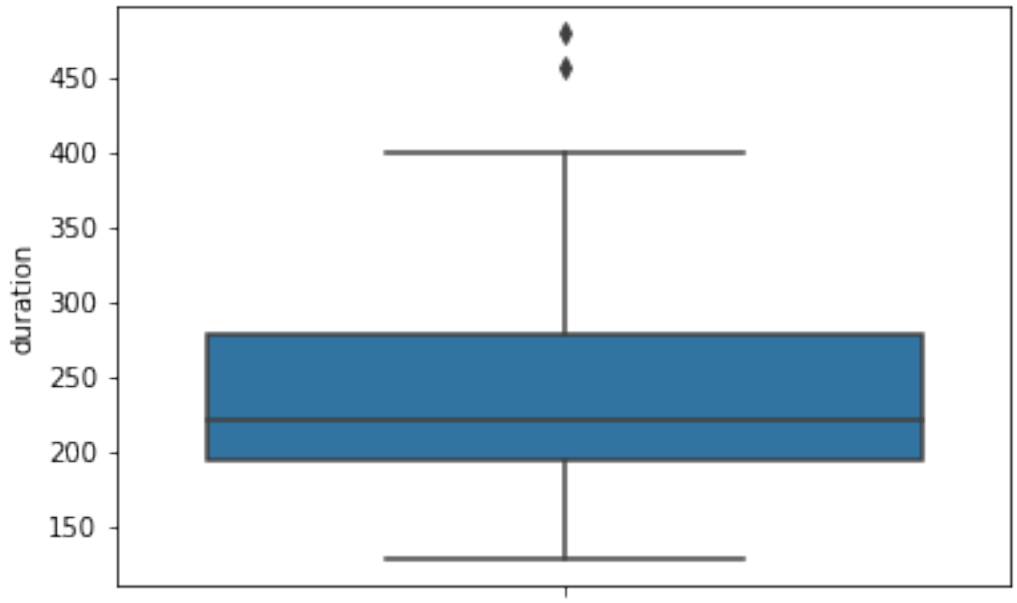
```
import os
import sys
import glob
import sqlite3
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

%matplotlib inline

#draw all the 7 boxplots we require for each column
def drawBoxPlotsByColumn(df, colName, successType):
    print('Global boxplot by',colName)
    sns.boxplot(y = df[colName])
    plt.show()
    if 'e_grammy' in successType:
        print('Boxplot grouped by grammy success')
        sns.boxplot(x = 'e_grammy', y = colName, data = df[[colName,
'e_grammy']] )
        plt.show()
    if 'e_billboard' in successType:
        print('Boxplot grouped by billboard success')
        sns.boxplot(x = 'e_billboard', y = colName, data =
df[[colName, 'e_billboard']] )
        plt.show()
    if 'e_spotify' in successType:
        print('Boxplot grouped by spotify success')
        sns.boxplot(x = 'e_spotify', y = colName, data = df[[colName,
'e_spotify']] )
        plt.show()

def drawBoxPlotsNoOutlierByColumn(df,colName):
    print('Global boxplot by',colName)
    ax = sns.boxplot(y = df[colName], fliersize=0)
    plt.show()
    print('Boxplot grouped by grammy success')
    sns.boxplot(x = 'e_grammy', y = colName, data = df[[colName,
'e_grammy']], fliersize=0)
    plt.show()
    print('Boxplot grouped by billboard success')
    sns.boxplot(x = 'e_billboard', y = colName, data = df[[colName,
'e_billboard']], fliersize=0)
    plt.show()
    print('Boxplot grouped by spotify success')
    sns.boxplot(x = 'e_spotify', y = colName, data = df[[colName,
'e_spotify']], fliersize=0)
    plt.show()

def drawAllBoxPlots(df, successType):
```

```
    columns = df.columns[:-5]
    for col in columns:
        drawBoxPlotsByColumn(df, col, successType)

def drawAllBoxPlotsNoOutlier(df):
    columns = df.columns[:-5]
    for col in columns:
        drawBoxPlotsNoOutlierByColumn(df, col)

df = pd.read_csv('./binary_final_data.csv')
```

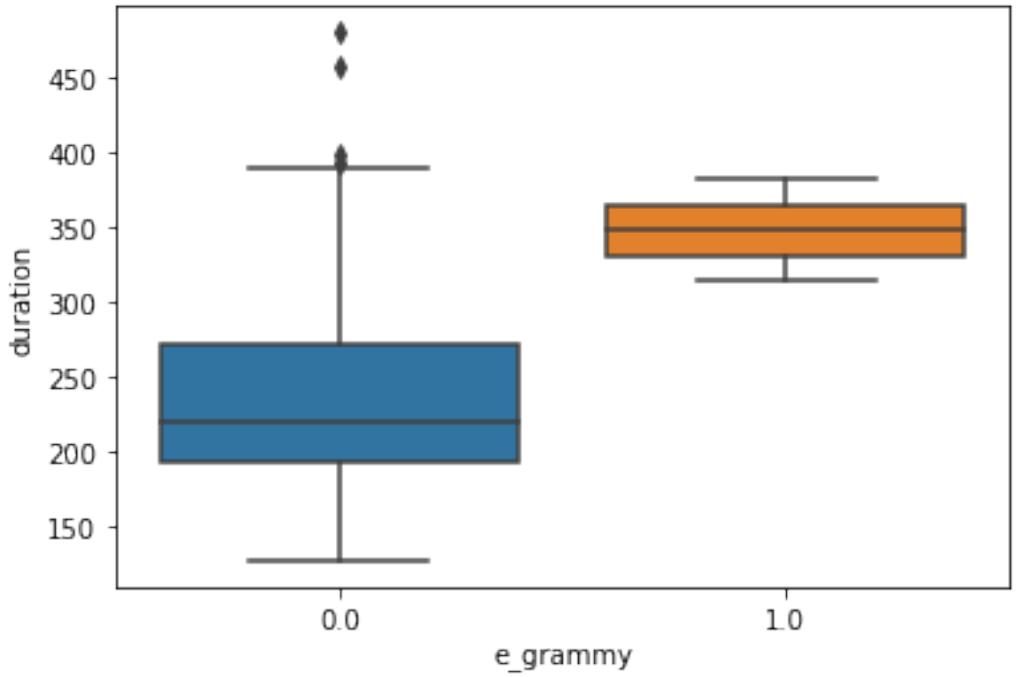
Michael Jackson

```
mickey_df = df[df['artist_name'] == 'Michael Jackson']
drawAllBoxPlots(mickey_df, ['e_grammy'])
drawAllBoxPlots(mickey_df, ['e_billboard'])
drawAllBoxPlots(mickey_df, ['e_spotify'])
```

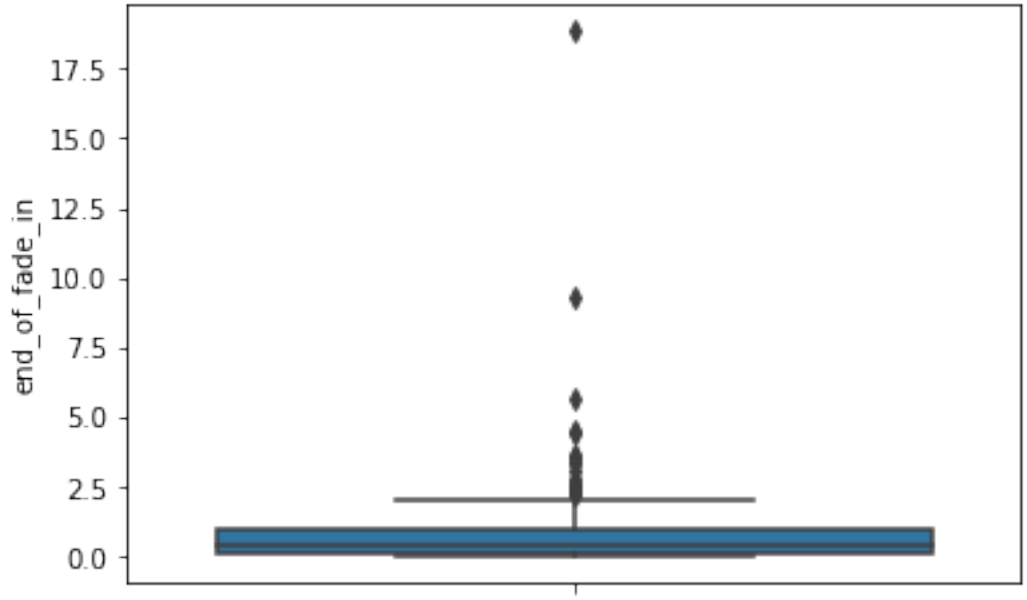
```
Global boxplot by duration
```



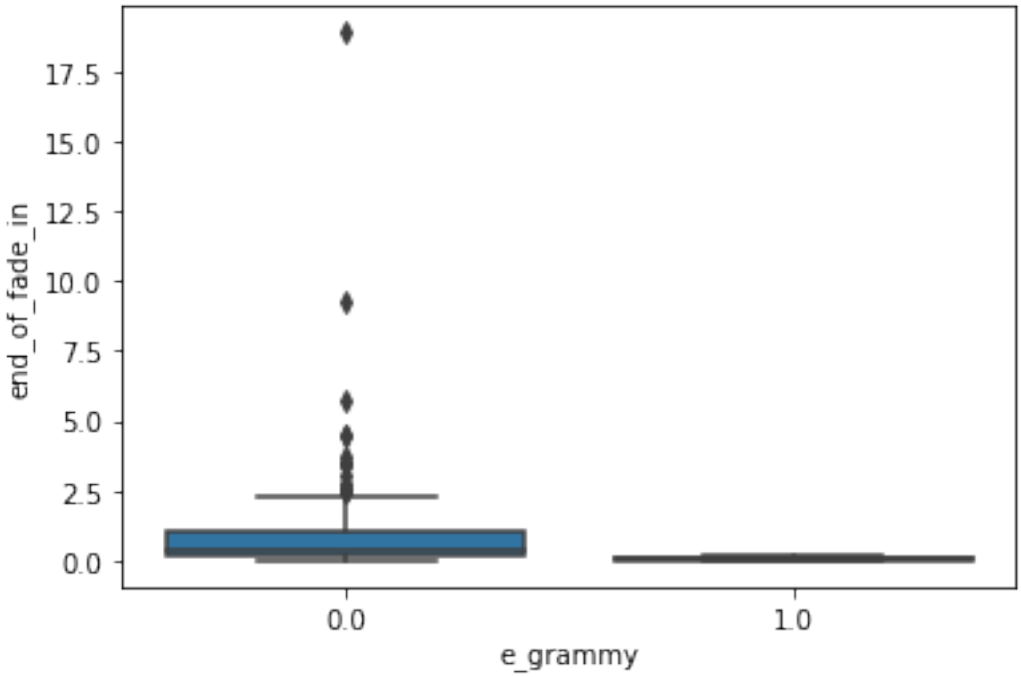

```
Boxplot grouped by grammy success
```



