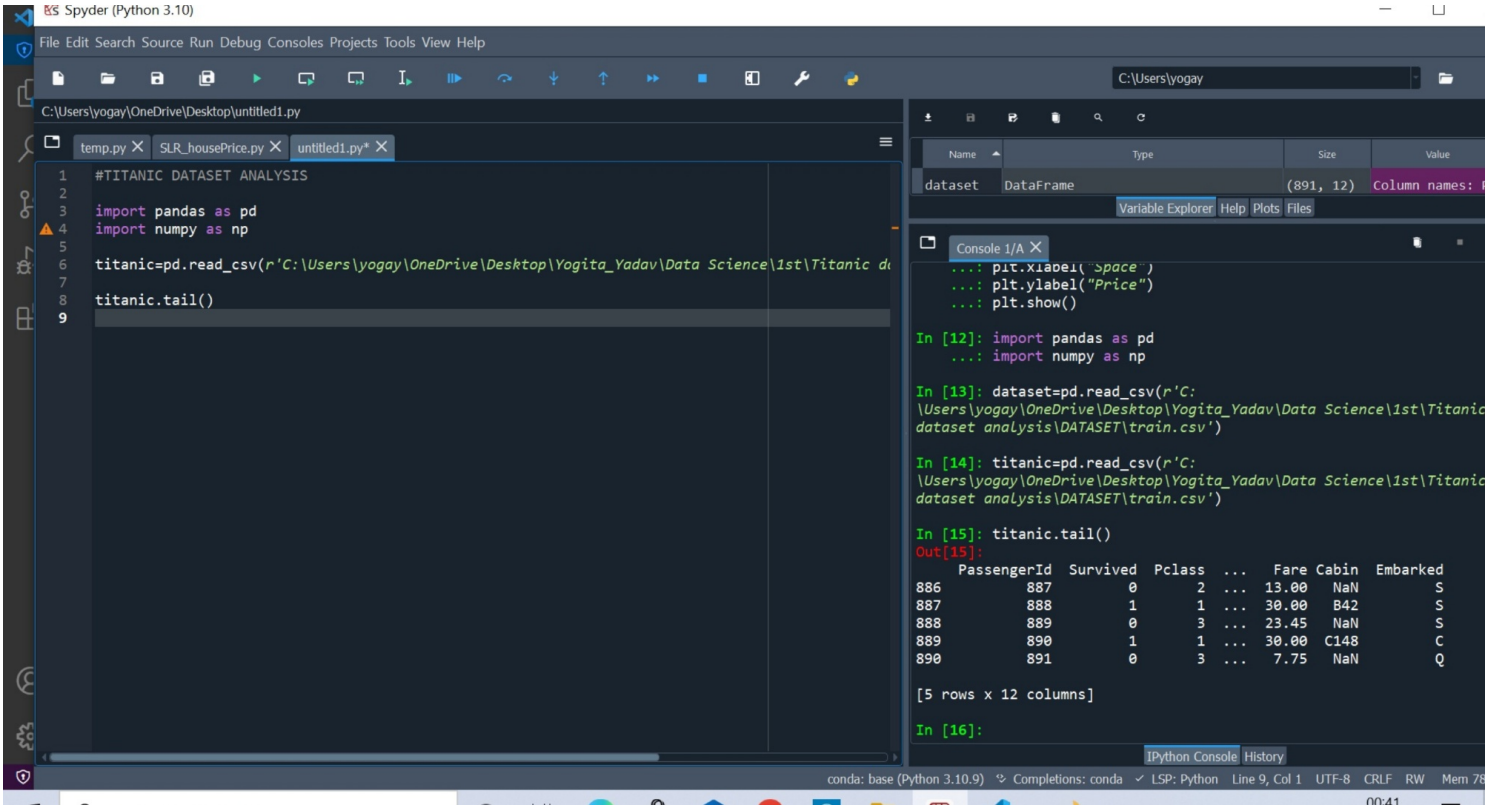
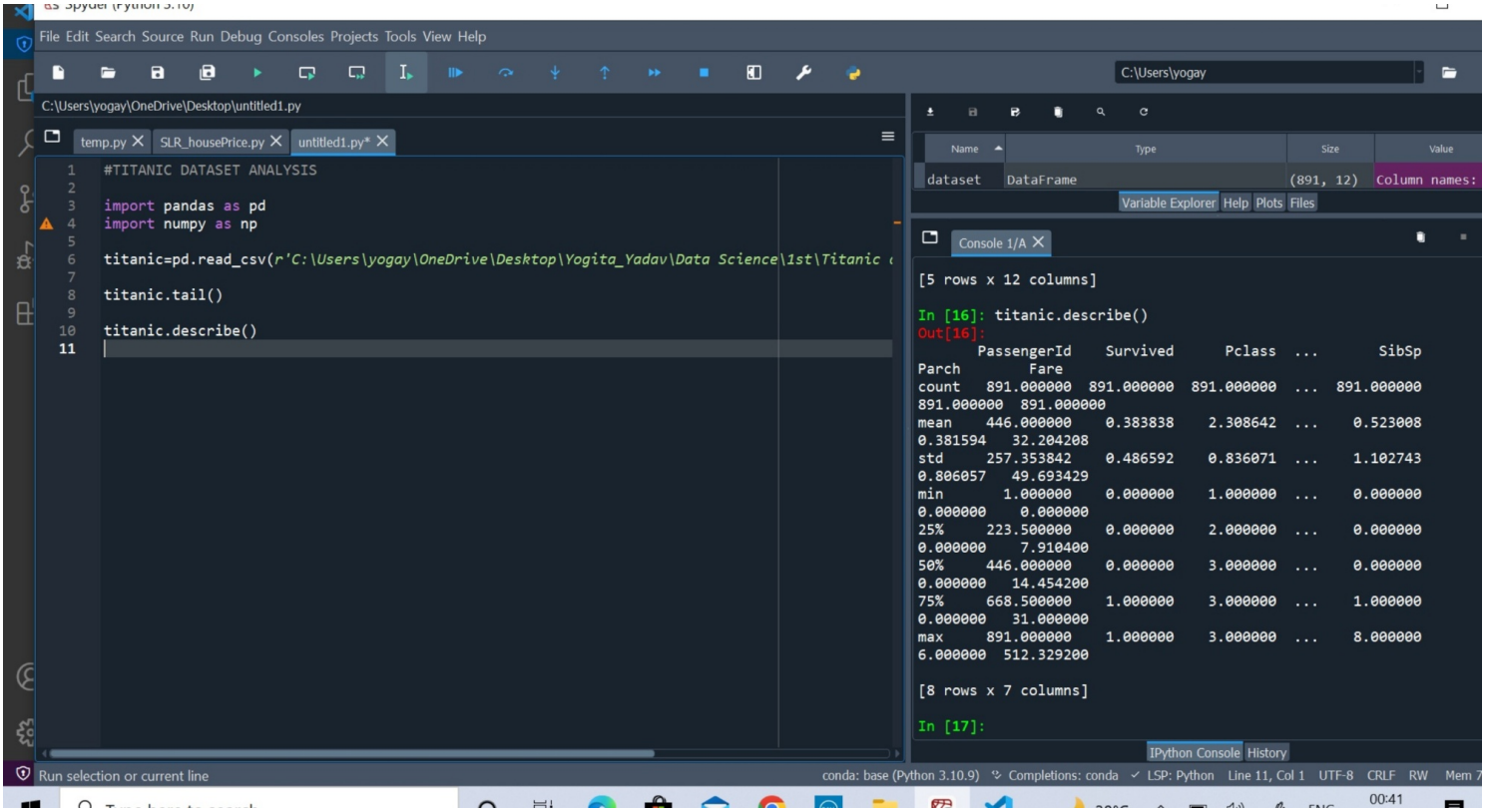
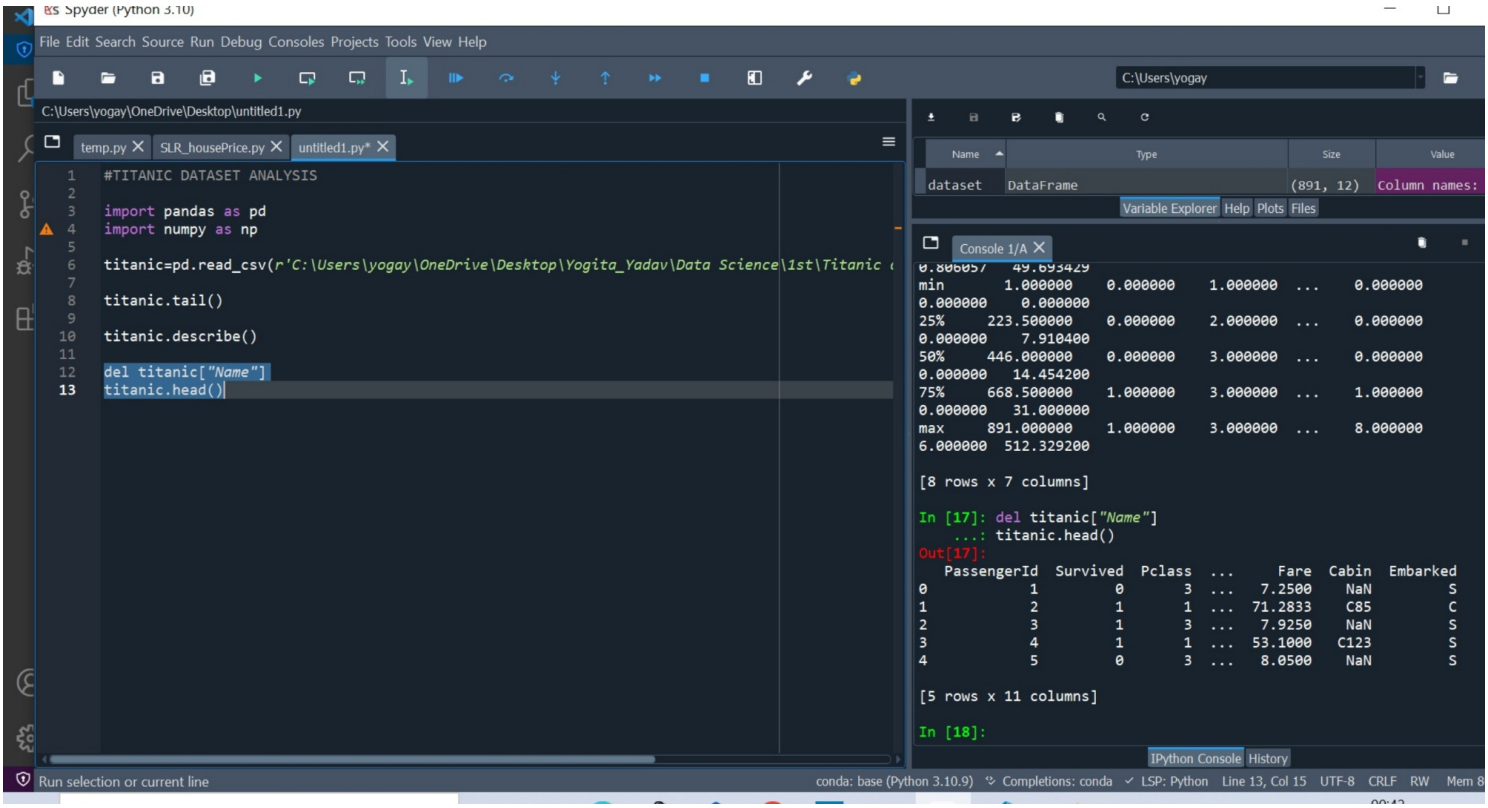
