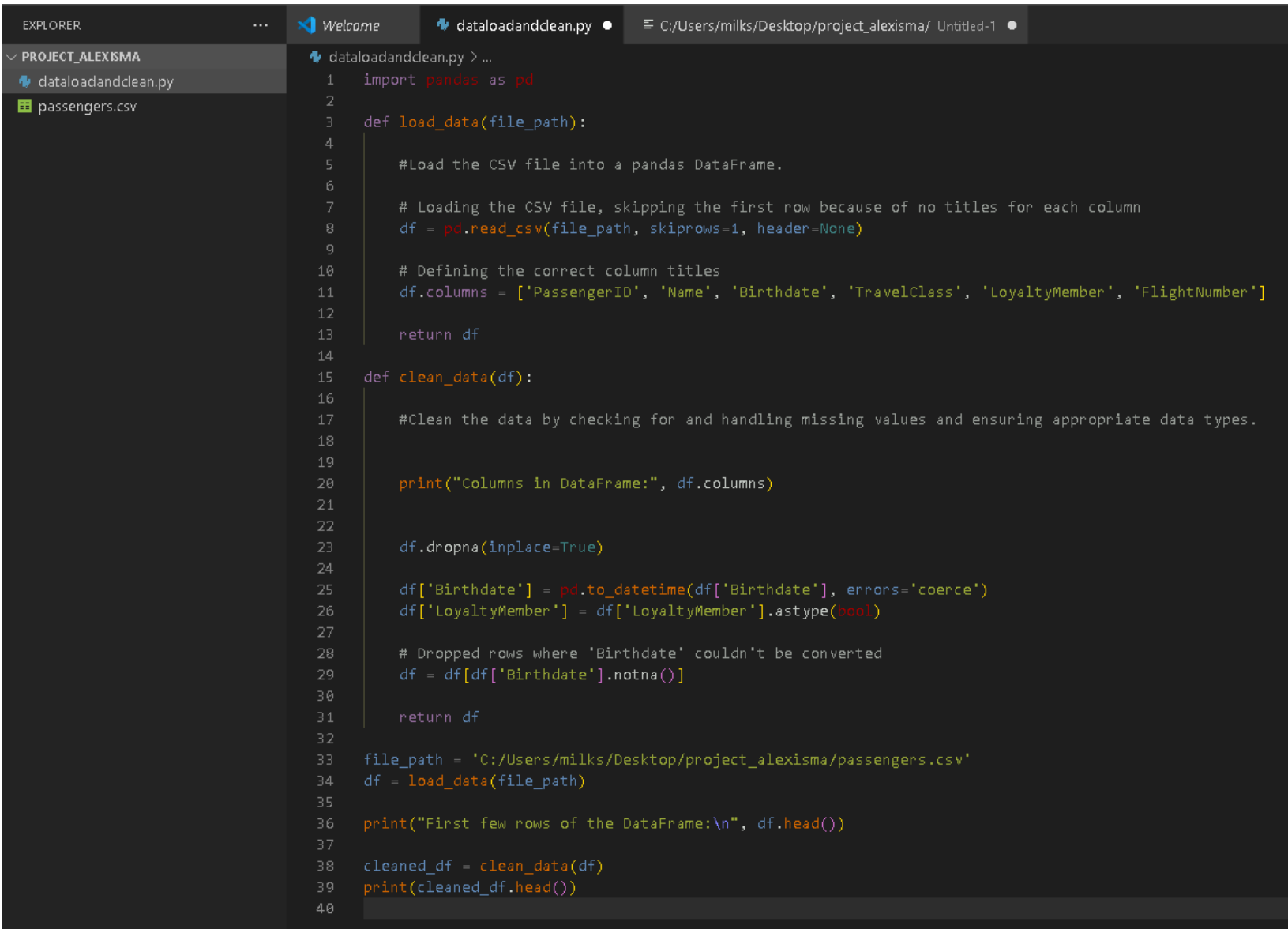
Marly Alexis

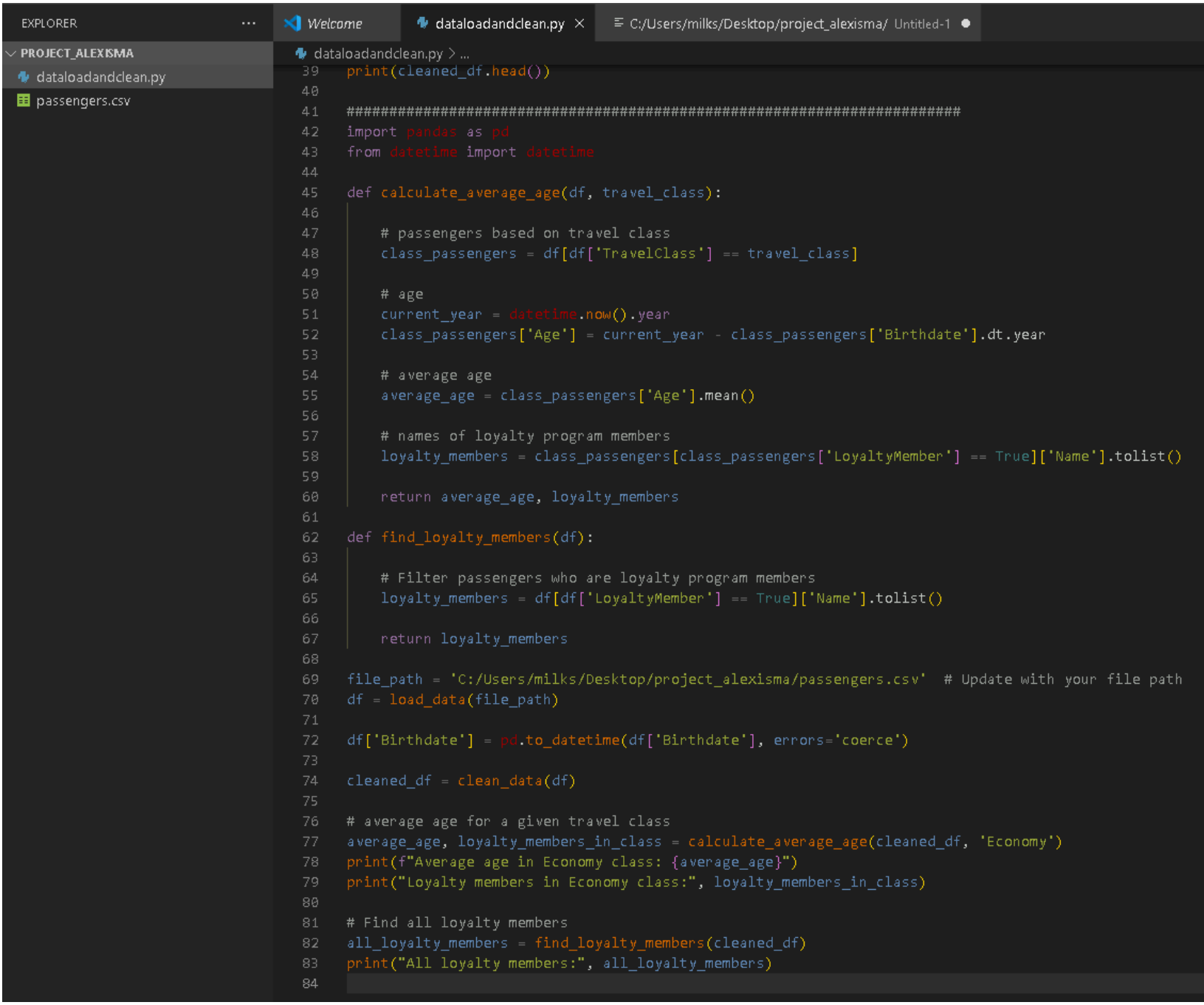
TASK 1:



PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

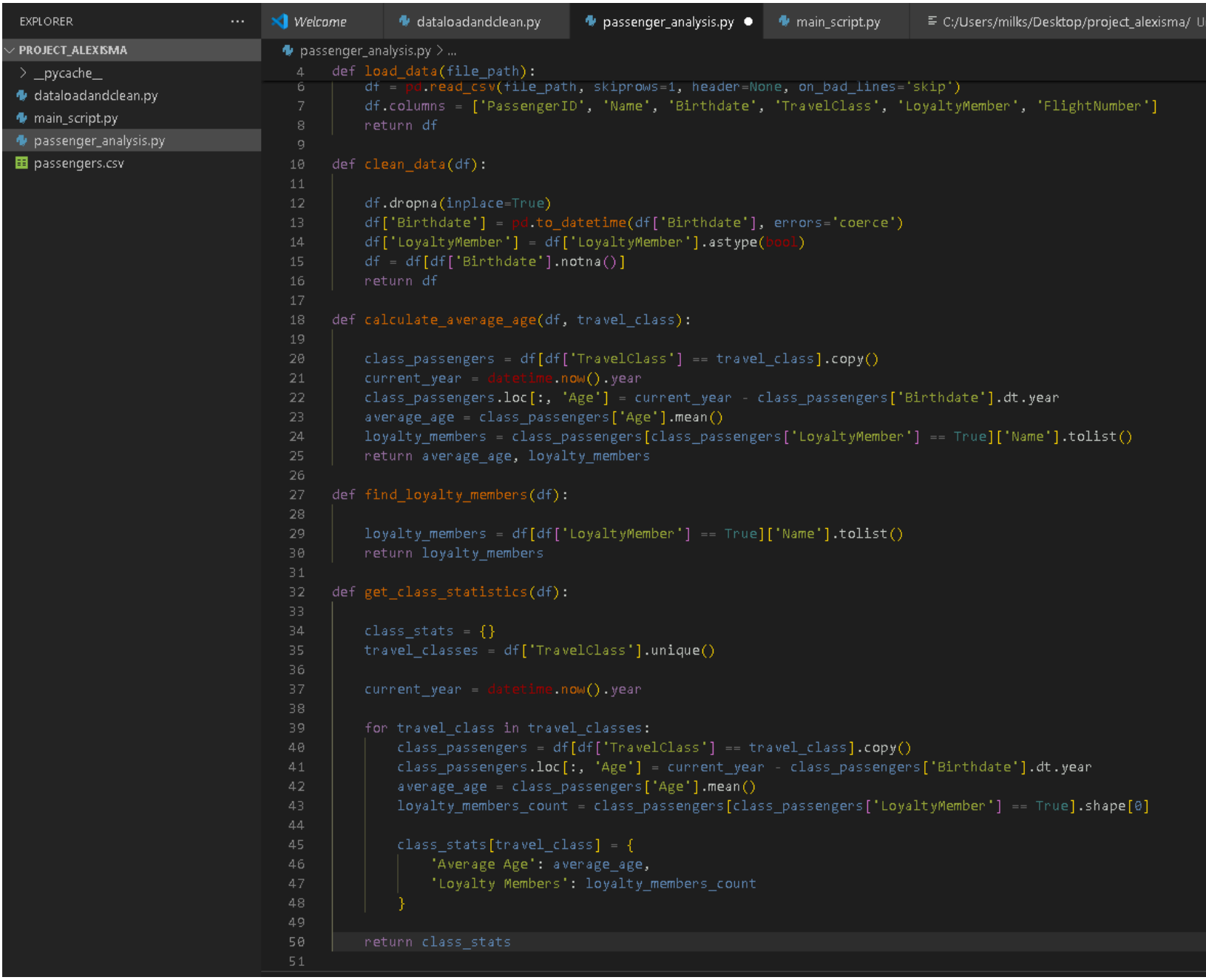
```
PS C:\Users\milks\Desktop\project_alexisma> python dataloadandclean.py
First few rows of the DataFrame:
    PassengerID            Name   Birthdate  TravelClass  LoyaltyMember FlightNumber
0             1        John Doe   5/21/1987  FIRST_CLASS           True        BA249
1             2      Jane Smith  11/12/1980     BUSINESS          False        AA100
2             3        Mia Wong    3/8/1992      ECONOMY           True        BA255
3             4    Noah Johnson   7/19/1995      ECONOMY          False        AA110
4             5  Isabella Rossi   8/30/1982  FIRST_CLASS           True        BA249
Columns in DataFrame: Index(['PassengerID', 'Name', 'Birthdate', 'TravelClass', 'LoyaltyMember',
       'FlightNumber'],
      dtype='object')
   PassengerID            Name  Birthdate  TravelClass  LoyaltyMember FlightNumber
0             1        John Doe 1987-05-21  FIRST_CLASS           True        BA249
1             2      Jane Smith 1980-11-12     BUSINESS          False        AA100
2             3        Mia Wong 1992-03-08      ECONOMY           True        BA255
3             4    Noah Johnson 1995-07-19      ECONOMY          False        AA110
4             5  Isabella Rossi 1982-08-30  FIRST_CLASS           True        BA249
PS C:\Users\milks\Desktop\project_alexisma>
```

