

A financial candlestick chart with a blue grid background. A white curved line, likely a Fibonacci retracement, is drawn across the chart. A large purple asterisk is positioned over the chart. The chart includes several numerical labels: '104.19' in a box, '86.72' in a box, '72.48' in a box, and '61.6%' near the top of the curve.

Online shoppers clustering EDA

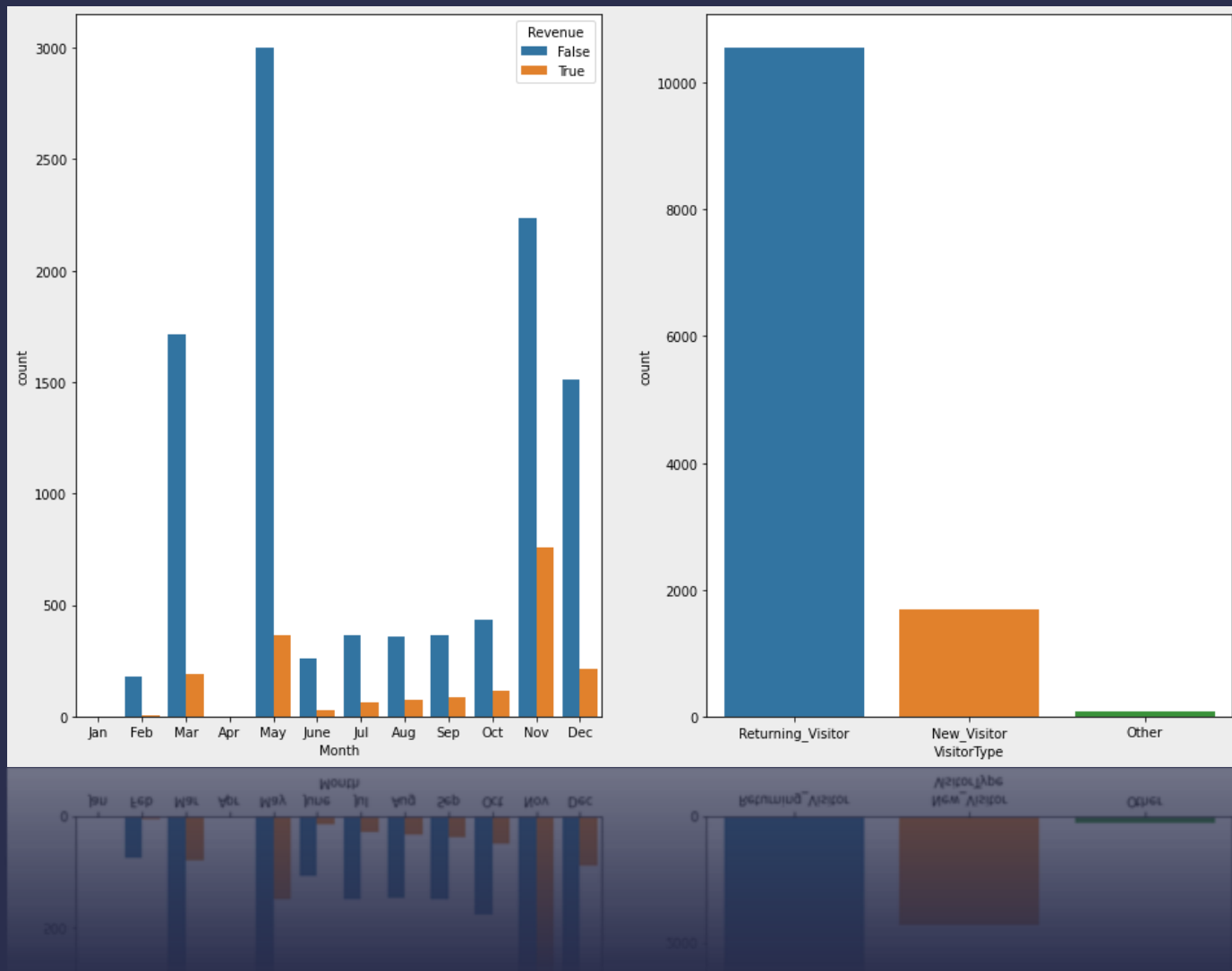
Hubert Ruczyński

Mateusz Sperkowski

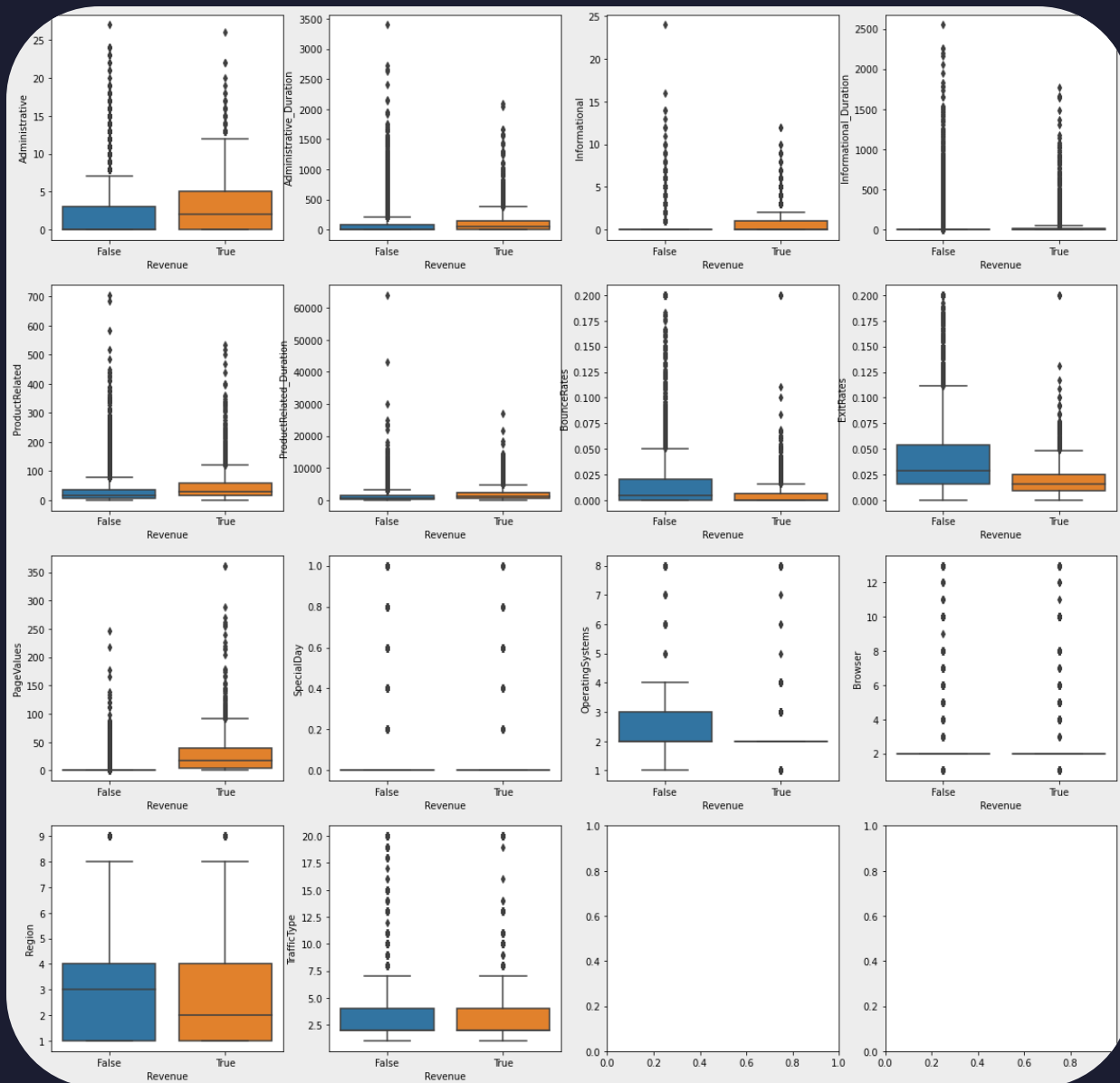
Prosty opis danych

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 12330 entries, 0 to 12329  
