```
import pandas as pd
from sklearn.preprocessing import MinMaxScaler, OneHotEncoder

# Load Dataset (you can replace 'your_dataset.csv' with your file)
file_path = 'annual-enterprise-survey-2023-financial-year-provisional.csv'
df = pd.read_csv(file_path)

# Display dataset details
print("Dataset Details:")
print(f"Number of rows: {df.shape[0]}")
print(f"Number of columns: {df.shape[1]}")
print("\nFirst five rows:")
print(df.head())
print("\nDataset size (total elements):", df.size)

# Check for missing values
missing_values = df.isnull().sum()
print("\nNumber of missing values per column:")
print(missing_values)

# Summarize numerical columns
numerical_cols = df.select_dtypes(include='number')
print("\nSummary of numerical columns:")
print("Sum:")
print(numerical_cols.sum())
print("Average:")
print(numerical_cols.mean())
print("Minimum:")
print(numerical_cols.min())
print("Maximum:")
print(numerical_cols.max())

# Handle missing values (if any) by filling with the mean
df.fillna(df.mean(), inplace=True)

# Feature Scaling: Normalize numerical columns using Min-Max Scaling
scaler = MinMaxScaler()
scaled_columns = scaler.fit_transform(numerical_cols)
scaled_df = pd.DataFrame(scaled_columns, columns=numerical_cols.columns)

# One-Hot Encoding for categorical columns
categorical_cols = df.select_dtypes(include='object').columns
encoder = OneHotEncoder(sparse=False)
encoded_columns = encoder.fit_transform(df[categorical_cols])

# Convert encoded columns into a DataFrame
encoded_df = pd.DataFrame(encoded_columns, columns=encoder.get_feature_names_out(categorical_cols))

# Combine scaled numerical and one-hot-encoded categorical columns
final_df = pd.concat([scaled_df, encoded_df], axis=1)

# Export the final preprocessed dataset
output_file_path = 'preprocessed_dataset.csv'
final_df.to_csv(output_file_path, index=False)

print("\nPreprocessing complete. The processed dataset has been saved as 'preprocessed_dataset.csv'.")
```

```
First five rows:
    ear Industry_aggregation_NZSIOC  Industry_code_NZSIOC  \
0  2023                     Level 1                 99999   
1  2023                     Level 1                 99999   
2  2023                     Level 1                 99999   
3  2023                     Level 1                 99999   
4  2023                     Level 1                 99999   

  Industry_name_NZSIOC               Units Variable_code  \
0       All industries  Dollars (millions)           H01   
1       All industries  Dollars (millions)           H04   
2       All industries  Dollars (millions)           H05   
3       All industries  Dollars (millions)           H07   
4       All industries  Dollars (millions)           H08   

                                     Variable_name      Variable_category  \
0                                     Total income  Financial performance   
1  Sales, government funding, grants and subsidies  Financial performance   
2                Interest, dividends and donations  Financial performance   
3                             Non-operating income  Financial performance   
4                                Total expenditure  Financial performance   

    Value                         Industry_code_ANZSIC06  
0  930995  ANZSIC06 divisions A-S (excluding classes K633...  
1  821630  ANZSIC06 divisions A-S (excluding classes K633...  
2   84354  ANZSIC06 divisions A-S (excluding classes K633...  
3   25010  ANZSIC06 divisions A-S (excluding classes K633...  
4  832964  ANZSIC06 divisions A-S (excluding classes K633...  

Dataset size (total elements): 290

Number of missing values per column:
ear                            0
Industry_aggregation_NZSIOC    0
Industry_code_NZSIOC           0
Industry_name_NZSIOC           0
Units                          0
Variable_code                  0
Variable_name                  0
Variable_category              0
Value                          0
Industry_code_ANZSIC06         0
dtype: int64

Summary of numerical columns:
Sum:
ear                        58667
Industry_code_NZSIOC     2899971
Value                   15647549
dtype: int64
Average:
ear                       2023.000000
Industry_code_NZSIOC     99999.000000
Value                   539570.655172
dtype: float64
Minimum:
ear                      2023
Industry_code_NZSIOC    99999
Value                      11
dtype: int64
Maximum:
```