

Saída: Leitura e Info do Dataset

SEÇÃO 1: Importando bibliotecas e lendo o CSV
Carregando arquivo CSV...

===== Primeiras linhas do dataset =====

	img	title	console	genre	publisher	...	jp_sales	pal_sales	other_sales	release_date	last_update
0	/games/boxart/full_6510540AmericaFrontccc.jpg	Grand Theft Auto V	PS3	Action	Rockstar Games	...	0.99	9.85	3.12	2013-09-17	NaN
1	/games/boxart/full_5563178AmericaFrontccc.jpg	Grand Theft Auto V	PS4	Action	Rockstar Games	...	0.60	9.71	3.02	2014-11-18	2018-01-03
2	/games/boxart/827563ccc.jpg	Grand Theft Auto: Vice City	PS2	Action	Rockstar Games	...	0.47	5.49	1.78	2002-10-28	NaN
3	/games/boxart/full_9218923AmericaFrontccc.jpg	Grand Theft Auto V	X360	Action	Rockstar Games	...	0.06	5.33	1.42	2013-09-17	NaN
4	/games/boxart/full_4990510AmericaFrontccc.jpg	Call of Duty: Black Ops 3	PS4	Shooter	Activision	...	0.41	6.05	2.44	2015-11-06	2018-01-14

[5 rows x 14 columns]

===== Informações gerais =====

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 64016 entries, 0 to 64015
Data columns (total 14 columns):
#   Column          Non-Null Count  Dtype
---  -
0   img              64016 non-null object
1   title            64016 non-null object
2   console          64016 non-null object
3   genre            64016 non-null object
4   publisher        64016 non-null object
5   developer        63999 non-null object
6   critic_score     6678 non-null  float64
7   total_sales      18922 non-null float64
8   na_sales         12637 non-null float64
9   jp_sales         6726 non-null  float64
10  pal_sales        12824 non-null float64
11  other_sales      15128 non-null float64
12  release_date     56965 non-null object
13  last_update      17879 non-null object
dtypes: float64(6), object(8)
memory usage: 6.8+ MB
None
```

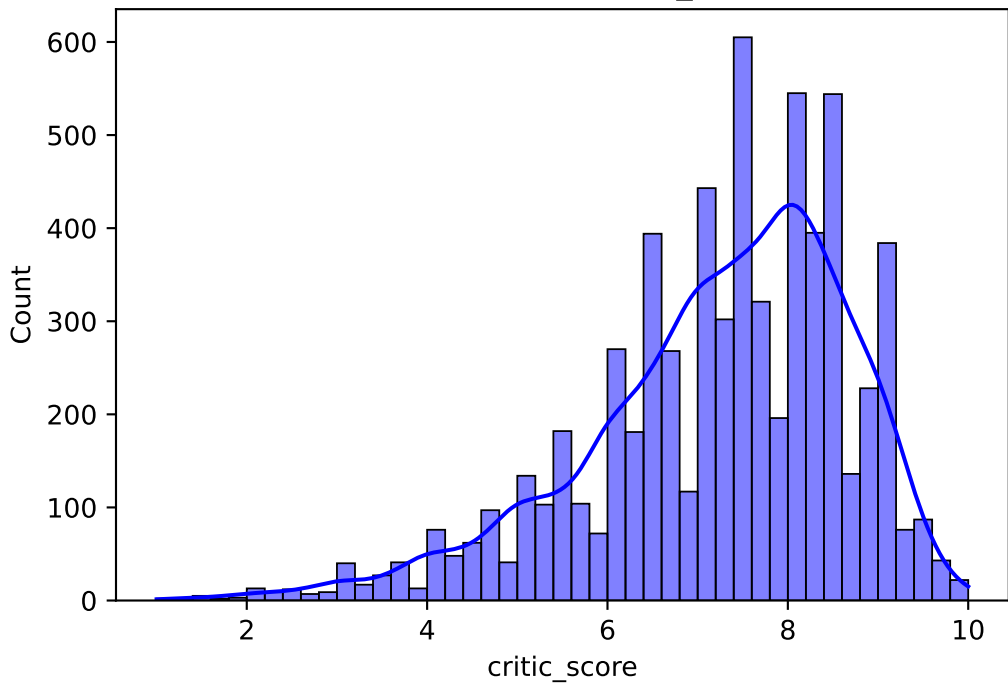
===== Quantidade de linhas e colunas =====
(64016, 14)

===== Verificando valores nulos =====

