lab_1

March 15, 2021

1 Adam Jochna

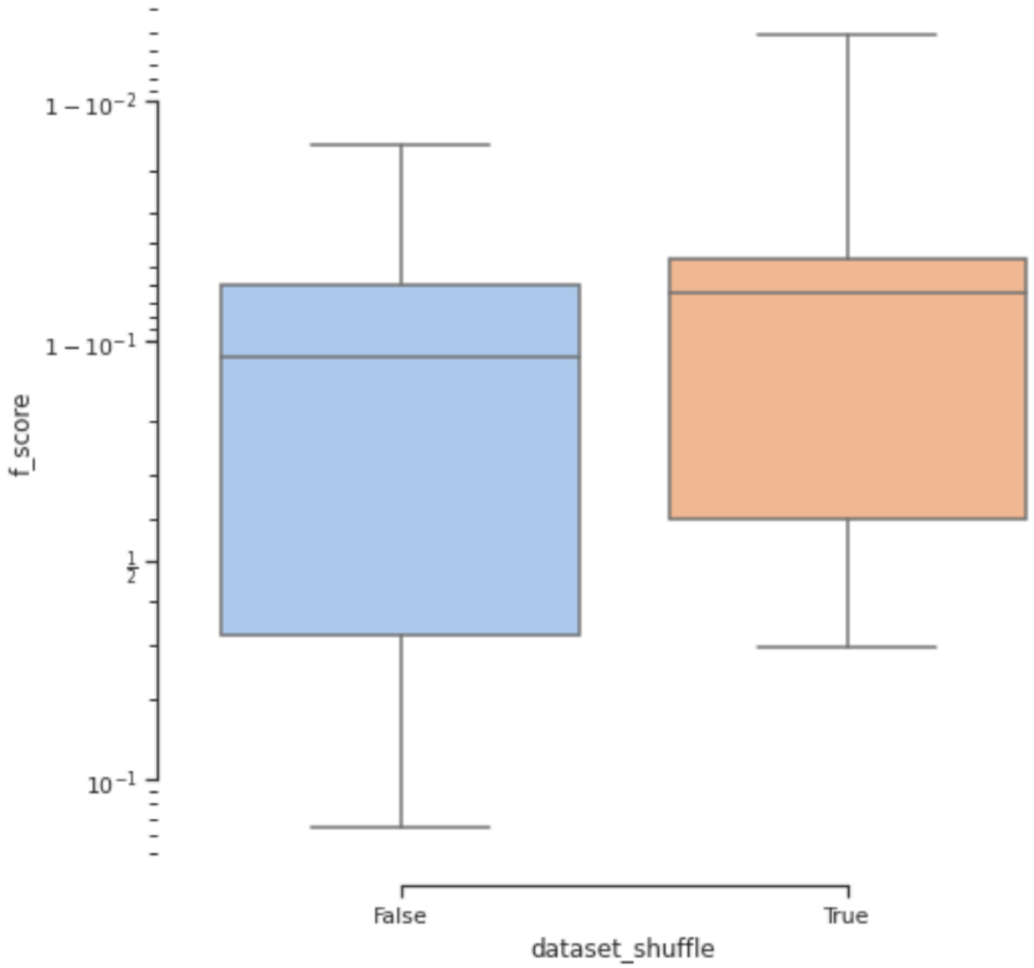
[67]:
```
import pandas as pd
from config import PROJECT_PATH
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
sns.set(rc={'figure.figsize':(8, 8)})
```

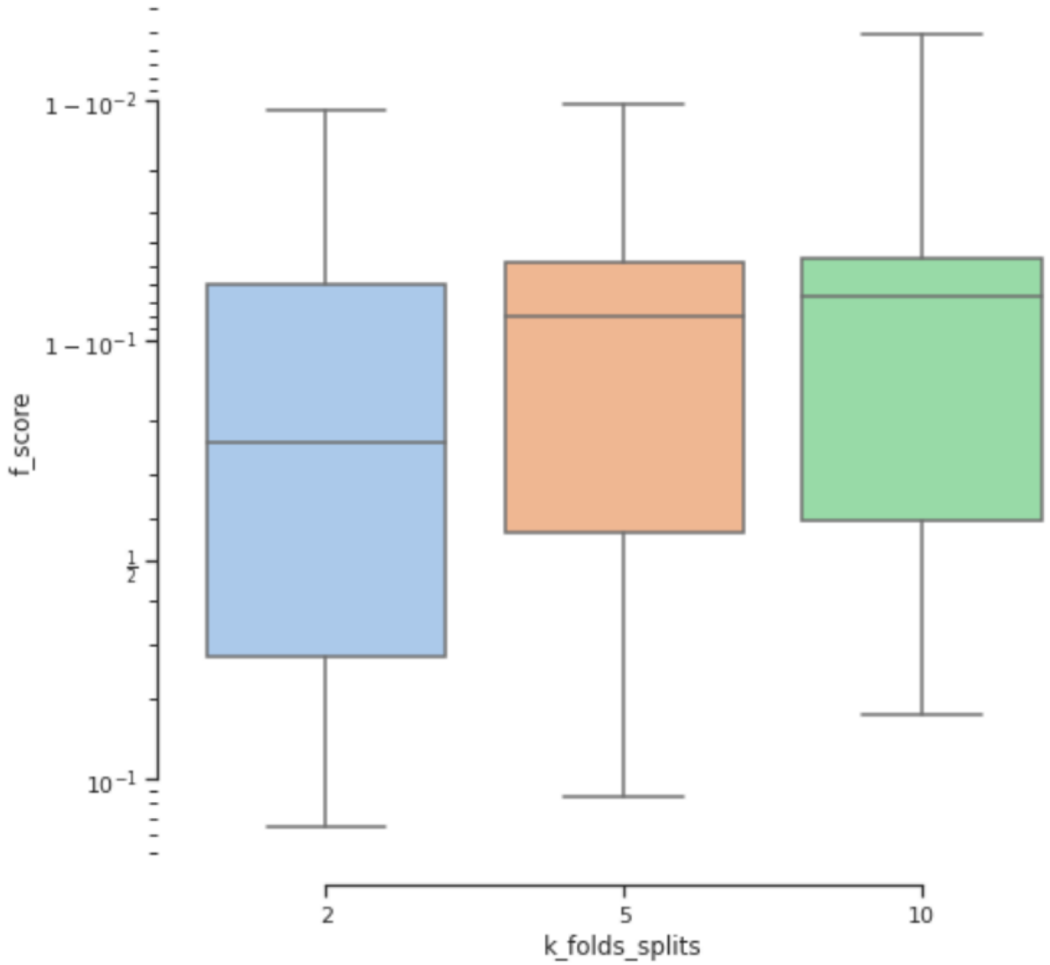
[68]:
```
def plot_0(df, x_arg):
    sns.set_theme(style="ticks", palette="pastel")

    g = sns.boxplot(
        x=x_arg,
        y="f_score",
        data=df
    )
    g.set_yscale("logit")

    sns.despine(offset=10, trim=True)
    plt.show()
```