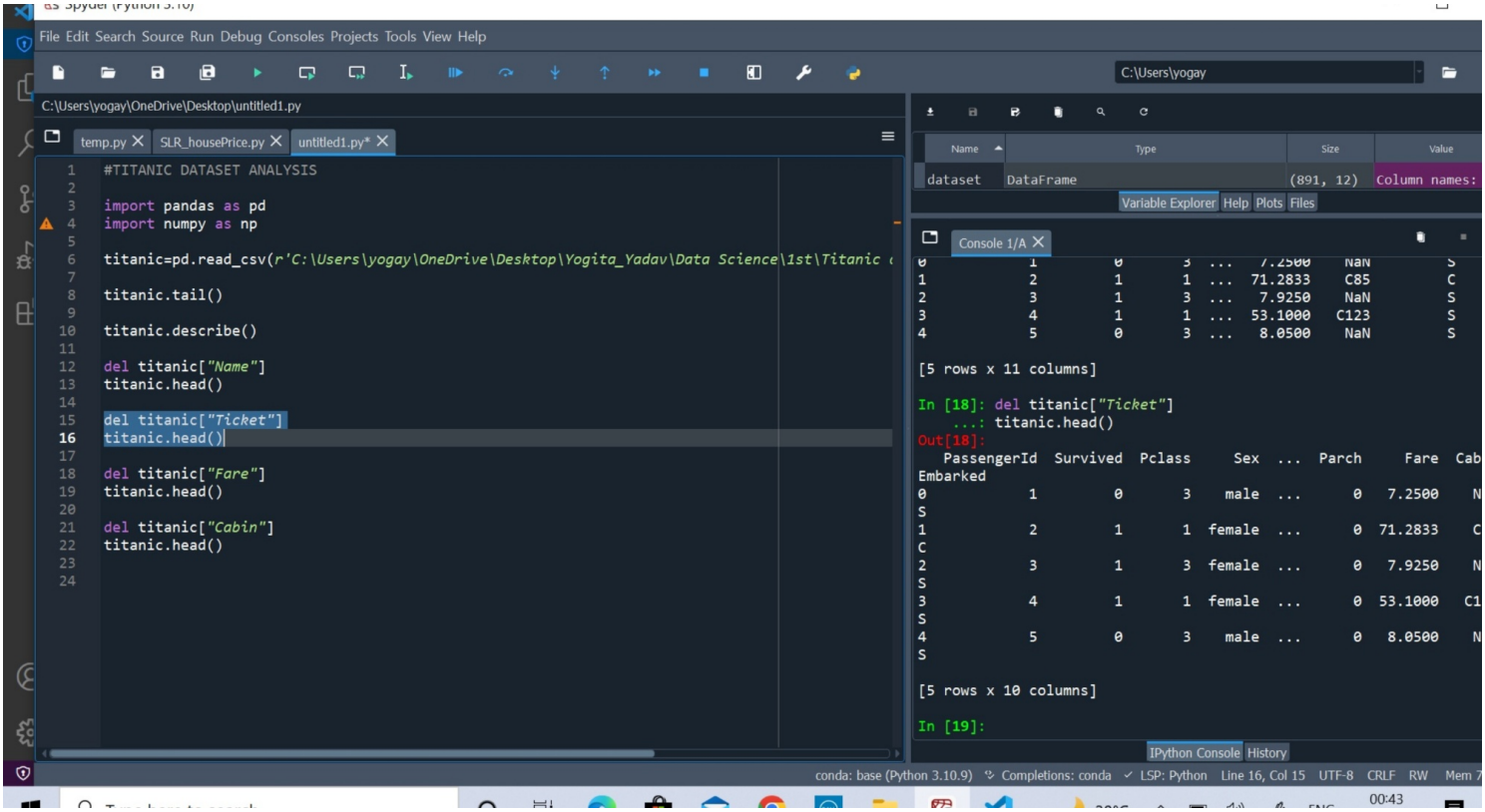
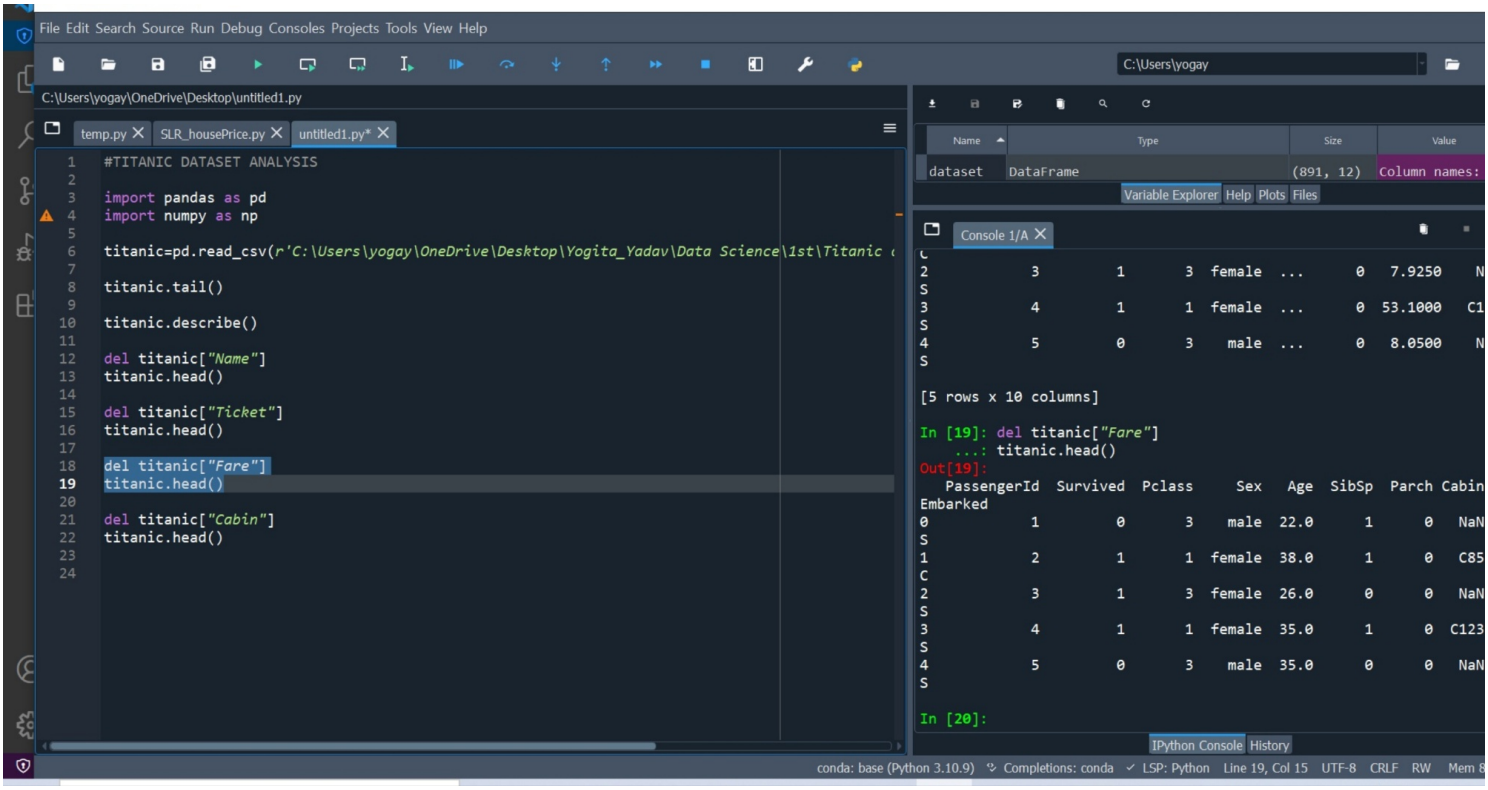
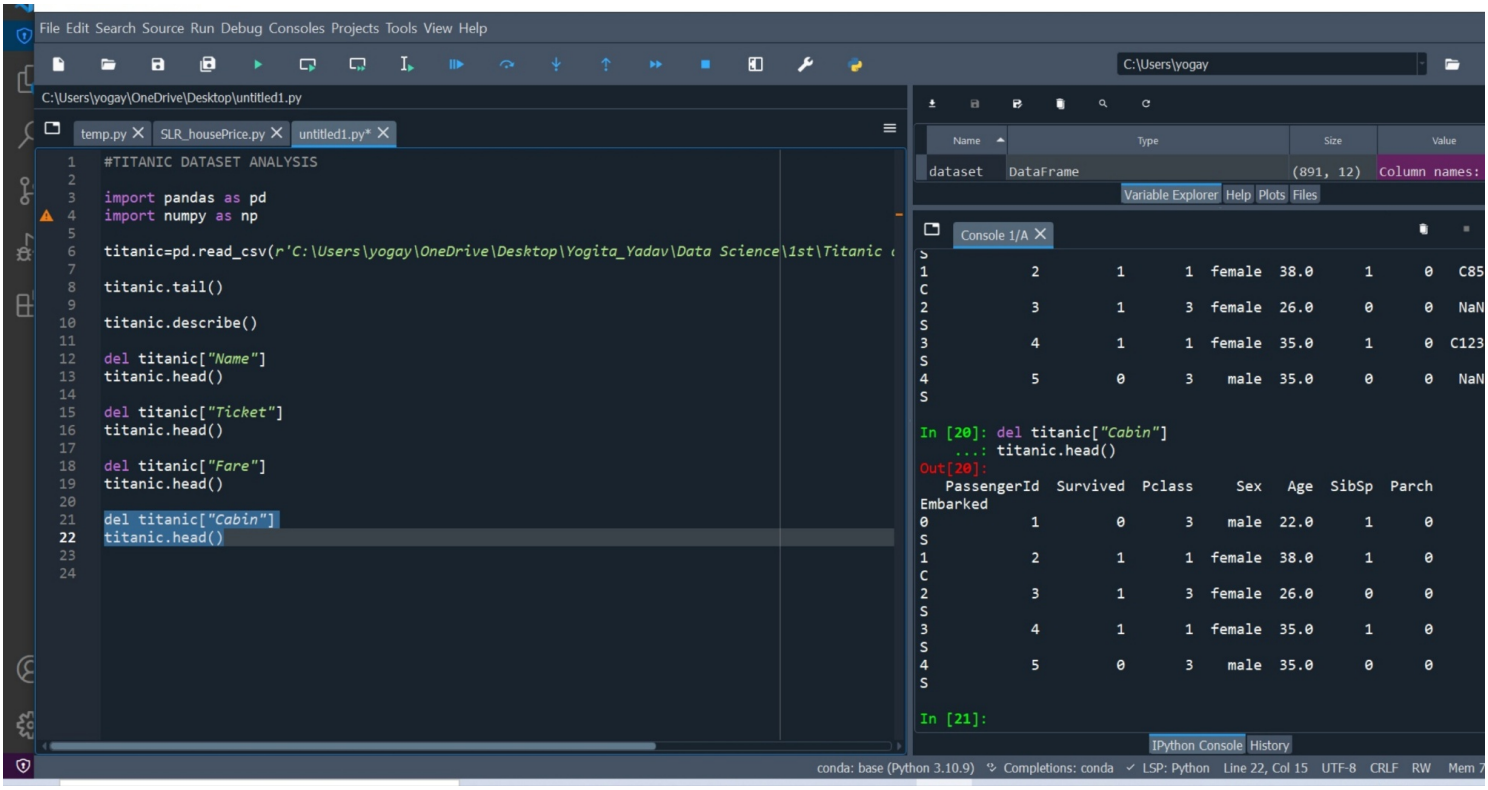

