



**Министерство науки и высшего образования Российской  
Федерации  
Федеральное государственное бюджетное образовательное  
учреждение  
высшего образования  
«Московский государственный технический университет  
имени Н.Э. Баумана  
(национальный исследовательский университет)»  
(МГТУ им. Н.Э. Баумана)**

**Факультет «Информатика и системы управления»  
Кафедра «Системы обработки информации и управления»**

Отчет по дополнительной лабораторной работе №1  
«Анализ и прогнозирование временного ряда»  
по дисциплине «Технологии машинного обучения»

Выполнил:  
студент группы ИУ5-64Б  
Стукалов И.Д.  
10.05.2024

Проверил:  
Гапанюк Ю.Е.

2024 г.

### Задание.

Выберите набор данных (датасет) для решения задачи прогнозирования временного ряда.

Визуализируйте временной ряд и его основные характеристики.

Разделите временной ряд на обучающую и тестовую выборку.

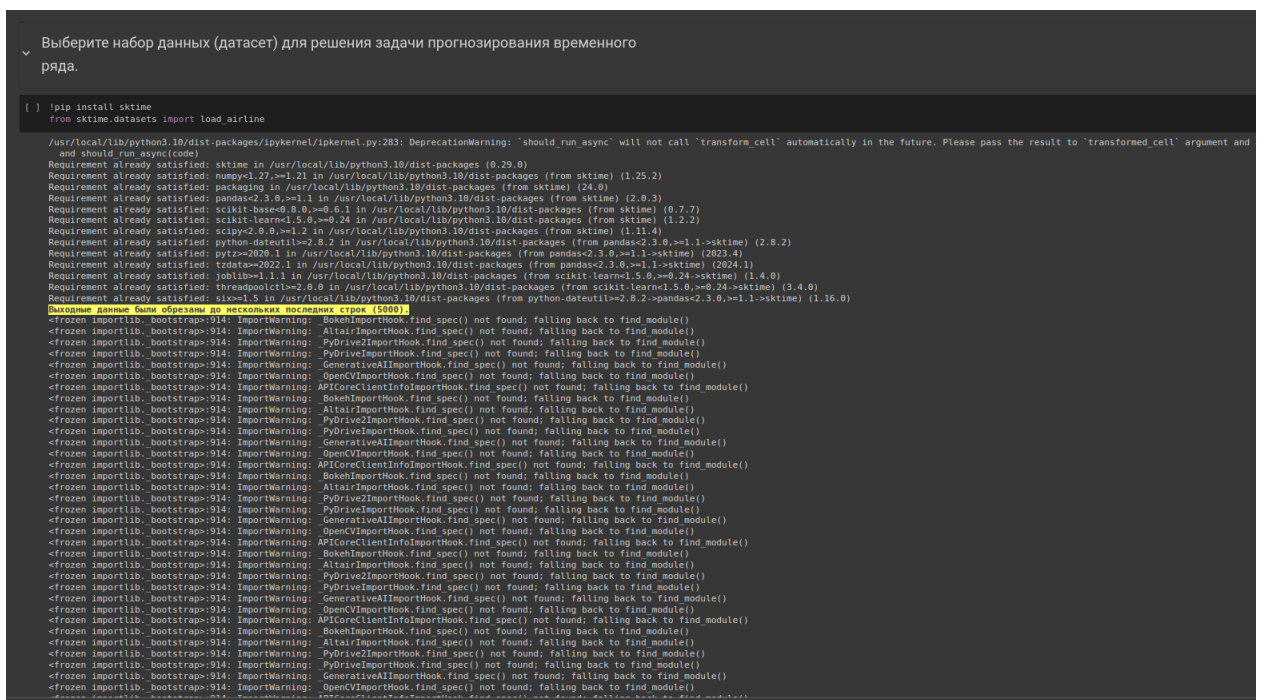
Произведите прогнозирование временного ряда с использованием следующих методов:

- один из авторегрессионных методов (ARMA, ARIMA, ...);
- метод символьной регрессии;
- двумя методами на выбор из семейства МГУА (один из линейных методов COMBI / MULTI + один из нелинейных методов MIA / RIA) с использованием библиотеки gmdh.

Визуализируйте тестовую выборку и каждый из прогнозов.

Оцените качество прогноза в каждом случае с помощью подходящей метрики.