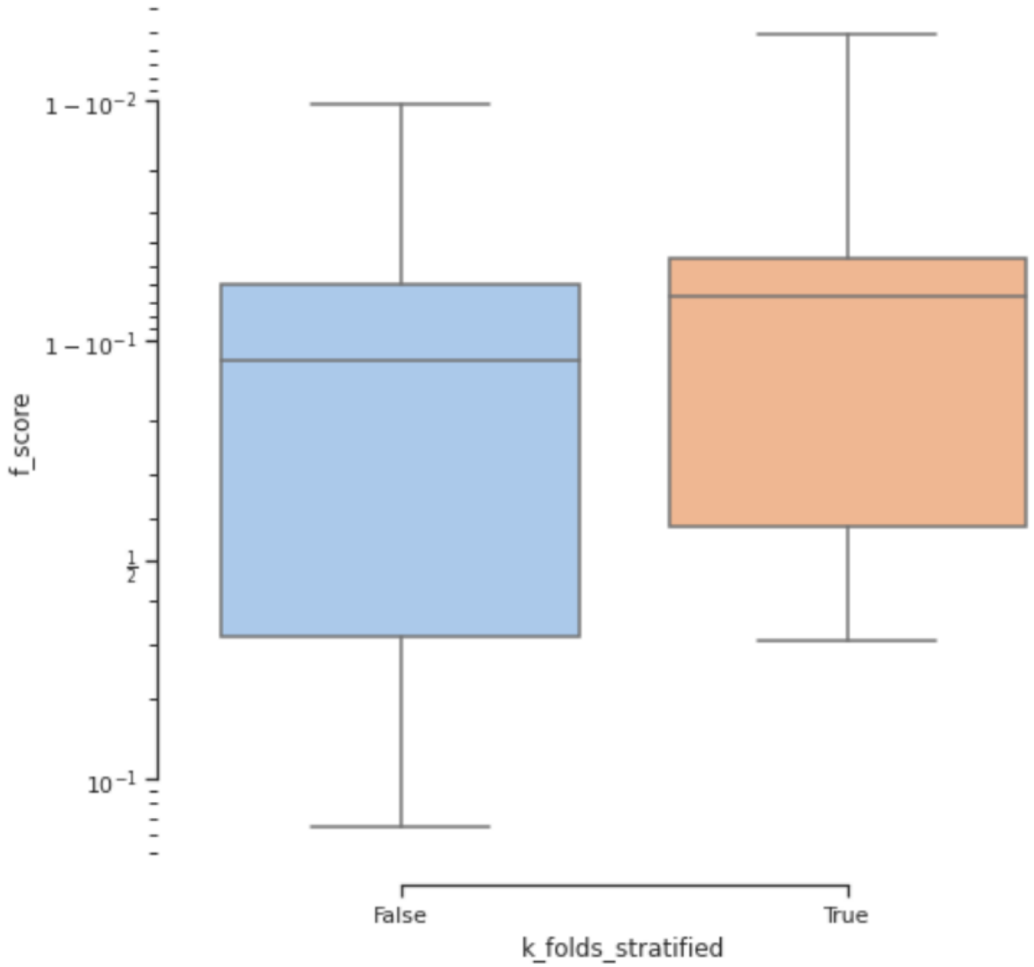
[69]:
```
def plots_for_dataset(df, dataset_type):
    df_tmp = df.loc[df['dataset_type'] == dataset_type]

    x_args = list(df_tmp.columns)
    x_args.remove('dataset_type')
    x_args.remove('f_score')

    df_tmp = df_tmp[x_args]
    df_tmp = df.groupby(x_args, as_index=False).mean()
    df_std = df.groupby(x_args, as_index=False).std()
    df_tmp['f_score_std'] = df_std['f_score']

    for x_arg in x_args:
        plot_0(df, x_arg)
```

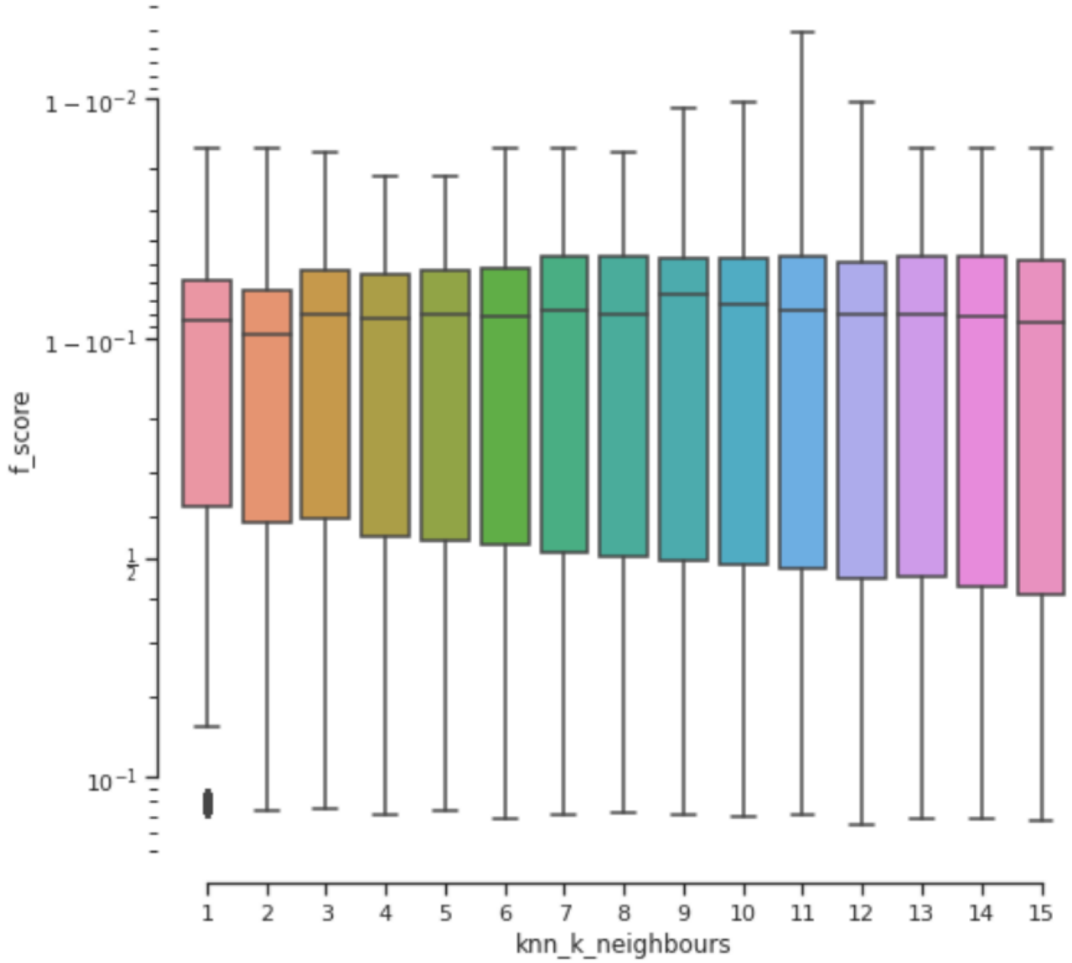
[70]:
```
df = pd.read_csv('{}/lab_1/results.csv'.format(PROJECT_PATH))
```

