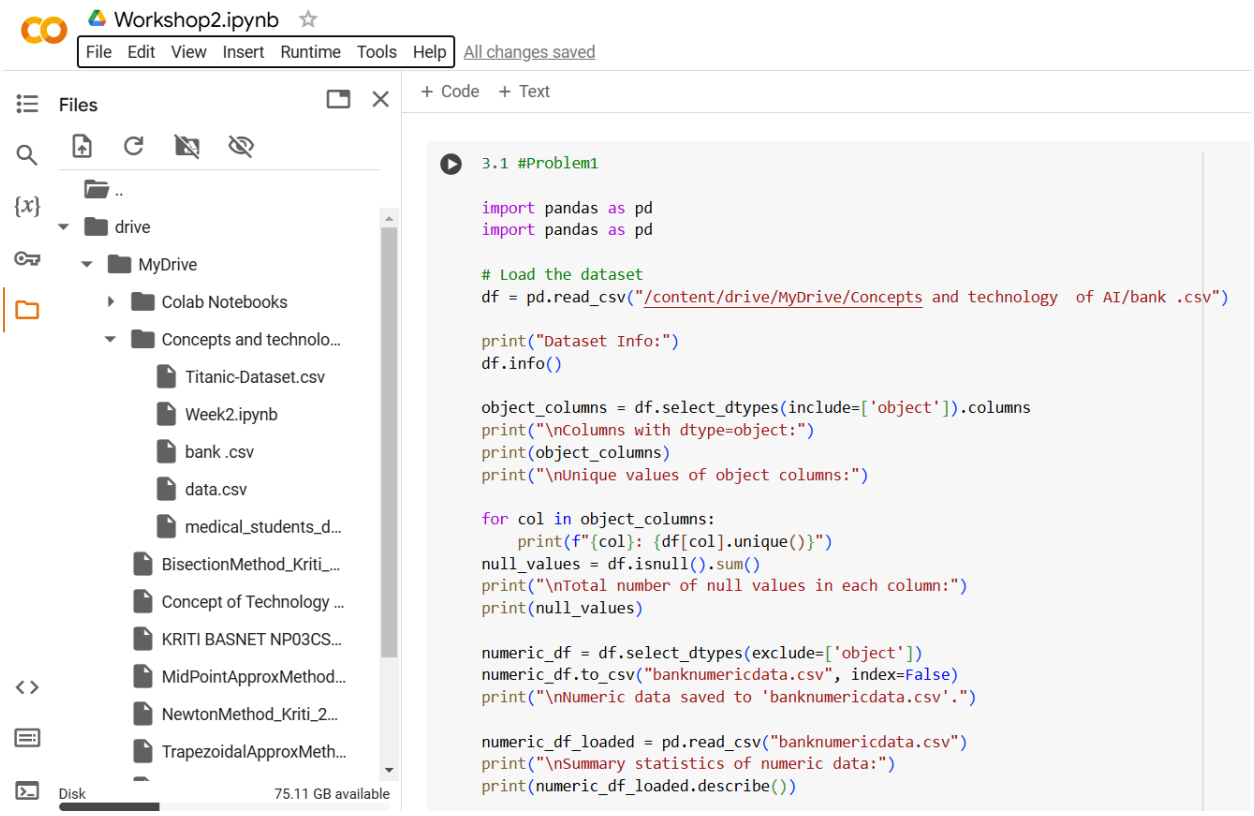
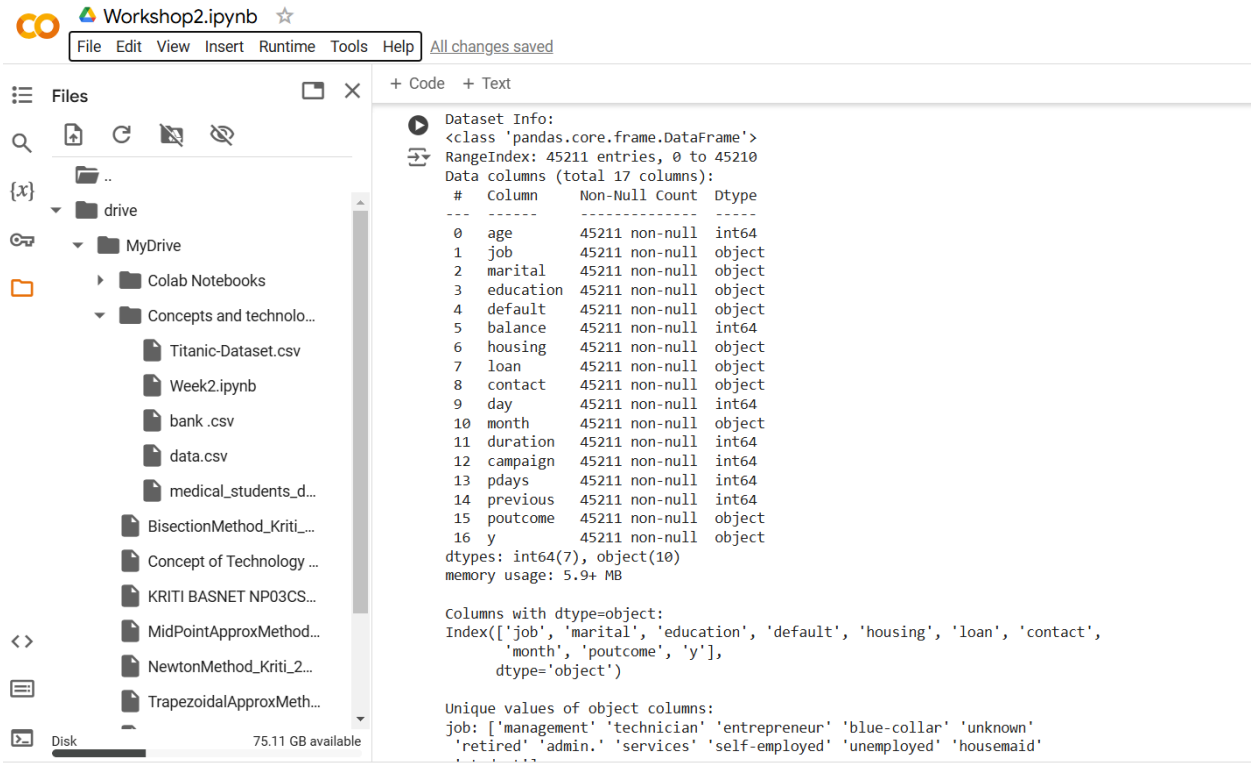
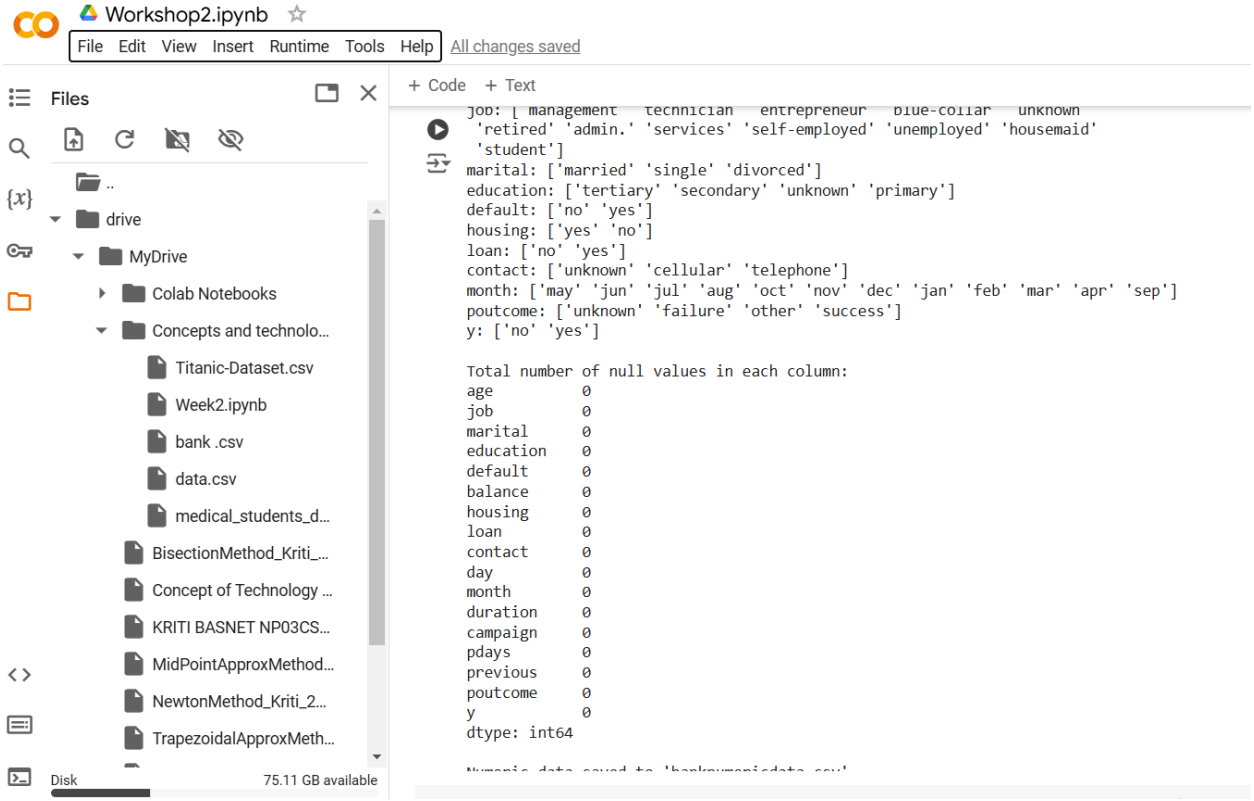
Concept and Technology of AI

