In [3]:
```
!pip install pycaret
```

```
WARNING: Keyring is skipped due to an exception: Failed to unlock the collect
ion!
Collecting pycaret
  Using cached pycaret-2.3.10-py3-none-any.whl (320 kB)
Collecting wordcloud
  Downloading wordcloud-1.8.2.2-cp38-cp38-manylinux_2_17_x86_64.manylinux2014
_x86_64.whl (458 kB)
     |████████████████████████████████| 458 kB 1.3 MB/s eta 0:00:01
Requirement already satisfied: numba<0.55 in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (0.51.2)
Requirement already satisfied: scikit-learn==0.23.2 in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from pycare
t) (0.23.2)
Collecting spacy<2.4.0
  Downloading spacy-2.3.8-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_6
4.whl (5.0 MB)
     |████████████████████████████████| 5.0 MB 1.2 MB/s eta 0:00:01
Collecting pandas-profiling>=2.8.0
  Using cached pandas_profiling-3.4.0-py2.py3-none-any.whl (315 kB)
Requirement already satisfied: pyyaml<6.0.0 in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (5.3.1)
Collecting textblob
  Using cached textblob-0.17.1-py2.py3-none-any.whl (636 kB)
Requirement already satisfied: matplotlib in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (3.3.2)
Collecting pyLDAvis
  Using cached pyLDAvis-3.3.1.tar.gz (1.7 MB)
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Installing backend dependencies ... done
    Preparing wheel metadata ... done
Requirement already satisfied: pandas in /home/damar.pramuditya/.pyenv/versio
ns/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (1.1.3)
Collecting umap-learn
  Using cached umap-learn-0.5.3.tar.gz (88 kB)
Collecting yellowbrick>=1.0.1
  Using cached yellowbrick-1.5-py3-none-any.whl (282 kB)
Requirement already satisfied: seaborn in /home/damar.pramuditya/.pyenv/versi
ons/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (0.11.0)
Collecting gensim<4.0.0
  Downloading gensim-3.8.3-cp38-cp38-manylinux1_x86_64.whl (24.2 MB)
     |████████████████████████████████| 24.2 MB 1.2 MB/s eta 0:00:01
Collecting cufflinks>=0.17.0
  Using cached cufflinks-0.17.3.tar.gz (81 kB)
Collecting mlxtend>=0.17.0
  Using cached mlxtend-0.21.0-py2.py3-none-any.whl (1.3 MB)
Collecting mlflow
  Using cached mlflow-1.30.0-py3-none-any.whl (17.0 MB)
Requirement already satisfied: joblib in /home/damar.pramuditya/.pyenv/versio
ns/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (0.17.0)
Collecting plotly>=4.4.1
  Downloading plotly-5.11.0-py2.py3-none-any.whl (15.3 MB)
     |████████████████████████████████| 15.3 MB 1.1 MB/s eta 0:00:01     |███
████████████| 8.2 MB 935 kB/s eta 0:00:08
Requirement already satisfied: scipy<=1.5.4 in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (1.5.2)
Requirement already satisfied: ipywidgets in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (7.5.1)
Collecting pyod
  Using cached pyod-1.0.6.tar.gz (141 kB)
Requirement already satisfied: IPython in /home/damar.pramuditya/.pyenv/versi
ons/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (7.19.0)
Requirement already satisfied: nltk in /home/damar.pramuditya/.pyenv/version
s/anaconda3-2020.11/lib/python3.8/site-packages (from pycaret) (3.5)
Collecting imbalanced-learn==0.7.0
  Using cached imbalanced_learn-0.7.0-py3-none-any.whl (167 kB)
```

```
Collecting Boruta
  Using cached Boruta-0.3-py3-none-any.whl (56 kB)
Collecting kmodes>=0.10.1
  Using cached kmodes-0.12.2-py2.py3-none-any.whl (20 kB)
Collecting scikit-plot
  Using cached scikit_plot-0.3.7-py3-none-any.whl (33 kB)
Collecting lightgbm>=2.3.1
  Using cached lightgbm-3.3.3-py3-none-manylinux1_x86_64.whl (2.0 MB)
Requirement already satisfied: pillow in /home/damar.pramuditya/.pyenv/versio
ns/anaconda3-2020.11/lib/python3.8/site-packages (from wordcloud->pycaret)
 (8.0.1)
Requirement already satisfied: numpy>=1.6.1 in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from wordcloud->pycar
et) (1.19.2)
Requirement already satisfied: llvmlite<0.35,>=0.34.0.dev0 in /home/damar.pra
muditya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from n
umba<0.55->pycaret) (0.34.0)
Requirement already satisfied: setuptools in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from numba<0.55->pycare
t) (50.3.1.post20201107)
Requirement already satisfied: threadpoolctl>=2.0.0 in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from scikit-
learn==0.23.2->pycaret) (2.1.0)
Collecting blis<0.8.0,>=0.4.0
  Downloading blis-0.7.9-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_6
4.whl (10.2 MB)
     |████████████████████████████████| 10.2 MB 1.3 MB/s eta 0:00:01
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from spacy<
2.4.0->pycaret) (4.50.2)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /home/damar.pramudi
tya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from spacy
<2.4.0->pycaret) (2.24.0)
Collecting thinc<7.5.0,>=7.4.1
  Downloading thinc-7.4.6-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_6
4.whl (1.1 MB)
     |████████████████████████████████| 1.1 MB 775 kB/s eta 0:00:01
Collecting cymem<2.1.0,>=2.0.2
  Downloading cymem-2.0.7-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_6
4.whl (36 kB)
Collecting catalogue<1.1.0,>=0.0.7
  Using cached catalogue-1.0.2-py2.py3-none-any.whl (16 kB)
Collecting preshed<3.1.0,>=3.0.2
  Downloading preshed-3.0.8-cp38-cp38-manylinux_2_5_x86_64.manylinux1_x86_64.
manylinux_2_17_x86_64.manylinux2014_x86_64.whl (130 kB)
     |████████████████████████████████| 130 kB 1.2 MB/s eta 0:00:01
Collecting srsly<1.1.0,>=1.0.2
  Downloading srsly-1.0.6-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_6
4.whl (211 kB)
     |████████████████████████████████| 211 kB 1.4 MB/s eta 0:00:01
Collecting plac<1.2.0,>=0.9.6
  Using cached plac-1.1.3-py2.py3-none-any.whl (20 kB)
Collecting wasabi<1.1.0,>=0.4.0
  Using cached wasabi-0.10.1-py3-none-any.whl (26 kB)
Collecting murmurhash<1.1.0,>=0.28.0