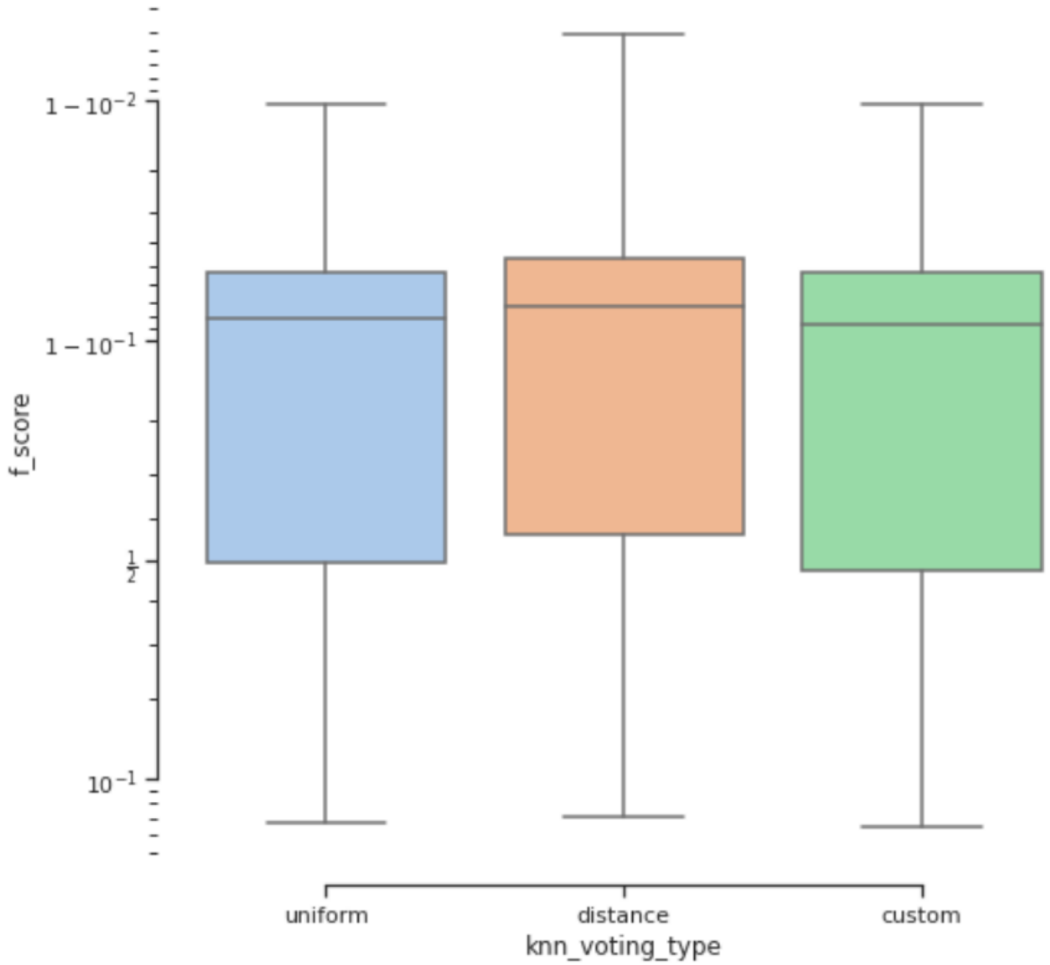
[71]:
```
plots_for_dataset(df.copy(), 'iris')
```

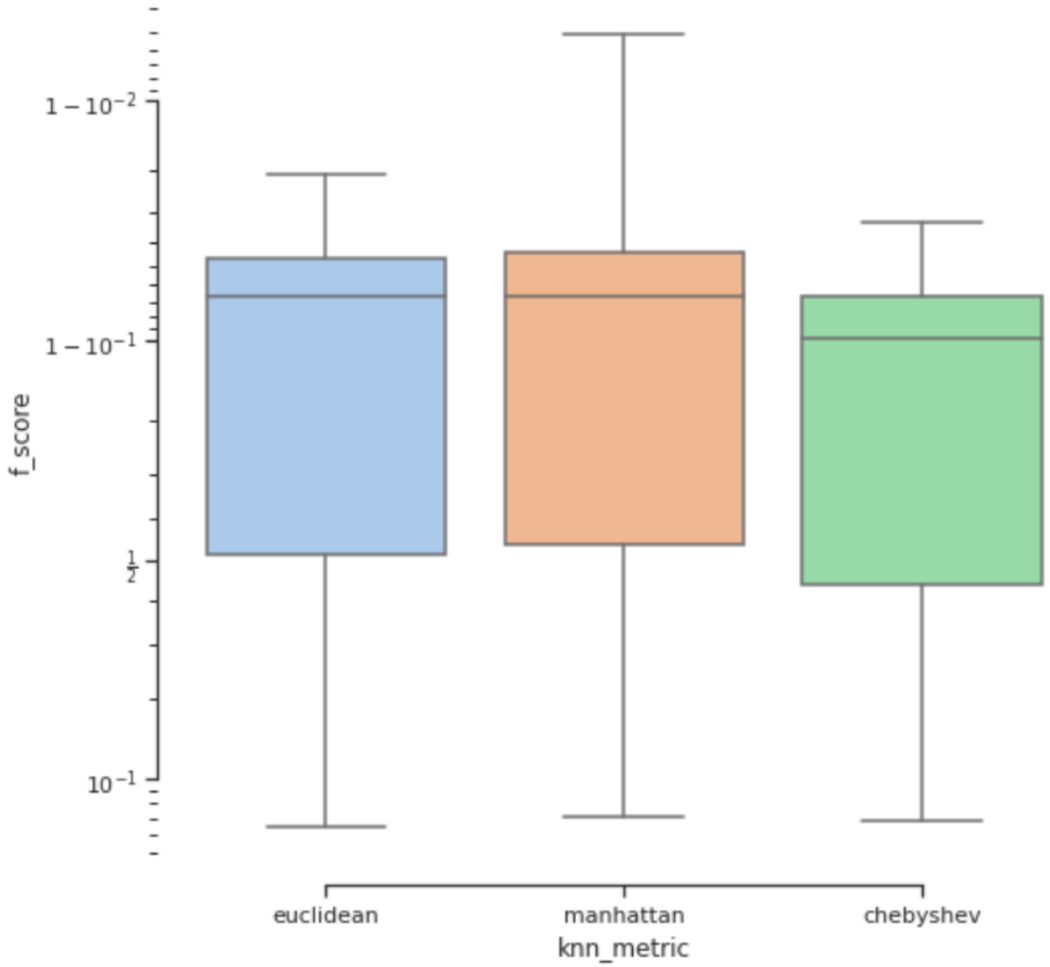



$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
2
5
10
k_folds_splits

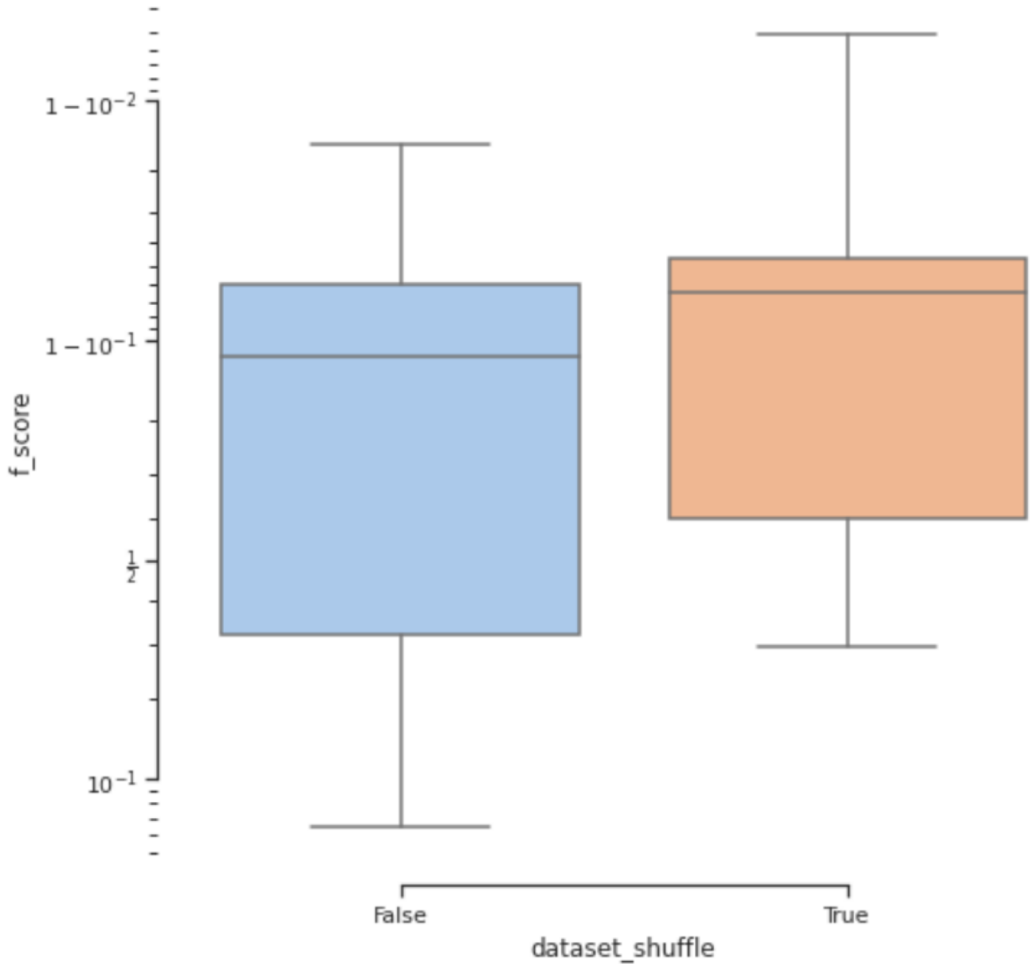

$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
False
True
k_folds_stratified