Data columns (total 18 columns):  
#   Column                                Non-Null Count  Dtype  ---  ---  
0   Administrative                        12330 non-null  int64  
1   Administrative_Duration               12330 non-null  float64  
2   Informational                        12330 non-null  int64  
3   Informational_Duration               12330 non-null  float64  
4   ProductRelated                      12330 non-null  int64  
5   ProductRelated_Duration             12330 non-null  float64  
6   BounceRates                         12330 non-null  float64  
7   ExitRates                           12330 non-null  float64  
8   PageValues                          12330 non-null  float64  
9   SpecialDay                          12330 non-null  float64  
10  Month                               12330 non-null  object  
11  OperatingSystems                   12330 non-null  int64  
12  Browser                           12330 non-null  int64  
13  Region                            12330 non-null  int64  
14  TrafficType                       12330 non-null  int64  
15  VisitorType                       12330 non-null  object  
16  Weekend                           12330 non-null  bool  
17  Revenue                           12330 non-null  bool  
dtypes: bool(2), float64(7), int64(7), object(2)  
memory usage: 1.5+ MB
```

	Administrative	Administrative_Duration	Informational	Informational_Duration	ProductRelated	ProductRelated_Duration	BounceRates	ExitRates	PageValues	SpecialDay	OperatingSystems	Browser	Region	TrafficType
count	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000	12330.000000
