installing necessary libraries

In [ ]:
```
!pip install pandas numpy
%pip install pycaret
```

```
Requirement already satisfied: pandas in c:\users\teste\appdata\local\programs\pytho
n\python311\lib\site-packages (1.5.2)
Requirement already satisfied: numpy in c:\users\teste\appdata\local\programs\python
\python311\lib\site-packages (1.24.1)
Requirement already satisfied: python-dateutil>=2.8.1 in c:\users\teste\appdata\roam
ing\python\python311\site-packages (from pandas) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\teste\appdata\local\programs
\python\python311\lib\site-packages (from pandas) (2022.7)
Requirement already satisfied: six>=1.5 in c:\users\teste\appdata\roaming\python\pyt
hon311\site-packages (from python-dateutil>=2.8.1->pandas) (1.16.0)
[notice] A new release of pip is available: 23.2.1 -> 23.3.1
[notice] To update, run: python.exe -m pip install --upgrade pip
```

```
Requirement already satisfied: pycaret in c:\users\teste\documents\3rd year\cloud ai
\cloud ai github\cloudai\venv_caret\lib\site-packages (3.1.0)
Requirement already satisfied: ipython>=5.5.0 in c:\users\teste\documents\3rd year\c
loud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (8.16.1)
Requirement already satisfied: ipywidgets>=7.6.5 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (8.1.
1)
Requirement already satisfied: tqdm>=4.62.0 in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (4.66.1)
Requirement already satisfied: numpy<1.24,>=1.21 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (1.2
3.5)
Requirement already satisfied: pandas<2.0.0,>=1.3.0 in c:\users\teste\documents\3rd
year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret)
(1.5.3)
Requirement already satisfied: jinja2>=1.2 in c:\users\teste\documents\3rd year\clou
d ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (3.1.2)
Requirement already satisfied: scipy~=1.10.1 in c:\users\teste\documents\3rd year\cl
oud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (1.10.1)
Requirement already satisfied: joblib>=1.2.0 in c:\users\teste\documents\3rd year\cl
oud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (1.3.2)
Requirement already satisfied: scikit-learn<1.3.0,>=1.0 in c:\users\teste\documents
\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycare
t) (1.2.2)
Requirement already satisfied: pyod>=1.0.8 in c:\users\teste\documents\3rd year\clou
d ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (1.1.0)
Requirement already satisfied: imbalanced-learn>=0.8.1 in c:\users\teste\documents\3
rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret)
(0.11.0)
Requirement already satisfied: category-encoders>=2.4.0 in c:\users\teste\documents
\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycare
t) (2.6.2)
Requirement already satisfied: lightgbm>=3.0.0 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (4.1.
0)
Requirement already satisfied: numba>=0.55.0 in c:\users\teste\documents\3rd year\cl
oud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (0.58.0)
Requirement already satisfied: requests>=2.27.1 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (2.31.
0)
Requirement already satisfied: psutil>=5.9.0 in c:\users\teste\documents\3rd year\cl
oud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (5.9.5)
Requirement already satisfied: markupsafe>=2.0.1 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (2.1.
3)
Requirement already satisfied: importlib-metadata>=4.12.0 in c:\users\teste\document
s\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycar
et) (6.8.0)
Requirement already satisfied: nbformat>=4.2.0 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (5.9.
2)
Requirement already satisfied: cloudpickle in c:\users\teste\documents\3rd year\clou
d ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (3.0.0)
Requirement already satisfied: deprecation>=2.1.0 in c:\users\teste\documents\3rd ye
ar\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (2.
1.0)
```

```
Requirement already satisfied: xxhash in c:\users\teste\documents\3rd year\cloud ai
\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (3.4.1)
Requirement already satisfied: matplotlib>=3.3.0 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (3.8.
0)
Requirement already satisfied: scikit-plot>=0.3.7 in c:\users\teste\documents\3rd ye
ar\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (0.
3.7)
Requirement already satisfied: yellowbrick>=1.4 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (1.5)
Requirement already satisfied: plotly>=5.0.0 in c:\users\teste\documents\3rd year\cl
oud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (5.17.0)
Requirement already satisfied: kaleido>=0.2.1 in c:\users\teste\documents\3rd year\c
loud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (0.2.1)
Requirement already satisfied: schemdraw==0.15 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (0.15)
Requirement already satisfied: plotly-resampler>=0.8.3.1 in c:\users\teste\documents
\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycare
t) (0.9.1)
Requirement already satisfied: statsmodels>=0.12.1 in c:\users\teste\documents\3rd y
ear\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (0.
14.0)
Requirement already satisfied: sktime!=0.17.1,!=0.17.2,!=0.18.0,<0.22.0,>=0.16.1 in
c:\users\teste\documents\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\si
te-packages (from pycaret) (0.21.1)
Requirement already satisfied: tbats>=1.1.3 in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pycaret) (1.1.3)
Requirement already satisfied: pmdarima!=1.8.1,<3.0.0,>=1.8.0 in c:\users\teste\docu
ments\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from p
ycaret) (2.0.3)
Requirement already satisfied: patsy>=0.5.1 in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from category-encoders>=
2.4.0->pycaret) (0.5.3)
Requirement already satisfied: packaging in c:\users\teste\documents\3rd year\cloud
ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from deprecation>=2.1.0->py
caret) (23.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in c:\users\teste\documents\3rd
year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from imbalanced-
learn>=0.8.1->pycaret) (3.2.0)
Requirement already satisfied: zipp>=0.5 in c:\users\teste\documents\3rd year\cloud
ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from importlib-metadata>=4.
12.0->pycaret) (3.17.0)
Requirement already satisfied: backcall in c:\users\teste\documents\3rd year\cloud a
i\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0->pycare
t) (0.2.0)
Requirement already satisfied: decorator in c:\users\teste\documents\3rd year\cloud
ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0->pycare
t) (5.1.1)
Requirement already satisfied: jedi>=0.16 in c:\users\teste\documents\3rd year\cloud
ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0->pycare
t) (0.19.1)
Requirement already satisfied: matplotlib-inline in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0
->pycaret) (0.1.6)
Requirement already satisfied: pickleshare in c:\users\teste\documents\3rd year\clou
d ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0->pyca
```

```
ret) (0.7.5)
Requirement already satisfied: prompt-toolkit!=3.0.37,<3.1.0,>=3.0.30 in c:\users\te
ste\documents\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages
(from ipython>=5.5.0->pycaret) (3.0.39)
Requirement already satisfied: pygments>=2.4.0 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0-
>pycaret) (2.16.1)
Requirement already satisfied: stack-data in c:\users\teste\documents\3rd year\cloud
ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0->pycare
t) (0.6.3)
Requirement already satisfied: traitlets>=5 in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0->pyc
aret) (5.11.2)
Requirement already satisfied: exceptiongroup in c:\users\teste\documents\3rd year\c
loud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0->p
ycaret) (1.1.3)
Requirement already satisfied: colorama in c:\users\teste\documents\3rd year\cloud a
i\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipython>=5.5.0->pycare
t) (0.4.6)
Requirement already satisfied: comm>=0.1.3 in c:\users\teste\documents\3rd year\clou
d ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipywidgets>=7.6.5->p
ycaret) (0.1.4)
Requirement already satisfied: widgetsnbextension~=4.0.9 in c:\users\teste\documents
\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipywid
gets>=7.6.5->pycaret) (4.0.9)
Requirement already satisfied: jupyterlab-widgets~=3.0.9 in c:\users\teste\documents
\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from ipywid
gets>=7.6.5->pycaret) (3.0.9)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from matplotlib>=3.
3.0->pycaret) (1.1.1)
Requirement already satisfied: cycler>=0.10 in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from matplotlib>=3.3.0->
pycaret) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from matplotlib>=3.
3.0->pycaret) (4.43.1)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from matplotlib>=3.
3.0->pycaret) (1.4.5)
Requirement already satisfied: pillow>=6.2.0 in c:\users\teste\documents\3rd year\cl
oud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from matplotlib>=3.3.0-
>pycaret) (10.1.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from matplotlib>=3.
3.0->pycaret) (3.1.1)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\teste\documents\3rd
year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from matplotlib>
=3.3.0->pycaret) (2.8.2)
Requirement already satisfied: fastjsonschema in c:\users\teste\documents\3rd year\c
loud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from nbformat>=4.2.0->
pycaret) (2.18.1)
Requirement already satisfied: jsonschema>=2.6 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from nbformat>=4.2.0
->pycaret) (4.19.1)
Requirement already satisfied: jupyter-core in c:\users\teste\documents\3rd year\clo
```

```
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from nbformat>=4.2.0->py
caret) (5.4.0)
Requirement already satisfied: llvmlite<0.42,>=0.41.0dev0 in c:\users\teste\document
s\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from numba
>=0.55.0->pycaret) (0.41.0)
Requirement already satisfied: pytz>=2020.1 in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pandas<2.0.0,>=1.3.
0->pycaret) (2023.3.post1)
Requirement already satisfied: tenacity>=6.2.0 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from plotly>=5.0.0->
pycaret) (8.2.3)
Requirement already satisfied: dash<3.0.0,>=2.11.0 in c:\users\teste\documents\3rd y
ear\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from plotly-resam
pler>=0.8.3.1->pycaret) (2.14.0)
Requirement already satisfied: orjson<4.0.0,>=3.8.0 in c:\users\teste\documents\3rd
year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from plotly-resa
mpler>=0.8.3.1->pycaret) (3.9.9)
Requirement already satisfied: trace-updater>=0.0.8 in c:\users\teste\documents\3rd
year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from plotly-resa
mpler>=0.8.3.1->pycaret) (0.0.9.1)
Requirement already satisfied: tsdownsample==0.1.2 in c:\users\teste\documents\3rd y
ear\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from plotly-resam
pler>=0.8.3.1->pycaret) (0.1.2)
Requirement already satisfied: Cython!=0.29.18,!=0.29.31,>=0.29 in c:\users\teste\do
cuments\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from
pmdarima!=1.8.1,<3.0.0,>=1.8.0->pycaret) (3.0.3)
Requirement already satisfied: urllib3 in c:\users\teste\documents\3rd year\cloud ai
\cloud ai github\cloudai\venv_caret\lib\site-packages (from pmdarima!=1.8.1,<3.0.0,>
=1.8.0->pycaret) (2.0.6)
Requirement already satisfied: setuptools!=50.0.0,>=38.6.0 in c:\users\teste\documen
ts\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from pmda
rima!=1.8.1,<3.0.0,>=1.8.0->pycaret) (65.5.0)
