

GSA-5859

SISTEMAS DE INFORMAÇÃO GEOGRÁFICA EM SOFTWARE LIVRE

CARLOS HENRIQUE GROHMANN

INSTITUTO DE ENERGIA E AMBIENTE - USP

Dados vetoriais

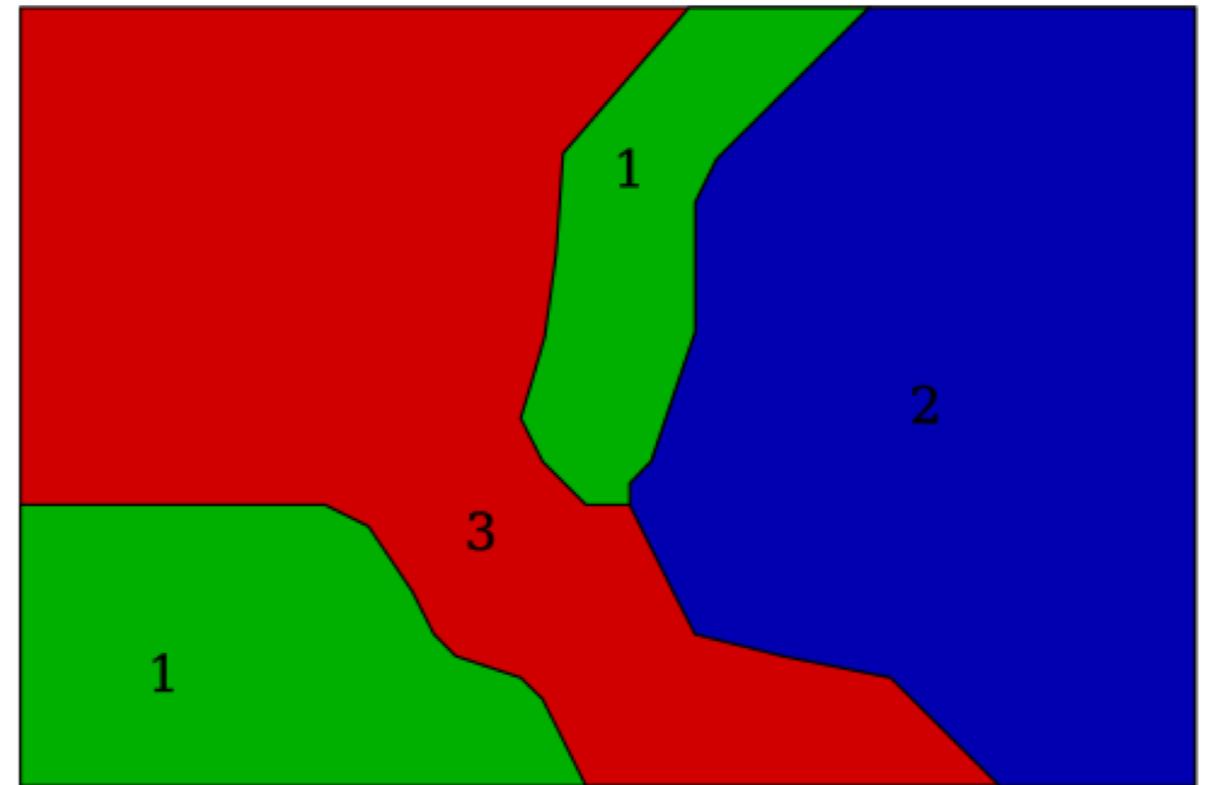
TIPOS DE DADOS

- ▶ Vetoriais
 - ▶ são usados para informações de objetos lineares ou áreas definidas por linhas fechadas (polígonos). Uma linha conecta dois pontos extremos (end points -- nós ou vértices), cada um com suas respectivas coordenadas XY ou XYZ. Cada objeto vetorial pode possuir vários (ou nenhum) atributos, armazenados em um banco de dados

TIPOS DE DADOS

Representações Raster e Vetorial dos mesmos objetos em SIG

3	3	3	3	3	1	2	2	2
3	3	3	3	1	2	2	2	2
3	3	3	3	1	2	2	2	2
3	3	3	3	1	2	2	2	2
1	1	1	3	3	2	2	2	2
1	1	1	1	3	3	3	2	2



1 = área vegetada

2 = água

3 = área urbana

REPRESENTANDO O MUNDO EM SIG

- ▶ o mundo é infinitamente complexo
- ▶ o conteúdo de um banco de dados espacial representa uma visão limitada da realidade
- ▶ o banco de dados espacial é um modelo da realidade
- ▶ o usuário vê o mundo real por meio do banco de dados

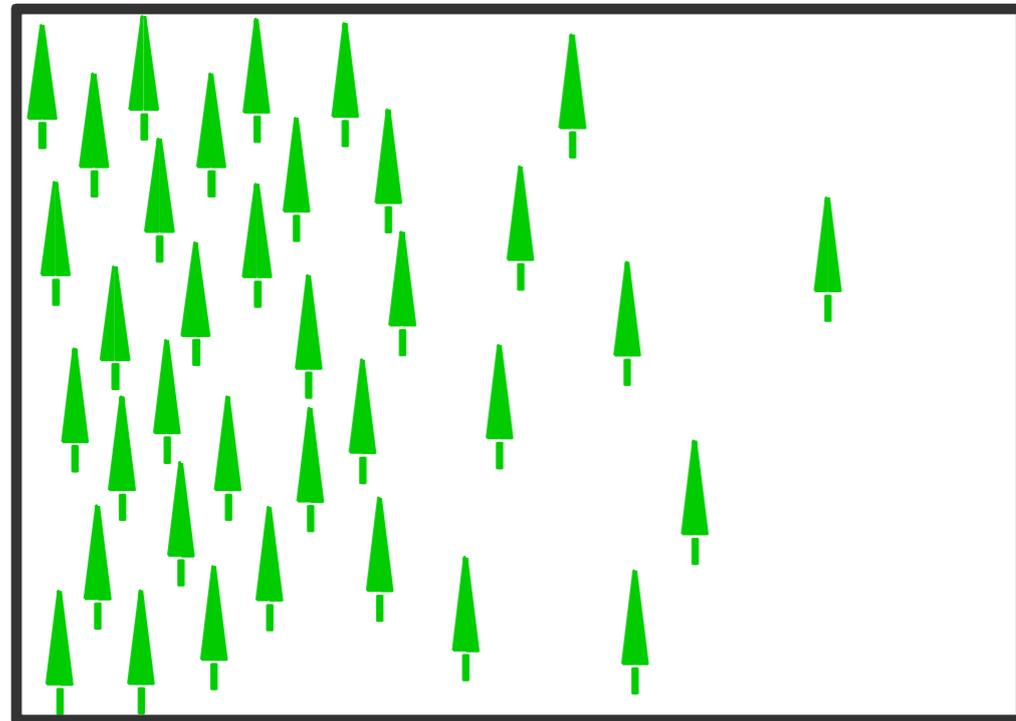
REPRESENTANDO O MUNDO EM SIG

- ▶ Um banco de dados pode incluir
 - ▶ versões digitais de objetos reais
 - ▶ por ex. casas, florestas, rios
 - ▶ versões digitais de objetos fictícios (inventados)
 - ▶ por ex. limites políticos

