

# Klastrowanie



# Preprocessing

Miesiące zakodowaliśmy za pomocą cyklicznego kodowania,

resztę danych kategorycznych zakodowaliśmy za pomocą OneHotEncoding.

TSNE, którego używamy do obrazowania klastrowań zostało utworzone na podstawie danych użytych do danego klastrowania.

# Dane

Nie modyfikowane

Normalizowane

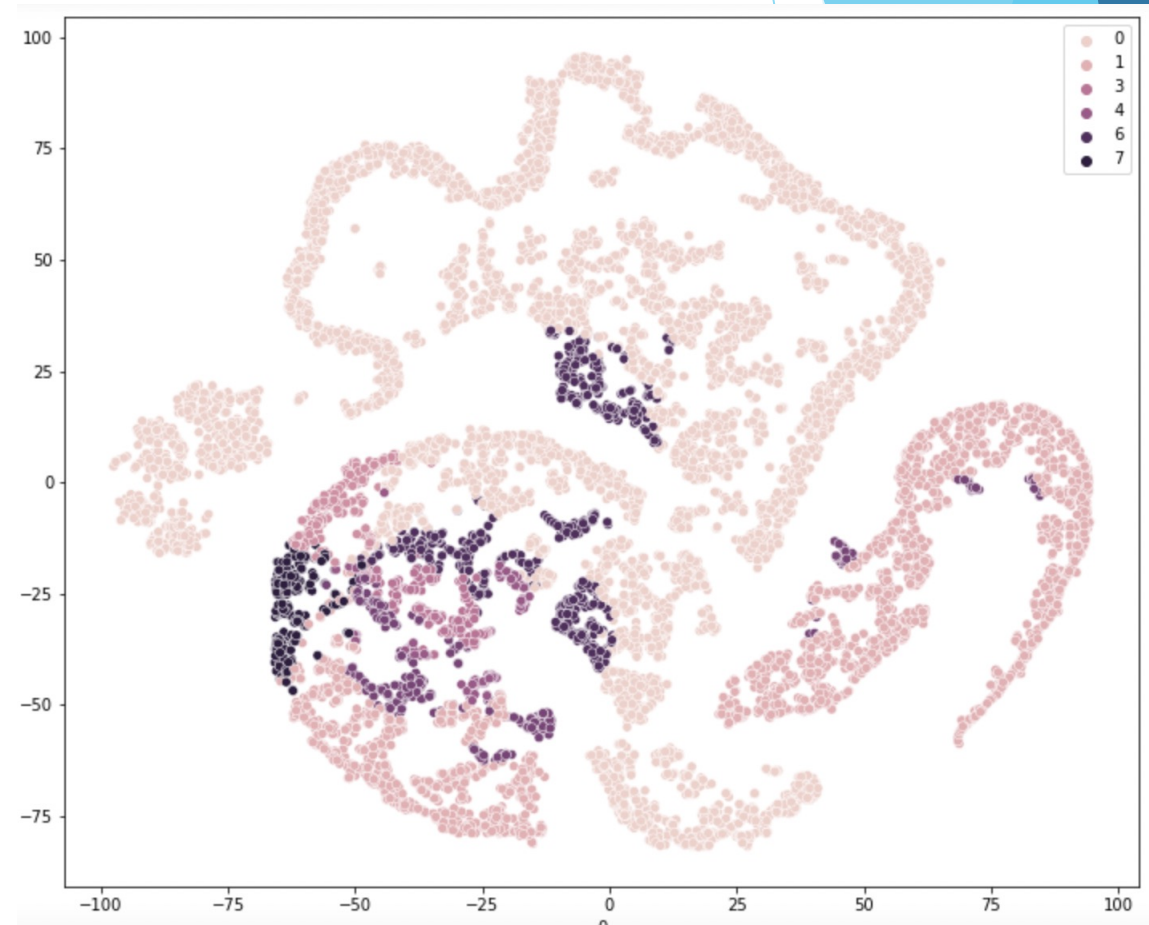
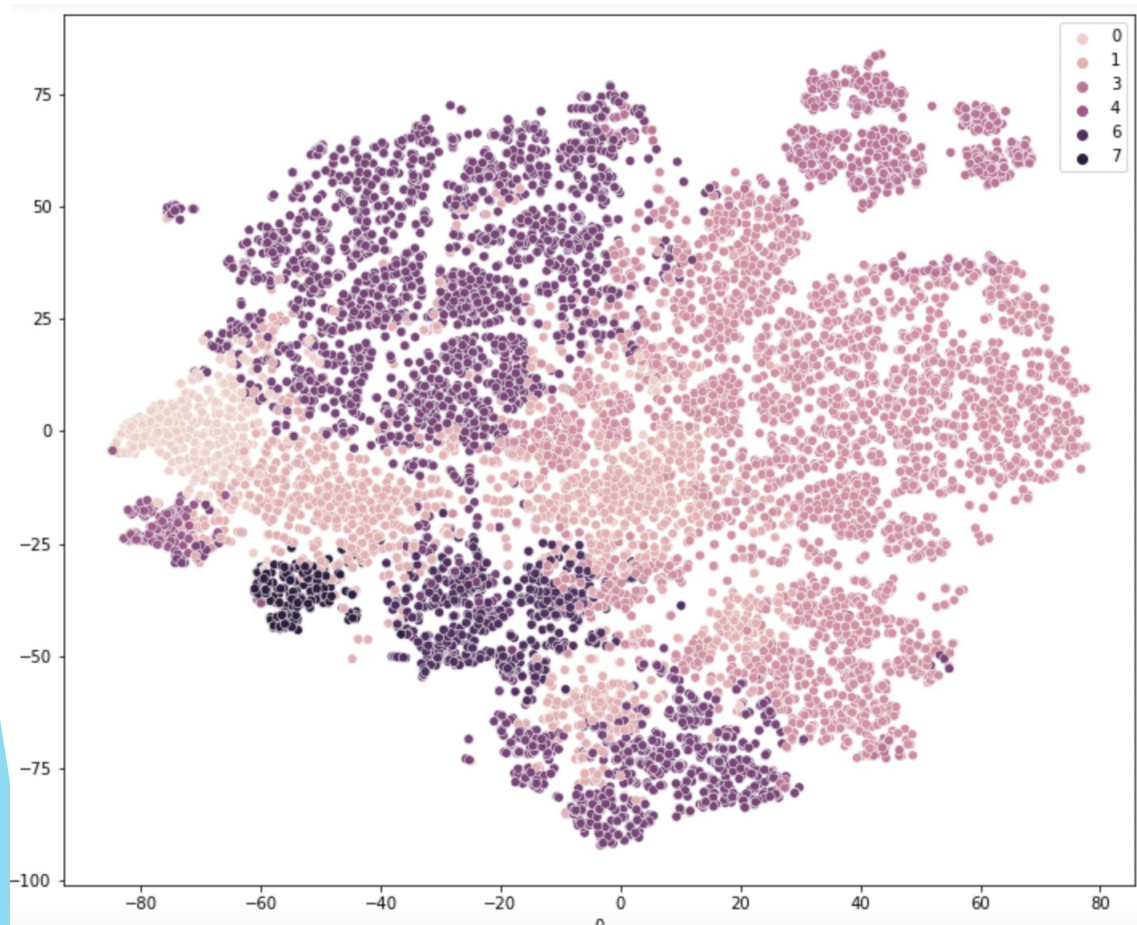
Skalowane

