PYTHON CODE :

```
import pandas as pd

# Load the Excel file
file_path = '/content/Amazon_Sales_data.csv'
data = pd.read_csv(file_path)

# Display the first few rows of the dataset
data.head()
```

Output:

	Region	Country	Item Type	Sales Channel	Order Priority	Order Date	Order ID	Ship Date	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost	Total Profit
0	Australia and Oceania	Tuvalu	Baby Food	Offline	H	5/28/2010	669165933	6/27/2010	9925	255.28	159.42	2533654.00	1582243.50	951410.50
1	Central America and the Caribbean	Grenada	Cereal	Online	C	8/22/2012	963881480	9/15/2012	2804	205.70	117.11	576782.80	328376.44	248406.36
2	Europe	Russia	Office Supplies	Offline	L	05-02-2014	341417157	05-08-2014	1779	651.21	524.96	1158502.59	933903.84	224598.75
3	Sub-Saharan Africa	Sao Tome and Principe	Fruits	Online	C	6/20/2014	514321792	07-05-2014	8102	9.33	6.92	75591.66	56065.84	19525.82
4	Sub-Saharan Africa	Rwanda	Office Supplies	Offline	L	02-01-2013	115456712	02-06-2013	5062	651.21	524.96	3296425.02	2657347.52	639077.50

```
[15] # Convert 'Order Date' and 'Ship Date' to datetime format
     data['Order Date'] = pd.to_datetime(data['Order Date'], errors='coerce')
     data['Ship Date'] = pd.to_datetime(data['Ship Date'], errors='coerce')

     # Extract year and month from 'Order Date'
     data['Order Year'] = data['Order Date'].dt.year
     data['Order Month'] = data['Order Date'].dt.month
     data['Order Year-Month'] = data['Order Date'].dt.to_period('M')

     # Further transformations as needed (e.g., fill missing values, remove duplicates, normalize data)
     data.fillna(0, inplace=True)

     # Display the transformed data
     data.head()
```

Output:

	Region	Country	Item Type	Sales Channel	Order Priority	Order Date	Order ID	Ship Date	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost	Total Profit	Order Year	Order Month	Order Year-Month
0	Australia and Oceania	Tuvalu	Baby Food	Offline	H	2010-05-28	669165933	2010-06-27	9925	255.28	159.42	2533654.00	1582243.50	951410.50	2010	5	2010-05
1	Central America and the Caribbean	Grenada	Cereal	Online	C	2012-08-22	963881480	2012-09-15	2804	205.70	117.11	576782.80	328376.44	248406.36	2012	8	2012-08
2	Europe	Russia	Office Supplies	Offline	L	1970-01-01	341417157	1970-01-01	1779	651.21	524.96	1158502.59	933903.84	224598.75	1970	1	1970-01
3	Sub-Saharan Africa	Sao Tome and Principe	Fruits	Online	C	2014-06-20	514321792	1970-01-01	8102	9.33	6.92	75591.66	56065.84	19525.82	2014	6	2014-06
4	Sub-Saharan Africa	Rwanda	Office Supplies	Offline	L	1970-01-01	115456712	1970-01-01	5062	651.21	524.96	3296425.02	2657347.52	639077.50	1970	1	1970-01

```
# Save the transformed data to a new CSV file
output_path = '/content/Transformed_Amazon_Sales_Data.csv'
data.to_csv(output_path, index=False)

# Confirm the file is saved
output_path
```

```
'/content/Transformed_Amazon_Sales_Data.csv'
```

```
from google.colab import files

# Download the file to your local system
files.download(output_path)
```