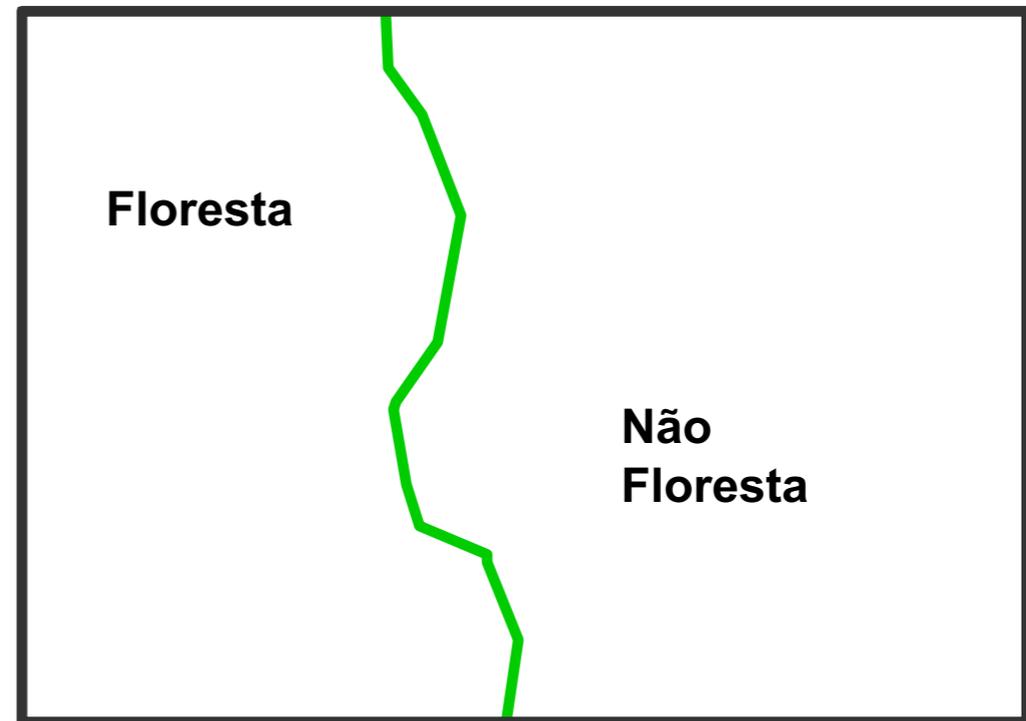
REPRESENTANDO O MUNDO EM SIG

- ▶ computadores são bons para armazenar dados espaciais discretos, maus para dados contínuos
- ▶ algumas coisas são entidades discretas
 - ▶ não são problema para representar no computador
- ▶ outras existem em todo o espaço e variam continuamente (ex. temperatura)
 - ▶ a variação precisa ser aproximada usando representações discretas

REPRESENTANDO O MUNDO EM SIG



Realidade



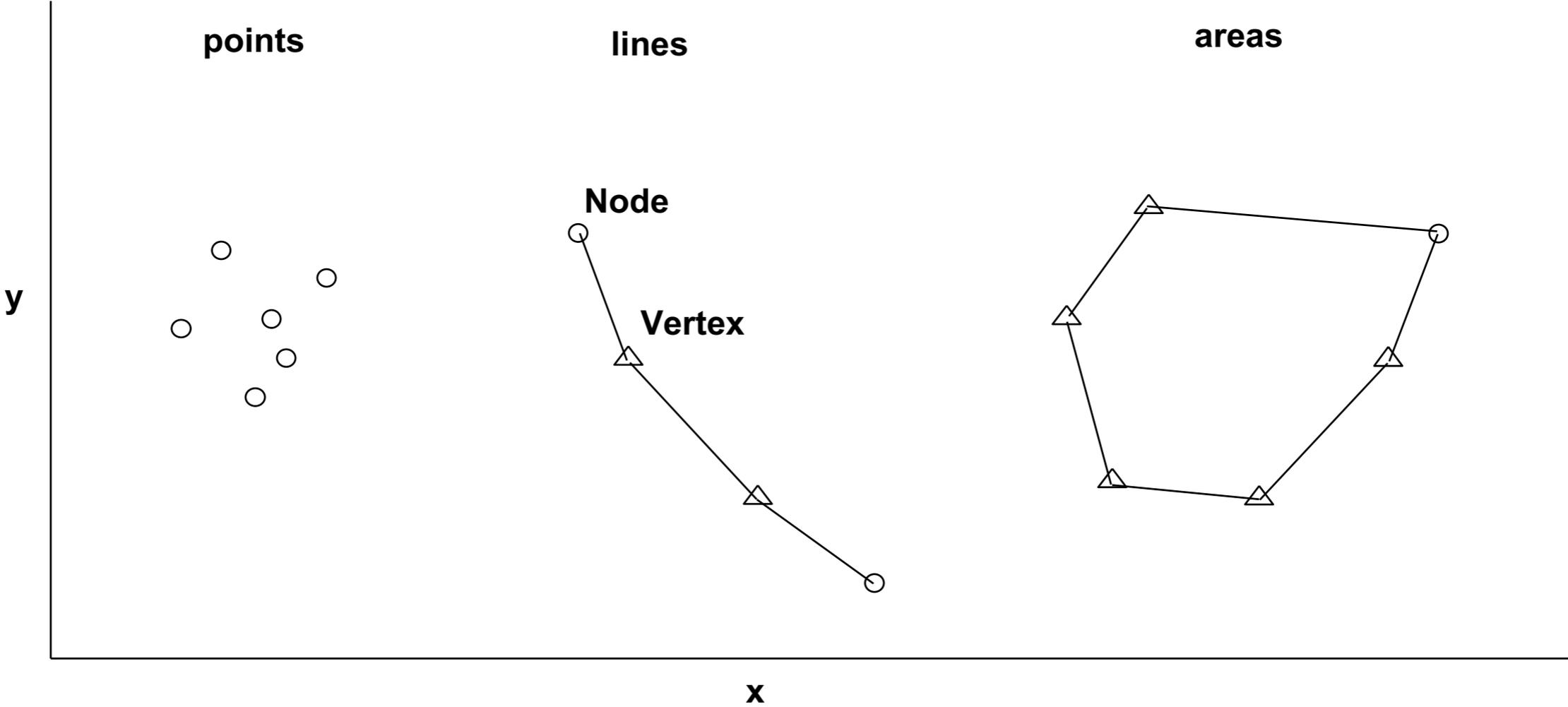
Representação no SIG

às vezes a distinção entre discreta e contínua não é muito clara

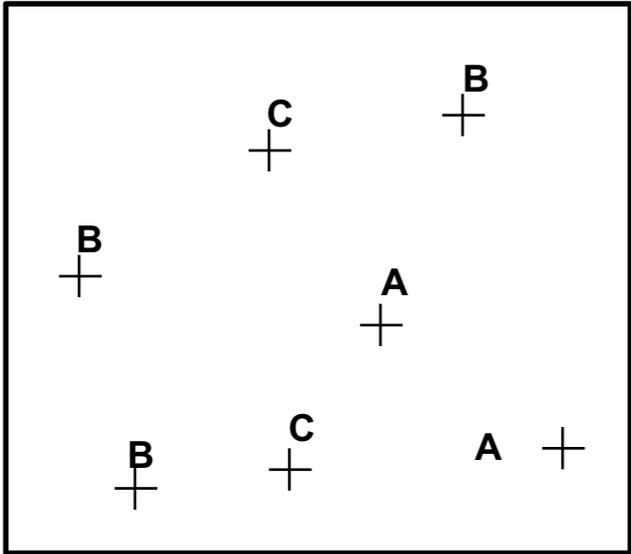
MODELO VETORIAL

- ▶ objetos do mundo real representados por pontos, linhas e áreas.
- ▶ pontos identificam localizações
- ▶ linhas conectam pontos
- ▶ arcos - linhas orientadas (vetores s.s.)
- ▶ polígonos consistem de segmentos de linhas conectados