Po PCA

Po PCA normalizowane

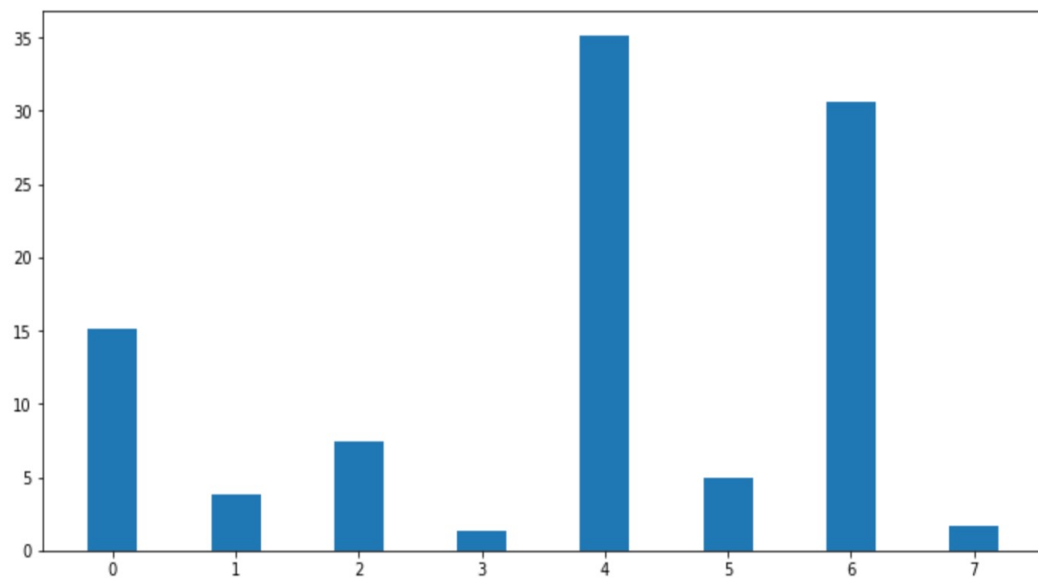
Po PCA skalowane

# Prezentacja finalistów

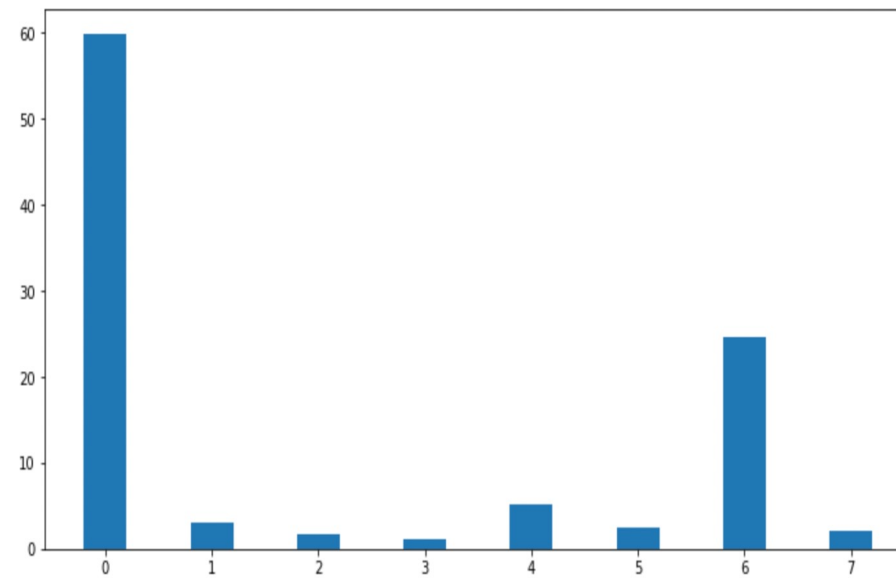


# Porównanie wielkości klastrow

Dla danych skalowanych



Dla danych normalizowanych po PCA

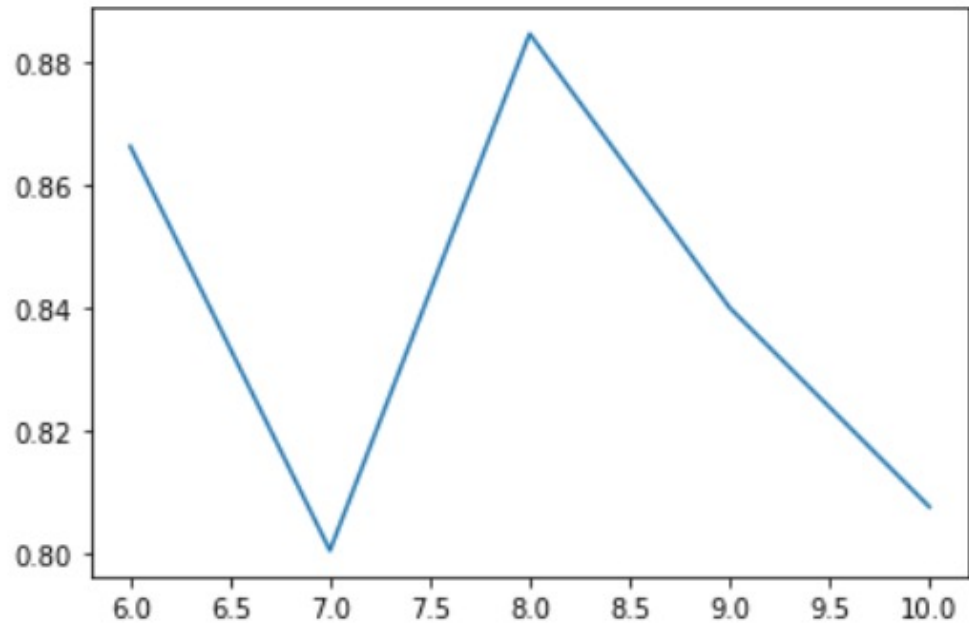


# Inne metryki

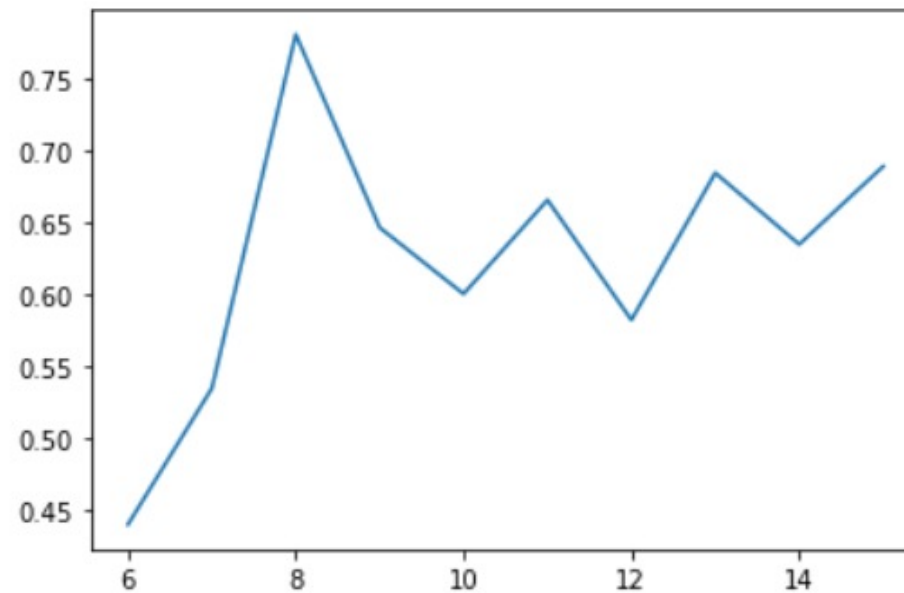
Metryka	Normalizowane po PCA	Skalowane