Requirement already satisfied: six in c:\users\teste\documents\3rd year\cloud ai\clo
ud ai github\cloudai\venv_caret\lib\site-packages (from pyod>=1.0.8->pycaret) (1.16.
0)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\teste\documents
\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from reques
ts>=2.27.1->pycaret) (3.3.0)
Requirement already satisfied: idna<4,>=2.5 in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from requests>=2.27.1->p
ycaret) (3.4)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\teste\documents\3rd ye
ar\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from requests>=2.2
7.1->pycaret) (2023.7.22)
Requirement already satisfied: deprecated>=1.2.13 in c:\users\teste\documents\3rd ye
ar\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from sktime!=0.17.
1,!=0.17.2,!=0.18.0,<0.22.0,>=0.16.1->pycaret) (1.2.14)
Requirement already satisfied: scikit-base<0.6.0 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from sktime!=0.17.
1,!=0.17.2,!=0.18.0,<0.22.0,>=0.16.1->pycaret) (0.5.2)
Requirement already satisfied: Flask<2.3.0,>=1.0.4 in c:\users\teste\documents\3rd y
ear\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash<3.0.0,>
=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (2.2.5)
Requirement already satisfied: Werkzeug<2.3.0 in c:\users\teste\documents\3rd year\c
loud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash<3.0.0,>=2.1
1.0->plotly-resampler>=0.8.3.1->pycaret) (2.2.3)
```

```
Requirement already satisfied: dash-html-components==2.0.0 in c:\users\teste\documen
ts\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash
<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (2.0.0)
Requirement already satisfied: dash-core-components==2.0.0 in c:\users\teste\documen
ts\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash
<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (2.0.0)
Requirement already satisfied: dash-table==5.0.0 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash<3.0.0,>=
2.11.0->plotly-resampler>=0.8.3.1->pycaret) (5.0.0)
Requirement already satisfied: typing-extensions>=4.1.1 in c:\users\teste\documents
\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash<
3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (4.8.0)
Requirement already satisfied: retrying in c:\users\teste\documents\3rd year\cloud a
i\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash<3.0.0,>=2.11.0->pl
otly-resampler>=0.8.3.1->pycaret) (1.3.4)
Requirement already satisfied: ansi2html in c:\users\teste\documents\3rd year\cloud
ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash<3.0.0,>=2.11.0->p
lotly-resampler>=0.8.3.1->pycaret) (1.8.0)
Requirement already satisfied: nest-asyncio in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from dash<3.0.0,>=2.11.0
->plotly-resampler>=0.8.3.1->pycaret) (1.5.8)
Requirement already satisfied: wrapt<2,>=1.10 in c:\users\teste\documents\3rd year\c
loud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from deprecated>=1.2.1
3->sktime!=0.17.1,!=0.17.2,!=0.18.0,<0.22.0,>=0.16.1->pycaret) (1.15.0)
Requirement already satisfied: parso<0.9.0,>=0.8.3 in c:\users\teste\documents\3rd y
ear\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from jedi>=0.16->
ipython>=5.5.0->pycaret) (0.8.3)
Requirement already satisfied: attrs>=22.2.0 in c:\users\teste\documents\3rd year\cl
oud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from jsonschema>=2.6->n
bformat>=4.2.0->pycaret) (23.1.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in c:\users\test
e\documents\3rd year\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages
(from jsonschema>=2.6->nbformat>=4.2.0->pycaret) (2023.7.1)
Requirement already satisfied: referencing>=0.28.4 in c:\users\teste\documents\3rd y
ear\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from jsonschema>=
2.6->nbformat>=4.2.0->pycaret) (0.30.2)
Requirement already satisfied: rpds-py>=0.7.1 in c:\users\teste\documents\3rd year\c
loud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from jsonschema>=2.6->
nbformat>=4.2.0->pycaret) (0.10.6)
Requirement already satisfied: wcwidth in c:\users\teste\documents\3rd year\cloud ai
\cloud ai github\cloudai\venv_caret\lib\site-packages (from prompt-toolkit!=3.0.37,<
3.1.0,>=3.0.30->ipython>=5.5.0->pycaret) (0.2.8)
Requirement already satisfied: platformdirs>=2.5 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from jupyter-core->
nbformat>=4.2.0->pycaret) (3.11.0)
Requirement already satisfied: pywin32>=300 in c:\users\teste\documents\3rd year\clo
ud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from jupyter-core->nbfor
mat>=4.2.0->pycaret) (306)
Requirement already satisfied: executing>=1.2.0 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from stack-data->ipy
thon>=5.5.0->pycaret) (2.0.0)
Requirement already satisfied: asttokens>=2.1.0 in c:\users\teste\documents\3rd year
\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from stack-data->ipy
thon>=5.5.0->pycaret) (2.4.0)
Requirement already satisfied: pure-eval in c:\users\teste\documents\3rd year\cloud
ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from stack-data->ipython>=
```

```
5.5.0->pycaret) (0.2.2)
Requirement already satisfied: itsdangerous>=2.0 in c:\users\teste\documents\3rd yea
r\cloud ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from Flask<2.3.0,>=
1.0.4->dash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (2.1.2)
Requirement already satisfied: click>=8.0 in c:\users\teste\documents\3rd year\cloud
ai\cloud ai github\cloudai\venv_caret\lib\site-packages (from Flask<2.3.0,>=1.0.4->d
ash<3.0.0,>=2.11.0->plotly-resampler>=0.8.3.1->pycaret) (8.1.7)
Note: you may need to restart the kernel to use updated packages.
[notice] A new release of pip is available: 23.2.1 -> 23.3.1
[notice] To update, run: c:\Users\teste\Documents\3rd Year\Cloud Ai\Cloud AI Github
\CloudAI\venv_caret\Scripts\python.exe -m pip install --upgrade pip
```