Spyder (Python 3.10)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
#TITANIC DATASET ANALYSIS
import pandas as pd
import numpy as np
titanic=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic d
titanic.tail()
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
...: plt.xlabel("Space")
...: plt.ylabel("Price")
...: plt.show()
In [12]: import pandas as pd
...: import numpy as np
In [13]: dataset=pd.read_csv(r'C:
\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic
dataset analysis\DATASET\train.csv')
In [14]: titanic=pd.read_csv(r'C:
\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic
dataset analysis\DATASET\train.csv')
In [15]: titanic.tail()
Out[15]:
PassengerId Survived Pclass ... Fare Cabin Embarked
886 887 0 2 ... 13.00 NaN S
887 888 1 1 ... 30.00 B42 S
888 889 0 3 ... 23.45 NaN S
889 890 1 1 ... 30.00 C148 C
890 891 0 3 ... 7.75 NaN Q
[5 rows x 12 columns]
In [16]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 9, Col 1
UTF-8
CRLF
RW
Mem 78

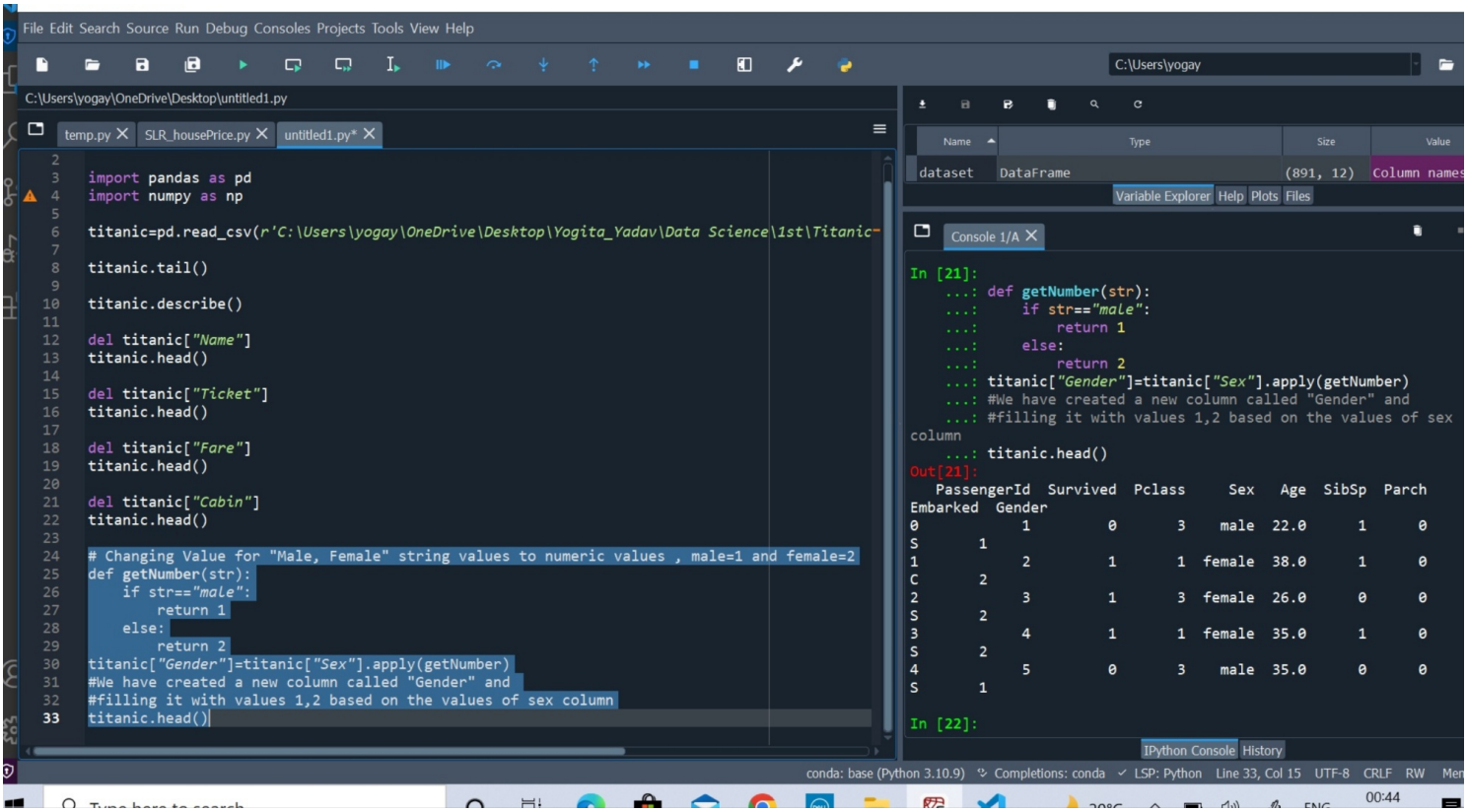

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
#TITANIC DATASET ANALYSIS
import pandas as pd
import numpy as np
titanic=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic
titanic.tail()
titanic.describe()
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
[5 rows x 12 columns]
In [16]: titanic.describe()
Out[16]:
PassengerId Survived Pclass ... SibSp
Parch Fare
count 891.000000 891.000000 891.000000 ... 891.000000
891.000000 891.000000
mean 446.000000 0.383838 2.308642 ... 0.523008
0.381594 32.204208
std 257.353842 0.486592 0.836071 ... 1.102743
0.806057 49.693429
min 1.000000 0.000000 1.000000 ... 0.000000
0.000000 0.000000
25% 223.500000 0.000000 2.000000 ... 0.000000
0.000000 7.910400
50% 446.000000 0.000000 3.000000 ... 0.000000
0.000000 14.454200
75% 668.500000 1.000000 3.000000 ... 1.000000
0.000000 31.000000
max 891.000000 1.000000 3.000000 ... 8.000000
6.000000 512.329200
[8 rows x 7 columns]
In [17]:
IPython Console
History
Run selection or current line
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 11, Col 1
UTF-8
CRLF
RW
Mem 7
00:41

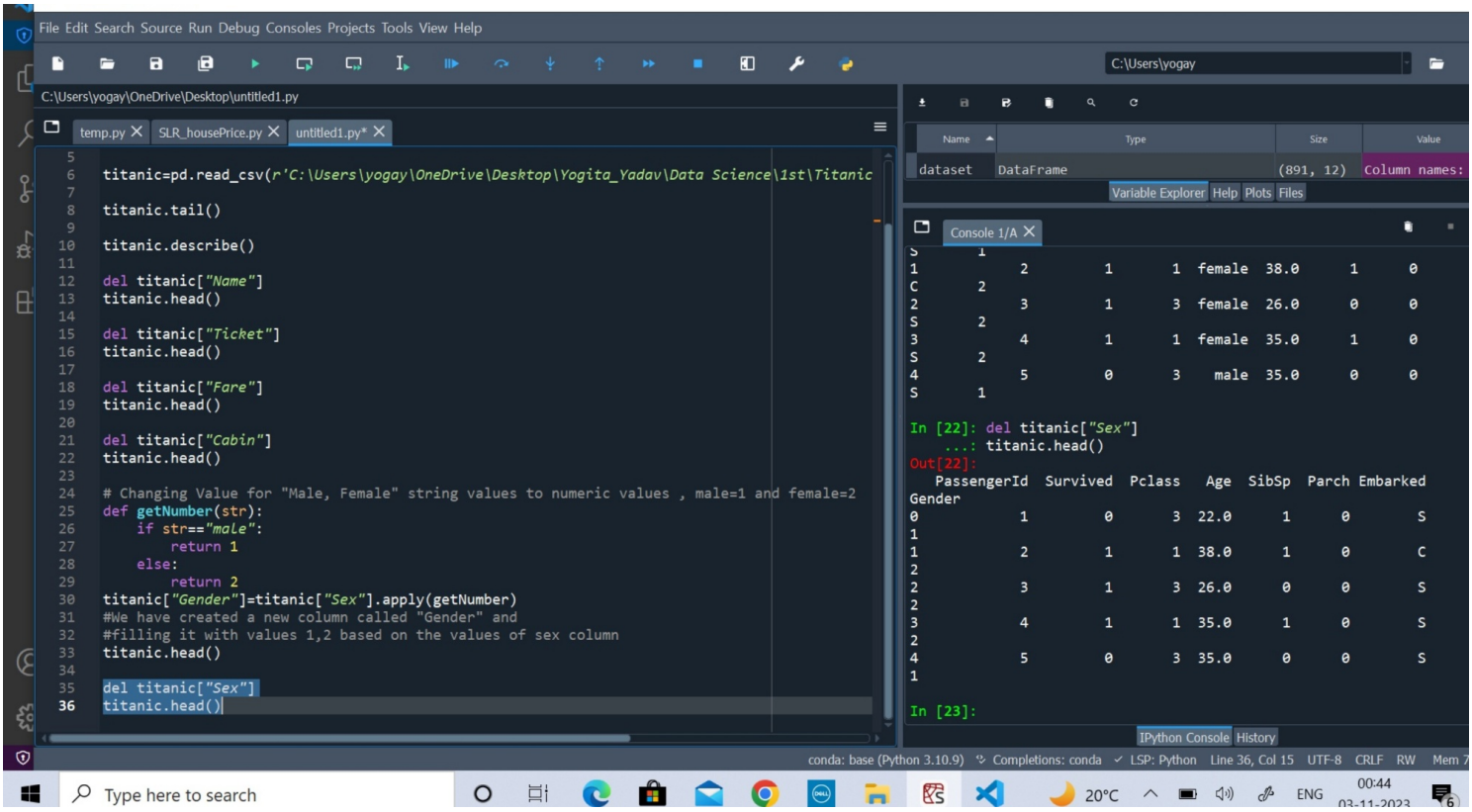

Spyder (Python 3.10)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
#TITANIC DATASET ANALYSIS
import pandas as pd
import numpy as np
titanic=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic
titanic.tail()
titanic.describe()
del titanic["Name"]
titanic.head()
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
min 1.000000 0.000000 1.000000 ... 0.000000
0.000000 0.000000
25% 223.500000 0.000000 2.000000 ... 0.000000
0.000000 7.910400
50% 446.000000 0.000000 3.000000 ... 0.000000
0.000000 14.454200
75% 668.500000 1.000000 3.000000 ... 1.000000
0.000000 31.000000
max 891.000000 1.000000 3.000000 ... 8.000000
6.000000 512.329200
[8 rows x 7 columns]
In [17]: del titanic["Name"]
...: titanic.head()
Out[17]:
PassengerId Survived Pclass ... Fare Cabin Embarked
0 1 0 3 ... 7.2500 NaN S
1 2 1 1 ... 71.2833 C85 C
2 3 1 3 ... 7.9250 NaN S
3 4 1 1 ... 53.1000 C123 S
4 5 0 3 ... 8.0500 NaN S
[5 rows x 11 columns]
In [18]:
IPython Console
History
Run selection or current line
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 13, Col 15
UTF-8
CRLF
RW
Mem 8

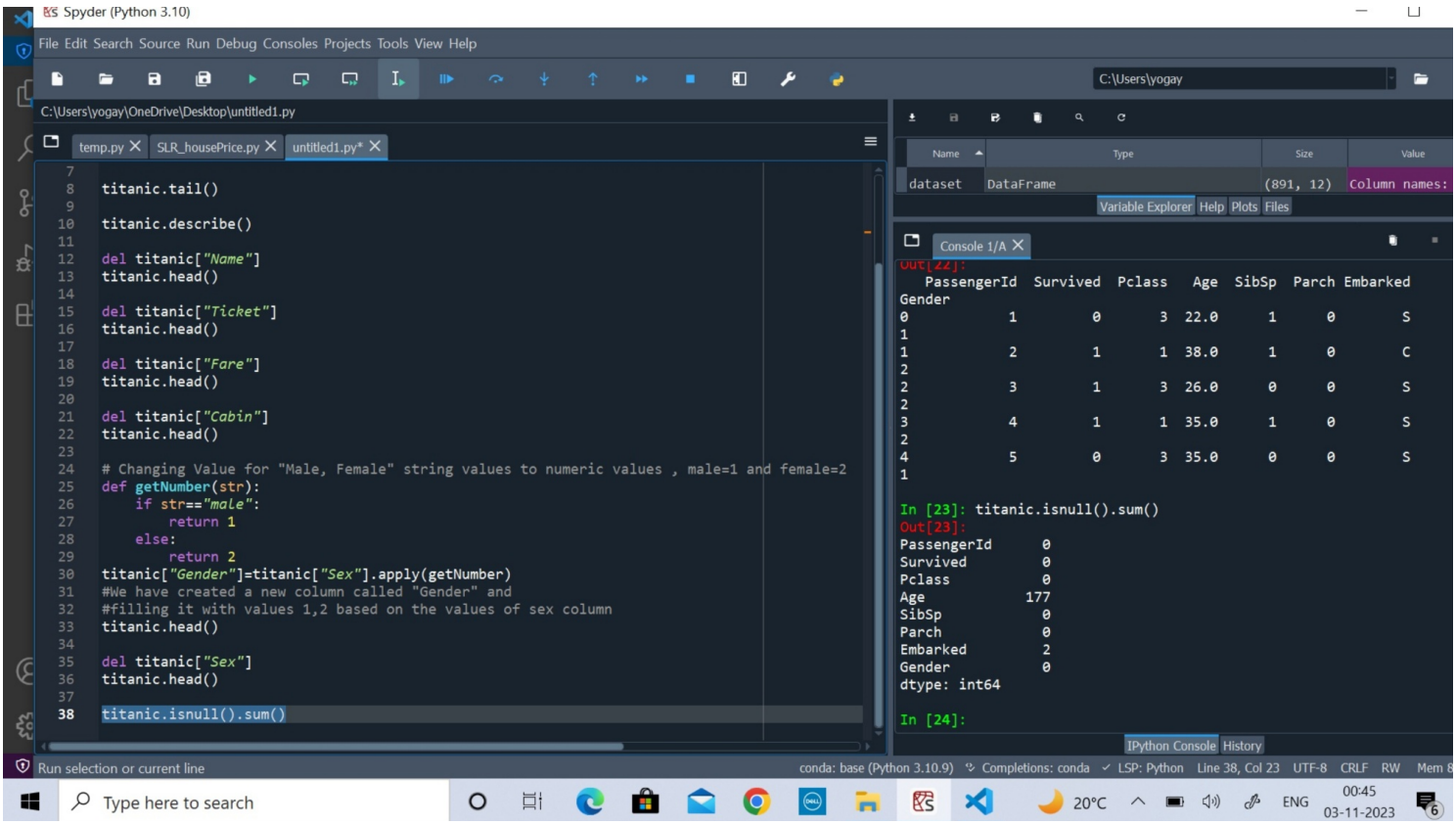

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
#TITANIC DATASET ANALYSIS
import pandas as pd
import numpy as np
titanic=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic
titanic.tail()
titanic.describe()
del titanic["Name"]
titanic.head()
del titanic["Ticket"]
titanic.head()
del titanic["Fare"]
titanic.head()
del titanic["Cabin"]
titanic.head()
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
1 2 1 1 ... 71.2833 C85 C
2 3 1 3 ... 7.9250 NaN S
3 4 1 1 ... 53.1000 C123 S
4 5 0 3 ... 8.0500 NaN S
[5 rows x 11 columns]
In [18]: del titanic["Ticket"]
...: titanic.head()
Out[18]:
PassengerId Survived Pclass Sex ... Parch Fare Cab
Embarked
0 1 0 3 male ... 0 7.2500 N
S
1 2 1 1 female ... 0 71.2833 C
C
2 3 1 3 female ... 0 7.9250 N
S
3 4 1 1 female ... 0 53.1000 C1
S
4 5 0 3 male ... 0 8.0500 N
S
[5 rows x 10 columns]
In [19]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 16, Col 15
UTF-8
CRLF
RW
00:43