Minimalny dystans pomiędzy klastrami	0.00	0.21
Średni dystans pomiędzy punktami w tej samej klasie	0.41	4.56
Odchylenie standardowe odległości pomiędzy punktami w tej samej klasie	0.182	1.325
Średni dystans do środka klastra	0.29	3.19

# Stabilność

Normalizowane po PCA



Skalowane

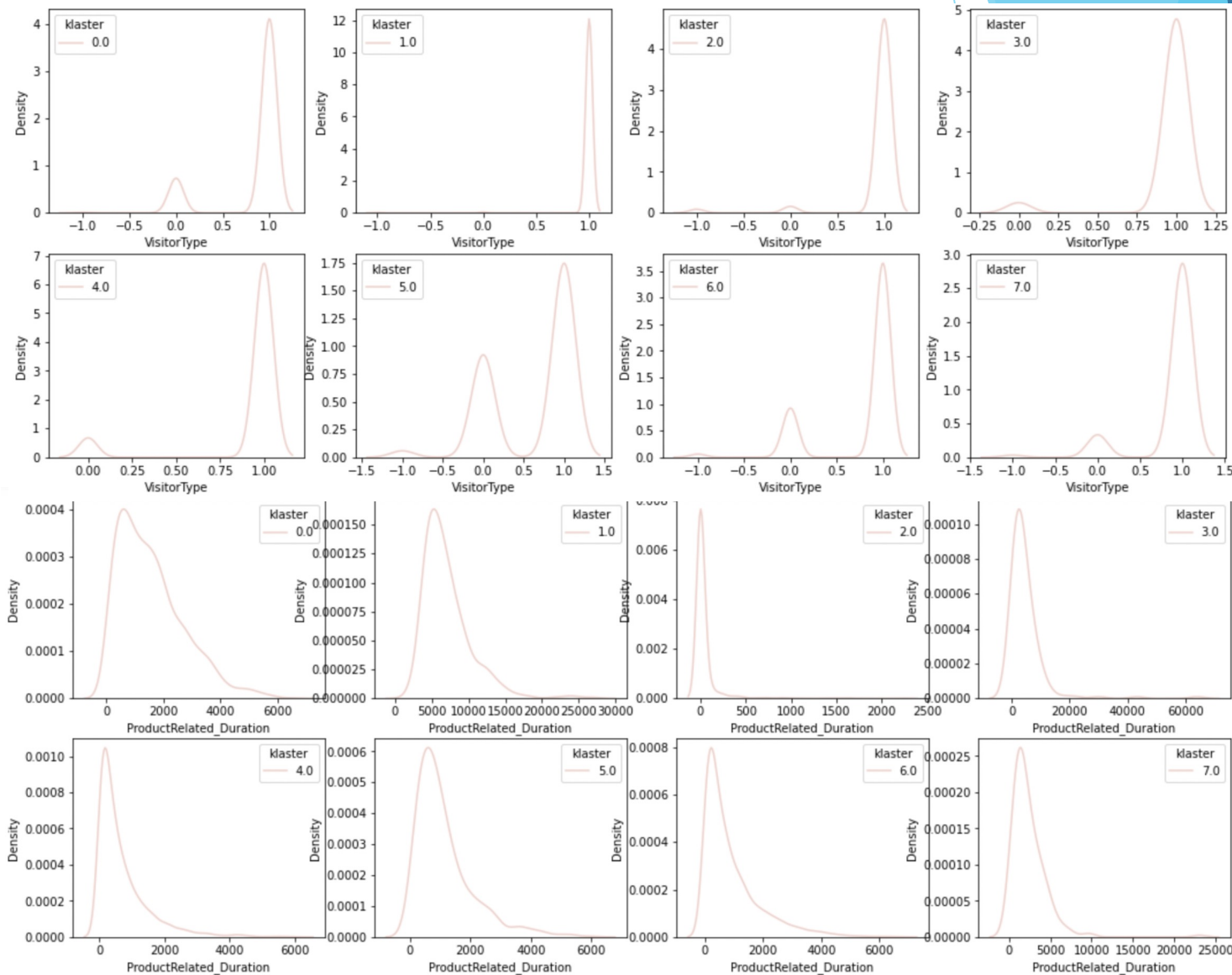
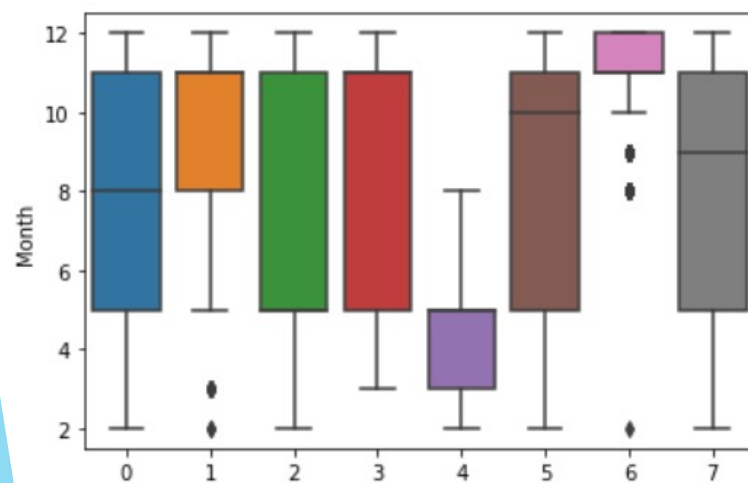
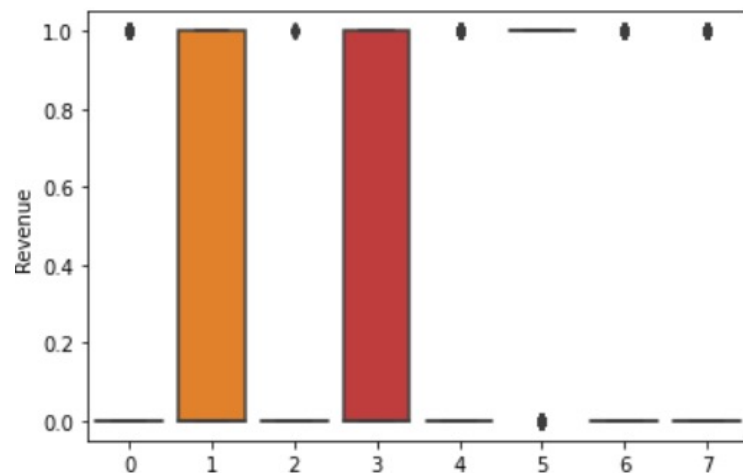