  Downloading murmurhash-1.0.9-cp38-cp38-manylinux_2_5_x86_64.manylinux1_x86_
64.manylinux_2_17_x86_64.manylinux2014_x86_64.whl (21 kB)
Collecting visions[type_image_path]==0.7.5
  Using cached visions-0.7.5-py3-none-any.whl (102 kB)
Collecting htmlmin==0.1.12
  Using cached htmlmin-0.1.12.tar.gz (19 kB)
Collecting phik<0.13,>=0.11.1
  Downloading phik-0.12.2-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_6
4.whl (696 kB)
     |████████████████████████████████| 696 kB 1.4 MB/s eta 0:00:01
Collecting multimethod<1.10,>=1.4
  Using cached multimethod-1.9-py3-none-any.whl (10 kB)
Collecting pydantic<1.11,>=1.8.1
  Downloading pydantic-1.10.2-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x
```

```
86_64.whl (13.6 MB)
     |████████████████████████████████| 13.6 MB 291 kB/s eta 0:00:01
Collecting missingno<0.6,>=0.4.2
  Using cached missingno-0.5.1-py3-none-any.whl (8.7 kB)
Collecting statsmodels<0.14,>=0.13.2
  Downloading statsmodels-0.13.2-cp38-cp38-manylinux_2_17_x86_64.manylinux201
4_x86_64.whl (9.9 MB)
     |████████████████████████████████| 9.9 MB 406 kB/s eta 0:00:01
Requirement already satisfied: jinja2<3.2,>=2.11.1 in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from pandas-
profiling>=2.8.0->pycaret) (2.11.2)
Requirement already satisfied: kiwisolver>=1.0.1 in /home/damar.pramuditya/.p
yenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from matplotlib-
>pycaret) (1.3.0)
Requirement already satisfied: certifi>=2020.06.20 in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from matplot
lib->pycaret) (2020.6.20)
Requirement already satisfied: cycler>=0.10 in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from matplotlib->pyca
ret) (0.10.0)
Requirement already satisfied: python-dateutil>=2.1 in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from matplot
lib->pycaret) (2.8.1)
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.3 in /h
ome/damar.pramuditya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-pac
kages (from matplotlib->pycaret) (2.4.7)
Requirement already satisfied: numexpr in /home/damar.pramuditya/.pyenv/versi
ons/anaconda3-2020.11/lib/python3.8/site-packages (from pyLDAvis->pycaret)
 (2.7.1)
Requirement already satisfied: future in /home/damar.pramuditya/.pyenv/versio
ns/anaconda3-2020.11/lib/python3.8/site-packages (from pyLDAvis->pycaret) (0.
18.2)
Collecting sklearn
  Downloading sklearn-0.0.tar.gz (1.1 kB)
Collecting funcy
  Using cached funcy-1.17-py2.py3-none-any.whl (33 kB)
Requirement already satisfied: pytz>=2017.2 in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from pandas->pycaret)
(2020.1)
Collecting pynndescent>=0.5
  Using cached pynndescent-0.5.7.tar.gz (1.1 MB)
Collecting smart-open>=1.8.1
  Downloading smart_open-6.2.0-py3-none-any.whl (58 kB)
     |████████████████████████████████| 58 kB 99 kB/s eta 0:00:01
Requirement already satisfied: six>=1.5.0 in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from gensim<4.0.0->pyca
ret) (1.15.0)
Collecting colorlover>=0.2.1
  Using cached colorlover-0.3.0-py3-none-any.whl (8.9 kB)
Collecting alembic<2
  Using cached alembic-1.8.1-py3-none-any.whl (209 kB)
Collecting sqlalchemy<2,>=1.4.0
  Downloading SQLAlchemy-1.4.42-cp38-cp38-manylinux_2_5_x86_64.manylinux1_x86
_64.manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.6 MB)
     |████████████████████████████████| 1.6 MB 579 kB/s eta 0:00:01
Requirement already satisfied: cloudpickle<3 in /home/damar.pramuditya/.pyen
v/versions/anaconda3-2020.11/lib/python3.8/site-packages (from mlflow->pycare
t) (1.6.0)
Collecting protobuf<5,>=3.12.0
  Downloading protobuf-4.21.9-cp37-abi3-manylinux2014_x86_64.whl (408 kB)
     |████████████████████████████████| 408 kB 1.3 MB/s eta 0:00:01
Collecting databricks-cli<1,>=0.8.7
  Using cached databricks-cli-0.17.3.tar.gz (77 kB)
Requirement already satisfied: click<9,>=7.0 in /home/damar.pramuditya/.pyen
v/versions/anaconda3-2020.11/lib/python3.8/site-packages (from mlflow->pycare
t) (7.1.2)
Collecting sqlparse<1,>=0.4.0
  Using cached sqlparse-0.4.3-py3-none-any.whl (42 kB)
Collecting prometheus-flask-exporter<1
```

```
          Using cached prometheus_flask_exporter-0.20.3-py3-none-any.whl (18 kB)
Requirement already satisfied: packaging<22 in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from mlflow->pycaret)
(20.4)
Requirement already satisfied: entrypoints<1 in /home/damar.pramuditya/.pyen
v/versions/anaconda3-2020.11/lib/python3.8/site-packages (from mlflow->pycare
t) (0.3)
Collecting querystring-parser<2
  Using cached querystring_parser-1.2.4-py2.py3-none-any.whl (7.9 kB)
Collecting gitpython<4,>=2.1.0
  Using cached GitPython-3.1.29-py3-none-any.whl (182 kB)
Collecting docker<7,>=4.0.0
  Using cached docker-6.0.0-py3-none-any.whl (147 kB)
Collecting gunicorn<21; platform_system != "Windows"
  Using cached gunicorn-20.1.0-py3-none-any.whl (79 kB)
Collecting importlib-metadata!=4.7.0,<6,>=3.7.0
  Downloading importlib_metadata-5.0.0-py3-none-any.whl (21 kB)
Requirement already satisfied: Flask<3 in /home/damar.pramuditya/.pyenv/versi
ons/anaconda3-2020.11/lib/python3.8/site-packages (from mlflow->pycaret) (1.
1.2)
Collecting tenacity>=6.2.0
  Downloading tenacity-8.1.0-py3-none-any.whl (23 kB)
Requirement already satisfied: nbformat>=4.2.0 in /home/damar.pramuditya/.pye
nv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from ipywidgets->p
ycaret) (5.0.8)
Requirement already satisfied: ipykernel>=4.5.1 in /home/damar.pramuditya/.py
env/versions/anaconda3-2020.11/lib/python3.8/site-packages (from ipywidgets->
pycaret) (5.3.4)
Requirement already satisfied: traitlets>=4.3.1 in /home/damar.pramuditya/.py
env/versions/anaconda3-2020.11/lib/python3.8/site-packages (from ipywidgets->
pycaret) (5.0.5)
Requirement already satisfied: widgetsnbextension~=3.5.0 in /home/damar.pramu
ditya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from ipy
widgets->pycaret) (3.5.1)