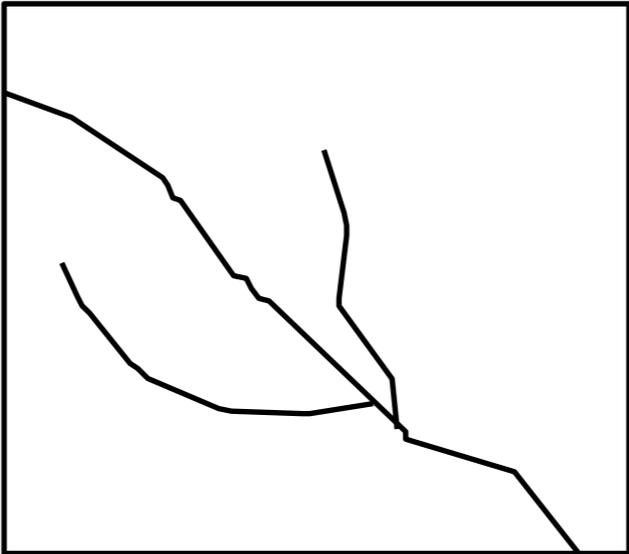
MODELO VETORIAL



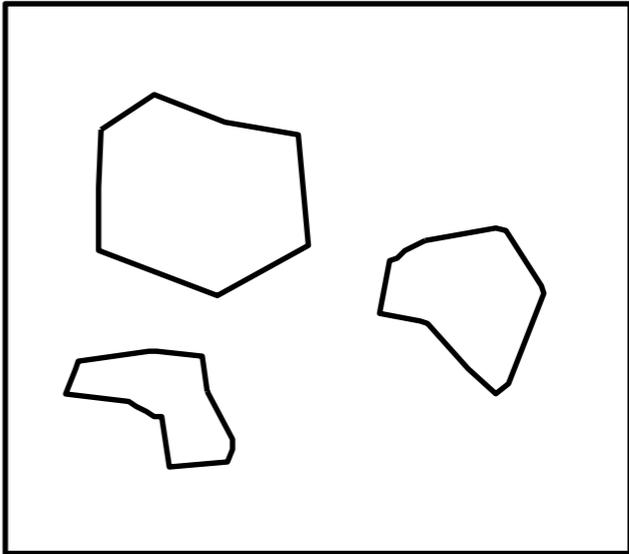
MODELO VETORIAL



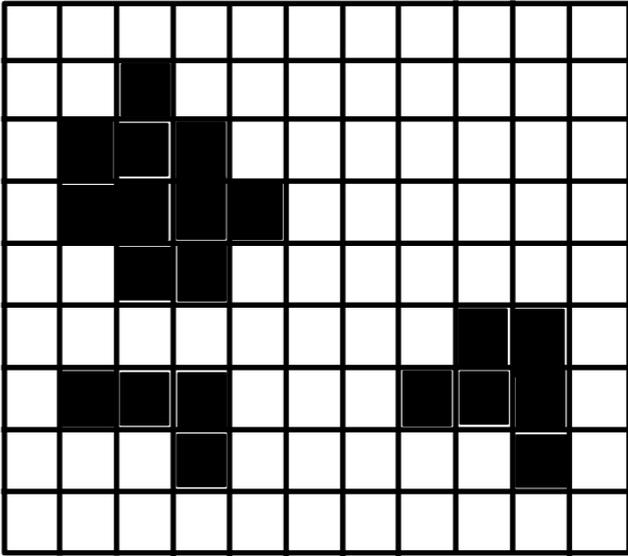
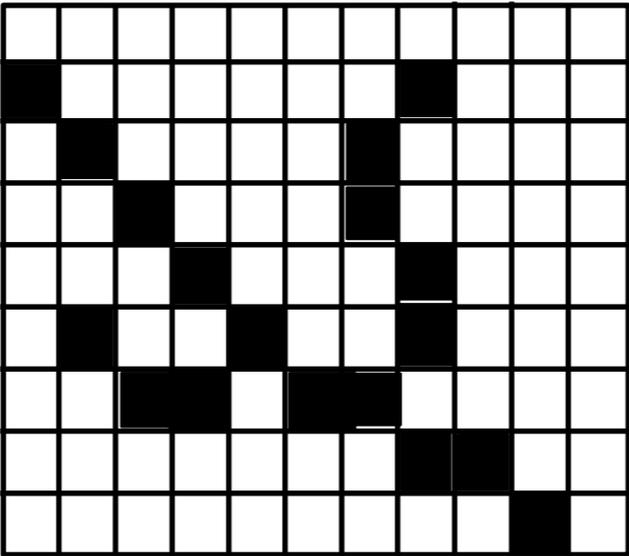
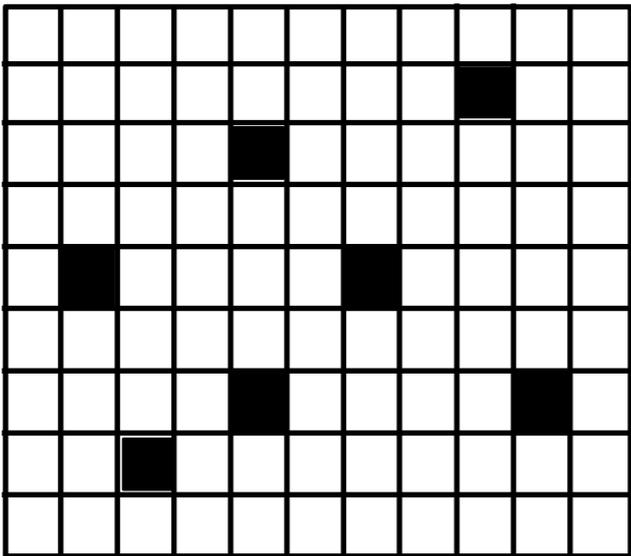
Pontos