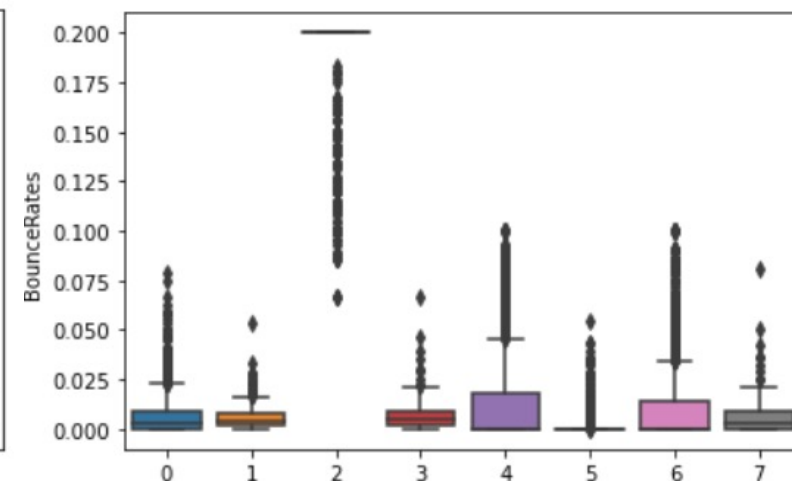
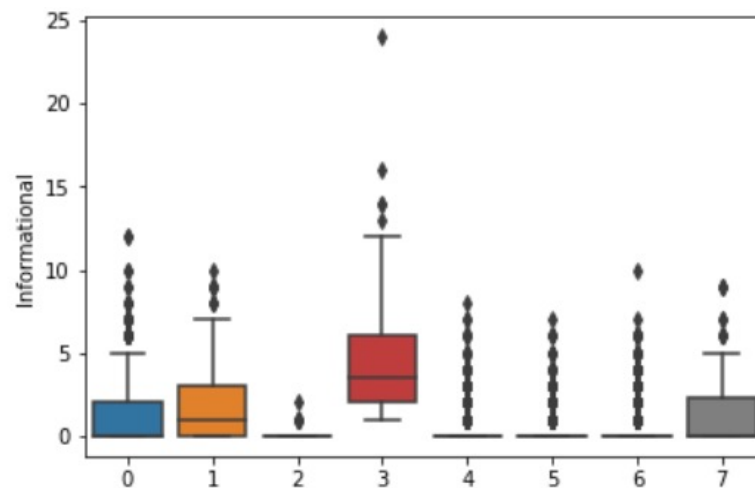
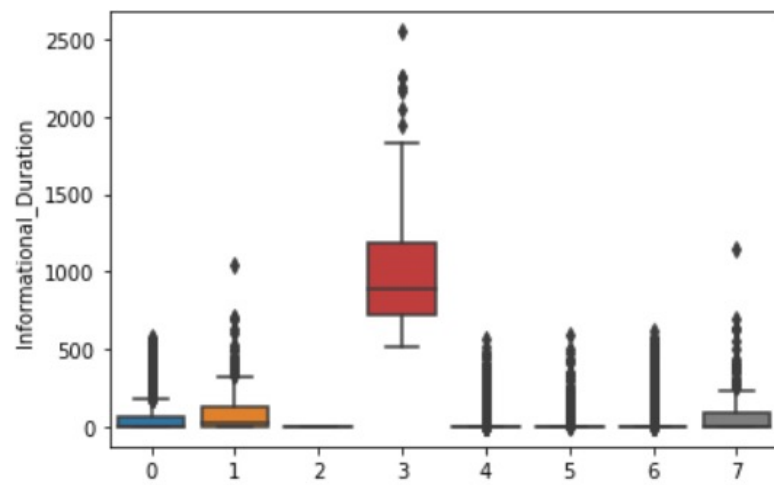
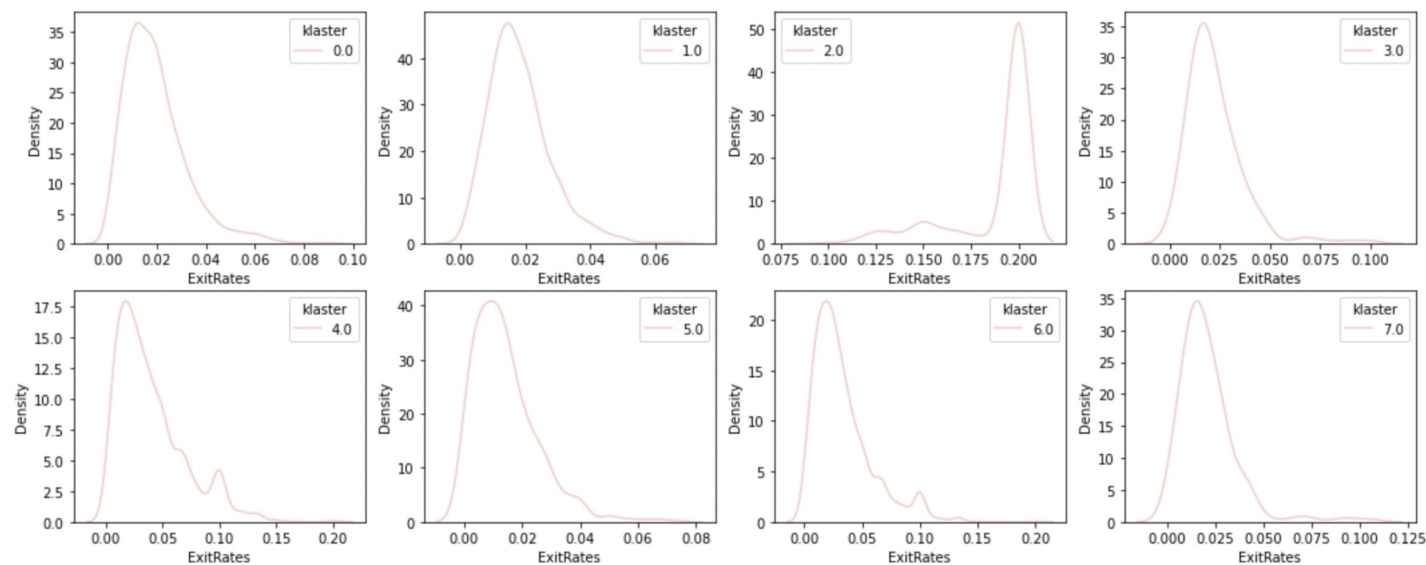
Klastry