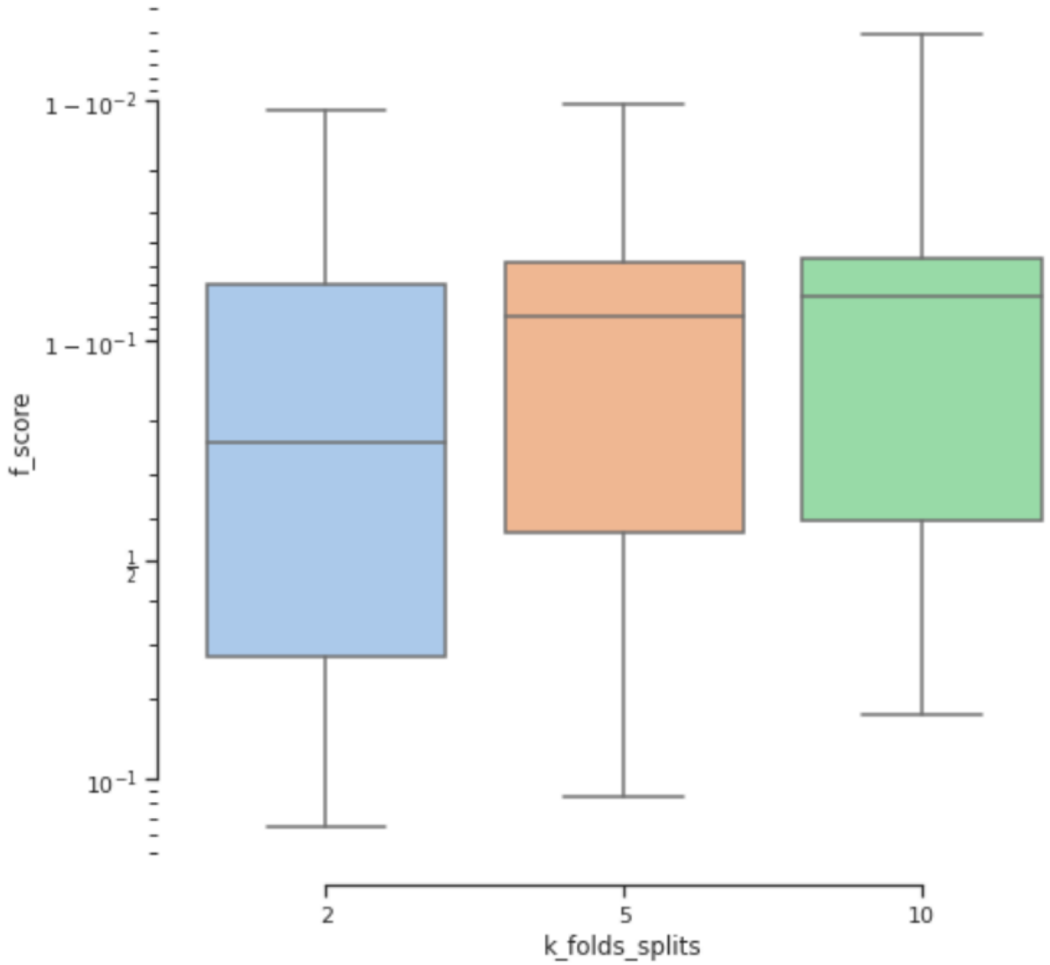

f_score
$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
knn_k_neighbours

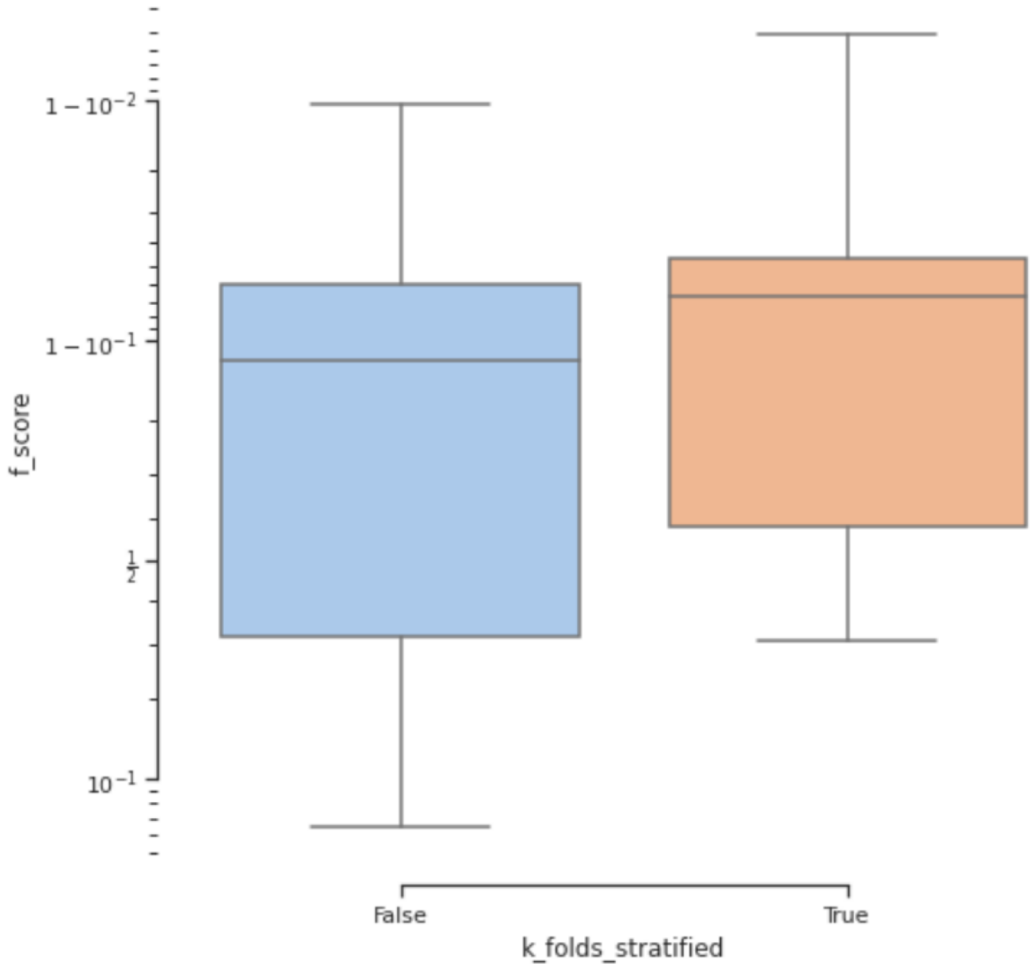

$1 - 10^{-2}$
$1 - 10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
uniform
distance
custom
knn_voting_type