Requirement already satisfied: decorator in /home/damar.pramuditya/.pyenv/ver
sions/anaconda3-2020.11/lib/python3.8/site-packages (from IPython->pycaret)
 (4.4.2)
Requirement already satisfied: pexpect>4.3; sys_platform != "win32" in /home/
damar.pramuditya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-package
s (from IPython->pycaret) (4.8.0)
Requirement already satisfied: jedi>=0.10 in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from IPython->pycaret)
 (0.17.1)
Requirement already satisfied: pygments in /home/damar.pramuditya/.pyenv/vers
ions/anaconda3-2020.11/lib/python3.8/site-packages (from IPython->pycaret)
 (2.7.2)
Requirement already satisfied: prompt-toolkit!=3.0.0,!=3.0.1,<3.1.0,>=2.0.0 i
n /home/damar.pramuditya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site
-packages (from IPython->pycaret) (3.0.8)
Requirement already satisfied: pickleshare in /home/damar.pramuditya/.pyenv/v
ersions/anaconda3-2020.11/lib/python3.8/site-packages (from IPython->pycaret)
 (0.7.5)
Requirement already satisfied: backcall in /home/damar.pramuditya/.pyenv/vers
ions/anaconda3-2020.11/lib/python3.8/site-packages (from IPython->pycaret)
 (0.2.0)
Requirement already satisfied: regex in /home/damar.pramuditya/.pyenv/version
s/anaconda3-2020.11/lib/python3.8/site-packages (from nltk->pycaret) (2020.1
0.15)
Requirement already satisfied: wheel in /home/damar.pramuditya/.pyenv/version
s/anaconda3-2020.11/lib/python3.8/site-packages (from lightgbm>=2.3.1->pycare
t) (0.35.1)
Requirement already satisfied: idna<3,>=2.5 in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from requests<3.0.0,>
=2.13.0->spacy<2.4.0->pycaret) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /ho
me/damar.pramuditya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-pack
ages (from requests<3.0.0,>=2.13.0->spacy<2.4.0->pycaret) (1.25.11)
Requirement already satisfied: chardet<4,>=3.0.2 in /home/damar.pramuditya/.p
yenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from requests<3.
```

```
0.0,>=2.13.0->spacy<2.4.0->pycaret) (3.0.4)
Requirement already satisfied: networkx>=2.4 in /home/damar.pramuditya/.pyen
v/versions/anaconda3-2020.11/lib/python3.8/site-packages (from visions[type_i
mage_path]==0.7.5->pandas-profiling>=2.8.0->pycaret) (2.5)
Collecting tangled-up-in-unicode>=0.0.4
  Using cached tangled_up_in_unicode-0.2.0-py3-none-any.whl (4.7 MB)
Requirement already satisfied: attrs>=19.3.0 in /home/damar.pramuditya/.pyen
v/versions/anaconda3-2020.11/lib/python3.8/site-packages (from visions[type_i
mage_path]==0.7.5->pandas-profiling>=2.8.0->pycaret) (20.3.0)
Collecting imagehash; extra == "type_image_path"
  Using cached ImageHash-4.3.1-py2.py3-none-any.whl (296 kB)
Collecting typing-extensions>=4.1.0
  Downloading typing_extensions-4.4.0-py3-none-any.whl (26 kB)
Collecting patsy>=0.5.2
  Downloading patsy-0.5.3-py2.py3-none-any.whl (233 kB)
     |████████████████████████████████| 233 kB 1.1 MB/s eta 0:00:01
Requirement already satisfied: MarkupSafe>=0.23 in /home/damar.pramuditya/.py
env/versions/anaconda3-2020.11/lib/python3.8/site-packages (from jinja2<3.2,>
=2.11.1->pandas-profiling>=2.8.0->pycaret) (1.1.1)
Collecting Mako
  Using cached Mako-1.2.3-py3-none-any.whl (78 kB)
Collecting importlib-resources; python_version < "3.9"
  Downloading importlib_resources-5.10.0-py3-none-any.whl (34 kB)
Collecting greenlet!=0.4.17; python_version >= "3" and (platform_machine ==
 "aarch64" or (platform_machine == "ppc64le" or (platform_machine == "x86_64"
or (platform_machine == "amd64" or (platform_machine == "AMD64" or (platform_
machine == "win32" or platform_machine == "WIN32"))))))
  Downloading greenlet-1.1.3.post0-cp38-cp38-manylinux_2_17_x86_64.manylinux2
014_x86_64.whl (157 kB)
     |████████████████████████████████| 157 kB 1.5 MB/s eta 0:00:01
Collecting pyjwt>=1.7.0
  Downloading PyJWT-2.6.0-py3-none-any.whl (20 kB)
Collecting oauthlib>=3.1.0
  Downloading oauthlib-3.2.2-py3-none-any.whl (151 kB)
     |████████████████████████████████| 151 kB 1.4 MB/s eta 0:00:01
Collecting tabulate>=0.7.7
  Downloading tabulate-0.9.0-py3-none-any.whl (35 kB)
Requirement already satisfied: prometheus-client in /home/damar.pramuditya/.p
yenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from prometheus-
flask-exporter<1->mlflow->pycaret) (0.8.0)
Collecting gitdb<5,>=4.0.1
  Using cached gitdb-4.0.9-py3-none-any.whl (63 kB)
Collecting websocket-client>=0.32.0
  Downloading websocket_client-1.4.1-py3-none-any.whl (55 kB)
     |████████████████████████████████| 55 kB 1.3 MB/s eta 0:00:01
Requirement already satisfied: zipp>=0.5 in /home/damar.pramuditya/.pyenv/ver
sions/anaconda3-2020.11/lib/python3.8/site-packages (from importlib-metadata!
=4.7.0,<6,>=3.7.0->mlflow->pycaret) (3.4.0)
Requirement already satisfied: itsdangerous>=0.24 in /home/damar.pramuditya/.
pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from Flask<3->m
lflow->pycaret) (1.1.0)
Requirement already satisfied: Werkzeug>=0.15 in /home/damar.pramuditya/.pyen
v/versions/anaconda3-2020.11/lib/python3.8/site-packages (from Flask<3->mlflo
w->pycaret) (1.0.1)
Requirement already satisfied: jupyter-core in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbformat>=4.2.0-
>ipywidgets->pycaret) (4.6.3)
Requirement already satisfied: jsonschema!=2.5.0,>=2.4 in /home/damar.pramudi
tya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbfor
mat>=4.2.0->ipywidgets->pycaret) (3.2.0)
Requirement already satisfied: ipython-genutils in /home/damar.pramuditya/.py
env/versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbformat>=4.
2.0->ipywidgets->pycaret) (0.2.0)
Requirement already satisfied: jupyter-client in /home/damar.pramuditya/.pyen
v/versions/anaconda3-2020.11/lib/python3.8/site-packages (from ipykernel>=4.
5.1->ipywidgets->pycaret) (6.1.7)
Requirement already satisfied: tornado>=4.2 in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from ipykernel>=4.5.1