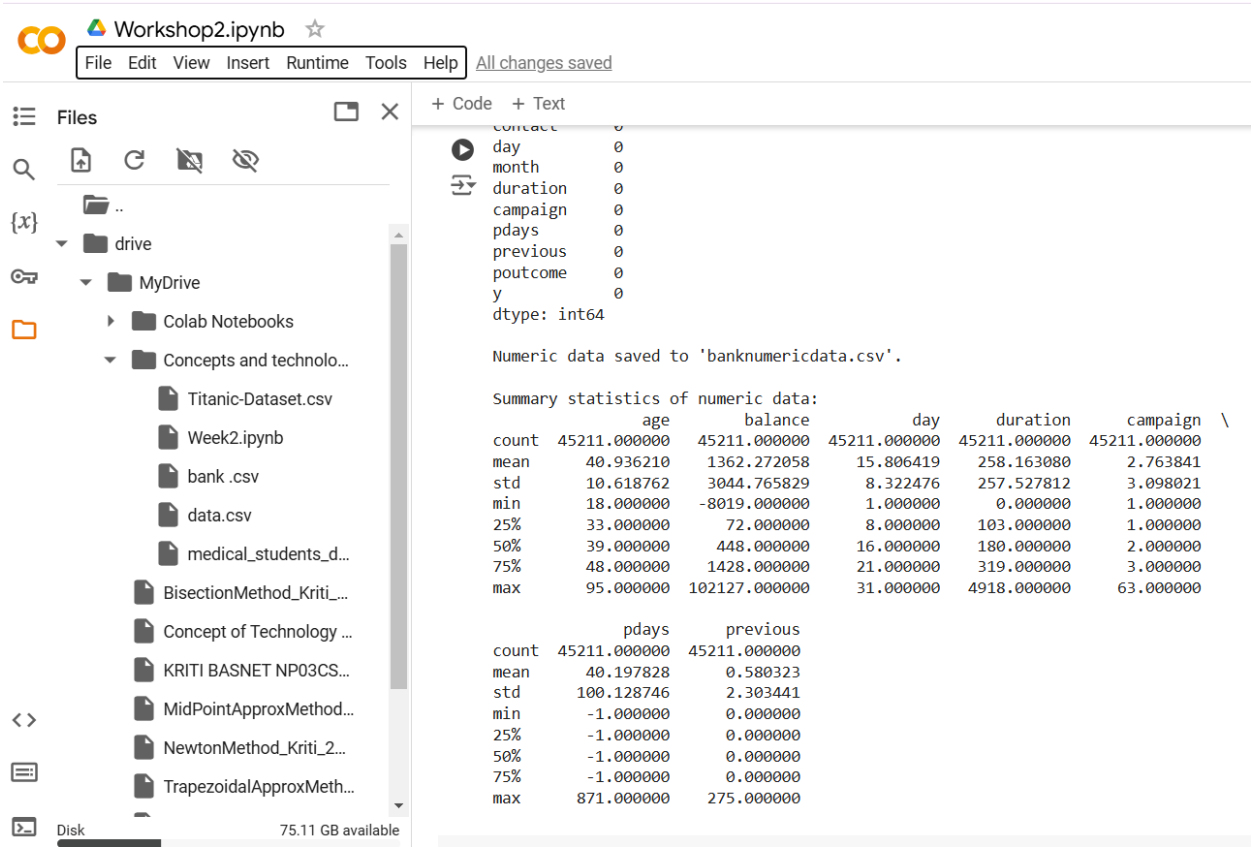
Workshop2

3.1 #Problem1

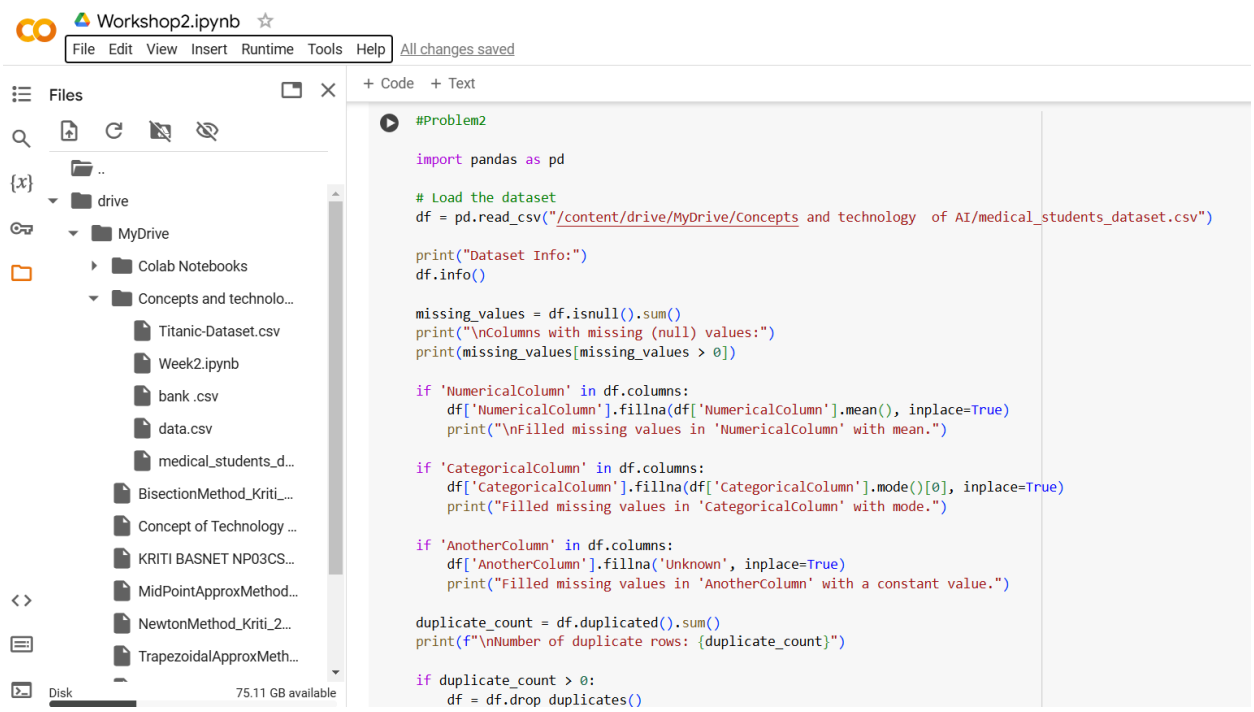


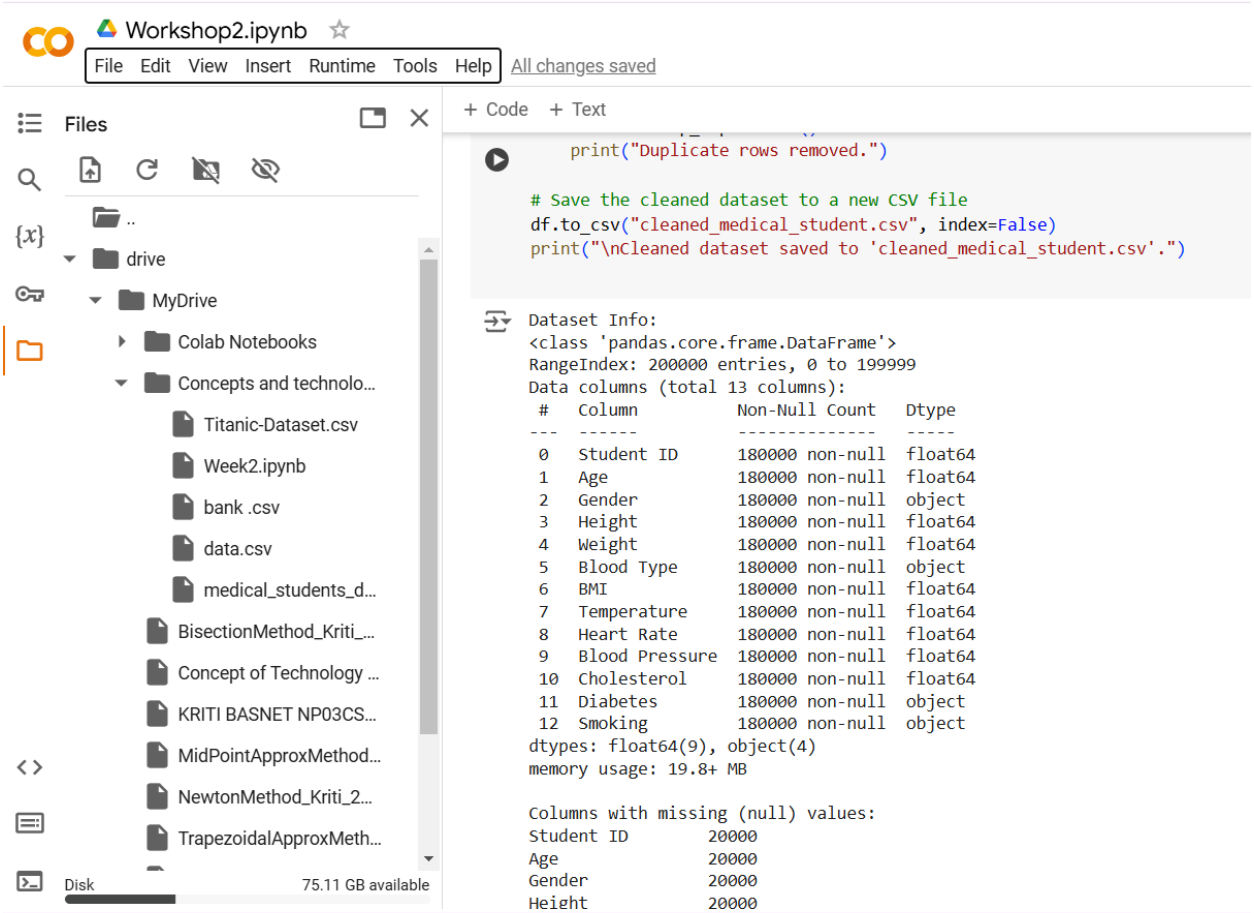

Workshop2.ipynb
File Edit View Insert Runtime Tools Help
All changes saved
Files
+ Code + Text
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk
75.11 GB available
Dataset Info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 45211 entries, 0 to 45210
Data columns (total 17 columns):
Column Non-Null Count Dtype
0 age 45211 non-null int64
1 job 45211 non-null object
2 marital 45211 non-null object
3 education 45211 non-null object
4 default 45211 non-null object
5 balance 45211 non-null int64
6 housing 45211 non-null object
7 loan 45211 non-null object
8 contact 45211 non-null object
9 day 45211 non-null int64
10 month 45211 non-null object
11 duration 45211 non-null int64
12 campaign 45211 non-null int64
13 pdays 45211 non-null int64
14 previous 45211 non-null int64
15 poutcome 45211 non-null object
16 y 45211 non-null object
dtypes: int64(7), object(10)
memory usage: 5.9+ MB
Columns with dtype=object:
Index(['job', 'marital', 'education', 'default', 'housing', 'loan', 'contact',
'month', 'poutcome', 'y'],
dtype='object')
Unique values of object columns:
job: ['management' 'technician' 'entrepreneur' 'blue-collar' 'unknown'
'retired' 'admin.' 'services' 'self-employed' 'unemployed' 'housemaid'

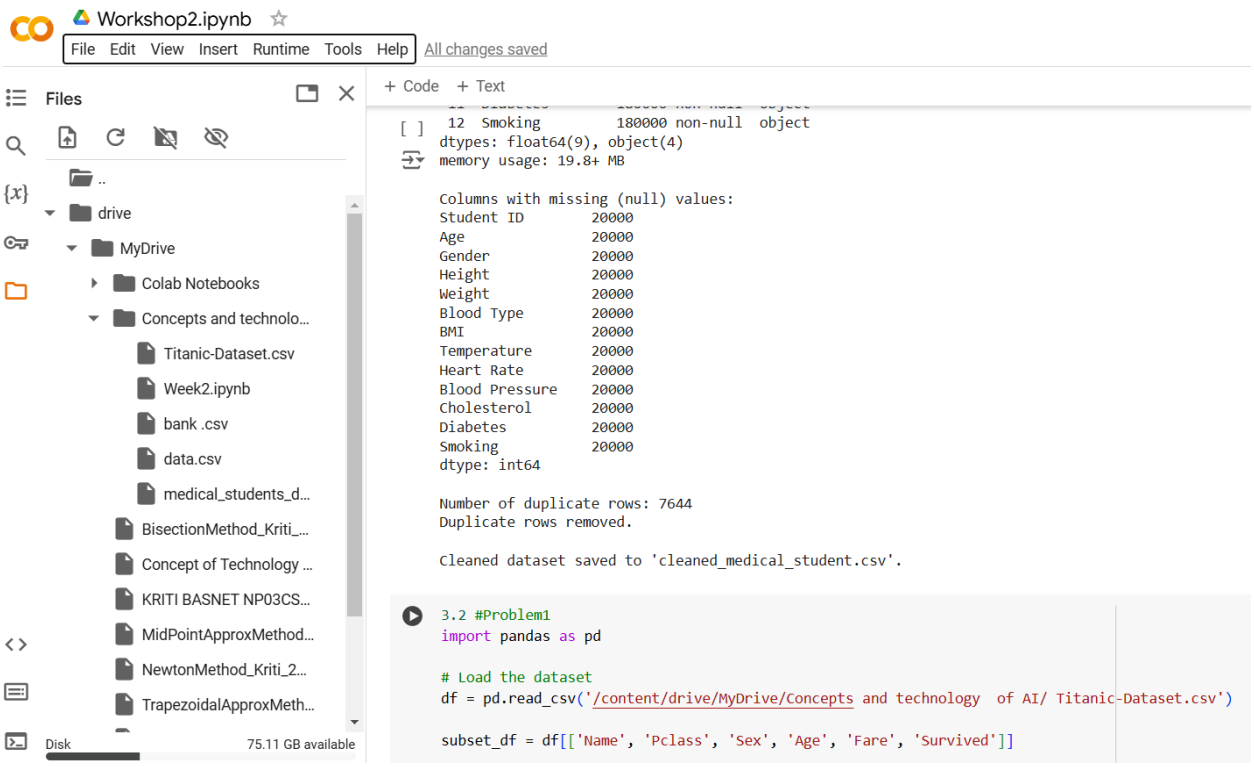

Workshop2.ipynb
File Edit View Insert Runtime Tools Help
All changes saved
Files
+ Code + Text
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk
75.11 GB available
'retired' 'admin.' 'services' 'self-employed' 'unemployed' 'housemaid'
'student']
marital: ['married' 'single' 'divorced']
education: ['tertiary' 'secondary' 'unknown' 'primary']
default: ['no' 'yes']
housing: ['yes' 'no']
loan: ['no' 'yes']
contact: ['unknown' 'cellular' 'telephone']
month: ['may' 'jun' 'jul' 'aug' 'oct' 'nov' 'dec' 'jan' 'feb' 'mar' 'apr' 'sep']
poutcome: ['unknown' 'failure' 'other' 'success']
y: ['no' 'yes']
Total number of null values in each column:
age 0
job 0
marital 0
education 0
default 0
balance 0
housing 0
loan 0
contact 0
day 0
month 0
duration 0
campaign 0
pdays 0
previous 0
poutcome 0
y 0
dtype: int64


Workshop2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk 75.11 GB available
day 0
month 0
duration 0
campaign 0
pdays 0
previous 0
poutcome 0
y 0
dtype: int64
Numeric data saved to 'banknumericdata.csv'.
Summary statistics of numeric data:
age balance day duration campaign \
count 45211.000000 45211.000000 45211.000000 45211.000000 45211.000000
mean 40.936210 1362.272058 15.806419 258.163080 2.763841
std 10.618762 3044.765829 8.322476 257.527812 3.098021
min 18.000000 -8019.000000 1.000000 0.000000 1.000000
25% 33.000000 72.000000 8.000000 103.000000 1.000000
50% 39.000000 448.000000 16.000000 180.000000 2.000000
75% 48.000000 1428.000000 21.000000 319.000000 3.000000
max 95.000000 102127.000000 31.000000 4918.000000 63.000000
pdays previous
count 45211.000000 45211.000000
mean 40.197828 0.580323
std 100.128746 2.303441
min -1.000000 0.000000
25% -1.000000 0.000000
50% -1.000000 0.000000
75% -1.000000 0.000000
max 871.000000 275.000000