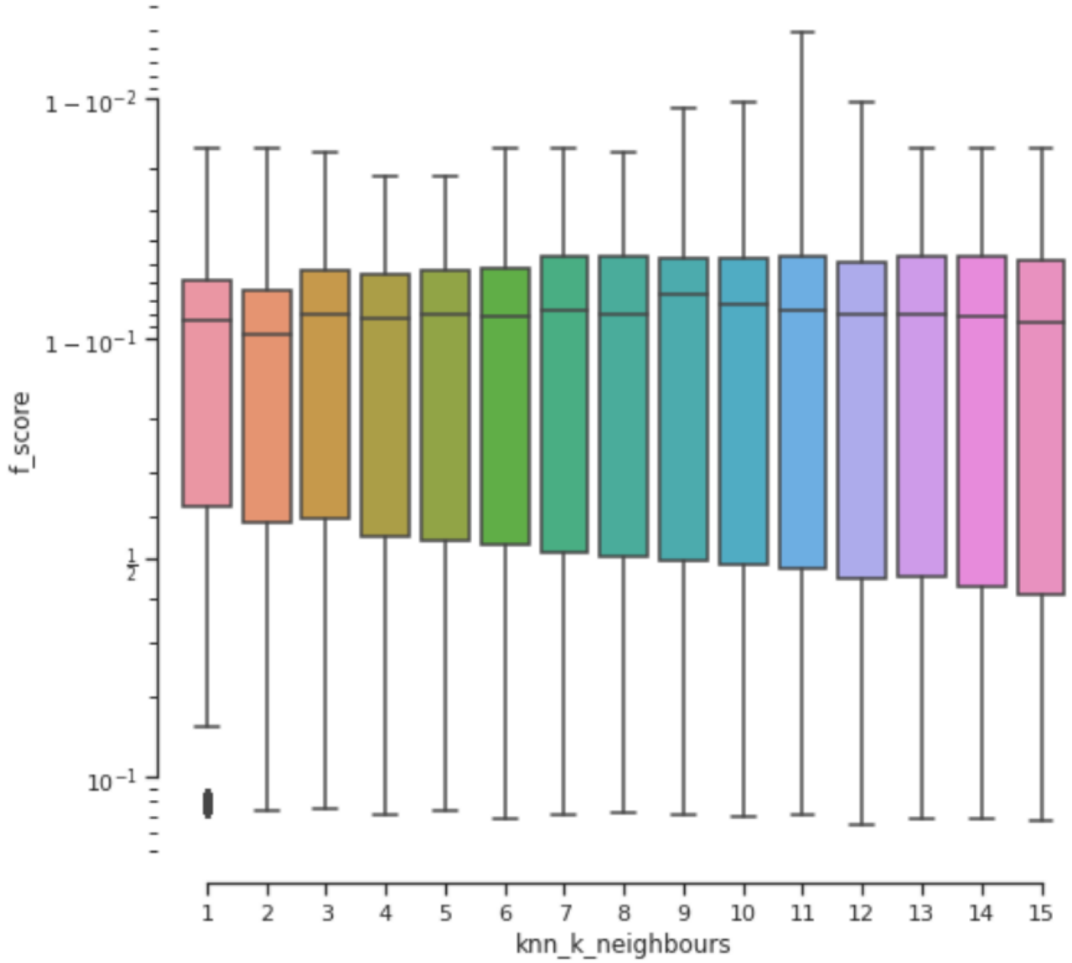


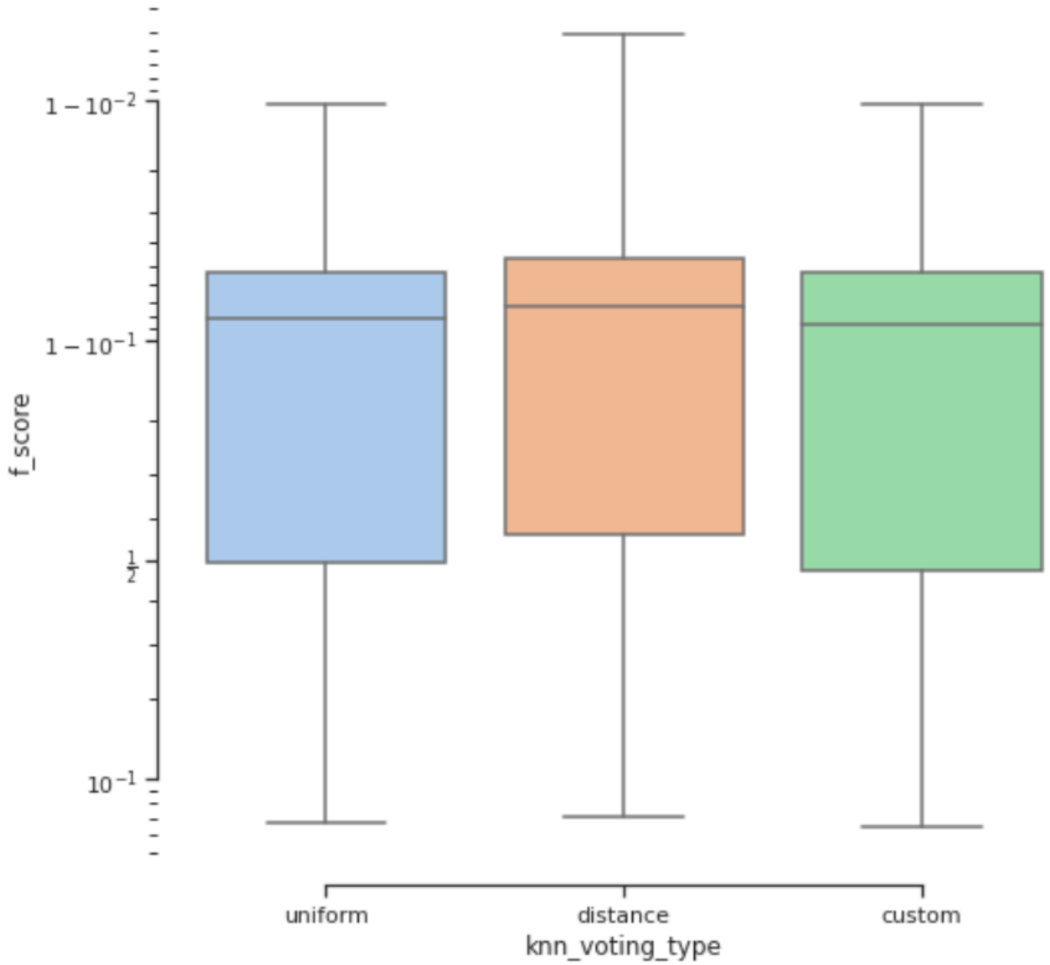
[72]:
```
plots_for_dataset(df.copy(), 'wine')
```