Linhas



Polígonos

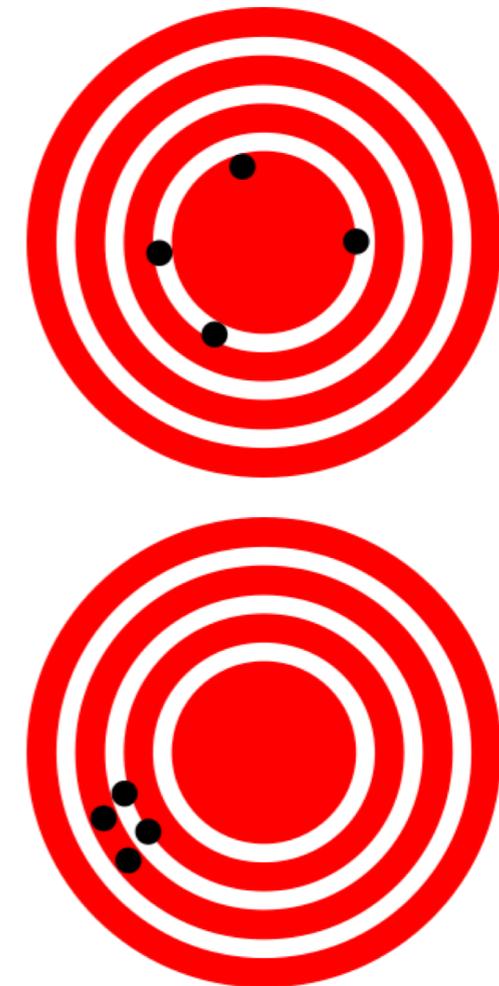
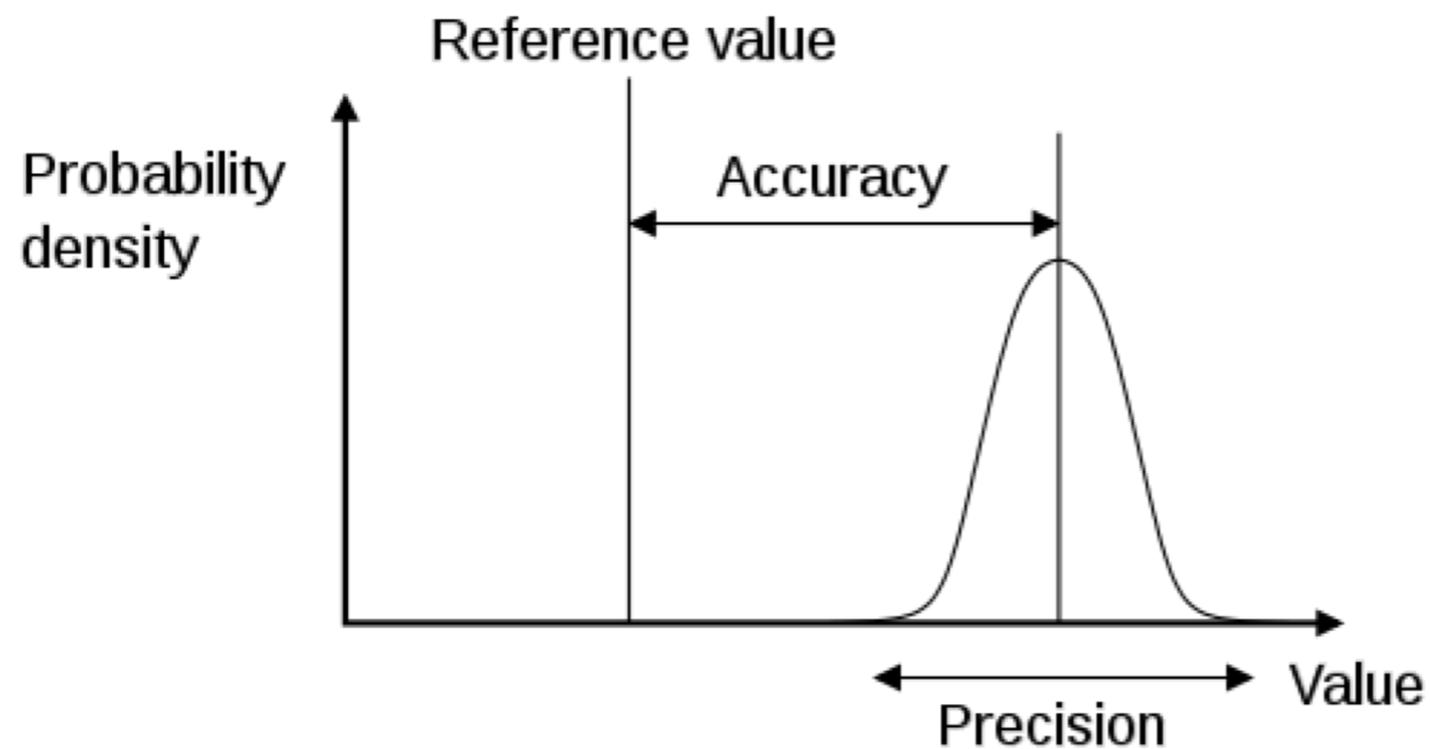


MODELO VETORIAL

- ▶ Objetos definidos por coordenadas planares x-y (cartesianas)
- ▶ Precisão das coordenadas virtualmente infinita – só depende do equipamento
 - ▶ mas a exatidão é limitada
- ▶ linha (arco): sequência de pares x,y
- ▶ nó: pontos extremos das linhas
- ▶ vértices: pontos intermediários

EXATIDÃO E PRECISÃO

- ▶ Exatidão (acurácia): quão **correta** é a medida
- ▶ Precisão (reprodutibilidade): indica a **dispersão** de um conjunto de dados

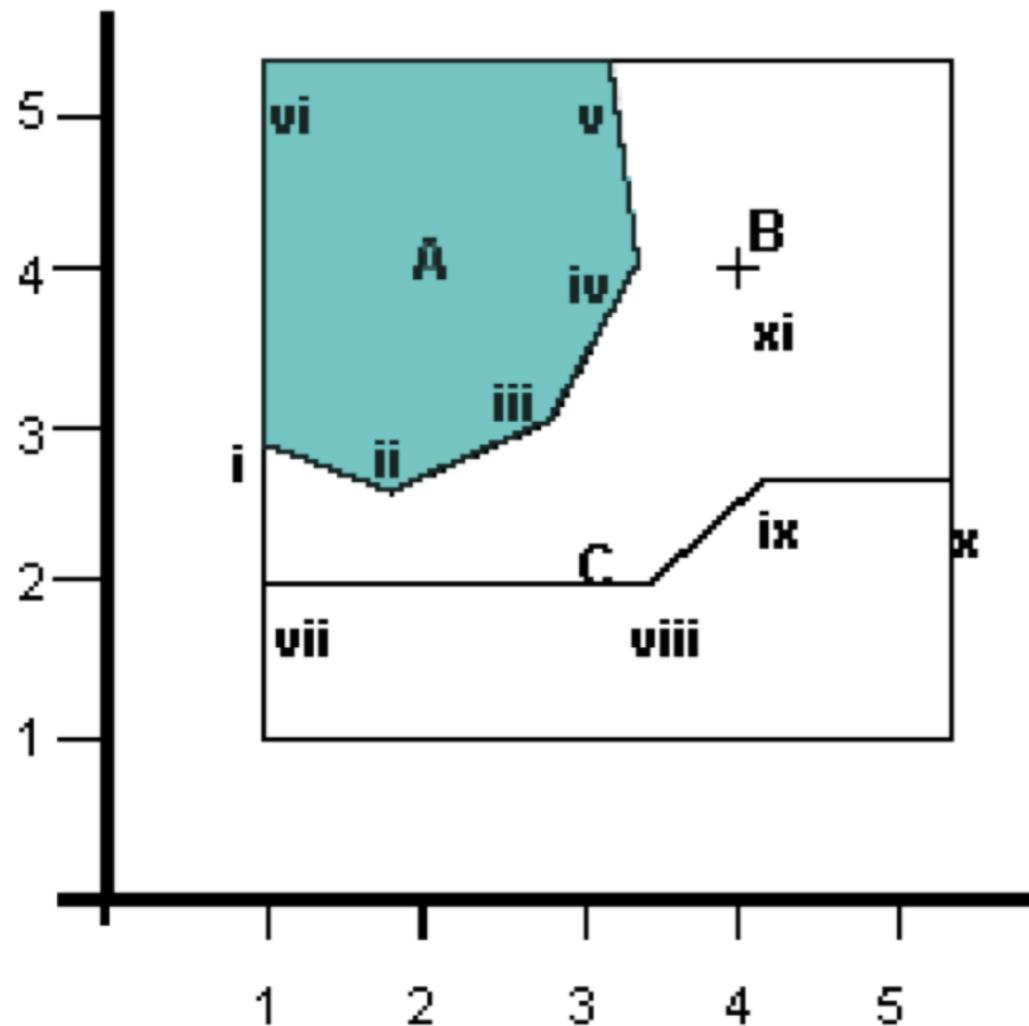


MODELO VETORIAL “SPAGHETTI”