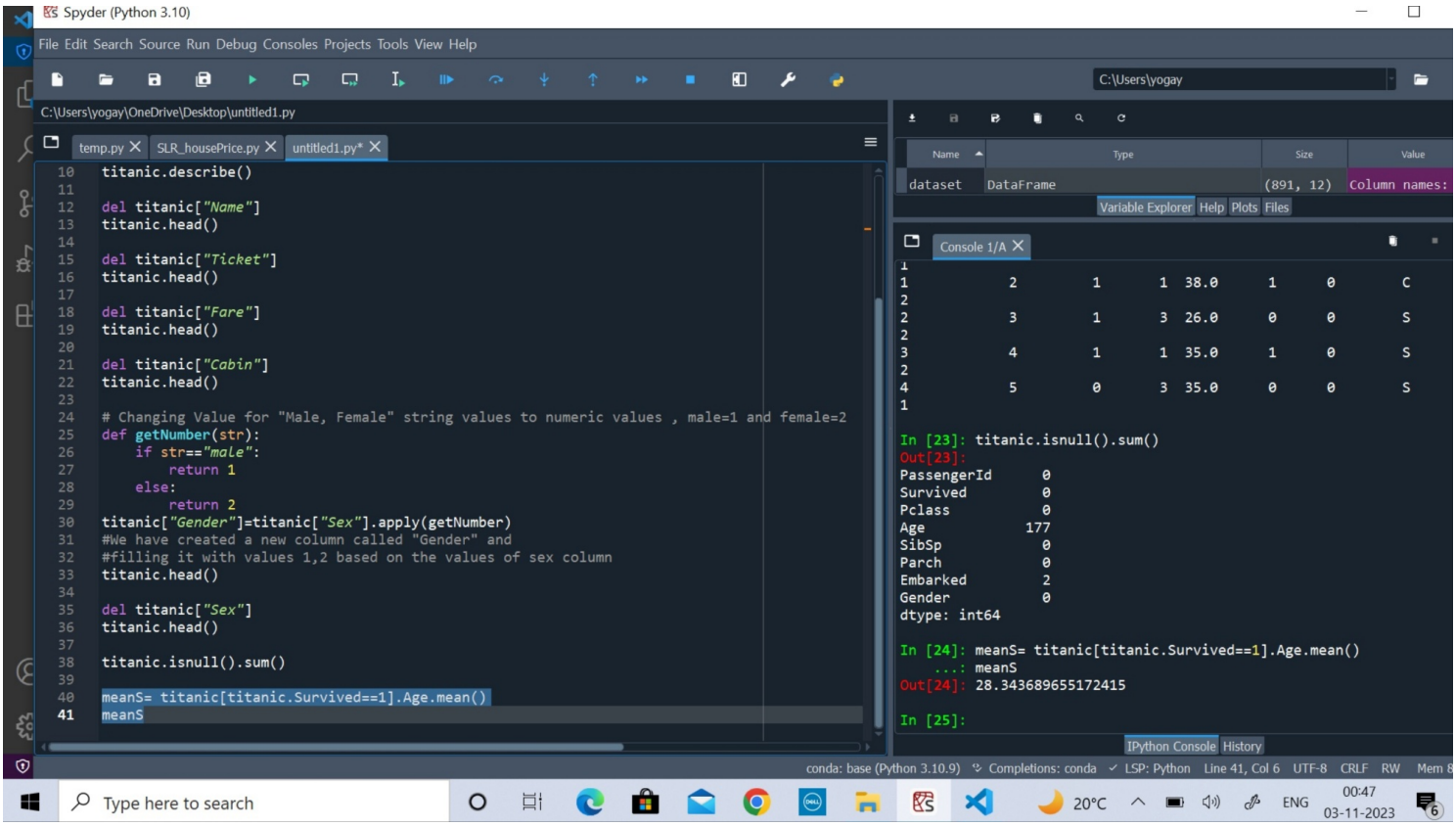

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
#TITANIC DATASET ANALYSIS
import pandas as pd
import numpy as np
titanic=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic
titanic.tail()
titanic.describe()
del titanic["Name"]
titanic.head()
del titanic["Ticket"]
titanic.head()
del titanic["Fare"]
titanic.head()
del titanic["Cabin"]
titanic.head()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
2 3 1 3 female ... 0 7.9250 N
S
3 4 1 1 female ... 0 53.1000 C1
S
4 5 0 3 male ... 0 8.0500 N
S
[5 rows x 10 columns]
In [19]: del titanic["Fare"]
...: titanic.head()
Out[19]:
PassengerId Survived Pclass Sex Age SibSp Parch Cabin
Embarked
0 1 0 3 male 22.0 1 0 NaN
S
1 2 1 1 female 38.0 1 0 C85
C
2 3 1 3 female 26.0 0 0 NaN
S
3 4 1 1 female 35.0 1 0 C123
S
4 5 0 3 male 35.0 0 0 NaN
S
In [20]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 19, Col 15
UTF-8
CRLF
RW

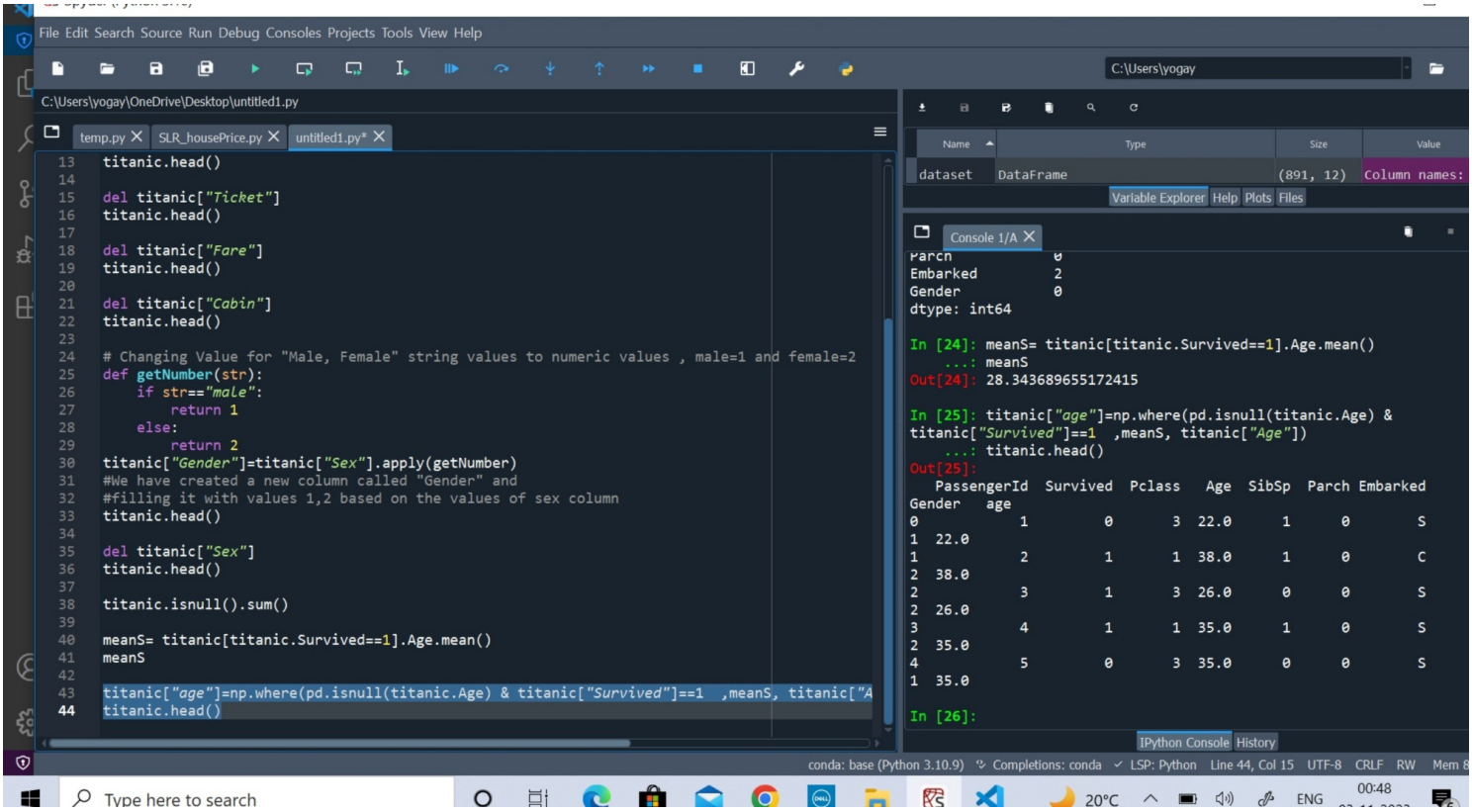

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
#TITANIC DATASET ANALYSIS
import pandas as pd
import numpy as np
titanic=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic
titanic.tail()
titanic.describe()
del titanic["Name"]
titanic.head()
del titanic["Ticket"]
titanic.head()
del titanic["Fare"]
titanic.head()
del titanic["Cabin"]
titanic.head()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
In [20]: del titanic["Cabin"]
...: titanic.head()
Out[20]:
PassengerId Survived Pclass Sex Age SibSp Parch
Embarked
In [21]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 22, Col 15
UTF-8
CRLF
RW

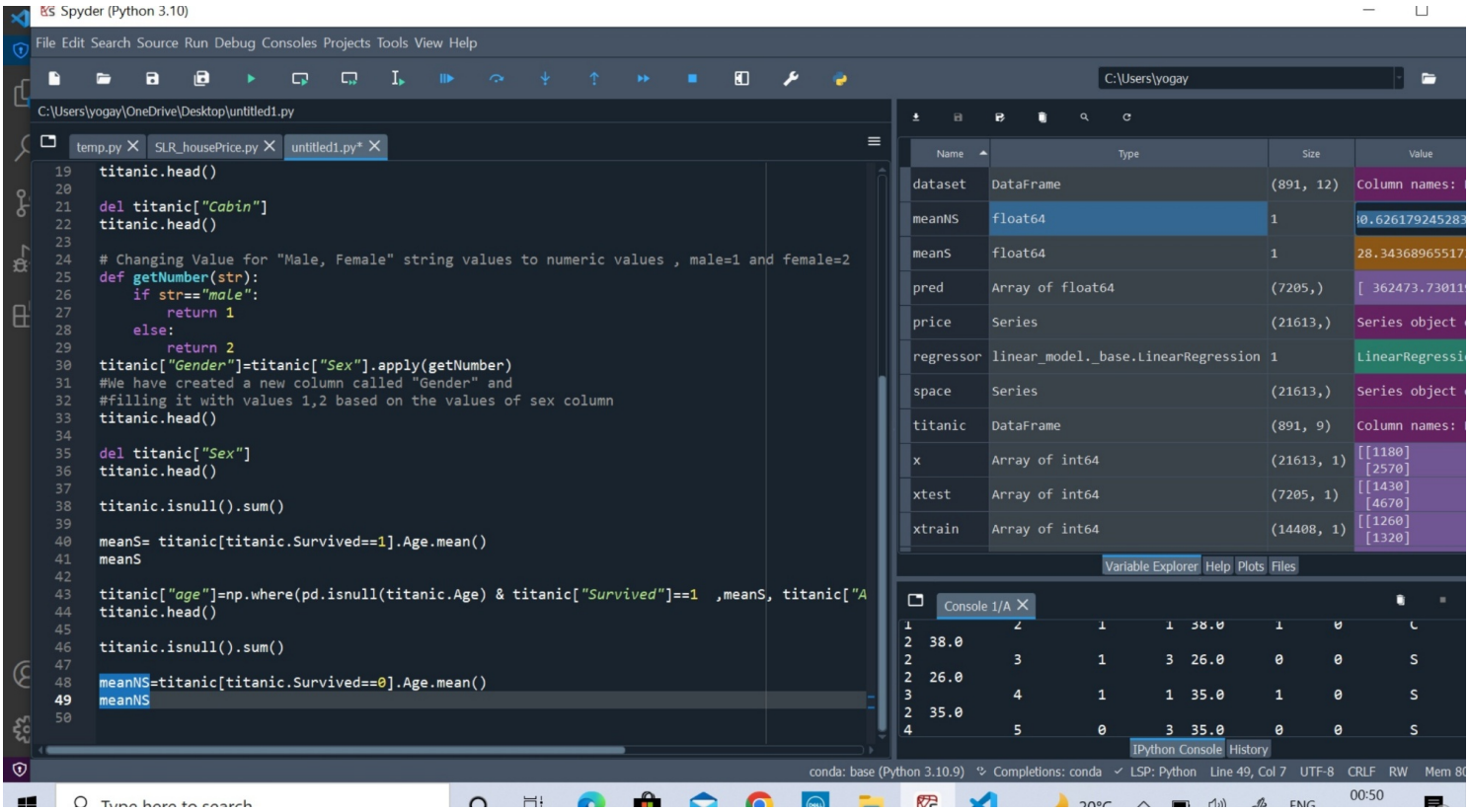

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
import pandas as pd
import numpy as np
titanic=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic-
titanic.tail()
titanic.describe()
del titanic["Name"]
titanic.head()
del titanic["Ticket"]
titanic.head()
del titanic["Fare"]
titanic.head()
del titanic["Cabin"]
titanic.head()
# Changing Value for "Male, Female" string values to numeric values , male=1 and female=2
def getNumber(str):
if str=="male":
return 1
else:
return 2
titanic["Gender"]=titanic["Sex"].apply(getNumber)
#We have created a new column called "Gender" and
#filling it with values 1,2 based on the values of sex column
titanic.head()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names
Variable Explorer
Help
Plots
Files
Console 1/A
In [21]:
... def getNumber(str):
... if str=="male":
... return 1
... else:
... return 2
... titanic["Gender"]=titanic["Sex"].apply(getNumber)
... #We have created a new column called "Gender" and
... #filling it with values 1,2 based on the values of sex column
... titanic.head()
Out[21]:
PassengerId Survived Pclass Sex Age SibSp Parch Embarked Gender
0 1 0 3 male 22.0 1 0 S 1
1 2 1 1 female 38.0 1 0 C 2
2 3 1 3 female 26.0 0 0 S 2
3 4 1 1 female 35.0 1 0 S 2
4 5 0 3 male 35.0 0 0 S 1
In [22]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 33, Col 15
UTF-8
CRLF
RW
00:44