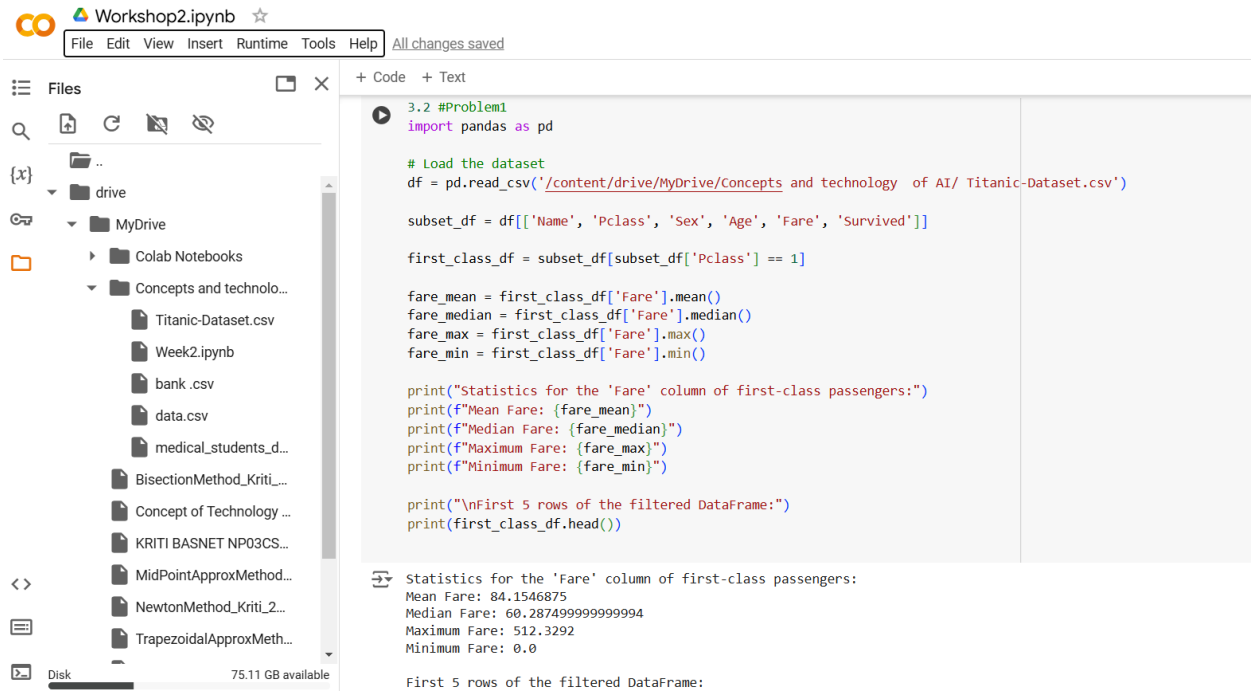
#Problem2

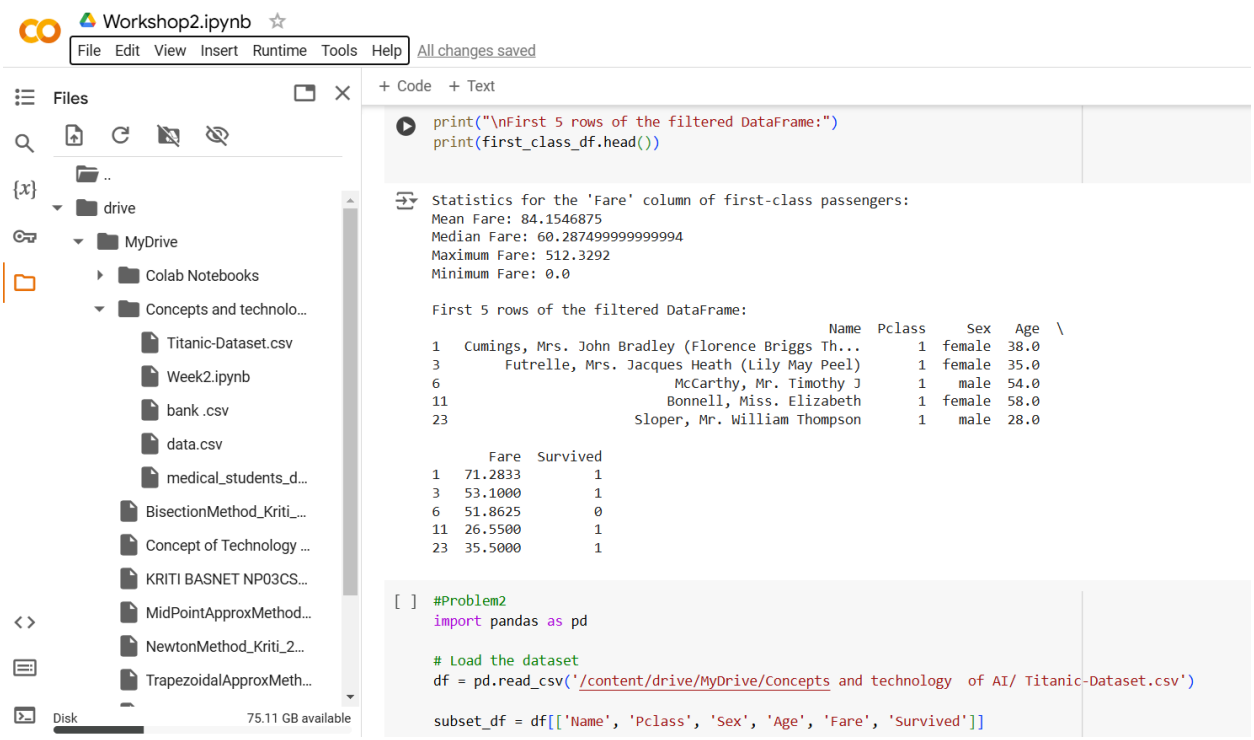

Workshop2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk 75.11 GB available
#Problem2
import pandas as pd
# Load the dataset
df = pd.read_csv("/content/drive/MyDrive/Concepts and technology of AI/medical_students_dataset.csv")
print("Dataset Info:")
df.info()
missing_values = df.isnull().sum()
print("\nColumns with missing (null) values:")
print(missing_values[missing_values > 0])
if 'NumericalColumn' in df.columns:
df['NumericalColumn'].fillna(df['NumericalColumn'].mean(), inplace=True)
print("\nFilled missing values in 'NumericalColumn' with mean.")
if 'CategoricalColumn' in df.columns:
df['CategoricalColumn'].fillna(df['CategoricalColumn'].mode()[0], inplace=True)
print("Filled missing values in 'CategoricalColumn' with mode.")
if 'AnotherColumn' in df.columns:
df['AnotherColumn'].fillna('Unknown', inplace=True)
print("Filled missing values in 'AnotherColumn' with a constant value.")
duplicate_count = df.duplicated().sum()
print(f"\nNumber of duplicate rows: {duplicate_count}")
if duplicate_count > 0:
df = df.drop_duplicates()


Workshop2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
print("Duplicate rows removed.")
# Save the cleaned dataset to a new CSV file
df.to_csv("cleaned_medical_student.csv", index=False)
print("\nCleaned dataset saved to 'cleaned_medical_student.csv'.")
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk 75.11 GB available
Dataset Info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200000 entries, 0 to 199999
Data columns (total 13 columns):
Column Non-Null Count Dtype
0 Student ID 180000 non-null float64
1 Age 180000 non-null float64
2 Gender 180000 non-null object
3 Height 180000 non-null float64
4 Weight 180000 non-null float64
5 Blood Type 180000 non-null object
6 BMI 180000 non-null float64
7 Temperature 180000 non-null float64
8 Heart Rate 180000 non-null float64
9 Blood Pressure 180000 non-null float64
10 Cholesterol 180000 non-null float64
11 Diabetes 180000 non-null object
12 Smoking 180000 non-null object
dtypes: float64(9), object(4)
memory usage: 19.8+ MB
Columns with missing (null) values:
Student ID 20000
Age 20000
Gender 20000
Height 20000