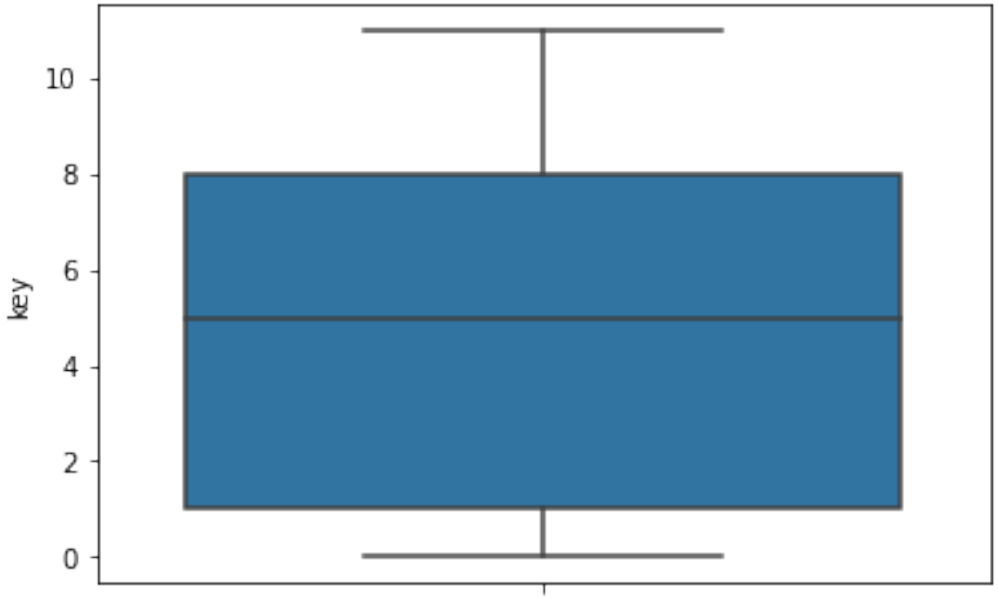
Global boxplot by end_of_fade_in



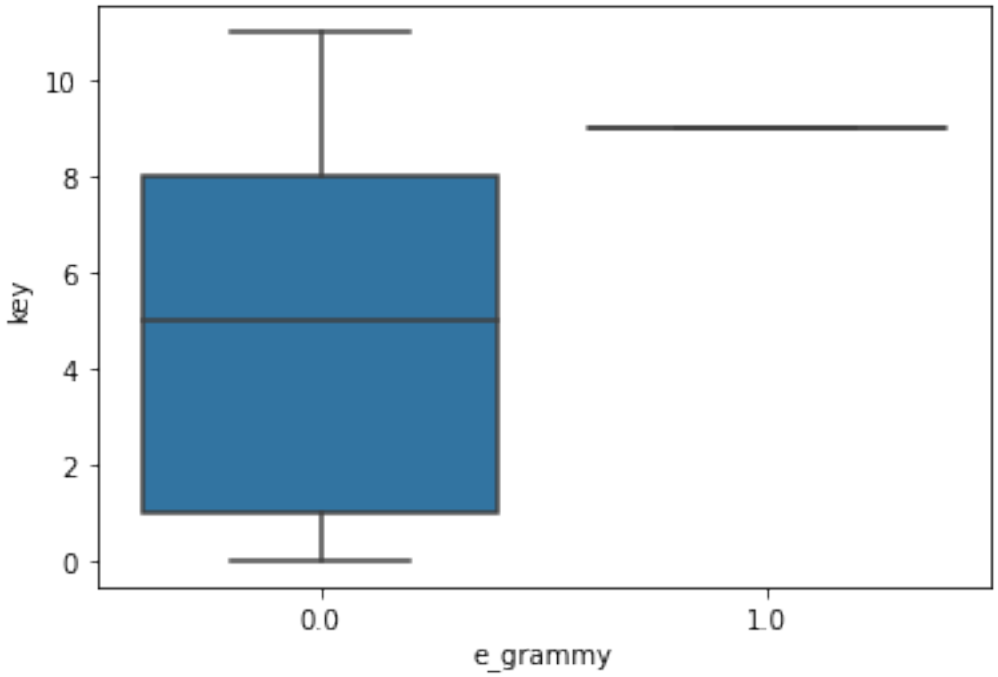
Boxplot grouped by grammy success



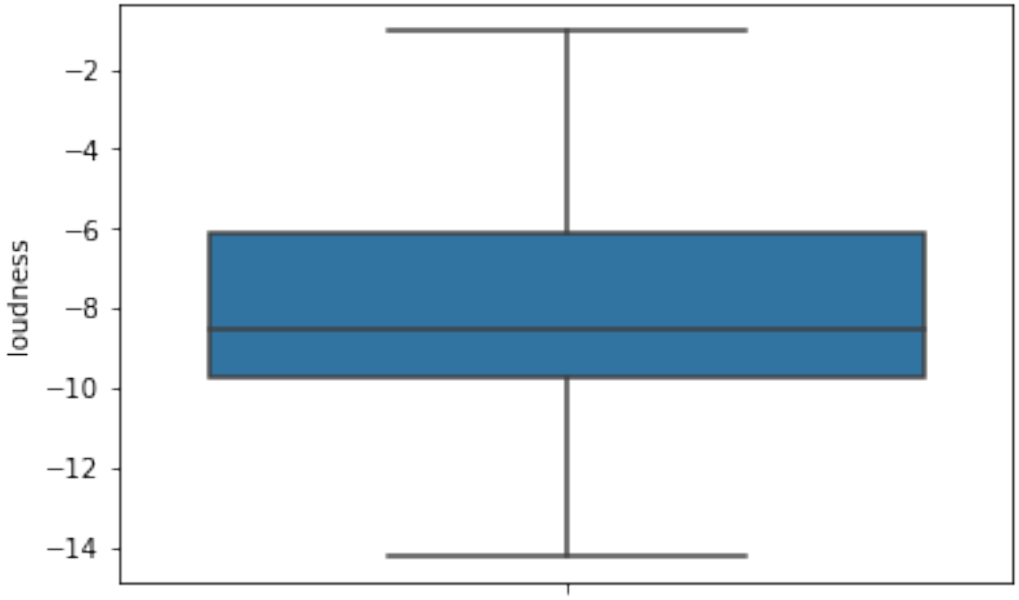
Global boxplot by key



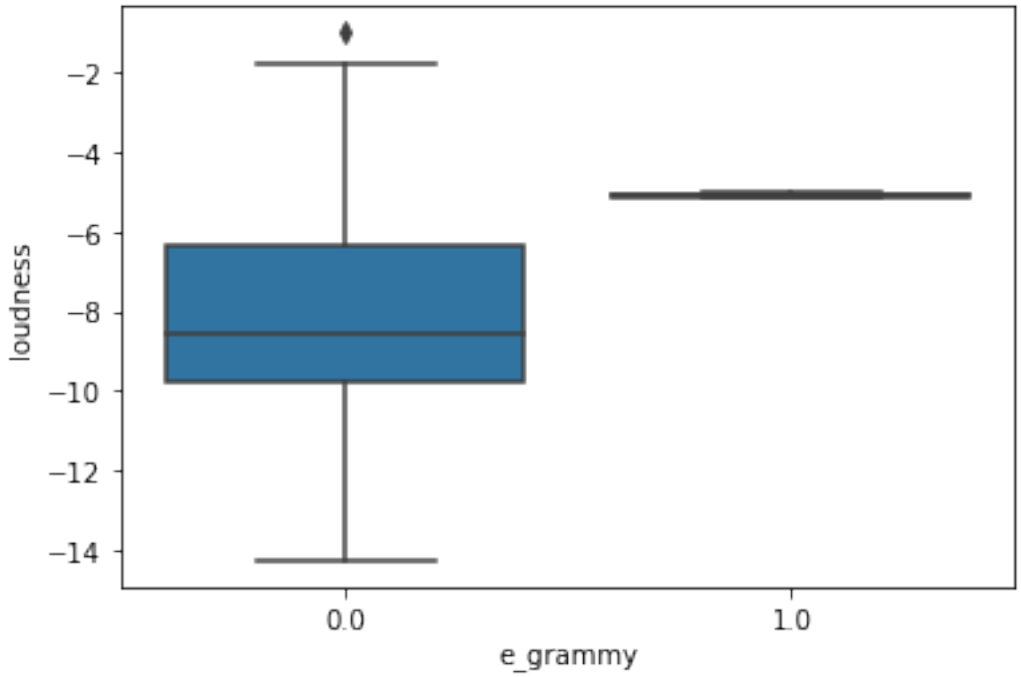
Boxplot grouped by grammy success



Global boxplot by loudness

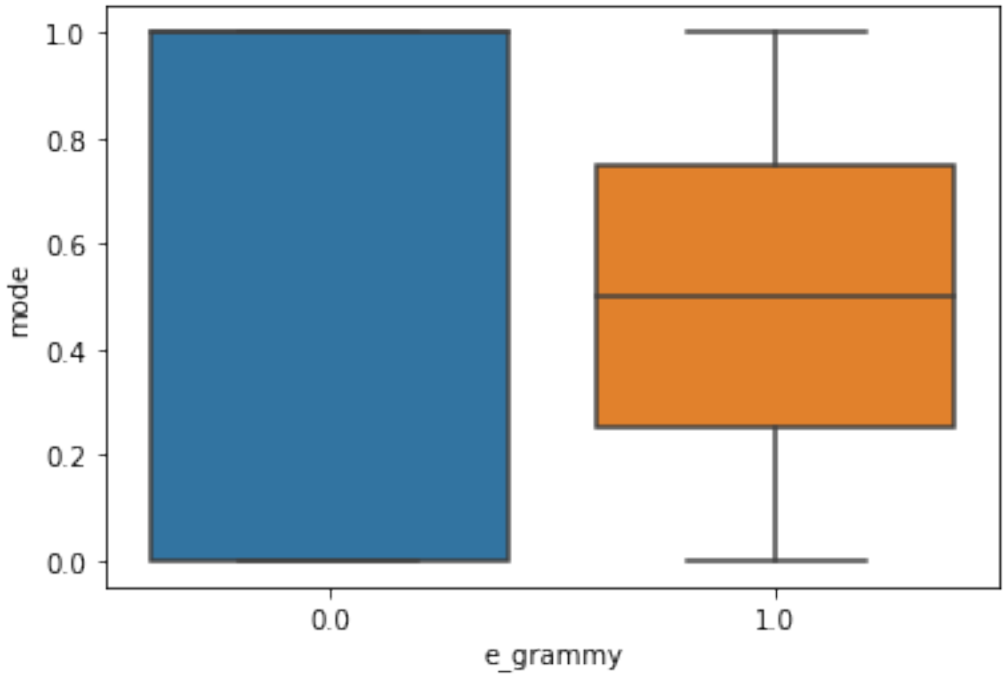


Boxplot grouped by grammy success

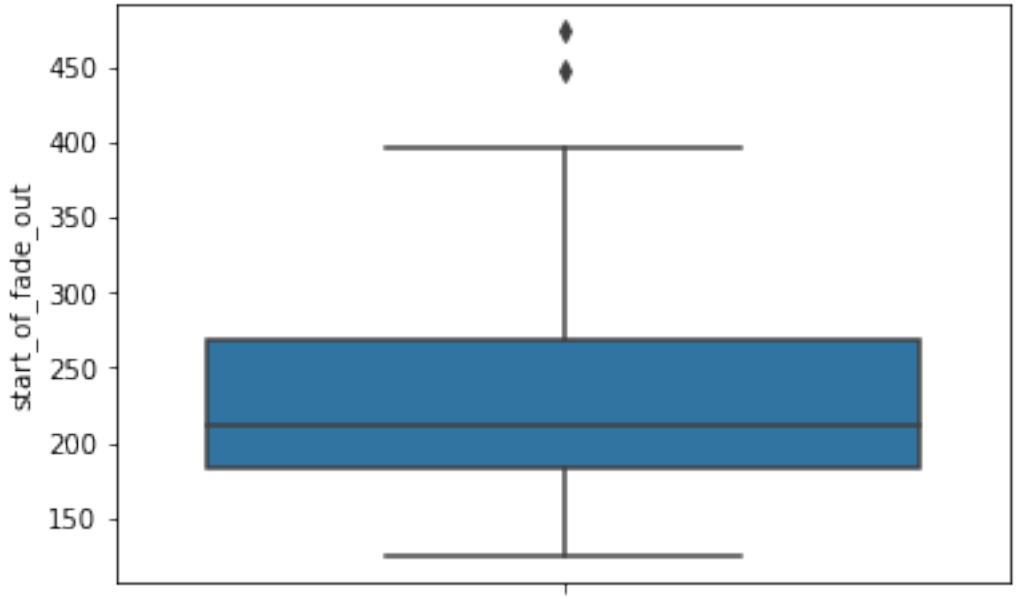


Global boxplot by mode

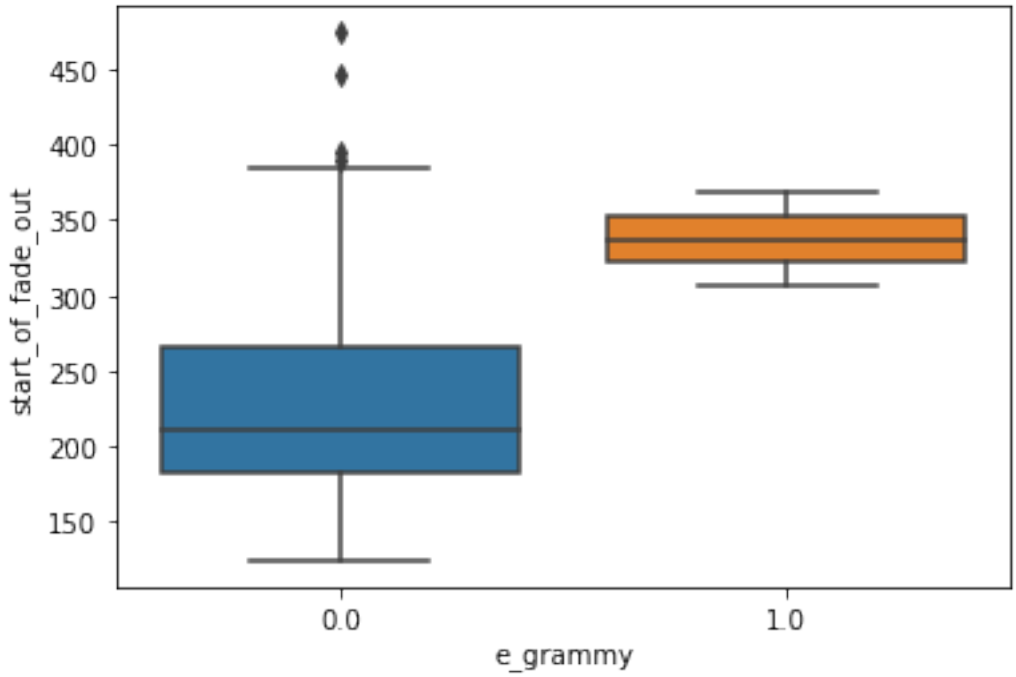
Boxplot grouped by grammy success



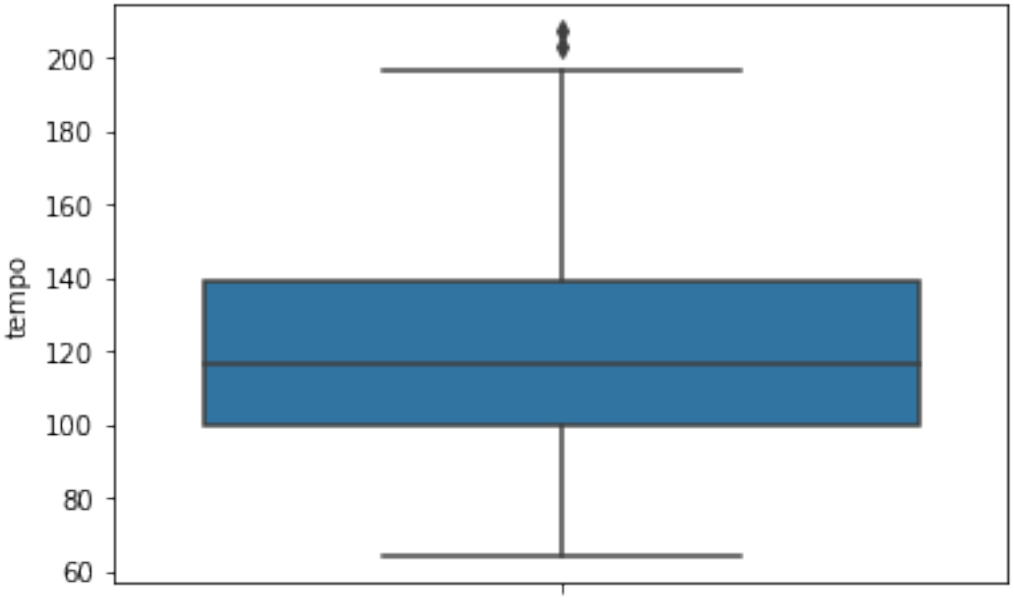
Global boxplot by start_of_fade_out



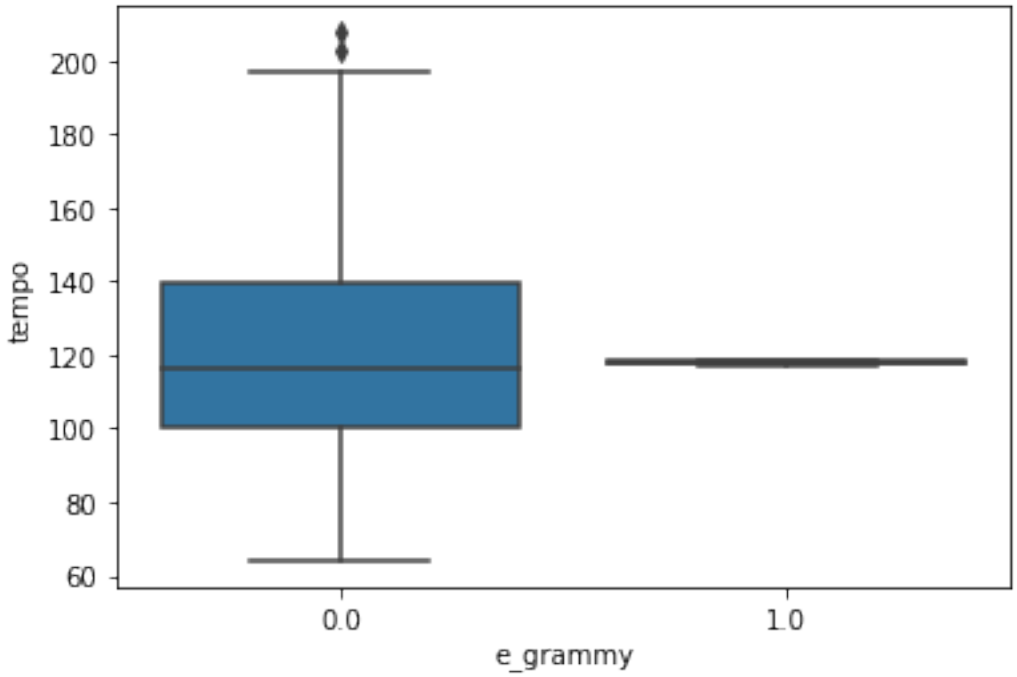
Boxplot grouped by grammy success



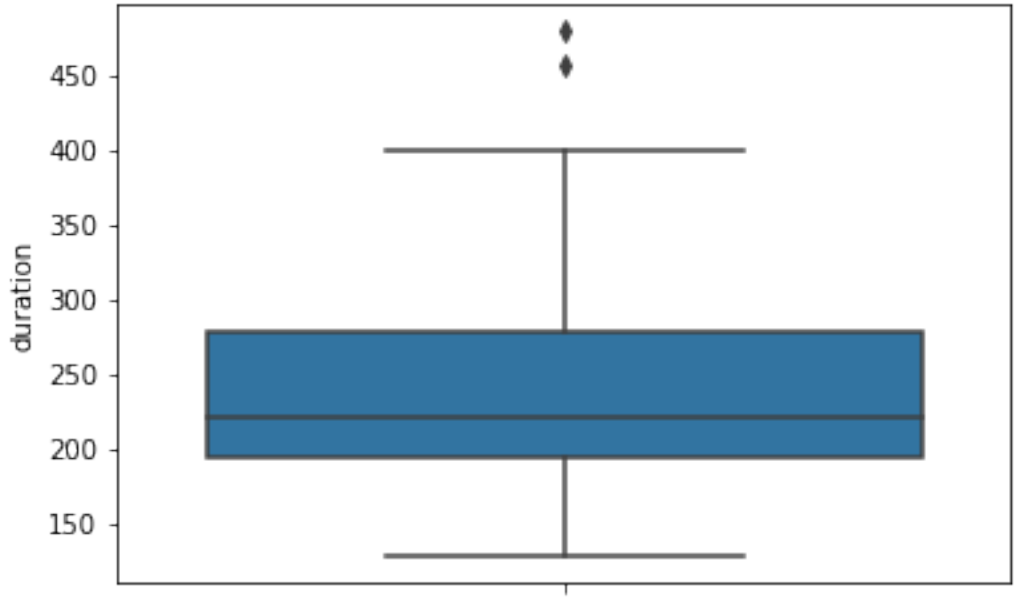
Global boxplot by tempo



