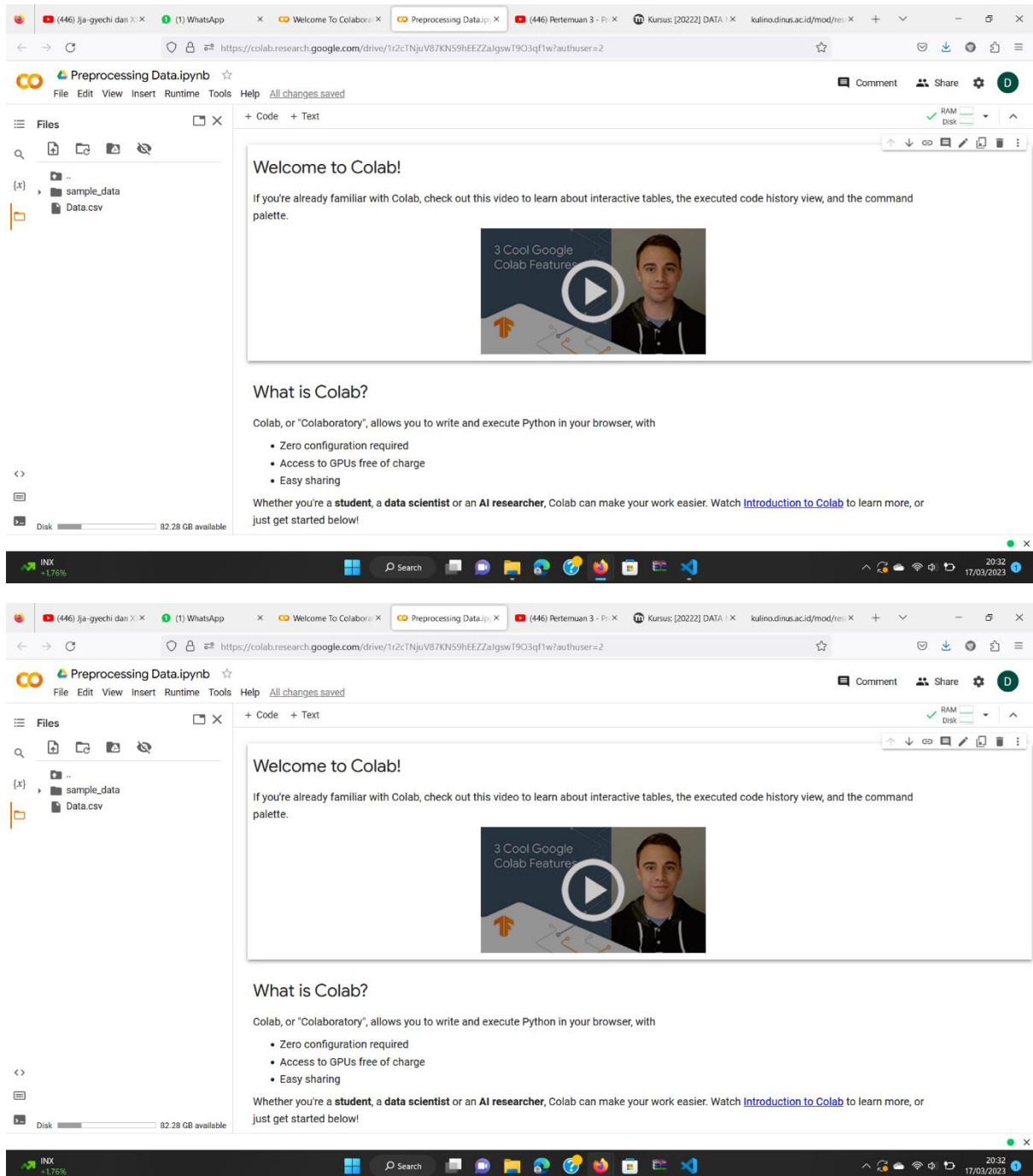


NAMA : DINA ETIKASARI

NIM : A11.2020.13191

KELOMPOK : A11.4616



Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help Saving...

