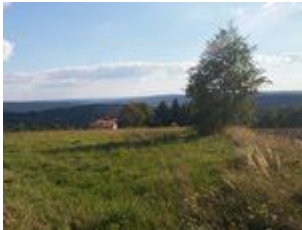


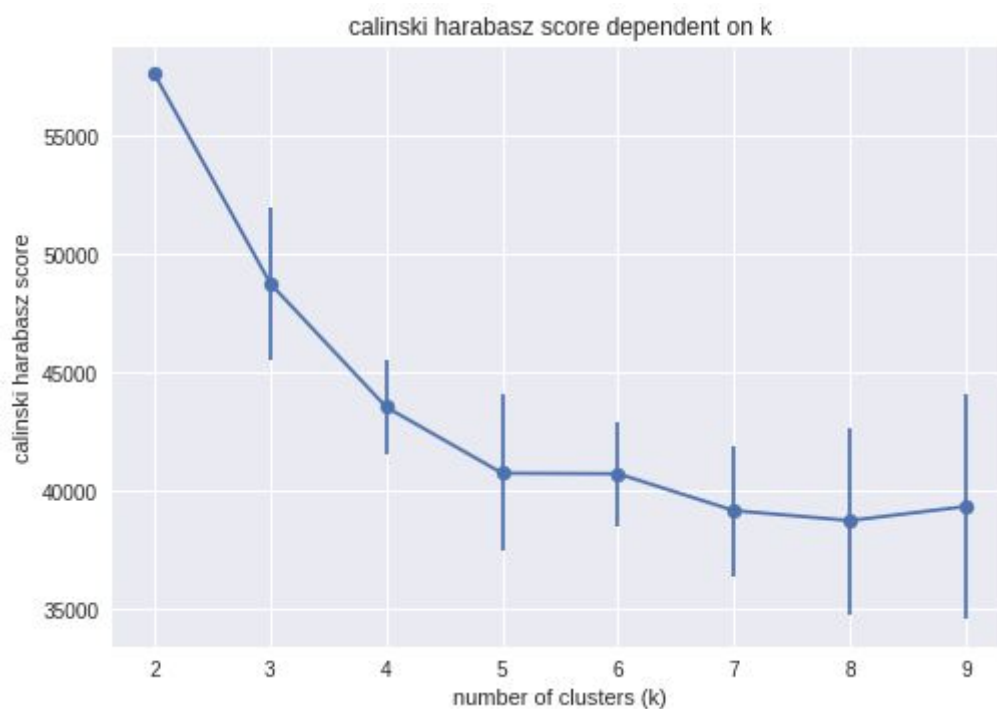
Anna Marciniak

Podstawy Uczenia Maszynowego lab 4 - Wykrywanie palety kolorów Raport

Do wykonania tego zadania użyłam poniższego zdjęcia w wielkości: 150px x 114px.

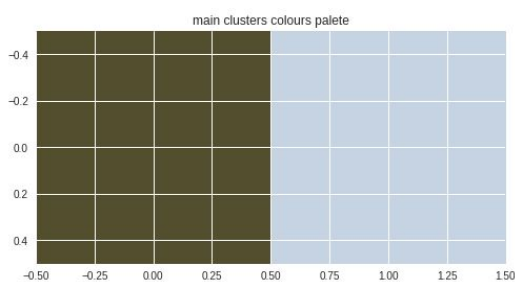


Usunęłam ze zbioru duplikaty. [Odległość kolorów](#) liczyłam metryką euklidesa. Do klasteryzacji użyłam metody k-medoids, aby wybrać k, posłużyłam się calinski harabasz score, obliczenia powtórzyłam 20 razy wyliczyłam średnią i odchylenie standardowe jako error bar.