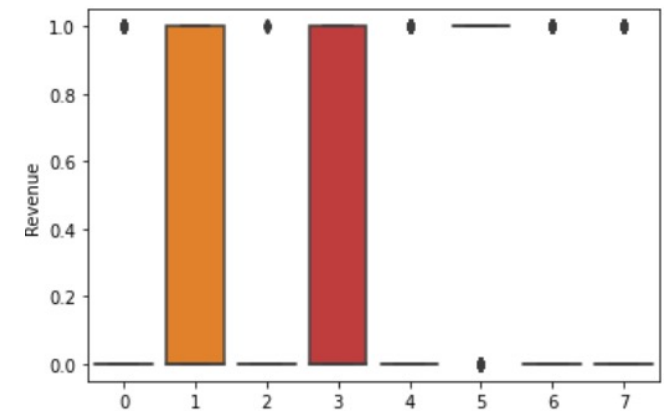
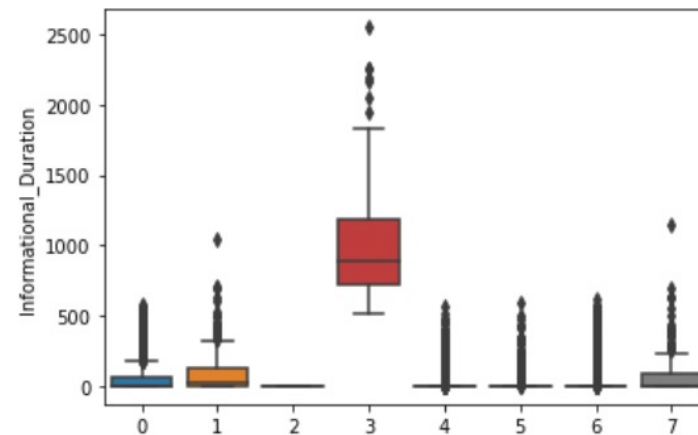
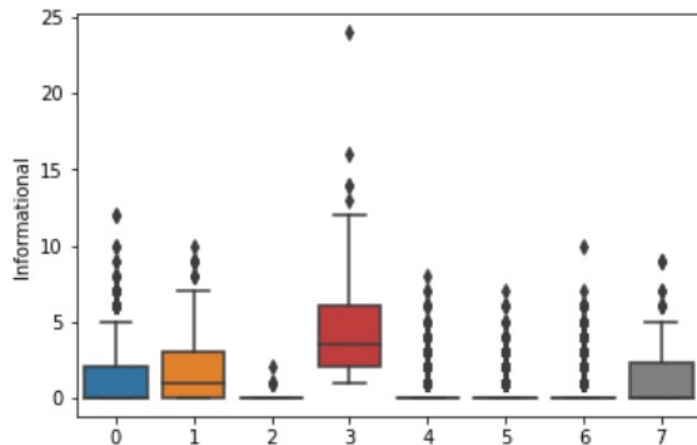
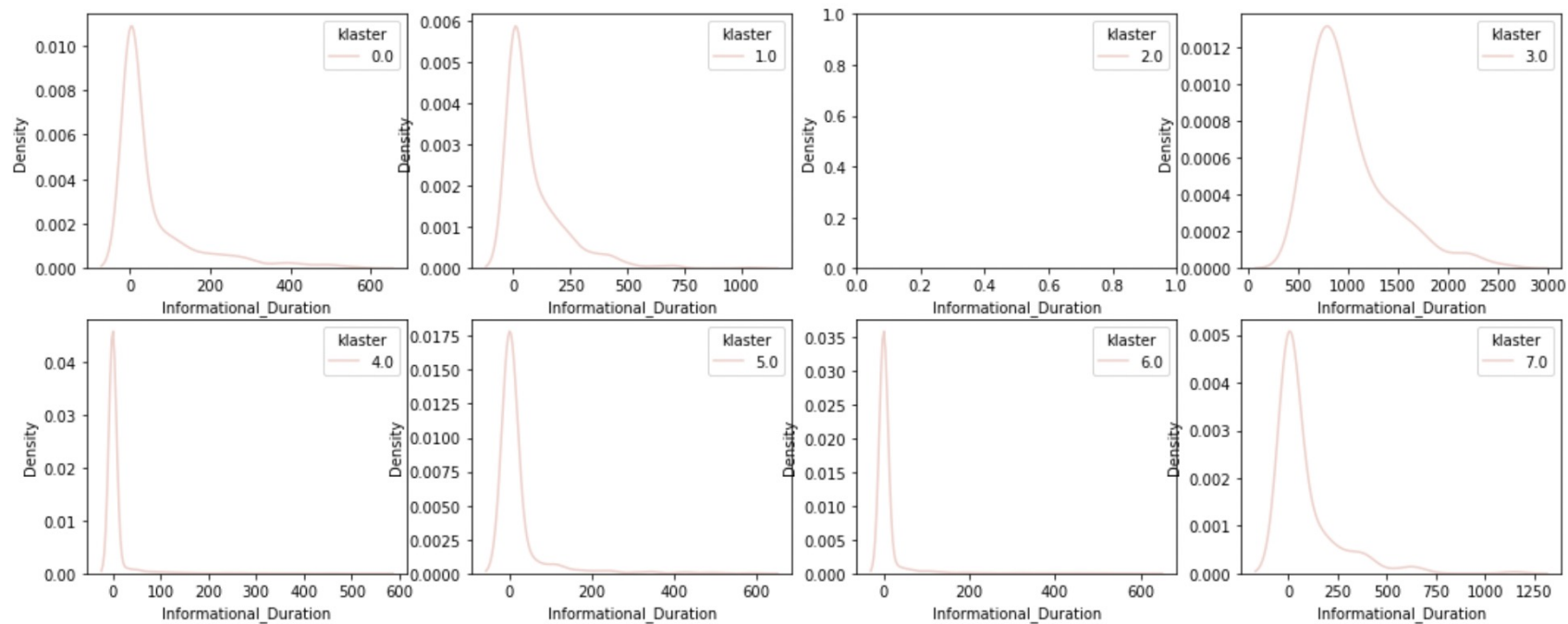
# Klaster 1- Uczeń



