

Tugas Analisis Visualisasi Data

Metode Klasifikasi dan Clustering



Oleh :

Ferry Triwantono – 082111633094

FAKULTAS SAINS DAN TEKNOLOGI
PROGRAM STUDI SISTEM INFORMASI
UNIVERSITAS AIRLANGGA

2022

Metode Klasifikasi dan Clustering

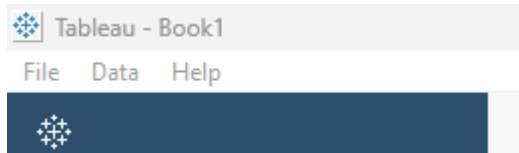
1. Pelajari materi pada Chapter 7 tentang membuat perhitungan di Tableau dengan skrip Python untuk perhitungan *coefficient correlation* antara Profit and Sales (Alexander Loth, 2019, Visual Analytics with Tableau®, John Wiley & Sons).

- Install TabPy

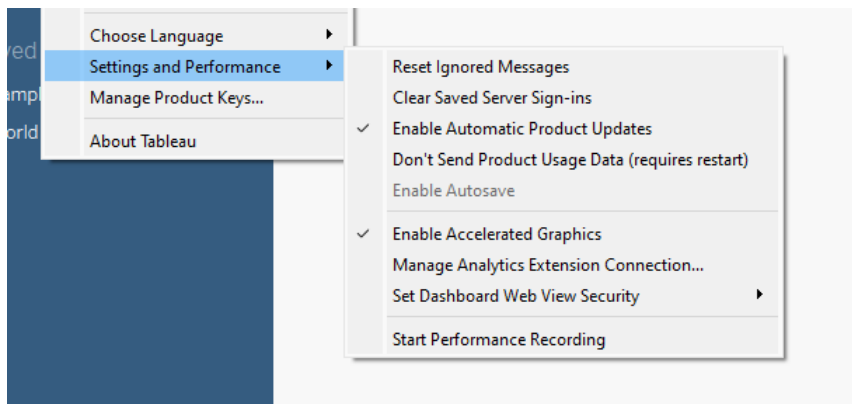
```
Using cached nltk-3.7-py3-none-any.whl (1.5 MB)
collecting hypothesis
  Downloading hypothesis-6.56.4-py3-none-any.whl (395 kB)
    395.3/395.3 kB 26.4 kB/s eta 0:00:00
collecting genson
  Downloading genson-1.2.2.tar.gz (34 kB)
  Preparing metadata (setup.py) ... done
collecting coveralls
  Using cached coveralls-3.3.1-py2.py3-none-any.whl (14 kB)
collecting pyopenssl
  Downloading pyOpenSSL-22.1.0-py3-none-any.whl (57 kB)
    57.0/57.0 kB 35.2 kB/s eta 0:00:00
collecting sklearn
  Using cached sklearn-0.0.tar.gz (1.1 kB)
  Preparing metadata (setup.py) ... done
collecting textblob
  Downloading textblob-0.17.1-py2.py3-none-any.whl (636 kB)
    636.0/636.0 kB 26.4 kB/s eta 0:00:00
Requirement already satisfied: pandas in c:\users\fikri\appdata\local\programs\python\python38-32\lib\site-packages (from tabpy) (1.
collecting mock
  Using cached mock-4.0.3-py3-none-any.whl (28 kB)
collecting docopt
  Downloading docopt-0.6.2.tar.gz (25 kB)
  Preparing metadata (setup.py) ... done
Requirement already satisfied: tornado in c:\users\fikri\appdata\local\programs\python\python38-32\lib\site-packages (from tabpy) (6
collecting urllib3
  Downloading urllib3-1.26.12-py2.py3-none-any.whl (140 kB)
    140.4/140.4 kB 24.6 kB/s eta 0:00:00
collecting cloudpickle
  Downloading cloudpickle-2.2.0-py3-none-any.whl (25 kB)
collecting pytest-cov
  Downloading pytest_cov-4.0.0-py3-none-any.whl (21 kB)
collecting future
  Downloading future-0.18.2.tar.gz (829 kB)
    829.2/829.2 kB 25.3 kB/s eta 0:00:00
  Preparing metadata (setup.py) ...
```

