suhu Temperature pada tahun 2011

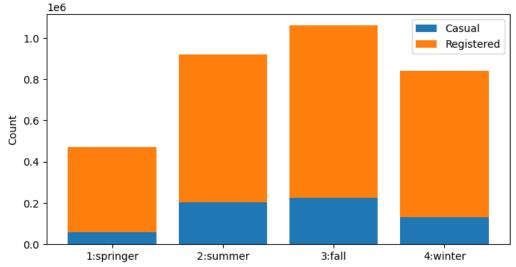
Dari Data diatas menunjukan bahwa pada musim gugur banyak orang bersepeda hal tersebut menunjukan bahwa dengan suhu yang cukup panas dapat membakar banyak kalori sehingga banyak yang bersepeda pada bulan july dan bertepatan pada musim gugur

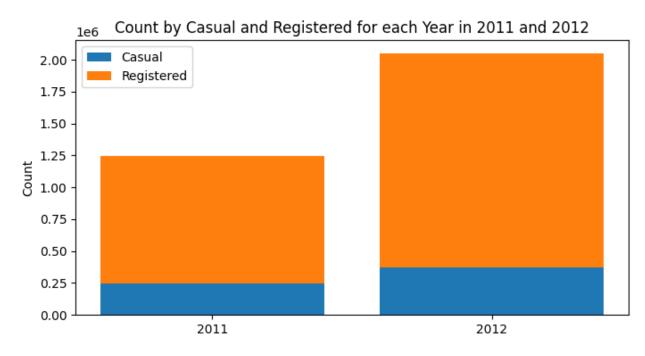
Dan pada data diatas yaitu ketika musim panas dan bertepatan pada bulan april banyak orang juga melakukan bersepeda karena angin dan suhu yang cukup baik untuk bersepeda pada musim tersebut

Pertanyaan 2: Berapa jumlah penyewa sepeda yang telah terdaftar dan yang belum terdaftar di tiap tahun

```
plt.figure(figsize=(8,4))
p1 = plt.bar(day df['season'].unique(),
             day_df.groupby(['season'])['casual'].sum())
p2 = plt.bar(day df['season'].unique(), # the x locations for the
groups
             day df.groupby(['season'])['registered'].sum(), # Count
of Registered per season
             bottom = day df.groupby(['season'])['casual'].sum()) #
Count of casual per season
plt.ylabel('Count')
plt.title("Count by Casual and Registered for each Season in
1:springer', '2:summer', '3:fall', '4:winter Data")
plt.xticks(day df['season'].unique(), ('1:springer', '2:summer',
'3:fall', '4:winter')) # Name of unique values in columns
plt.legend((p1[0], p2[0]), ('Casual', 'Registered')) #setting legends
as per target
plt.show()
```

Count by Casual and Registered for each Season in 1:springer', '2:summer', '3:fall', '4:winter Data





```
day_df.groupby(by="yr").casual.nunique().sort_values(ascending=False)

yr
1     332
0     328
Name: casual, dtype: int64
```

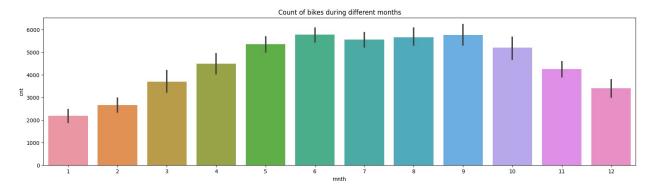
Pertanyaan 2

Berapa jumlah penyewa sepeda yang telah terdaftar dan yang belum terdaftar di tiap tahun

Terlihat pada data diatas menunjukan bahwa jumlah penyewa sepeda terus meningkat setiap tahun pada tahun 2011 jumlah penyewa sepeda terlihat cukup banyak yaitu sekitar 328 jumlah yang menyewa sepeda kemudian pada tahun berikutnya terjadi

peningkatan yaitu sekitar 328 orang yang menyewa sepeda hal ini menunjukan bahwa jumlah peminat bersepada terus meningkat setiap tahunnya.

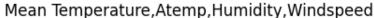
```
fig, ax = plt.subplots(figsize=(20,5))
sns.barplot(data=day_df, x='mnth', y='cnt', ax=ax)
ax.set(title='Count of bikes during different months')
[Text(0.5, 1.0, 'Count of bikes during different months')]
```

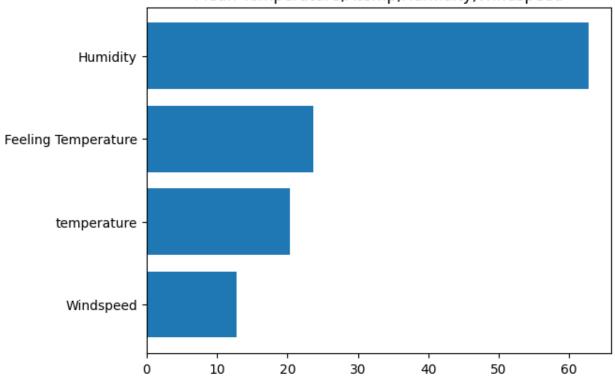


```
rfm_df = day_df.groupby(by=['temp', 'atemp', 'hum',
    'windspeed']).agg({
        "temp": "mean",
        "atemp": "mean",
        "windspeed": "mean"
})
rfm_df.columns = ["temperature", "Feeling Temperature",
    "Humidity", "Windspeed"]

rfm_df = rfm_df.mean().sort_values(ascending=True)

plt.barh(rfm_df.index, rfm_df.values)
plt.title(label="Mean Temperature,Atemp,Humidity,Windspeed")
plt.xlabel("")
plt.ylabel("")
plt.show()
```





Conclusion

- Conclution pertanyaan 1: Berapa Kecepatan Rata rata angin pada tahun 2011
 Pada data yang telah diatas bisa kita liat bahwa rata rata kecepatan angin tiap bulan berbeda-beda Tingkat paling tinggi yaitu pada bulan April yaitu ketika musim panas. Kemudian kecepatan angin kembali menurun dibulan berikutnya seiring bergantinya musim dan pada musim panas ini banyak orang yang menyewa sepeda pada bulan dan musim tersebut.
- conclution pertanyaan 2 : Berapa jumlah penyewa sepeda yang telah terdaftar dan yang belum terdaftar di tiap tahun

Setiap Tahunnya penyewa sepeda memiliki peningkatan.Pada tahun 2011 banyak penyewa sepeda yang sudah menjadi member dan ada beberapa penyewa sepeda yang belum menjadi member.Kemudian ditahun selanjutnya pada tahun 2012 jumlah penyewa sepeda meningkat dan jumlah penyewa sudah menjadi member pun memiliki peningkatan