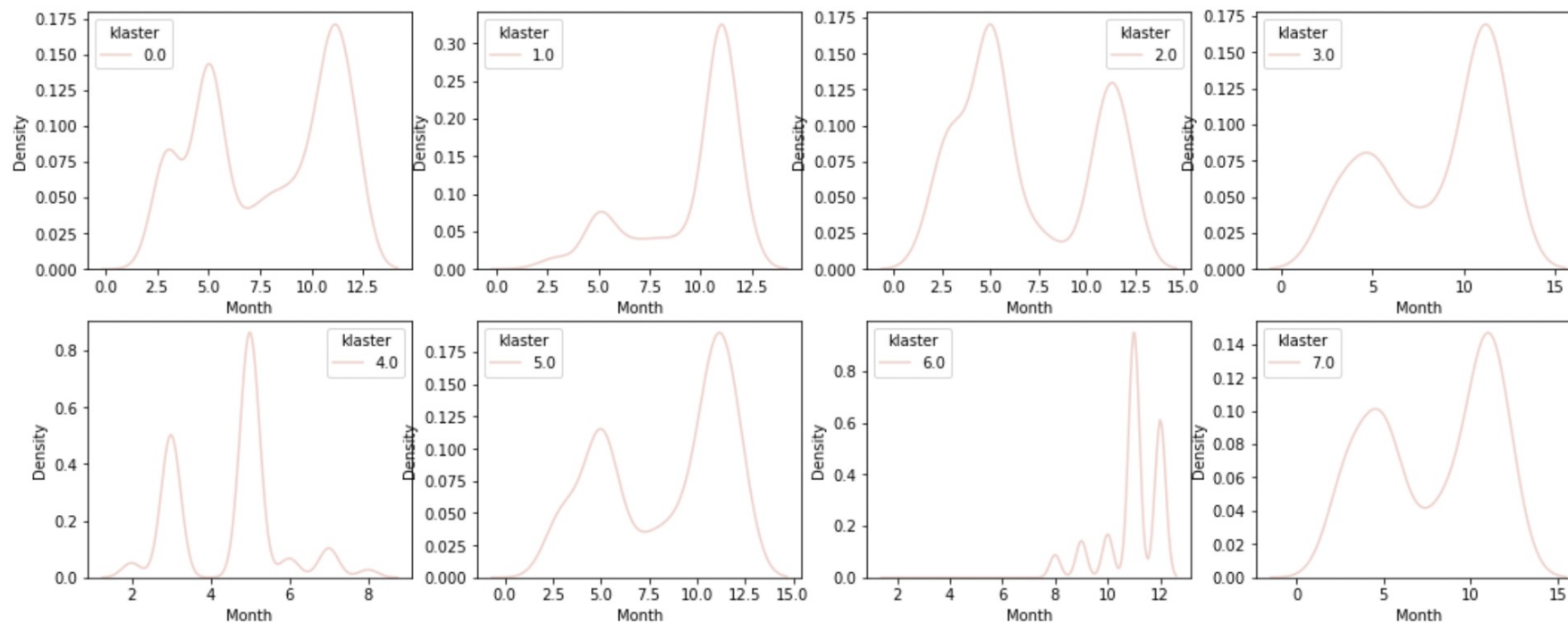
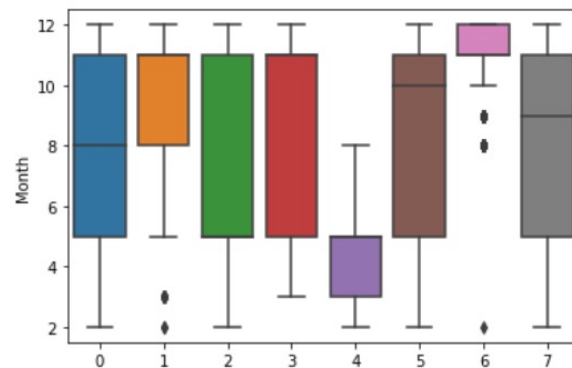
# Klaster 2- Skoczek