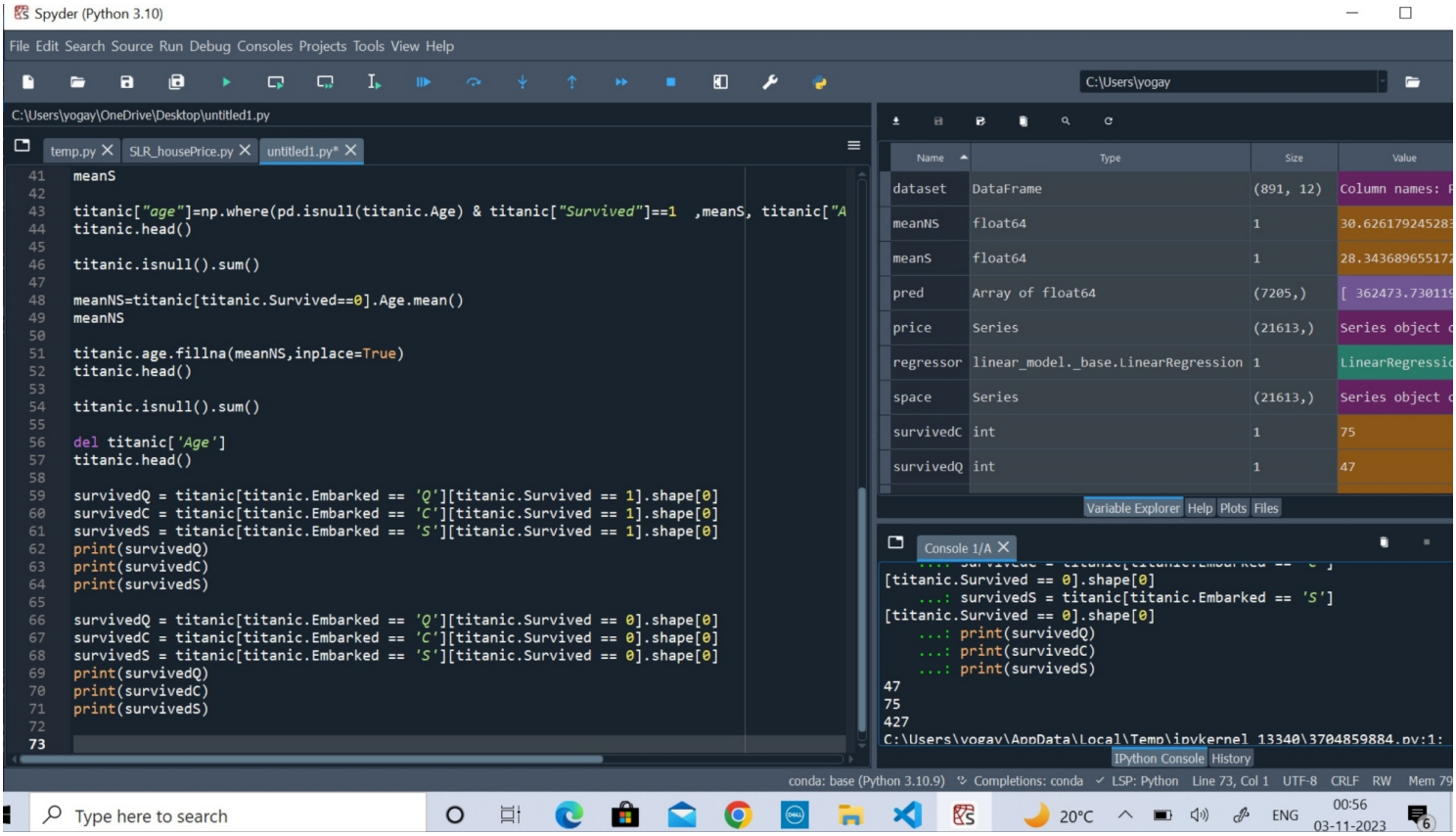

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
titanic=pd.read_csv(r'C:\Users\yogay\OneDrive\Desktop\Yogita_Yadav\Data Science\1st\Titanic
titanic.tail()
titanic.describe()
del titanic["Name"]
titanic.head()
del titanic["Ticket"]
titanic.head()
del titanic["Fare"]
titanic.head()
del titanic["Cabin"]
titanic.head()
# Changing Value for "Male, Female" string values to numeric values , male=1 and female=2
def getNumber(str):
if str=="male":
return 1
else:
return 2
titanic["Gender"]=titanic["Sex"].apply(getNumber)
#We have created a new column called "Gender" and
#filling it with values 1,2 based on the values of sex column
titanic.head()
del titanic["Sex"]
titanic.head()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
In [22]: del titanic["Sex"]
...: titanic.head()
Out[22]:
PassengerId Survived Pclass Age SibSp Parch Embarked
Gender
In [23]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 36, Col 15
UTF-8
CRLF
RW
Type here to search
20°C
ENG
00:44

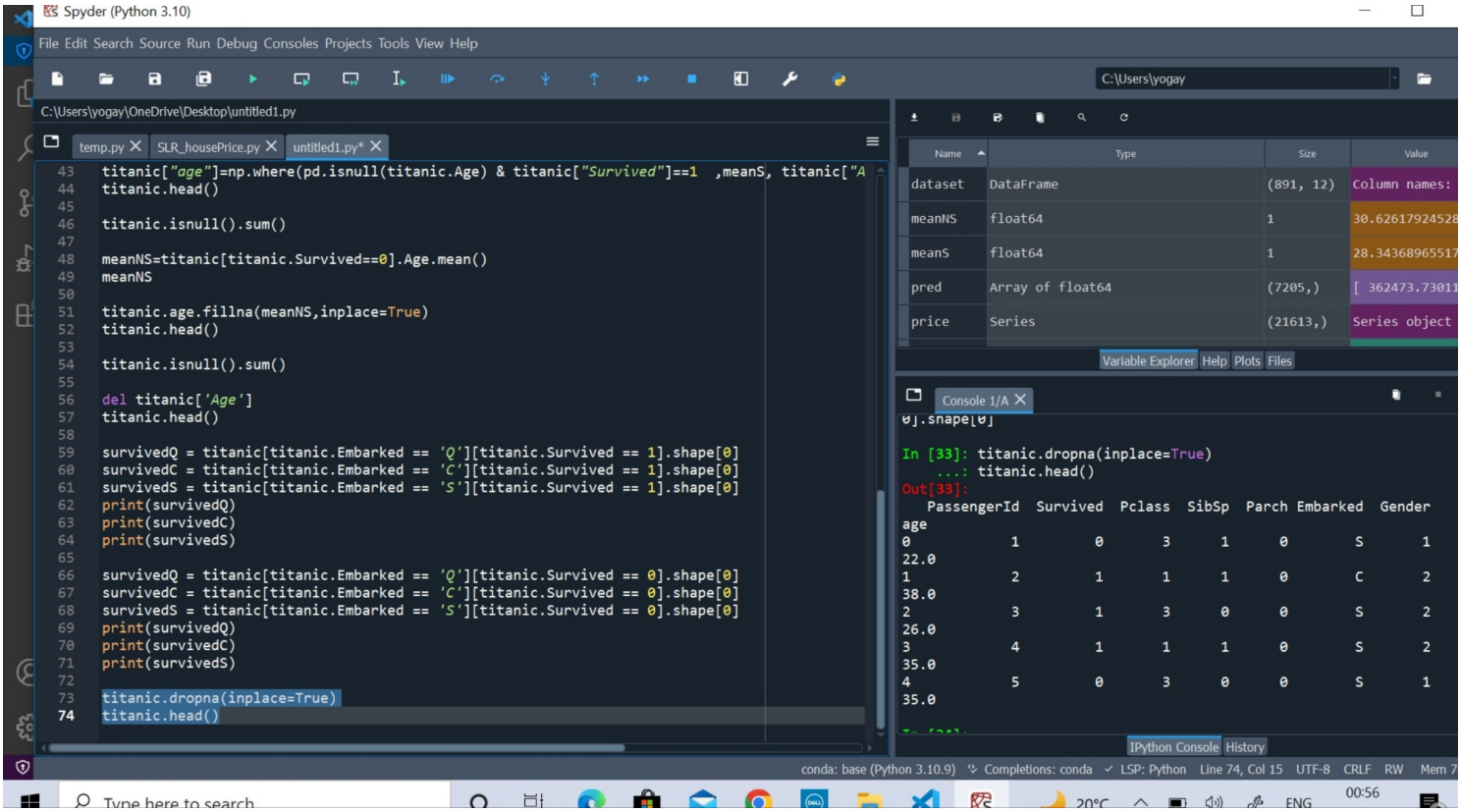

Spyder (Python 3.10)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
titanic.tail()
titanic.describe()
del titanic["Name"]
titanic.head()
del titanic["Ticket"]
titanic.head()
del titanic["Fare"]
titanic.head()
del titanic["Cabin"]
titanic.head()
# Changing Value for "Male, Female" string values to numeric values , male=1 and female=2
def getNumber(str):
if str=="male":
return 1
else:
return 2
titanic["Gender"]=titanic["Sex"].apply(getNumber)
#We have created a new column called "Gender" and
#filling it with values 1,2 based on the values of sex column
titanic.head()
del titanic["Sex"]
titanic.head()
titanic.isnull().sum()
C:\Users\yogay
Name Type Size Value
dataset DataFrame (891, 12) Column names:
Variable Explorer Help Plots Files
Console 1/A
PassengerId Survived Pclass Age SibSp Parch Embarked
Gender
0 1 0 3 22.0 1 0 S
1
1 2 1 1 38.0 1 0 C
2
2 3 1 3 26.0 0 0 S
2
3 4 1 1 35.0 1 0 S
2
4 5 0 3 35.0 0 0 S
1
In [23]: titanic.isnull().sum()
Out[23]:
PassengerId 0
Survived 0
Pclass 0
Age 177
SibSp 0
Parch 0
Embarked 2
Gender 0
dtype: int64
In [24]:
IPython Console History
Run selection or current line
conda: base (Python 3.10.9) Completions: conda LSP: Python Line 38, Col 23 UTF-8 CRLF RW
Type here to search
20°C
ENG
00:45
03-11-2023

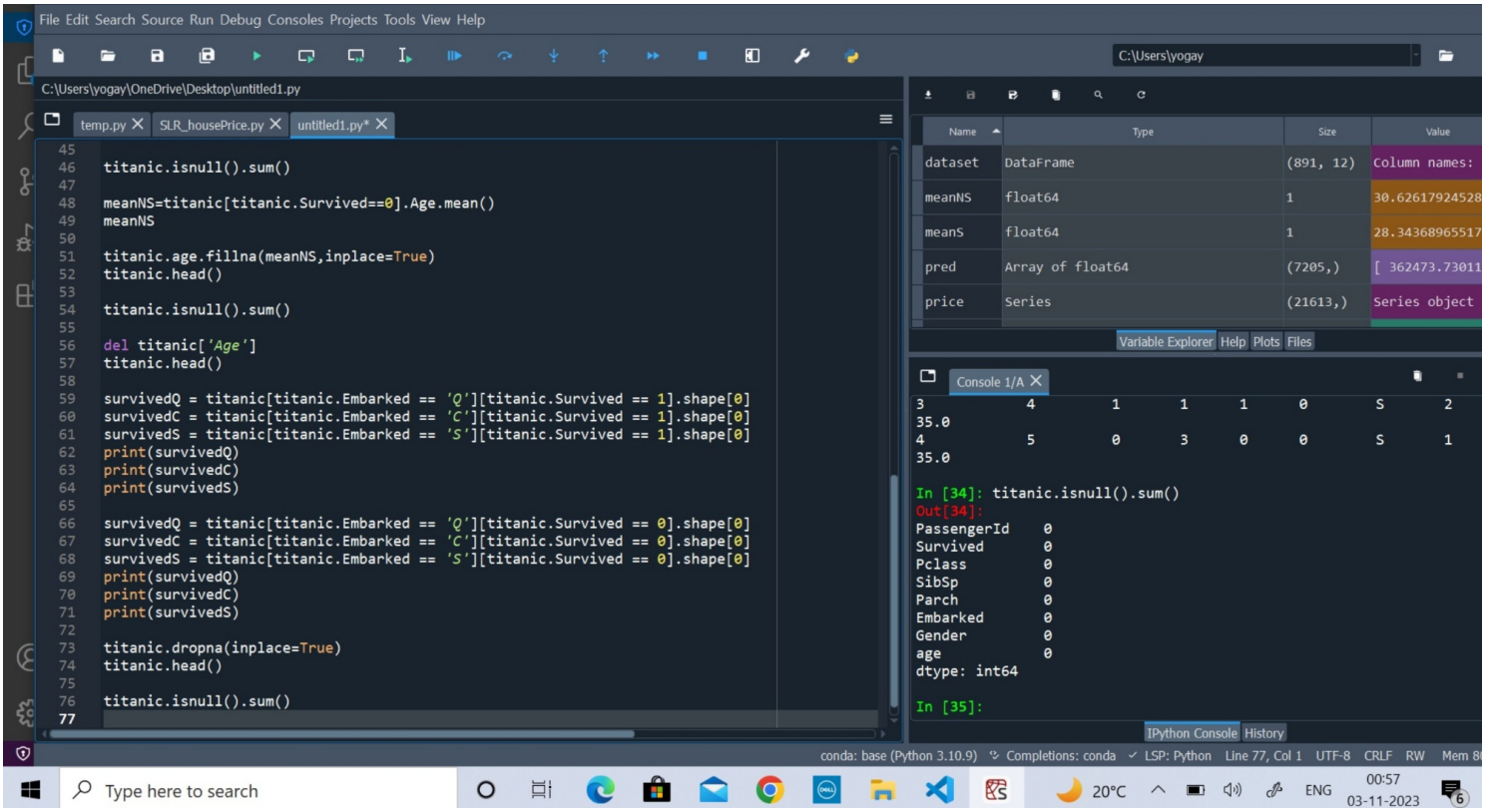

Spyder (Python 3.10)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
titanic.describe()
del titanic["Name"]
titanic.head()
del titanic["Ticket"]
titanic.head()
del titanic["Fare"]
titanic.head()
del titanic["Cabin"]
titanic.head()
# Changing Value for "Male, Female" string values to numeric values , male=1 and female=2
def getNumber(str):
if str=="male":
return 1
else:
return 2
titanic["Gender"]=titanic["Sex"].apply(getNumber)
#We have created a new column called "Gender" and
#filling it with values 1,2 based on the values of sex column
titanic.head()
del titanic["Sex"]
titanic.head()
titanic.isnull().sum()
meanS= titanic[titanic.Survived==1].Age.mean()
meanS
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
1 2 1 1 38.0 1 0 C
2 3 1 3 26.0 0 0 S
3 4 1 1 35.0 1 0 S
4 5 0 3 35.0 0 0 S
In [23]: titanic.isnull().sum()
Out[23]:
PassengerId 0
Survived 0
Pclass 0
Age 177
SibSp 0
Parch 0
Embarked 2
Gender 0
dtype: int64
In [24]: meanS= titanic[titanic.Survived==1].Age.mean()
...: meanS
Out[24]: 28.343689655172415
In [25]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 41, Col 6
UTF-8
CRLF
RW
Type here to search
20°C
ENG
00:47
03-11-2023