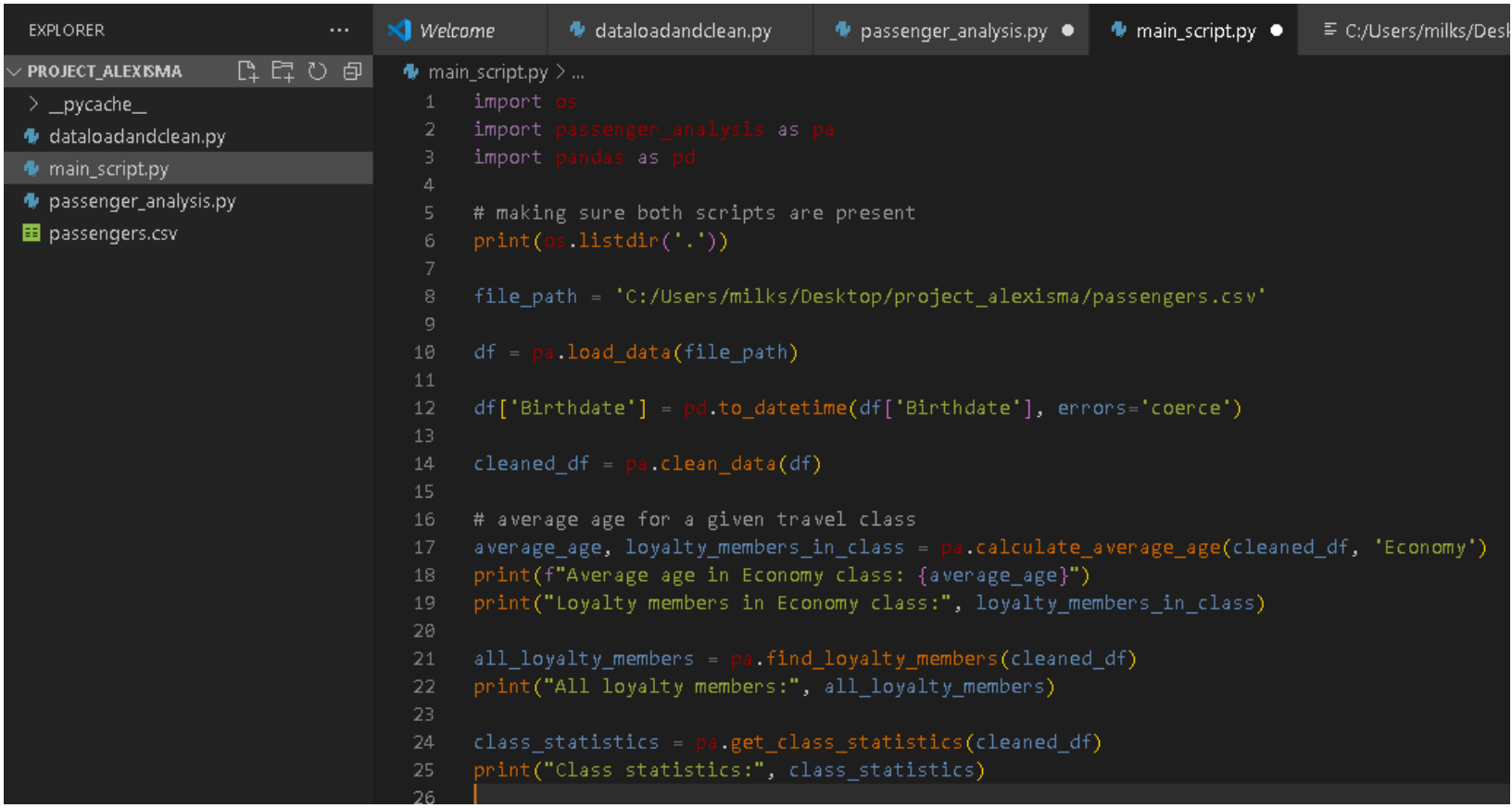
TASK 2:


EXPLORER
Welcome
dataloadandclean.py
C:/Users/milks/Desktop/project_alexisma/ Untitled-1
PROJECT_ALEXISMA
dataloadandclean.py
passengers.csv
dataloadandclean.py > ...
39 print(cleaned_df.head())
40
41 ##
42 import pandas as pd
43 from datetime import datetime
44
45 def calculate_average_age(df, travel_class):
46
47 # passengers based on travel class
48 class_passengers = df[df['TravelClass'] == travel_class]
49
50 # age
51 current_year = datetime.now().year
52 class_passengers['Age'] = current_year - class_passengers['Birthdate'].dt.year
53
54 # average age
55 average_age = class_passengers['Age'].mean()
56
57 # names of loyalty program members
58 loyalty_members = class_passengers[class_passengers['LoyaltyMember'] == True]['Name'].tolist()
59
60 return average_age, loyalty_members
61
62 def find_loyalty_members(df):
63
64 # Filter passengers who are loyalty program members
65 loyalty_members = df[df['LoyaltyMember'] == True]['Name'].tolist()
66
67 return loyalty_members
68
69 file_path = 'C:/Users/milks/Desktop/project_alexisma/passengers.csv' # Update with your file path
70 df = load_data(file_path)
71
72 df['Birthdate'] = pd.to_datetime(df['Birthdate'], errors='coerce')
73
74 cleaned_df = clean_data(df)
75
76 # average age for a given travel class
77 average_age, loyalty_members_in_class = calculate_average_age(cleaned_df, 'Economy')
78 print(f"Average age in Economy class: {average_age}")
79 print("Loyalty members in Economy class:", loyalty_members_in_class)
80
81 # Find all loyalty members
82 all_loyalty_members = find_loyalty_members(cleaned_df)
83 print("All loyalty members:", all_loyalty_members)
84

```
Columns in DataFrame: Index(['PassengerID', 'Name', 'Birthdate', 'TravelClass', 'LoyaltyMember',
       'FlightNumber'],
      dtype='object')
Average age in Economy class: nan
Loyalty members in Economy class: []
All loyalty members: ['John Doe', 'Mia Wong', 'Isabella Rossi', 'Liam Davis', 'Emma Miller', 'Sophia Taylor', 'Amelia Thompson', 'Oliver Martinez', 'Charlotte Clark', 'Ethan Harris', 'James Robinson', 'Victoria Lee', 'Elizabeth Jackson', 'Michael Hill', 'Thomas Turner', 'Steven Campbell', 'Kevin Evans', 'Jessica Edwards', 'Laura Walker', 'Julia Lee', 'Megan Clark', 'Amy Robinson', 'Mary Davis', 'Anthony Johnson', 'Paula Thomas', 'Linda Taylor', 'Mark Anderson', 'Donald Harris', 'George Lewis', 'Sharon Clark', 'Jennifer Allen', 'Diane Hernandez', 'Barbara Jackson', 'Samantha Thompson', 'Nancy Martinez', 'Deborah Clark', 'Lisa Walker', 'Kevin Adams', 'Carol Anderson', 'Nicole Mitchell', 'Ryan Lee', 'Amy Brown', 'Julie Taylor', 'Gloria White', 'Edward King', 'Michael Harris', 'Robert Nelson', 'Jason Clark', 'Sarah Baker', 'Betty Gonzalez', 'Daniel Lewis', 'Gregory Young', 'Christopher Moore', 'Justin Baker', 'Nicholas Adams', 'Brandon Martinez', 'Joshua Scott', 'Justin Clark', 'Natalie Rodriguez', 'Kelly Martinez', 'Kimberly Davis', 'Christopher Allen', 'Nathan Wright', 'Matthew Gonzalez', 'Daniel Martinez', 'Scott Thompson', 'Adam Rodriguez', 'Andrew Hernandez', 'Nicholas Martinez', 'David Johnson', 'James Garcia', 'Ryan Martinez', 'Michael Clark', 'Robert Lopez', 'John Rodriguez', 'William Martin', 'Elizabeth Wright', 'Emily Allen', 'Alexander King', 'Matthew Davis', 'Brandon Garcia', 'David Young', 'Lauren Moore', 'Samantha Lee', 'Nicole Harris', 'Ashley Thompson', 'Kimberly Davis', 'Emily Allen', 'Stephanie Miller', 'Emily Hernandez', 'Jessica Wilson', 'Amanda Martinez', 'Lauren Taylor', 'Sarah Brown', 'Rebecca Hernandez', 'Nicole Taylor', 'Ashley Johnson', 'Samantha Rodriguez', 'Stephanie Scott', 'Lauren Thompson', 'Alexander Garcia', 'Christopher Johnson', 'Jonathan Hernandez', 'Nicholas Wilson', 'Michael Clark', 'Matthew Rodriguez', 'Brandon Lee', 'John Taylor', 'Lauren White', 'Jessica Davis', 'Rebecca Rodriguez', 'Amanda Johnson', 'Emily Perez', 'Brandon Young', 'Michael Thompson', 'Matthew Taylor', 'David Clark', 'Christopher Brown', 'Brandon Garcia', 'Michael Davis', 'David Wilson', 'Matthew Johnson', 'Michael Hernandez', 'Brandon Thomas', 'Matthew Garcia', 'David Rodriguez', 'Christopher Brown', 'Brandon Hernandez', 'Michael Lee', 'David Taylor', 'Matthew Thompson', 'Brandon Wilson', 'Michael Garcia', 'David Hernandez', 'Christopher Clark', 'Matthew Rodriguez', 'Michael Davis', 'Brandon Martinez', 'David Allen', 'Christopher Taylor', 'Matthew Young', 'Michael Clark', 'Brandon Davis', 'David Moore', 'Christopher Martin', 'Matthew Taylor', 'Michael Scott', 'Brandon Thompson', 'David Martinez', 'Christopher Hernandez', 'Matthew Gonzalez', 'Michael Rodriguez', 'Brandon Wilson', 'David Hernandez']
PS C:\Users\milks\Desktop\project_alexisma>
```