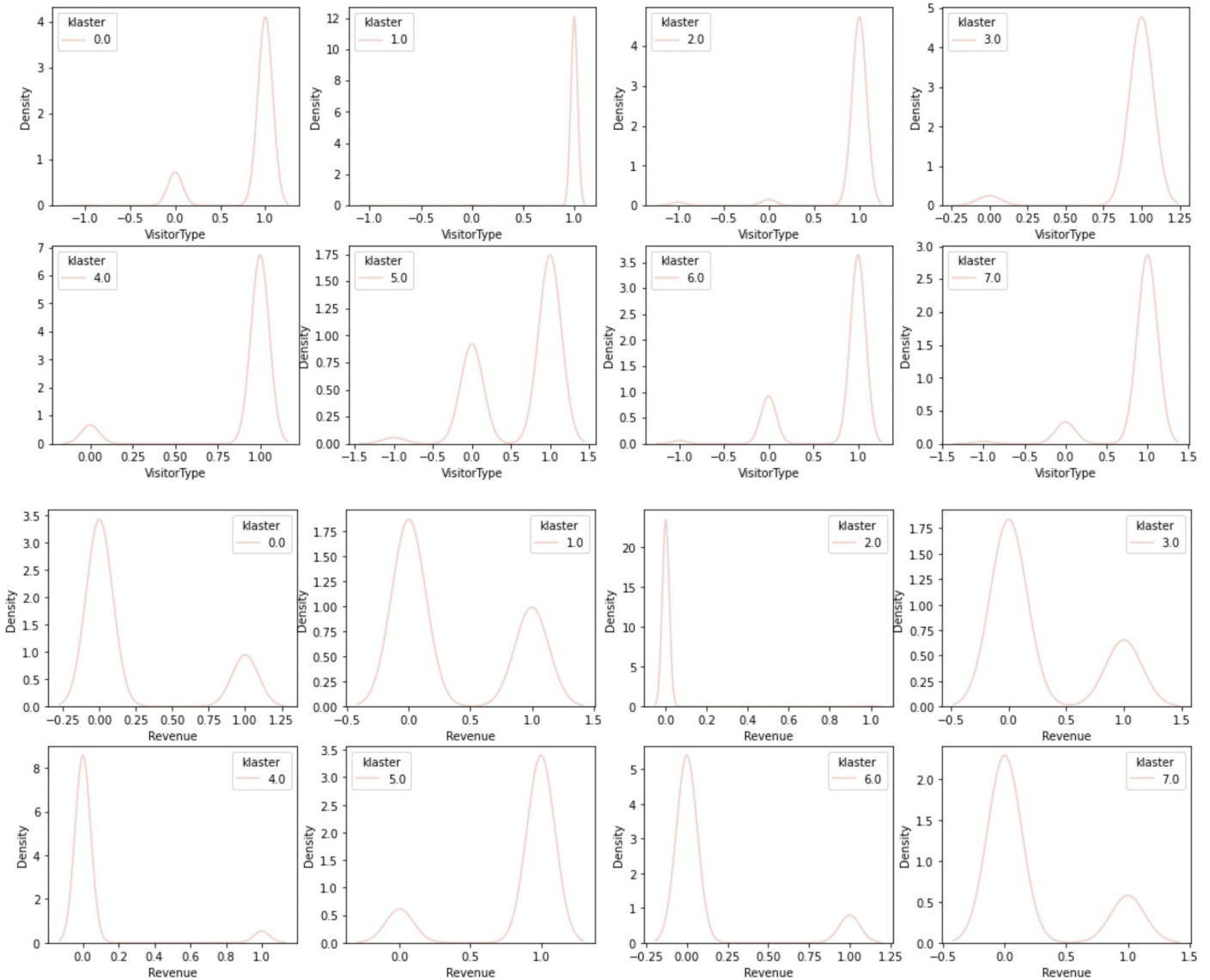
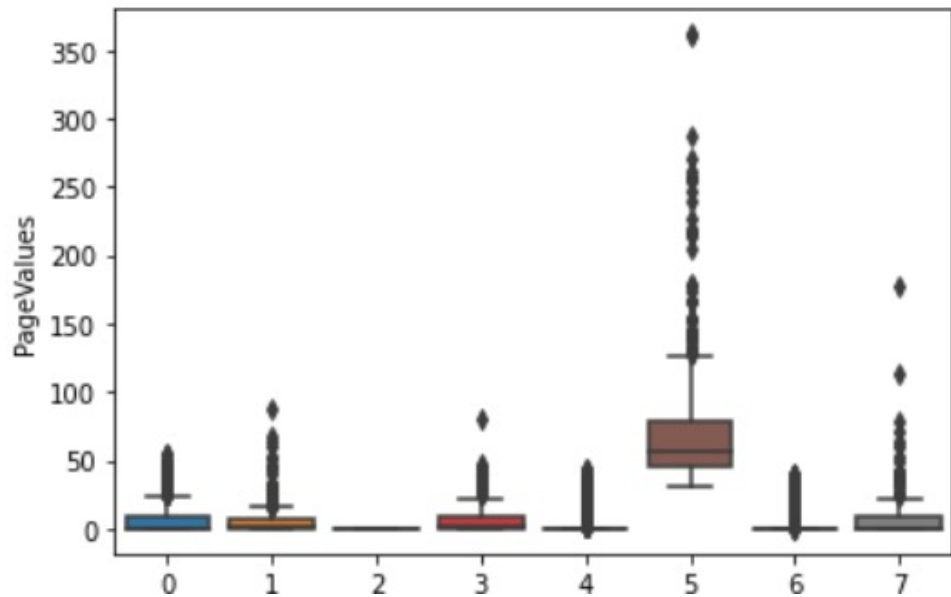
# Klaster 3 - Informacionista



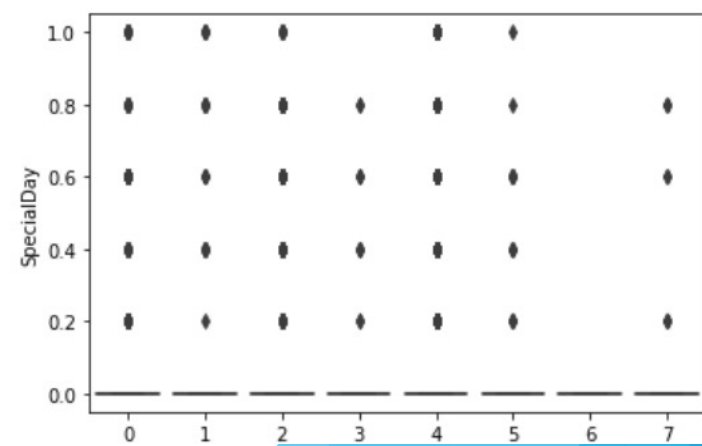
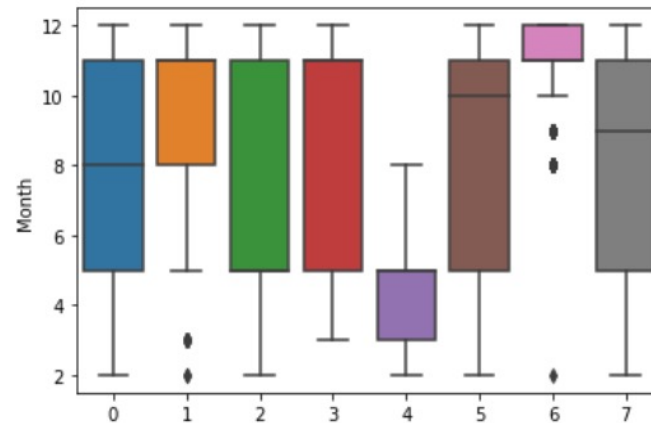
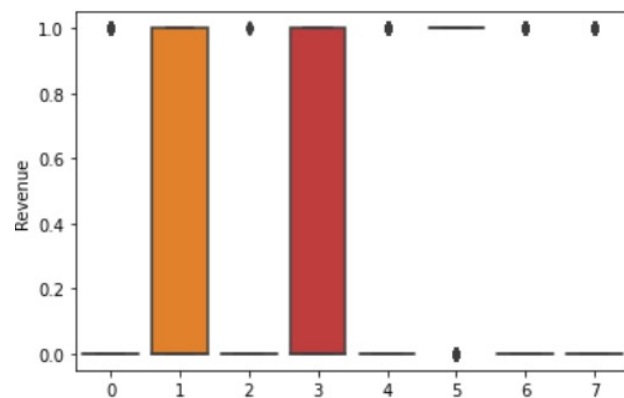
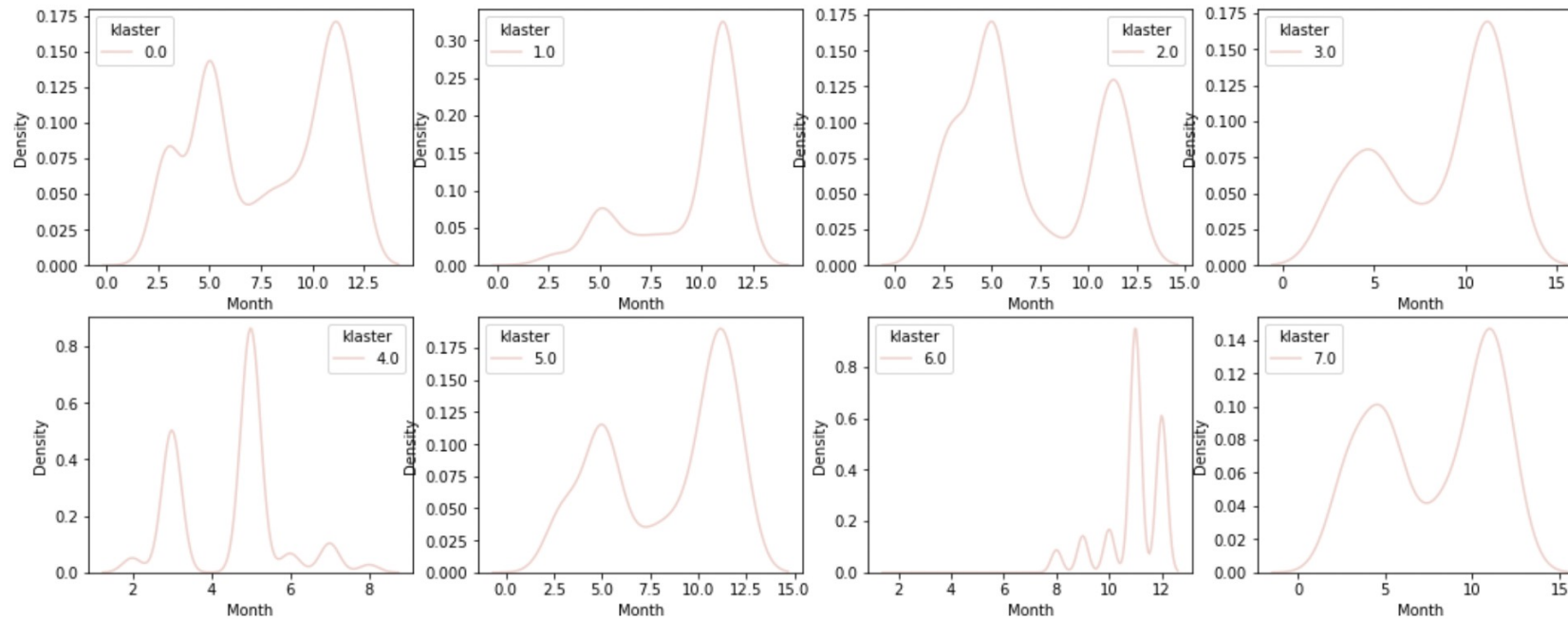
# Klaster 4 - wiosenny klient