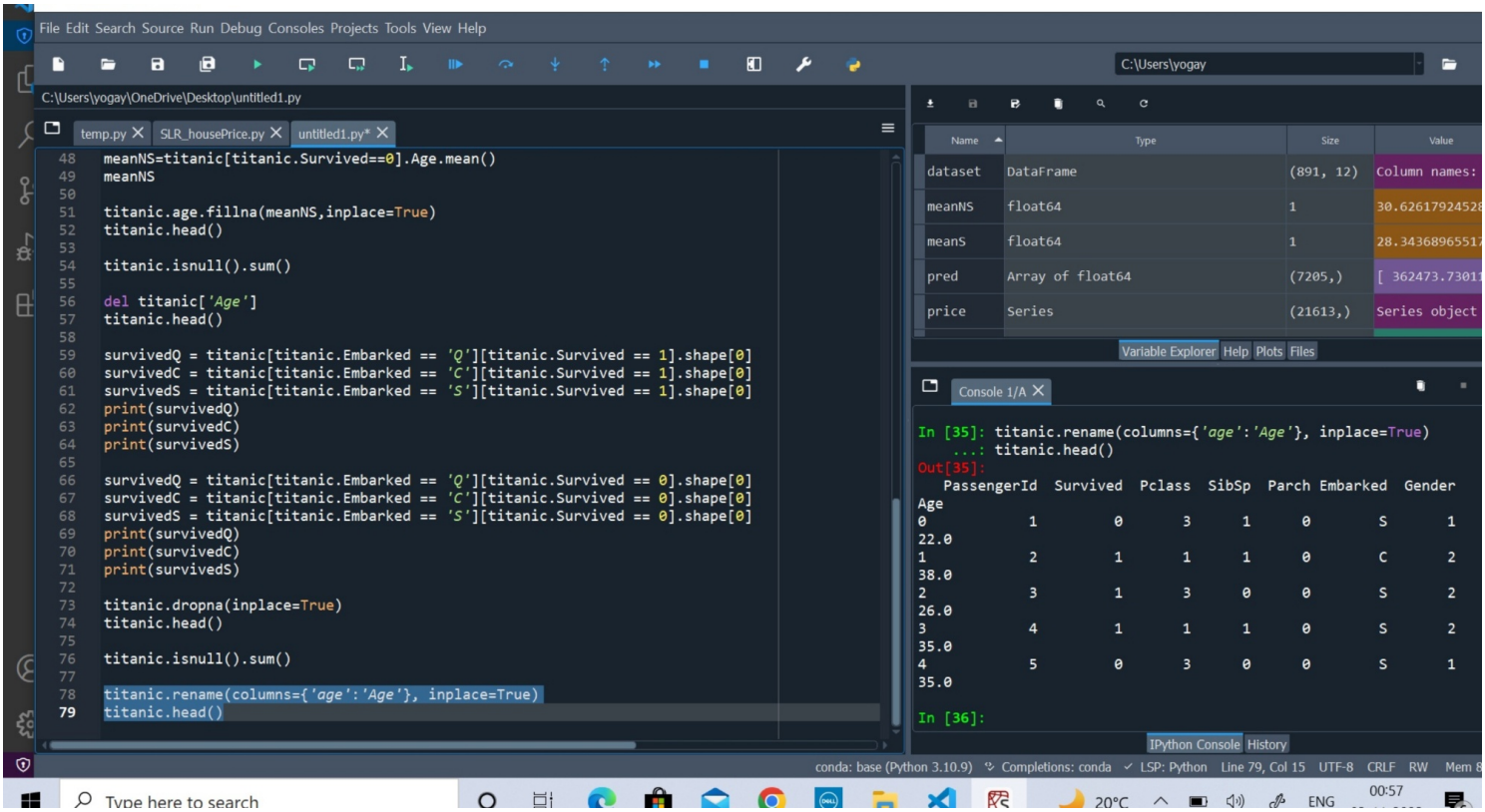

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
titanic.head()
del titanic["Ticket"]
titanic.head()
del titanic["Fare"]
titanic.head()
del titanic["Cabin"]
titanic.head()
# Changing Value for "Male, Female" string values to numeric values , male=1 and female=2
def getNumber(str):
if str=="male":
return 1
else:
return 2
titanic["Gender"]=titanic["Sex"].apply(getNumber)
#We have created a new column called "Gender" and
#filling it with values 1,2 based on the values of sex column
titanic.head()
del titanic["Sex"]
titanic.head()
titanic.isnull().sum()
meanS= titanic[titanic.Survived==1].Age.mean()
meanS
titanic["age"]=np.where(pd.isnull(titanic.Age) & titanic["Survived"]==1 ,meanS, titanic["A
titanic.head()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
Variable Explorer
Help
Plots
Files
Console 1/A
Embarked 2
Gender 0
dtype: int64
In [24]: meanS= titanic[titanic.Survived==1].Age.mean()
...: meanS
Out[24]: 28.343689655172415
In [25]: titanic["age"]=np.where(pd.isnull(titanic.Age) &
titanic["Survived"]==1 ,meanS, titanic["Age"])
...: titanic.head()
Out[25]:
PassengerId Survived Pclass Age SibSp Parch Embarked
Gender age
0 1 0 3 22.0 1 0 S
1 22.0
1 2 1 1 38.0 1 0 C
2 38.0
2 3 1 3 26.0 0 0 S
2 26.0
3 4 1 1 35.0 1 0 S
2 35.0
4 5 0 3 35.0 0 0 S
1 35.0
In [26]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 44, Col 15
UTF-8
CRLF
RW
Type here to search
20°C
ENG
00:48

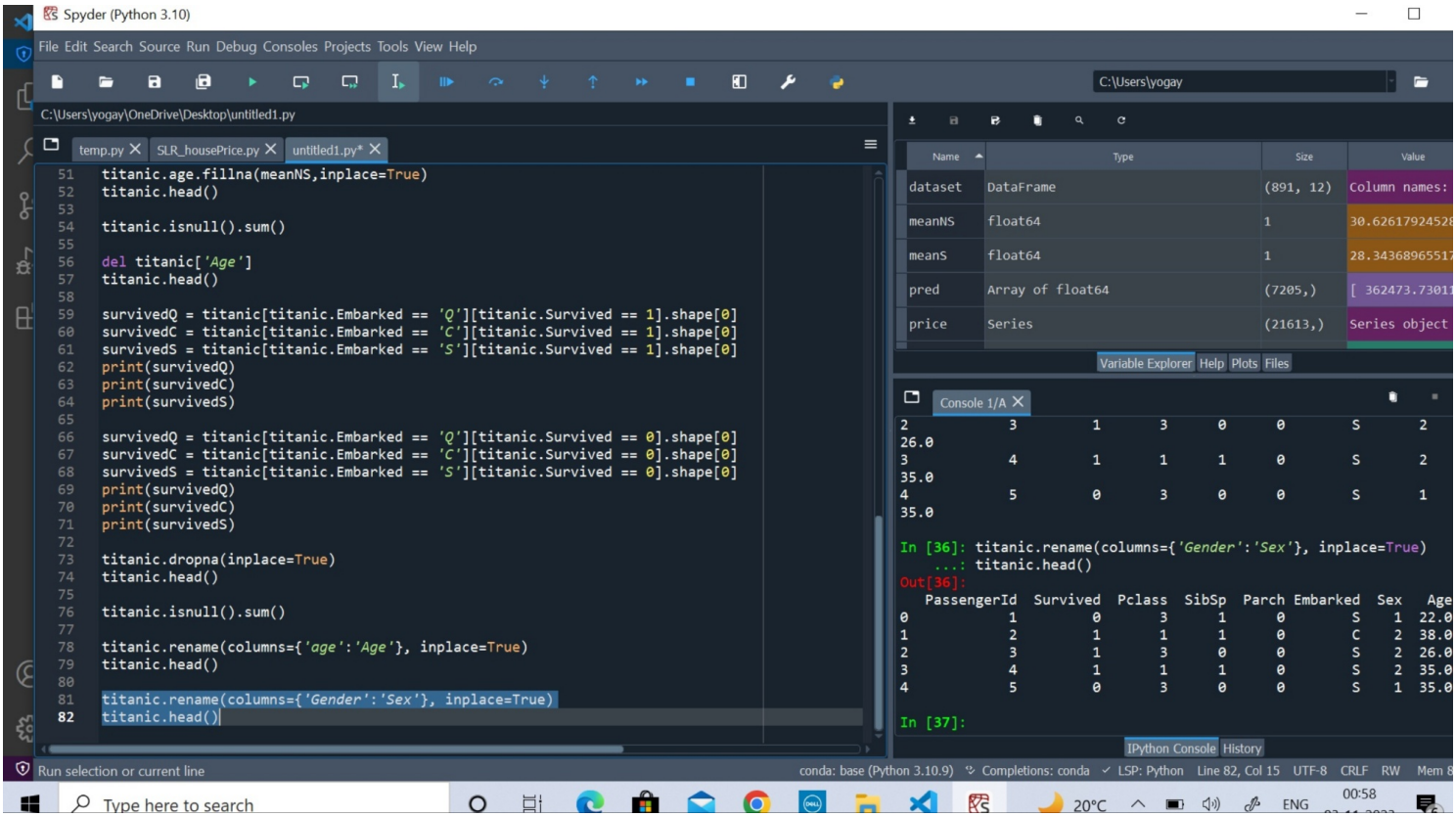

Spyder (Python 3.10)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
titanic.head()
del titanic["Cabin"]
titanic.head()
# Changing Value for "Male, Female" string values to numeric values , male=1 and female=2
def getNumber(str):
if str=="male":
return 1
else:
return 2
titanic["Gender"]=titanic["Sex"].apply(getNumber)
#We have created a new column called "Gender" and
#filling it with values 1,2 based on the values of sex column
titanic.head()
del titanic["Sex"]
titanic.head()
titanic.isnull().sum()
meanS= titanic[titanic.Survived==1].Age.mean()
meanS
titanic["age"]=np.where(pd.isnull(titanic.Age) & titanic["Survived"]==1 ,meanS, titanic["A
titanic.head()
titanic.isnull().sum()
meanNS=titanic[titanic.Survived==0].Age.mean()
meanNS
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
meanNS
float64
1
meanS
float64
1
28.34368965517
pred
Array of float64
(7205,)
price
Series
(21613,)
Series object
regressor
linear_model._base.LinearRegression
1
LinearRegressi
space
Series
(21613,)
Series object
titanic
DataFrame
(891, 9)
Column names:
x
Array of int64
(21613, 1)
xtest
Array of int64
(7205, 1)
xtrain
Array of int64
(14408, 1)
Variable Explorer
Help
Plots
Files
Console 1/A
2 38.0
2 26.0
2 35.0
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 49, Col 7
UTF-8
CRLF
RW
00:50