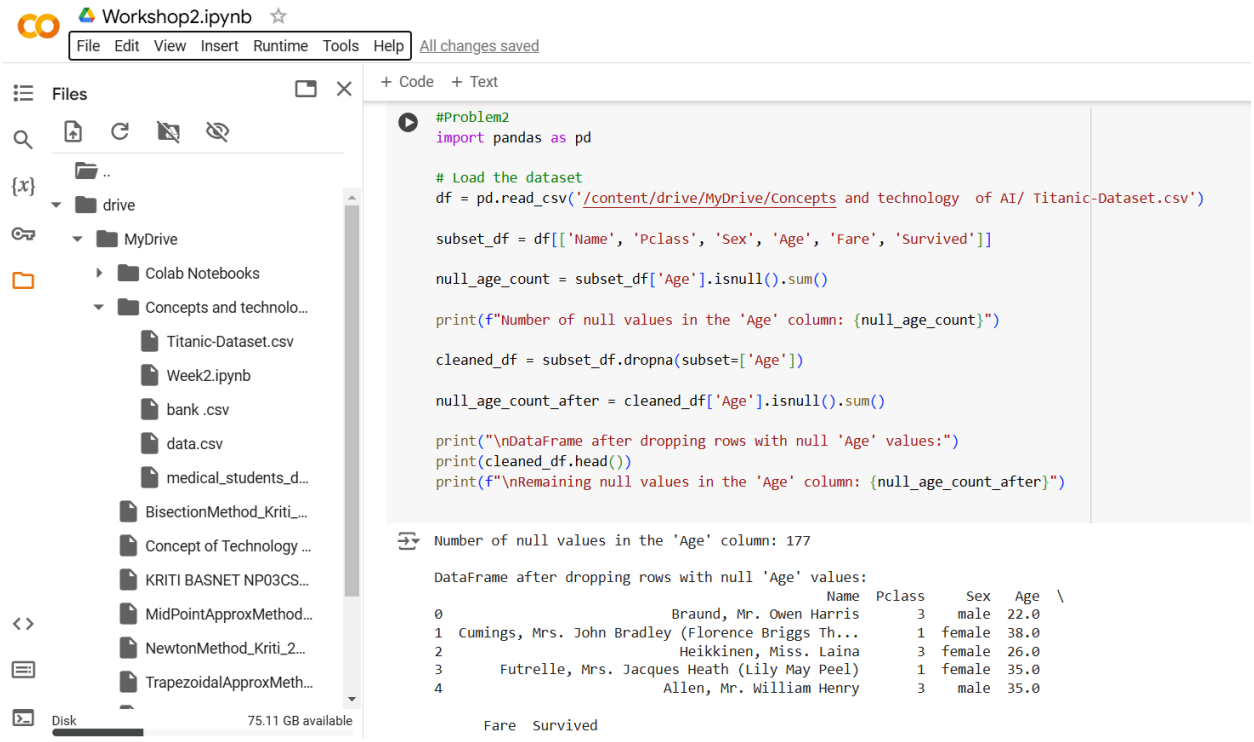

Workshop2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
12 Smoking 180000 non-null object
dtypes: float64(9), object(4)
memory usage: 19.8+ MB
Columns with missing (null) values:
Student ID 20000
Age 20000
Gender 20000
Height 20000
Weight 20000
Blood Type 20000
BMI 20000
Temperature 20000
Heart Rate 20000
Blood Pressure 20000
Cholesterol 20000
Diabetes 20000
Smoking 20000
dtype: int64
Number of duplicate rows: 7644
Duplicate rows removed.
Cleaned dataset saved to 'cleaned_medical_student.csv'.
3.2 #Problem1
import pandas as pd
# Load the dataset
df = pd.read_csv('/content/drive/MyDrive/Concepts and technology of AI/ Titanic-Dataset.csv')
subset_df = df[['Name', 'Pclass', 'Sex', 'Age', 'Fare', 'Survived']]