```
img              0
title            0
console          0
genre            0
publisher        0
developer        17
critic_score     57338
total_sales      45094
na_sales         51379
jp_sales         57290
pal_sales        51192
other_sales      48888
release_date     7051
last_update      46137
dtype: int64
```

Gerando histograma de 'critic_score'...

Distribuição de critic_score

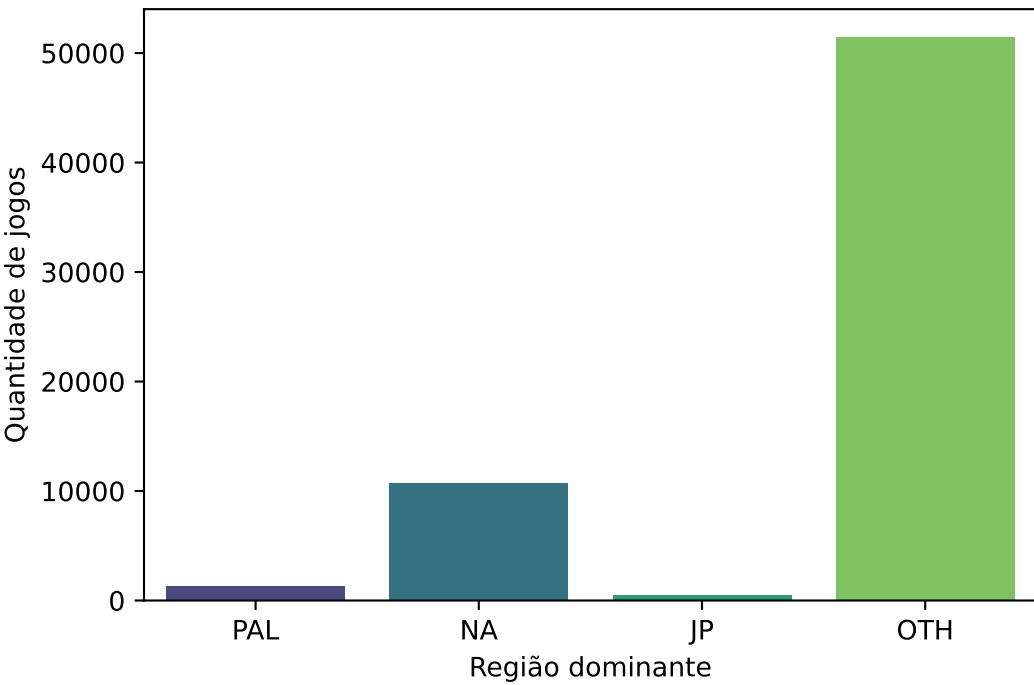


Saída: Region Preference

SEÇÃO 3: Criando a coluna 'region_preference'