In [ ]:
```
import pandas as pd
import pycaret
crop_df = pd.read_csv('cropStats.csv',delimiter=',')
crop_df
```

Out[ ]:

	Country	Location	1900	1901	1902	1903	1904	
0	United States	ALABAMA	1064308.40	1044074.40	1084542.40	1088589.20	1052168.00	1
1	United States	ARIZONA	4451.48	4856.16	4046.80	4856.16	4046.80	
2	United States	ARKANSAS	922670.40	886249.20	906483.20	870062.00	837687.60	
3	United States	CALIFORNIA	25090.16	25494.84	25899.52	25090.16	23876.12	
4	United States	COLORADO	38039.92	44110.12	53013.08	57869.24	66367.52	
5	United States	DELAWARE	78912.60	78103.24	77698.56	80126.64	78912.60	
6	United States	FLORIDA	232691.00	230667.60	238761.20	240784.60	240784.60	
7	United States	GEORGIA	1444707.60	1388052.40	1388052.40	1363771.60	1339490.80	1
8	United States	IDAHO	4046.80	4046.80	3642.12	4046.80	3237.44	
9	United States	ILLINOIS	4232952.80	4330076.00	4390778.00	4269374.00	4249140.00	4
10	United States	INDIANA	2033517.00	2043634.00	2114453.00	2084102.00	2205506.00	2
11	United States	IOWA	3682588.00	3828272.80	3868740.80	3403358.80	3864694.00	3
12	United States	KANSAS	3023769.00	2766392.50	2889415.20	2714188.80	2714188.80	2
13	United States	KENTUCKY	1363771.60	1375912.00	1416380.00	1386029.00	1497316.00	1
14	United States	LOUISIANA	526084.00	509896.80	505850.00	489662.80	473475.60	
15	United States	MARYLAND	270326.24	271135.60	273159.00	267088.80	269112.20	
16	United States	MICHIGAN	635347.60	647488.00	655581.60	655581.60	647488.00	
17	United States	MINNESOTA	598926.40	651534.80	728424.00	724377.20	789126.00	
18	United States	MISSISSIPPI	849828.00	857921.60	870062.00	882202.40	829594.00	