Files

- sample_data
- Data.csv

Welcome to Colab!

If you're already familiar with Colab, check out this video to learn about interactive tables, the executed code history view, and the command palette.

3 Cool Google Colab Features

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

What is Colab?

Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge

Disk 82.28 GB available

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help Saving...

Files

- sample_data
- Data.csv

Welcome to Colab!

If you're already familiar with Colab, check out this video to learn about interactive tables, the executed code history view, and the command palette.

3 Cool Google Colab Features

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

What is Colab?

Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge

OneDrive

Screenshot saved

The screenshot was added to your OneDrive.

completed at 8:37 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data

Data.csv

RAM

Disk

Welcome to Colab!

If you're already familiar with Colab, check out this video to learn about interactive tables, the executed code history view, and the command palette.

3 Cool Google Colab Features

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

Numpy merupakan library python digunakan untuk komputasi matriks.

Matplotlib merupakan library python untuk presentasi data berupa grafik atau plot.

Numpy merupakan library python digunakan untuk komputasi matriks.

Matplotlib merupakan library python untuk presentasi data berupa grafik atau plot.

completed at 8:37 PM

27°C

Hujan ringan

20:40

17/03/2023

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

Welcome to Colab!

If you're already familiar with Colab, check out this video to learn about interactive tables, the executed code history view, and the command palette.

3 Cool Google Colab Features

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

Numpy merupakan library python digunakan untuk komputasi matriks. Matplotlib merupakan library python untuk presentasi data berupa grafik atau plot.

Disk 82.28 GB available

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help

Files

- sample_data
- Data.csv

Welcome to Colab!

If you're already familiar with Colab, check out this video to learn about interactive tables, the executed code history view, and the command palette.

3 Cool Google Colab Features

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

Numpy merupakan library python digunakan untuk komputasi matriks. Matplotlib merupakan library python untuk presentasi data berupa grafik atau plot.

```
dataset = pd.read_csv('Data.csv')
```

Disk 82.28 GB available

Preprocessing Data.ipynb


File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

Welcome to Colab!

If you're already familiar with Colab, check out this video to learn about interactive tables, the executed code history view, and the command palette.



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

Numpy merupakan library python digunakan untuk komputasi matriks. Matplotlib merupakan library python untuk presentasi data berupa grafik atau plot.

```
dataset = pd.read_csv('Data.csv')
x = dataset.iloc[:, :-1].values
y = dataset.iloc[:, -1].values
```

0s completed at 10:38 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

atau plot.

```
[4] dataset = pd.read_csv('Data.csv')
x = dataset.iloc[:, :-1].values
y = dataset.iloc[:, -1].values
```

```
print(x)
```

```
[['France' 44.0 72000.0]
['Spain' 27.0 48000.0]
['Germany' 30.0 54000.0]
['Spain' 38.0 61000.0]
['Germany' 40.0 nan]
['France' 35.0 58000.0]
['Spain' nan 52000.0]
['France' 48.0 79000.0]
['Germany' 50.0 83000.0]
['France' 37.0 67000.0]]
```

What is Colab?

Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge
- Easy sharing

0s completed at 10:44 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help Saving...

Files

- sample_data
- Data.csv

```
[4] dataset = pd.read_csv('Data.csv')
x = dataset.iloc[:, 1:1].values
y = dataset.iloc[:, -1].values

[5] print(x)

[['France' 44.0 72000.0]
 ['Spain' 27.0 48000.0]
 ['Germany' 30.0 54000.0]
 ['Spain' 38.0 61000.0]
 ['Germany' 40.0 nan]
 ['France' 35.0 58000.0]
 ['Spain' nan 52000.0]
 ['France' 48.0 79000.0]
 ['Germany' 50.0 83000.0]
 ['France' 37.0 67000.0]]

print(y)

['No' 'Yes' 'No' 'No' 'Yes' 'Yes' 'No' 'Yes' 'No' 'Yes']
```

What is Colab?

Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

0s completed at 10:46 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

```
[5] [['France' 44.0 72000.0]
 ['Spain' 27.0 48000.0]
 ['Germany' 30.0 54000.0]
 ['Spain' 38.0 61000.0]
 ['Germany' 40.0 nan]
 ['France' 35.0 58000.0]
 ['Spain' nan 52000.0]
 ['France' 48.0 79000.0]
 ['Germany' 50.0 83000.0]
 ['France' 37.0 67000.0]]

[6] print(y)

['No' 'Yes' 'No' 'No' 'Yes' 'Yes' 'No' 'Yes' 'No' 'Yes']

from sklearn.impute import SimpleImputer
imputer = SimpleImputer(missing_values=np.nan, strategy='mean')
imputer.fit(x[:, 1:3])
x[:, 1:3] = imputer.transform(x[:, 1:3])
```

What is Colab?

Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge
- Easy sharing

Whether you're a **student**, a **data scientist** or an **AI researcher**, Colab can make your work easier. Watch [Introduction to Colab](#) to learn more, or

0s completed at 10:53 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

```
[6] print(y)

['No' 'Yes' 'No' 'No' 'Yes' 'Yes' 'No' 'Yes' 'No' 'Yes']

[7] from sklearn.impute import SimpleImputer
imputer = SimpleImputer(missing_values=np.nan, strategy='mean')
imputer.fit(x[:, 1:3])
x[:, 1:3] = imputer.transform(x[:, 1:3])

[8] print(x)

[['France' 44.0 72000.0]
 ['Spain' 27.0 48000.0]
 ['Germany' 30.0 54000.0]
 ['Spain' 38.0 61000.0]
 ['Germany' 40.0 63777.77777777778]
 ['France' 35.0 58000.0]
 ['Spain' 38.77777777777778 52000.0]
 ['France' 48.0 79000.0]
 ['Germany' 50.0 83000.0]
 ['France' 37.0 67000.0]]
```

What is Colab?

Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge

0s completed at 10:59 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help

Files

- sample_data
- Data.csv

```
[8] [['France' 44.0 72000.0]
 ['Spain' 27.0 48000.0]
 ['Germany' 30.0 54000.0]
 ['Spain' 38.0 61000.0]
 ['Germany' 40.0 63777.77777777778]
 ['France' 35.0 58000.0]
 ['Spain' 38.77777777777778 52000.0]
 ['France' 48.0 79000.0]
 ['Germany' 50.0 83000.0]
 ['France' 37.0 67000.0]]

[9] from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OneHotEncoder
ct = ColumnTransformer(transformers=[('encoder', OneHotEncoder(), [0])], remainder='passthrough')
x = np.array(ct.fit_transform(x))
```

What is Colab?

Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge
- Easy sharing

Whether you're a **student**, a **data scientist** or an **AI researcher**, Colab can make your work easier. Watch [Introduction to Colab](#) to learn more, or just get started below!

0s completed at 11:13 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help

Files

- sample_data
- Data.csv

```
[10]: from sklearn.compose import ColumnTransformer
      from sklearn.preprocessing import OneHotEncoder
      ct = ColumnTransformer(transformers=[('endoder', OneHotEncoder(), [0])], remainder='passthrough')
      x = np.array(ct.fit_transform(x))

[11]: print(x)

[[1.0 0.0 0.0 44.0 72000.0]
 [0.0 0.0 1.0 27.0 48000.0]
 [0.0 1.0 0.0 30.0 54000.0]
 [0.0 0.0 1.0 38.0 61000.0]
 [0.0 1.0 0.0 40.0 63777.77777777778]
 [1.0 0.0 0.0 35.0 58000.0]
 [0.0 0.0 1.0 38.77777777777778 52000.0]
 [1.0 0.0 0.0 48.0 79000.0]
 [0.0 1.0 0.0 50.0 83000.0]
 [1.0 0.0 0.0 37.0 67000.0]]
```

What is Colab?