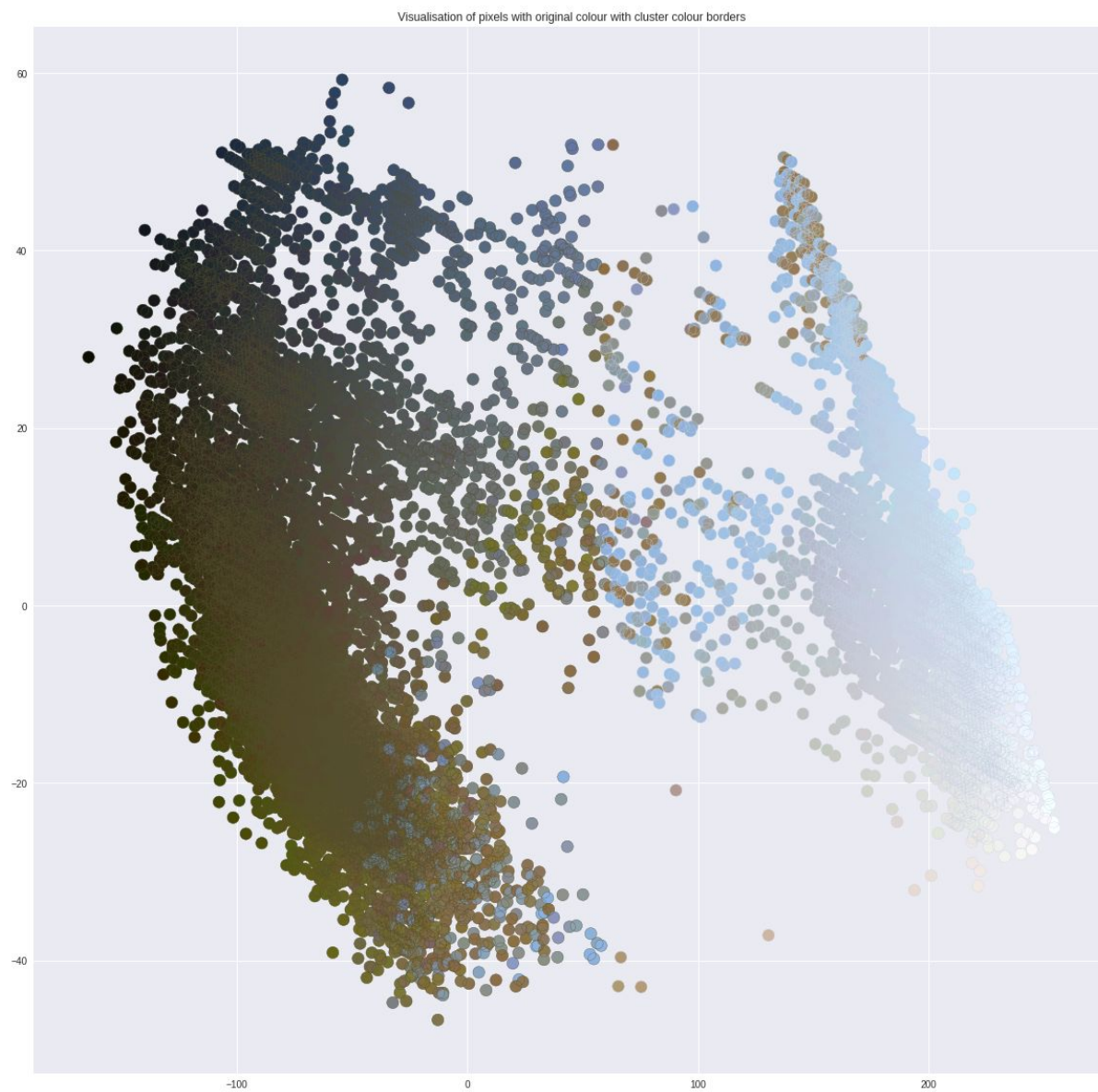


Największy score jest dla $k=2$, czyli powinno najlepiej klasteryzować dla 2, na podstawie tych wyników wybrałam $k=2$.