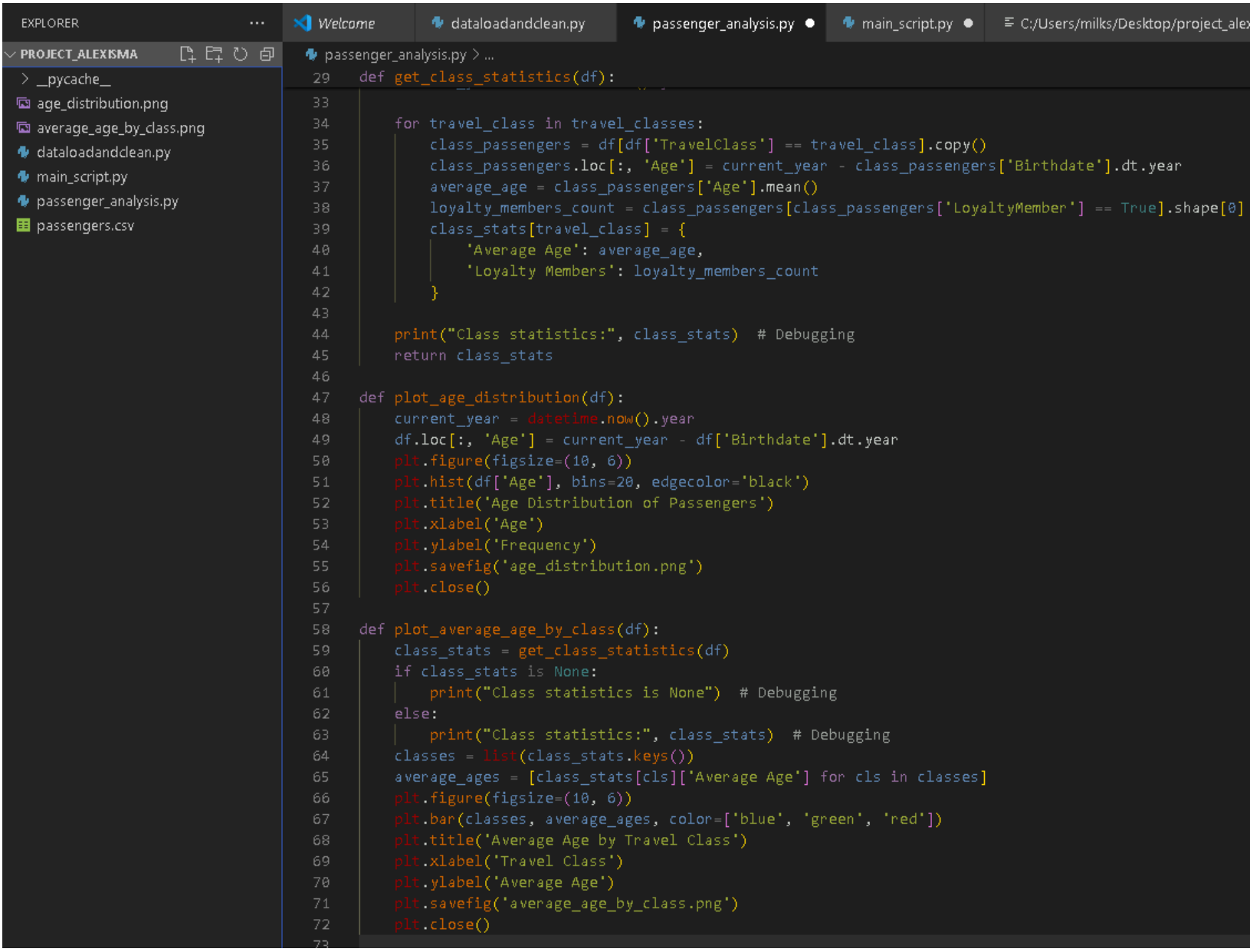
TASK 3:

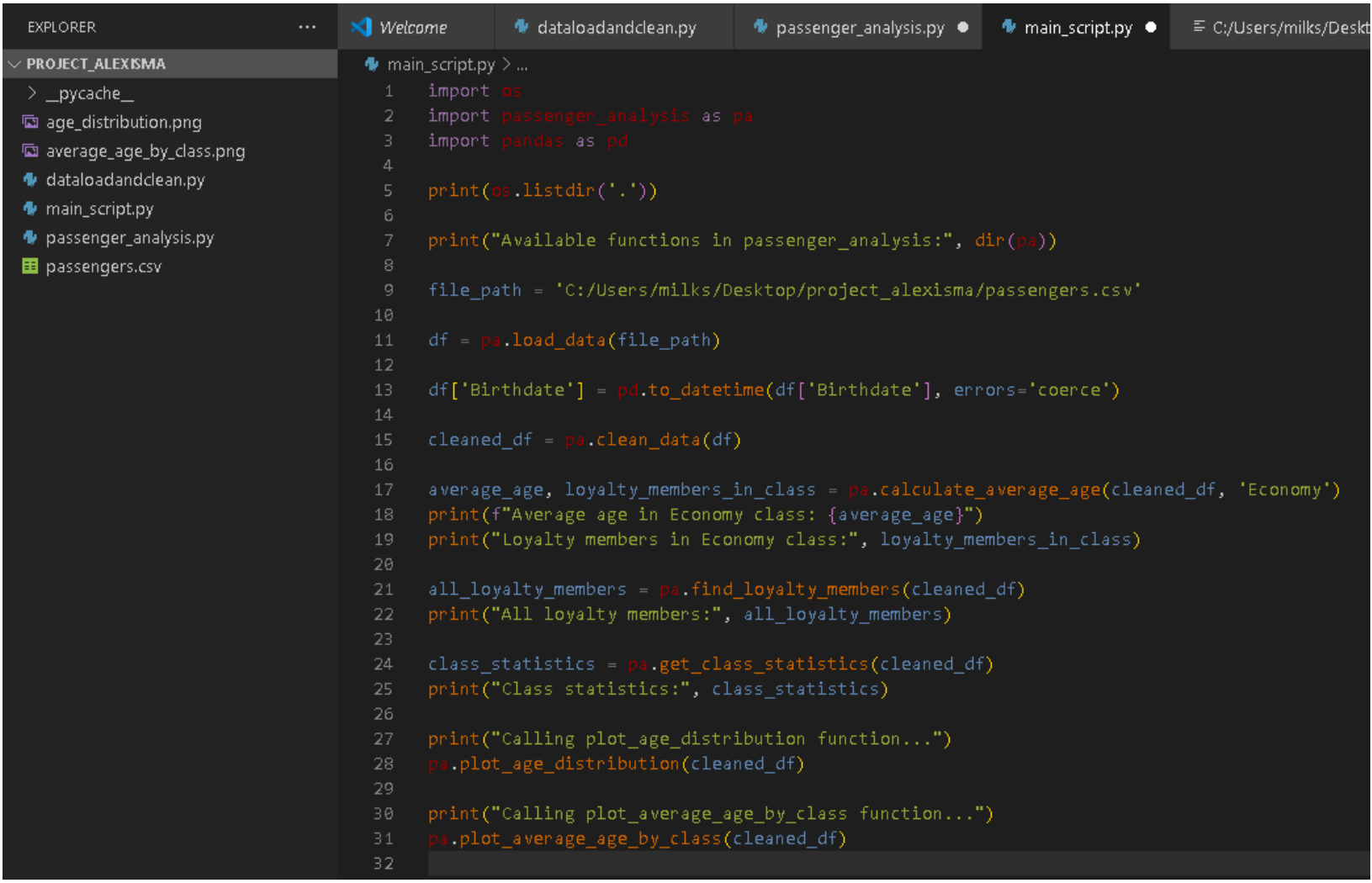

EXPLORER
PROJECT_ALEXISMA
__pycache__
dataloadandclean.py
main_script.py
passenger_analysis.py
passengers.csv
Welcome
dataloadandclean.py
passenger_analysis.py
main_script.py
C:/Users/milks/Desktop/project_alexisma/
passenger_analysis.py > ...
def load_data(file_path):
df = pd.read_csv(file_path, skiprows=1, header=None, on_bad_lines='skip')
df.columns = ['PassengerID', 'Name', 'Birthdate', 'TravelClass', 'LoyaltyMember', 'FlightNumber']
return df
def clean_data(df):
df.dropna(inplace=True)
df['Birthdate'] = pd.to_datetime(df['Birthdate'], errors='coerce')
df['LoyaltyMember'] = df['LoyaltyMember'].astype(bool)
df = df[df['Birthdate'].notna()]
return df
def calculate_average_age(df, travel_class):
class_passengers = df[df['TravelClass'] == travel_class].copy()
current_year = datetime.now().year
class_passengers.loc[:, 'Age'] = current_year - class_passengers['Birthdate'].dt.year
average_age = class_passengers['Age'].mean()
loyalty_members = class_passengers[class_passengers['LoyaltyMember'] == True]['Name'].tolist()
return average_age, loyalty_members
def find_loyalty_members(df):
loyalty_members = df[df['LoyaltyMember'] == True]['Name'].tolist()
return loyalty_members
def get_class_statistics(df):
class_stats = {}
travel_classes = df['TravelClass'].unique()
current_year = datetime.now().year
for travel_class in travel_classes:
class_passengers = df[df['TravelClass'] == travel_class].copy()
class_passengers.loc[:, 'Age'] = current_year - class_passengers['Birthdate'].dt.year
average_age = class_passengers['Age'].mean()
loyalty_members_count = class_passengers[class_passengers['LoyaltyMember'] == True].shape[0]
class_stats[travel_class] = {
'Average Age': average_age,
'Loyalty Members': loyalty_members_count
}
return class_stats

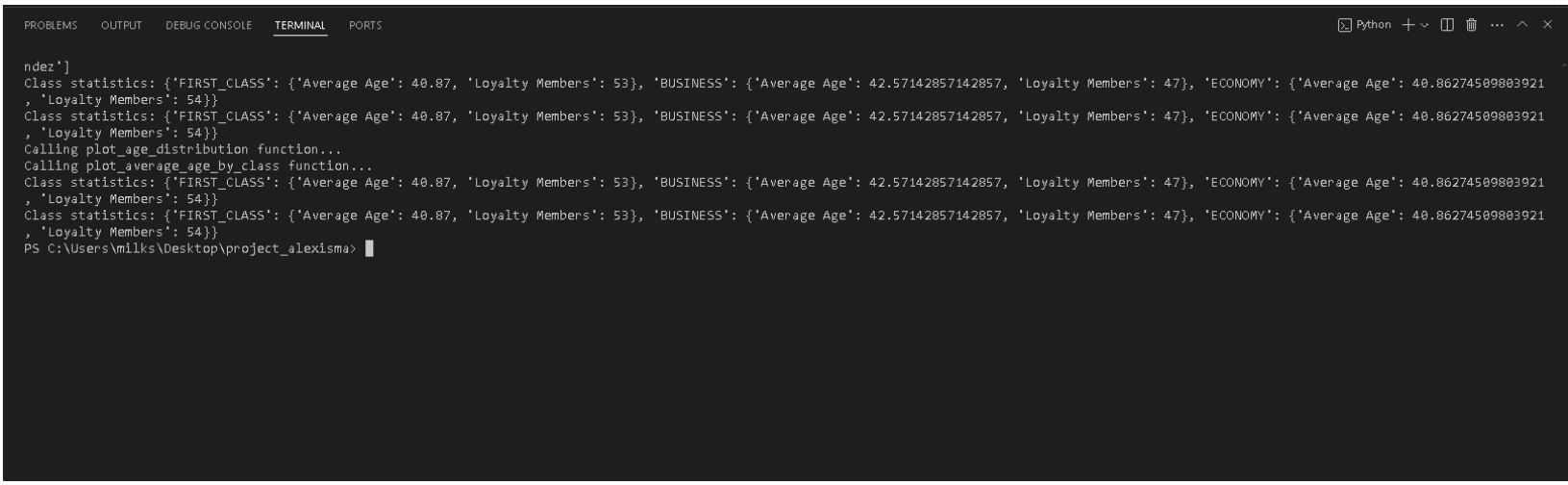

EXPLORER
Welcome
dataloadandclean.py
passenger_analysis.py
main_script.py
C:/Users/milks/Desk
PROJECT_ALEXISMA
pycache
dataloadandclean.py
main_script.py
passenger_analysis.py
passengers.csv
main_script.py > ...
1 import os
2 import passenger_analysis as pa
3 import pandas as pd
4
5 # making sure both scripts are present
6 print(os.listdir('.'))
7
8 file_path = 'C:/Users/milks/Desktop/project_alexisma/passengers.csv'
9
10 df = pa.load_data(file_path)
11
12 df['Birthdate'] = pd.to_datetime(df['Birthdate'], errors='coerce')
13
14 cleaned_df = pa.clean_data(df)
15
16 # average age for a given travel class
17 average_age, loyalty_members_in_class = pa.calculate_average_age(cleaned_df, 'Economy')
18 print(f"Average age in Economy class: {average_age}")
19 print("Loyalty members in Economy class:", loyalty_members_in_class)
20
21 all_loyalty_members = pa.find_loyalty_members(cleaned_df)
22 print("All loyalty members:", all_loyalty_members)
23
24 class_statistics = pa.get_class_statistics(cleaned_df)
25 print("Class statistics:", class_statistics)
26