## Текст программы



- Визуализируйте временной ряд и его основные характеристики.

```
<frozen importlib._bootstrap>:914: ImportWarning: APICoreClientInfoImportHook.find_spec() not found; falling back to find_module()
```

```
# Размер временного ряда
a1.shape[0]
```

- Разделите временной ряд на обучающую и тестовую выборку.

```
from sklearn.metrics import mean_squared_error
from statsmodels.tsa.arima.model import ARIMA
from statsmodels.tsa.holtwinters import ExponentialSmoothing

# Итоговая метрика качества прогноза
xnum = list(range(1, shape[0]))

# Разделение выборки на обучающую и тестовую
Y = al.values
train_size = int(len(Y) * 0.7)
xnum_train, xnum_test = xnum[0:train_size], xnum[train_size:]
train, test = Y[0:train_size], Y[train_size:]
history_arima = [x for x in train]
history_es = [x for x in train]
```

✓ Произведите прогнозирование временного ряда с использованием как минимум двух методов.

✓ ARIMA

```
# Параметры модели (p,d,q)
arima_order = (6,1,0)
# Формирование предсказаний
predictions_arima = list()
for t in range(len(test)):
    model_arima = ARIMA(history_arima, order=arima_order)
    model_arima_fit = model_arima.fit()
    yhat_arima = model_arima_fit.forecast()[0]
    predictions_arima.append(yhat_arima)
    history_arima.append(test[t])
# Вычисление метрики RMSE
error_arima = mean_squared_error(test, predictions_arima, squared=False)

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and should_run_async(code)

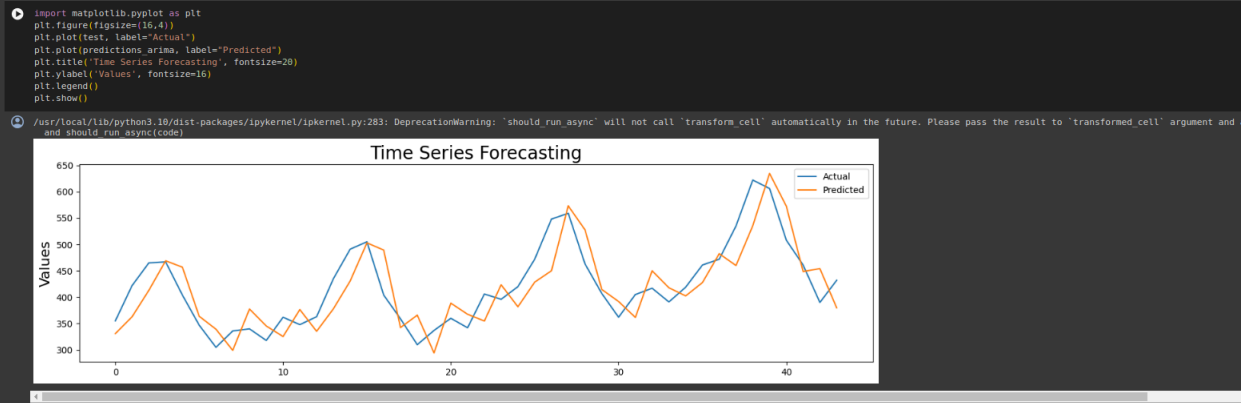
[ ] # Формирование предсказаний
predictions_es = list()
for t in range(len(test)):
    model_es = ExponentialSmoothing(history_es)
    model_es_fit = model_es.fit()
    yhat_es = model_es_fit.forecast()[0]
    predictions_es.append(yhat_es)
    history_es.append(test[t])
# Вычисление метрики RMSE
error_es = mean_squared_error(test, predictions_es, squared=False)

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and should_run_async(code)

[ ] # Выписка результатов
np.mean(V), error_arima, error_es

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and should_run_async(code)
(280.2988111111111, 45.28242088258871, 49.447888816444894)

[ ] import matplotlib.pyplot as plt
plt.figure(figsize=(16,4))
plt.plot(test, label="Actual")
plt.plot(predictions_arima, label="Predicted")
plt.title("Time Series Forecasting - ARIMA")
```