->ipywidgets->pycaret) (6.0.4)
```

```
Requirement already satisfied: notebook>=4.4.1 in /home/damar.pramuditya/.pye
nv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from widgetsnbexte
nsion~=3.5.0->ipywidgets->pycaret) (6.1.4)
Requirement already satisfied: ptyprocess>=0.5 in /home/damar.pramuditya/.pye
nv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from pexpect>4.3;
 sys_platform != "win32"->IPython->pycaret) (0.6.0)
Requirement already satisfied: parso<0.8.0,>=0.7.0 in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from jedi>=
0.10->IPython->pycaret) (0.7.0)
Requirement already satisfied: wcwidth in /home/damar.pramuditya/.pyenv/versi
ons/anaconda3-2020.11/lib/python3.8/site-packages (from prompt-toolkit!=3.0.
0,!=3.0.1,<3.1.0,>=2.0.0->IPython->pycaret) (0.2.5)
Requirement already satisfied: PyWavelets in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from imagehash; extra =
= "type_image_path"->visions[type_image_path]==0.7.5->pandas-profiling>=2.8.0
->pycaret) (1.1.1)
Collecting smmap<6,>=3.0.1
  Using cached smmap-5.0.0-py3-none-any.whl (24 kB)
Requirement already satisfied: pyrsistent>=0.14.0 in /home/damar.pramuditya/.
pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from jsonschem
a!=2.5.0,>=2.4->nbformat>=4.2.0->ipywidgets->pycaret) (0.17.3)
Requirement already satisfied: pyzmq>=13 in /home/damar.pramuditya/.pyenv/ver
sions/anaconda3-2020.11/lib/python3.8/site-packages (from jupyter-client->ipy
kernel>=4.5.1->ipywidgets->pycaret) (19.0.2)
Requirement already satisfied: Send2Trash in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from notebook>=4.4.1->w
idgetsnbextension~=3.5.0->ipywidgets->pycaret) (1.5.0)
Requirement already satisfied: nbconvert in /home/damar.pramuditya/.pyenv/ver
sions/anaconda3-2020.11/lib/python3.8/site-packages (from notebook>=4.4.1->wi
dgetsnbextension~=3.5.0->ipywidgets->pycaret) (6.0.7)
Requirement already satisfied: terminado>=0.8.3 in /home/damar.pramuditya/.py
env/versions/anaconda3-2020.11/lib/python3.8/site-packages (from notebook>=4.
4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (0.9.1)
Requirement already satisfied: argon2-cffi in /home/damar.pramuditya/.pyenv/v
ersions/anaconda3-2020.11/lib/python3.8/site-packages (from notebook>=4.4.1->
widgetsnbextension~=3.5.0->ipywidgets->pycaret) (20.1.0)
Requirement already satisfied: bleach in /home/damar.pramuditya/.pyenv/versio
ns/anaconda3-2020.11/lib/python3.8/site-packages (from nbconvert->notebook>=
4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (3.2.1)
Requirement already satisfied: nbclient<0.6.0,>=0.5.0 in /home/damar.pramudit
ya/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbconv
ert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (0.5.1)
Requirement already satisfied: jupyterlab-pygments in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbconve
rt->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (0.1.2)
Requirement already satisfied: testpath in /home/damar.pramuditya/.pyenv/vers
ions/anaconda3-2020.11/lib/python3.8/site-packages (from nbconvert->notebook>
=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (0.4.4)
Requirement already satisfied: pandocfilters>=1.4.1 in /home/damar.pramudity
a/.pyenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbconve
rt->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (1.4.3)
Requirement already satisfied: defusedxml in /home/damar.pramuditya/.pyenv/ve
rsions/anaconda3-2020.11/lib/python3.8/site-packages (from nbconvert->noteboo
k>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (0.6.0)
Requirement already satisfied: mistune<2,>=0.8.1 in /home/damar.pramuditya/.p
yenv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbconvert->
notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (0.8.4)
Requirement already satisfied: cffi>=1.0.0 in /home/damar.pramuditya/.pyenv/v
ersions/anaconda3-2020.11/lib/python3.8/site-packages (from argon2-cffi->note
book>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (1.14.3)
Requirement already satisfied: webencodings in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from bleach->nbconver
t->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (0.5.1)
Requirement already satisfied: nest-asyncio in /home/damar.pramuditya/.pyenv/
versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbclient<0.6.0,>
=0.5.0->nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->py
caret) (1.4.2)
Requirement already satisfied: async-generator in /home/damar.pramuditya/.pye
nv/versions/anaconda3-2020.11/lib/python3.8/site-packages (from nbclient<0.6.
```

```
0,>=0.5.0->nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets-
>pycaret) (1.10)
Requirement already satisfied: pycparser in /home/damar.pramuditya/.pyenv/ver
sions/anaconda3-2020.11/lib/python3.8/site-packages (from cffi>=1.0.0->argon2
-cffi->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets->pycaret) (2.2
0)
Building wheels for collected packages: pyLDAvis, umap-learn, cufflinks, pyo
d, htmlmin, sklearn, pynndescent, databricks-cli
  Building wheel for pyLDAvis (PEP 517) ... done
  Created wheel for pyLDAvis: filename=pyLDAvis-3.3.1-py2.py3-none-any.whl si
ze=136882 sha256=791eb67696817ffc038ab65497991285882e52bc39d0df6c42db73cb3424
75cf
  Stored in directory: /home/damar.pramuditya/.cache/pip/wheels/90/61/ec/9dbe
9efc3acf9c4e37ba70fbbcc3f3a0ebd121060aa593181a
  Building wheel for umap-learn (setup.py) ... done
  Created wheel for umap-learn: filename=umap_learn-0.5.3-py3-none-any.whl si
ze=82820 sha256=01deed179b6ed428d53f92f23c9a8b7c18eb423cff6d83ce94fcec65206ef
112
  Stored in directory: /home/damar.pramuditya/.cache/pip/wheels/a9/3a/67/06a8
950e053725912e6a8c42c4a3a241410f6487b8402542ea
  Building wheel for cufflinks (setup.py) ... done
  Created wheel for cufflinks: filename=cufflinks-0.17.3-py3-none-any.whl siz