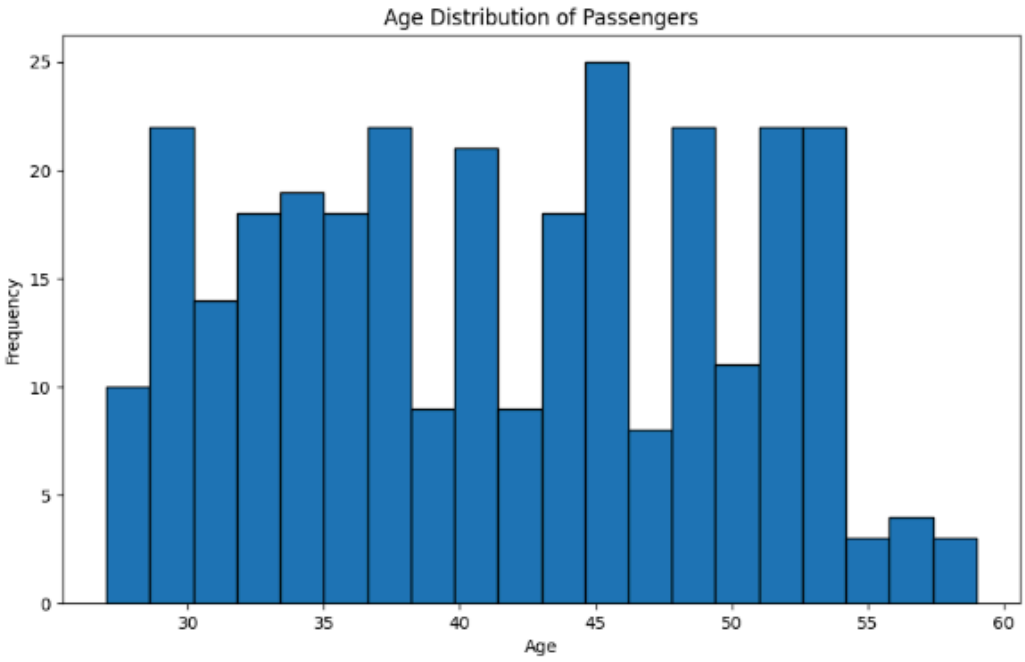

PROBLEMS
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
Python
PS C:\Users\milks\Desktop\project_alexisma> & C:/Users/milks/AppData/Local/Microsoft/WindowsApps/python3.11.exe c:/Users/milks/Desktop/project_alexisma/main_script.py
['dataloadandclean.py', 'main_script.py', 'passengers.csv', 'passenger_analysis.py', '__pycache__']
Average age in Economy class: nan
Loyalty members in Economy class: []
All loyalty members: ['John Doe', 'Mia Wong', 'Isabella Rossi', 'Liam Davis', 'Emma Miller', 'Sophia Taylor', 'Amelia Thompson', 'Oliver Martinez', 'Charlotte Clark', 'Ethan Harris', 'James Robinson', 'Victoria Lee', 'Elizabeth Jackson', 'Michael Hill', 'Thomas Turner', 'Steven Campbell', 'Kevin Evans', 'Jessica Edwards', 'Laura Walker', 'Julia Lee', 'Megan Clark', 'Amy Robinson', 'Mary Davis', 'Anthony Johnson', 'Paula Thomas', 'Linda Taylor', 'Mark Anderson', 'Donald Harris', 'George Lewis', 'Sharon Clark', 'Jennifer Allen', 'Diane Hernandez', 'Barbara Jackson', 'Samantha Thompson', 'Nancy Martinez', 'Deborah Clark', 'Lisa Walker', 'Kevin Adams', 'Carol Anderson', 'Nicole Mitchell', 'Ryan Lee', 'Amy Brown', 'Julie Taylor', 'Gloria White', 'Edward King', 'Michael Harris', 'Robert Nelson', 'Jason Clark', 'Sarah Baker', 'Betty Gonzalez', 'Daniel Lewis', 'Gregory Young', 'Christopher Moore', 'Justin Baker', 'Nicholas Adams', 'Brandon Martinez', 'Joshua Scott', 'Justin Clark', 'Natalie Rodriguez', 'Kelly Martinez', 'Kimberly Davis', 'Christopher Allen', 'Nathan Wright', 'Matthew Gonzalez', 'Daniel Martinez', 'Scott Thompson', 'Adam Rodriguez', 'Andrew Hernandez', 'Nicholas Martinez', 'David Johnson', 'James Garcia', 'Ryan Martinez', 'Michael Clark', 'Robert Lopez', 'John Rodriguez', 'William Martin', 'Elizabeth Wright', 'Emily Allen', 'Alexander King', 'Matthew Davis', 'Brandon Garcia', 'David Young', 'Lauren Moore', 'Samantha Lee', 'Nicole Harris', 'Ashley Thompson', 'Kimberly Davis', 'Emily Allen', 'Stephanie Miller', 'Emily Hernandez', 'Jessica Wilson', 'Amanda Martinez', 'Lauren Taylor', 'Sarah Brown', 'Rebecca Hernandez', 'Nicole Taylor', 'Ashley Johnson', 'Samantha Rodriguez', 'Stephanie Scott', 'Lauren Thompson', 'Alexander Garcia', 'Christopher Johnson', 'Jonathan Hernandez', 'Nicholas Wilson', 'Michael Clark', 'Matthew Rodriguez', 'Brandon Lee', 'John Taylor', 'Lauren White', 'Jessica Davis', 'Rebecca Rodriguez', 'Amanda Johnson', 'Emily Perez', 'Brandon Young', 'Michael Thompson', 'Matthew Taylor', 'David Clark', 'Christopher Brown', 'Brandon Garcia', 'Michael Davis', 'David Wilson', 'Matthew Johnson', 'Michael Hernandez', 'Brandon Thomas', 'Matthew Garcia', 'David Rodriguez', 'Christopher Brown', 'Brandon Hernandez', 'Michael Lee', 'David Taylor', 'Matthew Thompson', 'Brandon Wilson', 'Michael Garcia', 'David Hernandez', 'Christopher Clark', 'Matthew Rodriguez', 'Michael Davis', 'Brandon Martinez', 'David Allen', 'Christopher Taylor', 'Matthew Young', 'Michael Clark', 'Brandon Davis', 'David Moore', 'Christopher Martin', 'Matthew Taylor', 'Michael Scott', 'Brandon Thompson', 'David Martinez', 'Christopher Hernandez', 'Matthew Gonzalez', 'Michael Rodriguez', 'Brandon Wilson', 'David Hernandez']
Class statistics: {'FIRST_CLASS': {'Average Age': 40.87, 'Loyalty Members': 53}, 'BUSINESS': {'Average Age': 42.57142857142857, 'Loyalty Members': 47}, 'ECONOMY': {'Average Age': 40.86274509803921, 'Loyalty Members': 54}}
PS C:\Users\milks\Desktop\project_alexisma>