✓ Символьная регрессия

```
[ ] !pip install gplearn
from gplearn.genetic import SymbolicRegressor

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and should_run_async(code)
Requirement already satisfied: gplearn in /usr/local/lib/python3.10/dist-packages (0.4.2)
Requirement already satisfied: scikit-learn=1.0.2 in /usr/local/lib/python3.10/dist-packages (from gplearn) (1.2.2)
Requirement already satisfied: joblib=1.0.0 in /usr/local/lib/python3.10/dist-packages (from gplearn) (1.4.0)
Requirement already satisfied: numpy=1.17.3 in /usr/local/lib/python3.10/dist-packages (from scikit-learn=1.0.2->gplearn) (1.25.2)
Requirement already satisfied: scipy=1.3.2 in /usr/local/lib/python3.10/dist-packages (from scikit-learn=1.0.2->gplearn) (1.11.4)
Requirement already satisfied: threadpoolctl=2.0.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn=1.0.2->gplearn) (3.4.0)
<frozen importlib._bootstrap-914: ImportError: PyDriveImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap-914: ImportError: GenerativeAIImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap-914: ImportError: OpenVImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap-914: ImportError: APICoreClientInfoImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap-914: ImportError: BokaImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap-914: ImportError: AltairImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap-914: ImportError: DufrinaImportHook.find_spec() not found; falling back to find module()
```

```
[ ] function set = ['add', 'sub', 'mul', 'div', 'sin']
est_gp = SymbolicRegressor(population_size=500, metric='mse',
                           generations=70, stopping_criteria=0.01,
                           init_depth=(4, 10), verbose=1, function_set=function_set,
                           const_range=(-100, 100), random_state=0)

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and
and should_run_async(code)

[ ] est_gp.fit(np.array(xnum_train).reshape(-1, 1), train.reshape(-1, 1))

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and
and should_run_async(code)
/usr/local/lib/python3.10/dist-packages/sklearn/utils/validation.py:1143: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples, 1), for example using
y = column_or_1d(y, warn=True)
| Population Average | Best Individual |
-----|-----|
Gen Length Fitness Length Fitness OOB Fitness Time Left
0 263.65 8.79976e+46 59 4935.14 N/A 4.49m
1 141.10 4.16457e+14 72 4638.92 N/A 1.17m
2 73.88 2.25505e+09 78 4579.99 N/A 36.57s
3 64.76 1.21852e+15 71 3864.76 N/A 52.18s
4 69.55 1.25496e+11 85 3715.16 N/A 1.21m
5 68.89 4.26607e+10 81 3012.4 N/A 1.12m
6 78.37 1.49839e+11 81 3012.4 N/A 1.06m
7 79.09 8.26805e+13 92 3009.99 N/A 1.16m
8 92.35 5.15147e+10 129 2754 N/A 1.30m
9 97.63 7.93413e+09 160 2702.01 N/A 1.30m
10 116.63 3.03425e+11 128 1507.72 N/A 1.50m
11 141.08 2.65532e+10 120 1507.69 N/A 3.30m
12 146.34 4.87157e+10 112 1251.72 N/A 4.09m
13 130.19 3.9615e+09 92 1251.68 N/A 2.27m
14 124.99 6.31431e+09 94 1121.53 N/A 1.31m
15 112.53 1.96229e+12 76 1101.53 N/A 1.14m
16 106.53 9.16678e+14 93 1098.81 N/A 1.15m
17 95.16 3.36262e+10 122 1091.06 N/A 1.32m
18 92.79 3.91346e+10 145 1090.23 N/A 2.41m
19 108.63 5.17089e+10 173 1089.28 N/A 1.67m
20 129.91 9.54207e+13 151 1070.74 N/A 1.25m
21 140.75 2.99526e+10 147 1065.74 N/A 2.00m
22 157.96 1.114e+10 157 1021.75 N/A 1.02m
23 150.68 1.77331e+10 267 1017.98 N/A 34.72s
24 162.40 2.36847e+10 129 1007.37 N/A 36.25s
25 183.90 4.04029e+10 171 977.83 N/A 37.23s
26 204.07 3.19638e+10 171 977.829 N/A 37.54s
27 234.07 1.37035e+10 151 966.887 N/A 1.07m
28 188.93 5.05123e+12 175 966.461 N/A 56.66s
29 192.81 7.17435e+09 177 965.649 N/A 39.19s
30 183.40 7.64131e+12 327 937.59 N/A 34.12s
31 194.48 4.10029e+09 309 905.978 N/A 32.65s
32 240.94 3.8342e+13 335 905.978 N/A 37.54s
33 307.25 2.18532e+10 175 879.024 N/A 41.64s