	Country	Location	1900	1901	1902	1903	1904	
19	United States	MISSOURI	3116036.00	3055334.00	3055334.00	2852994.00	2691122.00	2
20	United States	MONTANA	1214.04	1618.72	2023.40	2023.40	2428.08	
21	United States	NEBRASKA	2974398.00	2903579.00	2974398.00	2863111.00	2944047.00	2
22	United States	NEW JERSEY	127878.88	120189.96	125046.12	114929.12	114929.12	
23	United States	NEW MEXICO	18210.60	19829.32	21448.04	23876.12	22662.08	
24	United States	NEW YORK	341954.60	319697.20	327790.80	307556.80	311603.60	
25	United States	NORTH CAROLINA	1092636.00	1040027.60	1080495.60	1027887.20	1027887.20	1
26	United States	NORTH DAKOTA	35611.84	55845.84	50989.68	55036.48	60702.00	
27	United States	OHIO	1618720.00	1568135.00	1618720.00	1578252.00	1608603.00	1
28	United States	OKLAHOMA	1064308.40	1116916.80	1278788.80	1250461.20	1456848.00	1
29	United States	OREGON	9712.32	10926.36	11331.04	11735.72	11735.72	
30	United States	PENNSYLVANIA	635347.60	629277.40	635347.60	617137.00	598926.40	
31	United States	SOUTH CAROLINA	704143.20	667722.00	692002.80	663675.20	639394.40	
32	United States	SOUTH DAKOTA	517990.40	586786.00	671768.80	687956.00	708190.00	
33	United States	TENNESSEE	1311163.20	1343537.60	1428520.40	1392099.20	1384005.60	1
34	United States	TEXAS	1954604.40	1938417.20	1974838.40	1974838.40	1954604.40	2
35	United States	UTAH	4451.48	4451.48	4451.48	4451.48	4046.80	
36	United States	VIRGINIA	768892.00	768892.00	776985.60	748658.00	748658.00	
37	United States	WASHINGTON	6070.20	6474.88	6879.56	7284.24	7688.92	

	Country	Location	1900	1901	1902	1903	1904
38	United States	WEST VIRGINIA	297439.80	291369.60	295416.40	279229.20	279229.20
39	United States	WISCONSIN	637371.00	659628.40	679862.40	667722.00	659628.40
40	United States	WYOMING	1618.72	2023.40	2428.08	2428.08	2428.08

41 rows × 121 columns

we only need the location and the year so I will remove the country column and the total

In [ ]:
```
# Dropping the first and last columns
crop_df = crop_df.iloc[:, 1:-1]  # Selects all rows, and columns from index 1 to th

crop_df
```

Out[ ]:

	Location	1900	1901	1902	1903	1904	1905
0	ALABAMA	1064308.40	1044074.40	1084542.40	1088589.20	1052168.00	1072402.00
1	ARIZONA	4451.48	4856.16	4046.80	4856.16	4046.80	5260.84
2	ARKANSAS	922670.40	886249.20	906483.20	870062.00	837687.60	821500.40
3	CALIFORNIA	25090.16	25494.84	25899.52	25090.16	23876.12	24685.48
4	COLORADO	38039.92	44110.12	53013.08	57869.24	66367.52	72842.40
5	DELAWARE	78912.60	78103.24	77698.56	80126.64	78912.60	76889.20
6	FLORIDA	232691.00	230667.60	238761.20	240784.60	240784.60	248878.20
7	GEORGIA	1444707.60	1388052.40	1388052.40	1363771.60	1339490.80	1408286.40
8	IDAHO	4046.80	4046.80	3642.12	4046.80	3237.44	3642.12
9	ILLINOIS	4232952.80	4330076.00	4390778.00	4269374.00	4249140.00	4249140.00
10	INDIANA	2033517.00	2043634.00	2114453.00	2084102.00	2205506.00	2023400.00
11	IOWA	3682588.00	3828272.80	3868740.80	3403358.80	3864694.00	3884928.00
12	KANSAS	3023769.00	2766392.50	2889415.20	2714188.80	2714188.80	2849756.60
13	KENTUCKY	1363771.60	1375912.00	1416380.00	1386029.00	1497316.00	1436614.00
14	LOUISIANA	526084.00	509896.80	505850.00	489662.80	473475.60	473475.60
15	MARYLAND	270326.24	271135.60	273159.00	267088.80	269112.20	267088.80
16	MICHIGAN	635347.60	647488.00	655581.60	655581.60	647488.00	631300.80
17	MINNESOTA	598926.40	651534.80	728424.00	724377.20	789126.00	785079.20
18	MISSISSIPPI	849828.00	857921.60	870062.00	882202.40	829594.00	821500.40
19	MISSOURI	3116036.00	3055334.00	3055334.00	2852994.00	2691122.00	2865134.40
20	MONTANA	1214.04	1618.72	2023.40	2023.40	2428.08	2428.08
21	NEBRASKA	2974398.00	2903579.00	2974398.00	2863111.00	2944047.00	2913696.00
22	NEW JERSEY	127878.88	120189.96	125046.12	114929.12	114929.12	114929.12
23	NEW MEXICO	18210.60	19829.32	21448.04	23876.12	22662.08	28327.60
24	NEW YORK	341954.60	319697.20	327790.80	307556.80	311603.60	307556.80
25	NORTH CAROLINA	1092636.00	1040027.60	1080495.60	1027887.20	1027887.20	1015746.80
26	NORTH DAKOTA	35611.84	55845.84	50989.68	55036.48	60702.00	60297.32
27	OHIO	1618720.00	1568135.00	1618720.00	1578252.00	1608603.00	1517550.00
28	OKLAHOMA	1064308.40	1116916.80	1278788.80	1250461.20	1456848.00	1598486.00