mean	2.315166	80.818611	0.503569	34.472398	31.731468	1194.746220	0.022191	0.043073	5.889258	0.061427	2.124006	2.357097	3.147364	4.069586
std	3.321784	176.779107	1.270156	140.749294	44.475503	1913.669288	0.048488	0.048597	18.568437	0.198917	0.911325	1.717277	2.401591	4.025169
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	1.000000	1.000000	1.000000
25%	0.000000	0.000000	0.000000	0.000000	7.000000	184.137500	0.000000	0.014286	0.000000	0.000000	2.000000	2.000000	1.000000	2.000000
50%	1.000000	7.500000	0.000000	0.000000	18.000000	598.936905	0.003112	0.025156	0.000000	0.000000	2.000000	2.000000	3.000000	2.000000
75%	4.000000	93.256250	0.000000	0.000000	38.000000	1464.157213	0.016813	0.050000	0.000000	0.000000	3.000000	2.000000	4.000000	4.000000
max	27.000000	3398.750000	24.000000	2549.375000	705.000000	63973.522230	0.200000	0.200000	361.763742	1.000000	8.000000	13.000000	9.000000	20.000000

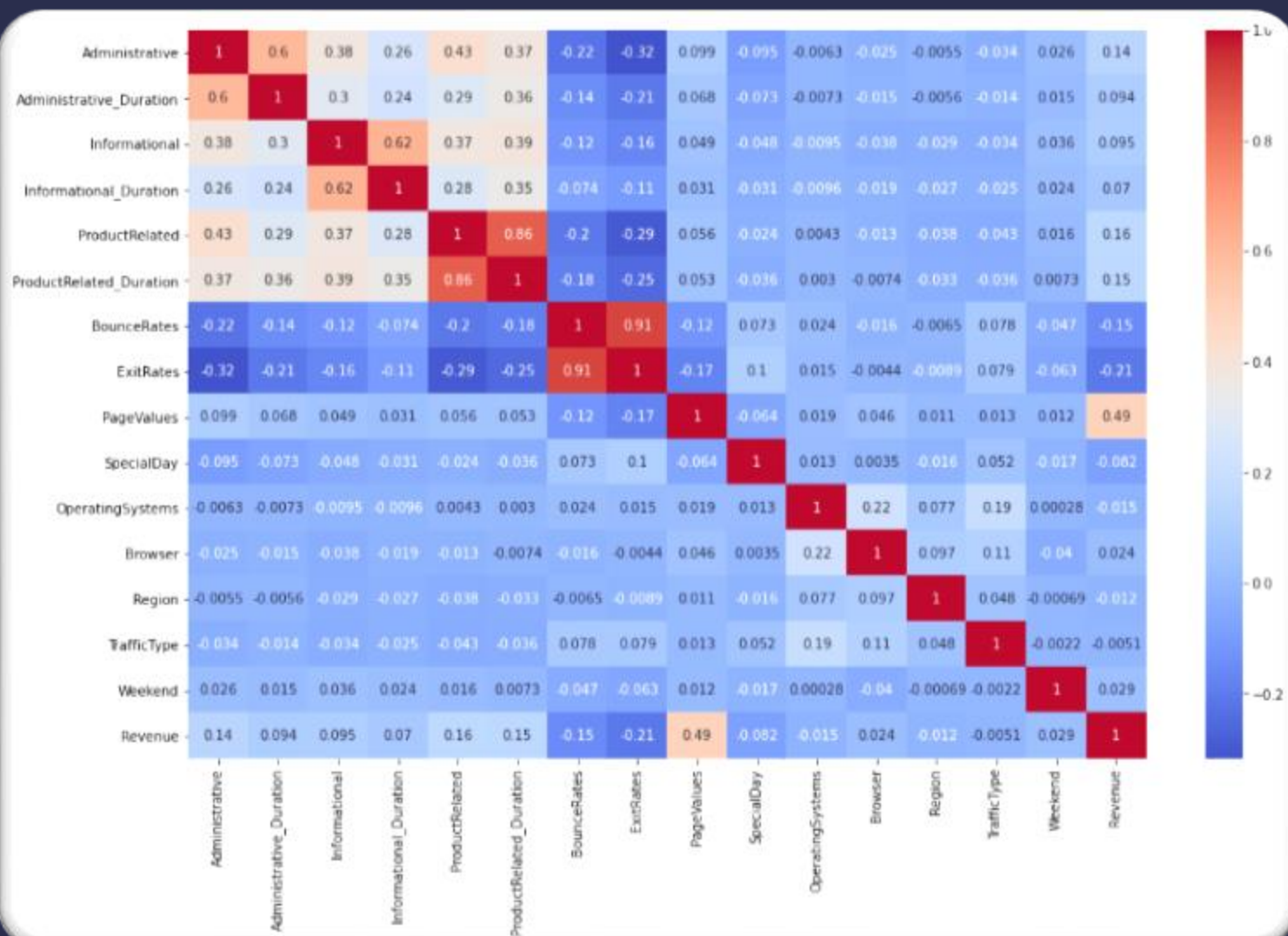


Wykresy zliczeń danych Kategorycznych

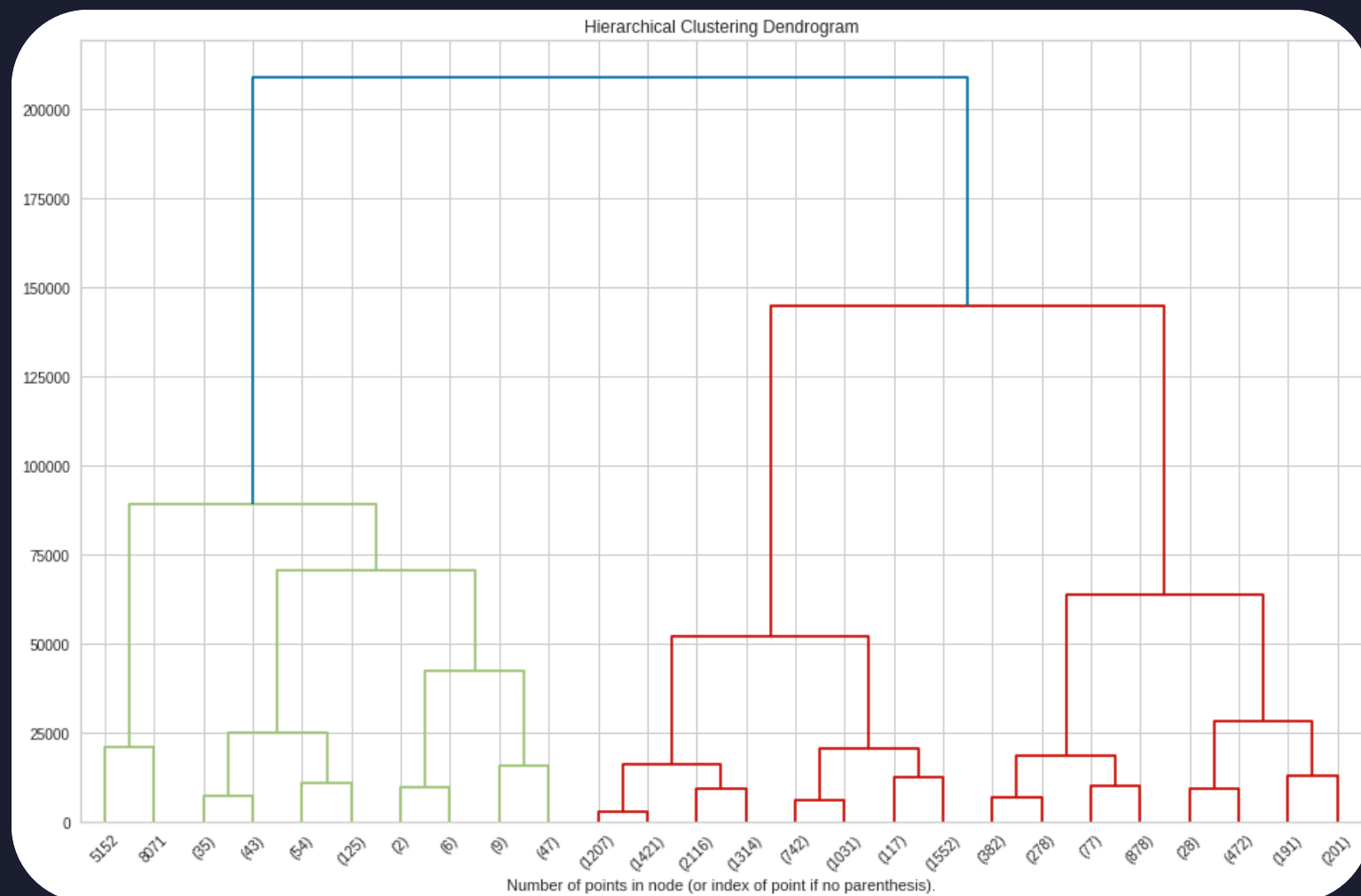


Wykresy Skrzynkowe

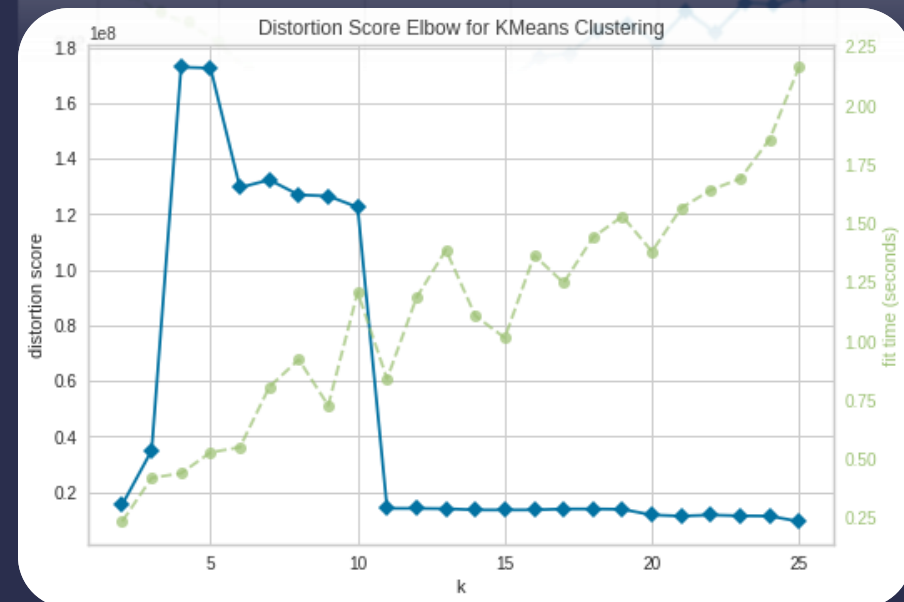
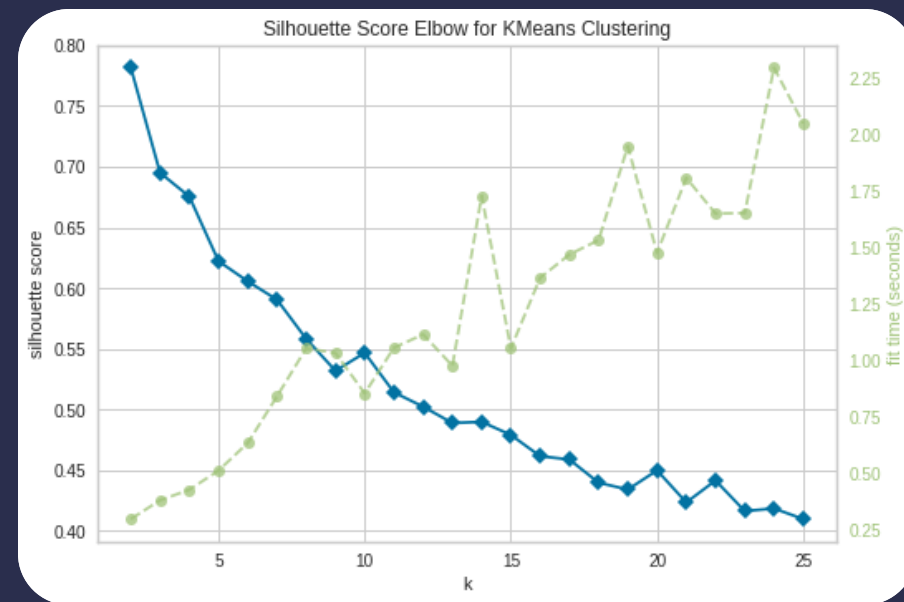
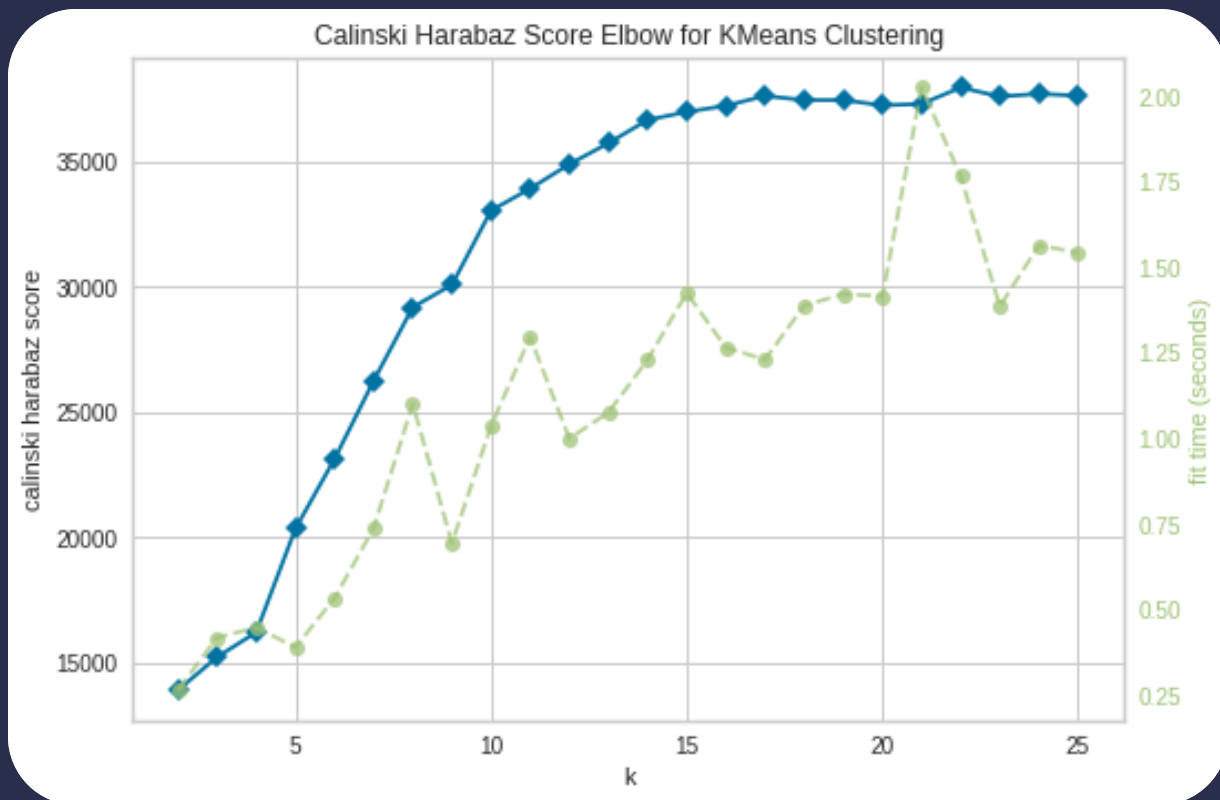
Macierz Korelacji Zmiennych