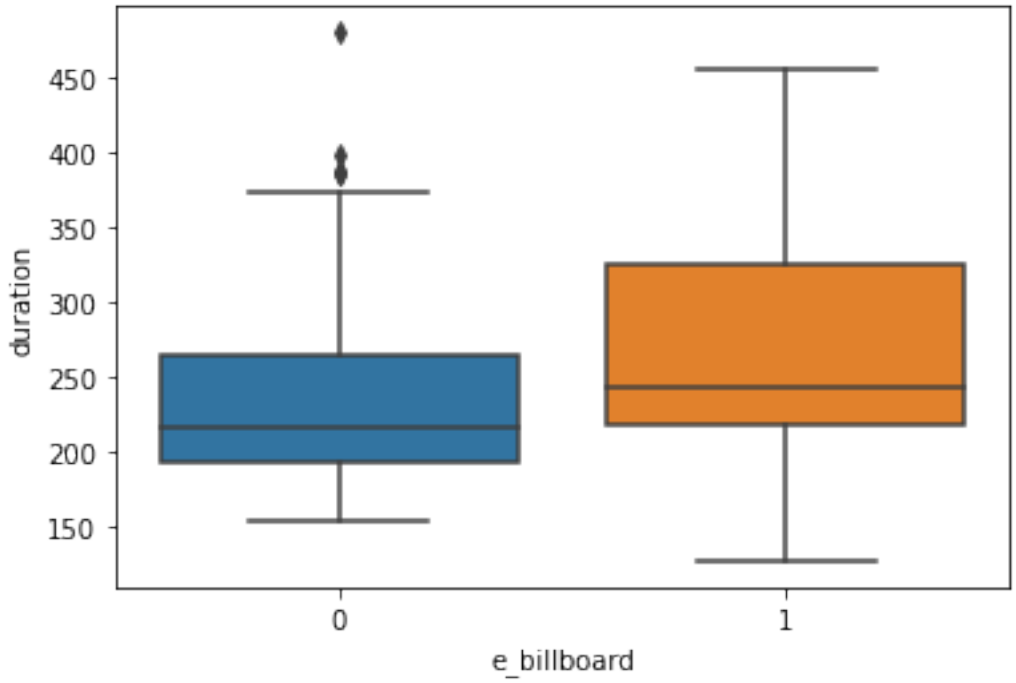
Boxplot grouped by grammy success



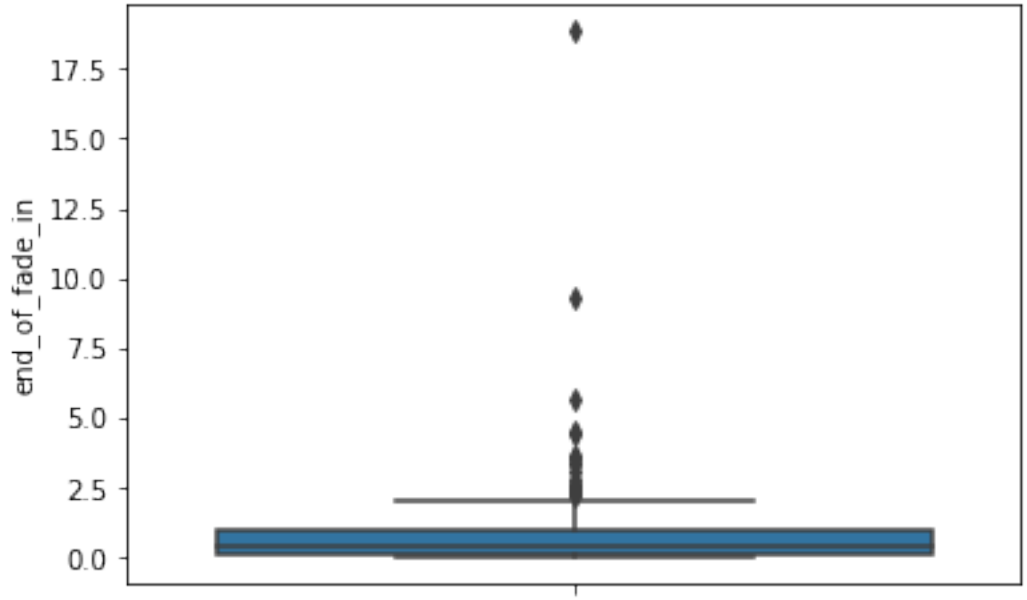
Global boxplot by duration



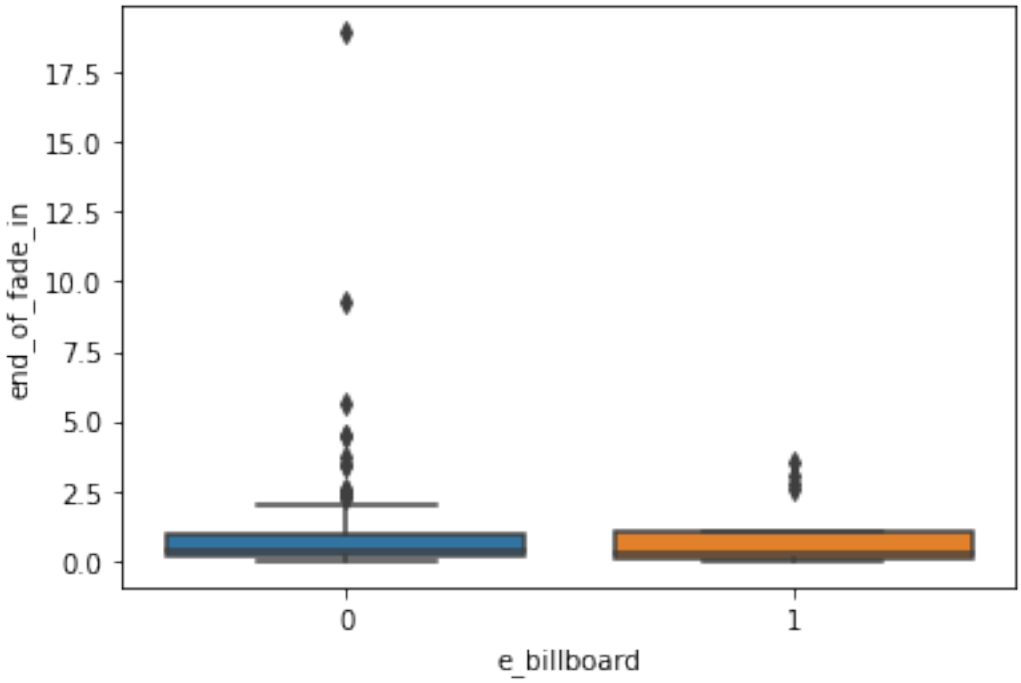
Boxplot grouped by billboard success



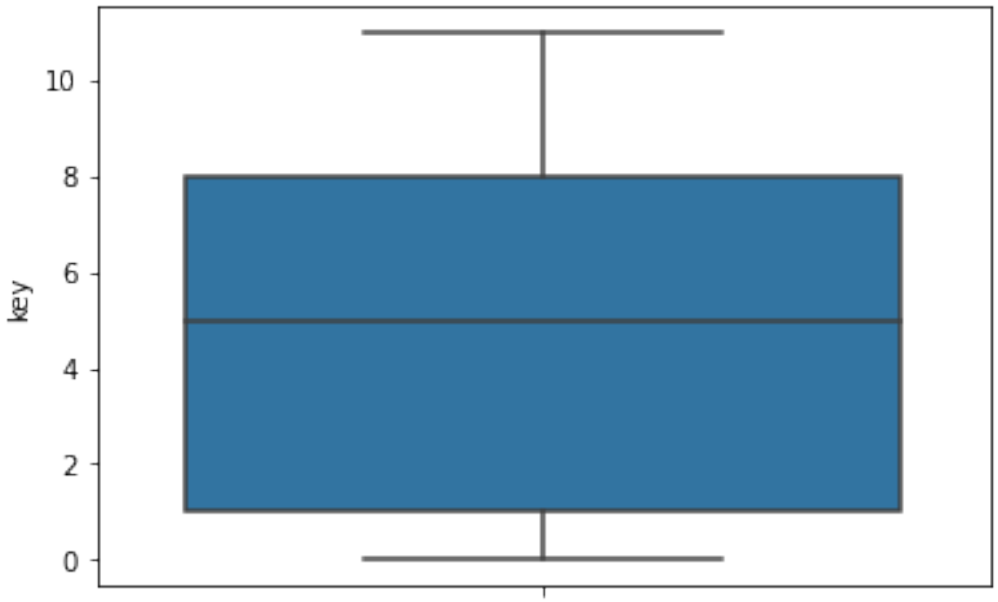
Global boxplot by end_of_fade_in



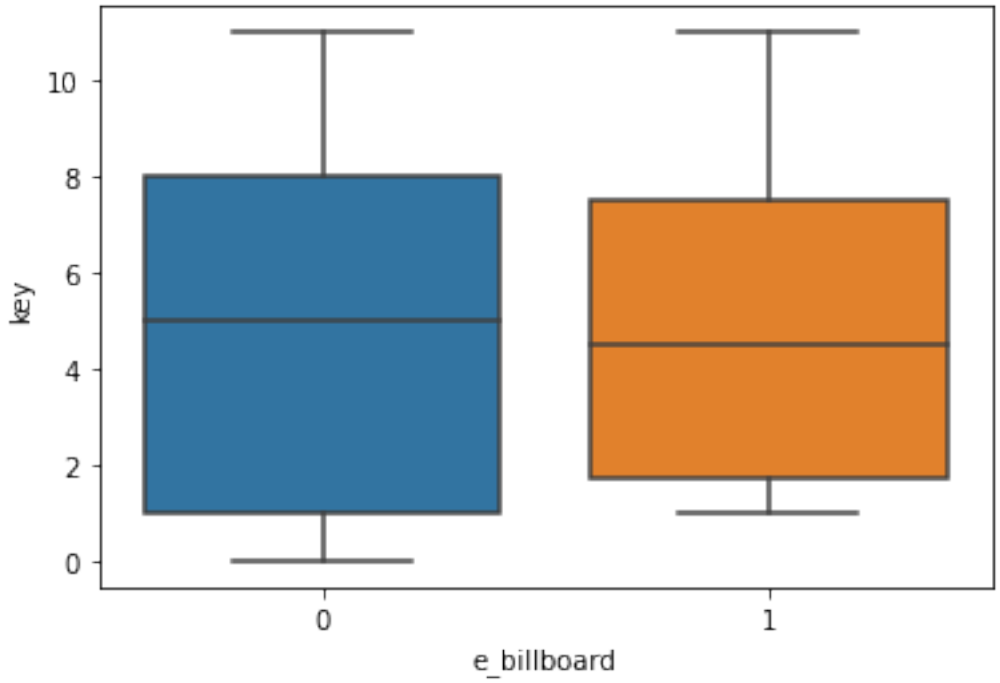
Boxplot grouped by billboard success



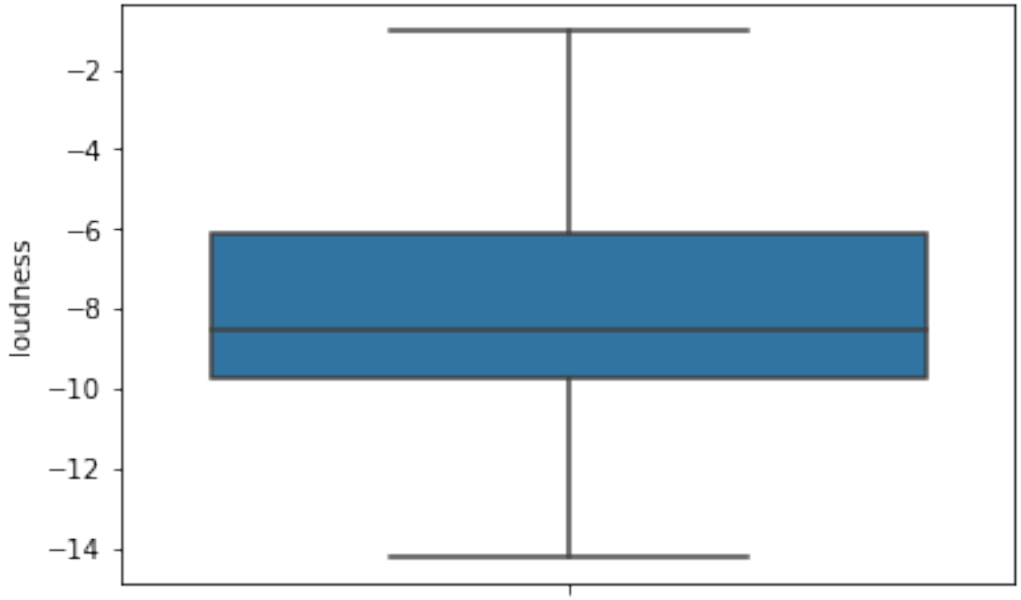
Global boxplot by key



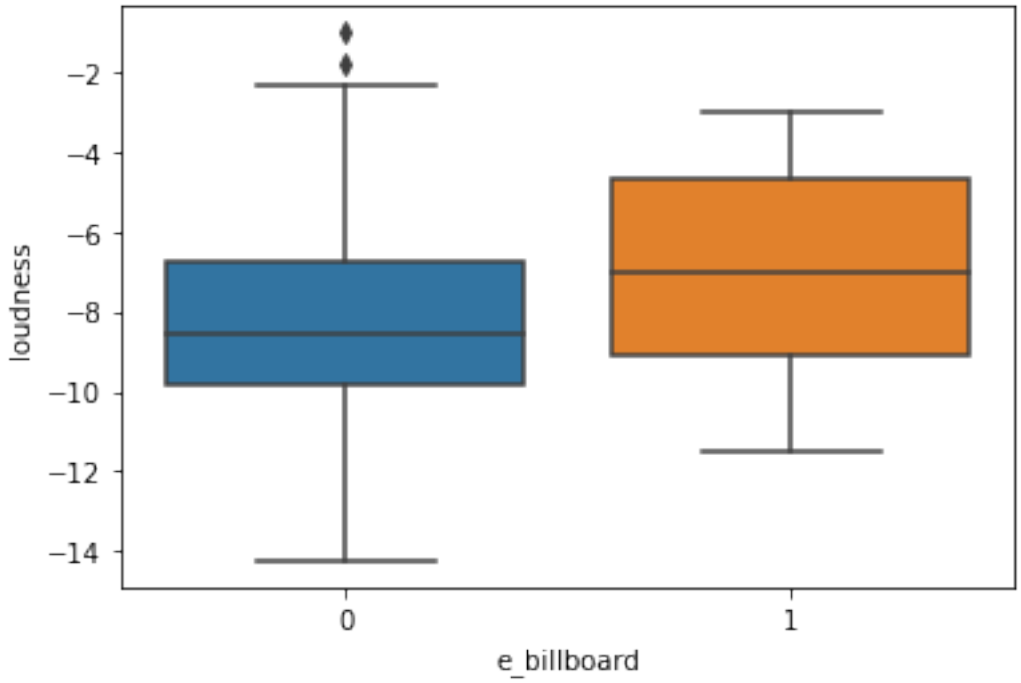
Boxplot grouped by billboard success



Global boxplot by loudness

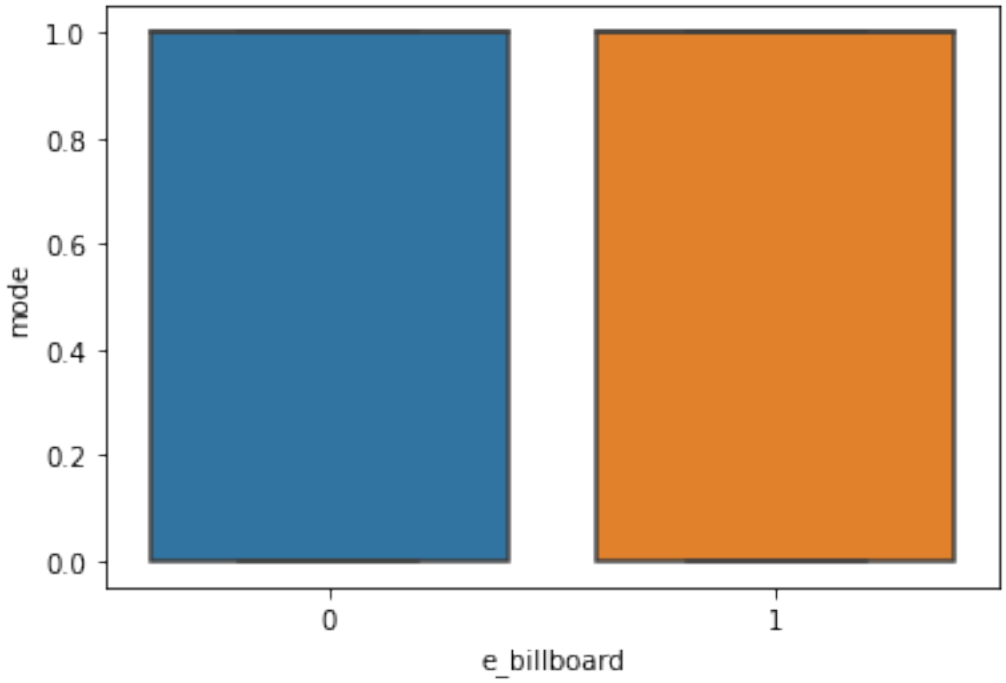


Boxplot grouped by billboard success

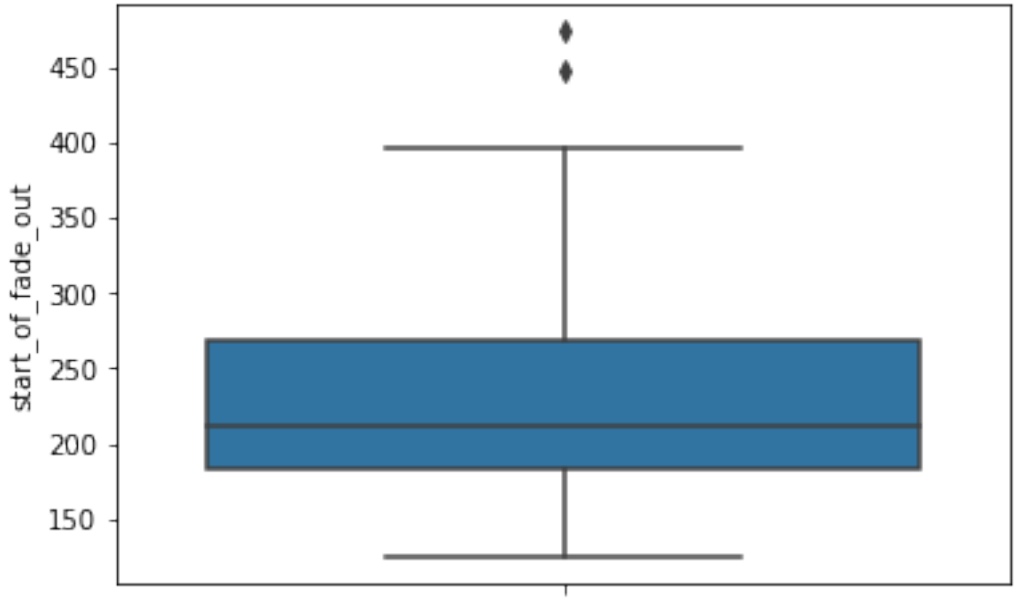


Global boxplot by mode

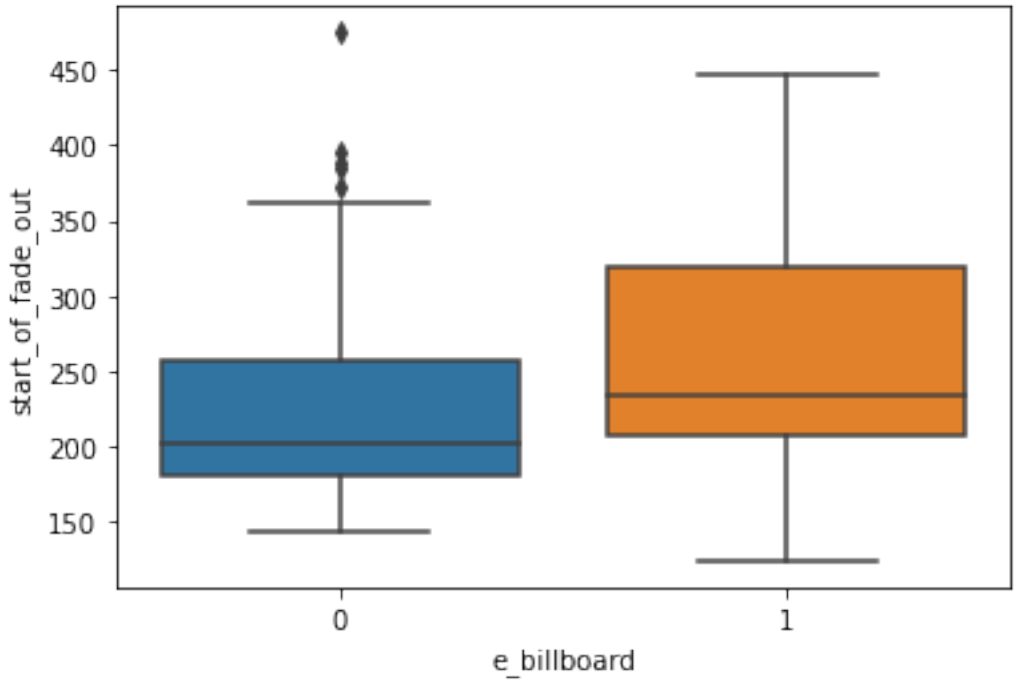