	Location	1900	1901	1902	1903	1904	1905
29	OREGON	9712.32	10926.36	11331.04	11735.72	11735.72	12140.40
30	PENNSYLVANIA	635347.60	629277.40	635347.60	617137.00	598926.40	598926.40
31	SOUTH CAROLINA	704143.20	667722.00	692002.80	663675.20	639394.40	647488.00
32	SOUTH DAKOTA	517990.40	586786.00	671768.80	687956.00	708190.00	740564.40
33	TENNESSEE	1311163.20	1343537.60	1428520.40	1392099.20	1384005.60	1367818.40
34	TEXAS	1954604.40	1938417.20	1974838.40	1974838.40	1954604.40	2015306.40
35	UTAH	4451.48	4451.48	4451.48	4451.48	4046.80	4046.80
36	VIRGINIA	768892.00	768892.00	776985.60	748658.00	748658.00	748658.00
37	WASHINGTON	6070.20	6474.88	6879.56	7284.24	7688.92	8902.96
38	WEST VIRGINIA	297439.80	291369.60	295416.40	279229.20	279229.20	275182.40
39	WISCONSIN	637371.00	659628.40	679862.40	667722.00	659628.40	633324.20
40	WYOMING	1618.72	2023.40	2428.08	2428.08	2428.08	3237.44

41 rows × 119 columns

In [ ]:
```
crop_df.dtypes
```

Out[ ]:
```
Location     object
1900        float64
1901        float64
1902        float64
1903        float64
             ...
2013        float64
2014        float64
2015        float64
2016        float64
2017        float64
Length: 119, dtype: object
```

Column with categorical features is the location column, so I will set up an experiment, with the target column as the year 2017

In [ ]:
```
from pycaret.regression import *
reg = setup(crop_df, target = '2017')
```

	Description	Value
0	Session id	6455
1	Target	2017
2	Target type	Regression
3	Original data shape	(41, 119)
4	Transformed data shape	(41, 119)
5	Transformed train set shape	(28, 119)
6	Transformed test set shape	(13, 119)
7	Numeric features	117
8	Categorical features	1
9	Preprocess	True
10	Imputation type	simple
11	Numeric imputation	mean
12	Categorical imputation	mode
13	Maximum one-hot encoding	25
14	Encoding method	None
15	Fold Generator	KFold
16	Fold Number	10
17	CPU Jobs	-1
18	Use GPU	False
19	Log Experiment	False
20	Experiment Name	reg-default-name
21	USI	53ae

In [ ]:
```
# # functional API
# best = compare_models()

# create linear regression model
lr = create_model('lr')
```

Initiated	. . . . .	03:27:26
Status	. . . . .	Loading Dependencies
Estimator	. . . . .	Compiling Library

Fold	MAE	MSE	RMSE	R2	RMSLE	MAPE
0	69288.1301	7496764785.9589	86583.8598	0.9866	0.0519	0.0470
1	79659.3930	14985162076.9725	122413.8966	0.9964	0.3072	0.2023
2	85631.1954	16927838972.3718	130107.0289	0.9971	0.1708	0.1543
3	14625.5167	401520387.8357	20037.9736	0.9352	0.1638	0.1671
4	89047.1346	12397346189.3150	111343.3707	0.9654	0.1271	0.0831
5	38744.3063	1887060747.3154	43440.3125	0.9937	0.2322	0.2109
6	19336.0038	575749784.5598	23994.7866	0.9983	0.4193	0.4455
7	10954.2918	293946476.5959	17144.8674	0.9788	0.6298	0.6664
8	167529.3850	39113040939.8778	197770.1720	0.8718	0.2520	0.2585
9	120168.8750	19494254046.0989	139621.8251	0.9942	0.5144	0.2835
Mean	69498.4232	11357268440.6902	89245.8093	0.9718	0.2869	0.2519
Std	47774.2686	11597943985.0874	58244.7762	0.0382	0.1735	0.1737