e=67921 sha256=8faf42f116d09d1e1d00652d77e0b583ee5792a9f2b67c3544ad57ac88a9d7
f5
  Stored in directory: /home/damar.pramuditya/.cache/pip/wheels/6b/76/62/6da9
7734911ffcbdd559fd1a3f28526321f0ae699182a23866
  Building wheel for pyod (setup.py) ... done
  Created wheel for pyod: filename=pyod-1.0.6-py3-none-any.whl size=175085 sh
a256=b488326e8151171c6817069a52a2d55b950263a5a7754fd76e3c078385312c97
  Stored in directory: /home/damar.pramuditya/.cache/pip/wheels/98/93/e6/6d40
410d9635ecde42d06041a1ba7f2ee7396e036fcf702e73
  Building wheel for htmlmin (setup.py) ... done
  Created wheel for htmlmin: filename=htmlmin-0.1.12-py3-none-any.whl size=27
084 sha256=4ea923e97c5c2d04b374585bc05f83f425078062eb94b67712daa20093755101
  Stored in directory: /home/damar.pramuditya/.cache/pip/wheels/23/14/6e/4be5
bfeeb027f4939a01764b48edd5996acf574b0913fe5243
  Building wheel for sklearn (setup.py) ... done
  Created wheel for sklearn: filename=sklearn-0.0-py2.py3-none-any.whl size=1
316 sha256=86ac01eed3671d8163de32d68b26b5167d7e236de5374b3d1c74b4b2966065a1
  Stored in directory: /home/damar.pramuditya/.cache/pip/wheels/22/0b/40/fd3f
795caaa1fb4c6cb738bc1f56100be1e57da95849bfc897
  Building wheel for pynndescent (setup.py) ... done
  Created wheel for pynndescent: filename=pynndescent-0.5.7-py3-none-any.whl
 size=54272 sha256=18fed5f6cf357c94b1d3d17909ab36df56ce0f4aa8750f1f1e3f1b528d
b01421
  Stored in directory: /home/damar.pramuditya/.cache/pip/wheels/1b/38/fe/99e2
2fbae88abd1c5e8d99253cba6d1c590cc7a94408bff3bf
  Building wheel for databricks-cli (setup.py) ... done
  Created wheel for databricks-cli: filename=databricks_cli-0.17.3-py3-none-a
ny.whl size=139100 sha256=cd5a9a6c9077669720647cb3d2693b5d3de106c9924ed5ce091
b073c0335f023
  Stored in directory: /home/damar.pramuditya/.cache/pip/wheels/58/40/7c/d021
d51dac18bfd095fb6837572ad2e6f1a34d221f4b1d976b
Successfully built pyLDAvis umap-learn cufflinks pyod htmlmin sklearn pynndes
cent databricks-cli
Installing collected packages: wordcloud, blis, cymem, catalogue, wasabi, srs
ly, plac, murmurhash, preshed, thinc, spacy, multimethod, tangled-up-in-unico
de, imagehash, visions, htmlmin, phik, typing-extensions, pydantic, missingn
o, patsy, statsmodels, pandas-profiling, textblob, smart-open, gensim, sklear
n, funcy, pyLDAvis, pynndescent, umap-learn, yellowbrick, tenacity, plotly, c
olorlover, cufflinks, mlxtend, importlib-metadata, greenlet, sqlalchemy, Mak
o, importlib-resources, alembic, protobuf, pyjwt, oauthlib, tabulate, databri
cks-cli, sqlparse, prometheus-flask-exporter, querystring-parser, smmap, gitd
b, gitpython, websocket-client, docker, gunicorn, mlflow, pyod, imbalanced-le
arn, Boruta, kmodes, scikit-plot, lightgbm, pycaret
  Attempting uninstall: typing-extensions
    Found existing installation: typing-extensions 3.7.4.3
    Uninstalling typing-extensions-3.7.4.3:
      Successfully uninstalled typing-extensions-3.7.4.3
```

```
   Attempting uninstall: patsy
     Found existing installation: patsy 0.5.1
     Uninstalling patsy-0.5.1:
       Successfully uninstalled patsy-0.5.1
   Attempting uninstall: statsmodels
     Found existing installation: statsmodels 0.12.0
     Uninstalling statsmodels-0.12.0:
       Successfully uninstalled statsmodels-0.12.0
   Attempting uninstall: importlib-metadata
     Found existing installation: importlib-metadata 2.0.0
     Uninstalling importlib-metadata-2.0.0:
       Successfully uninstalled importlib-metadata-2.0.0
   Attempting uninstall: greenlet
     Found existing installation: greenlet 0.4.17
     Uninstalling greenlet-0.4.17:
       Successfully uninstalled greenlet-0.4.17
   Attempting uninstall: sqlalchemy
     Found existing installation: SQLAlchemy 1.3.20
     Uninstalling SQLAlchemy-1.3.20:
       Successfully uninstalled SQLAlchemy-1.3.20
 ERROR: After October 2020 you may experience errors when installing or updati
 ng packages. This is because pip will change the way that it resolves depende
 ncy conflicts.

 We recommend you use --use-feature=2020-resolver to test your packages with t
 he new resolver before it becomes the default.

 statsmodels 0.13.2 requires packaging>=21.3, but you'll have packaging 20.4 w
 hich is incompatible.
 pyldavis 3.3.1 requires numpy>=1.20.0, but you'll have numpy 1.19.2 which is
  incompatible.
 pyldavis 3.3.1 requires pandas>=1.2.0, but you'll have pandas 1.1.3 which is
  incompatible.
 yellowbrick 1.5 requires scikit-learn>=1.0.0, but you'll have scikit-learn 0.
 23.2 which is incompatible.
 mlxtend 0.21.0 requires scikit-learn>=1.0.2, but you'll have scikit-learn 0.2
 3.2 which is incompatible.
 docker 6.0.0 requires requests>=2.26.0, but you'll have requests 2.24.0 which
 is incompatible.
 docker 6.0.0 requires urllib3>=1.26.0, but you'll have urllib3 1.25.11 which
  is incompatible.
 Successfully installed Boruta-0.3 Mako-1.2.3 alembic-1.8.1 blis-0.7.9 catalog
 ue-1.0.2 colorlover-0.3.0 cufflinks-0.17.3 cymem-2.0.7 databricks-cli-0.17.3
  docker-6.0.0 funcy-1.17 gensim-3.8.3 gitdb-4.0.9 gitpython-3.1.29 greenlet-
 1.1.3.post0 gunicorn-20.1.0 htmlmin-0.1.12 imagehash-4.3.1 imbalanced-learn-
 0.7.0 importlib-metadata-5.0.0 importlib-resources-5.10.0 kmodes-0.12.2 light
 gbm-3.3.3 missingno-0.5.1 mlflow-1.30.0 mlxtend-0.21.0 multimethod-1.9 murmur
 hash-1.0.9 oauthlib-3.2.2 pandas-profiling-3.4.0 patsy-0.5.3 phik-0.12.2 plac
 -1.1.3 plotly-5.11.0 preshed-3.0.8 prometheus-flask-exporter-0.20.3 protobuf-
 4.21.9 pyLDAvis-3.3.1 pycaret-2.3.10 pydantic-1.10.2 pyjwt-2.6.0 pynndescent-
 0.5.7 pyod-1.0.6 querystring-parser-1.2.4 scikit-plot-0.3.7 sklearn-0.0 smart
 -open-6.2.0 smmap-5.0.0 spacy-2.3.8 sqlalchemy-1.4.42 sqlparse-0.4.3 srsly-1.
 0.6 statsmodels-0.13.2 tabulate-9.0.0 tangled-up-in-unicode-0.2.0 tenacity-8.
 1.0 textblob-0.17.1 thinc-7.4.6 typing-extensions-4.4.0 umap-learn-0.5.3 visi
 ons-0.7.5 wasabi-0.10.1 websocket-client-1.4.1 wordcloud-1.8.2.2 yellowbrick-
 1.5
```