- ▶ Ponto gravado como par de coordenadas x,y
- ▶ Linha é série de pares x,y
- ▶ Polígono = série de pares, com último igual ao primeiro (fechamento)

- ▶ simples
- ▶ fácil de manipular
- ▶ sem topologia
- ▶ duplicação de dados

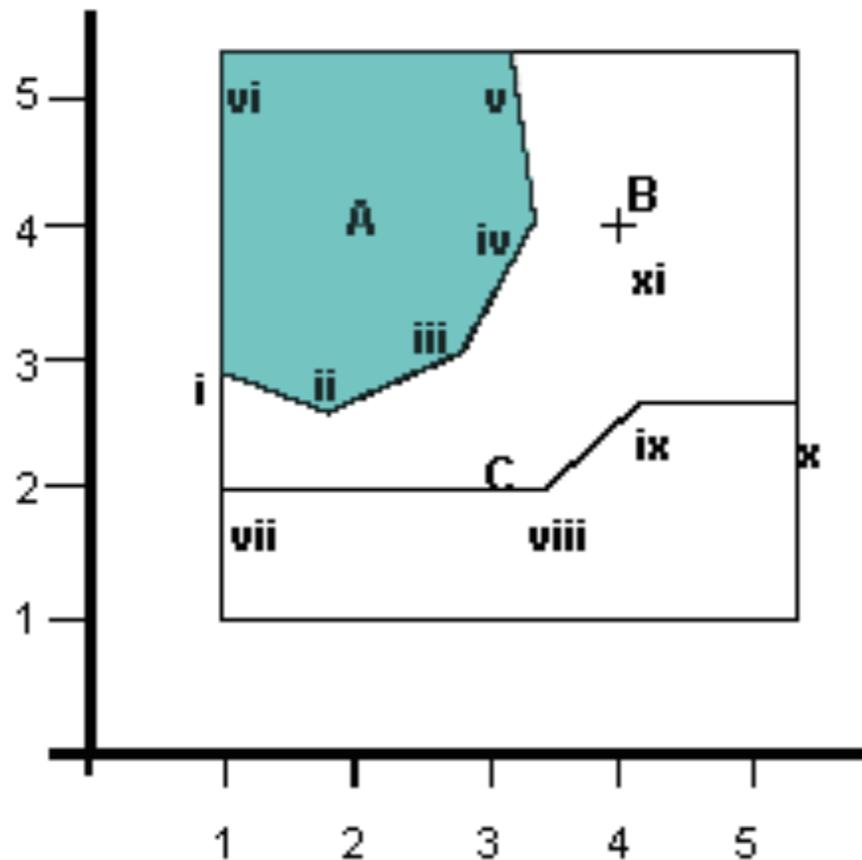
MODELO VETORIAL "SPAGHETTI"



A, 6 (identifier of polygon and number of vertex)
 1, 3 (coordinates of the first vertex)
 1.8, 2.6
 2.8, 3
 3.3, 4
 3.2, 5.2
 1, 5.2
 1, 3 (coordinates of the first vertex again)
 B, 1 (identifier of the point and number of vertex)
 4, 4
 C, 4 (identifier of the line and number of vertex)
 1, 2
 3.5, 2
 4.2, 2.7
 5.2, 2.7

MODELO VETORIAL DE DICIONÁRIO DE VÉRTICES

- ▶ modelo simples, sem duplicação
 - ▶ sem topologia



	<u>file 1</u>	
vertex	X	Y
i	1	3
ii	1.8	2.6
iii	2.8	3
iv	3.3	4
v	3.2	5.2
vi	1	5.2
vii	1	2
viii	3.5	2
ix	4.2	2.7
x	5.2	2.7
xi	4	4

file 2
 polygon A: i, ii, iii, iv, v, vi
 point B: xi
 line C: vii, viii, ix, x

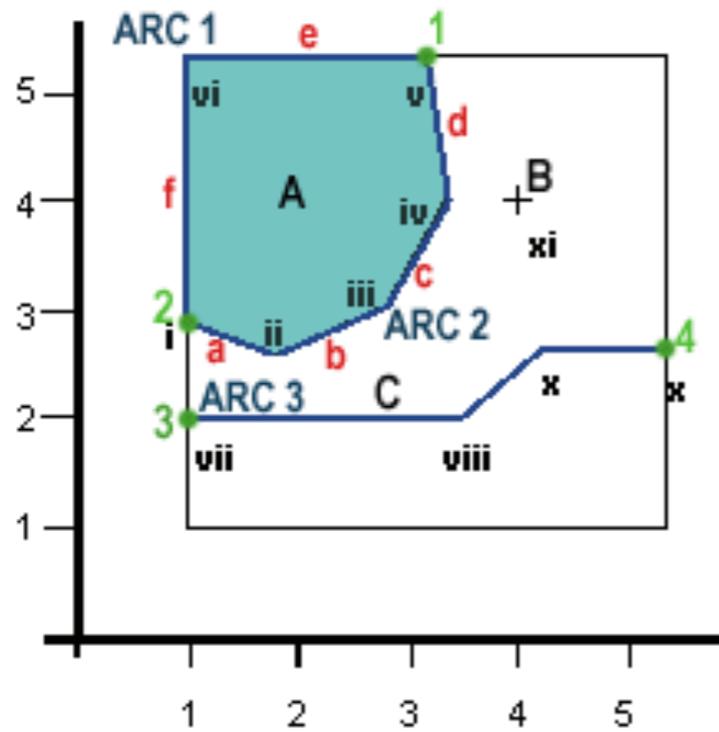
MODELO VETORIAL DE ARCO-NÓ (TOPOLÓGICO)

- ▶ Topologia: ramo da matemática que lida com propriedades geométricas
- ▶ Relações de vizinhança se mantêm depois de transformações
- ▶ Topologia é o que distingue os modelos vetoriais mais complexos

MODELO VETORIAL DE ARCO-NÓ (TOPOLÓGICO)

- ▶ Registra coordenadas x , y e códigos de relações espaciais
- ▶ Arco: linha com sentido
- ▶ Nó: ponto inicial ou final
- ▶ Vértice: ponto intermediário
- ▶ Pode responder a perguntas:
 - ▶ Quais estradas conectam com a praça central ?
 - ▶ Quais rios cruzam uma determinada área ?
 - ▶ Quais unidades geológicas fazem contato com esta ?