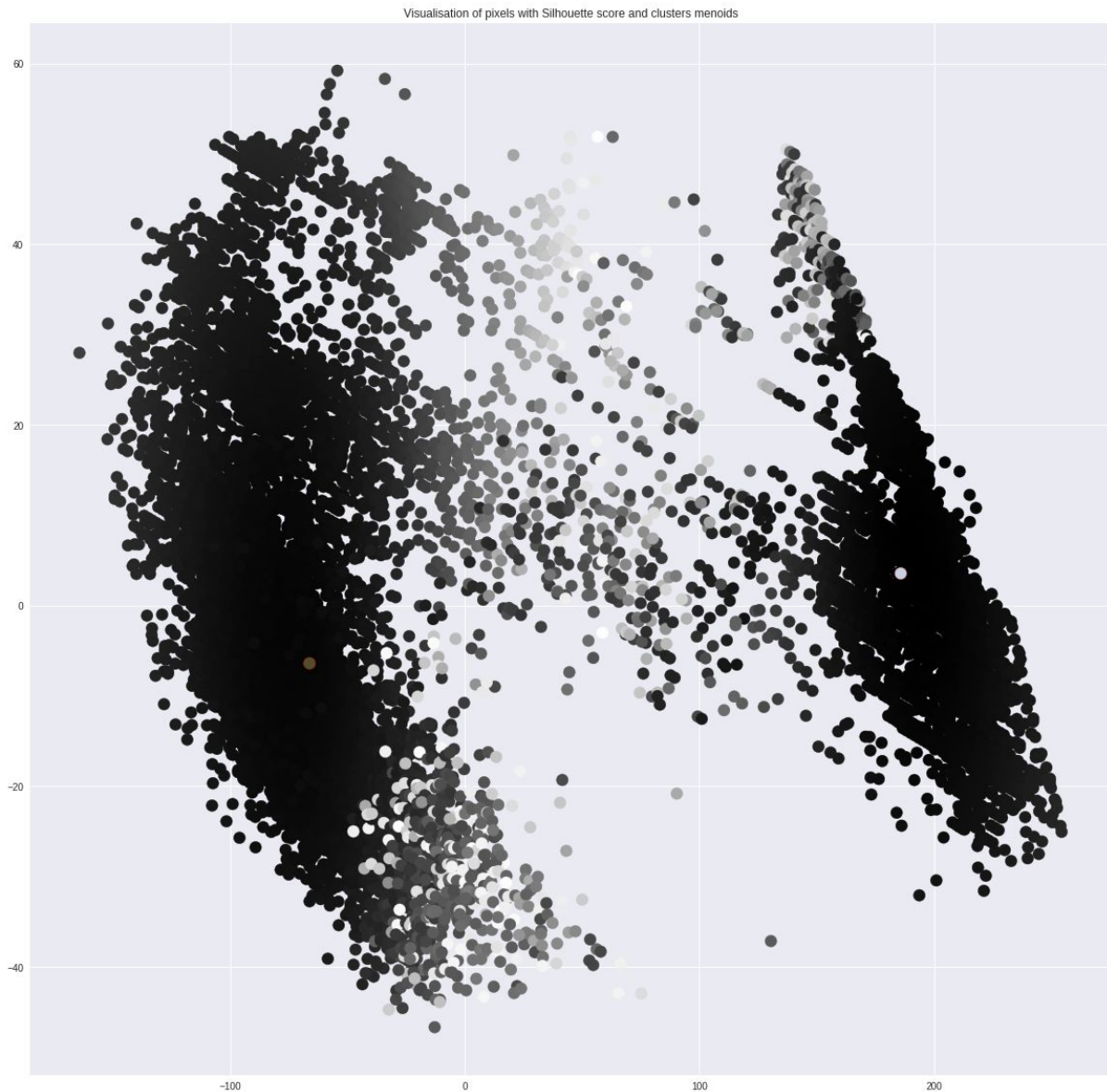
Wybrane centra posiadają taką paletę.



Zrzutowałam punkty na 2D, dałam oryginalny kolor jako środek i kolor klastra jako obwódka punktu.



Następnie policzyłam miarę Silhouette dla każdego z punktów, i zwizualizowałam to miarą szarości, nałożyłam centra klastrów w oryginalnych kolorach. Dodana czerwona obwódka to tych punktów, żeby je łatwiej było zobaczyć.



Wnioski:

- $k = 2$ wydaje się dobrym wyborem, gdyż mamy 2 dominujące kolory nieba i ziemi, które przedstawia paleta.
- k-monoids nawet dobrze podzieliło kolory na klastry, na rzucie na 2d widać, że są prawie rozdzielone.
- wartość miary Silhouette jest niska na granicy, i jest niska tam gdzie kolor niebieski zaklasyfikowany jako kolor zielony, więc można dzięki niej wykryć anomalie.