Boxplot grouped by billboard success



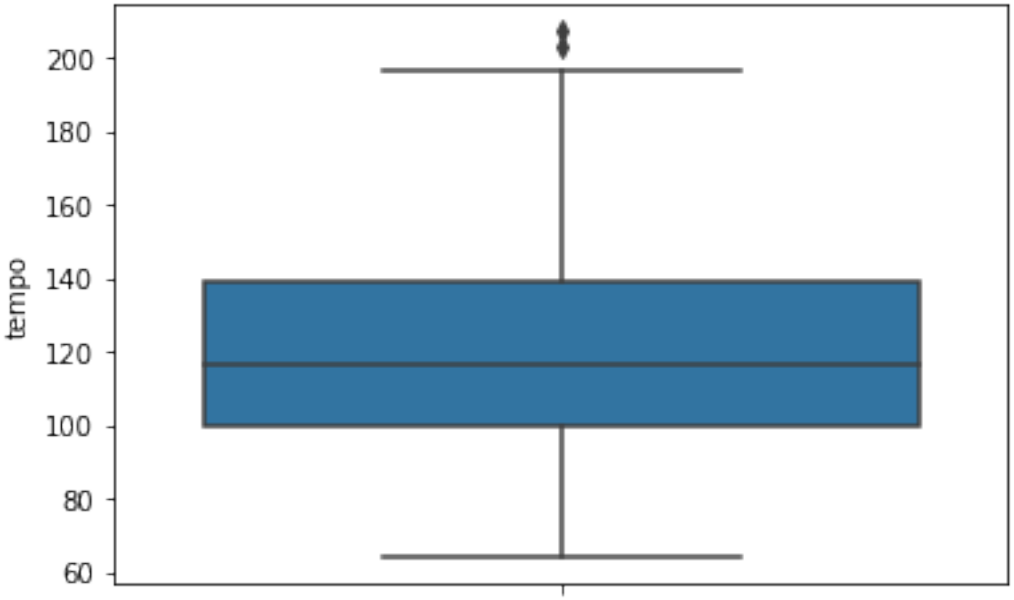
Global boxplot by start_of_fade_out



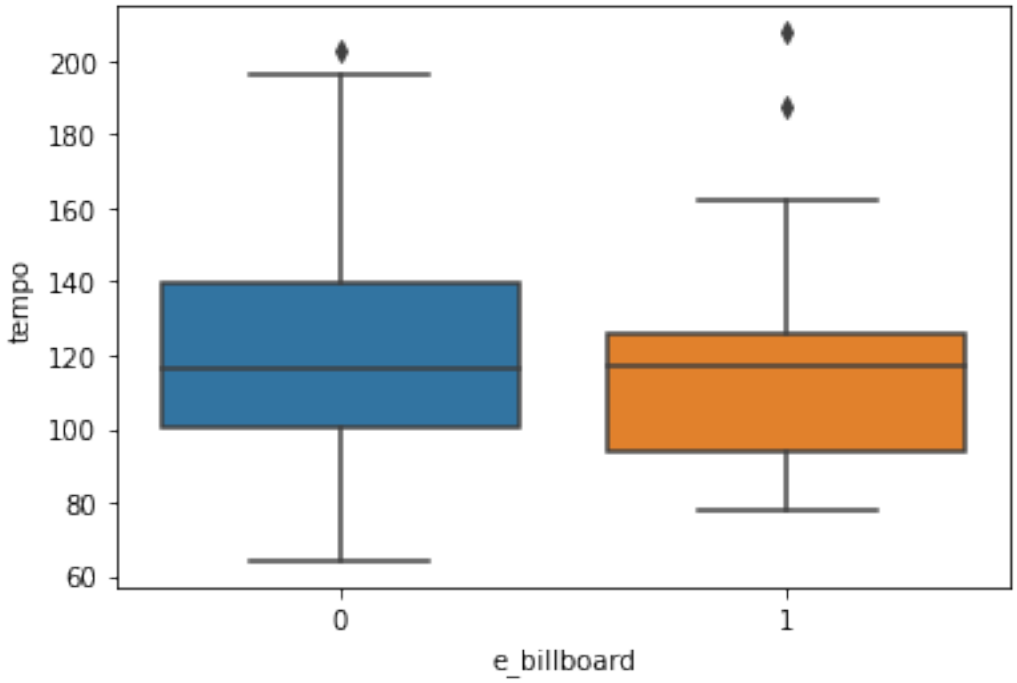
Boxplot grouped by billboard success



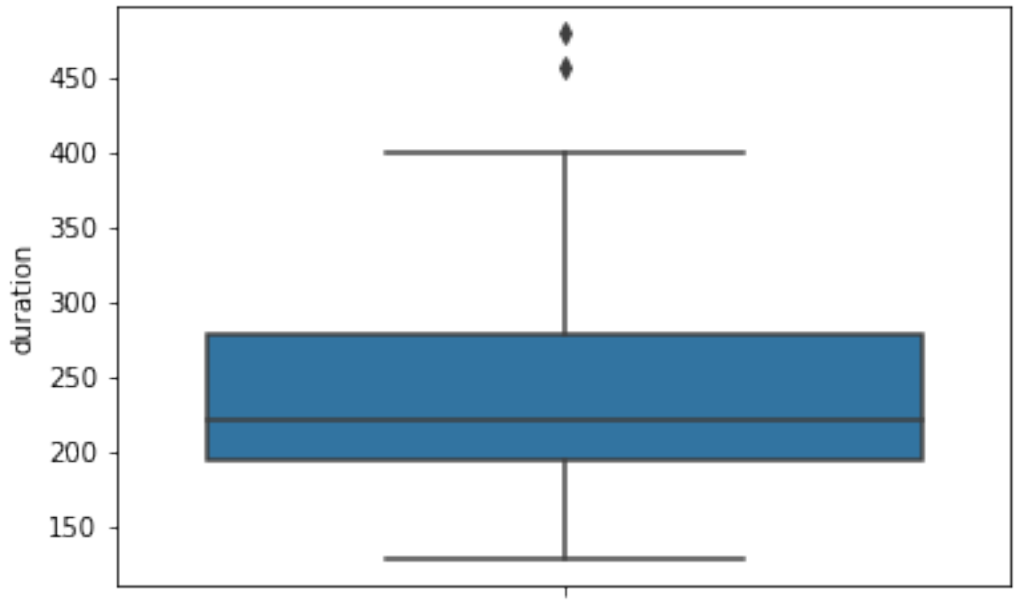
Global boxplot by tempo



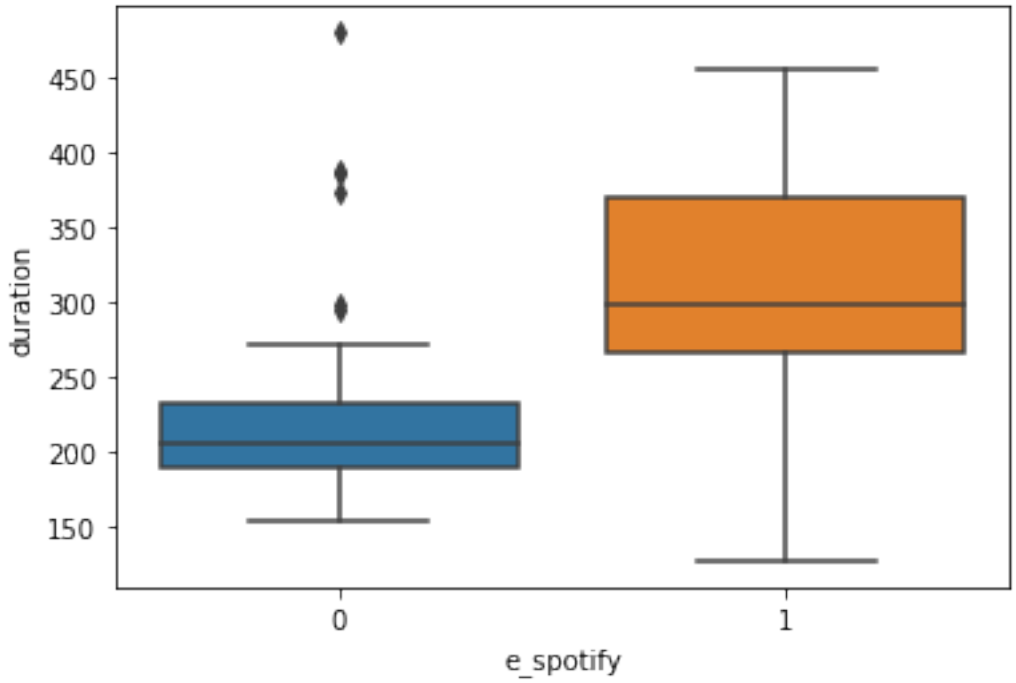
Boxplot grouped by billboard success



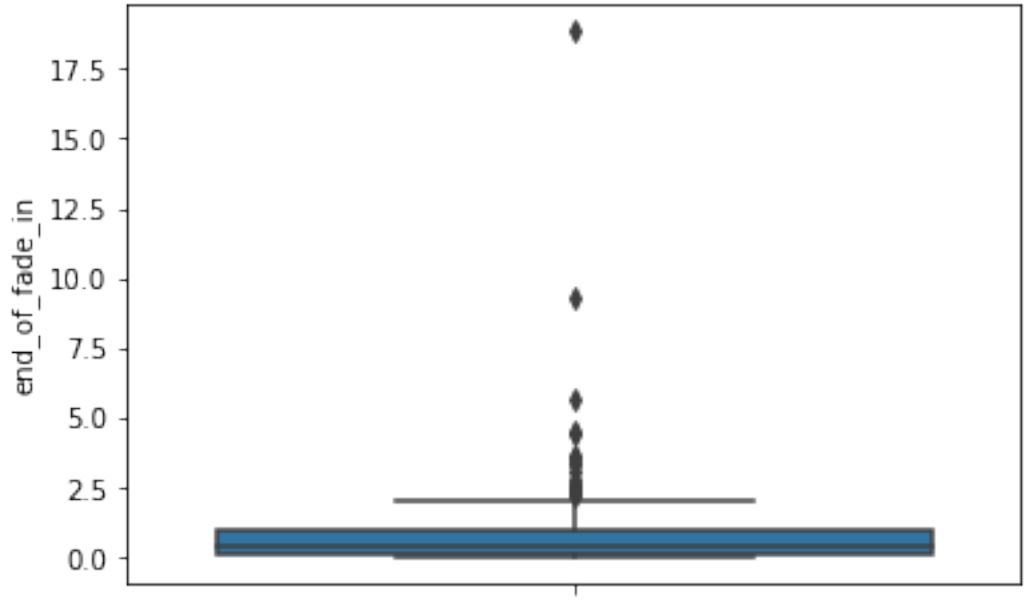
Global boxplot by duration



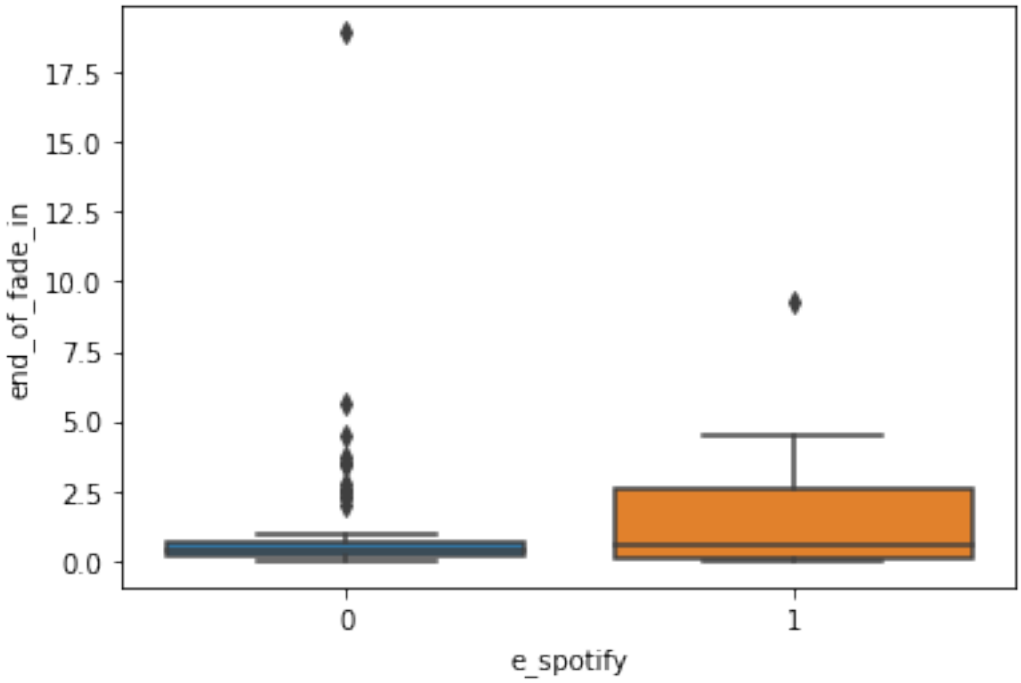
Boxplot grouped by spotify success



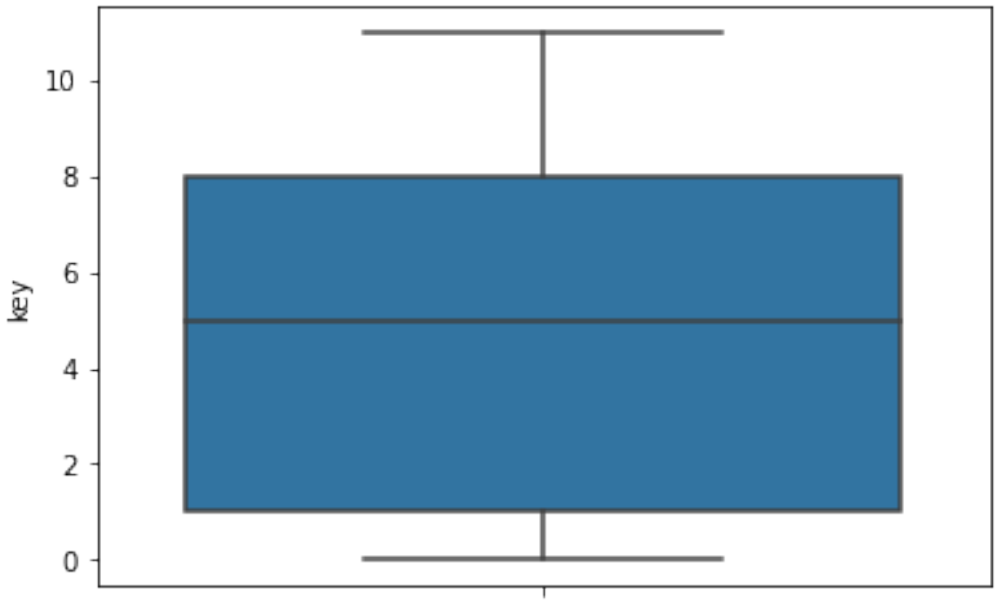
Global boxplot by end_of_fade_in



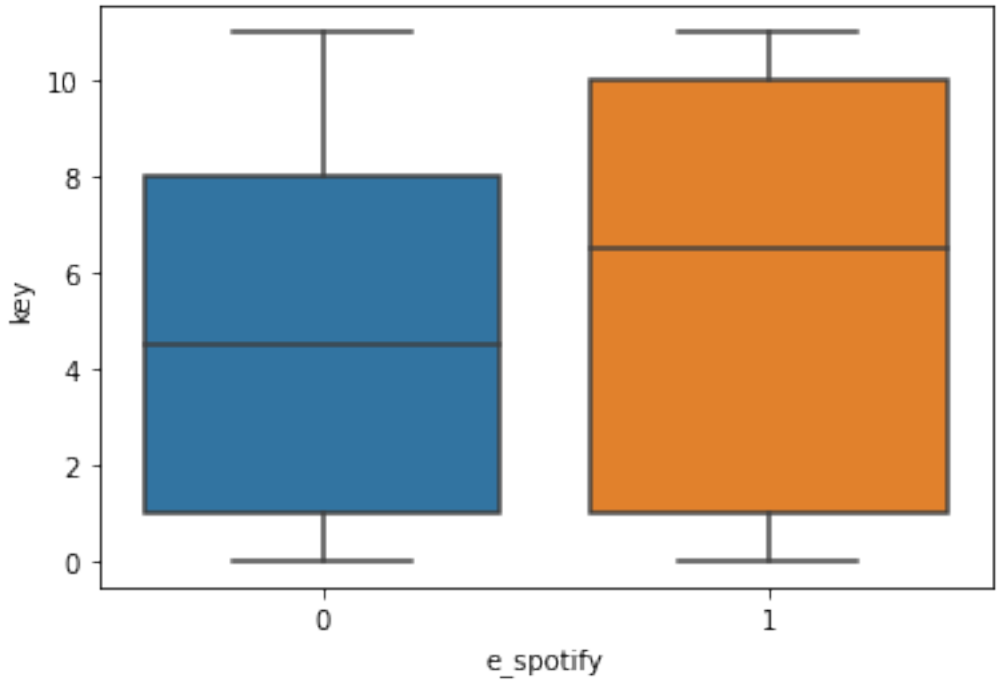
Boxplot grouped by spotify success



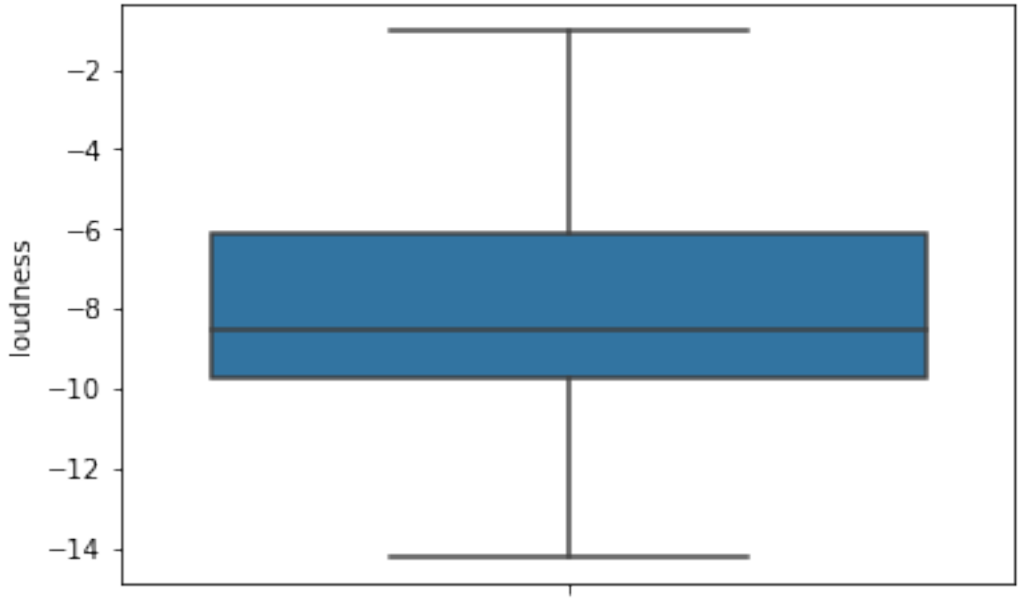
Global boxplot by key



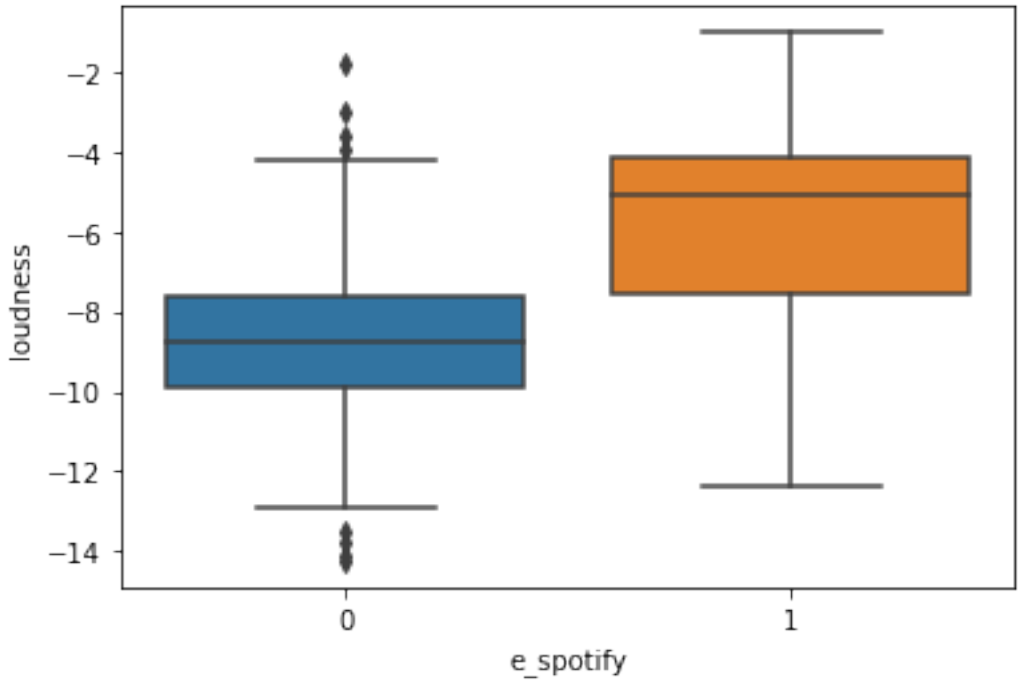