In [ ]:
```
# print(best)

# evaluate model performance
evaluate_model(lr)
```

```
interactive(children=(ToggleButtons(description='Plot Type:', icons=('',), options=
(('Pipeline Plot', 'pipelin…
```

In [ ]:
```
# evaluate_model(best)
# tune model hyperparameters
tuned_lr = tune_model(lr, optimize = 'r2')
```

Initiated	. . . . .	03:27:45
Status	. . . . .	Searching Hyperparameters
Estimator	. . . . .	Linear Regression

Fold	MAE	MSE	RMSE	R2	RMSLE	MAPE
0	68724.7770	7061309798.4133	84031.6000	0.9874	0.0493	0.0455
1	89274.5361	18005054208.6295	134182.9133	0.9957	0.8082	0.3329
2	77606.4363	13664812252.2493	116896.5879	0.9976	0.1222	0.1112
3	11586.0469	282441584.5676	16805.9985	0.9544	0.1269	0.1198
4	87979.3610	12285282220.3031	110838.9923	0.9658	0.1279	0.0831
5	38884.4926	2076736748.6901	45571.2272	0.9931	0.1584	0.1424
6	16107.5180	564747553.4156	23764.4178	0.9984	0.2271	0.1966
7	10903.3220	204078959.0070	14285.6207	0.9853	0.5510	0.5627
8	166074.7231	38933850675.2334	197316.6254	0.8724	0.2462	0.2514
9	131294.5448	23212813886.1193	152357.5200	0.9931	0.5944	0.3116
Mean	69843.5758	11629112788.6628	89605.1503	0.9743	0.3012	0.2157
Std	49516.2194	11911210423.1174	60000.2486	0.0367	0.2429	0.1473

```
Fitting 10 folds for each of 2 candidates, totalling 20 fits
```

In [ ]:
```
# predict_model(best)
# compare models
compare_models([lr, tuned_lr])

# # plot model performance
# plot_model(tuned_lr, plot = 'learning')
# plot_model(tuned_lr, plot = 'error')
```

Out[ ]:
```
LinearRegression
LinearRegression(fit_intercept=False, n_jobs=-1)
```

As you can see the newly tuned model does better than the previous in terms of Mean absolute error, Root Mean Squared Error and Root squared error but when it came down to Mean Squared Error the previous model performed better.

Predicting number of hectares for 2018

This show the last 5 years with 2018 added, and the predictions for 2018

In [ ]:
```
new_data = pd.read_csv('2018_Cropdata.csv',delimiter=',')
new_data
predictions = predict_model(tuned_lr, data=new_data)
predictions = predictions.rename(columns={'prediction_label': 'predicted_2018'})
predictions = predictions[['2015', '2016', '2017', '2018', 'predicted_2018']]  # Re
```

```
last_5_columns = predictions.iloc[:,-5:]
last_5_columns
```

	Model	MAE	MSE	RMSE	R2	RMSLE	MAPE
0	Linear Regression	15247.1748	1728054942.2512	41569.8802	0.9989	0.2025	0.0998

Out[ ]:

	2015	2016	2017	2018	predicted_2018
0	9.914660e+04	1.274742e+05	9.509980e+04	9.509980e+04	9.509980e+04
1	1.456848e+04	2.023400e+04	1.294976e+04	NaN	1.294976e+04
2	1.800826e+05	3.014866e+05	2.407846e+05	2.630420e+05	2.556849e+05
3	2.428080e+04	4.046800e+04	3.237440e+04	3.035100e+04	3.237440e+04
4	3.844460e+05	4.734756e+05	5.260840e+05	5.463180e+05	4.355032e+05
5	6.636752e+04	6.636752e+04	6.920028e+04	6.434412e+04	6.920028e+04
6	2.023400e+04	1.618720e+04	1.497316e+04	NaN	1.497316e+04
7	1.153338e+05	1.375912e+05	9.914660e+04	1.112870e+05	1.709077e+05
8	2.832760e+04	4.046800e+04	4.653820e+04	5.260840e+04	4.450071e+04
9	4.653820e+06	4.633586e+06	4.431246e+06	4.390778e+06	4.431246e+06
10	2.217646e+06	2.213600e+06	2.100289e+06	2.092196e+06	2.100289e+06
11	5.281074e+06	5.463180e+06	5.220372e+06	5.159670e+06	5.220372e+06
12	1.586346e+06	1.991026e+06	2.104336e+06	2.063868e+06	2.104336e+06
13	5.301308e+05	5.665520e+05	4.937096e+05	5.018032e+05	4.937096e+05
14	1.578252e+05	2.225740e+05	1.982932e+05	1.821060e+05	1.982932e+05
15	1.537784e+05	1.618720e+05	1.699656e+05	1.578252e+05	1.549252e+05
16	8.376876e+05	8.255472e+05	7.648452e+05	7.688920e+05	7.648452e+05
17	3.075568e+06	3.237440e+06	3.087708e+06	3.014866e+06	3.050662e+06
18	1.982932e+05	2.913696e+05	2.023400e+05	1.861528e+05	2.023400e+05
19	1.246414e+06	1.416380e+06	1.315210e+06	1.355678e+06	1.315210e+06
20	2.023400e+04	2.225740e+04	2.630420e+04	NaN	2.630420e+04
21	3.702822e+06	3.864694e+06	3.763524e+06	3.743290e+06	3.763524e+06
22	2.913696e+04	2.873228e+04	2.832760e+04	NaN	2.456799e+04
23	1.618720e+04	1.659188e+04	1.740124e+04	NaN	1.740124e+04
24	2.387612e+05	2.306676e+05	1.962698e+05	2.549484e+05	1.962698e+05
25	2.954164e+05	3.803992e+05	3.399312e+05	3.399312e+05	3.399312e+05
26	1.035981e+06	1.323304e+06	1.307116e+06	1.181666e+06	1.307116e+06
27	1.319257e+06	1.335444e+06	1.266648e+06	1.319257e+06	1.266648e+06
28	1.133104e+05	1.416380e+05	1.234274e+05	1.133104e+05	3.280613e+05
29	1.214040e+04	1.578252e+04	1.780592e+04	NaN	1.602628e+04