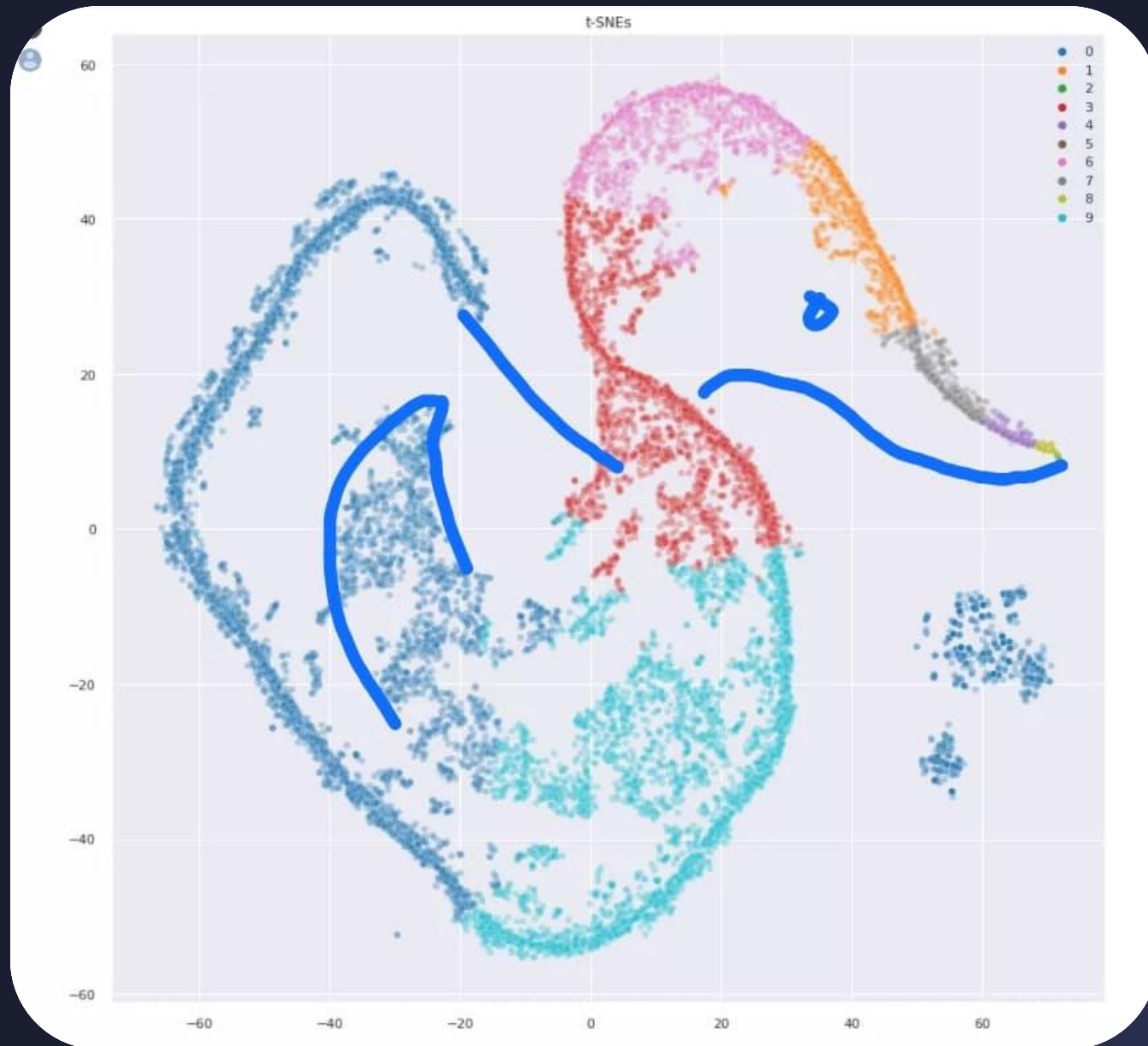
Hierarchiczny Dendrogram aglomeracyjnego klastrowania



Wykresy łokciowe



T-SNE po zmiennych numerycznych



T-SNE po one
hotcie



The background is a solid dark blue color. Scattered across the background are several bright pink, pill-shaped or capsule-shaped graphics. These shapes are oriented in various directions, some horizontally and some diagonally. They are distributed around the central text, with some appearing near the top and bottom edges and others closer to the center.

Dziękujemy za
Uwagę!