# Klaster 5 - stały bywalec



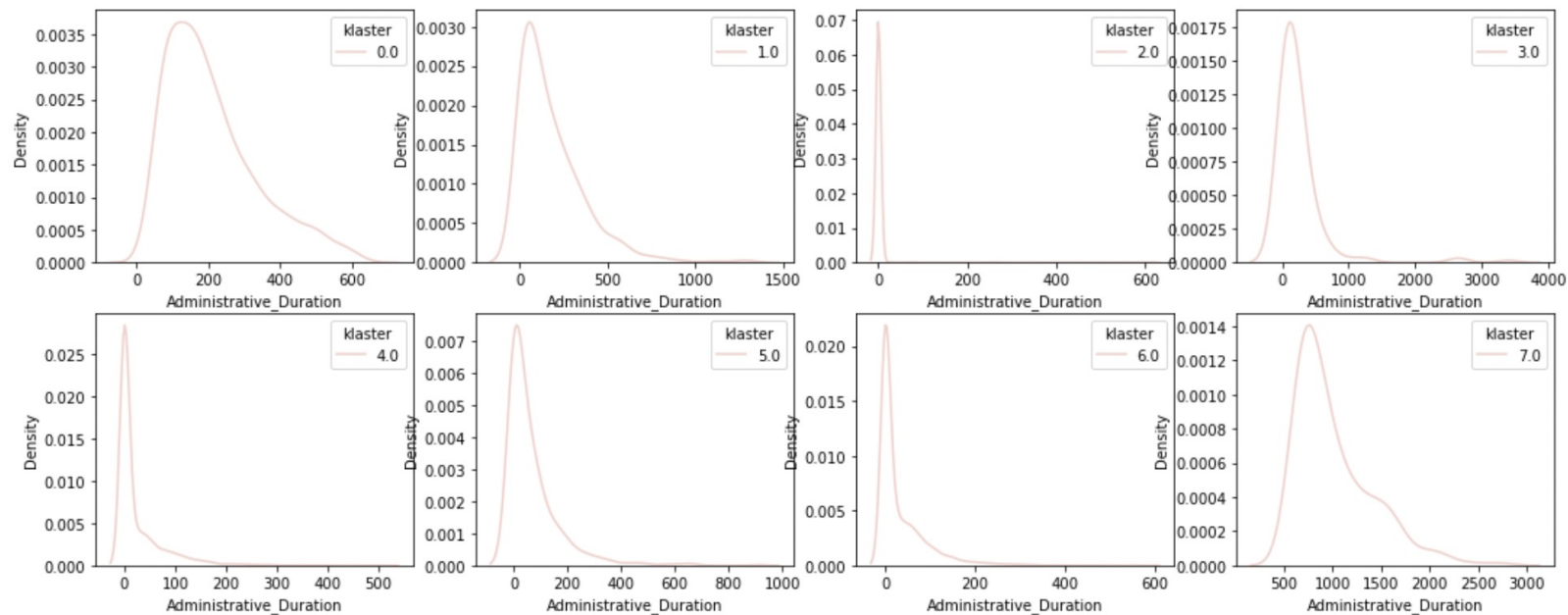
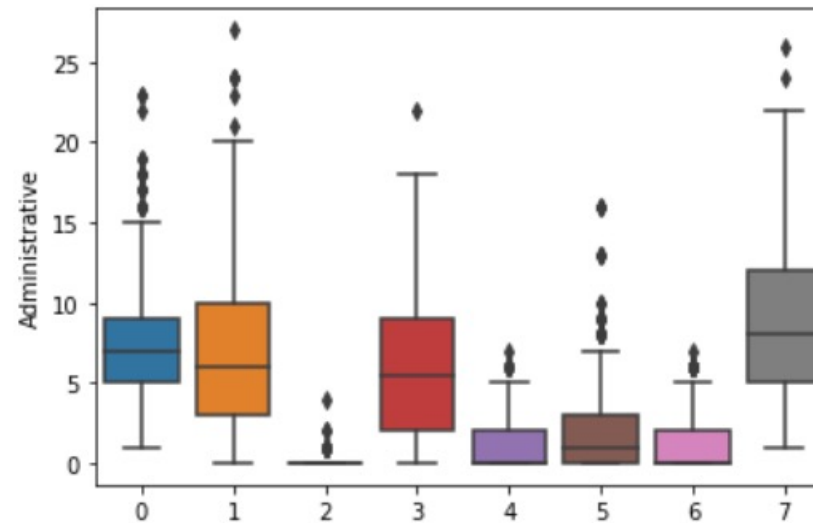
# Klaster 6- Święty Mikołaj





# Klaster 7

## Administrator



# Klaster 0- klient nijaki