34 203.38 1.07338e+10 175 879.024 N/A 37.21s
35 246.97 3.58219e+10 181 876.484 N/A 33.98s
36 206.65 3.26092e+10 163 870.93 N/A 29.56s
37 188.63 2.7412e+10 163 870.93 N/A 26.90s
38 183.49 4.22489e+13 171 870.757 N/A 26.06s
39 184.69 5.11402e+10 161 869.956 N/A 33.41s
```

```
<frozen importlib._bootstrap:914: ImportWarning: OpenCVImportHook.find_spec() not found; falling back to find_module()
<frozen importlib._bootstrap:914: ImportWarning: APICoreClientInfoImportHook.find_spec() not found; falling back to find_module()
<frozen importlib._bootstrap:914: ImportWarning: BokehImportHook.find_spec() not found; falling back to find_module()
<frozen importlib._bootstrap:914: ImportWarning: AltairImportHook.find_spec() not found; falling back to find_module()

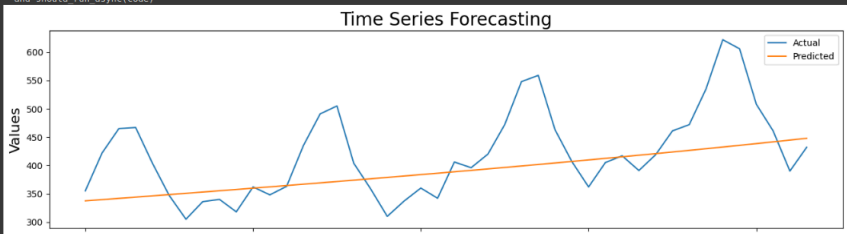
div(add(add(mul(sub(26.428, 26.637), mul(-91.195, 51.329)), div(mul(-70.060, sub(div(X0, 20.271), div(mul(div(mul(div(X0, -92.726), add(X0, add(X0, div(-64.562, -73.213)))), sub(X0, X0)), add(div(mul(mul(sub(26.428, 26.637),

[ ] # fpackcasnew
y_gp = est_gp.predict(np.array(xnum_test).reshape(-1, 1))
y_gp

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and
and should_run_async(code)
array([[337.4146303, 339.60335004, 341.80088045, 344.00917373,
       346.2293616, 348.4621579, 350.70806389, 352.96747737,
       355.24019207, 357.5282224, 359.83026919, 362.14723222,
       364.47951942, 366.82755023, 369.19177705, 371.57268081,
       373.97077202, 376.38659062, 378.82070573, 381.27371534,
       383.74024615, 386.23095342, 388.75252092, 391.20766102,
       393.84511483, 396.42565251, 399.03007368, 401.65920788,
       404.31391519, 406.99508693, 409.70364651, 412.44055026,
       415.20670851, 418.00338669, 420.83140651, 423.69394734,
       426.58614759, 429.51518626, 432.48028462, 435.48270796,
       438.52376748, 441.60482236, 444.72728185, 447.89260766])

plt.figure(figsize=(16,4))
plt.plot(test, label='Actual')
plt.plot(y_gp, label='Predicted')
plt.title('Time Series Forecasting', fontsize=20)
plt.ylabel('Values', fontsize=16)
plt.legend()
plt.show()

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and
and should_run_async(code)
```



```
4

▼ МГУА

[ ] | pip install gmdh
import gmdh

Requirement already satisfied: gmdh in /usr/local/lib/python3.10/dist-packages (1.0.3)
Requirement already satisfied: docstring-inheritance in /usr/local/lib/python3.10/dist-packages (from gmdh) (2.2.0)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from gmdh) (1.25.2)

1 | pip install sktime
from sklearn.metrics import mean_squared_error
from statsmodels.tsa.arima.model import ARIMA
from statsmodels.tsa.holtwinters import ExponentialSmoothing
from sktime.datasets import load_airline
# Зарядим исходные данные
al = load_airline()
al.info()

# Исчисленная метрика ошибки времени
xnum = list(range(al.shape[0]))
# Разделение выборки на обучающую и тестовую
Y = al.values
train_size = int(len(Y) * 0.7)
xnum_train, xnum_test = xnum[0:train_size], xnum[train_size:]
train, test = Y[0:train_size], Y[train_size:]
history_arima = [x for x in train]
history_es = [x for x in train]

2 | /usr/local/lib/python3.10/dist-packages/IPykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and
and should_run_async(code)
Requirement already satisfied: sktime in /usr/local/lib/python3.10/dist-packages (0.29.0)
Requirement already satisfied: numpy<1.27,>=1.21 in /usr/local/lib/python3.10/dist-packages (from sktime) (1.25.2)
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (from sktime) (24.0)
Requirement already satisfied: pandas<2.3.0,>=1.1 in /usr/local/lib/python3.10/dist-packages (from sktime) (2.0.3)
Requirement already satisfied: scikit-base<0.8.0,>=0.6.1 in /usr/local/lib/python3.10/dist-packages (from sktime) (0.7.7)
Requirement already satisfied: scikit-learn<1.5.0,>=0.24 in /usr/local/lib/python3.10/dist-packages (from sktime) (1.2.2)