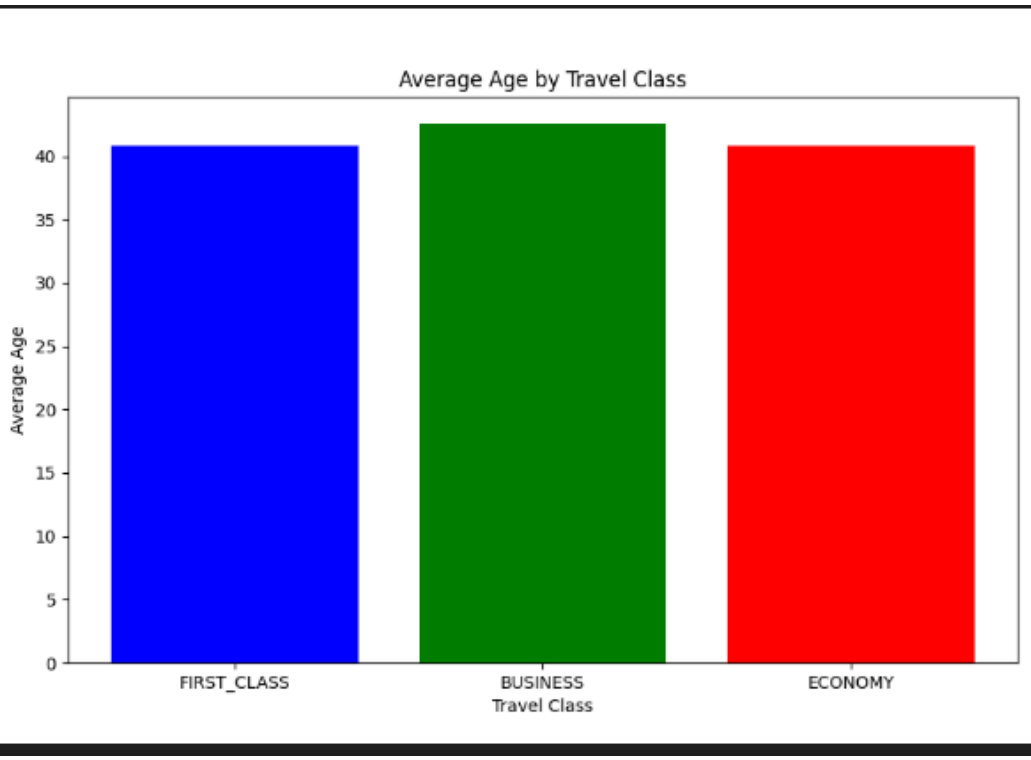
TASK 4:


EXPLORER
PROJECT_ALEXISMA
__pycache__
age_distribution.png
average_age_by_class.png
dataloadandclean.py
main_script.py
passenger_analysis.py
passengers.csv
Welcome
dataloadandclean.py
passenger_analysis.py
main_script.py
C:/Users/milks/Desktop/project_alex
passenger_analysis.py > ...
29 def get_class_statistics(df):
33
34 for travel_class in travel_classes:
35 class_passengers = df[df['TravelClass'] == travel_class].copy()
36 class_passengers.loc[:, 'Age'] = current_year - class_passengers['Birthdate'].dt.year
37 average_age = class_passengers['Age'].mean()
38 loyalty_members_count = class_passengers[class_passengers['LoyaltyMember'] == True].shape[0]
39 class_stats[travel_class] = {
40 'Average Age': average_age,
41 'Loyalty Members': loyalty_members_count
42 }
43
44 print("Class statistics:", class_stats) # Debugging
45 return class_stats
46
47 def plot_age_distribution(df):
48 current_year = datetime.now().year
49 df.loc[:, 'Age'] = current_year - df['Birthdate'].dt.year
50 plt.figure(figsize=(10, 6))
51 plt.hist(df['Age'], bins=20, edgecolor='black')
52 plt.title('Age Distribution of Passengers')
53 plt.xlabel('Age')
54 plt.ylabel('Frequency')
55 plt.savefig('age_distribution.png')
56 plt.close()
57
58 def plot_average_age_by_class(df):
59 class_stats = get_class_statistics(df)
60 if class_stats is None:
61 print("Class statistics is None") # Debugging
62 else:
63 print("Class statistics:", class_stats) # Debugging
64 classes = list(class_stats.keys())
65 average_ages = [class_stats[cls]['Average Age'] for cls in classes]
66 plt.figure(figsize=(10, 6))
67 plt.bar(classes, average_ages, color=['blue', 'green', 'red'])
68 plt.title('Average Age by Travel Class')
69 plt.xlabel('Travel Class')
70 plt.ylabel('Average Age')
71 plt.savefig('average_age_by_class.png')
72 plt.close()