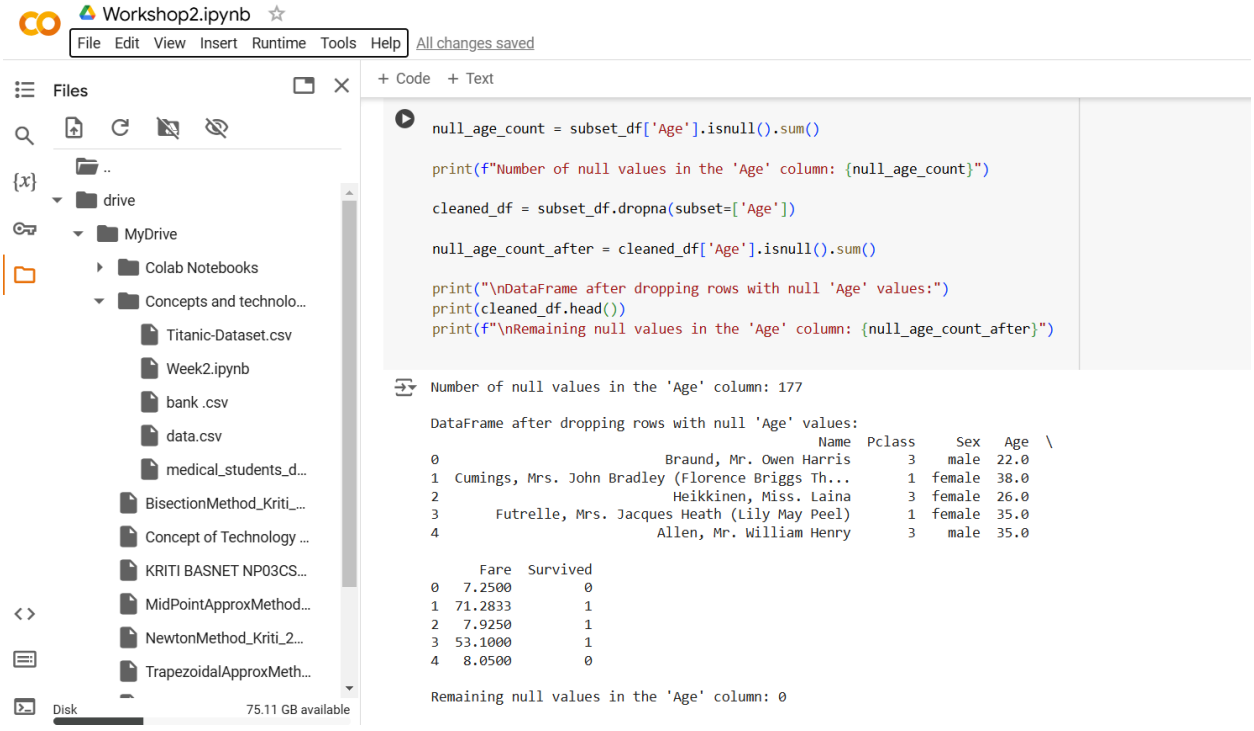
3.2 #Problem1



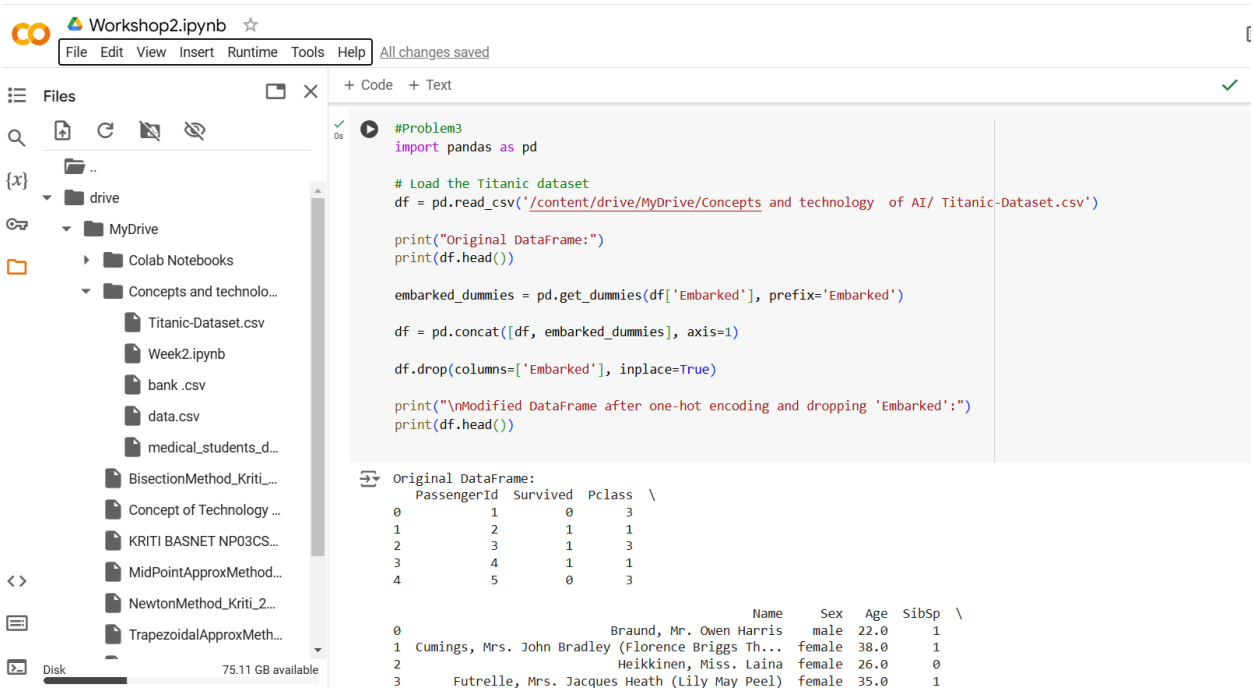


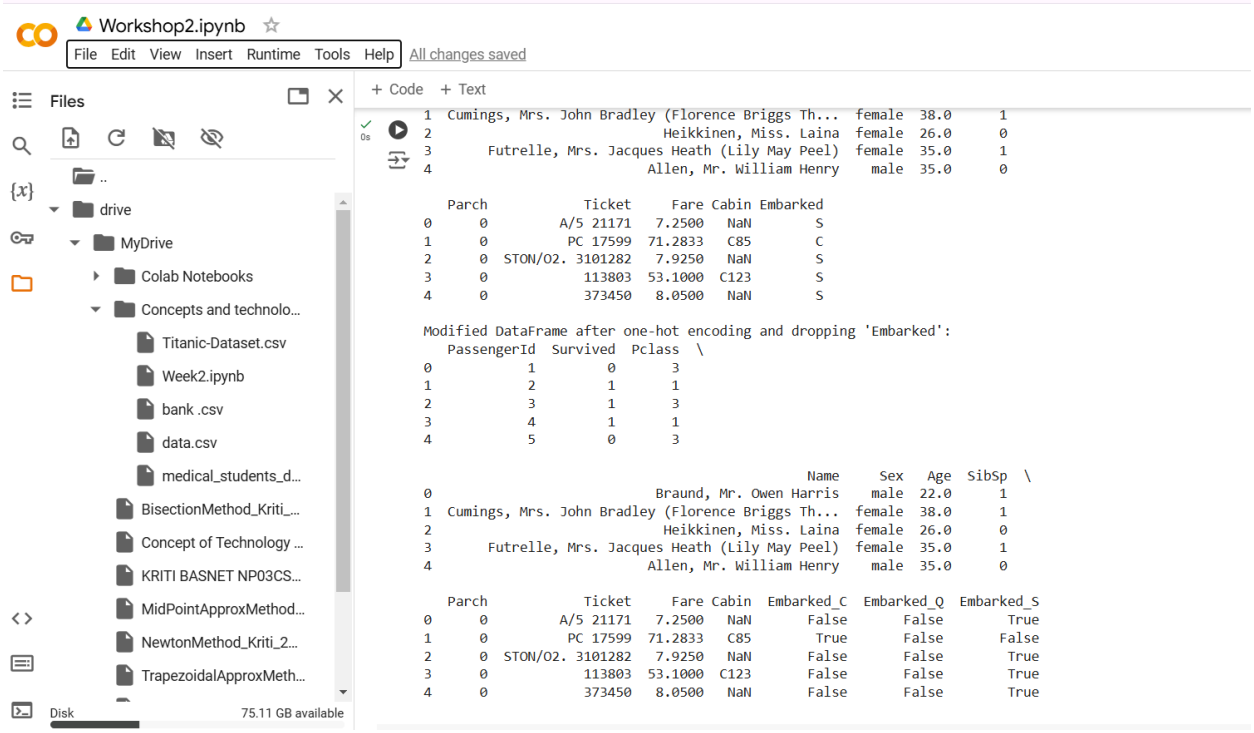
#Problem2


Workshop2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk 75.11 GB available
#Problem2
import pandas as pd
# Load the dataset
df = pd.read_csv('/content/drive/MyDrive/Concepts and technology of AI/ Titanic-Dataset.csv')
subset_df = df[['Name', 'Pclass', 'Sex', 'Age', 'Fare', 'Survived']]
null_age_count = subset_df['Age'].isnull().sum()
print(f"Number of null values in the 'Age' column: {null_age_count}")
cleaned_df = subset_df.dropna(subset=['Age'])
null_age_count_after = cleaned_df['Age'].isnull().sum()
print("\nDataFrame after dropping rows with null 'Age' values:")
print(cleaned_df.head())
print(f"\nRemaining null values in the 'Age' column: {null_age_count_after}")
Number of null values in the 'Age' column: 177
DataFrame after dropping rows with null 'Age' values:
Name Pclass Sex Age \
0 Braund, Mr. Owen Harris 3 male 22.0
1 Cumings, Mrs. John Bradley (Florence Briggs Th... 1 female 38.0
2 Heikkinen, Miss. Laina 3 female 26.0
3 Futrelle, Mrs. Jacques Heath (Lily May Peel) 1 female 35.0
4 Allen, Mr. William Henry 3 male 35.0
Fare Survived


Workshop2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk 75.11 GB available
null_age_count = subset_df['Age'].isnull().sum()
print(f"Number of null values in the 'Age' column: {null_age_count}")
cleaned_df = subset_df.dropna(subset=['Age'])
null_age_count_after = cleaned_df['Age'].isnull().sum()
print("\nDataFrame after dropping rows with null 'Age' values:")
print(cleaned_df.head())
print(f"\nRemaining null values in the 'Age' column: {null_age_count_after}")
Number of null values in the 'Age' column: 177
DataFrame after dropping rows with null 'Age' values:
Name Pclass Sex Age \
0 Braund, Mr. Owen Harris 3 male 22.0
1 Cumings, Mrs. John Bradley (Florence Briggs Th... 1 female 38.0
2 Heikkinen, Miss. Laina 3 female 26.0
3 Futrelle, Mrs. Jacques Heath (Lily May Peel) 1 female 35.0
4 Allen, Mr. William Henry 3 male 35.0
Fare Survived
0 7.2500 0
1 71.2833 1
2 7.9250 1
3 53.1000 1
4 8.0500 0
Remaining null values in the 'Age' column: 0