Boxplot grouped by spotify success



Global boxplot by loudness

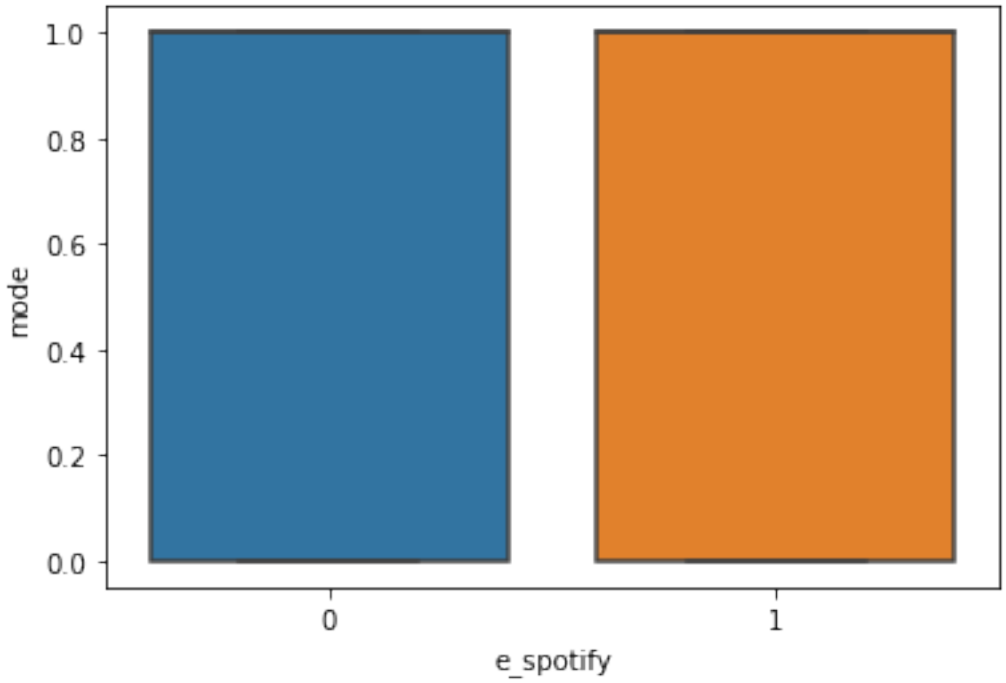


Boxplot grouped by spotify success

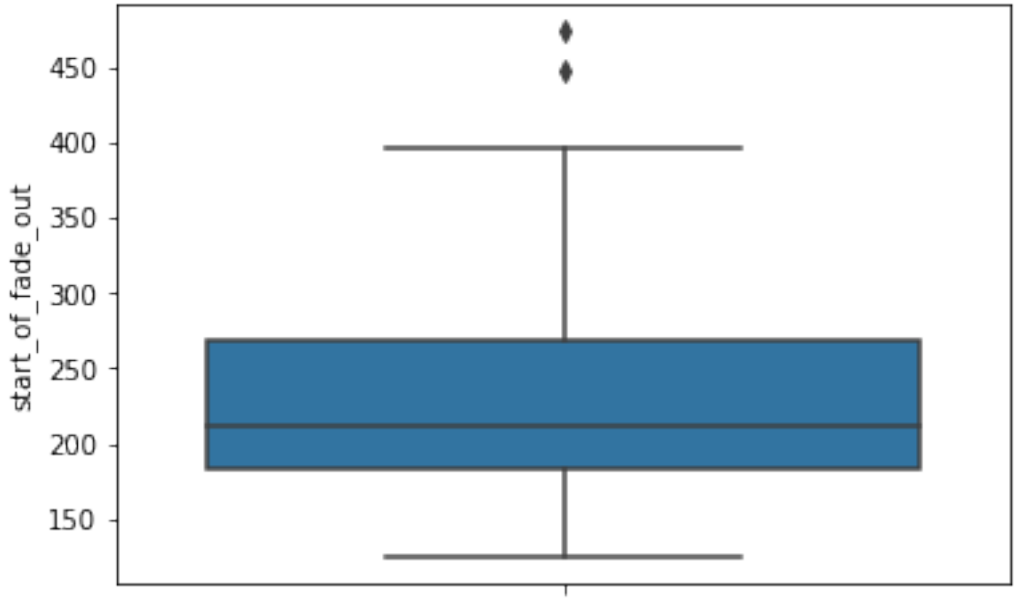


Global boxplot by mode

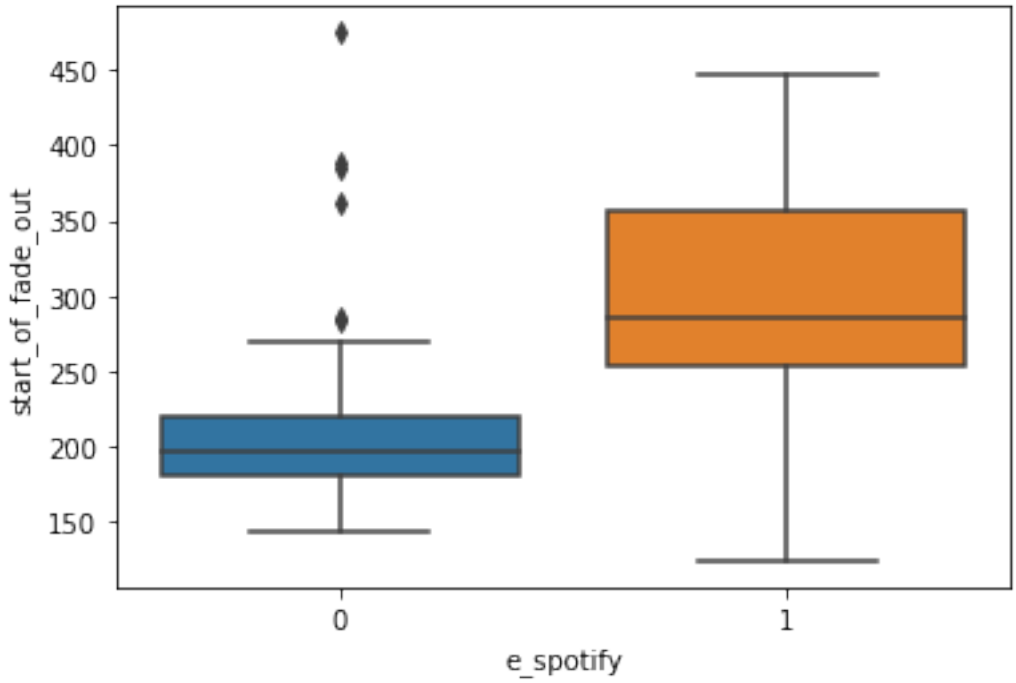
Boxplot grouped by spotify success



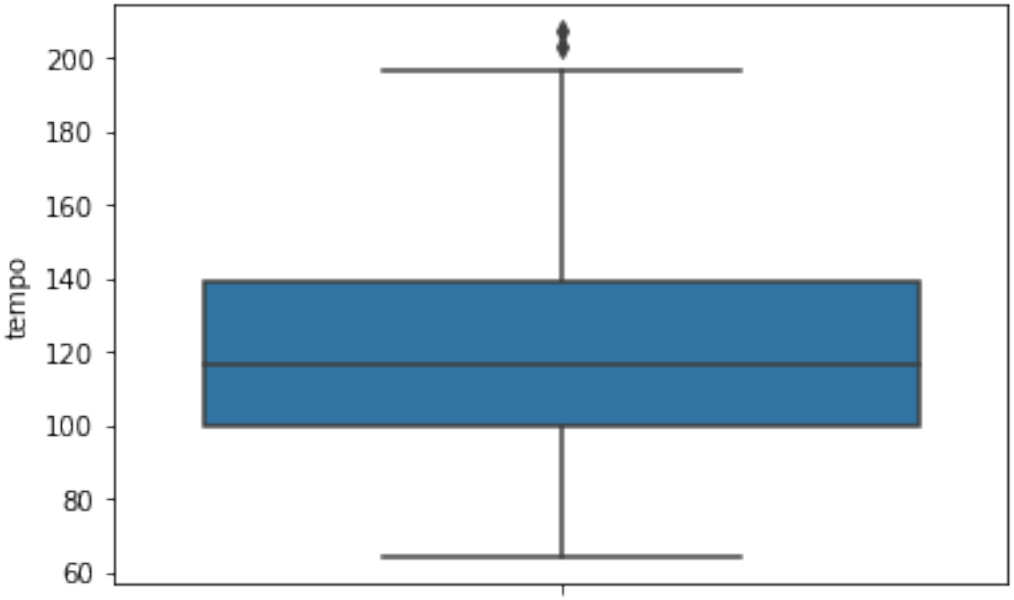
Global boxplot by start_of_fade_out



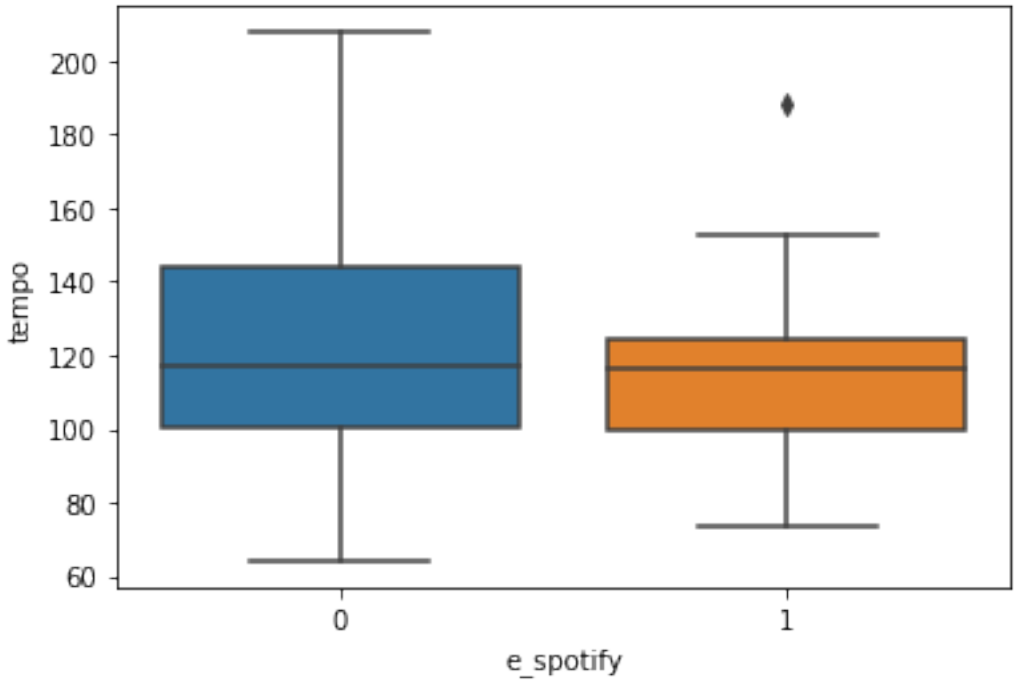
Boxplot grouped by spotify success



Global boxplot by tempo



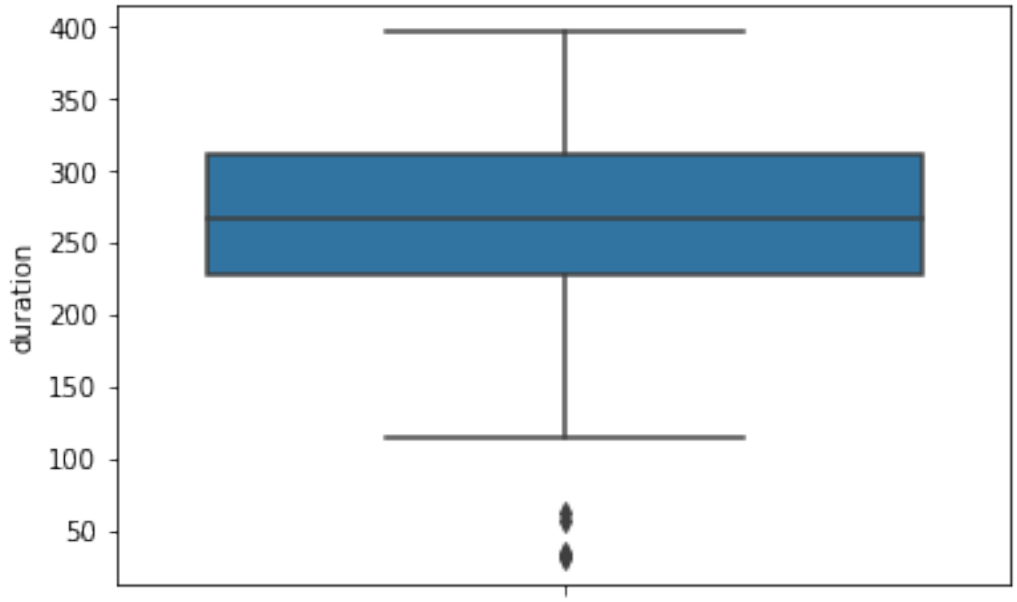
Boxplot grouped by spotify success



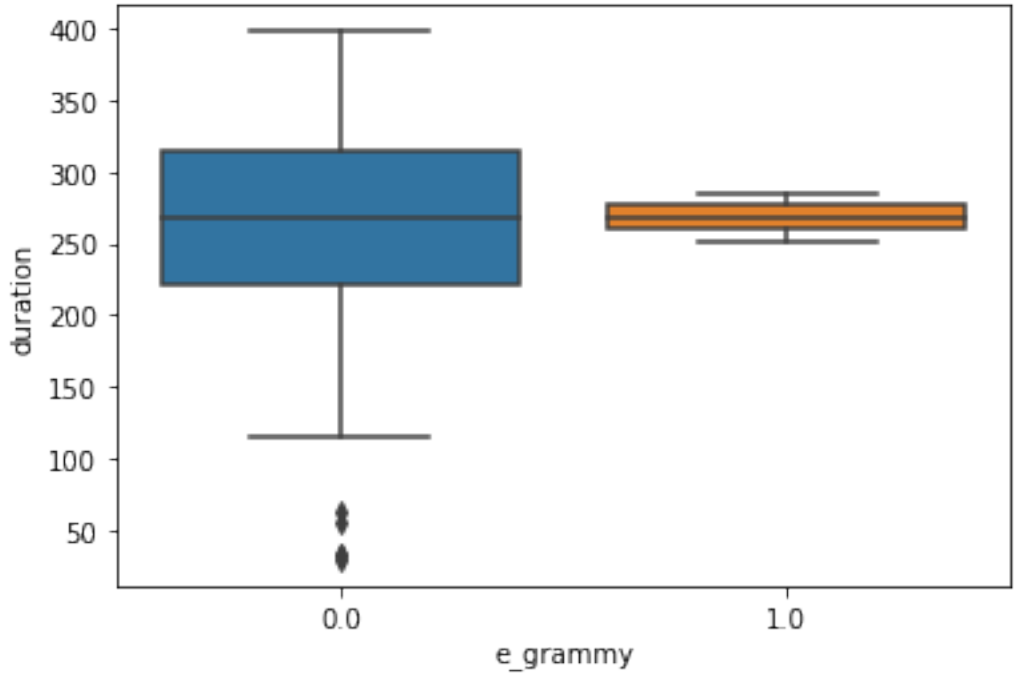
Eminem

```
eminem_df = df[df['artist_name'] == 'Eminem']
drawAllBoxPlots(eminem_df, ['e_grammy'])
drawAllBoxPlots(eminem_df, ['e_billboard'])

drawAllBoxPlots(eminem_df, ['e_spotify'])
```