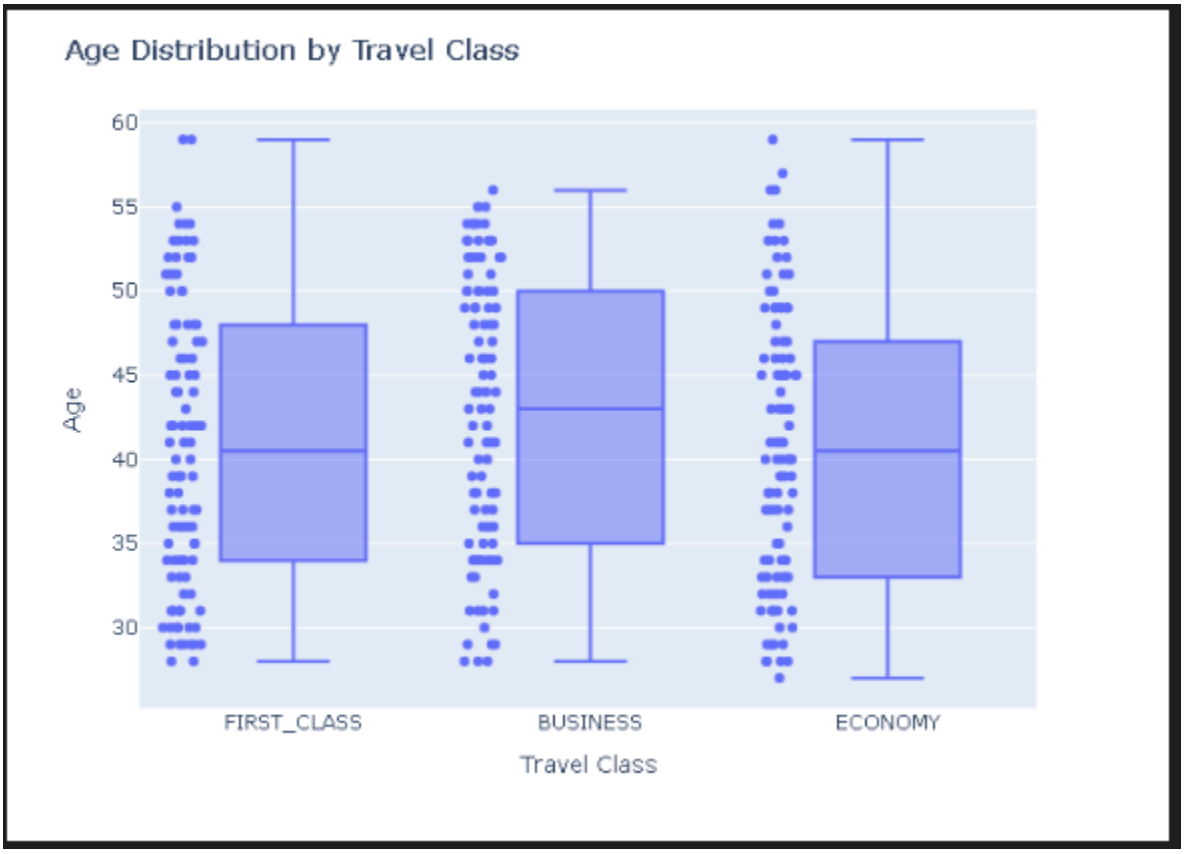

EXPLORER
Welcome
dataloadandclean.py
passenger_analysis.py
main_script.py
C:/Users/milks/Deskt
PROJECT_ALEXISMA
__pycache__
age_distribution.png
average_age_by_class.png
dataloadandclean.py
main_script.py
passenger_analysis.py
passengers.csv
main_script.py > ...
1 import os
2 import passenger_analysis as pa
3 import pandas as pd
4
5 print(os.listdir('.'))
6
7 print("Available functions in passenger_analysis:", dir(pa))
8
9 file_path = 'C:/Users/milks/Desktop/project_alexisma/passengers.csv'
10
11 df = pa.load_data(file_path)
12
13 df['Birthdate'] = pd.to_datetime(df['Birthdate'], errors='coerce')
14
15 cleaned_df = pa.clean_data(df)
16
17 average_age, loyalty_members_in_class = pa.calculate_average_age(cleaned_df, 'Economy')
18 print(f"Average age in Economy class: {average_age}")
19 print("Loyalty members in Economy class:", loyalty_members_in_class)
20
21 all_loyalty_members = pa.find_loyalty_members(cleaned_df)
22 print("All loyalty members:", all_loyalty_members)
23
24 class_statistics = pa.get_class_statistics(cleaned_df)
25 print("Class statistics:", class_statistics)
26
27 print("Calling plot_age_distribution function...")
28 pa.plot_age_distribution(cleaned_df)
29
30 print("Calling plot_average_age_by_class function...")
31 pa.plot_average_age_by_class(cleaned_df)
32

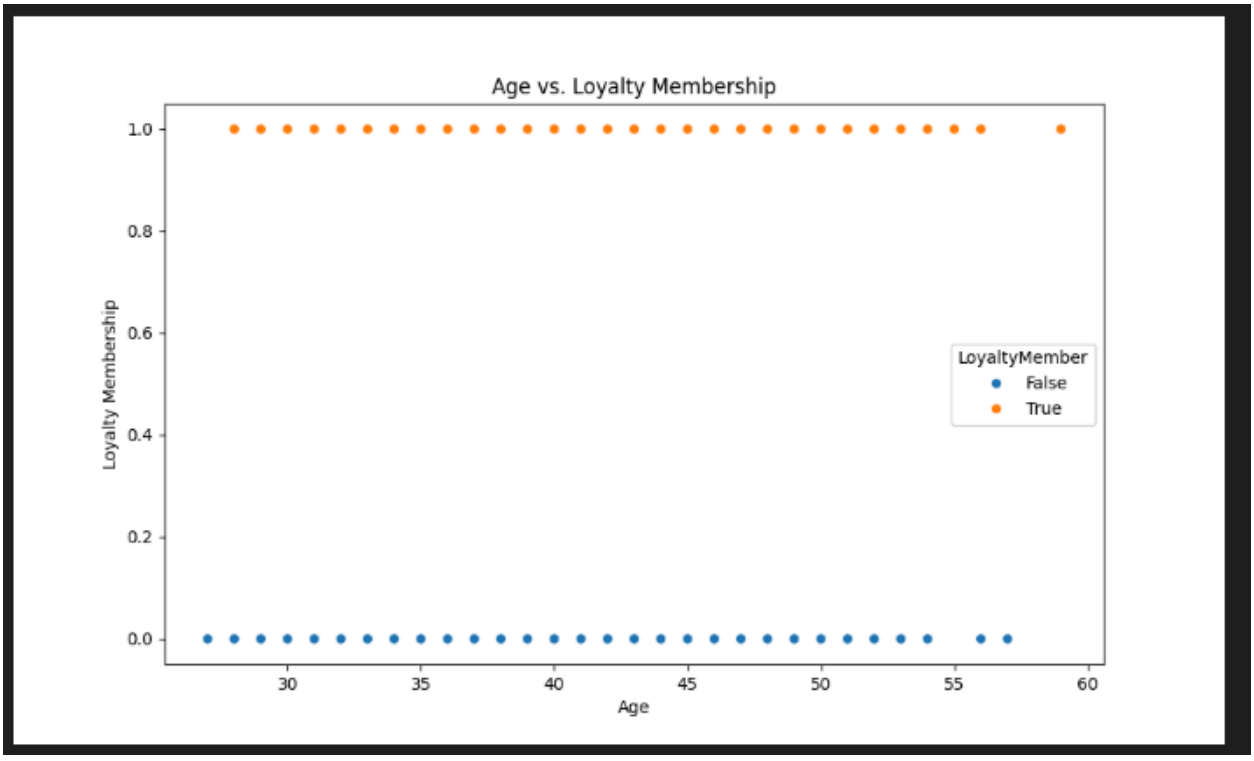

PROBLEMS
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
Python
ndez']
Class statistics: {'FIRST_CLASS': {'Average Age': 40.87, 'Loyalty Members': 53}, 'BUSINESS': {'Average Age': 42.57142857142857, 'Loyalty Members': 47}, 'ECONOMY': {'Average Age': 40.86274509803921, 'Loyalty Members': 54}}
Class statistics: {'FIRST_CLASS': {'Average Age': 40.87, 'Loyalty Members': 53}, 'BUSINESS': {'Average Age': 42.57142857142857, 'Loyalty Members': 47}, 'ECONOMY': {'Average Age': 40.86274509803921, 'Loyalty Members': 54}}
Calling plot_age_distribution function...
Calling plot_average_age_by_class function...
Class statistics: {'FIRST_CLASS': {'Average Age': 40.87, 'Loyalty Members': 53}, 'BUSINESS': {'Average Age': 42.57142857142857, 'Loyalty Members': 47}, 'ECONOMY': {'Average Age': 40.86274509803921, 'Loyalty Members': 54}}
Class statistics: {'FIRST_CLASS': {'Average Age': 40.87, 'Loyalty Members': 53}, 'BUSINESS': {'Average Age': 42.57142857142857, 'Loyalty Members': 47}, 'ECONOMY': {'Average Age': 40.86274509803921, 'Loyalty Members': 54}}
PS C:\Users\milks\Desktop\project_alexisma>


Age Distribution of Passengers
Frequency
Age
0
5
10
15
20
25
30
35
40
45
50
55
60


Average Age by Travel Class
Average Age
Travel Class
0
5
10
15
20
25
30
35
40
FIRST_CLASS
BUSINESS
ECONOMY

TASK 5:


Age Distribution by Travel Class
Age
60
55
50
45
40
35
30
FIRST_CLASS
BUSINESS
ECONOMY
Travel Class


Age vs. Loyalty Membership
Loyalty Membership
1.0
0.8
0.6
0.4
0.2
0.0
30
35
40
45
50
55
60
Age
LoyaltyMember
False
True