#Problem3


Workshop2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk 75.11 GB available
#Problem3
import pandas as pd
# Load the Titanic dataset
df = pd.read_csv('/content/drive/MyDrive/Concepts and technology of AI/ Titanic-Dataset.csv')
print("Original DataFrame:")
print(df.head())
embarked_dummies = pd.get_dummies(df['Embarked'], prefix='Embarked')
df = pd.concat([df, embarked_dummies], axis=1)
df.drop(columns=['Embarked'], inplace=True)
print("\nModified DataFrame after one-hot encoding and dropping 'Embarked':")
print(df.head())
Original DataFrame:
PassengerId Survived Pclass \
0 1 0 3
1 2 1 1
2 3 1 3
3 4 1 1
4 5 0 3
Name Sex Age SibSp \
0 Braund, Mr. Owen Harris male 22.0 1
1 Cumings, Mrs. John Bradley (Florence Briggs Th... female 38.0 1
2 Heikkinen, Miss. Laina female 26.0 0
3 Futrelle, Mrs. Jacques Heath (Lily May Peel) female 35.0 1


Workshop2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Files
+ Code + Text
drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...
Disk 75.11 GB available
1 Cumings, Mrs. John Bradley (Florence Briggs Th... female 38.0 1
2 Heikkinen, Miss. Laina female 26.0 0
3 Futrelle, Mrs. Jacques Heath (Lily May Peel) female 35.0 1
4 Allen, Mr. William Henry male 35.0 0
Parch Ticket Fare Cabin Embarked
0 0 A/5 21171 7.2500 NaN S
1 0 PC 17599 71.2833 C85 C
2 0 STON/O2. 3101282 7.9250 NaN S
3 0 113803 53.1000 C123 S
4 0 373450 8.0500 NaN S
Modified DataFrame after one-hot encoding and dropping 'Embarked':
PassengerId Survived Pclass \
0 1 0 3
1 2 1 1
2 3 1 3
3 4 1 1
4 5 0 3
Name Sex Age SibSp \
0 Braund, Mr. Owen Harris male 22.0 1
1 Cumings, Mrs. John Bradley (Florence Briggs Th... female 38.0 1
2 Heikkinen, Miss. Laina female 26.0 0
3 Futrelle, Mrs. Jacques Heath (Lily May Peel) female 35.0 1
4 Allen, Mr. William Henry male 35.0 0
Parch Ticket Fare Cabin Embarked_C Embarked_Q Embarked_S
0 0 A/5 21171 7.2500 NaN False False True
1 0 PC 17599 71.2833 C85 True False False
2 0 STON/O2. 3101282 7.9250 NaN False False True
3 0 113803 53.1000 C123 False False True
4 0 373450 8.0500 NaN False False True

#Problem4

Workshop2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

drive

MyDrive

Colab Notebooks

Concepts and technolo...

Titanic-Dataset.csv

Week2.ipynb

bank .csv

data.csv

medical_students_d...

BisectionMethod_Kriti_...

Concept of Technology ...

KRITI BASNET NP03CS...

MidPointApproxMethod...

NewtonMethod_Kriti_2...

TrapezoidalApproxMeth...

Disk 75.11 GB available

+ Code + Text

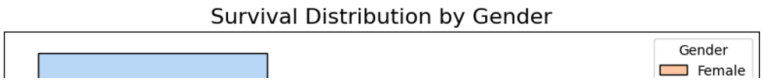
```
#Problem4
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load the Titanic dataset
df = pd.read_csv('/content/drive/MyDrive/Concepts and technology  of AI/ Titanic-Dataset.csv')

mean_survival_rates = df.groupby('Sex')['Survived'].mean()
print("Mean Survival Rates by Gender:")
print(mean_survival_rates)

plt.figure(figsize=(10, 6))
sns.histplot(data=df, x='Survived', hue='Sex', multiple='stack', bins=3, kde=False, palette='pastel')
plt.title('Survival Distribution by Gender', fontsize=16)
plt.xlabel('Survived (0 = No, 1 = Yes)', fontsize=12)
plt.ylabel('Count', fontsize=12)
plt.legend(title='Gender', labels=['Female', 'Male'])
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```

```
Mean Survival Rates by Gender:
Sex
female    0.742038
male      0.188908
Name: Survived, dtype: float64
```



Workshop2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files

drive

MyDrive

Colab Notebooks

Concepts and technolo...

Titanic-Dataset.csv

Week2.ipynb

bank .csv

data.csv

medical_students_d...

BisectionMethod_Kriti_...

Concept of Technology ...

KRITI BASNET NP03CS...

MidPointApproxMethod...

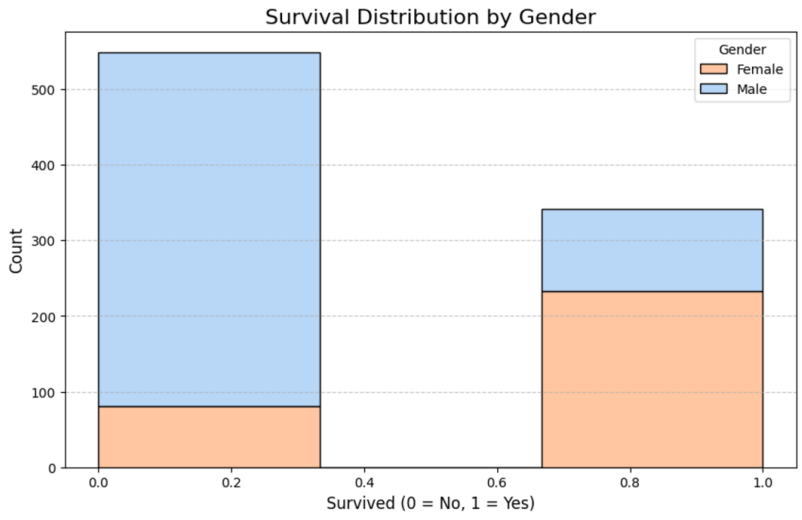
NewtonMethod_Kriti_2...

TrapezoidalApproxMeth...

+ Code + Text

RAM Disk

```
Name: Survived, dtype: float64
```



#Problem5

Workshop2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Share

Files

+ Code + Text

RAM Disk

Gemini

```
#Problem5
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load the Titanic dataset
df = pd.read_csv('/content/drive/MyDrive/Concepts and technology  of AI/ Titanic-Dataset.csv')

df = df.dropna(subset=['Embarked', 'Survived'])

plt.figure(figsize=(12, 6))
sns.histplot(
    data=df,
    x='Survived',
    hue='Embarked',
    multiple='stack',
    bins=3,
    kde=False,
    palette='Set2'
)
plt.title('Survival Distribution by Port of Embarkation', fontsize=16)
plt.xlabel('Survived (0 = No, 1 = Yes)', fontsize=12)
plt.ylabel('Count', fontsize=12)
plt.legend(title='Port of Embarkation', labels=['Cherbourg (C)', 'Queenstown (Q)', 'Southampton (S)'])
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```

drive
MyDrive
Colab Notebooks
Concepts and technolo...
Titanic-Dataset.csv
Week2.ipynb
bank .csv
data.csv
medical_students_d...
BisectionMethod_Kriti_...
Concept of Technology ...
KRITI BASNET NP03CS...
MidPointApproxMethod...
NewtonMethod_Kriti_2...
TrapezoidalApproxMeth...

Disk 75.11 GB available

Survival Distribution by Port of Embarkation