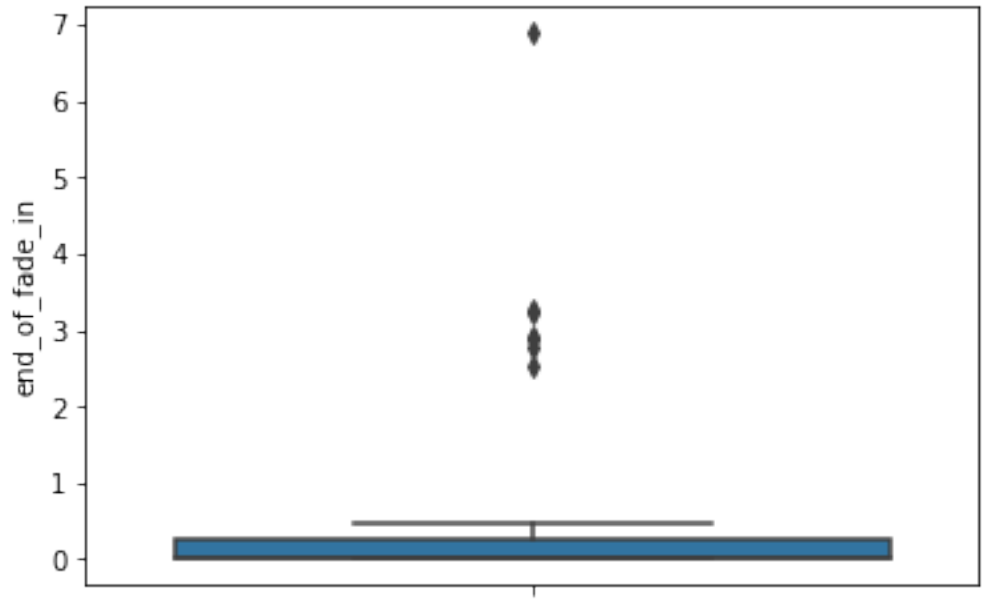
```
Global boxplot by duration
```



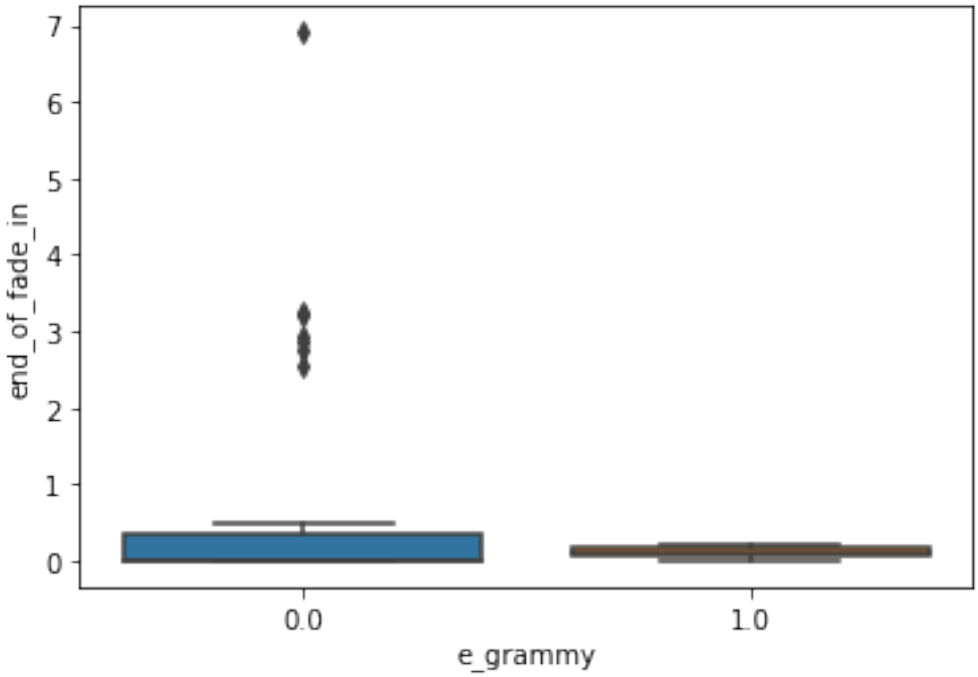
Boxplot grouped by grammy success



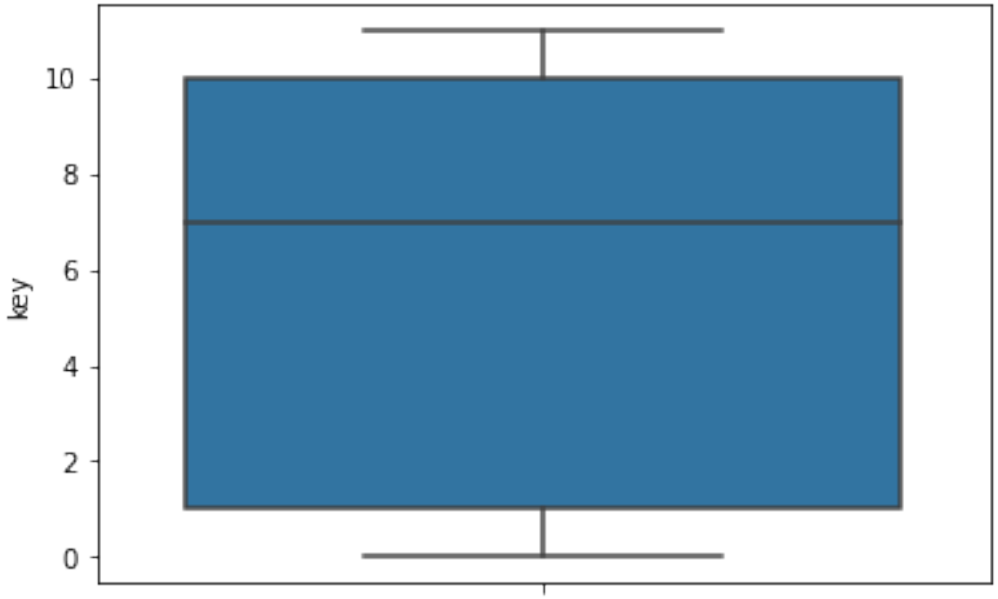
Global boxplot by end_of_fade_in



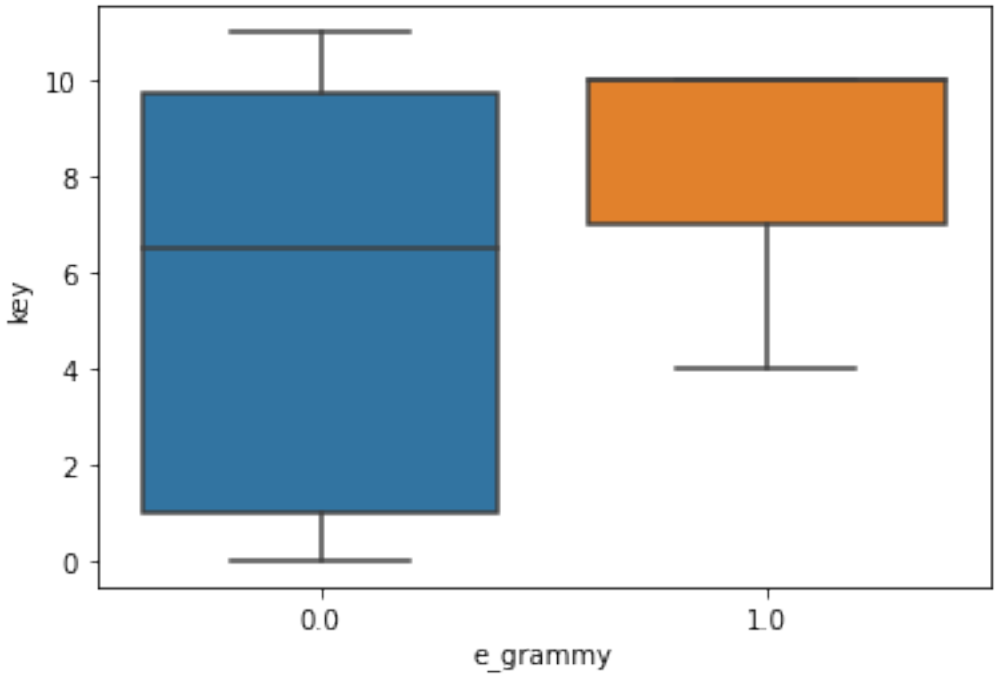
Boxplot grouped by grammy success



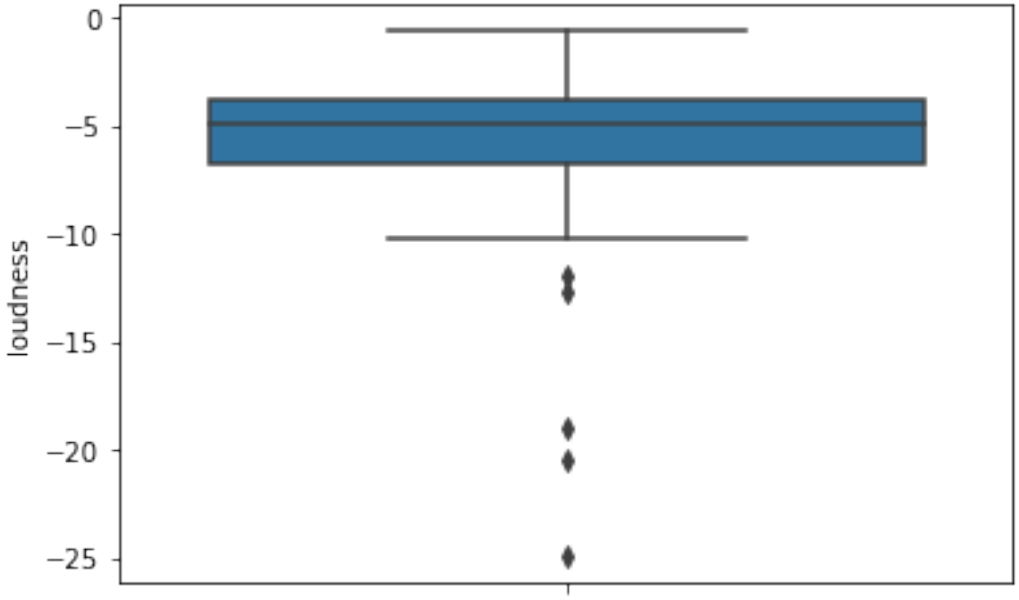
Global boxplot by key



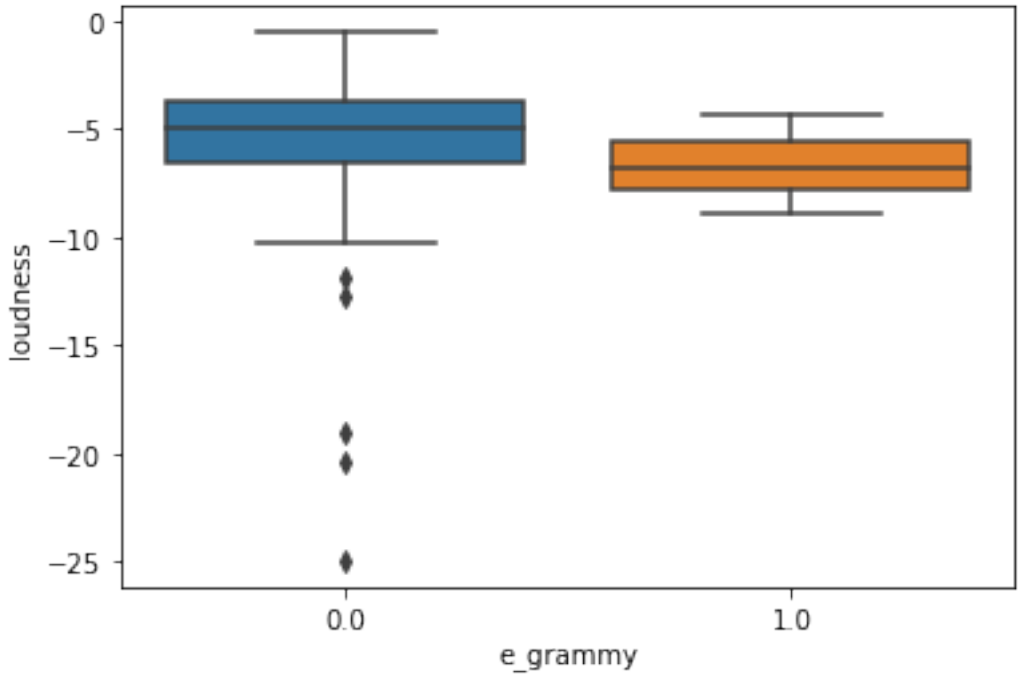
Boxplot grouped by grammy success



Global boxplot by loudness

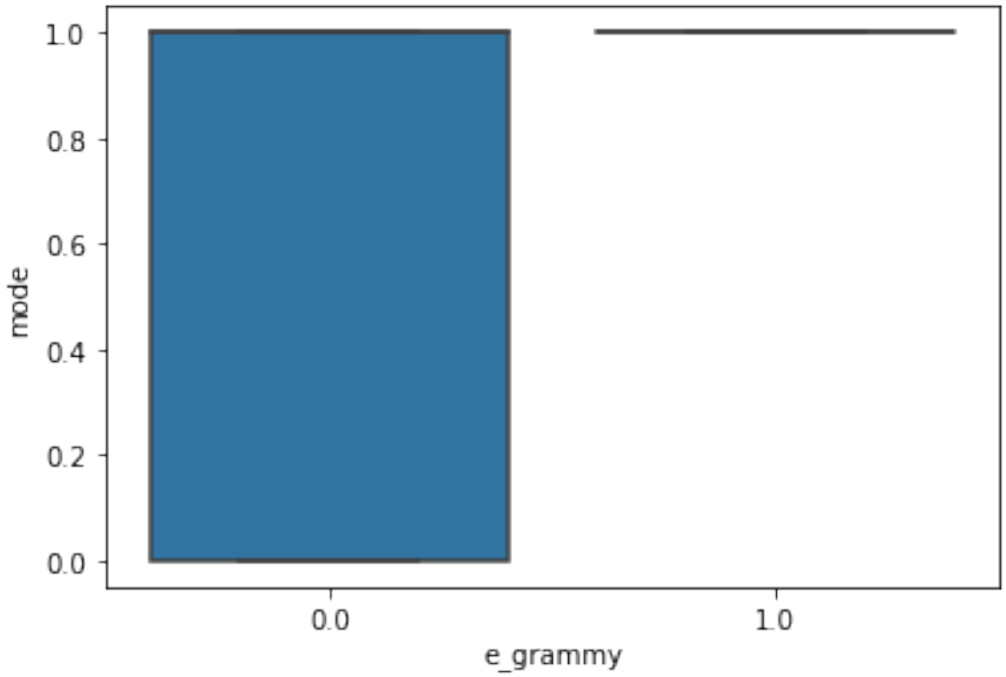


Boxplot grouped by grammy success

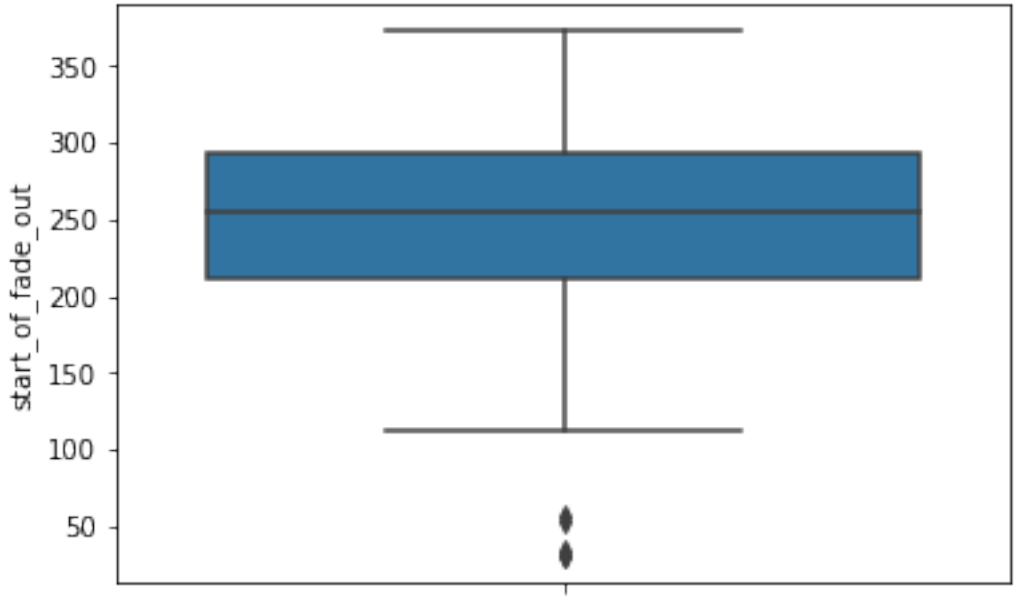


Global boxplot by mode

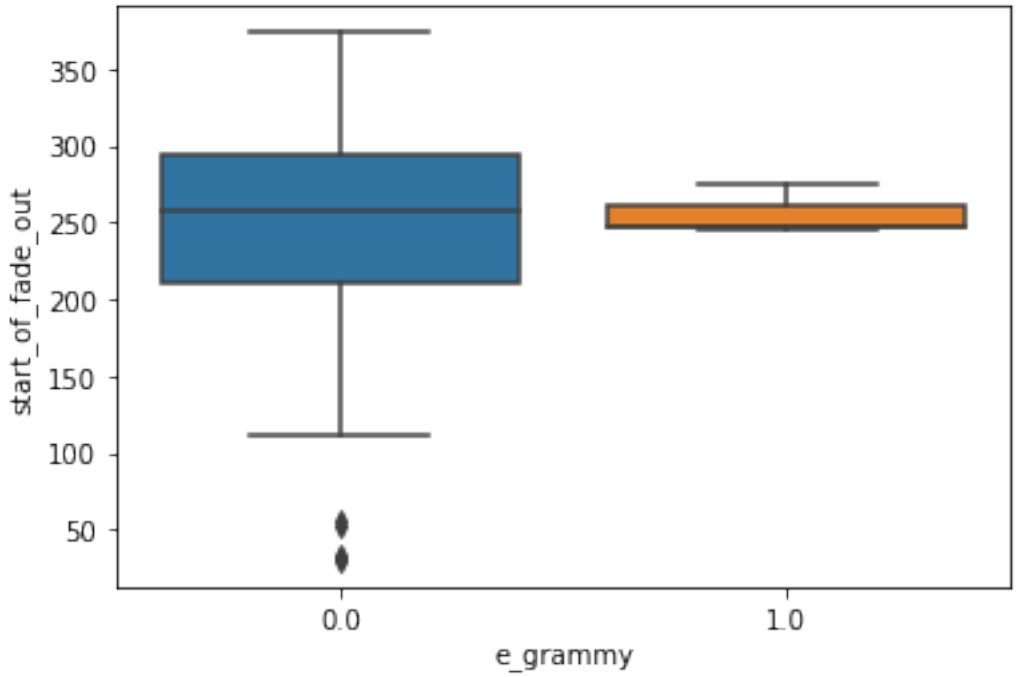
Boxplot grouped by grammy success



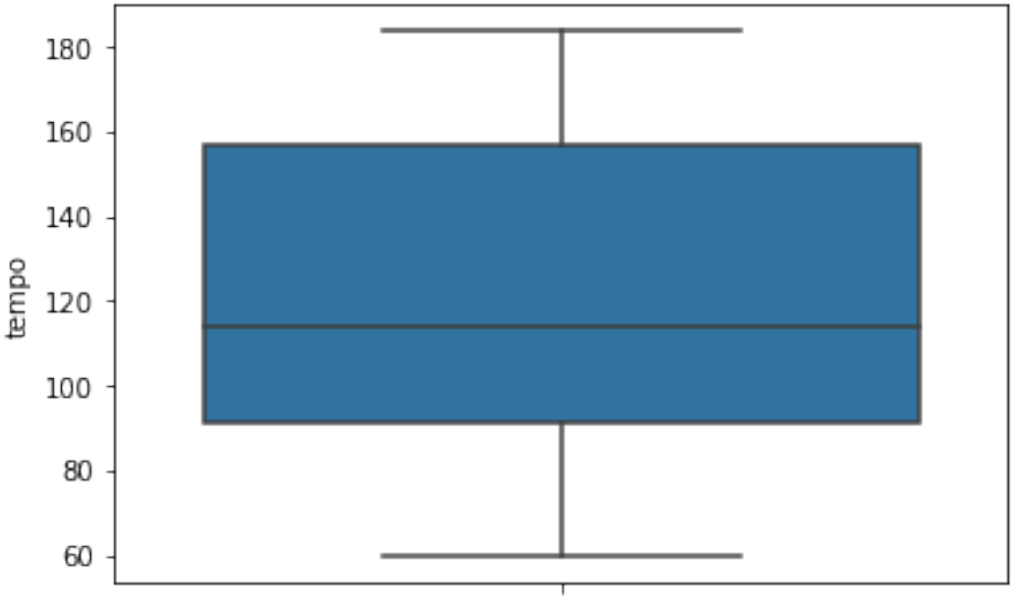
Global boxplot by start_of_fade_out



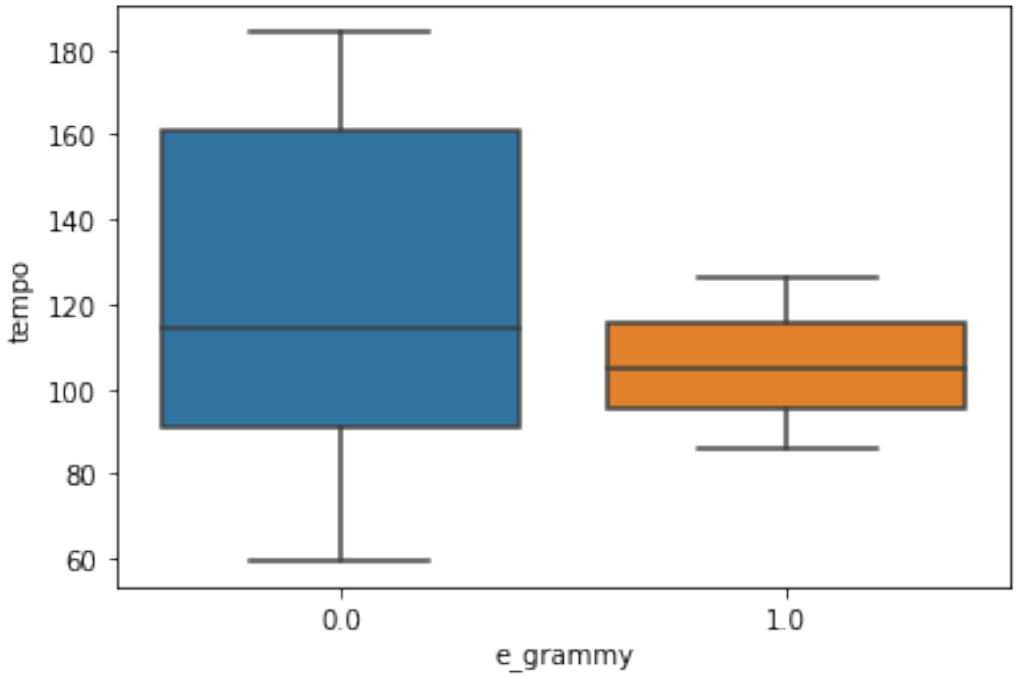
Boxplot grouped by grammy success



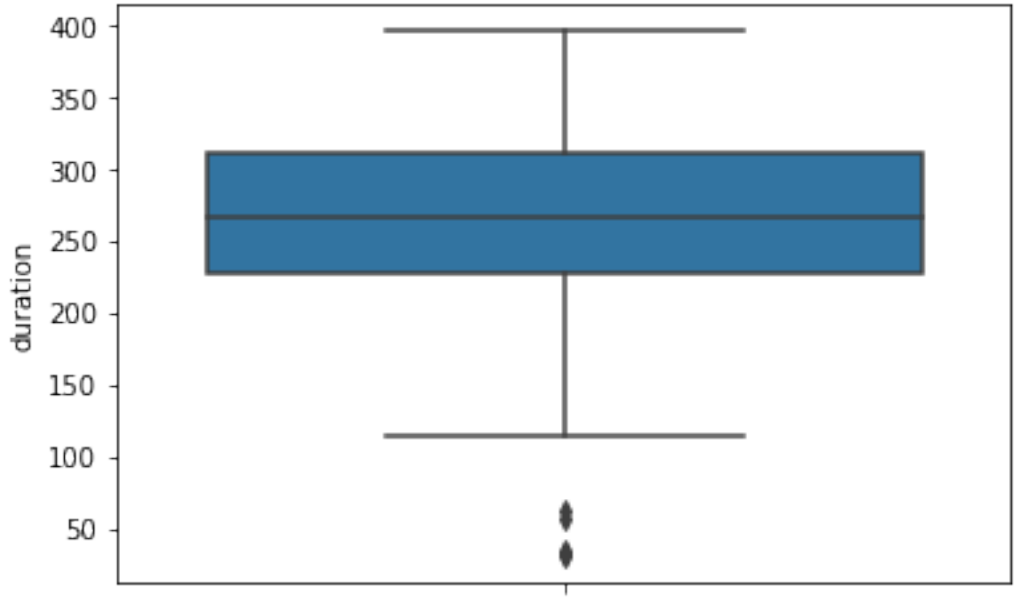
Global boxplot by tempo



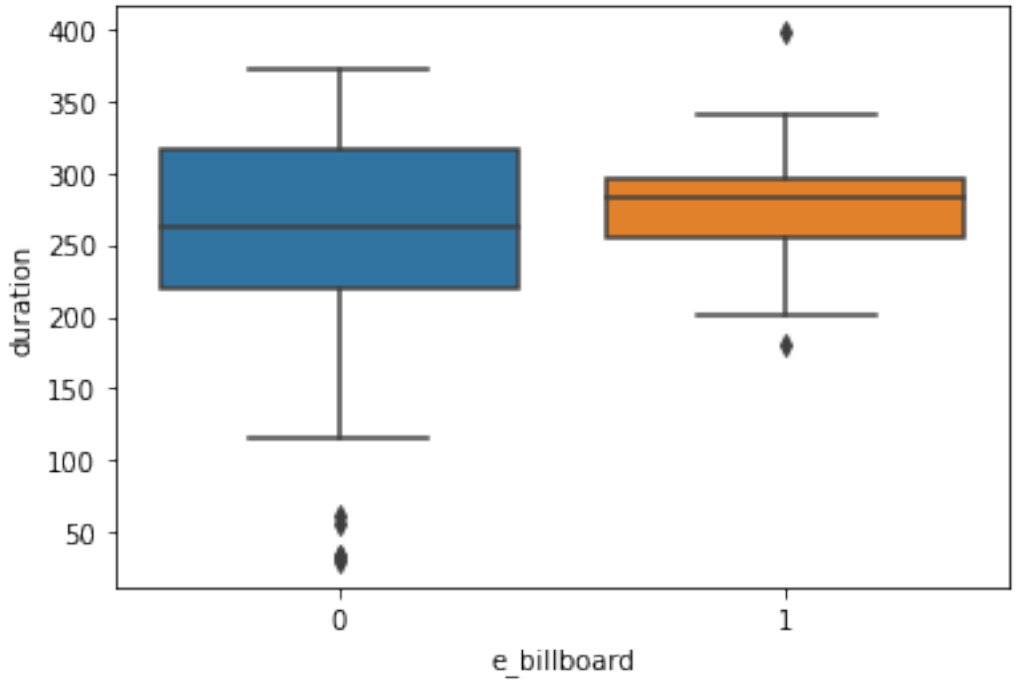
Boxplot grouped by grammy success



