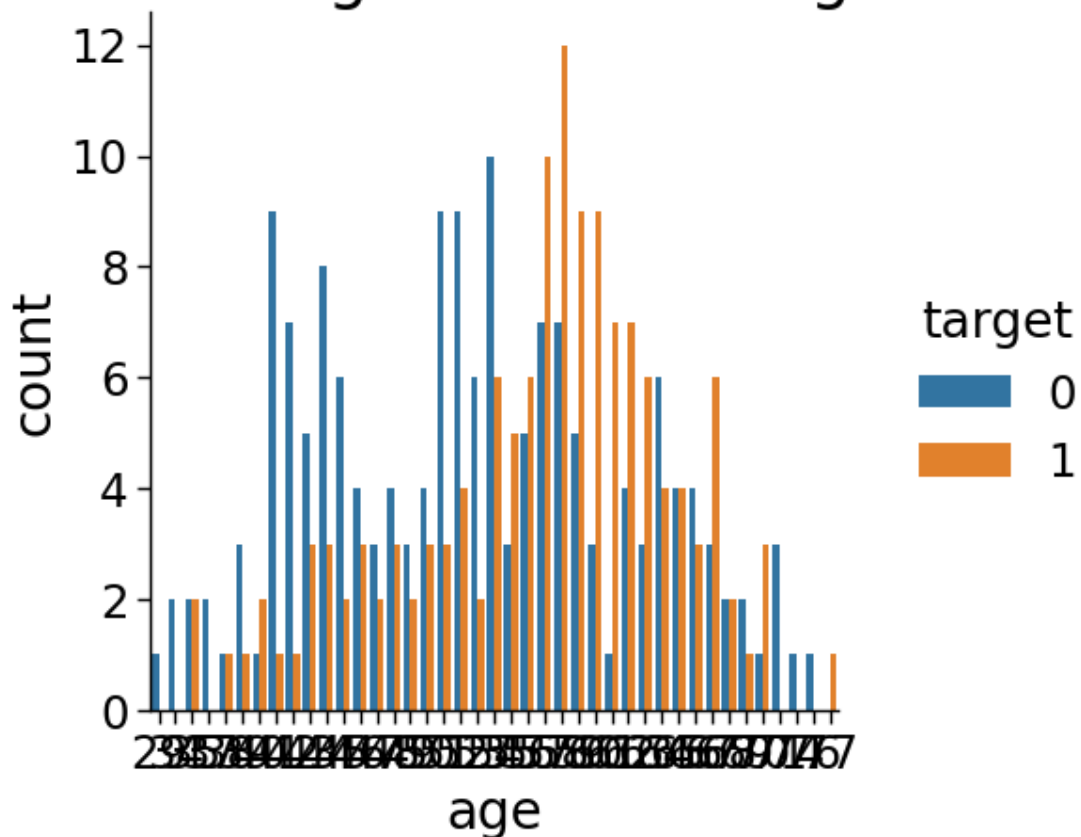
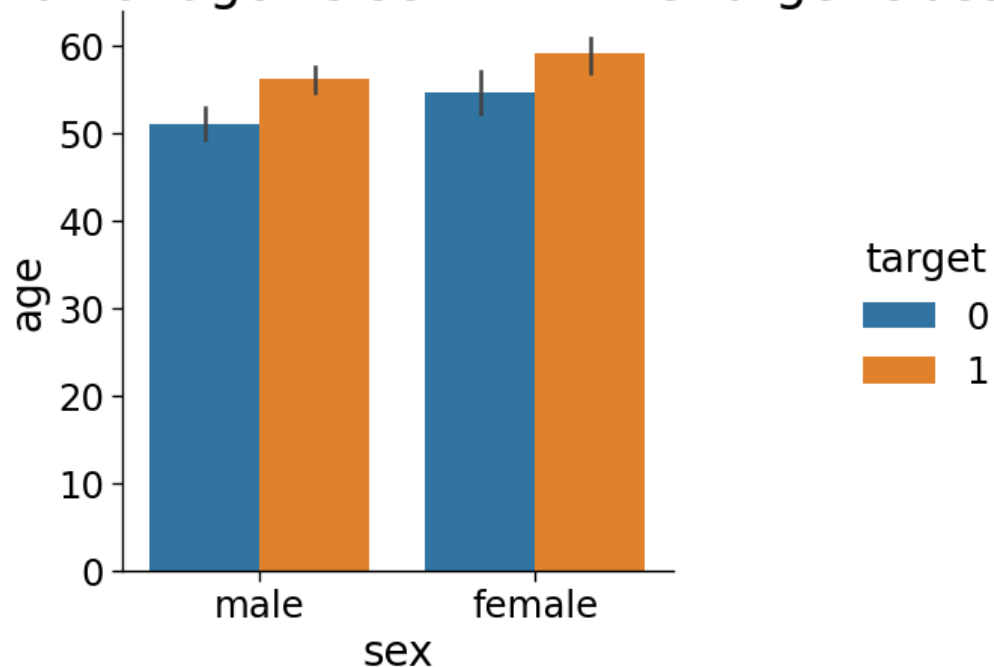


```
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-  
packages (2.0.3)  
Requirement already satisfied: python-dateutil>=2.8.2 in  
/usr/local/lib/python3.10/dist-packages (from pandas) (2.8.2)  
Requirement already satisfied: pytz>=2020.1 in  
/usr/local/lib/python3.10/dist-packages (from pandas) (2023.4)  
Requirement already satisfied: tzdata>=2022.1 in  
/usr/local/lib/python3.10/dist-packages (from pandas) (2024.1)  
Requirement already satisfied: numpy>=1.21.0 in  
/usr/local/lib/python3.10/dist-packages (from pandas) (1.25.2)  
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-  
packages (from python-dateutil>=2.8.2->pandas) (1.16.0)  
ERROR: Could not find a version that satisfies the requirement warings (from  
versions: none)  
ERROR: No matching distribution found for warings
```

Variation of Age for each target class



Distribution of age vs sex with the target class



Accuracy for training set for svm = 0.9256198347107438

Accuracy for test set for svm = 0.8032786885245902

Accuracy for training set for Naive Bayes = 0.8677685950413223

Accuracy for test set for Naive Bayes = 0.7868852459016393

Accuracy for training set for Logistic Regression = 0.8636363636363636

Accuracy for test set for Logistic Regression = 0.8032786885245902

Accuracy for training set for Decision Tree = 1.0

Accuracy for test set for Decision Tree = 0.7540983606557377

/usr/local/lib/python3.10/dist-packages/sklearn/linear_model/_logistic.py:458: ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:

<https://scikit-learn.org/stable/modules/preprocessing.html>

Please also refer to the documentation for alternative solver options:

https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression

```
n_iter_i = _check_optimize_result(
```

Accuracy for training set for Random Forest = 0.987603305785124

Accuracy for test set for Random Forest = 0.7540983606557377

[LightGBM] [Info] Auto-choosing row-wise multi-threading, the overhead of testing was 0.014137 seconds.

You can set `force_row_wise=true` to remove the overhead.

And if memory is not enough, you can set `force_col_wise=true`.

[LightGBM] [Info] Total Bins 249

```
[LightGBM] [Info] Number of data points in the train set: 242, number of
used features: 13
```

```
[LightGBM] [Info] Start training from score 0.466942
```

[illegible]

Accuracy for training set for LightGBM = 0.9958677685950413
Accuracy for test set for LightGBM = 0.7704918032786885

Accuracy for training set for XGBoost = 1.0
Accuracy for test set for XGBoost = 0.7704918032786885