

Tugas Materi 12

Zul Fauzi Oktaviansyah

2110181056

3 – D4 IT - B

Code 1

```
In [2]: dataset = pd.read_csv('transaction.csv')
dataset
```

Out[2]:

	InvoiceNo	StockCode	Qty	InvoiceDate	CustomerID	Country
0	537626	22725	830	12/7/2010 14:57	12347	Iceland
1	537626	22729	948	12/7/2010 14:57	12347	Iceland
2	537626	22195	695	12/7/2010 14:57	12347	Iceland
3	542237	22725	636	1/26/2011 14:30	12347	Iceland
4	542237	22729	536	1/26/2011 14:30	12347	Iceland
...
10541	543911	21700	455	2/14/2011 12:46	17829	United Arab Emirates
10542	543911	22111	578	2/14/2011 12:46	17829	United Arab Emirates
10543	543911	22112	163	2/14/2011 12:46	17829	United Arab Emirates
10544	564428	23296	545	8/25/2011 11:27	17844	Canada
10545	564428	23294	643	8/25/2011 11:27	17844	Canada

10546 rows × 6 columns

Membaca data csv transaction

Code 2

```
In [3]: data = dataset[dataset['Country'] == 'Germany']  
data
```

Out[3]:

	InvoiceNo	StockCode	Qty	InvoiceDate	CustomerID	Country
1185	554985	21746	628	5/29/2011 12:26	12426	Germany
1186	554985	21770	981	5/29/2011 12:26	12426	Germany
1187	554985	22329	212	5/29/2011 12:26	12426	Germany
1188	554985	22976	910	5/29/2011 12:26	12426	Germany
1189	554985	22845	668	5/29/2011 12:26	12426	Germany
...
8339	565430	22725	562	9/4/2011 13:07	14335	Germany
8340	565430	22729	692	9/4/2011 13:07	14335	Germany
8341	565430	22302	400	9/4/2011 13:07	14335	Germany
8342	578273	22725	769	11/23/2011 13:41	14335	Germany
8343	578273	23309	842	11/23/2011 13:41	14335	Germany

2269 rows × 6 columns

Menampilkan data country germany

Code 3

```
In [4]: data['Year'] = pd.DatetimeIndex(data['InvoiceDate']).year  
data['Month'] = pd.DatetimeIndex(data['InvoiceDate']).month
```

Menambah kolom year dan month

Code 4

```
In [5]: data
```

Out[5]:

	InvoiceNo	StockCode	Qty	InvoiceDate	CustomerID	Country	Year	Month
1185	554985	21746	628	5/29/2011 12:26	12426	Germany	2011	5
1186	554985	21770	981	5/29/2011 12:26	12426	Germany	2011	5
1187	554985	22329	212	5/29/2011 12:26	12426	Germany	2011	5
1188	554985	22976	910	5/29/2011 12:26	12426	Germany	2011	5
output; double click to hide		2845	668	5/29/2011 12:26	12426	Germany	2011	5
...
8339	565430	22725	562	9/4/2011 13:07	14335	Germany	2011	9
8340	565430	22729	692	9/4/2011 13:07	14335	Germany	2011	9
8341	565430	22302	400	9/4/2011 13:07	14335	Germany	2011	9
8342	578273	22725	769	11/23/2011 13:41	14335	Germany	2011	11
8343	578273	23309	842	11/23/2011 13:41	14335	Germany	2011	11

2269 rows × 8 columns

Menampilkan fitur month dan year

Code 5

```
In [6]: data = data[data['Year'] == 2011]
data
```

Out[6]:

	InvoiceNo	StockCode	Qty	InvoiceDate	CustomerID	Country	Year	Month
1185	554985	21746	628	5/29/2011 12:26	12426	Germany	2011	5
1186	554985	21770	981	5/29/2011 12:26	12426	Germany	2011	5
1187	554985	22329	212	5/29/2011 12:26	12426	Germany	2011	5
1188	554985	22976	910	5/29/2011 12:26	12426	Germany	2011	5
1189	554985	22845	668	5/29/2011 12:26	12426	Germany	2011	5
...
8339	565430	22725	562	9/4/2011 13:07	14335	Germany	2011	9
8340	565430	22729	692	9/4/2011 13:07	14335	Germany	2011	9
8341	565430	22302	400	9/4/2011 13:07	14335	Germany	2011	9
8342	578273	22725	769	11/23/2011 13:41	14335	Germany	2011	11
8343	578273	23309	842	11/23/2011 13:41	14335	Germany	2011	11