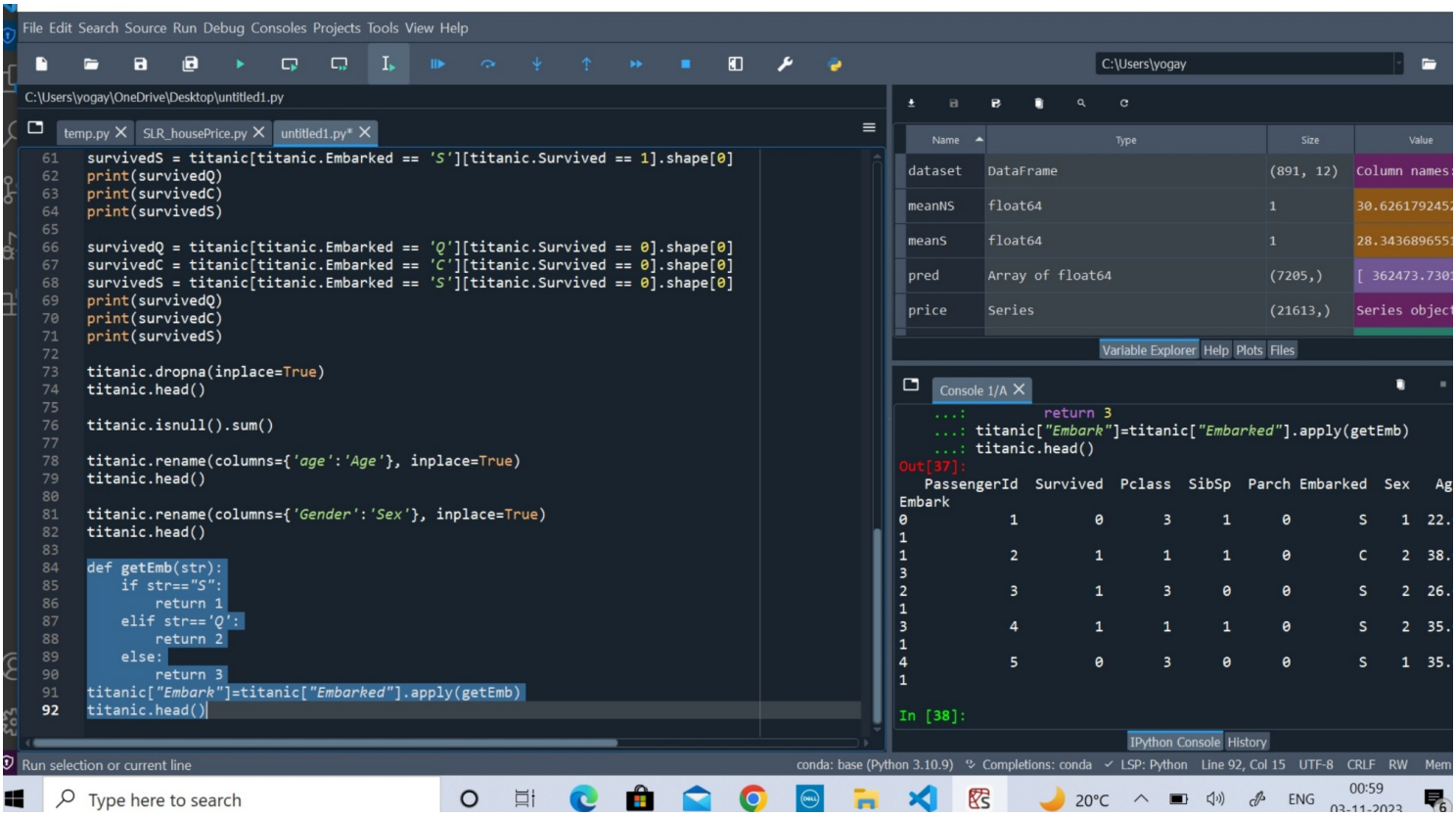

Spyder (Python 3.10)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
meanS
titanic["age"]=np.where(pd.isnull(titanic.Age) & titanic["Survived"]==1 ,meanS, titanic["A
titanic.head()
titanic.isnull().sum()
meanNS=titanic[titanic.Survived==0].Age.mean()
meanNS
titanic.age.fillna(meanNS,inplace=True)
titanic.head()
titanic.isnull().sum()
del titanic['Age']
titanic.head()
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 1].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 1].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 1].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 0].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 0].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 0].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
meanNS
float64
1
meanS
float64
1
pred
Array of float64
(7205,)
price
Series
(21613,)
regressor
linear_model._base.LinearRegression
1
space
Series
(21613,)
survivedC
int
1
75
survivedQ
int
1
47
Variable Explorer
Help
Plots
Files
Console 1/A
[titanic.Survived == 0].shape[0]
survivedS = titanic[titanic.Embarked == 'S']
[titanic.Survived == 0].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
47
75
427
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 73, Col 1
UTF-8
CRLF
RW
Type here to search
20°C
ENG
00:56
03-11-2023

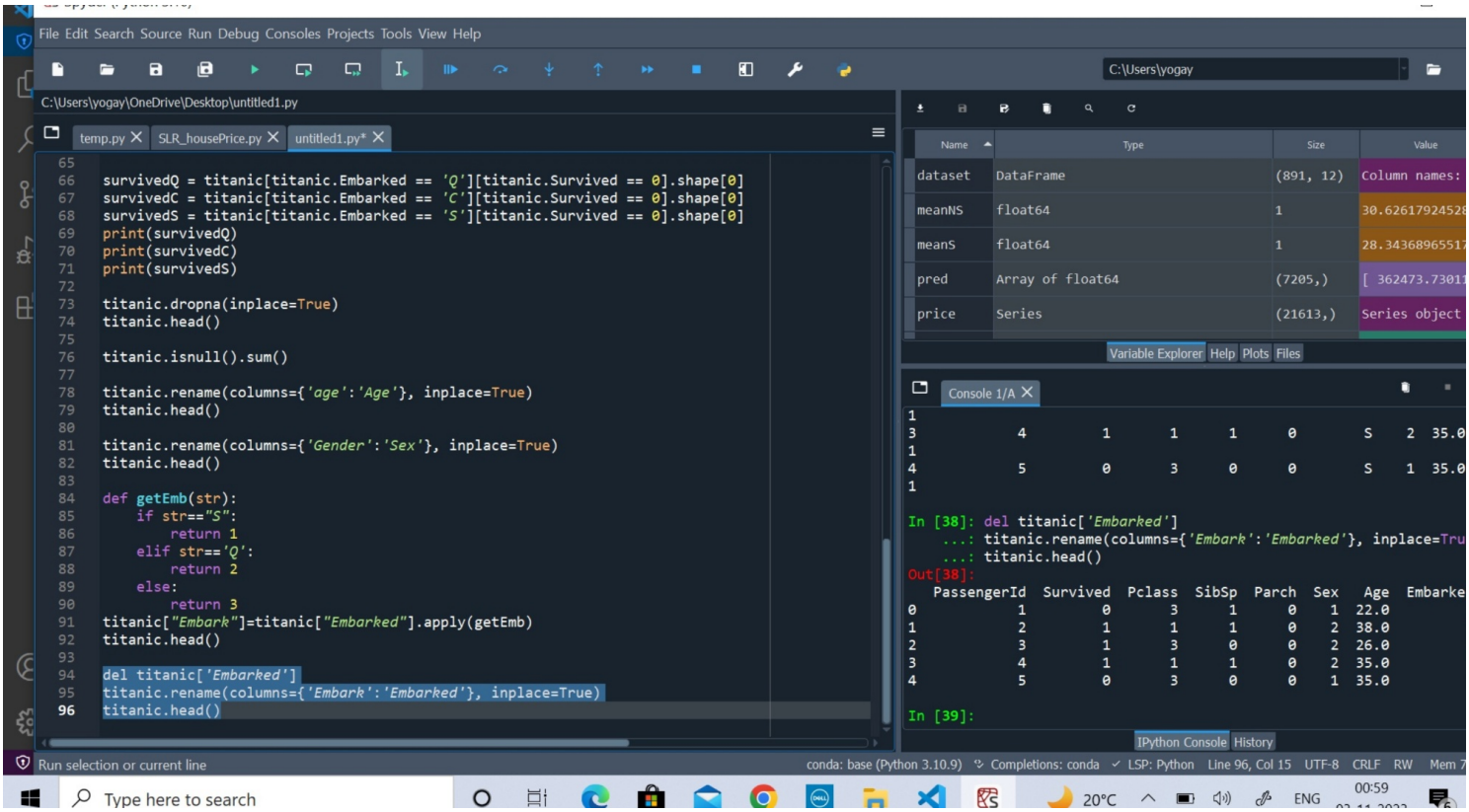

Spyder (Python 3.10)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
titanic["age"]=np.where(pd.isnull(titanic.Age) & titanic["Survived"]==1 ,meanS, titanic["A
titanic.head()
titanic.isnull().sum()
meanNS=titanic[titanic.Survived==0].Age.mean()
meanNS
titanic.age.fillna(meanNS,inplace=True)
titanic.head()
titanic.isnull().sum()
del titanic['Age']
titanic.head()
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 1].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 1].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 1].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 0].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 0].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 0].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
titanic.dropna(inplace=True)
titanic.head()
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
meanNS
float64
1
30.6261792452
meanS
float64
1
28.3436896551
pred
Array of float64
(7205,)
[362473.7301
price
Series
(21613,)
Series object
Variable Explorer
Help
Plots
Files
Console 1/A
In [33]: titanic.dropna(inplace=True)
...: titanic.head()
Out[33]:
PassengerId Survived Pclass SibSp Parch Embarked Gender
age
0 1 0 3 1 0 S 1
22.0
1 2 1 1 1 0 C 2
38.0
2 3 1 3 0 0 S 2
26.0
3 4 1 1 1 0 S 2
35.0
4 5 0 3 0 0 S 1
35.0
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 74, Col 15
UTF-8
CRLF
RW
Mem 7
Type here to search
20°C
ENG
00:56

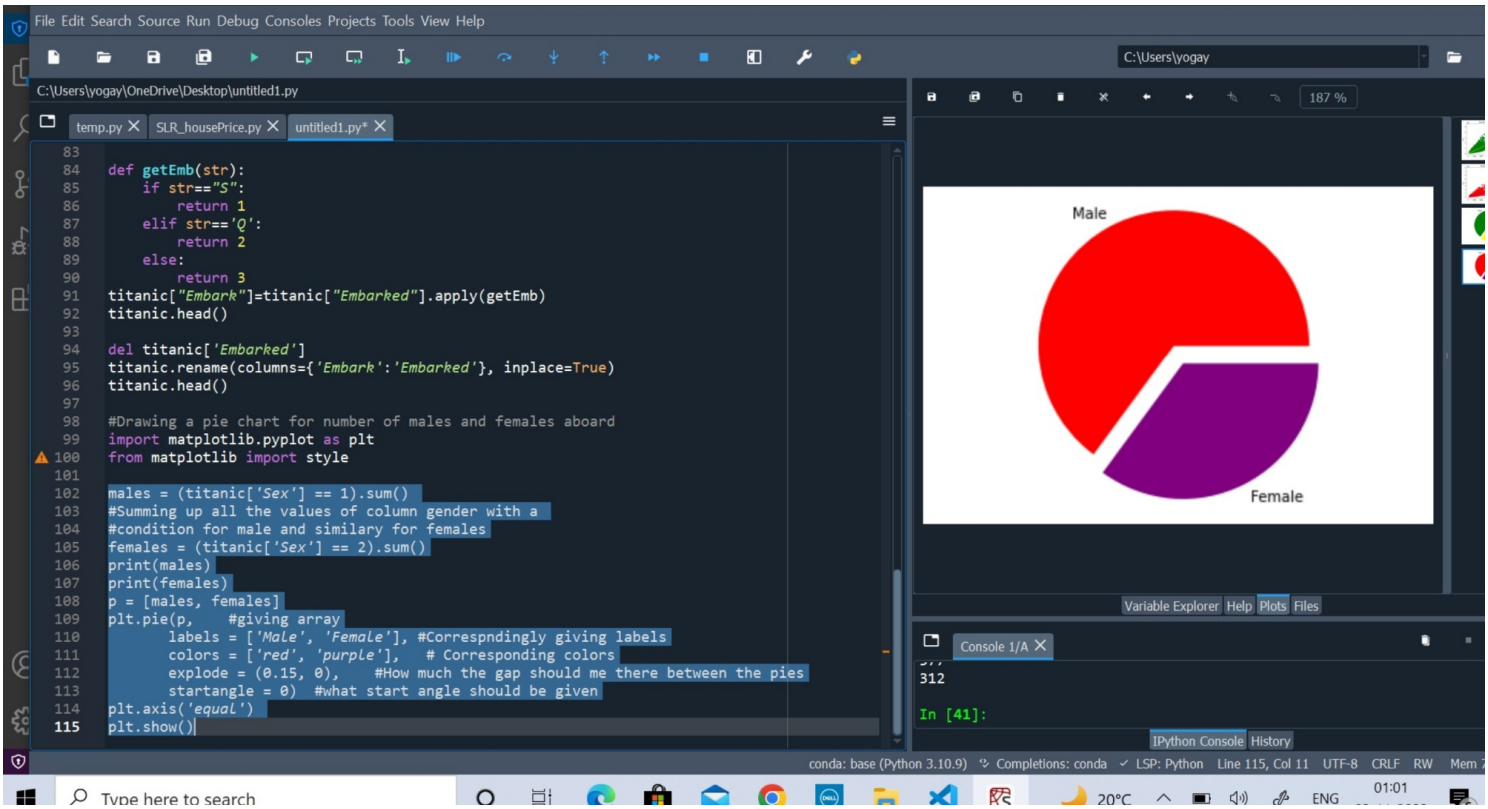

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
titanic.isnull().sum()
meanNS=titanic[titanic.Survived==0].Age.mean()
meanNS
titanic.age.fillna(meanNS,inplace=True)
titanic.head()
titanic.isnull().sum()
del titanic['Age']
titanic.head()
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 1].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 1].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 1].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 0].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 0].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 0].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
titanic.dropna(inplace=True)
titanic.head()
titanic.isnull().sum()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
meanNS
float64
1
30.62617924528
meanS
float64
1
28.34368965517
pred
Array of float64
(7205,)
[362473.73011
price
Series
(21613,)
Series object
Variable Explorer
Help
Plots
Files
Console 1/A
3 4 1 1 1 0 S 2
35.0
4 5 0 3 0 0 S 1
35.0
In [34]: titanic.isnull().sum()
Out[34]:
PassengerId 0
Survived 0
Pclass 0
SibSp 0
Parch 0
Embarked 0
Gender 0
age 0
dtype: int64
In [35]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 77, Col 1
UTF-8
CRLF
RW
Mem 8
Type here to search
20°C
ENG
00:57
03-11-2023

