ASSIGNMENT NO: 01

Name: Mihir Unmesh Patil

Roll No: TYCOC213

Batch: C-3

Subject: DMW

Code:

```
import numpy as np
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
df = pd.read_csv(r'C:\Users\mitpa\Downloads\archive\Iris.csv')
print("Original DataFrame:")
print(df.head(), "\n")
duplicate_df = df.iloc[0:1].copy()
df = pd.concat([df, duplicate_df], ignore_index=False)
print("DataFrame after duplicate entries:")
print(df.head(), "\n")
df = df.drop_duplicates()
print("DataFrame after removing duplicates:")
print(df.head(), "\n")
print("DataFrame after changing missing entries:")
np.random.seed(0)
df.loc[0, df.columns[1]] = np.nan
df.loc[1, df.columns[2]] = np.nan
print(df.head(), "\n")
df = df.fillna(df.mean(numeric_only=True))
print("DataFrame after handling missing values:")
print(df.head(), "\n")
if 'target' in df.columns:
    features = df.drop(columns=['target'])
    target = df['target']
else:
    features = df
```

```
    target = None

min_max_scaler = MinMaxScaler()

numeric_features = df.select_dtypes(include=[np.number])

features_normalized = pd.DataFrame(min_max_scaler.fit_transform(numeric_features),
columns=numeric_features.columns)

min_max_scaled = pd.DataFrame(min_max_scaler.fit_transform(numeric_features),
columns=numeric_features.columns)

print("DataFrame after Min-Max Scaling:")

print(min_max_scaled.head(), "\n")

print("Measures of Central Tendency and Dispersion:")

print(numeric_features.describe(), "\n")
```

Output:-

```
Original DataFrame:
   Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
0   1            5.1           3.5            1.4           0.2  Iris-setosa
1   2            4.9           3.0            1.4           0.2  Iris-setosa
2   3            4.7           3.2            1.3           0.2  Iris-setosa
3   4            4.6           3.1            1.5           0.2  Iris-setosa
4   5            5.0           3.6            1.4           0.2  Iris-setosa
DataFrame after duplicate entries:
   Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
0   1            5.1           3.5            1.4           0.2  Iris-setosa
1   2            4.9           3.0            1.4           0.2  Iris-setosa
2   3            4.7           3.2            1.3           0.2  Iris-setosa
3   4            4.6           3.1            1.5           0.2  Iris-setosa
4   5            5.0           3.6            1.4           0.2  Iris-setosa
DataFrame after removing duplicates:
   Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
0   1            5.1           3.5            1.4           0.2  Iris-setosa
1   2            4.9           3.0            1.4           0.2  Iris-setosa
2   3            4.7           3.2            1.3           0.2  Iris-setosa
3   4            4.6           3.1            1.5           0.2  Iris-setosa
4   5            5.0           3.6            1.4           0.2  Iris-setosa
DataFrame after changing missing entries:
   Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
0   1            NaN           3.5            1.4           0.2  Iris-setosa
1   2            4.9           NaN            1.4           0.2  Iris-setosa
2   3            4.7           3.2            1.3           0.2  Iris-setosa
3   4            4.6           3.1            1.5           0.2  Iris-setosa
4   5            5.0           3.6            1.4           0.2  Iris-setosa
DataFrame after handling missing values:
   Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
0   1       5.848322      3.500000            1.4           0.2  Iris-setosa
```

1	2	4.900000	3.054362	1.4	0.2	Iris-setosa
2	3	4.700000	3.200000	1.3	0.2	Iris-setosa
3	4	4.600000	3.100000	1.5	0.2	Iris-setosa
4	5	5.000000	3.600000	1.4	0.2	Iris-setosa

DataFrame after Min-Max Scaling:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
0	0.000000	0.430089	0.625000	0.067797	0.041667
1	0.006711	0.166667	0.439318	0.067797	0.041667
2	0.013423	0.111111	0.500000	0.050847	0.041667
3	0.020134	0.083333	0.458333	0.084746	0.041667
4	0.026846	0.194444	0.666667	0.067797	0.04166

Measures of Central Tendency and Dispersion:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	75.500000	5.848322	3.054362	3.758667	1.198667
std	43.445368	0.825809	0.433572	1.764420	0.763161
min	1.000000	4.300000	2.000000	1.000000	0.100000
25%	38.250000	5.100000	2.800000	1.600000	0.300000
50%	75.500000	5.800000	3.000000	4.350000	1.300000
75%	112.750000	6.400000	3.300000	5.100000	1.800000
max	150.000000	7.900000	4.400000	6.900000	2.500000