Colab, or 'Colaboratory', allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge
- Easy sharing

Whether you're a **student**, a **data scientist** or an **AI researcher**, Colab can make your work easier. Watch [Introduction to Colab](#) to learn more, or just get started below!

0s completed at 11:14 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help

Files

- sample_data
- Data.csv

```
[0.0 0.0 1.0 27.0 48000.0]
[0.0 1.0 0.0 30.0 54000.0]
[0.0 0.0 1.0 38.0 61000.0]
[0.0 1.0 0.0 40.0 63777.77777777778]
[1.0 0.0 0.0 35.0 58000.0]
[0.0 0.0 1.0 38.77777777777778 52000.0]
[1.0 0.0 0.0 48.0 79000.0]
[0.0 1.0 0.0 50.0 83000.0]
[1.0 0.0 0.0 37.0 67000.0]]

[ ] seconds_in_a_day = 24 * 60 * 60
seconds_in_a_day

86400

from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
y = le.fit_transform(y)

[ ] seconds_in_a_week = 7 * seconds_in_a_day
seconds_in_a_week

604800
```

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:

0s completed at 11:16 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

```
[11] [0.0 0.0 1.0 27.0 48000.0]
      [0.0 1.0 0.0 30.0 54000.0]
      [0.0 0.0 1.0 38.0 61000.0]
      [0.0 1.0 0.0 40.0 63777.77777777778]
      [1.0 0.0 0.0 35.0 58000.0]
      [0.0 0.0 1.0 38.77777777777778 52000.0]
      [1.0 0.0 0.0 48.0 79000.0]
      [0.0 1.0 0.0 50.0 83000.0]
      [1.0 0.0 0.0 37.0 67000.0]]

[12] from sklearn.preprocessing import LabelEncoder
      le = LabelEncoder()
      y = le.fit_transform(y)

[13] print(y)

[0 1 0 0 1 1 0 1 0 1]

[ ] seconds_in_a_day = 24 * 60 * 60
      seconds_in_a_day

86400

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:

[ ] seconds_in_a_week = 7 * seconds_in_a_day
```

0s completed at 11:18 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

```
[11] [0.0 0.0 1.0 27.0 48000.0]
      [0.0 1.0 0.0 30.0 54000.0]
      [0.0 0.0 1.0 38.0 61000.0]
      [0.0 1.0 0.0 40.0 63777.77777777778]
      [1.0 0.0 0.0 35.0 58000.0]
      [0.0 0.0 1.0 38.77777777777778 52000.0]
      [1.0 0.0 0.0 48.0 79000.0]
      [0.0 1.0 0.0 50.0 83000.0]
      [1.0 0.0 0.0 37.0 67000.0]]

[12] from sklearn.preprocessing import LabelEncoder
      le = LabelEncoder()
      y = le.fit_transform(y)

[13] print(y)

[0 1 0 0 1 1 0 1 0 1]

[14] print(x)

[[1.0 0.0 0.0 44.0 72000.0]
 [0.0 0.0 1.0 27.0 48000.0]
 [0.0 1.0 0.0 30.0 54000.0]
 [0.0 0.0 1.0 38.0 61000.0]
 [0.0 1.0 0.0 40.0 63777.77777777778]
 [1.0 0.0 0.0 35.0 58000.0]
 [0.0 0.0 1.0 38.77777777777778 52000.0]
 [1.0 0.0 0.0 48.0 79000.0]
 [0.0 1.0 0.0 50.0 83000.0]
 [1.0 0.0 0.0 37.0 67000.0]]

[ ] seconds_in_a_day = 24 * 60 * 60
      seconds_in_a_day

86400

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:

[ ] seconds_in_a_week = 7 * seconds_in_a_day
```