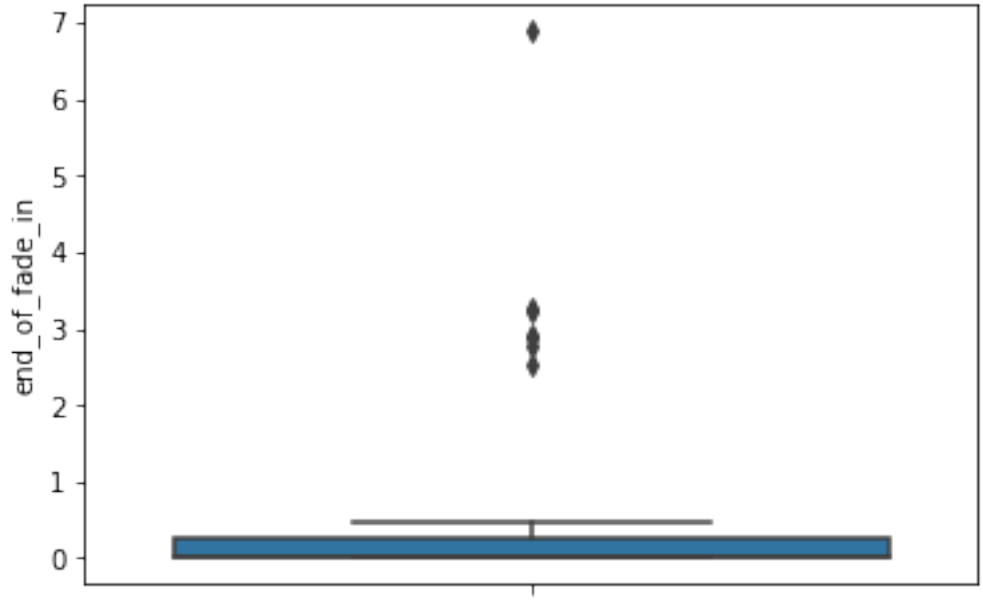
Global boxplot by duration



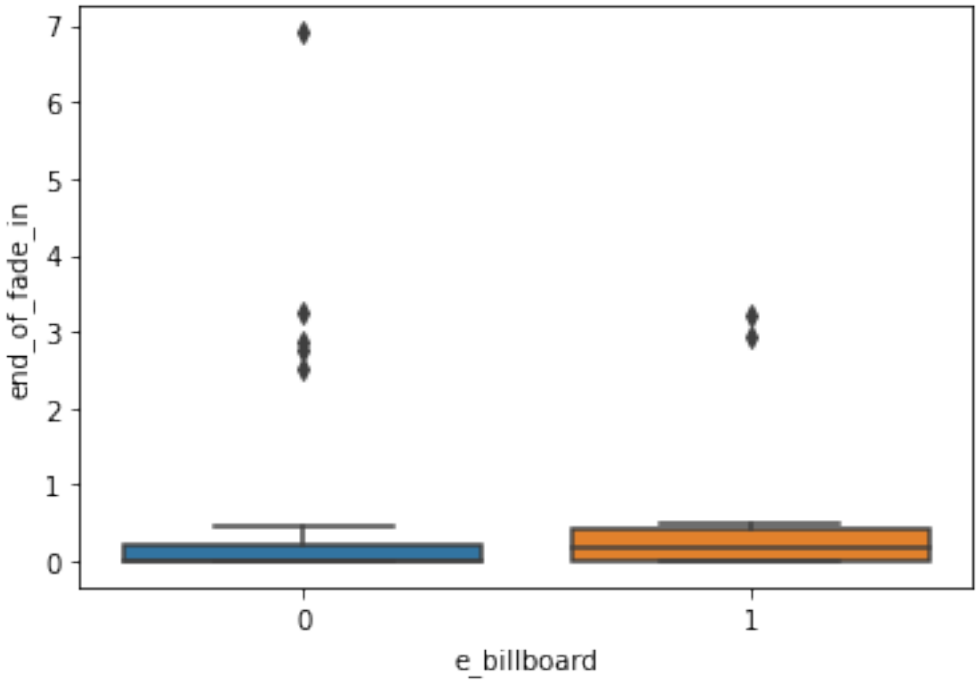
Boxplot grouped by billboard success



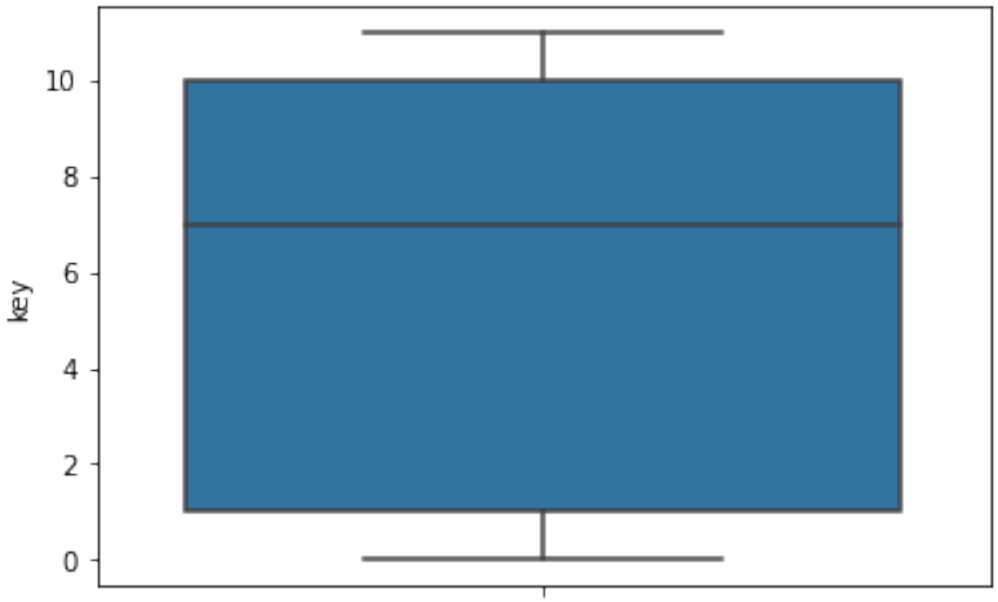
Global boxplot by end_of_fade_in



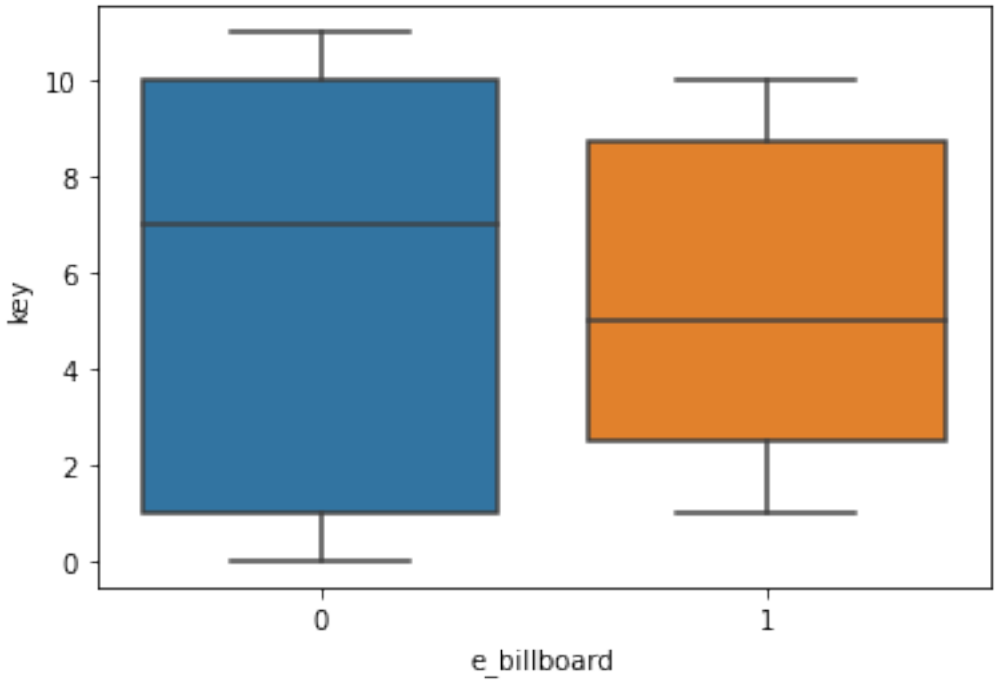
Boxplot grouped by billboard success



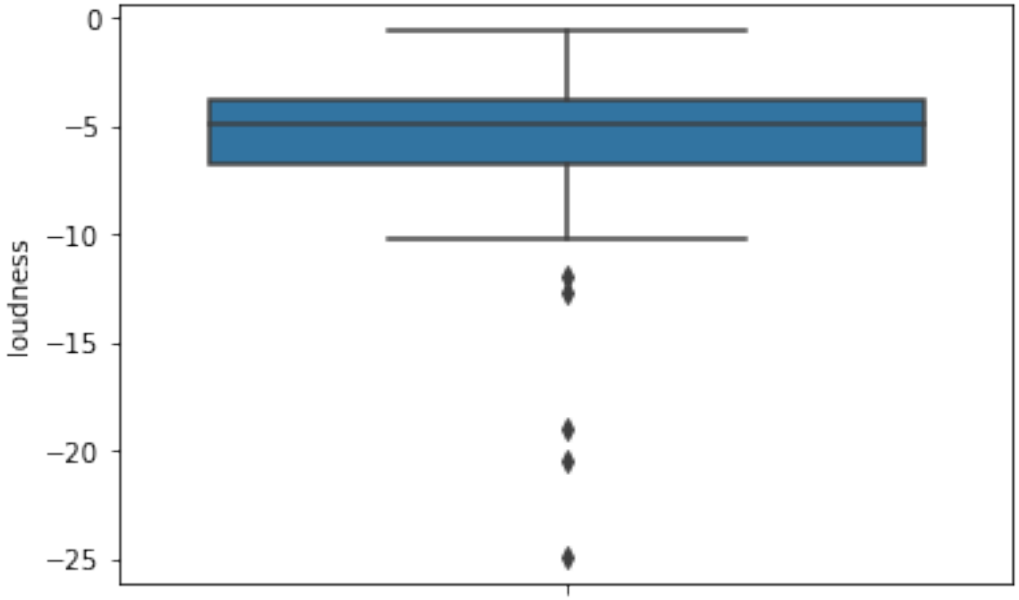
Global boxplot by key



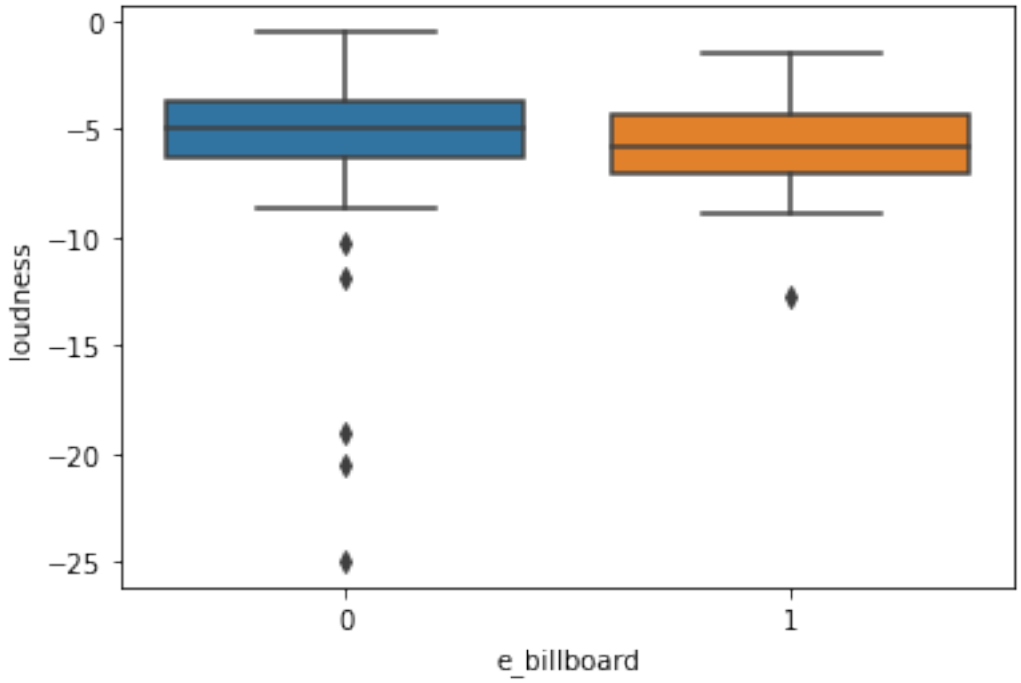
Boxplot grouped by billboard success



Global boxplot by loudness

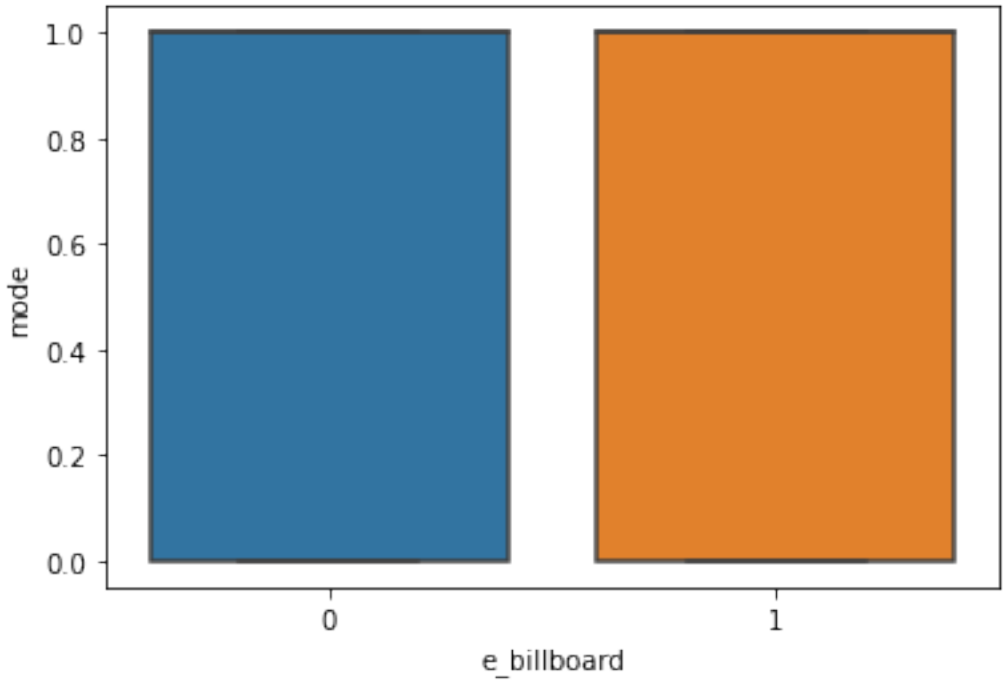


Boxplot grouped by billboard success

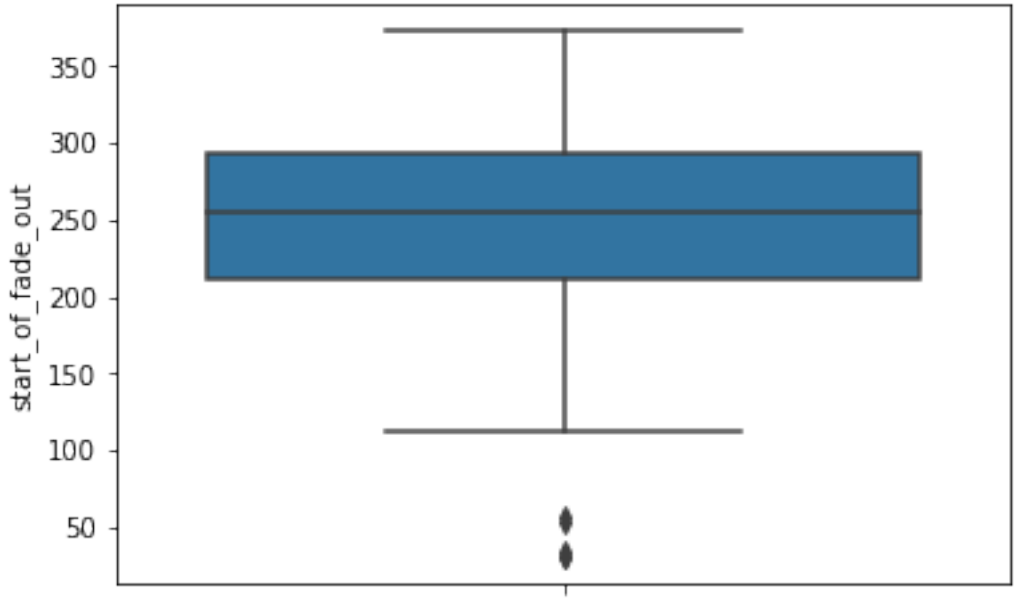


Global boxplot by mode

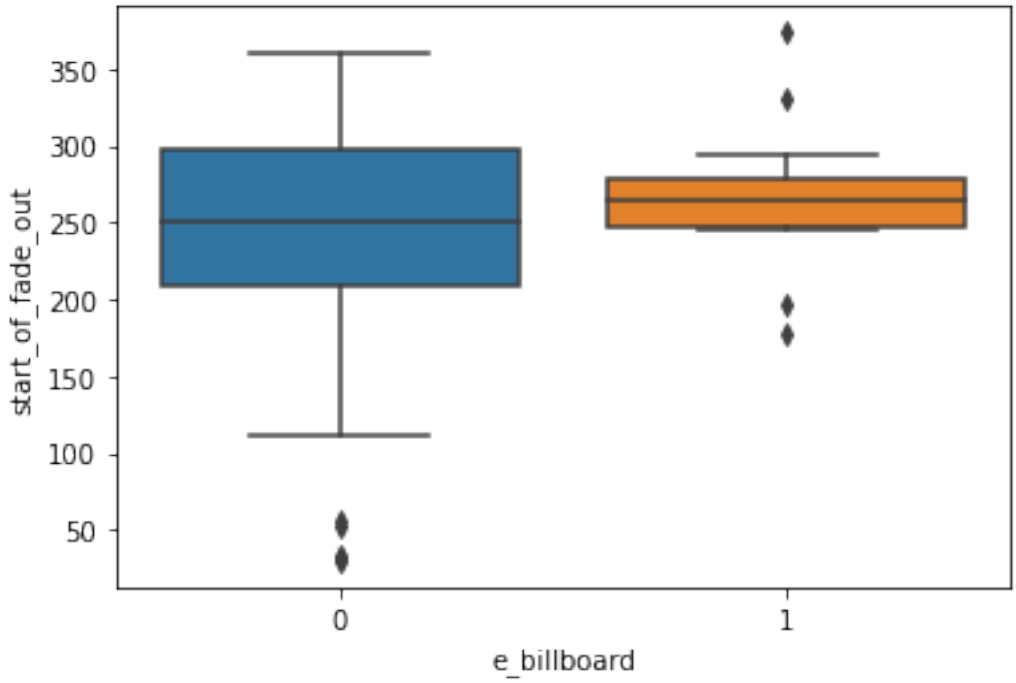
Boxplot grouped by billboard success



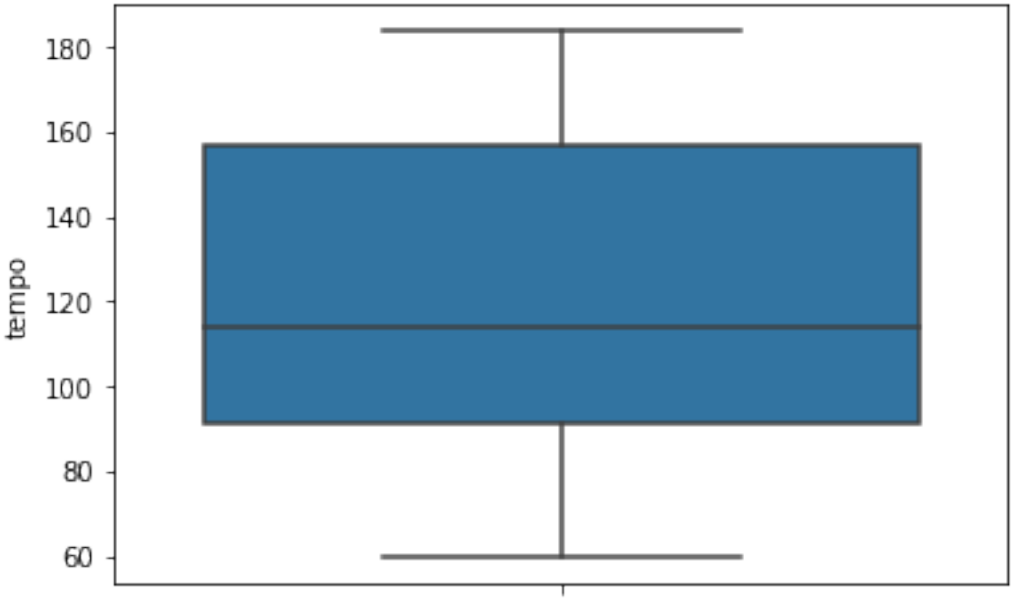
Global boxplot by start_of_fade_out



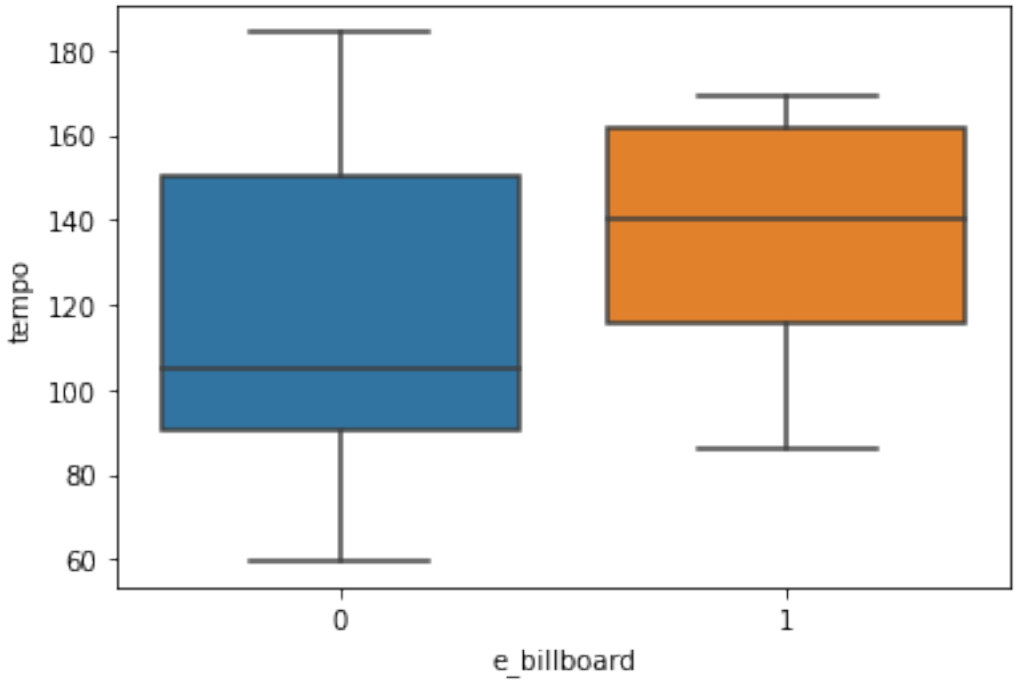
Boxplot grouped by billboard success



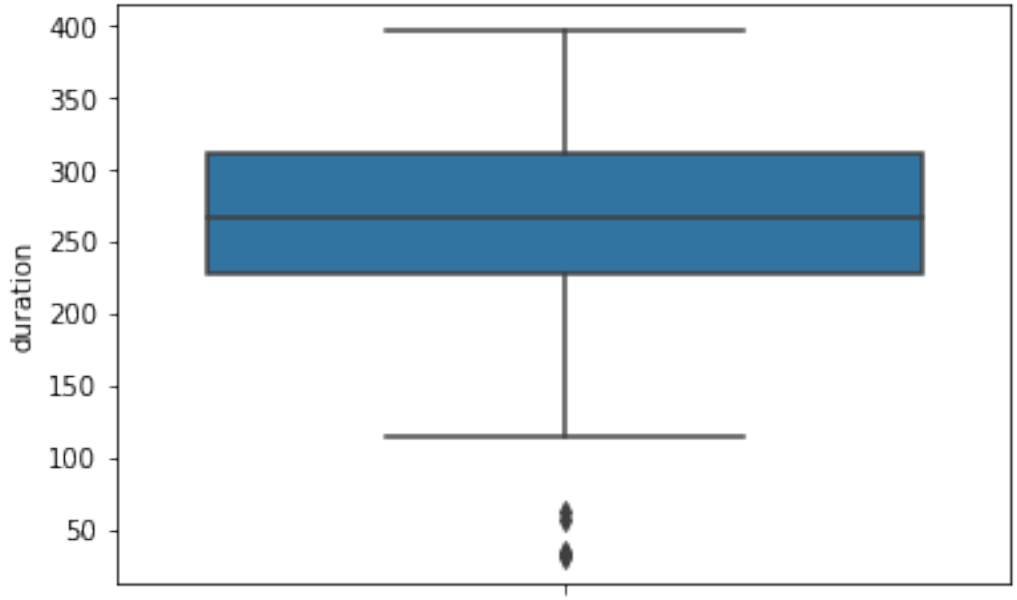
Global boxplot by tempo



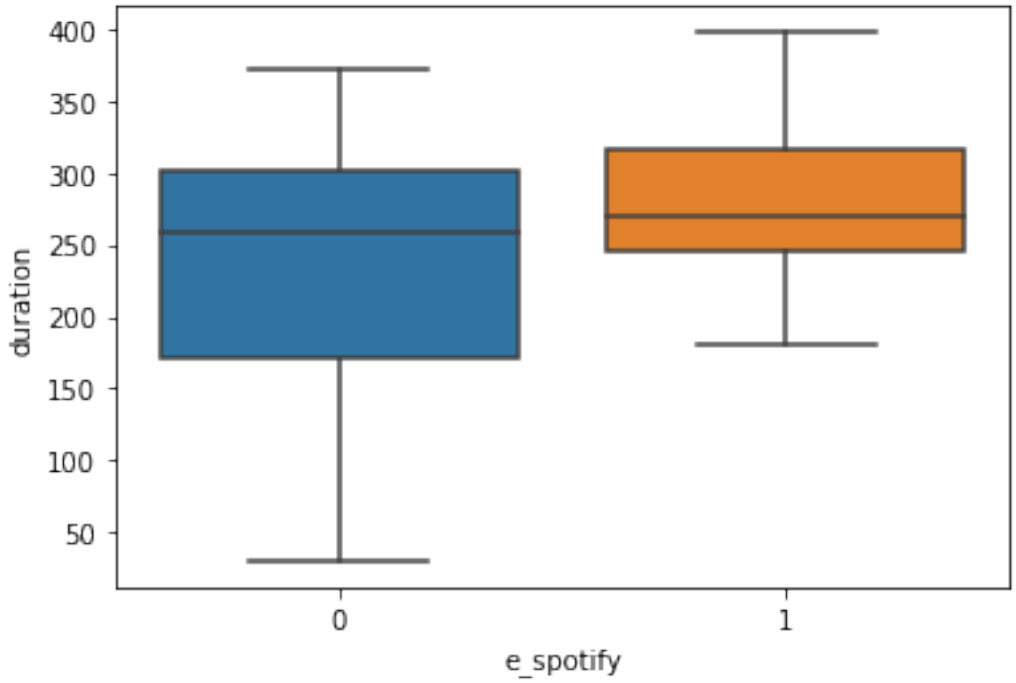
