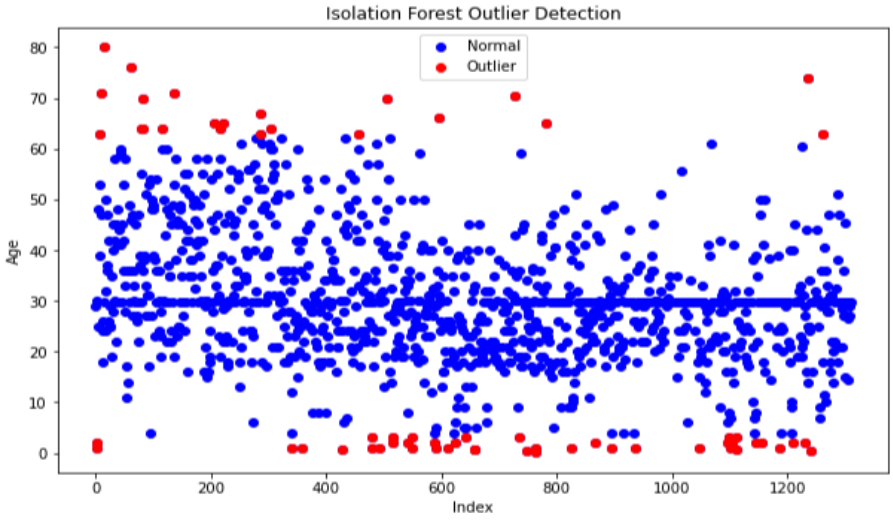
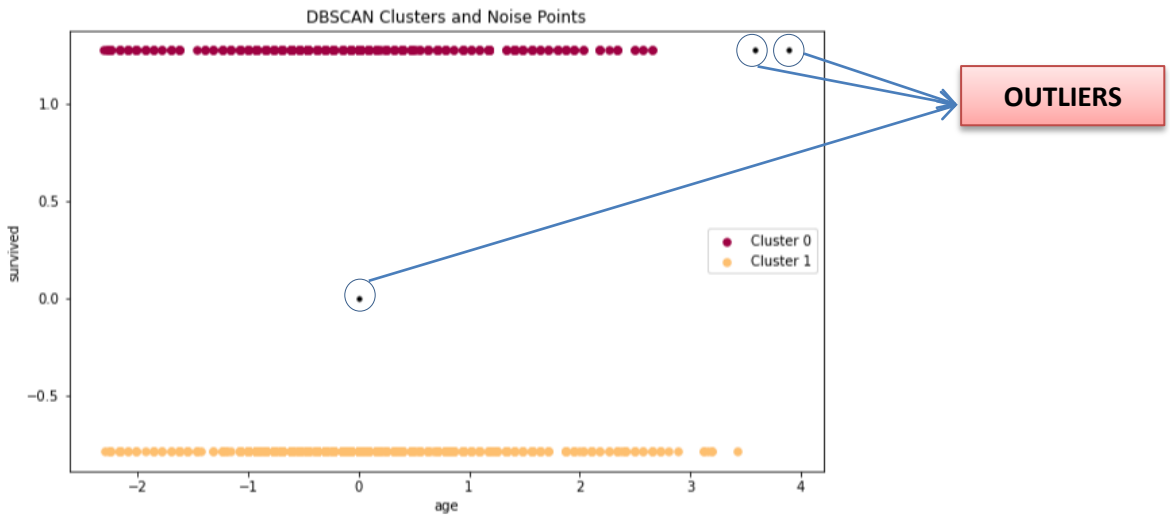
8 Methods to Detect Outliers in a Dataset – Visualization Plots (Histogram, Box Plot, Scatter Plot, KDE, Dist Plot, Frequency Plot), DBSCAN Clustering, Isolation Forest, Z-Statistics & Local Outlier Factor Algorithm