	2015	2016	2017	2018	predicted_2018
30	3.803992e+05	3.844460e+05	3.723056e+05	3.682588e+05	3.214976e+05
31	1.052168e+05	1.416380e+05	1.315210e+05	1.254508e+05	1.071585e+05
32	2.035540e+06	2.076008e+06	2.055774e+06	2.003166e+06	2.055774e+06
33	2.954164e+05	3.358844e+05	2.873228e+05	2.772058e+05	2.873228e+05
34	7.972196e+05	1.031934e+06	9.064832e+05	7.729388e+05	9.064832e+05
35	6.879560e+03	1.173572e+04	8.093600e+03	NaN	1.369661e+04
36	1.214040e+05	1.375912e+05	1.375912e+05	1.335444e+05	1.375912e+05
37	3.035100e+04	3.439780e+04	3.237440e+04	2.873228e+04	3.237440e+04
38	1.416380e+04	1.416380e+04	1.335444e+04	NaN	1.335444e+04
39	1.214040e+06	1.303070e+06	1.185712e+06	1.214040e+06	1.082891e+06
40	2.387612e+04	2.792292e+04	2.549484e+04	NaN	2.549484e+04

This shows the predicted columns where the original 2018 data was null

In [ ]:
```
null_2018_rows = last_5_columns[last_5_columns['2018'].isnull()]
print(null_2018_rows)
```

```
           2015          2016          2017  2018  predicted_2018
1   14568.480469  20234.000000  12949.759766   NaN    12949.759766
6   20234.000000  16187.200195  14973.160156   NaN    14973.160156
20  20234.000000  22257.400391  26304.199219   NaN    26304.199219
22  29136.960938  28732.279297  28327.599609   NaN    24567.985885
23  16187.200195  16591.880859  17401.240234   NaN    17401.240234
29  12140.400391  15782.519531  17805.919922   NaN    16026.279865
35   6879.560059  11735.719727   8093.600098   NaN    13696.609090
38  14163.799805  14163.799805  13354.440430   NaN    13354.440430
40  23876.119141  27922.919922  25494.839844   NaN    25494.839844
```

Saving the model

In [ ]:
```
save_model(tuned_lr, 'saved_models/pycaret_crop_model')
```

```
Transformation Pipeline and Model Successfully Saved
```

Out[ ]:
```
(Pipeline(memory=Memory(location=None),
          steps=[('numerical_imputer',
                  TransformerWrapper(include=['1900', '1901', '1902', '1903',
                                              '1904', '1905', '1906', '1907',
                                              '1908', '1909', '1910', '1911',
                                              '1912', '1913', '1914', '1915',
                                              '1916', '1917', '1918', '1919',
                                              '1920', '1921', '1922', '1923',
                                              '1924', '1925', '1926', '1927',
                                              '1928', '1929', ...],
                                     transformer=SimpleImputer())),
                 ('categorical_imputer',
                  TransformerWrapper(include=['Location'],
                                     transformer=SimpleImputer(strategy='most_freq
uent'))),
                 ('rest_encoding',
                  TransformerWrapper(include=['Location'],
                                     transformer=TargetEncoder(cols=['Location'],
                                                               handle_missing='ret
urn_nan'))),
                 ('trained_model',
                  LinearRegression(fit_intercept=False, n_jobs=-1))]),
 'saved_models/pycaret_crop_model.pkl')
```