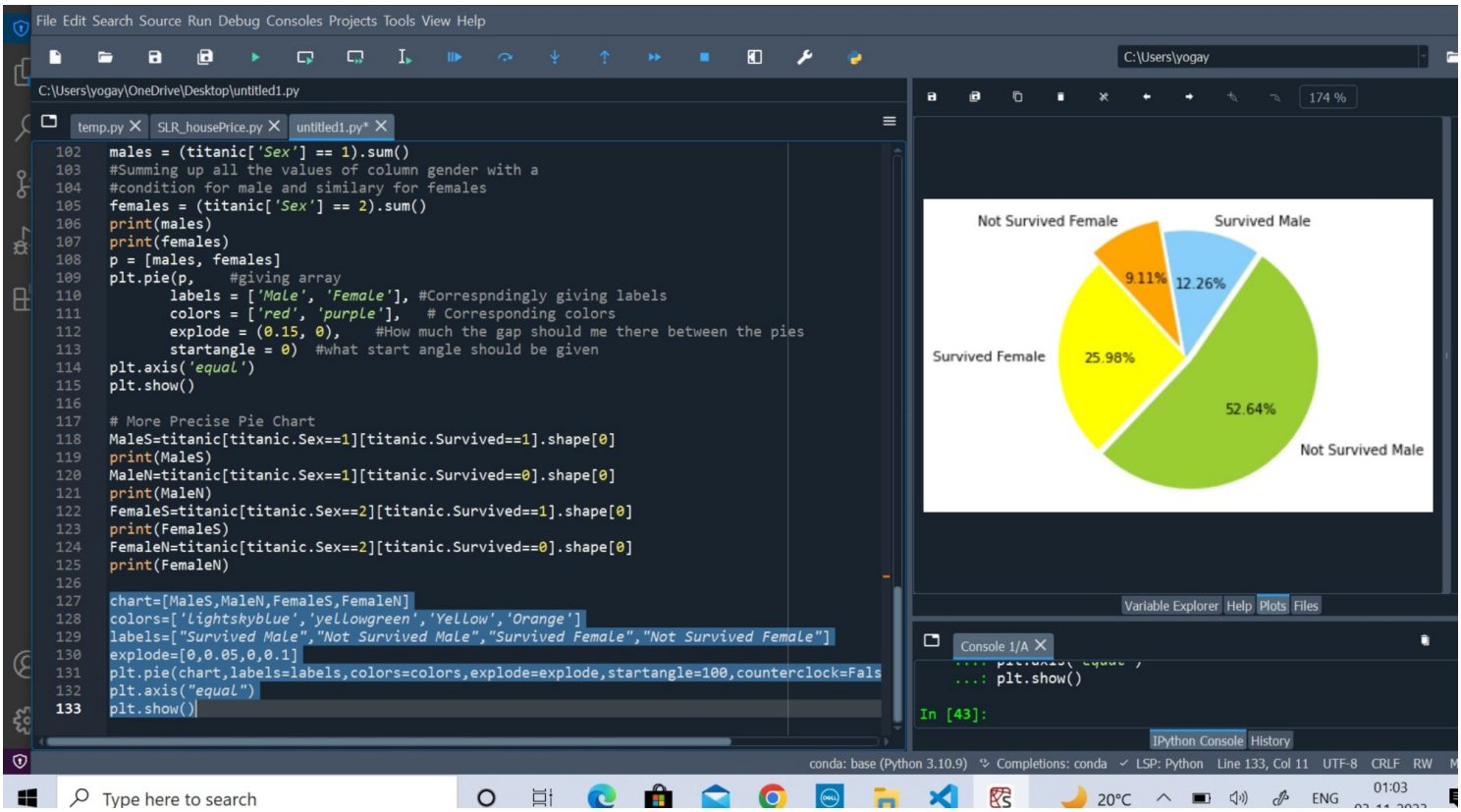

File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
meanNS=titanic[titanic.Survived==0].Age.mean()
meanNS
titanic.age.fillna(meanNS,inplace=True)
titanic.head()
titanic.isnull().sum()
del titanic['Age']
titanic.head()
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 1].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 1].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 1].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 0].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 0].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 0].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
titanic.dropna(inplace=True)
titanic.head()
titanic.isnull().sum()
titanic.rename(columns={'age':'Age'}, inplace=True)
titanic.head()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
meanNS
float64
1
30.62617924528
meanS
float64
1
28.34368965517
pred
Array of float64
(7205,)
[362473.7301
price
Series
(21613,)
Series object
Variable Explorer
Help
Plots
Files
Console 1/A
In [35]: titanic.rename(columns={'age':'Age'}, inplace=True)
...: titanic.head()
Out[35]:
PassengerId Survived Pclass SibSp Parch Embarked Gender
Age
0 1 0 3 1 0 S 1
22.0
1 2 1 1 1 0 C 2
38.0
2 3 1 3 0 0 S 2
26.0
3 4 1 1 1 0 S 2
35.0
4 5 0 3 0 0 S 1
35.0
In [36]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 79, Col 15
UTF-8
CRLF
RW
Type here to search
20°C
ENG
00:57


Spyder (Python 3.10)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
titanic.age.fillna(meanNS,inplace=True)
titanic.head()
titanic.isnull().sum()
del titanic['Age']
titanic.head()
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 1].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 1].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 1].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 0].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 0].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 0].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
titanic.dropna(inplace=True)
titanic.head()
titanic.isnull().sum()
titanic.rename(columns={'age':'Age'}, inplace=True)
titanic.head()
titanic.rename(columns={'Gender':'Sex'}, inplace=True)
titanic.head()
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
meanNS
float64
1
30.6261792452
meanS
float64
1
28.3436896551
pred
Array of float64
(7205,)
[362473.7301
price
Series
(21613,)
Series object
Variable Explorer
Help
Plots
Files
Console 1/A
2 3 1 3 0 0 S 2
26.0
3 4 1 1 1 0 S 2
35.0
4 5 0 3 0 0 S 1
35.0
In [36]: titanic.rename(columns={'Gender':'Sex'}, inplace=True)
...: titanic.head()
Out[36]:
PassengerId Survived Pclass SibSp Parch Embarked Sex Age
0 1 0 3 1 0 S 1 22.0
1 2 1 1 1 0 C 2 38.0
2 3 1 3 0 0 S 2 26.0
3 4 1 1 1 0 S 2 35.0
4 5 0 3 0 0 S 1 35.0
In [37]:
IPython Console
History
Run selection or current line
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 82, Col 15
UTF-8
CRLF
RW
Mem 8
Type here to search
20°C
ENG
00:58


File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 1].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 0].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 0].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 0].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
titanic.dropna(inplace=True)
titanic.head()
titanic.isnull().sum()
titanic.rename(columns={'age':'Age'}, inplace=True)
titanic.head()
titanic.rename(columns={'Gender':'Sex'}, inplace=True)
titanic.head()
def getEmb(str):
if str=="S":
return 1
elif str=='Q':
return 2
else:
return 3
titanic["Embark"]=titanic["Embarked"].apply(getEmb)
titanic.head()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
meanNS
float64
1
30.626179245
meanS
float64
1
28.343689655
pred
Array of float64
(7205,)
[362473.730
price
Series
(21613,)
Series object
Variable Explorer
Help
Plots
Files
Console 1/A
...: return 3
...: titanic["Embark"]=titanic["Embarked"].apply(getEmb)
...: titanic.head()
Out[37]:
PassengerId Survived Pclass SibSp Parch Embarked Sex Ag
Embark
0 1 0 3 1 0 S 1 22.
1
1 2 1 1 1 0 C 2 38.
3
2 3 1 3 0 0 S 2 26.
1
3 4 1 1 1 0 S 2 35.
1
4 5 0 3 0 0 S 1 35.
1
In [38]:
IPython Console
History
Run selection or current line
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 92, Col 15
UTF-8
CRLF
RW
Mem
Type here to search
20°C
ENG
00:59
03-11-2023


File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
survivedQ = titanic[titanic.Embarked == 'Q'][titanic.Survived == 0].shape[0]
survivedC = titanic[titanic.Embarked == 'C'][titanic.Survived == 0].shape[0]
survivedS = titanic[titanic.Embarked == 'S'][titanic.Survived == 0].shape[0]
print(survivedQ)
print(survivedC)
print(survivedS)
titanic.dropna(inplace=True)
titanic.head()
titanic.isnull().sum()
titanic.rename(columns={'age':'Age'}, inplace=True)
titanic.head()
titanic.rename(columns={'Gender':'Sex'}, inplace=True)
titanic.head()
def getEmb(str):
if str=="S":
return 1
elif str=='Q':
return 2
else:
return 3
titanic["Embark"]=titanic["Embarked"].apply(getEmb)
titanic.head()
del titanic['Embarked']
titanic.rename(columns={'Embark':'Embarked'}, inplace=True)
titanic.head()
C:\Users\yogay
Name
Type
Size
Value
dataset
DataFrame
(891, 12)
Column names:
meanNS
float64
1
30.62617924528
meanS
float64
1
28.34368965517
pred
Array of float64
(7205,)
[362473.7301
price
Series
(21613,)
Series object
Variable Explorer
Help
Plots
Files
Console 1/A
In [38]: del titanic['Embarked']
...: titanic.rename(columns={'Embark':'Embarked'}, inplace=Tru
...: titanic.head()
Out[38]:
PassengerId Survived Pclass SibSp Parch Sex Age Embarke
0 1 0 3 1 0 1 22.0
1 2 1 1 1 0 2 38.0
2 3 1 3 0 0 2 26.0
3 4 1 1 1 0 2 35.0
4 5 0 3 0 0 1 35.0
In [39]:
IPython Console
History
Run selection or current line
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 96, Col 15
UTF-8
CRLF
RW
Mem 7
Type here to search
20°C
ENG
00:59


File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
def getEmb(str):
if str=="S":
return 1
elif str=='Q':
return 2
else:
return 3
titanic["Embark"]=titanic["Embarked"].apply(getEmb)
titanic.head()
del titanic['Embarked']
titanic.rename(columns={'Embark':'Embarked'}, inplace=True)
titanic.head()
#Drawing a pie chart for number of males and females aboard
import matplotlib.pyplot as plt
from matplotlib import style
males = (titanic['Sex'] == 1).sum()
#Summing up all the values of column gender with a
#condition for male and similary for females
females = (titanic['Sex'] == 2).sum()
print(males)
print(females)
p = [males, females]
plt.pie(p, #giving array
labels = ['Male', 'Female'], #Correspndingly giving labels
colors = ['red', 'purple'], # Corresponding colors
explode = (0.15, 0), #How much the gap should me there between the pies
startangle = 0) #what start angle should be given
plt.axis('equal')
plt.show()
C:\Users\yogay
187 %
Male
Female
Variable Explorer
Help
Plots
Files
Console 1/A
312
In [41]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 115, Col 11
UTF-8
CRLF
RW
Type here to search
20°C
ENG
01:01


File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\yogay\OneDrive\Desktop\untitled1.py
temp.py
SLR_housePrice.py
untitled1.py*
males = (titanic['Sex'] == 1).sum()
#Summing up all the values of column gender with a
#condition for male and similary for females
females = (titanic['Sex'] == 2).sum()
print(males)
print(females)
p = [males, females]
plt.pie(p, #giving array
labels = ['Male', 'Female'], #Correspndingly giving labels
colors = ['red', 'purple'], # Corresponding colors
explode = (0.15, 0), #How much the gap should me there between the pies
startangle = 0) #what start angle should be given
plt.axis('equal')
plt.show()
# More Precise Pie Chart
MaleS=titanic[titanic.Sex==1][titanic.Survived==1].shape[0]
print(MaleS)
MaleN=titanic[titanic.Sex==1][titanic.Survived==0].shape[0]
print(MaleN)
FemaleS=titanic[titanic.Sex==2][titanic.Survived==1].shape[0]
print(FemaleS)
FemaleN=titanic[titanic.Sex==2][titanic.Survived==0].shape[0]
print(FemaleN)
chart=[MaleS,MaleN,FemaleS,FemaleN]
colors=['lightskyblue','yellowgreen','Yellow','Orange']
labels=["Survived Male","Not Survived Male","Survived Female","Not Survived Female"]
explode=[0,0.05,0,0.1]
plt.pie(chart,labels=labels,colors=colors,explode=explode,startangle=100,counterclock=Fals
plt.axis("equal")
plt.show()
C:\Users\yogay
174 %
Not Survived Female
Survived Male
9.11%
12.26%
Survived Female
25.98%
52.64%
Not Survived Male
Variable Explorer
Help
Plots
Files
Console 1/A
plt.show()
In [43]:
IPython Console
History
conda: base (Python 3.10.9)
Completions: conda
LSP: Python
Line 133, Col 11
UTF-8
CRLF
RW
Type here to search
20°C
ENG
01:03