In [1]:
```python
import nltk
import sklearn
import pycaret
import pandas as pd

print('The nltk version is {}.'.format(nltk.__version__))
print('The scikit-learn version is {}.'.format(sklearn.__version__))
print('The pycaret version is {}.'.format(pycaret.__version__))
print('The pandas version is {}.'.format(pd.__version__))
```

```
The nltk version is 3.5.
The scikit-learn version is 0.23.2.
```

```
The pycaret version is 2.3.10.
The pandas version is 1.1.3.
```

In [2]:
```
dataset = pd.read_csv('adult.csv')
#check the shape of data
dataset.shape
```

Out[2]: (32563, 15)

In [3]:
```
dataset.head(5)
```

Out[3]:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship
0	continuous	Private Self-emp-not-inc Self-emp-inc Federal-...	continuous	Bachelors Some-college 11th HS-grad Prof-schoo...	continuous	Married-civ-spouse Divorced Never-married Sepa...	Tech-support Craft-repair Other-service Sales ...	Wife Own-child Husband Not-in-family Other-rel...
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	39.0	State-gov	77516.0	Bachelors	13.0	Never-married	Adm-clerical	Not-in-family
3	50.0	Self-emp-not-inc	83311.0	Bachelors	13.0	Married-civ-spouse	Exec-managerial	Husband
4	38.0	Private	215646.0	HS-grad	9.0	Divorced	Handlers-cleaners	Not-in-family

In [5]:
```
data = dataset.sample(frac=0.95, random_state=786)
data_unseen = dataset.drop(data.index)

data.reset_index(drop=True, inplace=True)
data_unseen.reset_index(drop=True, inplace=True)

print('Data for Modeling: ' + str(data.shape))
print('Unseen Data For Predictions: ' + str(data_unseen.shape))
```

```
Data for Modeling: (30935, 15)
Unseen Data For Predictions: (1628, 15)
```

In [6]:
```
from pycaret.clustering import *

exp_clu101 = setup(data, normalize = True,
                   session_id = 123)
```

	Description	Value
0	session_id	123
1	Original Data	(30935, 15)
2	Missing Values	True
3	Numeric Features	0
4	Categorical Features	15
5	Ordinal Features	False
6	High Cardinality Features	False
7	High Cardinality Method	None

	Description	Value
8	Transformed Data	(30935, 21355)
9	CPU Jobs	-1
10	Use GPU	False
11	Log Experiment	False
12	Experiment Name	cluster-default-name
13	USI	8249
14	Imputation Type	simple
15	Iterative Imputation Iteration	None
16	Numeric Imputer	mean
17	Iterative Imputation Numeric Model	None
18	Categorical Imputer	mode
19	Iterative Imputation Categorical Model	None
20	Unknown Categoricals Handling	least_frequent
21	Normalize	True
22	Normalize Method	zscore
23	Transformation	False
24	Transformation Method	None
25	PCA	False
26	PCA Method	None
27	PCA Components	None
28	Ignore Low Variance	False
29	Combine Rare Levels	False
30	Rare Level Threshold	None
31	Numeric Binning	False
32	Remove Outliers	False
33	Outliers Threshold	None
34	Remove Multicollinearity	False
35	Multicollinearity Threshold	None
36	Remove Perfect Collinearity	False
37	Clustering	False
38	Clustering Iteration	None
39	Polynomial Features	False
40	Polynomial Degree	None
41	Trignometry Features	False
42	Polynomial Threshold	None
43	Group Features	False
44	Feature Selection	False
45	Feature Selection Method	classic
46	Features Selection Threshold	None

	Description	Value
47	Feature Interaction	False
48	Feature Ratio	False
49	Interaction Threshold	None

In [7]:
```
models()
```

Out[7]:

ID	Name	Reference
kmeans	K-Means Clustering	sklearn.cluster._kmeans.KMeans
ap	Affinity Propagation	sklearn.cluster._affinity_propagation.Affinity...
meanshift	Mean Shift Clustering	sklearn.cluster._mean_shift.MeanShift
sc	Spectral Clustering	sklearn.cluster._spectral.SpectralClustering
hclust	Agglomerative Clustering	sklearn.cluster._agglomerative.AgglomerativeCl...
dbscan	Density-Based Spatial Clustering	sklearn.cluster._dbscan.DBSCAN
optics	OPTICS Clustering	sklearn.cluster._optics.OPTICS
birch	Birch Clustering	sklearn.cluster._birch.Birch
kmodes	K-Modes Clustering	kmodes.kmodes.KModes

In [8]:
```
kmeans = create_model('kmeans')
```

	Silhouette	Calinski-Harabasz	Davies-Bouldin	Homogeneity	Rand Index	Completeness
0	0.0652	2035.138	3.4917	0	0	0

In [11]:
```
print(kmeans)
```

```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
       n_clusters=4, n_init=10, n_jobs=-1, precompute_distances='deprecated',
       random_state=123, tol=0.0001, verbose=0)
```
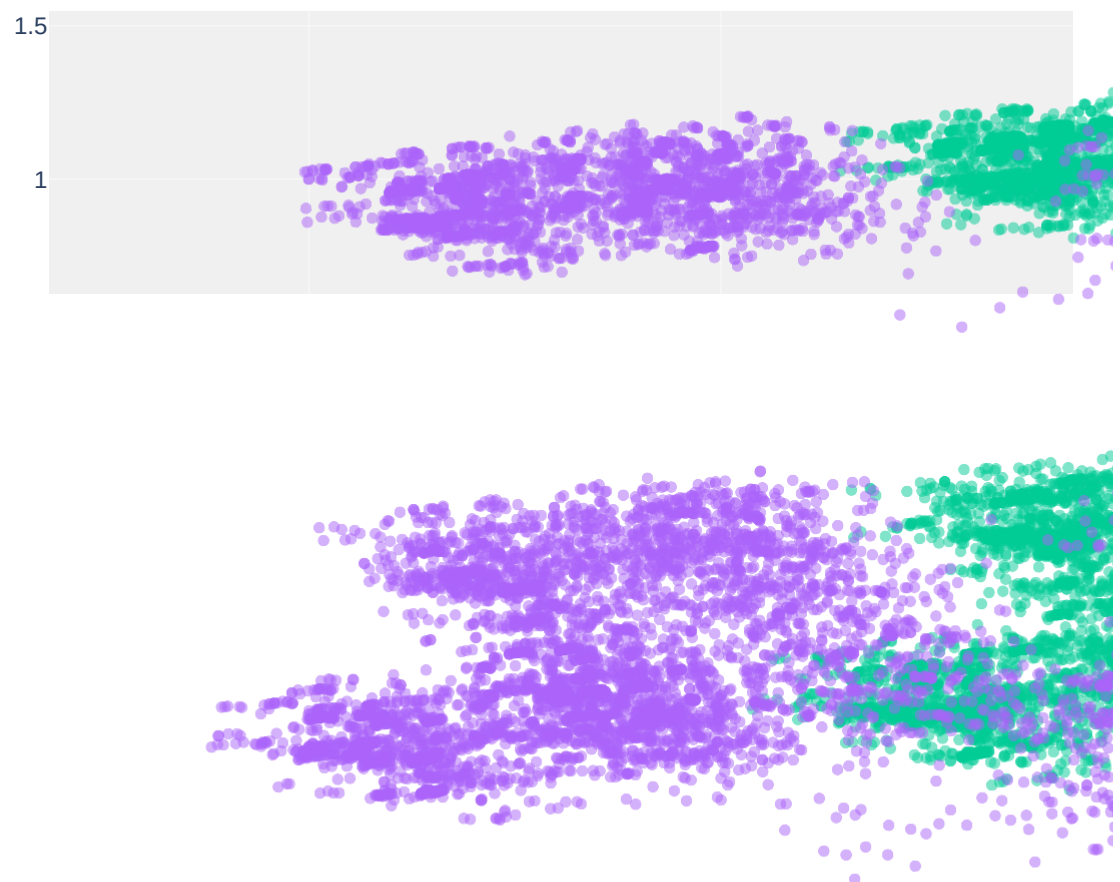
In [15]:
```
kmean_results = assign_model(kmeans)
kmean_results.head(5)
```