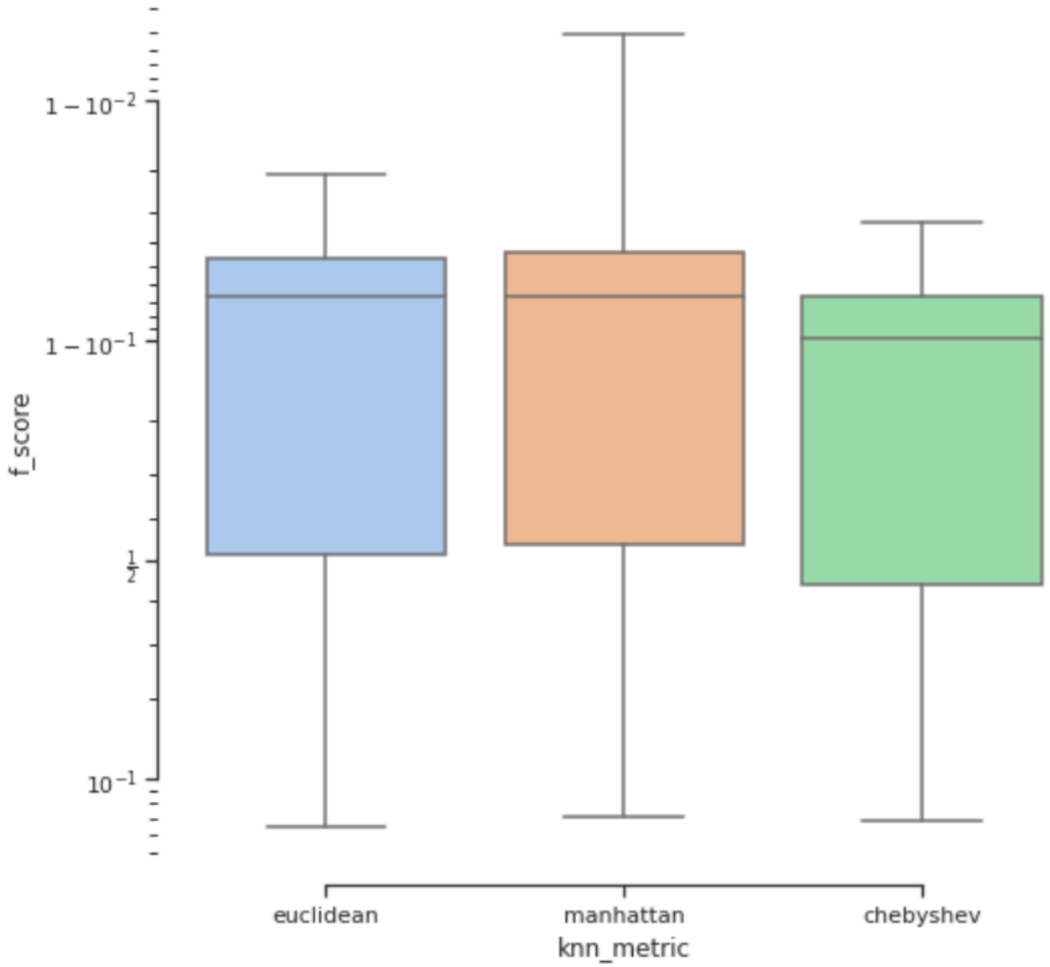

$1 - 10^{-2}$
$1 - 10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
False
True
dataset_shuffle


$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
2
5
10
k_folds_splits

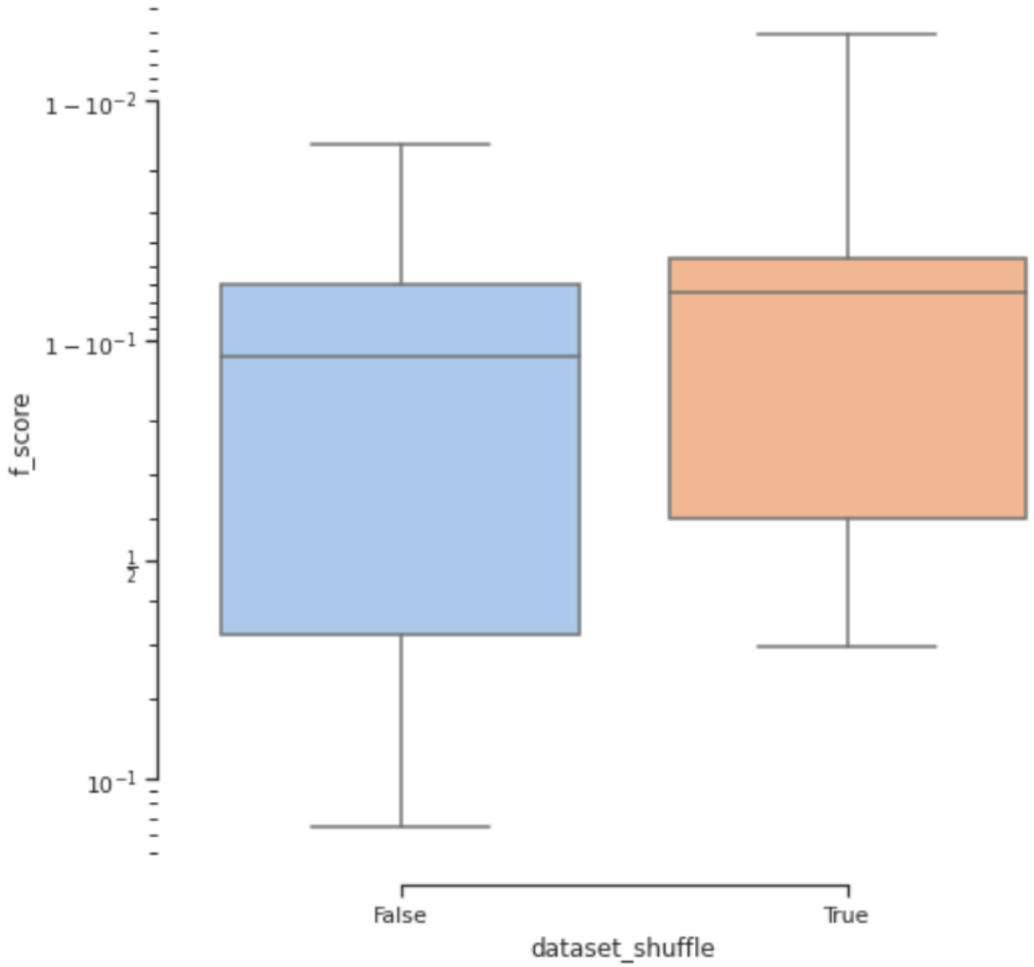

$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
False
True
k_folds_stratified