0s completed at 11:19 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

```
[13] print(y)

[0 1 0 0 1 1 0 1 0 1]

[14] print(x)

[[1.0 0.0 0.0 44.0 72000.0]
 [0.0 0.0 1.0 27.0 48000.0]
 [0.0 1.0 0.0 38.0 54000.0]
 [0.0 0.0 1.0 38.0 61000.0]
 [0.0 1.0 0.0 48.0 63777.77777777778]
 [1.0 0.0 0.0 35.0 58000.0]
 [0.0 0.0 1.0 38.77777777777778 52000.0]
 [1.0 0.0 0.0 48.0 79000.0]
 [0.0 1.0 0.0 50.0 83000.0]
 [1.0 0.0 0.0 37.0 67000.0]]

from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=1)

[ ] seconds_in_a_day = 24 * 60 * 60
seconds_in_a_day

86400
```

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:

0s completed at 11:24 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help

Files

- sample_data
- Data.csv

```
[14] [0.0 1.0 0.0 40.0 63777.77777777778]
      [1.0 0.0 0.0 35.0 58000.0]
      [0.0 0.0 1.0 38.77777777777778 52000.0]
      [1.0 0.0 0.0 48.0 79000.0]
      [0.0 1.0 0.0 50.0 83000.0]
      [1.0 0.0 0.0 37.0 67000.0]]

[18] from sklearn.model_selection import train_test_split
      x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=1)

[ ] print(x_train)

[[0.0 0.0 1.0 38.77777777777778 52000.0]
 [0.0 1.0 0.0 48.0 63777.77777777778]
 [1.0 0.0 0.0 44.0 72000.0]
 [0.0 0.0 1.0 38.0 61000.0]
 [0.0 0.0 1.0 27.0 48000.0]
 [1.0 0.0 0.0 48.0 79000.0]
 [0.0 1.0 0.0 50.0 83000.0]
 [1.0 0.0 0.0 35.0 58000.0]]

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:



```
[] seconds_in_a_week = 7 * seconds_in_a_day
seconds_in_a_week

604800
```



0s completed at 11:29 PM


```

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help

Files

- sample_data
- Data.csv

```
[14]: [[0.0 1.0 0.0 40.0 63777.777777777778]
       [1.0 0.0 0.0 35.0 58000.0]
       [0.0 0.0 1.0 38.77777777777778 52000.0]
       [1.0 0.0 0.0 48.0 79000.0]
       [0.0 1.0 0.0 50.0 83000.0]
       [1.0 0.0 0.0 37.0 67000.0]]

[18]: from sklearn.model_selection import train_test_split
       x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=1)

[19]: print(x_train)
[[0.0 0.0 1.0 38.77777777777778 52000.0]
 [0.0 1.0 0.0 40.0 63777.77777777778]
 [1.0 0.0 0.0 44.0 72000.0]
 [0.0 0.0 1.0 38.0 61000.0]
 [0.0 0.0 1.0 27.0 48000.0]
 [1.0 0.0 0.0 48.0 79000.0]
 [0.0 1.0 0.0 50.0 83000.0]
 [1.0 0.0 0.0 35.0 58000.0]]

[20]: print(x_test)
[[0.0 1.0 0.0 30.0 54000.0]
 [1.0 0.0 0.0 37.0 67000.0]]
```

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:

0s completed at 11:30 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help Save failed

Files

- sample_data
- Data.csv

```
[18]: from sklearn.model_selection import train_test_split
       x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=1)

[19]: print(x_train)
[[0.0 0.0 1.0 38.77777777777778 52000.0]
 [0.0 1.0 0.0 40.0 63777.77777777778]
 [1.0 0.0 0.0 44.0 72000.0]
 [0.0 0.0 1.0 38.0 61000.0]
 [0.0 0.0 1.0 27.0 48000.0]
 [1.0 0.0 0.0 48.0 79000.0]
 [0.0 1.0 0.0 50.0 83000.0]
 [1.0 0.0 0.0 35.0 58000.0]]

[20]: print(x_test)
[[0.0 1.0 0.0 30.0 54000.0]
 [1.0 0.0 0.0 37.0 67000.0]]

[21]: print(y_train)
[0 1 0 0 1 1 0 1]

[22]: print(y_test)
[0 1]
```