- Integrasi Tableau dengan TabPy

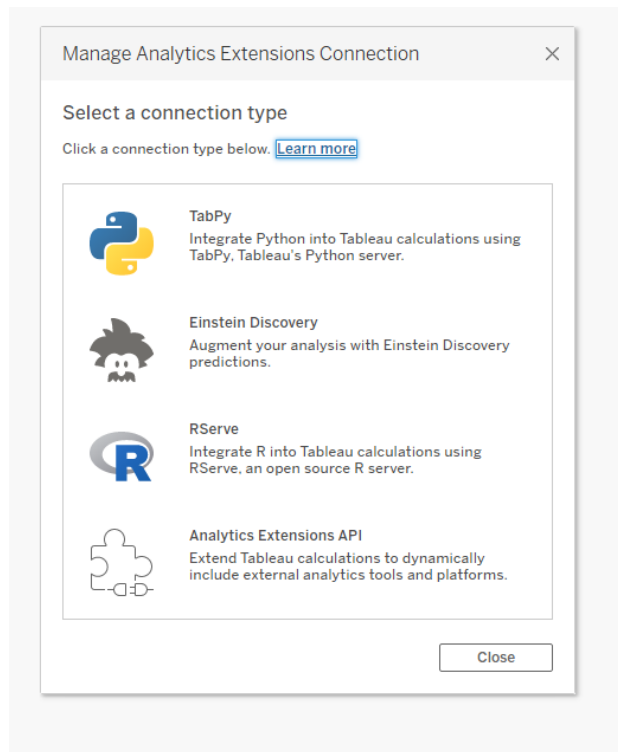
- a. Buka tab menu ‘Help’ di kiri atas



- b. Pilih ‘Settings and Performance’ lalu pilih ‘Manage External Service Connection’



c. Pilih Ektensi TabPy



d. Nyalakan TabPy melalui cmd dengan command tabpy

```
C:\Windows\System32>tabpy
Error processing line 1 of C:\Python311\Lib\site-packages\_editable___django-5.0.dev20230912095101.pth:

Traceback (most recent call last):
  File "<frozen site>", line 186, in addpackage
  File "<strings>", line 1, in <module>
ModuleNotFoundError: No module named '_editable___django_5_0_dev20230912095101_finder'


Remainder of file ignored
2023-11-04,22:34:13 [INFO] (app.py:app:308): Parsing config file C:\Python311\Lib\site-packages\tabpy\tabpy_server\app\..\common\default.conf
2023-11-04,22:34:14 [INFO] (app.py:app:363): Setting max request size to 104857600 bytes
2023-11-04,22:34:14 [INFO] (app.py:app:527): Loading state from state file C:\Python311\Lib\site-packages\tabpy\tabpy_server\state.ini

WARNING: This TabPy server is not currently configured for username/password authentication. This means that, because the TABPY_EVALUATE_ENABLE feature is enabled, there is the potential that unauthenticated individuals may be able to remotely execute code on this machine. We strongly advise against proceeding without authentication as it poses a significant security risk.

Do you wish to proceed without authentication? (y/N): y
2023-11-04,22:34:15 [INFO] (app.py:app:492): Password file is not specified: Authentication is not enabled
2023-11-04,22:34:15 [INFO] (app.py:app:411): Call context logging is disabled
2023-11-04,22:34:15 [INFO] (app.py:app:181): Initializing TabPy...
2023-11-04,22:34:15 [INFO] (callbacks.py:callbacks:43): Initializing TabPy Server...
2023-11-04,22:34:15 [INFO] (app.py:app:185): Done initializing TabPy.
2023-11-04,22:34:15 [INFO] (callbacks.py:callbacks:64): Initializing models...
2023-11-04,22:34:15 [INFO] (app.py:app:151): Web service listening on port 9004
```