```
===== Distribuição da classe (region_preference) =====  
region_preference  
OTH      51437  
NA       10755  
PAL       1329  
JP         495  
Name: count, dtype: int64
```

Contagem de jogos por preferência regional



SEÇÃO 4: Pré-processamento

===== Formato dos dados =====

X_train: (847, 9), y_train: (847,)

X_test : (363, 9), y_test : (363,)

Classes no treino:

region_preference

NA 550

PAL 182

JP 111

OTH 4

Name: count, dtype: int64

Classes no teste:

region_preference

NA 219

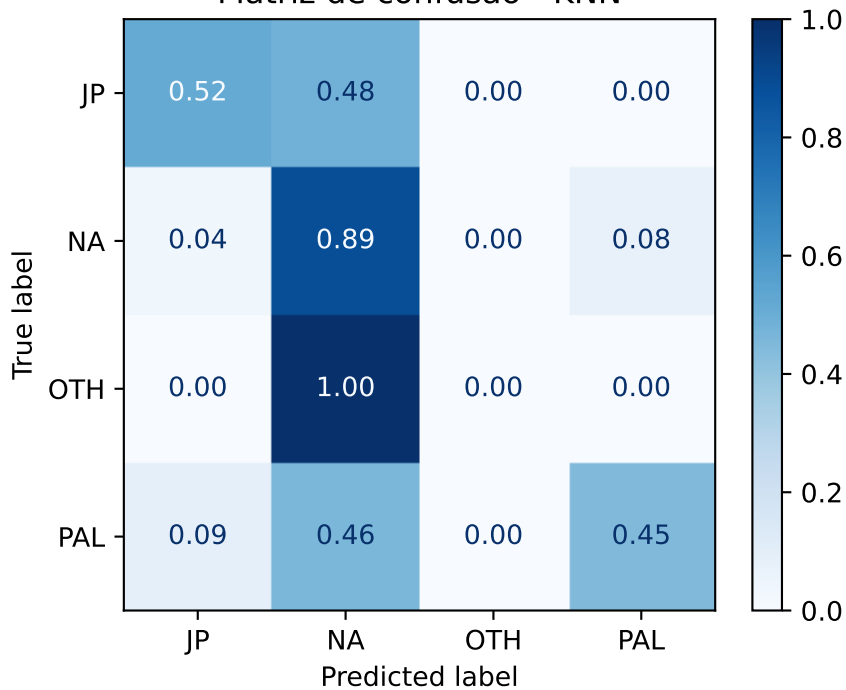
PAL 87

JP 56

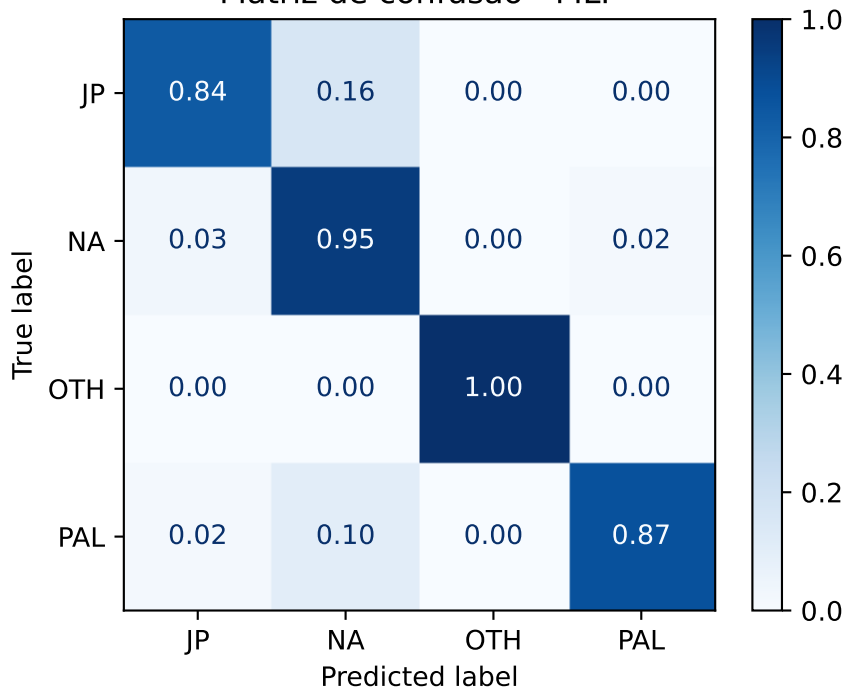
OTH 1

Name: count, dtype: int64

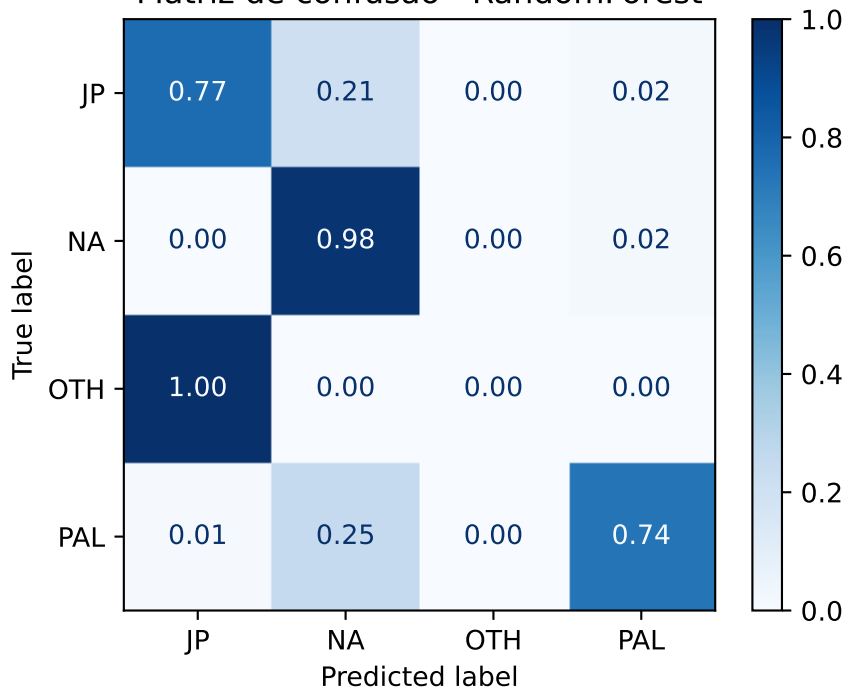
Matriz de confusão - KNN



Matriz de confusão - MLP



Matriz de confusão - RandomForest



Saída: Resultados Finais

===== Comparativo final =====

	Algoritmo	Acuracia	Precisao	Recall
0	KNN	0.721763	0.713054	0.721763
1	MLP	0.914601	0.915172	0.914601
2	RandomForest	0.884298	0.887134	0.884298

=== KNN ===

Acurácia : 0.7218
Precisão : 0.7131
Recall : 0.7218

=== MLP ===

Acurácia : 0.9146
Precisão : 0.9152
Recall : 0.9146

=== RandomForest ===

Acurácia : 0.8843
Precisão : 0.8871
Recall : 0.8843

Comparação de Acurácia por Algoritmo