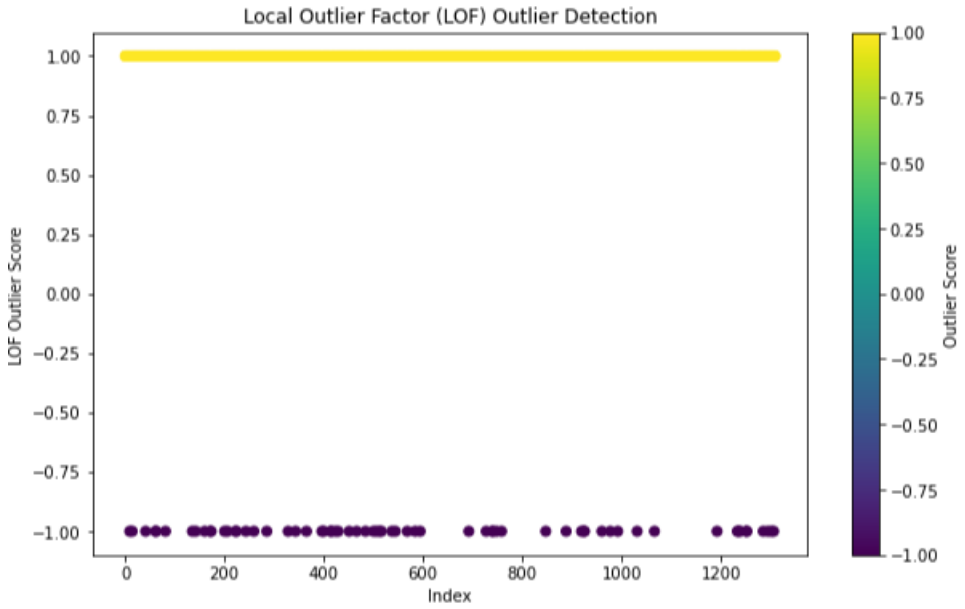
1) Isolation Forest Outlier Detection



2) DBSCAN Clustering Outlier Detection



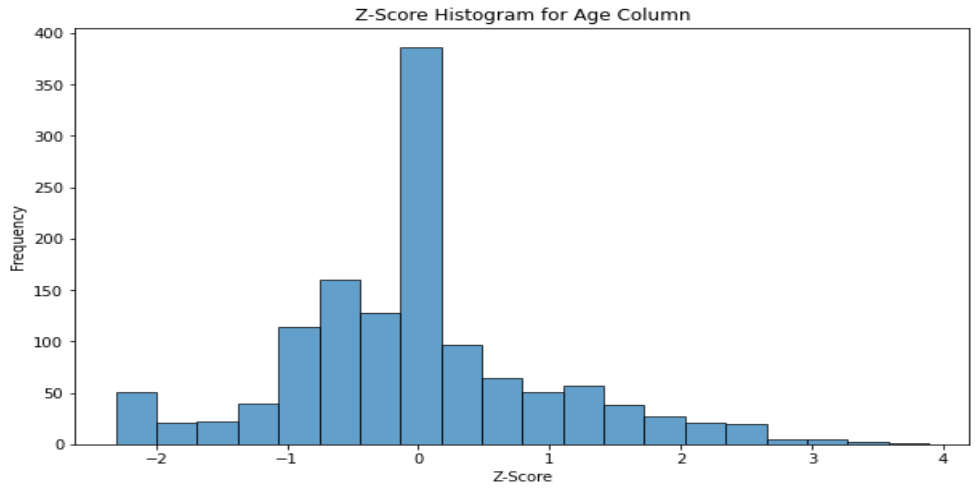
3) Local Outlier Factor Algorithm



4) Z-Statistics Outlier Algorithm

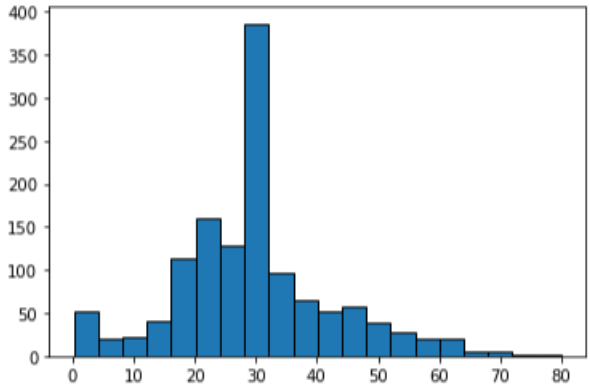
```
from scipy.stats import zscore
# Select the 'age' column
age_column = data['age']

# Calculate the Z-scores for the 'age' column
z_scores = zscore(age_column)

# Define a threshold for identifying outliers (e.g., z-score above 3 or below -3)
outlier_threshold = 3

# Identify the indices of data points with z-scores above the threshold
outlier_indices = [index for index, z_score in enumerate(z_scores) if abs(z_score) > outlier_threshold]

# Print the indices of potential outliers
print("Indices of potential outliers:", outlier_indices)
```

```
Indices of potential outliers: [9, 14, 61, 81, 135, 506, 727, 1235]
```