e. Masukkan hasil port pada konfigurasi

Manage Analytics Extensions Connection



New TabPy Connection

☐ Require SSL

Hostname

localhost

Port

9004

☐ Sign in with username and password

Username

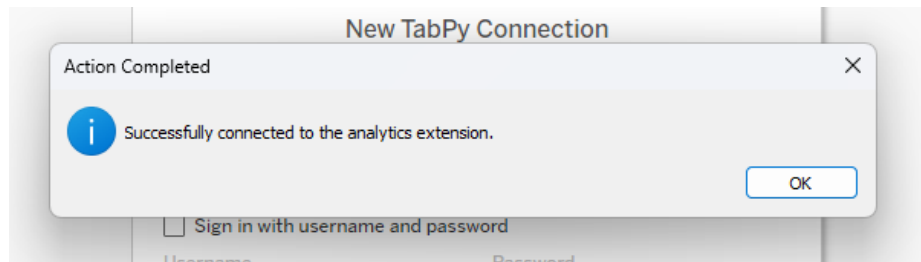
Password

Test Connection

Back

Save

f. Akan muncul dialog berhasil apabila ‘test connection’ berhasil



- Analisis dan Visualisasi Data Profit and Sales

a. Pilih Dataset superstore

File Data Window Help

Connections

Sample - Superstore

Recent Datas

Sheets

Use Data Interpreter

Data Interpreter might be able to clean your Microsoft Excel workbook

Orders

People

Returns

Orders

People

Returns

New Union

Go to Worksheet

Sample - Superstore

Connection

Live Extract

Filters

Add

Orders

Sort Fields

Data source order

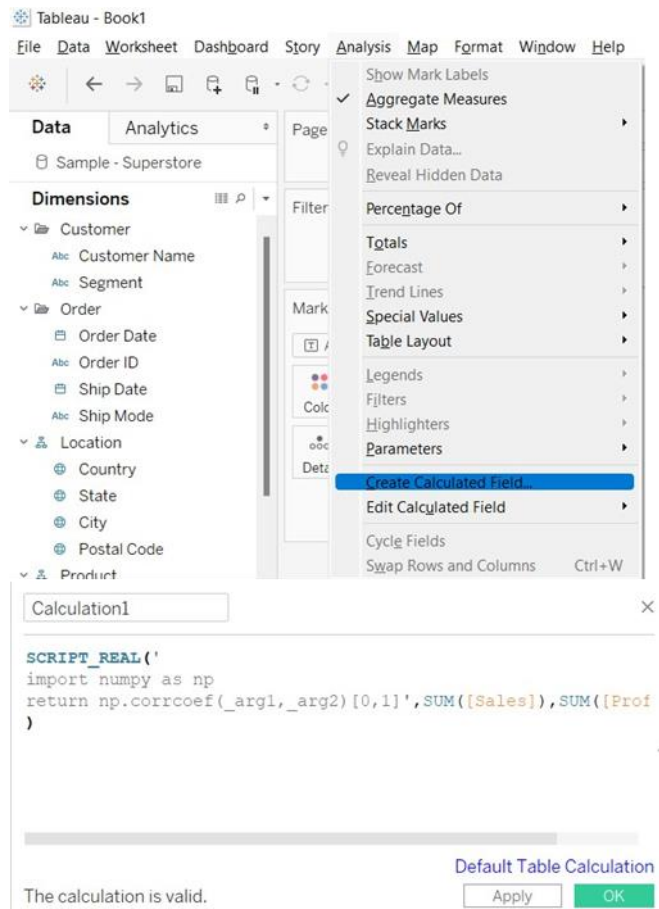
Show aliases

Show hidden fields

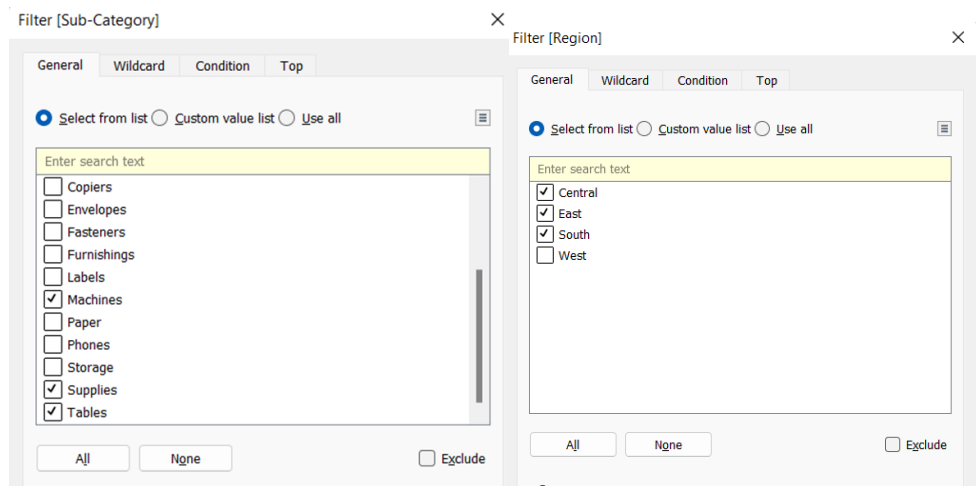
1,000 rows

Order ID	Order Date	Ship Date	Ship Mode	Customer Name	Segment	Country	City	State	Postal Code	Region	Category
CA-2018-132154	06/11/2018	11/11/2018	Second Class	Claira Gutie	Consumer	United States	Henderson	Kentucky	42420	South	Furniture