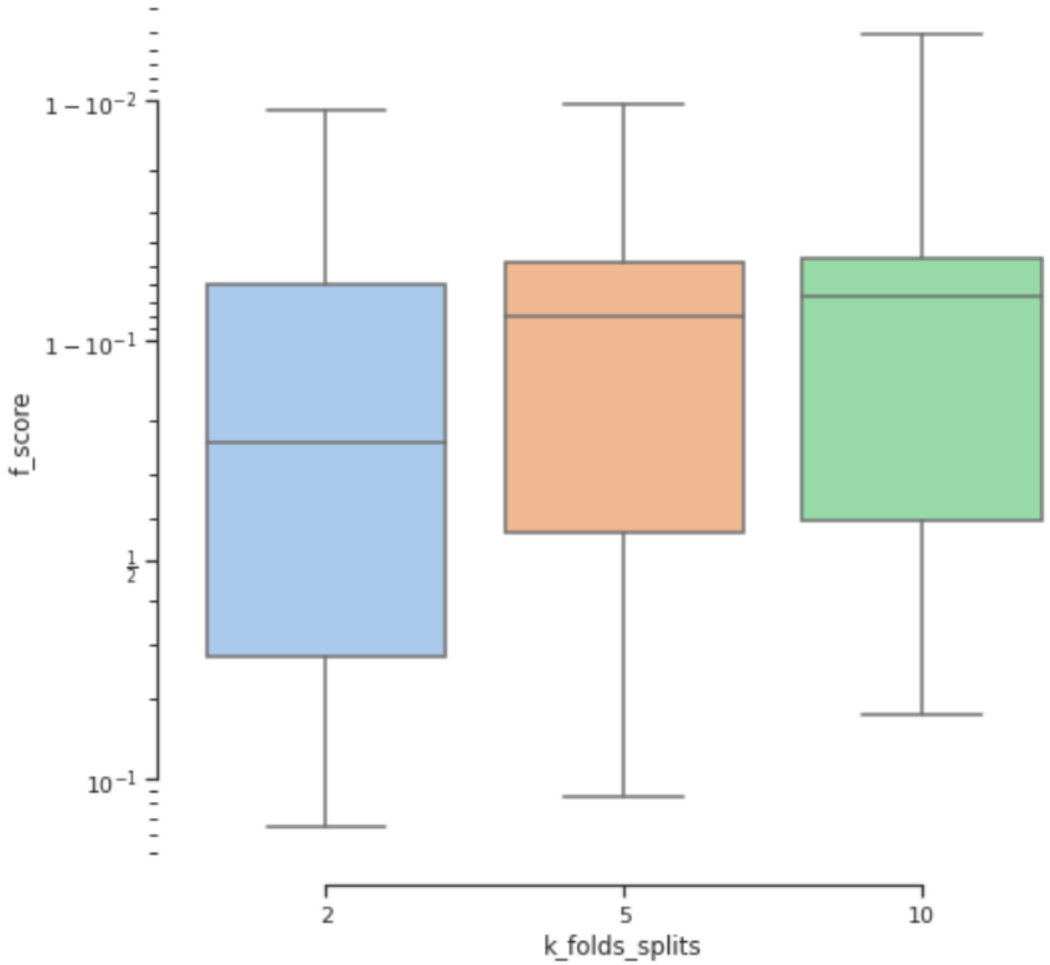

f_score
$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
knn_k_neighbours

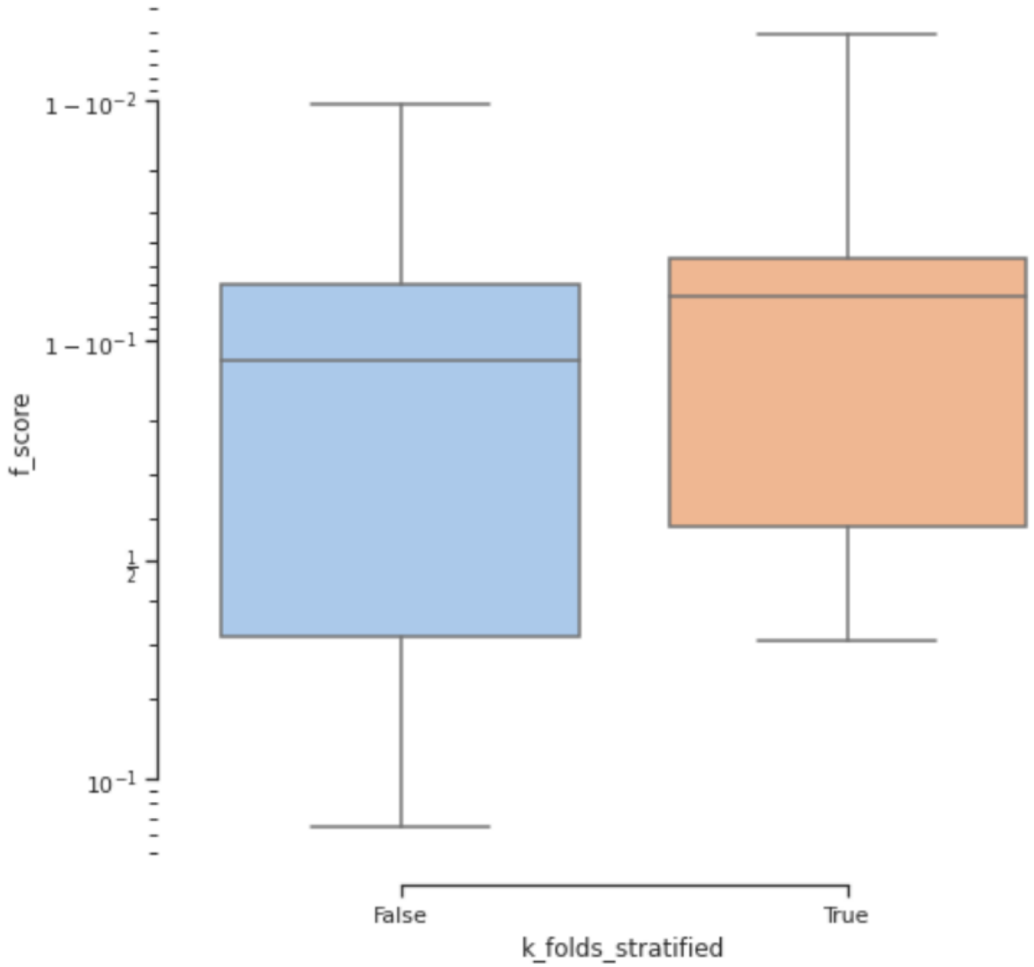

$1 - 10^{-2}$
$1 - 10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
uniform
distance
custom
knn_voting_type