ASSIGNMENT NO: 01

Name: Siddhesh Sardar Patil

Roll No: TYCOC218

Batch: C-4

Subject: DMW

Code:

```
import pandas as pd
from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori, association_rules
data = [
    ['Milk', 'Bread', 'Butter'],
    ['Beer', 'Bread', 'Diaper'],
    ['Milk', 'Diaper', 'Bread', 'Beer'],
    ['Milk', 'Bread'],
    ['Beer', 'Diaper', 'Bread']
]
te = TransactionEncoder()
te_ary = te.fit(data).transform(data)
df_encoded = pd.DataFrame(te_ary, columns=te.columns_)
frequent_itemsets = apriori(df_encoded, min_support=0.2, use_colnames=True)
print("Frequent Itemsets:")
print(frequent_itemsets)
rules = association_rules(frequent_itemsets, metric="confidence", min_threshold=0.7)
print("\nAssociation Rules:")
print(rules)
```

Output:

Frequent Itemsets:

	support	itemsets
0	0.6	(Beer)
1	1.0	(Bread)
2	0.2	(Butter)
3	0.6	(Diaper)
4	0.6	(Milk)
5	0.6	(Bread, Beer)
6	0.6	(Diaper, Beer)
7	0.2	(Milk, Beer)
8	0.2	(Bread, Butter)
9	0.6	(Bread, Diaper)
10	0.6	(Milk, Bread)
11	0.2	(Milk, Butter)
12	0.2	(Milk, Diaper)
13	0.6	(Bread, Diaper, Beer)
14	0.2	(Milk, Bread, Beer)
15	0.2	(Milk, Diaper, Beer)
16	0.2	(Milk, Bread, Butter)
17	0.2	(Milk, Bread, Diaper)
18	0.2	(Milk, Bread, Diaper, Beer)

Association Rules:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
0	(Beer)	(Bread)	0.6	1.0	0.6	1.0	1.000000	0.00	inf	0.0
1	(Diaper)	(Beer)	0.6	0.6	0.6	1.0	1.666667	0.24	inf	1.0
2	(Beer)	(Diaper)	0.6	0.6	0.6	1.0	1.666667	0.24	inf	1.0
3	(Butter)	(Bread)	0.2	1.0	0.2	1.0	1.000000	0.00	inf	0.0
4	(Diaper)	(Bread)	0.6	1.0	0.6	1.0	1.000000	0.00	inf	0.0

5	(Milk)	(Bread)	0.6	1.0	0.6	1.0	1.000000	0.00	inf	0.0
6	(Butter)	(Milk)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5
7	(Bread, Diaper)	(Beer)	0.6	0.6	0.6	1.0	1.666667	0.24	inf	1.0
8	(Bread, Beer)	(Diaper)	0.6	0.6	0.6	1.0	1.666667	0.24	inf	1.0
9	(Diaper, Beer)	(Bread)	0.6	1.0	0.6	1.0	1.000000	0.00	inf	0.0
10	(Diaper)	(Bread, Beer)	0.6	0.6	0.6	1.0	1.666667	0.24	inf	1.0
11	(Beer)	(Bread, Diaper)	0.6	0.6	0.6	1.0	1.666667	0.24	inf	1.0
12	(Milk, Beer)	(Bread)	0.2	1.0	0.2	1.0	1.000000	0.00	inf	0.0
13	(Milk, Diaper)	(Beer)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5
14	(Milk, Beer)	(Diaper)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5
15	(Milk, Butter)	(Bread)	0.2	1.0	0.2	1.0	1.000000	0.00	inf	0.0
16	(Bread, Butter)	(Milk)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5
17	(Butter)	(Milk, Bread)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5
18	(Milk, Diaper)	(Bread)	0.2	1.0	0.2	1.0	1.000000	0.00	inf	0.0
19	(Milk, Diaper, Bread)	(Beer)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5
20	(Milk, Bread, Beer)	(Diaper)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5
21	(Milk, Diaper, Beer)	(Bread)	0.2	1.0	0.2	1.0	1.000000	0.00	inf	0.0
22	(Milk, Diaper)	(Bread, Beer)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5
23	(Milk, Beer)	(Bread, Diaper)	0.2	0.6	0.2	1.0	1.666667	0.08	inf	0.5