CA-2018-138898	12/06/2018	16/06/2018	Second Class	Darinn Van Huff	Corporate	United States	Los Angeles	California	90036	West	Office Supplies
US-2017-108966	11/10/2017	18/10/2017	Standard Class	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	Furniture
US-2017-108966	11/10/2017	18/10/2017	Standard Class	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	Office Supplies
CA-2016-113612	06/06/2016	14/06/2016	Standard Class	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	Furniture
CA-2016-113612	06/06/2016	14/06/2016	Standard Class	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	Office Supplies
CA-2016-113612	06/06/2016	14/06/2016	Standard Class	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	Technology
CA-2016-113612	06/06/2016	14/06/2016	Standard Class	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	Office Supplies
CA-2016-113612	06/06/2016	14/06/2016	Standard Class	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	Office Supplies

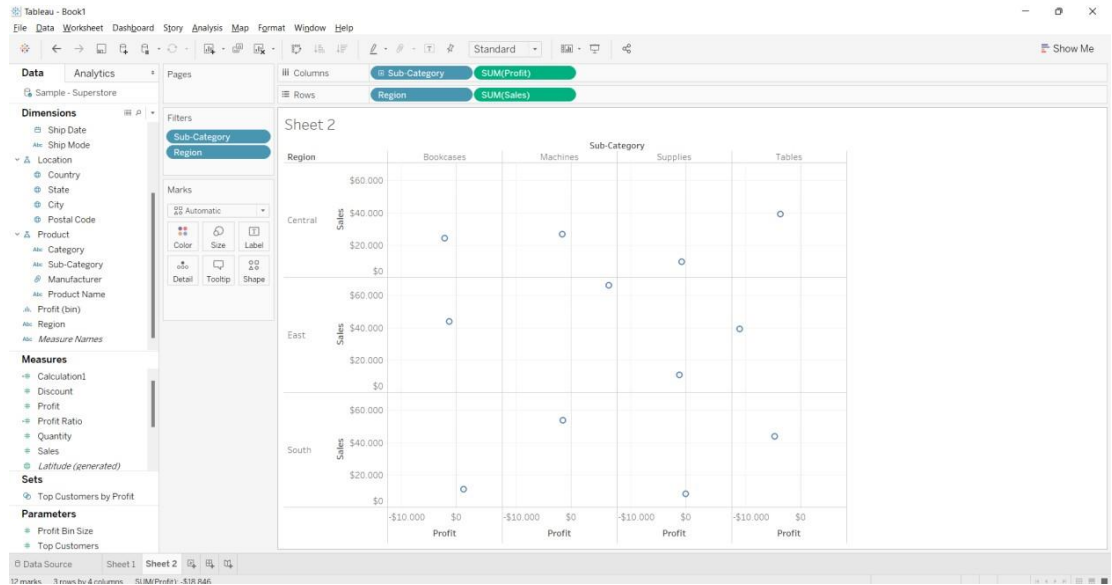
b. Buka dan buat Calculated Field di menu Analysis diatas