Out[15]:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	s
0	31.0	Private	352465.0	Some-college	10.0	Married-civ-spouse	Exec-managerial	Husband	White	Ma
1	46.0	Self-emp-inc	120902.0	Bachelors	13.0	Married-civ-spouse	Exec-managerial	Husband	White	Ma
2	28.0	Private	94880.0	Some-college	10.0	Never-married	Craft-repair	Not-in-family	White	Ma
3	33.0	Private	409172.0	Bachelors	13.0	Married-civ-spouse	Exec-managerial	Own-child	White	Ma
4	25.0	Private	176729.0	HS-grad	9.0	Married-civ-spouse	Craft-repair	Husband	White	Ma

In [16]:
```python
plot_model(kmeans)
```

2D Cluster PCA Plot

In [17]:
```python
plot_model(kmeans, plot = 'elbow')
```

Distortion Score Elbow for KMeans Clustering
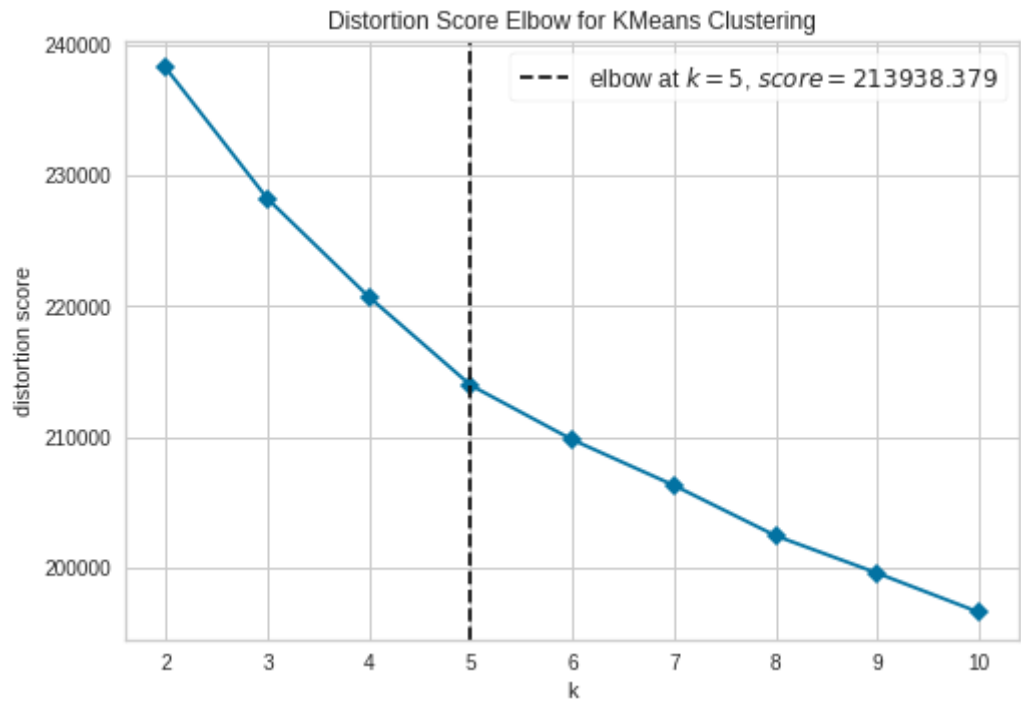

```
In [18]:   plot_model(kmeans, plot = 'silhouette')
```