Boxplot grouped by billboard success



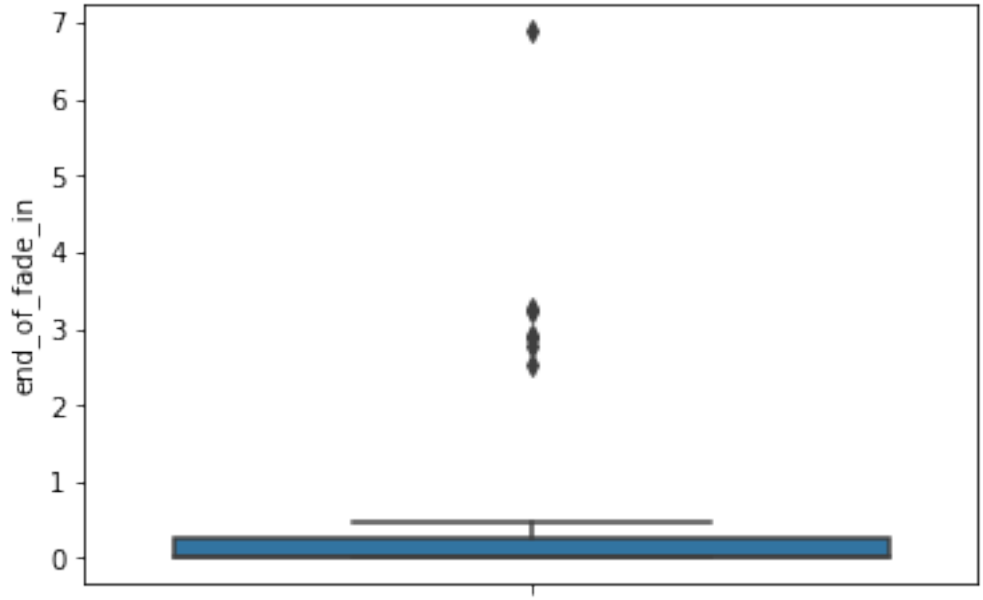
Global boxplot by duration



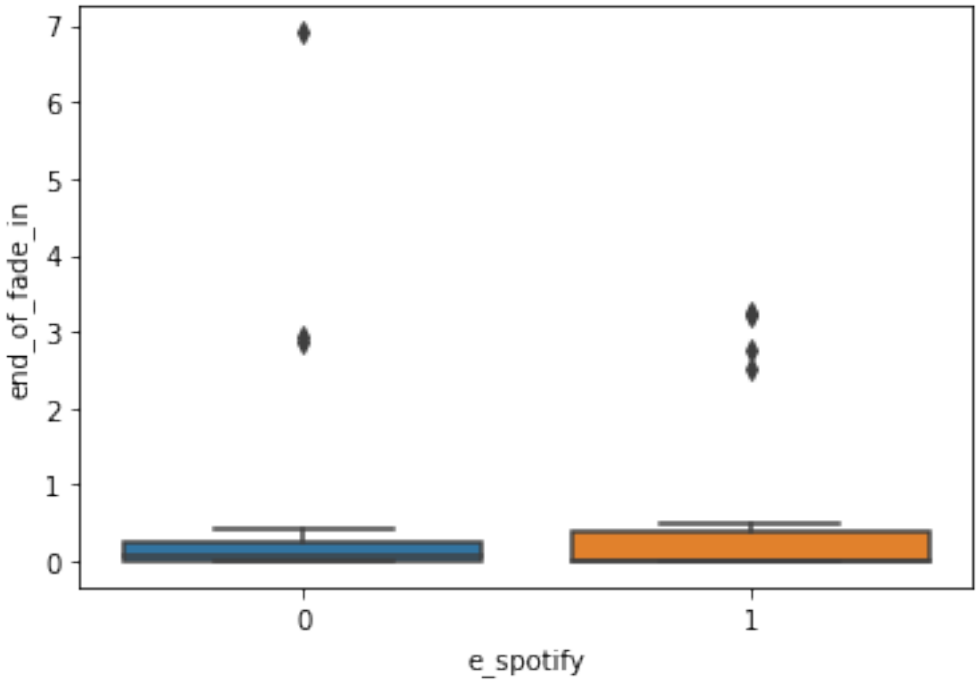
Boxplot grouped by spotify success



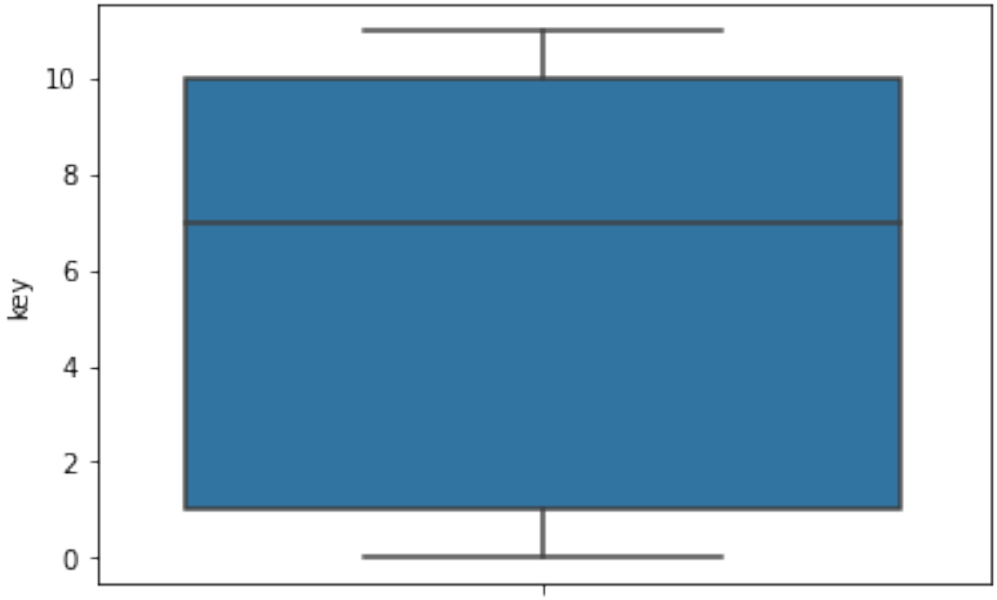
Global boxplot by end_of_fade_in



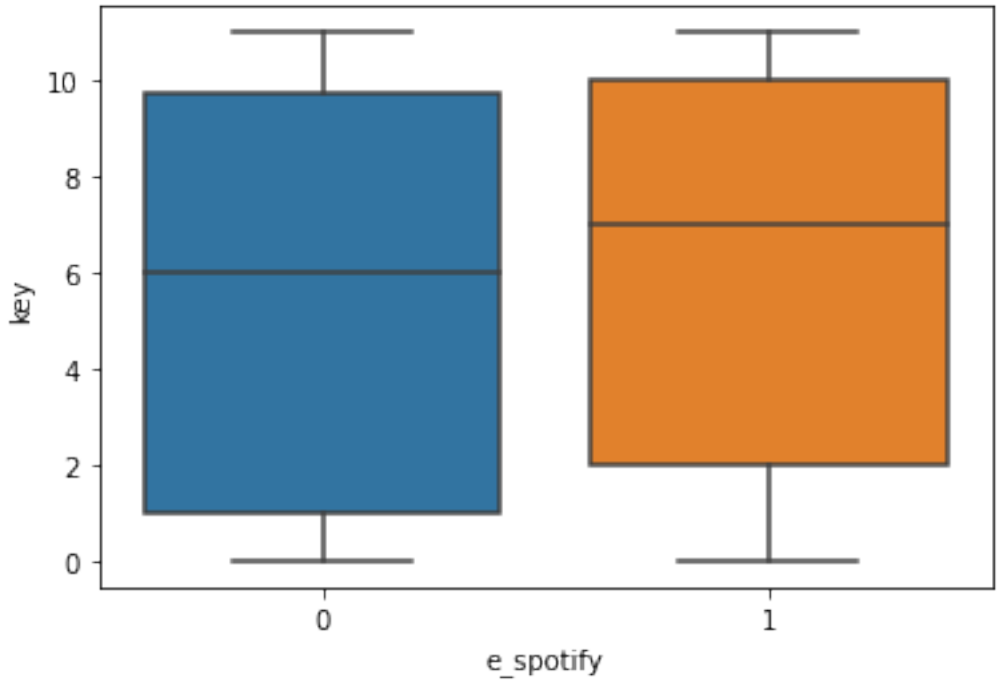
Boxplot grouped by spotify success



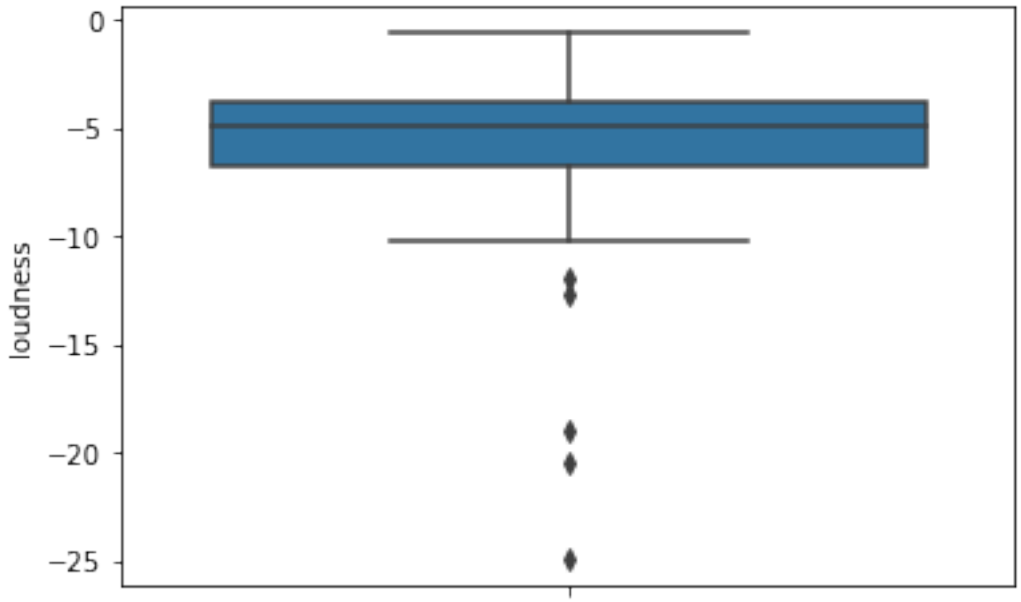
Global boxplot by key



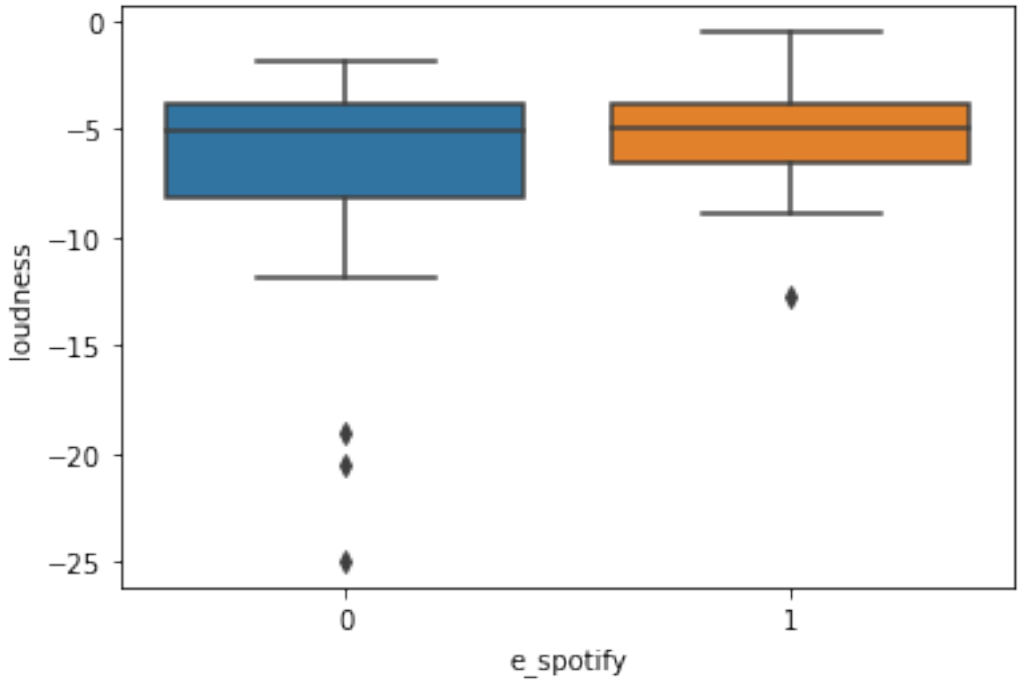
Boxplot grouped by spotify success



Global boxplot by loudness

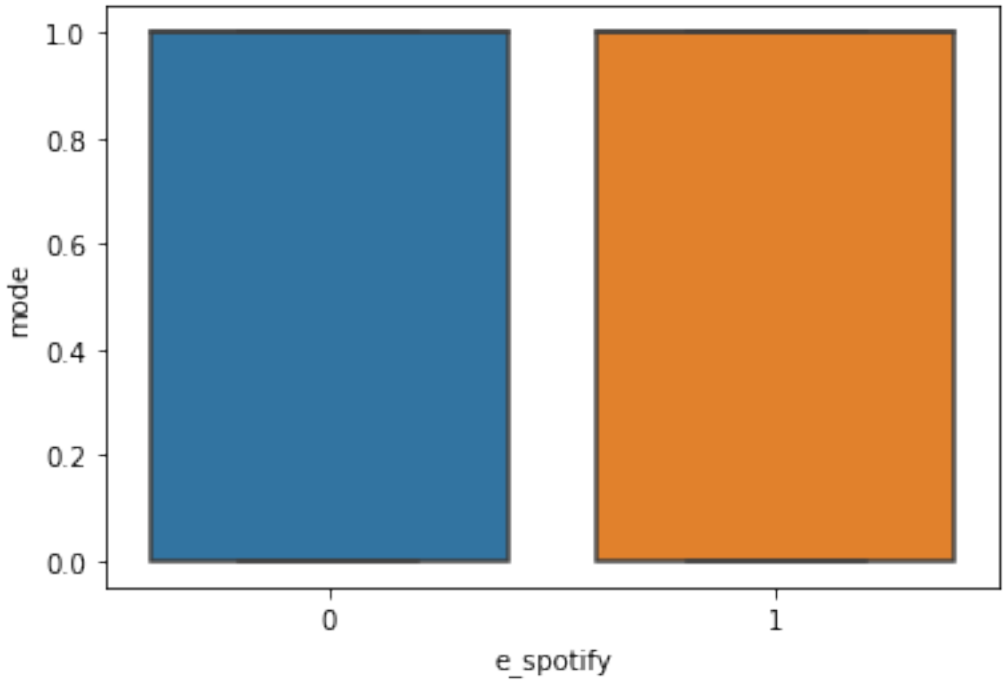


Boxplot grouped by spotify success

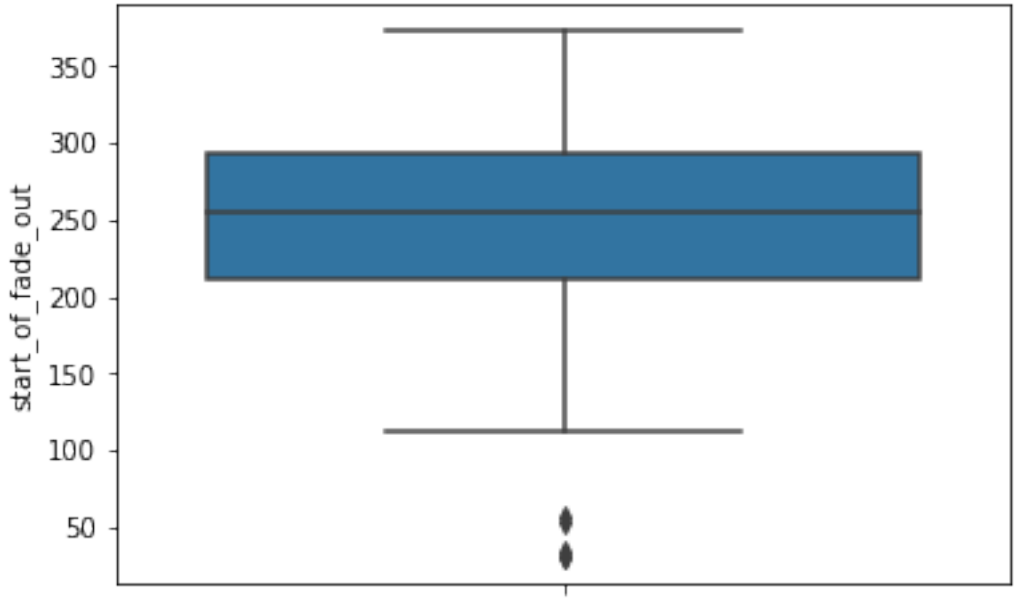


Global boxplot by mode

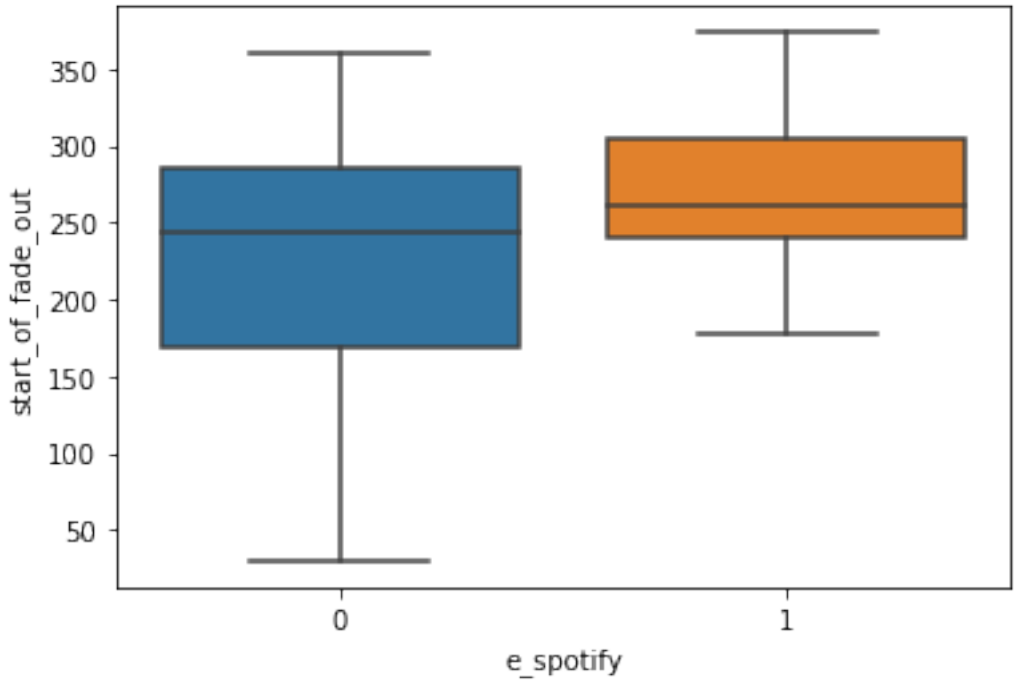
Boxplot grouped by spotify success



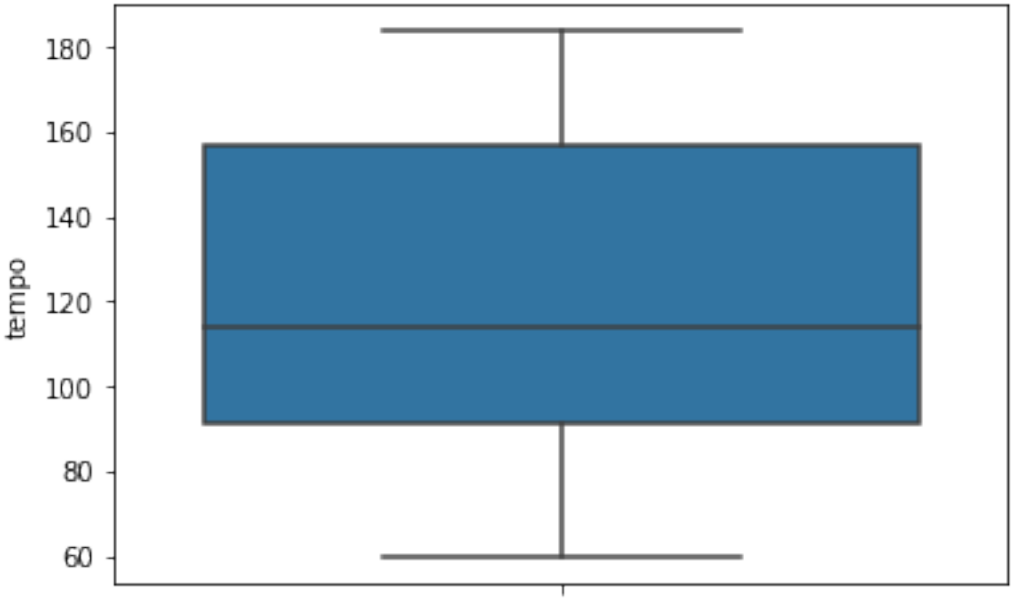
Global boxplot by start_of_fade_out



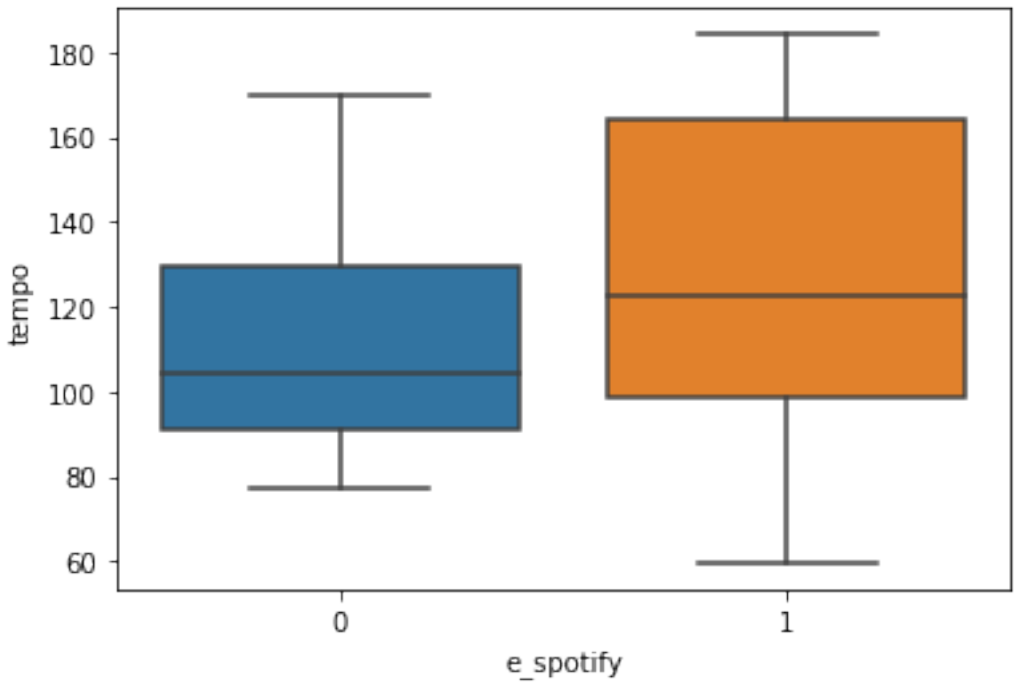
Boxplot grouped by spotify success



Global boxplot by tempo



Boxplot grouped by spotify success


180
160
140
120
100
80
60
tempo
0
1
e_spotify