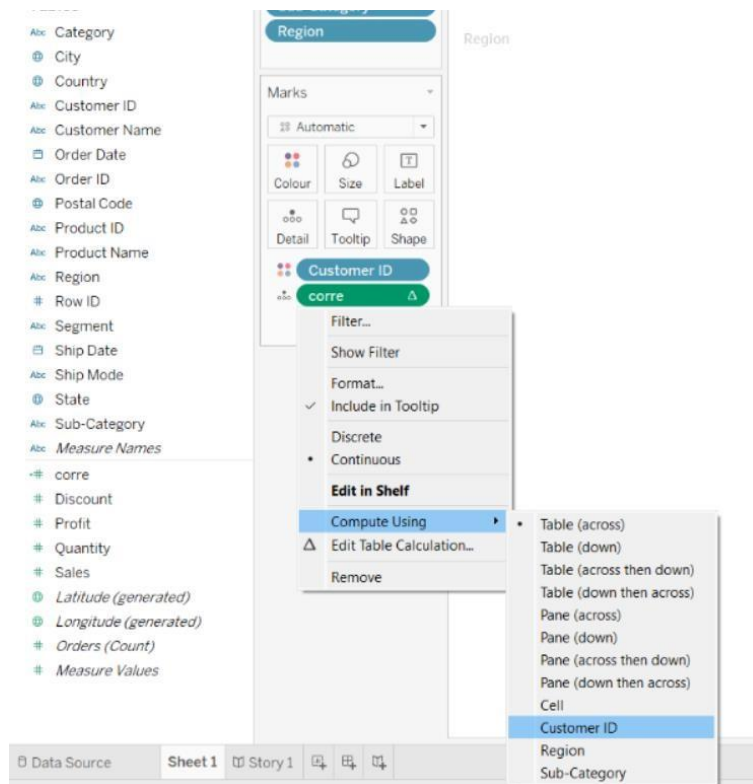
c. Ambil variabel yang dibutuhkan sebagai filter



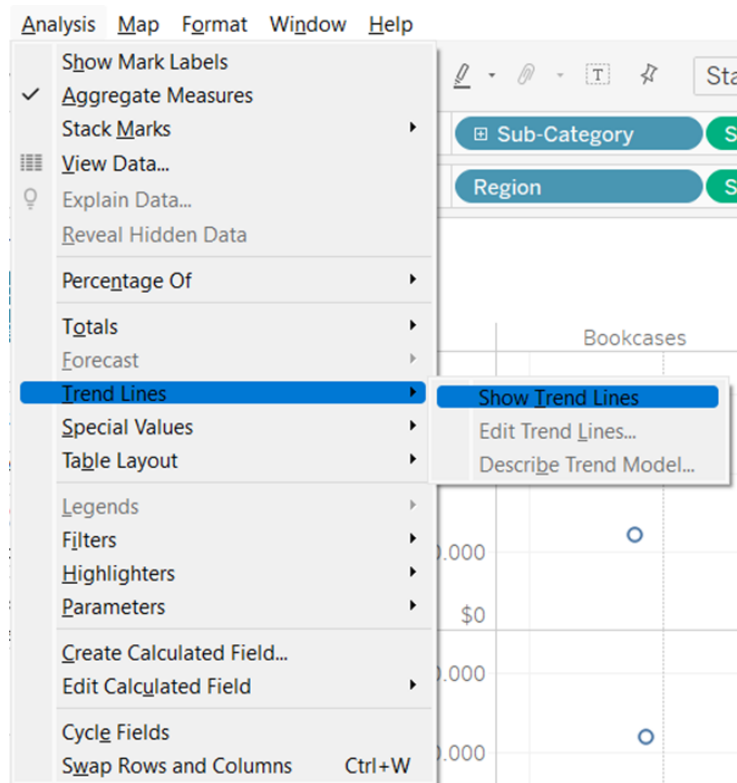
d. Hasil Filter



e. Menambahkan customer ID dan Correlation Coefficient pada marks



f. Menambahkan Trend-lines di menu Analysis



g. Output

