

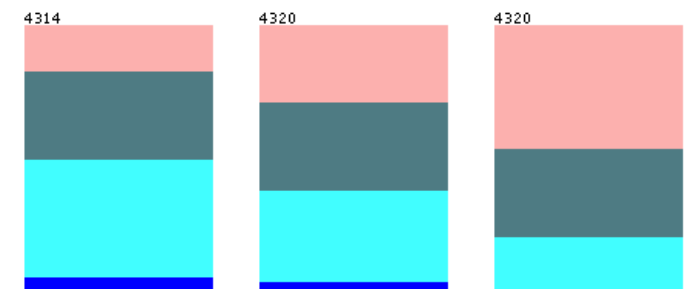
## Tugas 1

- (1) (50) Buatlah statistik yang menceritakan keadaan semua atribut  
Terdapat statistic data untuk
- parents { usual, pretentious, great\_pret }
  - has\_nurs { proper, less\_proper, improper, critical, very\_crit }
  - form { complete, completed, incomplete, foster }
  - children NUMERIC
  - housing { convenient, less\_conv, critical }
  - finance { convenient, inconv }
  - social { nonprob, slightly\_prob, problematic }
  - health { recommended, priority, not\_recom }
  - health2 { very\_recom, recommended, priority, not\_recom, spec\_prior }

contoh statistic

Name: parents		Type: Nominal	
Missing: 0 (0%)		Unique: 0 (0%)	
Distinct: 3			
No.	Label	Count	Weight
1	usual	4314	4314.0
2	pretentious	4320	4320.0
3	great_pret	4320	4320.0

Class: health2 (Nom) Visualize All



Name: has_nurs		Type: Nominal	
Missing: 0 (0%)		Unique: 0 (0%)	
Distinct: 5			
No.	Label	Count	Weight
1	proper	2587	2587.0
2	less_proper	2592	2592.0
3	improper	2592	2592.0
4	critical	2592	2592.0
5	very_crit	2591	2591.0

Name: has_nurs		Type: Nominal	
Missing: 0 (0%)		Unique: 0 (0%)	
		Distinct: 5	
No.	Label	Count	Weight
1	proper	2587	2587.0
2	less_proper	2592	2592.0
3	improper	2592	2592.0
4	critical	2592	2592.0
5	very_crit	2591	2591.0

Name: children		Type: Numeric	
Missing: 0 (0%)		Unique: 0 (0%)	
		Distinct: 4	
Statistic	Value		
Minimum	1		
Maximum	4		
Mean	2.5		
StdDev	1.118		

dst.

(2) (25) Adakah missing value? Outlier? Apakah yang anda lakukan terhadap data tersebut? Jelaskan.

Ada missing value di baris 55,97,134,443. Data (rows) dibuang.

Ada jumlah children = -1 baris 3714. Data dibuang.

(3) (25) Usulkanlah representasi dari setiap atribut

Bisa menggunakan representasi seperti adanya atau dengan diubah menjadi bilangan numeric. Misal untuk data parents: **usual**, **pretentious**, **great\_pret** bisa diganti menjadi 1,2,3.