MODELO VETORIAL DE ARCO-NÓ (TOPOLÓGICO)



File 1. Coordinates of nodes and vertex for all the arcs

ARC	F_node	Vertex	T_node
1	3.2, 5.2	1, 5.2	1,3
2	1,3	1.8,2.6 2.8,3 3.3,4	3.2, 5.2
3	1,2	3.5,2 4.2,2.7	5.2,2.7

File 2. Arcs topology

ARC	F_node	T_node	R_poly	L_poly
1	1	2	External	A
2	2	1	A	External
3	3	4	External	External

File 3. Polygons topology

Polygon	Arcs
A	1, 2

File 4. Nodes topology

Node	Arcs
1	1,2
2	1,2
3	3
4	4
5	5

BANCOS DE DADOS

- ▶ O banco de dados ideal em SIG é aquele que maximiza as propriedades singulares de cada feição ao mesmo tempo em que minimiza o volume de dados necessários

TIPOS DE ATRIBUTOS DO BD

- ▶ **Catagóricos (nomes):**
- ▶ **nominal**
 - ▶ sem ordenamento inerente
 - ▶ tipos de uso do solo, nomes de cidades
- ▶ **ordinal**
 - ▶ ordenamento inerente
 - ▶ classe de drenagem, classe de rodovia
- ▶ normalmente codificado por números, mas não é possível realizar aritmética

TIPOS DE ATRIBUTOS DO BD

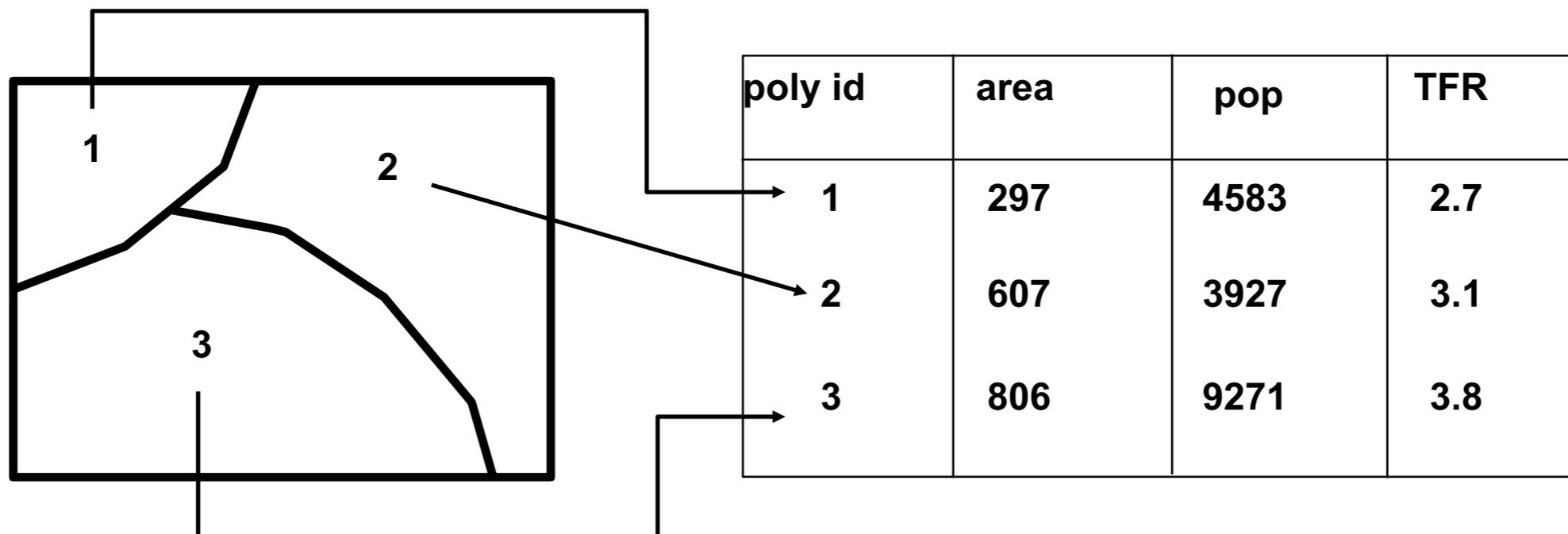
- ▶ **Numéricos**
- ▶ **intervalo**
 - ▶ sem “zero”
 - ▶ não existe “o dobro de”
- ▶ **razão**
 - ▶ existe um “zero”
 - ▶ razões fazem sentido (“o dobro de”)
- ▶ números inteiros ou decimais (ponto flutuante)

TIPOS DE BANCOS DE DADOS

- ▶ tabular ("flat file") – dados em um único arquivo
- ▶ hierárquico
- ▶ rede (network)
- ▶ relacional