Requirement already satisfied: scipy<2.0.0,>=1.2 in /usr/local/lib/python3.10/dist-packages (from sktime) (1.11.4)
Requirement already satisfied: python-dateutil<=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas<2.3.0,>=1.1->sktime) (2.8.2)
Requirement already satisfied: pytz<2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas<2.3.0,>=1.1->sktime) (2023.4)
Requirement already satisfied: tzdata<=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas<2.3.0,>=1.1->sktime) (2024.1)
Requirement already satisfied: joblib<1.1.1 in /usr/local/lib/python3.10/dist-packages (from scikit-learn<1.5.0,>=0.24->sktime) (1.4.0)
Requirement already satisfied: threadpoolctl<=2.0.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn<1.5.0,>=0.24->sktime) (3.4.0)
Requirement already satisfied: size<1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil<=2.8.2->pandas<2.3.0,>=1.1->sktime) (1.16.0)
Выходные данные были обрезаны до нескольких последних строк (5000).
<frozen importlib._bootstrap>:914: ImportWarning: BockelImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap>:914: ImportWarning: AltairImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap>:914: ImportWarning: PyDrive2ImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap>:914: ImportWarning: PyDriveImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap>:914: ImportWarning: GenerativeAIImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap>:914: ImportWarning: OpenAIImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap>:914: ImportWarning: APICoreClientInfoImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap>:914: ImportWarning: BockelImportHook.find_spec() not found; falling back to find module()
<frozen importlib._bootstrap>:914: ImportWarning: AltairImportHook.find_spec() not found; falling back to find module()
```

```
4

[ ] from sklearn.metrics import mean_absolute_error, mean_squared_error, mean_squared_log_error, median_absolute_error, r2_score

def print_metrics(y_test, y_pred, squared=False):
    print(f"R^2: {r2_score(y_test, y_pred)}")
    crit_name = "MSE" if squared else "RMSE"
    print(f"({crit_name}: {mean_squared_error(y_test, y_pred, squared=squared)})")
    print(f"MAE: {mean_absolute_error(y_test, y_pred)}")

/usr/local/lib/python3.10/dist-packages/IPykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and
and should_run_async(code)

4

▼ COMBI

1 | # xnum_train, xnum_test = xnum[0:train_size], xnum[train_size:]
# train, test = Y[0:train_size], Y[train_size:]
combi_model = gmdh.Combi()
combi_model.fit(xnum_train, train, verbose=1, n_jobs=-1, test_size=0.24, limit=0,
               criterion=gmdh.Criterion(gmdh.CriterionType.REGULARITY))

print()
print(combi_model.get_best_polynomial())
print()
y_pred_combi = combi_model.predict(xnum_test)

print_metrics(test, y_pred_combi)

2 | /usr/local/lib/python3.10/dist-packages/IPykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and
and should_run_async(code)
LEVEL 1 |=====| 100% [00m:00s] (1 combinations) error=69398.16782

y = 1.9626*x1 + 112.9815

R^2: -0.620705489345804
RMSE: 97.3866459642962
MAE: 73.58676256757596

4

[ ] import numpy as np
import pandas as pd
from matplotlib import pyplot
import matplotlib.pyplot as plt

plt.figure(figsize=(16,4))
plt.plot(test, label="Actual")
plt.plot(y_pred_combi, label="Predicted")
plt.title('Time Series Forecasting', fontsize=20)
plt.ylabel('Values', fontsize=16)
plt.legend()
plt.show()

/usr/local/lib/python3.10/dist-packages/IPykernel/ipkernel.py:283: DeprecationWarning: 'should_run_async' will not call 'transform_cell' automatically in the future. Please pass the result to 'transformed_cell' argument and
and should_run_async(code)
<frozen importlib._bootstrap>:914: ImportWarning: PyDrive2ImportHook.find_spec() not found; falling back to find module()
```

