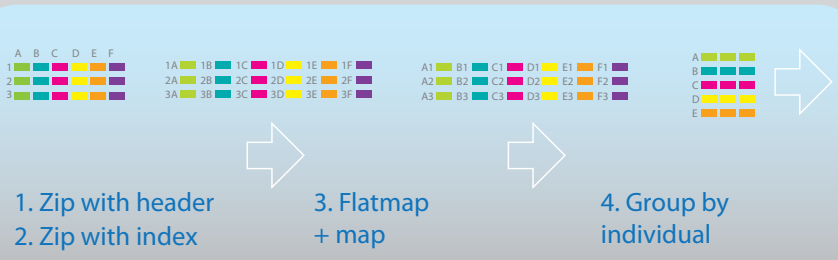


VCF file

#	CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO
#	1	100		A	G	100		
#	1	100		A	T	100		
#	1	100		A	C	100		
#	1	100		A	G	100		
#	1	100		A	T	100		
#	1	100		A	C	100		
#	1	100		A	G	100		
#	1	100		A	T	100		
#	1	100		A	C	100		

**custom function*



Interface: convert from VCF to sample-genotype vector

Kmeans



MLlib

Cluster
Visualization



Interface: visualize
genotype-based
clusters