BANCOS DE DADOS

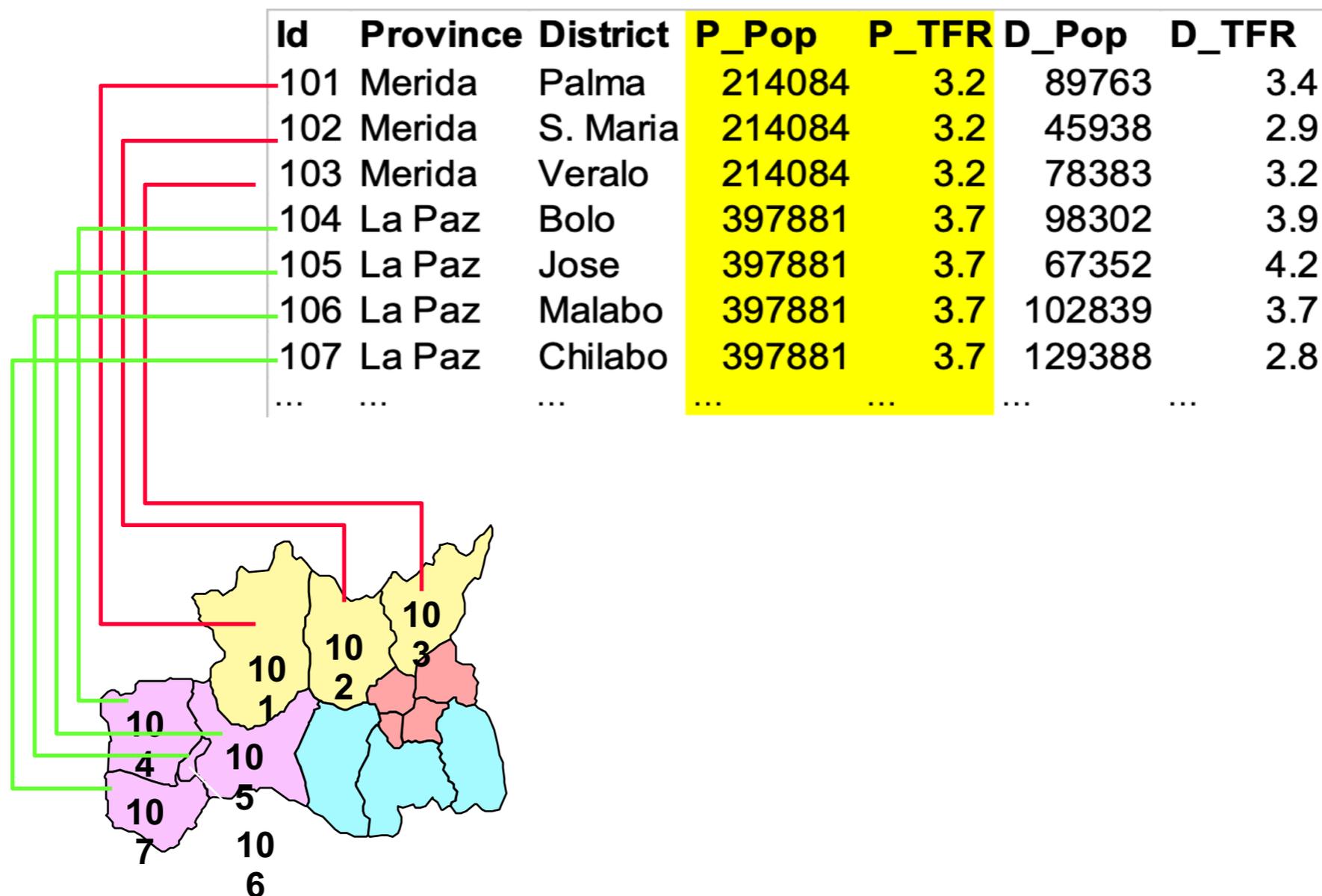
Tabela de atributos de polígonos



Da mesma forma tabelas para pontos e linhas

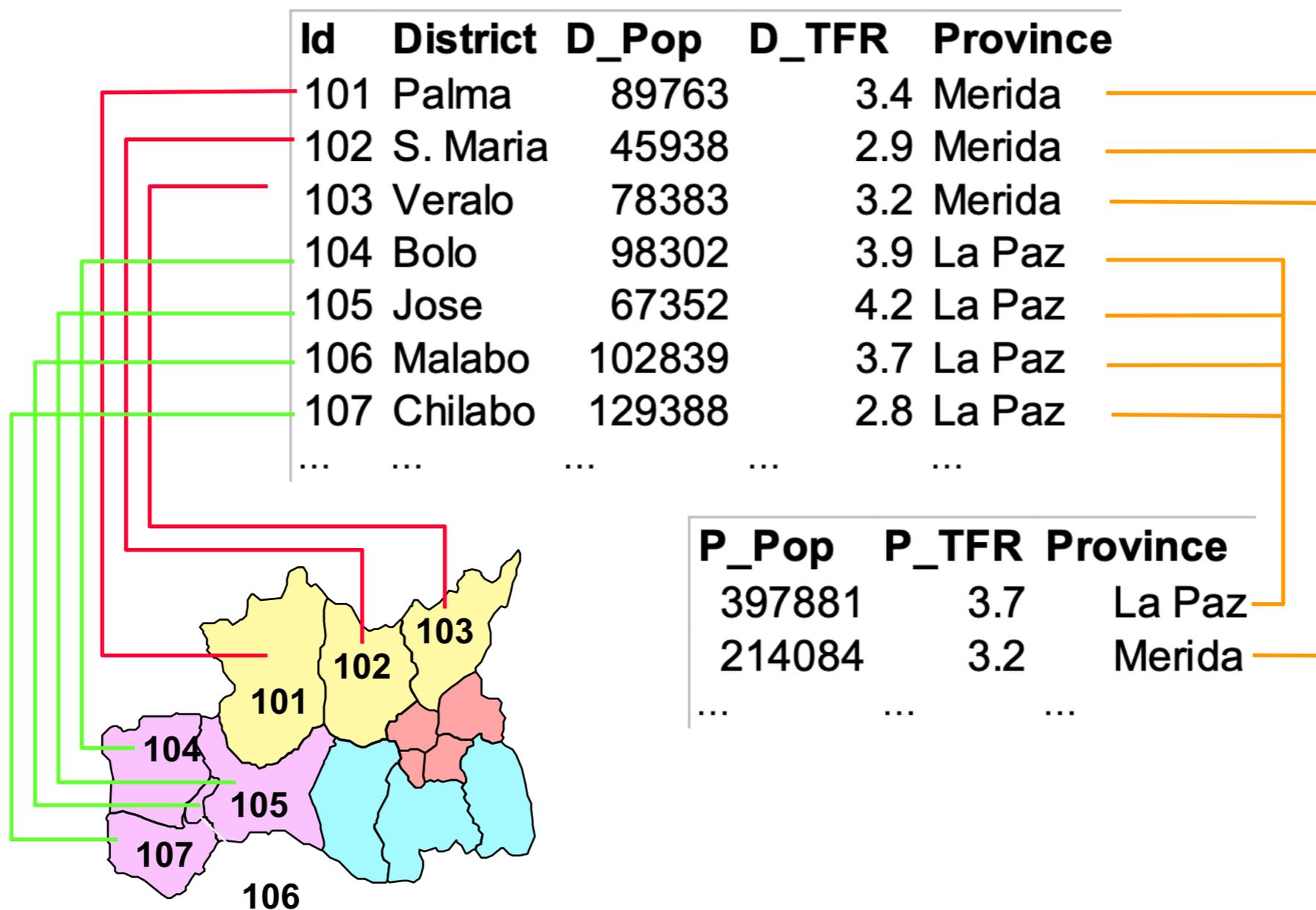
BANCOS DE DADOS

- ▶ Armazenar dados de províncias e distritos na mesma tabela é ineficiente – repetição dos dados de província em cada distrito