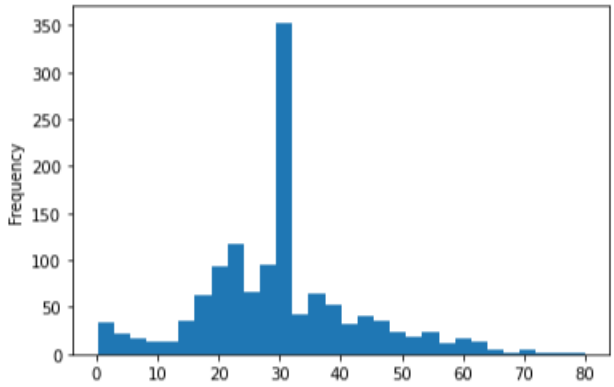


5) Visualization Plots

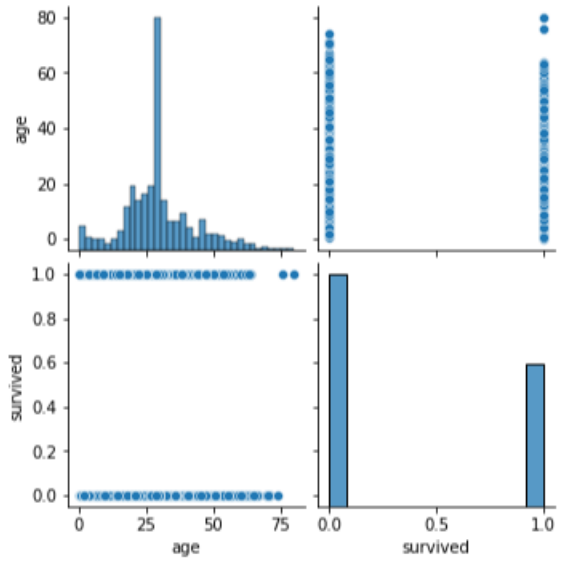
a) Histogram:



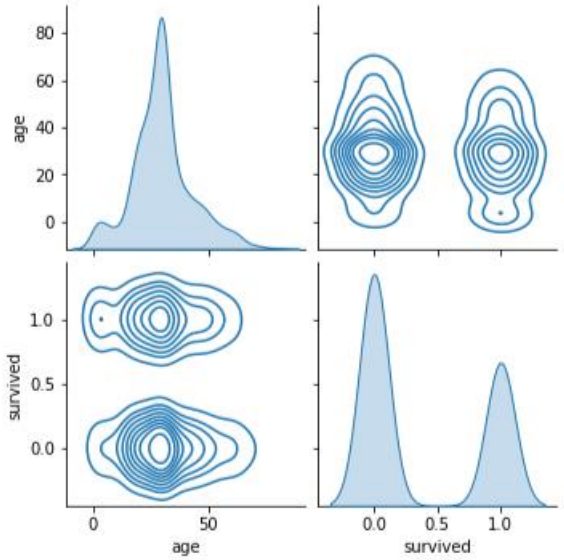
b) Frequency Plot:



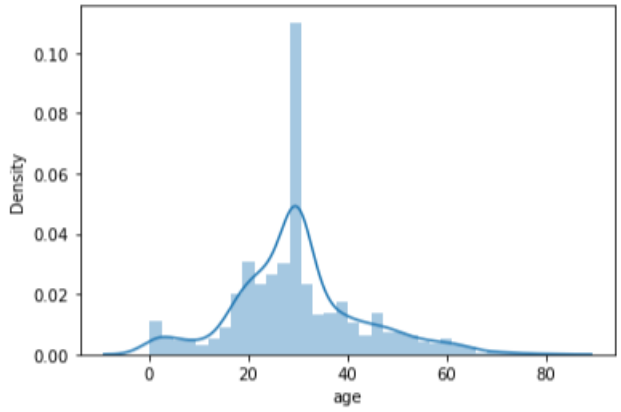
c) Scatter Plot:



d) KDE Plot:



e) Distribution Plot:



f) Box Plot: