```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns


from google.colab import files
uploaded = files.upload()
```

    Escolher Ficheiros luna_pets.csv
    • **luna_pets.csv**(text/csv) - 132539 bytes, last modified: 25/04/2024 - 100% done
    Saving luna_pets.csv to luna_pets.csv

```python
import io
df = pd.read_csv(io.BytesIO(uploaded['luna_pets.csv']))


print(df.head())
print(df.info())
print(df.describe())
```

```
                        brand         gtin13  \
0                      Fluval    15561173551
1            Kit 'N Kaboodle    17800130431
2  Chicken Soup for the Soul  8192390126506
3                     Top Fin   737257865426
4                     ZuPreem   762177850306

                                     description  product_id availability  \
0            Fluval&reg; Bug Bites Spirulina Flakes     5300528      InStock
1  Purina&reg; Kit &amp; Kaboodle&reg; Original A...     5139253      InStock
2  Chicken Soup for the Soul All Life Stages Dry ...     5329365      InStock
3      Top Fin&reg; Orange Octopus Aquarium Ornament     5295111   OutOfStock
4                      ZuPreem&reg; NutBlend Bird Food     5167132      InStock

  currency  price  avg_rating  reviews_count  \
0      USD   7.99      4.7778              9
1      USD  11.99      3.7430            284
2      USD  17.99         NaN              0
3      USD  19.99      5.0000              1
4      USD  17.99      4.2727             11

                                uniq_id            scraped_at
0  05537726-5061-5b90-843b-2d1f431bbdbd  30/07/2022 12:08:13
1  f0ff4c10-4dc6-54a3-80e0-886bf4c508cf  30/07/2022 12:08:13
2  db7ddaf3-99fc-50ea-9958-980bb09a7327  30/07/2022 12:08:13
3  f721f3fd-080b-5147-a215-b8808b9731ca  30/07/2022 12:08:13
4  50978e5c-3ea8-51dd-bb8c-854afbe5f191  30/07/2022 12:08:15
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 371 entries, 0 to 370
Data columns (total 14 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   url            371 non-null    object
 1   name           371 non-null    object
 2   sku            371 non-null    int64
 3   brand          367 non-null    object
 4   gtin13         371 non-null    int64
 5   description    371 non-null    object
 6   product_id     371 non-null    int64
 7   availability   371 non-null    object
 8   currency       371 non-null    object
```
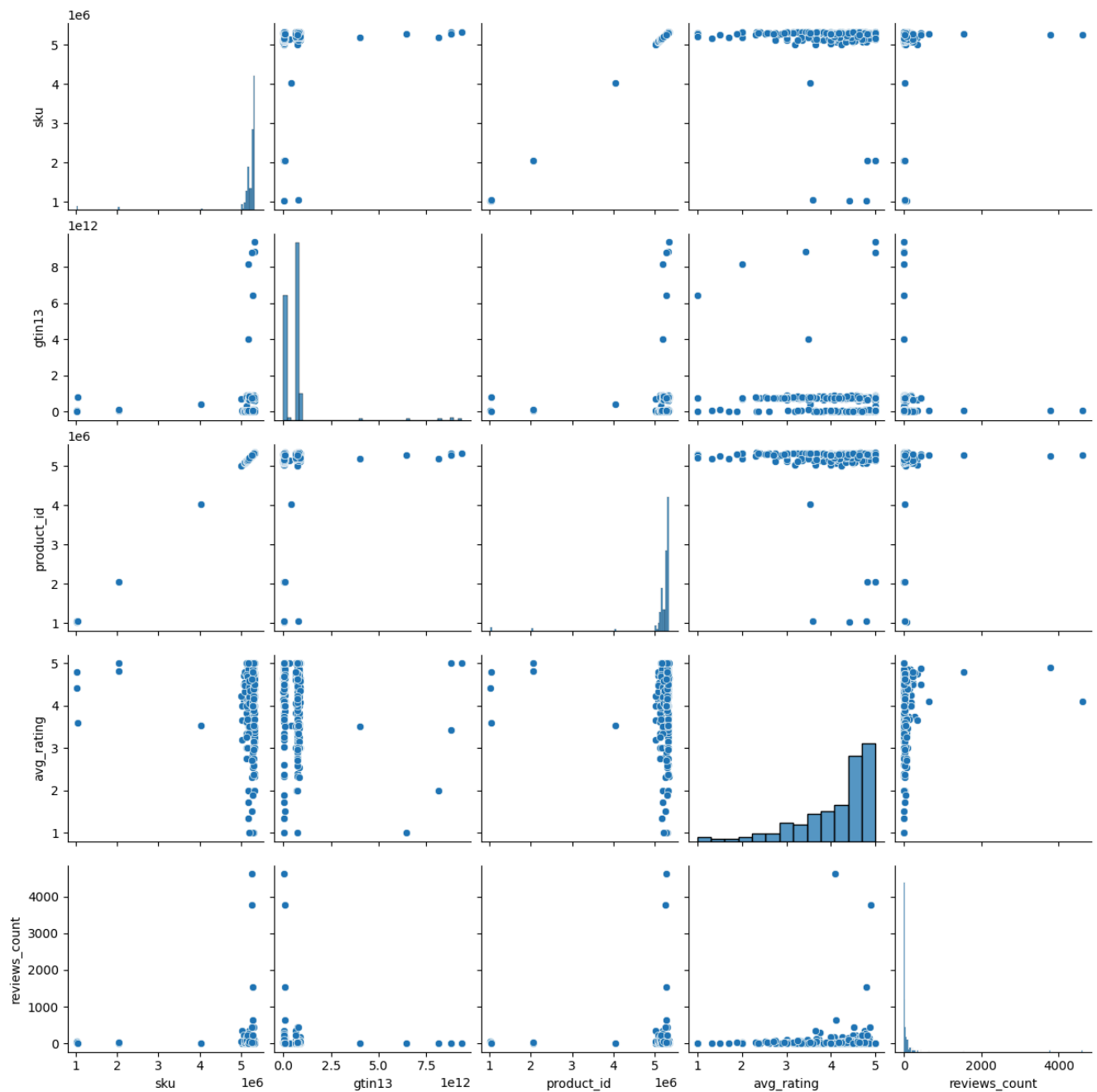
50%	5.28/5//e+06	7.3/25/9e+11	5.28/5//e+06	4.426/00	7.000000
75%	5.311680e+06	8.154585e+11	5.311680e+06	4.727300	32.000000
max	5.335307e+06	9.353482e+12	5.335307e+06	5.000000	4617.000000

```
df.dropna(inplace=True)
```

```
sns.pairplot(df)
plt.show()
```



```
df_encoded = pd.get_dummies(df, columns=['product_id'])
```

```
df_encoded.to_csv('dados_processados.csv', index=False)
files.download('dados_processados.csv')
```