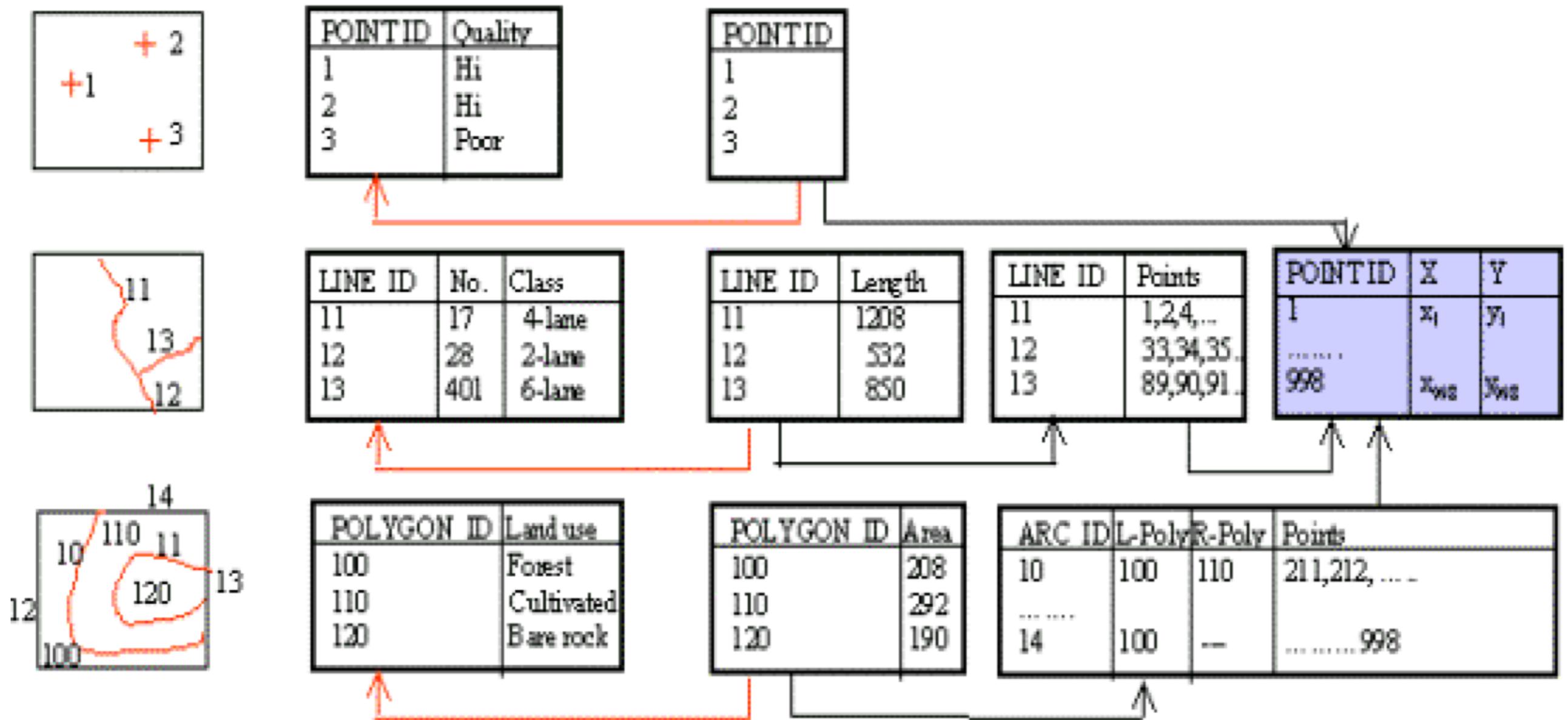


RE

- ▶ Banco de dados **relacional** permite mais eficiência



BANCOS DE DADOS RELACIONAL



GIS FILE FORMATS

- ▶ Raster, vector, TIN, etc. are generic models for representing spatial information in digital form
- ▶ GIS vendors implement these models in *file formats or structures* which may be
 - ▶ **Proprietary:** useable only with that vendor's software (e.g. ESRI coverage)
 - ▶ **Published:** specifications available for use by any vendor (e.g. ESRI shapefile)
 - ▶ **Transfer formats:** intended only for transfer of data
 - ▶ Between different vendor's systems (e.g. AutoCAD .dxf format, or SDTS)
 - ▶ Between different users of same vendors' software (e.g. ESRI's E00 format for coverages)
- ▶ One GIS vendor may be able to read another file format:
 - ▶ **By translation**, whereby format is converted externally to vendors own format
 - ▶ Usually requires user to carry out conversion prior to use of data
 - ▶ **On-the-fly**, whereby conversion is accomplished internally and "automatically"
 - ▶ No user action needed, but usually no ability to change data
 - ▶ **Natively, or transparently**, which normally implies
 - ▶ No special user action needed
 - ▶ Ability to read and write (change or edit) the data

COMMON GIS & CAD FILE FORMATS

▶ ESRI

- ▶ Coverages (vector--proprietary)
- ▶ E00 (“E-zero-zero”) for coverage exchange between ESRI users
- ▶ Shapefiles (vector--published) .shp
- ▶ Geodatabase (proprietary) .gdb
 - ▶ Based on current object-oriented software technology
- ▶ GRID (raster)

▶ Spatial Data Transfer Standard (SDTS)

- ▶ US federal standard for transfer of data
- ▶ Federal agencies legally required to conform
- ▶ embraces the philosophy of self-contained transfers, i.e. spatial data, attribute, georeferencing, data quality report, data dictionary, and other supporting metadata all included
- ▶ Not widely adopted ‘cos of competitive pressures, and complexity and perceived disutility derived from philosophy

▶ AutoCAD

- ▶ AutoCAD .DWG (native)
- ▶ AutoCAD .DXF for digital file exchange

▶ Intergraph/Bentley

- ▶ Bentley MicroStation .DGN
- ▶ Intergraph/Bentley .MGE

ESRI VECTOR FILE FORMATS: “GEORELATIONAL”

- ▶ **Shapefile:** native GIS data structure for a vector layer in *ArcView*
 - ▶ not fully topological
 - ▶ limited info about relationship of features one to another
 - ▶ draw faster
 - ▶ not as good for some fancy spatial analyses
 - ▶ is a ‘logical’ file which comprises several (at least 3) physical disk files, all of which must be present for AV to read the theme
 - ▶ layer.shp (geometric shape described by XY coords)
 - ▶ layer.shx (indices to improve performance)
 - ▶ layer.dbf (contains associated attribute data)
 - ▶ layer.sbn layer.sbx
 - ▶ not really a database, although ArcView presents files to user via relational concepts
 - ▶ openly published specs so other vendors can develop shape files and read them
- ▶ **Coverage:** native GIS data structure for a vector layer in *ArcInfo*
 - fully topological
 - better suited for large data sets
 - better suited for fancy spatial analyses
 - comprises multiple physical files (12 or so) per coverage
 - each coverage saved in a separate folder named same as the coverage
 - physical file set differs depending on type of coverage (point, line, polygon).
 - coverage folders stored in a “workspace” directory with an **info** folder for tracking
 - attribute tables stored there also
 - ARC/INFO required to make changes
 - proprietary: no published specs.
 - ▶ **E00 Export Files:** format for export of coverages to other ESRI users
 - IMPORT71 utility in ArcView Start Menu can read E00 files and convert them back to coverages
 - Must convert to *shapefile* or *AutoCAD .dxf* format to transfer to a non-ESRI GIS system

ARCGIS DATABASE ENVIRONMENT

- ▶ **Geodatabase**
 - ▶ The new term with ArcInfo 8 in 2000
 - ▶ Replacement for coverages, and support for
- ▶ **Simple features:** points, lines polygons
- ▶ **Complex features:** real world entities modeled as objects with properties, behavior, rules, & relationships
- ▶ **Personal Geodatabase**
 - ▶ Single-user editing
 - ▶ Stored as one .mdb file (but Access can't read)
- ▶ **Multiuser Geodatabase**
 - ▶ Supports versioning and long transactions
 - ▶ Uses ArcSDE 8 as middleware
 - ▶ Stores in standard db: ORACLE, MS SQL Server, Informix, Sybase, IBM DB2
- ▶ **Spatial Database Engine (SDE)**
 - ▶ ESRI “middleware” product designed to interface with industry-standard RDBMS for large scale spatial data bases
 - ▶ both attribute and spatial data is stored in the same RDBMS (such as Oracle, which supports SDE)
 - ▶ allows mass data capabilities, security and data integrity mechanisms of the RDBMS to be applied to the spatial data
 - ▶ data is grouped into:
 - ▶ sets, which share common security (e.g. all data for a city)
 - ▶ layers, similar to themes (e.g. road layer, parcel layer)
 - ▶ features, individual elements (e.g. single road)
 - ▶ advantages for large data sets include
 - ▶ layers are not tiled, so no re-assembly is required
 - ▶ features can be extracted as a complete element e.g. entire road

GEOPACKAGE

- ▶ Banco de dados em formato SQLite com extensão **.gpkg**
- ▶ raster + vetor
- ▶ 1 único arquivo
- ▶ <http://www.geopackage.org>
- ▶ <http://switchfromshapefile.org>