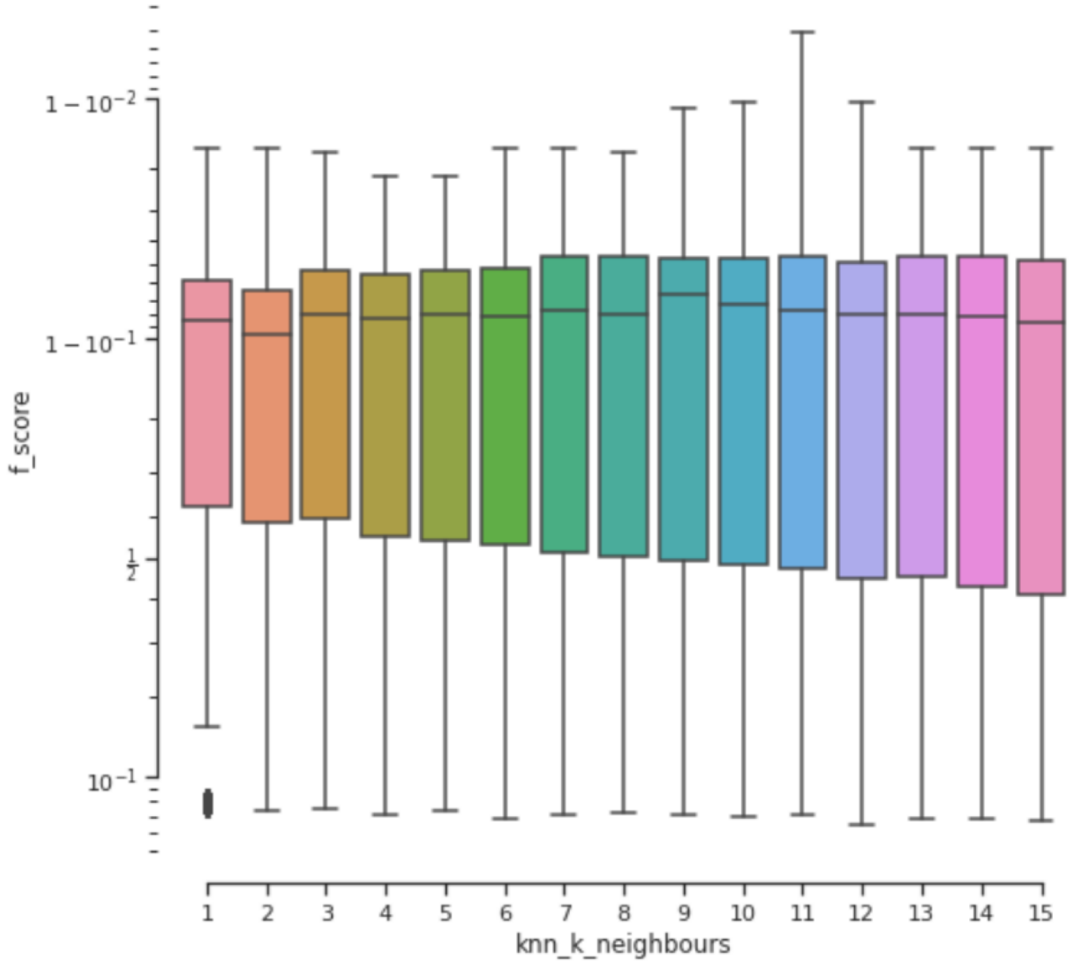


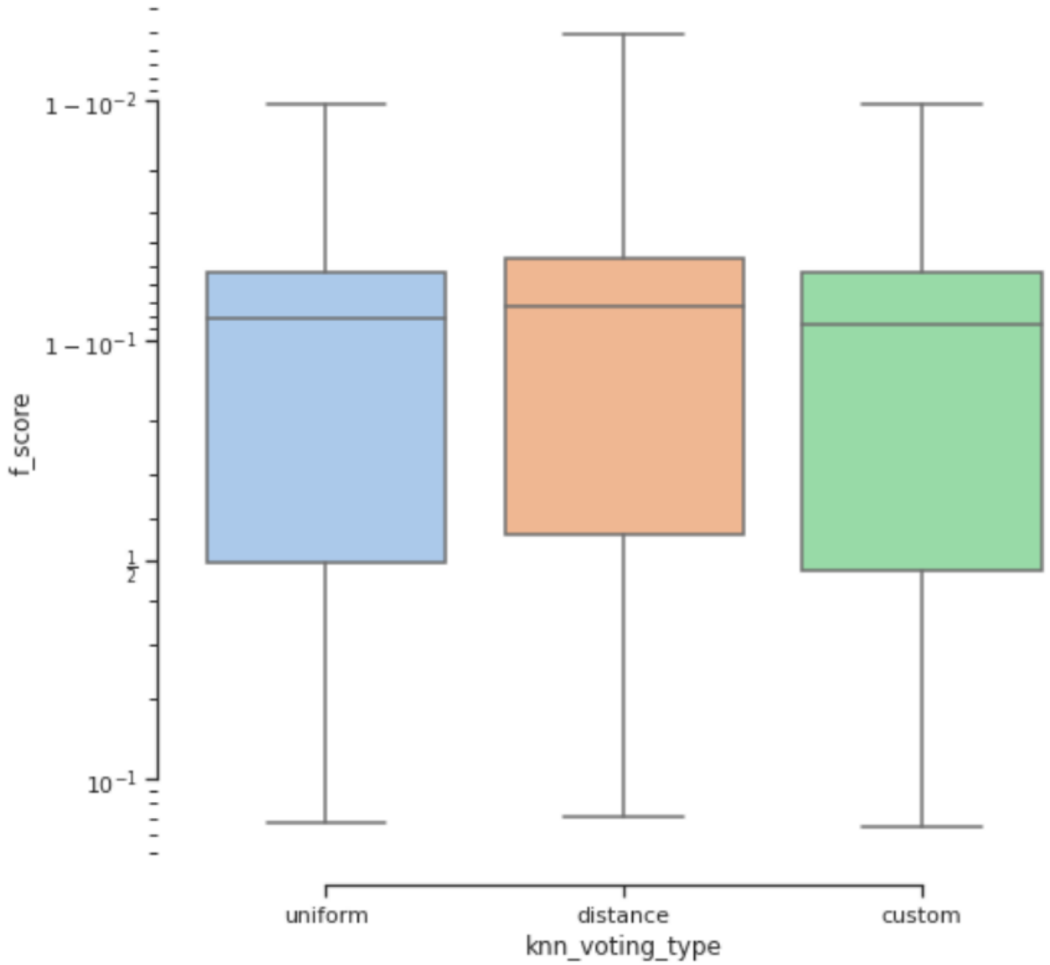
[73]:
```
plots_for_dataset(df.copy(), 'glass')
```

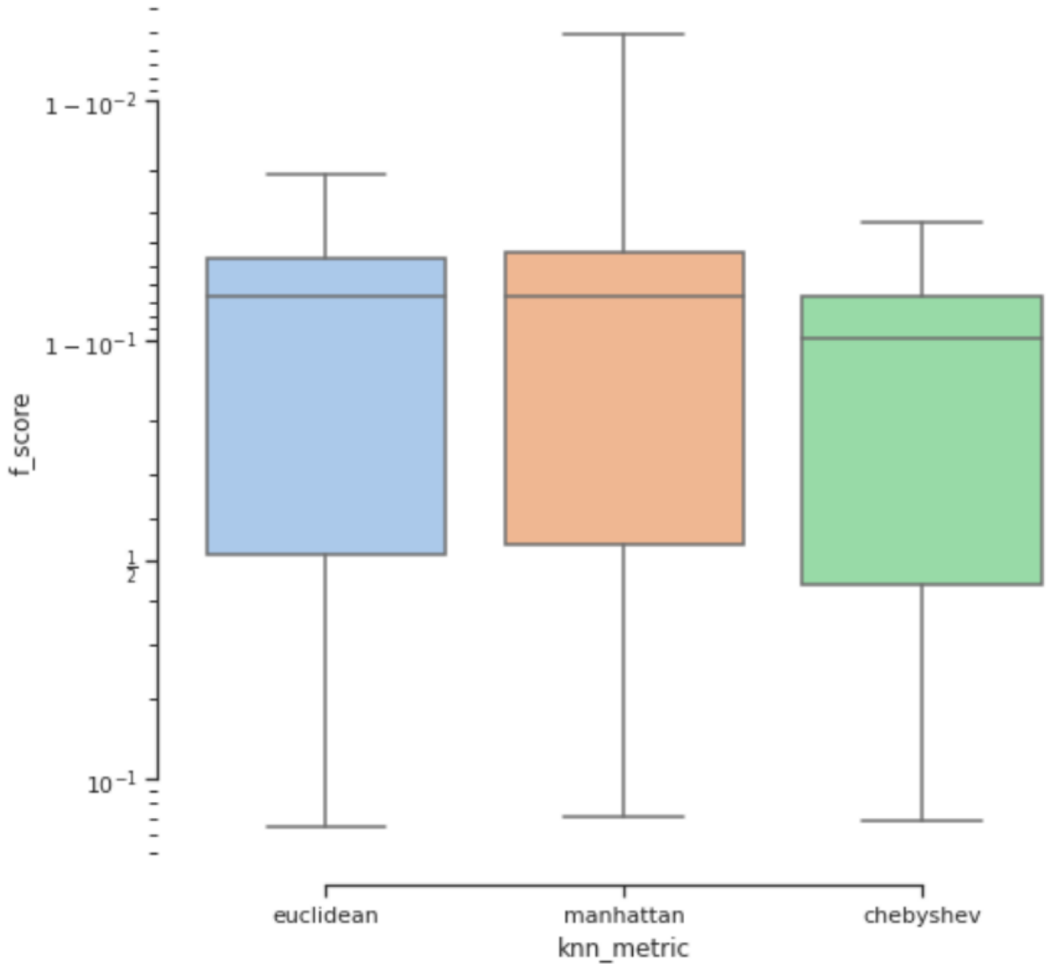

$1 - 10^{-2}$
$1 - 10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
False
True
dataset_shuffle


$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
2
5
10
k_folds_splits


$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
False
True
k_folds_stratified


f_score
$1-10^{-2}$
$1-10^{-1}$
$\frac{1}{2}$
10^{-1}
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
knn_k_neighbours


$1 - 10^{-2}$
$1 - 10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
uniform
distance
custom
knn_voting_type


$1 - 10^{-2}$
$1 - 10^{-1}$
$\frac{1}{2}$
10^{-1}
f_score
euclidean
manhattan
chebyshev
knn_metric