2148 rows × 8 columns

Mengambil data year = 2011

Code 6

```
In [7]: data = pd.DataFrame(data[['Month', 'Qty']])  
data
```

Out[7]:

	Month	Qty
1185	5	628
1186	5	981
1187	5	212
1188	5	910
1189	5	668
...
8339	9	562
8340	9	692
8341	9	400
8342	11	769
8343	11	842

2148 rows × 2 columns

Mengambil data month dan qty

Code 7

```
In [8]: totalQty = data.groupby('Month')['Qty'].sum()  
totalQty
```

```
Out[8]: Month  
1      101266  
2       37800  
3       72084  
4       60993  
5      103749  
6       79711  
7       96423  
8      120908  
9      104487  
10     159490  
11     143409  
12       33126  
Name: Qty, dtype: int64
```

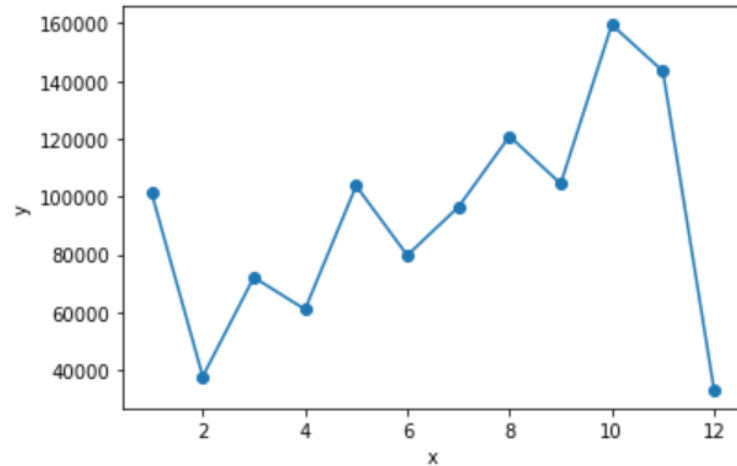
Mengambil total qty per month

Code 8

```
In [10]: x=totalQty.index  
y=totalQty.values
```

```
In [11]: plt.scatter(x, y)  
plt.plot(x, y)  
plt.xlabel('x')  
plt.ylabel('y')
```

```
Out[11]: Text(0, 0.5, 'y')
```



Menampilkan grafik total qty per month

Code 9

```
In [12]: linreg=LinearRegression()  
x=np.array(x).reshape(-1,1)  
linreg.fit(x, y)
```

```
Out[12]: LinearRegression()
```

```
In [13]: next_x=13  
next_x=np.array(next_x).reshape(-1,1)  
pred_x=linreg.predict(next_x)
```

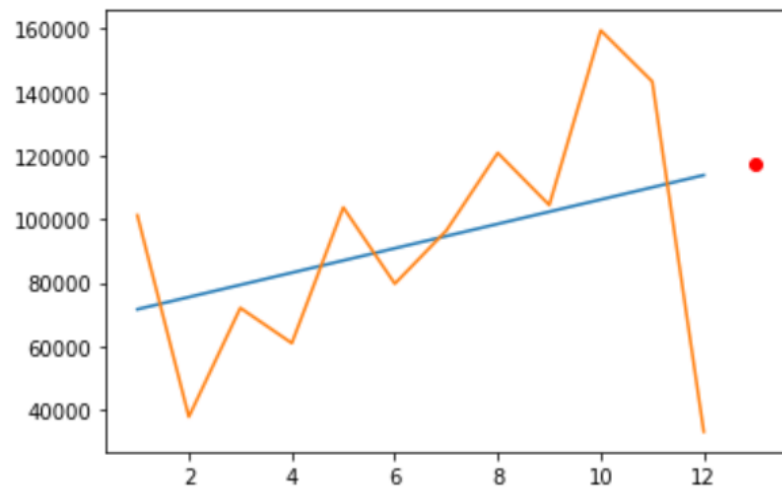
```
In [14]: print('\nPrediksi x \n', pred_x.item())
```

```
Prediksi x  
117751.75757575757
```

Memprediksi total qty pada bulan januari 2012

Code 10

```
In [15]: plt.scatter(next_x, pred_x, c='red')  
pred_y=linreg.predict(x)  
plt.plot(x, pred_y)  
plt.plot(x, y)  
plt.show()
```



Menampilkan hasil linear regresi

Code 11

```
In [16]: MSE=mean_squared_error(y,pred_y)
print('\nMSE = ', MSE)
```

```
MSE = 1185741672.860917
```

Menampilkan hasil error