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

0s completed at 11:32 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

```
[19] print(x_train)

[[0.0 0.0 1.0 38.77777777777778 52000.0]
 [0.0 1.0 0.0 40.0 63777.77777777778]
 [1.0 0.0 0.0 44.0 72000.0]
 [0.0 0.0 1.0 38.0 61000.0]
 [0.0 0.0 1.0 27.0 48000.0]
 [1.0 0.0 0.0 48.0 79000.0]
 [0.0 1.0 0.0 50.0 83000.0]
 [1.0 0.0 0.0 35.0 58000.0]]

[20] print(x_test)

[[0.0 1.0 0.0 30.0 54000.0]
 [1.0 0.0 0.0 37.0 67000.0]]

[21] print(y_train)

[0 1 0 0 1 1 0 1]

[22] print(y_test)

[0 1]

from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train[:, 3:] = sc.fit_transform(x_train[:, 3:])
x_test[:, 3:] = sc.transform(x_test[:, 3:])
```

0s completed at 11:37 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

```
[21] print(y_train)

[0 1 0 0 1 1 0 1]

[22] print(y_test)

[0 1]

[23] from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train[:, 3:] = sc.fit_transform(x_train[:, 3:])
x_test[:, 3:] = sc.transform(x_test[:, 3:])

print(x_train)

[[0.0 0.0 1.0 -0.19159184384578545 -1.0781259408412425]
 [0.0 1.0 0.0 -0.014117293757057777 -0.07013167641635372]
 [1.0 0.0 0.0 0.566708506533324 0.633562432710451]
 [0.0 0.0 1.0 -0.30453019390224867 -0.30786617274297867]
 [0.0 0.0 1.0 -1.9018011447007988 -1.420463615551582]
 [1.0 0.0 0.0 1.1475343068237058 1.232653363453549]
 [0.0 1.0 0.0 1.4379472069688968 1.5749910381638885]
 [1.0 0.0 0.0 -0.7401495441200351 -0.5646194287757332]]
```

To execute the code in the above cell, select it with a click and then either press the play button to the left of the code, or use the keyboard shortcut "Command/Ctrl+Enter". To edit the code, just click the cell and start editing.

Variables that you define in one cell can later be used in other cells:

0s completed at 11:38 PM

Preprocessing Data.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- Data.csv

```
[22] print(y_test)

[0 1]

[23] from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train[:, 3:] = sc.fit_transform(x_train[:, 3:])
x_test[:, 3:] = sc.transform(x_test[:, 3:])

[24] print(x_train)

[[0.0 0.0 1.0 -0.19159184384578545 -1.0781259408412425]
 [0.0 1.0 0.0 -0.014117293757057777 -0.07013167641635372]
 [1.0 0.0 0.0 0.566708506533324 0.633562432710455]
 [0.0 0.0 1.0 -0.30453019390224867 -0.30786617274297867]
 [0.0 0.0 1.0 -1.9018011447007980 -1.420463615551582]
 [1.0 0.0 0.0 1.1475343068237058 1.232653363453549]
 [0.0 1.0 0.0 1.4379472069688968 1.5749910381638885]
 [1.0 0.0 0.0 -0.7401495441200351 -0.5646194287757332]]

print(x_test)

[[0.0 1.0 0.0 -1.4661817944830124 -0.9069571034860727]
 [1.0 0.0 0.0 -0.44973664397484414 0.2056403393225306]]

[ ] seconds_in_a_week = 7 * seconds_in_a_day
seconds_in_a_week
```

0s completed at 11:40 PM

25°C Kabut 23:40 17/03/2023