



**AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE**

Geograficzne Systemy Informacji Przestrzennej

Klasyfikacja danych w ArcGIS

**Wydział Geologii, Geofizyki i Ochrony Środowiska
Katedra Geologii Podstawowej i Geoturystyki
Kraków, 2014**

Grupowanie wartości atrybutów

- Istnieje szereg klasycznych metod klasyfikacji danych.

Metoda naturalnych przerw

- (ang. Natural breaks (Jenks))
- Tworzony jest szereg rozdzielczy, a następnie grupy są identyfikowane na podstawie wyraźnych przerw występujących w utworzonym ciągu danych. Jest to domyślna metoda klasyfikacji danych w ArcGIS.

HEIGHT	
11	Class: 11 - 11
15	Break: 11
18	Class: 12 - 19
19	Break: 19
29	Class: 20 - 35
30	
35	Break: 35
44	Class: 36 - 44

Metoda równych rozstępów

- Jest to metoda którą można porównać do linijki. Szerokości przedziałów klasowych są zawsze takie same.
- Np. przedziały o szerokości 10% (1-10%, 11-20%, 21-30%, itd.)

HEIGHT	
11	
15	
18	Class: 11 - 22
19	Break: 22
29	Class: 23 - 33
30	Break: 33
35	Class: 34 - 44
44	

Metoda kwantyli

- każda klasa zawiera jednakową liczbę elementów badanej cechy.
- Np. możemy mieć 15 województw pogrupowanych w trzech klasach. Niezależnie od wartości atrybutów, w każdej klasie będzie się znajdowało po pięć województw.

HEIGHT	
11	Class: 10 - 15
15	
18	Class: 16 - 19
19	
29	Class: 20 - 30
30	
35	Class: 31 - 44
44	

Metoda manualna

- Zakres każdej klasy jest określany przez użytkownika. Metoda ta jest szczególnie przydatna, gdy mamy odzwierciedlić jakieś szczególne kryteria lub typy danych.
- Np. mamy dane dotyczące temperatury. Możemy zechcieć określić granicę przedziałów klasowych dokładne na 32° Fahrenheita (0°C).

HEIGHT	
11	
15	
18	Class: Less than 30
19	
29	
30	Break: 29
35	Class: More than 30
44	

Decyzja o wyborze schematu klasyfikacji

- Schemat klasyfikacji danych obejmuje dwa zagadnienia:
- 1. wybór metody klasyfikacji
- 2. wybór liczby klas służących do pogrupowania danych

Wybór metody klasyfikacji

- Która metoda klasyfikacji jest najlepsza?

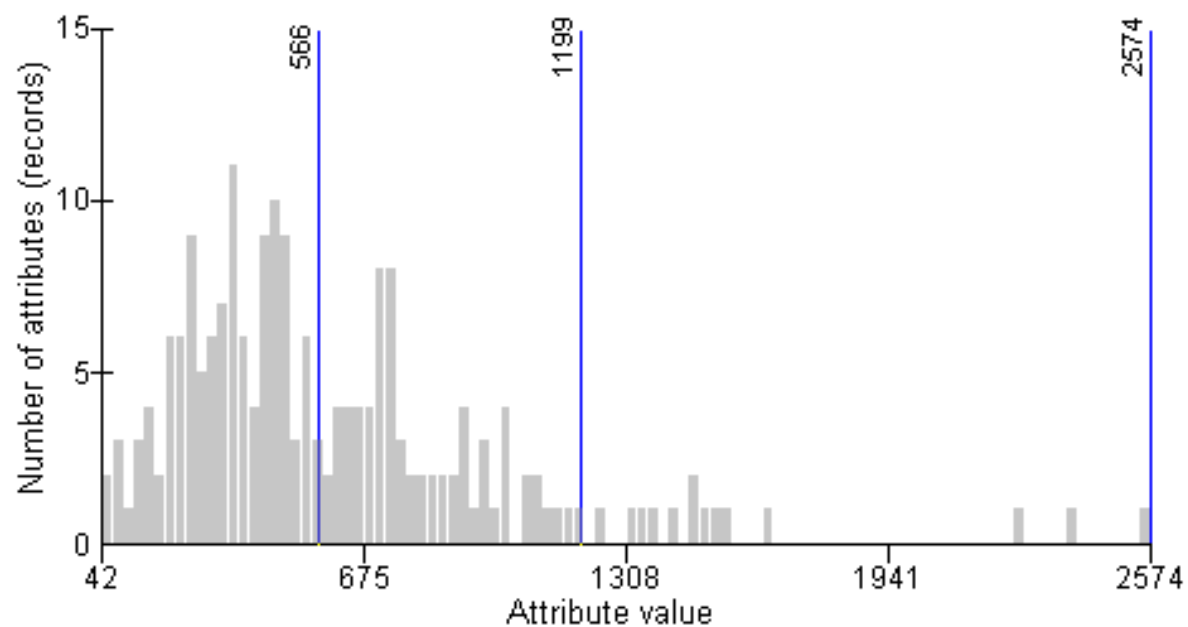
Nie ma „poprawnej” odpowiedzi na tak postawione pytanie.

- Najlepszy schemat klasyfikacji dla danej klasy zależy od przeznaczenia analizy, rodzaju analizowanych danych, przyjętej dokładności i innych.

Wybór metody klasyfikacji

- Jednym rozsądnym rozwiązaniem jest próba oparcia tej decyzji o same dane. Można np. spróbować różnych metod klasyfikacji i wizualnie ocenić mapy wynikowe, a następnie wybrać metodę, która wydaje się najlepiej ilustrować przedstawiane zjawisko.
- Do oceny schematów klasyfikacji, można także użyć histogramu klasyfikacji.

Histogram klasyfikacji



Wybór metody klasyfikacji

- Innym podejściem jest wybór schematu klasyfikacji na podstawie analizy (naukowej lub statystycznej) zmienności badanego atrybutu.
- Może wreszcie być tak, że możemy dysponować z góry określonymi standardami i kryteriami, określającymi metodę klasyfikacji lub liczbę klas.

Wybór metody klasyfikacji

Metoda	Kiedy używać	Ile klas
Naturalnej przerwy	Jeśli atrybuty są rozmieszczone nierównomiernie w całym zakresie zmienności analizowanego atrybutu (np. asymetria, wielomodalność rozkładów).	Wybierz liczbę, która najlepiej odzwierciedla naturalne grupy atrybutu, których zmienność chcemy zobrazować.
Równych rozstępów	Gdy chcemy, aby wszystkie klasy miały ten sam zakres.	Wybierz „okrągłą” liczbę, która będzie łatwa do zapamiętania i interpretacji np.: 2, 50, 1000, itd.,
Kwantyli	Jeśli atrybuty są rozmieszczone w sposób liniowy (równomierny rozkład w całym zakresie zmienności i niewielkie odchylenie liczebności elementów dla każdego przedziału klasowego).	Wybierz liczbę sensowną z punktu widzenia celów klasyfikacji.
Manualna	Gdy chcemy tworzyć klasy w oparciu o nasze doświadczenie i z uwzględnieniem specyfiki danych.	Wybierz liczbę adekwatną do zjawiska, które wizualizujemy. Np. potrzeba 2 klas, aby pokazać wartości powyżej i poniżej pewnej wartości progowej.

Wybór liczby przedziałów klasowych

- Przy podejmowaniu decyzji o liczbie klas, powinno się raczej używać ograniczonej liczby przedziałów klasowych. Najlepszym wyborem jest z reguły użycie 3-7 klas.