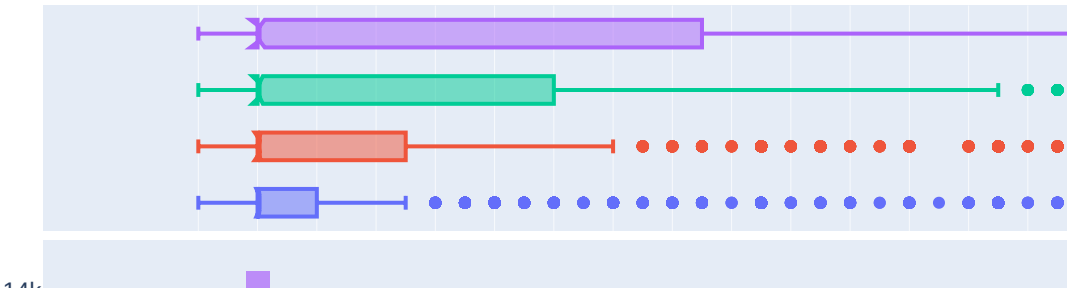Silhouette Plot of KMeans Clustering for 30935 Samples in 4 Centers
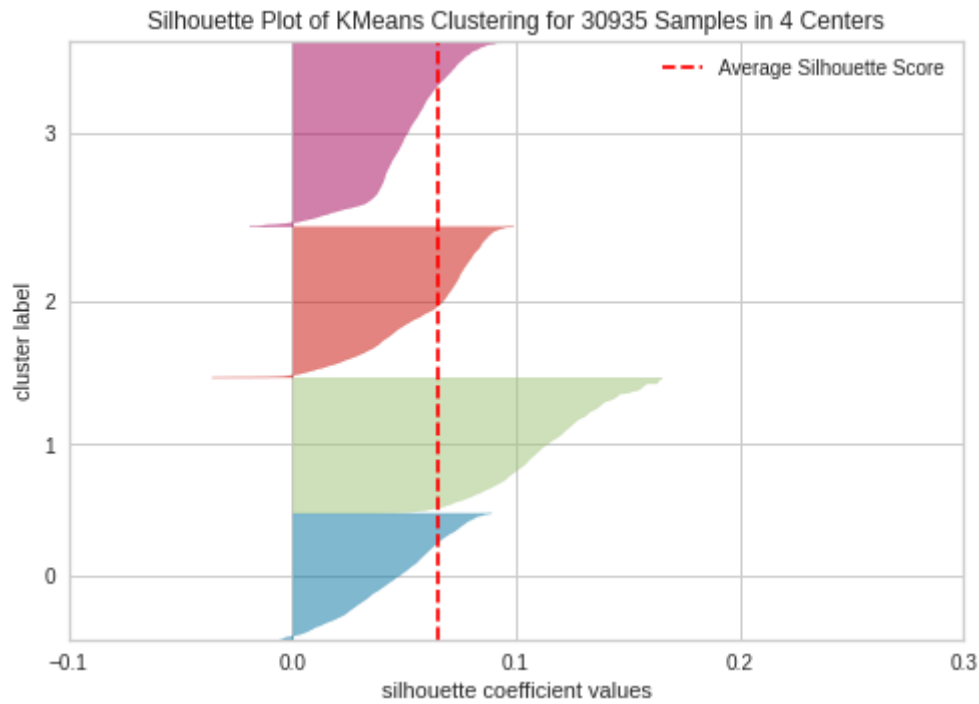

```
In [19]:   plot_model(kmeans, plot = 'distribution') #to see size of clusters
```
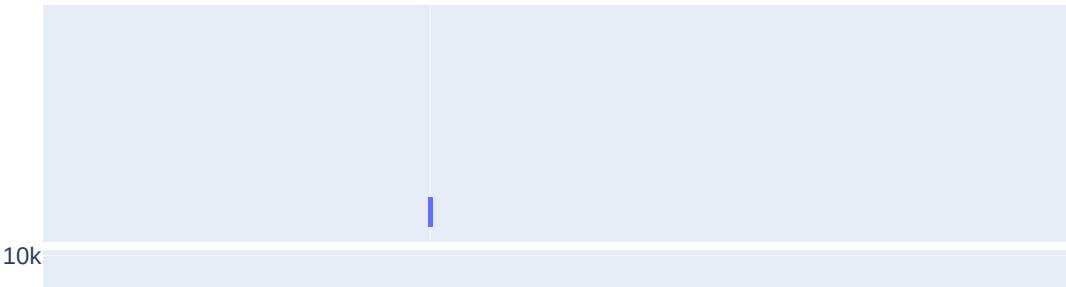
In [21]:
```
plot_model(kmeans, plot = 'distribution', feature = 'capital-gain')
```

In [22]: `plot_model(kmeans, plot = 'distribution', feature = 'hours-per-week')`
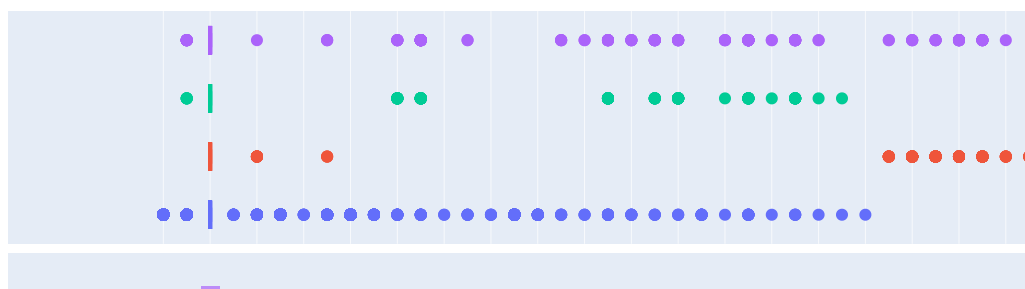
In [24]:
```
unseen_predictions = predict_model(kmeans, data=data_unseen)
unseen_predictions.head(5)
```

Out[24]:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	
0	53.0	Private	234721.0	11th	7.0	Married-civ-spouse	Handlers-cleaners	Husband	Black	
1	19.0	Private	544091.0	HS-grad	9.0	Married-AF-spouse	Adm-clerical	Wife	White	Fer
2	43.0	Private	237993.0	Some-college	10.0	Married-civ-spouse	Tech-support	Husband	White	
3	19.0	Private	101509.0	Some-college	10.0	Never-married	Prof-specialty	Own-child	White	
4	44.0	Private	198282.0	Bachelors	13.0	Married-civ-spouse	Exec-managerial	Husband	White	

In [25]:
```
save_model(kmeans,'Final KMeans Model 30-10-2022')
```

Transformation Pipeline and Model Successfully Saved

Out[25]:
```
(Pipeline(memory=None,
          steps=[('dtypes',
                  DataTypes_Auto_infer(categorical_features=[],
                                       display_types=True, features_todrop=
[],
                                       id_columns=[], ml_usecase='regressio
n',
                                       numerical_features=[],
                                       target='UNSUPERVISED_DUMMY_TARGET',
                                       time_features=[])),
                 ('imputer',
                  Simple_Imputer(categorical_strategy='most frequent',
                                 fill_value_categorical=None,
                                 fill_value_numerical=None...
                 ('fix_perfect', 'passthrough'),
                 ('clean_names', Clean_Colum_Names()),
                 ('feature_select', 'passthrough'), ('fix_multi', 'passthroug
h'),
                 ('dfs', 'passthrough'), ('pca', 'passthrough'),
                 ['trained_model',
                  KMeans(algorithm='auto', copy_x=True, init='k-means++',
                         max_iter=300, n_clusters=4, n_init=10, n_jobs=-1,
                         precompute_distances='deprecated', random_state=123,
                         tol=0.0001, verbose=0)]],
          verbose=False),
 'Final KMeans Model 30-10-2022.pkl')
```

In [26]:
```
saved_kmeans = load_model('Final KMeans Model 30-10-2022')
```

Transformation Pipeline and Model Successfully Loaded

In [28]:
```
new_prediction = predict_model(saved_kmeans, data=data_unseen)
new_prediction.head(5)
```

Out[28]:

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	
0	53.0	Private	234721.0	11th	7.0	Married-civ-spouse	Handlers-cleaners	Husband	Black	
1	19.0	Private	544091.0	HS-grad	9.0	Married-AF-spouse	Adm-clerical	Wife	White	Fei
2	43.0	Private	237993.0	Some-college	10.0	Married-civ-spouse	Tech-support	Husband	White	
3	19.0	Private	101509.0	Some-college	10.0	Never-married	Prof-specialty	Own-child	White	
4	44.0	Private	198282.0	Bachelors	13.0	Married-civ-spouse	Exec-managerial	Husband	White	

In [ ]: