HOLCOMBE COUGHLIN & ASSOCIATES

CONSULTANTS IN STRUCTURAL GEOLOGY & EXPLORATION

Database and GIS Fundamentals in Exploration

Or

There is a difference between a spreadsheet and a database

Is this a database?

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Drill data			Structure					Assay (- = below detection)				
2	DDH_ID	Depth	Plunge	Azimuth	Type	α	β	Calc Dip	Calc dirn	Cu (ppm)	Au1	Au (int)	Interval (m)	Sample ID
3	DDH225	12.3	-65.96	202.68	S0	45	326	37.0	175.0					
4		15.7	-65.96	202.68	S3	56	235	45.0	234.0					
5		25.2	-65.96	204.9						3	-		2.3	225-1
6		27.5	-65.75	206.9						98	0.03		1	225-2
7		28.5	-65.75	206.9	S0	32	125	36.0	146.0	120	1.1		1.3	225-3
8		29.8	-63.51	204.88	S0	57	225	37.0	218.0	3	1.5	2.87	1.4	225-4
9		31.2	-63.51	204.88	Sch	40	285	61.7	112.8	332	8.4		2.3	225-5
10		33.5	-63.51	204.88	QCV	42	275	55.9	106.6	302	3.3		1.3	225-6
11		34.8	-62.86	205.51	S0	33	300	73.5	120.8	134	-		2.5	225-7
12		37.3	-62.55	206.1	S3	52	0	66.0	170.0	10	-		1.9	225-8
13		39.2	-62.55	206.1	QCV	53	10	64.7	176.6	25	-		1	225-9
14		58.5	-62.55	206.1	S0	55	15	62.4	179.6	32	0.05		0.2	225-10
15		58.7	-62.55	206.1	Sch	15	0	77.0	350.0	10	-		0.1	225-11
16		58.7	-62.55	206.1	S0	25	5	87.1	354.5					
17		58.8	-62.19	206.52	F	30	290	72.4	111.4	34	0.01		0.3	225-12
18		59.1	-61.97	206.28	Sch	46	340	70.8	155.4	52	0.009		1.2	225-13
19		60.3	-61.79	206.2	S3	48	80	53.0	225.6	136	0.02		0.3	225-14
20		60.3	-61.79	206.2	S0	45	340	71.8	155.3					
21		60.3	-61.79	206.2	Sch	38	280	61.4	107.8					
22		60.6	-61.79	206.2	S3	55	350	62.7	163.6	136	0.009		0.9	225-15
23		60.6	-61.79	206.2	S0	32	345	85.2	157.3					
24		61.5	-61.79	206.2	S0	35	285	66.0	110.0	14	-		0.7	225-16
25	DDH23	23.5	-61.79	206.2	Sch	45	395	69.4	195.7					
26		45.6	-61.55	205.52	S0	43	330	72.3	147.4					
27		62.2	-61.55	205.52	Sch	45	330	70.3	147.9	25	0.01		1	23-1
28		62.2	-61.4	206.68	S0	46	0	72.0	170.0					
29		82	-61.4	206.68	S3	37	0	81.0	170.0	10	0.02		0.4	23-2
30		82	-61.4	206.68	F	32	355	85.9	165.8					
31		82.4	-61.14	206.5	S0	37	345	80.3	157.9	34	0.009		0.3	23-3
32		82.4	-61.14	206.5	Sch	55	110	35.4	238.6					
33		82.4	-61.14	206.5	S3	35	290	68.0	113.9					
34		82.7	-60.8	205.83	S0	36	0	82.0	170.0	56	0.009		0.5	23-4
35		82.7	-60.8	205.83	Sch	48	290	61.5	126.3					
36		83.2	-60.8	205.83	S3	45	40	74.4	200.2	102	3.29	1.585	0.6	23-5
37		83.2	-60.8	205.83	S0	45	50	71.4	206.9					
38		83.2	-60	206.25	QCV	30	40	88.3	205.8					
39		83.8	-60	206.25	S0	38	345	87.1	161.2	23	0.03		1	23-6
40		83.8	-59.9	206.2	Sch	47	280	58.5	121.1					
41		83.8	-59.9	206.2	S0	37	340	87.4	157.1					
42		112	-59.9	206.2	QCV	40	280	63.8	115.8	15	0.009		0.2	23-7
43		112	-59.9	206.2	F	35	90	62.4	240.6					

No!

- At least two 1-to-1 sets of data in a single flat spreadsheet
  - Assay
  - Structure
    - See ‘key’field comment below
- Split cells
  - bad enough in headings
  - disastrous in body of table
- Non-numeric data in numeric column
  - (- symbol)
- Alphanumeric numbers in non-sortable format
  - 225-1,….,225-11,….,23-1,…
  - DDH23,…,DDH225,…
- No distinct key field for related data
  - Not critical, but desirable

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Drill data			Structure					Assay (- = below detection)				
2	DDH_ID	Depth	Plunge	Azimuth	Type	α	β	Calc Dip	Calc dirn	Cu (ppm)	Au1	Au (int)	Interval (m)	Sample ID
3	DDH225	12.3	-65.96	202.68	S0	45	326	37.0	175.0					
4		15.7	-65.96	202.68	S3	56	235	45.0	234.0					
5		25.2	-65.96	204.9						3	-		2.3	225-1
6		27.5	-65.75	206.9						98	0.03		1	225-2
7		28.5	-65.75	206.9	S0	32	125	36.0	146.0	120	1.1		1.3	225-3
8		29.8	-63.51	204.88	S0	57	225	37.0	218.0	3	1.5	2.87	1.4	225-4
9		31.2	-63.51	204.88	Sch	40	285	61.7	112.8	332	8.4		2.3	225-5
10		33.5	-63.51	204.88	QCV	42	275	55.9	106.6	302	3.3		1.3	225-6
11		34.8	-62.86	205.51	S0	33	300	73.5	120.8	134	-		2.5	225-7
12		37.3	-62.55	206.1	S3	52	0	66.0	170.0	10	-		1.9	225-8
13		39.2	-62.55	206.1	QCV	53	10	64.7	176.6	25	-		1	225-9
14		58.5	-62.55	206.1	S0	55	15	62.4	179.6	32	0.05		0.2	225-10
15		58.7	-62.55	206.1	Sch	15	0	77.0	350.0	10	-		0.1	225-11
16		58.7	-62.55	206.1	S0	25	5	87.1	354.5					
17		58.8	-62.19	206.52	F	30	290	72.4	111.4	34	0.01		0.3	225-12
18		59.1	-61.97	206.28	Sch	46	340	70.8	155.4	52	0.009		1.2	225-13
19		60.3	-61.79	206.2	S3	48	80	53.0	225.6	136	0.02		0.3	225-14
20		60.3	-61.79	206.2	S0	45	340	71.8	155.3					
21		60.3	-61.79	206.2	Sch	38	280	61.4	107.8					
22		60.6	-61.79	206.2	S3	55	350	62.7	163.6	136	0.009		0.9	225-15
23		60.6	-61.79	206.2	S0	32	345	85.2	157.3					
24		61.5	-61.79	206.2	S0	35	285	66.0	110.0	14	-		0.7	225-16
25	DDH23	23.5	-61.79	206.2	Sch	45	395	69.4	195.7					
26		45.6	-61.55	205.52	S0	43	330	72.3	147.4					
27		62.2	-61.55	205.52	Sch	45	330	70.3	147.9	25	0.01		1	23-1
28		62.2	-61.4	206.68	S0	46	0	72.0	170.0					
29		82	-61.4	206.68	S3	37	0	81.0	170.0	10	0.02		0.4	23-2
30		82	-61.4	206.68	F	32	355	85.9	165.8					
31		82.4	-61.14	206.5	S0	37	345	80.3	157.9	34	0.009		0.3	23-3
32		82.4	-61.14	206.5	Sch	55	110	35.4	238.6					
33		82.4	-61.14	206.5	S3	35	290	68.0	113.9					
34		82.7	-60.8	205.83	S0	36	0	82.0	170.0	56	0.009		0.5	23-4
35		82.7	-60.8	205.83	Sch	48	290	61.5	126.3					
36		83.2	-60.8	205.83	S3	45	40	74.4	200.2	102	3.29	1.585	0.6	23-5
37		83.2	-60.8	205.83	S0	45	50	71.4	206.9					
38		83.2	-60	206.25	QCV	30	40	88.3	205.8					
39		83.8	-60	206.25	S0	38	345	87.1	161.2	23	0.03		1	23-6
40		83.8	-59.9	206.2	Sch	47	280	58.5	121.1					
41		83.8	-59.9	206.2	S0	37	340	87.4	157.1					
42		112	-59.9	206.2	QCV	40	280	63.8	115.8	15	0.009		0.2	23-7
43		112	-59.9	206.2	F	35	90	62.4	240.6					

Excel spreadsheet databases

- ‘Flat’ databases - Can only show 1:1 relationships
- Therefore need to include fields for all possibilities
- Thus get loads of empty space

E.g.:

All this empty space

(increases the file size – an issue for very large databases

All these fields are needed just in case there are 3 rocktypes, joints, or samples at the locn

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Field Num	Easting	Northing	AMG Zone	Location	Rocktype1	Rocktype2	Rocktype3	Bedding Dip	Bedding Dirn	Joint1 Dip	Joint1 dirn	Joint2 Dip	Joint2 dirn	Joint3 Dip	Joint3 dirn	Sample 1	Sample 2	Comment
2	1	336456	7769234	56	Lynd Hwy km 31.5	sandstone	rhyolite dyke		23	123	45	6o					RH009		
3	2	336336	7769434		Lynd Hwy km 42.4	sandstone			-	?							JJ001		weathered
4	2a	336466	7768124		Lynd Hwy km 51.7	siltstone	minor ss	minor cgl	56	234							RH010		
5	3	336456	7769134		Peach Rd 2.1km S of Lynd Hwy	sltst	ss												
6	4	336356	7769221		Peach Rd 500m S of Lynd Hwy	ss			78	342							RH234		

Do not mix numbers and characters in locns (can’t sort data easily)

Remember somewhere to indicate datums and UTM zones (else may not be able to plot on a digital map)

Be consistent with spelling within fields: else can’t query efficiently

Don’t put non-numeric characters into numeric fields: interferes with sorting and processing

Use sample numbers that identify the owner and are sortable

What is a Database?

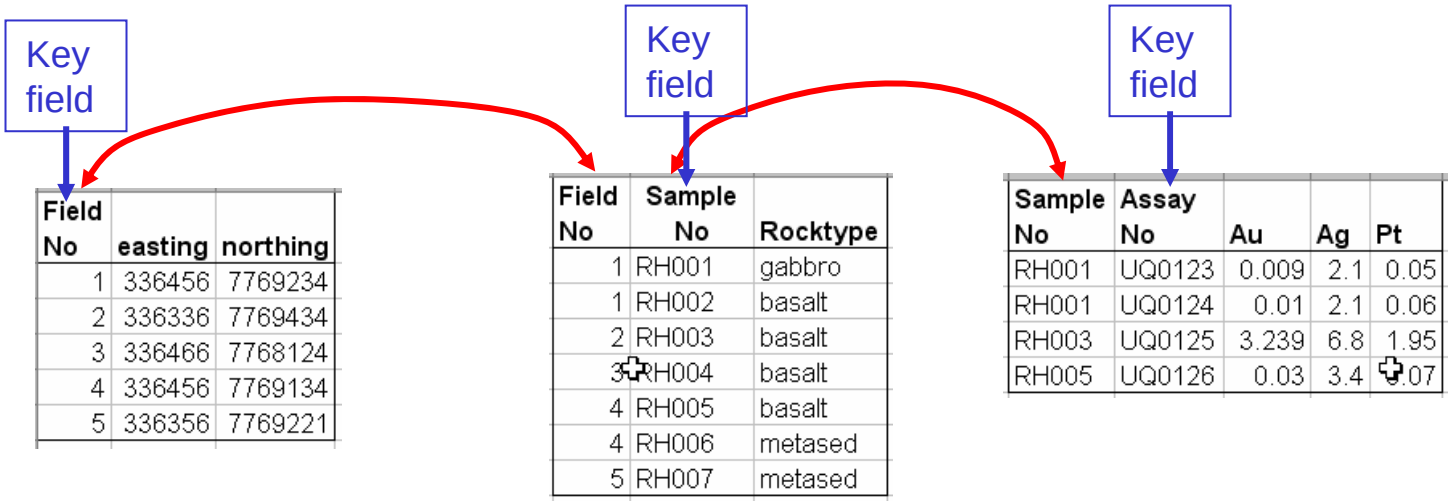
- **Consist of:**
  - One or more Tables of basic data
    - spreadsheet-like Fields (columns) and Records (rows)
    - structured and populated following database ‘rules’
    - e.g. Relational Databases:
      - each Record can be uniquely identified within any one Table
      - each Field contains the same Type of data (numeric, alphanumeric, date, object, etc)
      - Tables containing at least one of the same data columns can be linked to one another as if they were a single large table
        - » Uses Relationship ‘rules’ between the linked Tables
  - Queries: structured views of of the data, using selected fields chosen from one or more Tables, linked together using the relationships between the overlapping fields

Relational databases

- E.g. Microsoft Access, Oracle, etc
- Multiple tables (spreadsheets) for every 1-to-many relationship
- Every Table has one or more Key field(s)
  - Unique value (number or characters)
  - May combine more than one field to form the Key
- Tables can be linked by their Key values in Queries

Relational Database

- Allows new temporary tables (Queries) to be formed by linking separate 1-1 tables using their Key fields
- Tables:



Field No	easting	northing
1	336456	7769234
2	336336	7769434
3	336466	7768124
4	336456	7769134
5	336356	7769221

Field No	Sample No	Rocktype
1	RH001	gabbro
1	RH002	basalt
2	RH003	basalt
3	RH004	basalt
4	RH005	basalt
4	RH006	metased
5	RH007	metased

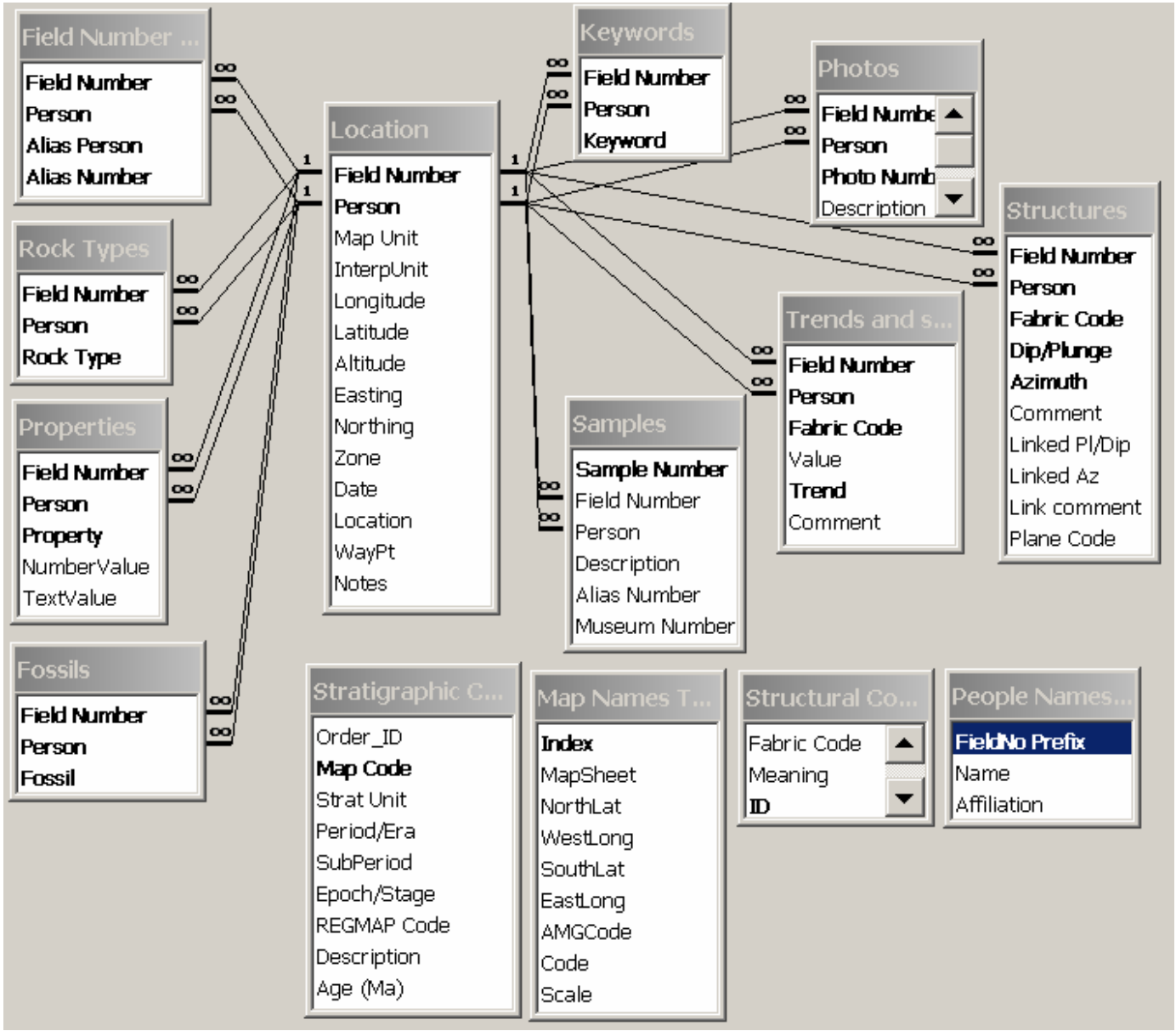
Sample No	Assay No	Au	Ag	Pt
RH001	UQ0123	0.009	2.1	0.05
RH001	UQ0124	0.01	2.1	0.06
RH003	UQ0125	3.239	6.8	1.95
RH005	UQ0126	0.03	3.4	3.07

- Query 'Table':

Rocktype	Au	easting	northing
gabbro	0.009	336456	7769234
gabbro	0.01	336456	7769234
basalt	3.239	336466	7768124
basalt	0.03	336356	7769221

  - A temporary joining of fields from the three tables

Relational Database Structure



Database design

- Poor design, even for a simple database, can lead to unwarranted cleaning-up at a later stage
- Think carefully about how you are going to Query your database - after all that's why you are doing it
- Only include those fields that will be involved in subsequent queries. The more fields you have, the more time-consuming it is to enter data.
  - in general, don't try to produce all-inclusive databases
  - databases tend to have a life that is limited to the project for which they were designed and hence extra effort may be wasted

Geographic Information Systems (GIS)

- GIS = Spatial relational databases
  - i.e exactly the same as any other relational database (RDBS) except that at least some of the Tables consist of entities with a geographic location
    - Data can be displayed as a map as well as a spreadsheet-like table
  - Separate Tables can be overlayed in map view as if they were a single map
    - This is the heart of a digital map system
  - Queries can link fields from different tables (as in any RDBS)
    - but can also select data on geographic criteria
      - such as 'Show only the data where map objects overlap'
  - Spatial Analysis of data is possible
    - Thematic mapping
      - Property distributions highlighted by colours, new symbols, contouring, etc
    - Spatial numerical analysis
      - spatial graphing

- Commonly used GIS software:
  - ArcInfo; ArcView; MapInfo

GIS - Spatial Relational Database

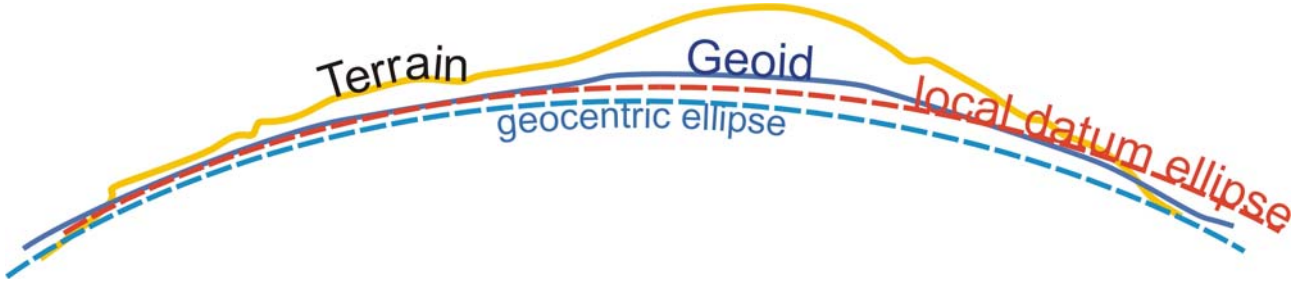
- E.g. MapInfo, ArcView, ArcInfo
- In addition to normal fields, Tables can have an extra field (commonly hidden) that contains geographic information about Geographic objects in the Table
  - *Point, Line, Arc, Polyline, Polygon, Region, etc*

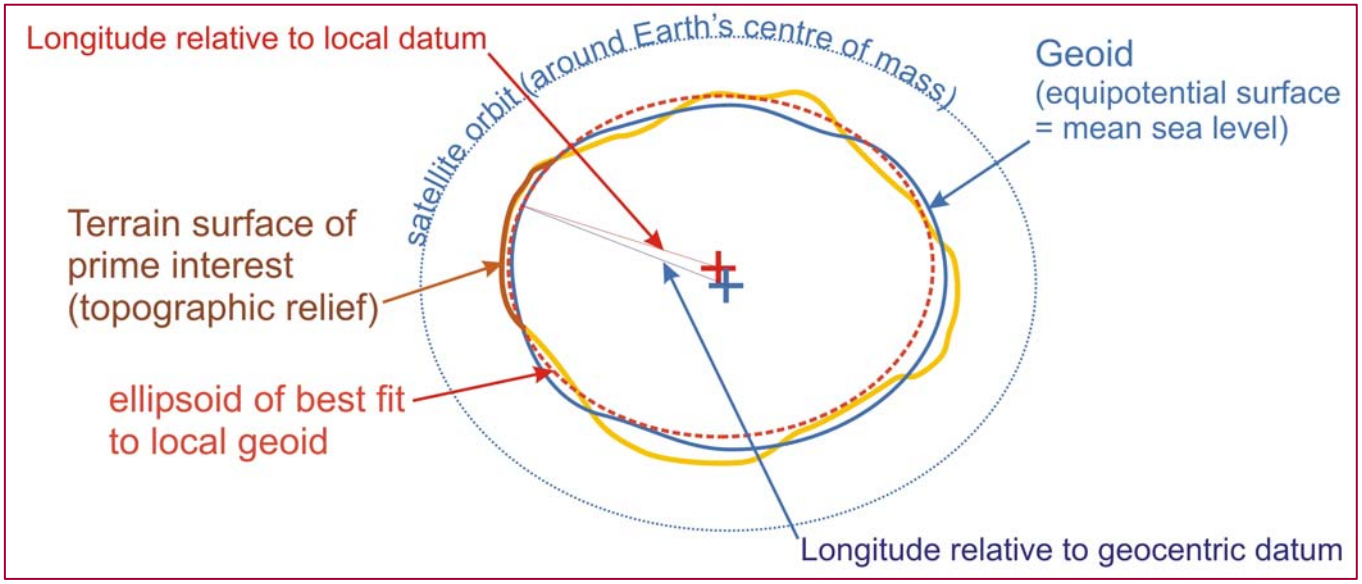
- Geographic Information:
  - Point location or Centroid location (if a polygon object)
  - Perimeter
  - Area
  - Object contained on the left/right side of line
  - Etc

- Location information is dependent on Datums and Projections

Map Datums

Earth is neither a sphere or a true ellipsoid





***Note: the latitude and longitude of a point on the earth depends on the datum used

WGS84-geoid

- WGS84-geoid (1996)
  - NB: Geoid definitions vary with time due to:
    - plate motions
    - increasing precision of the grid and measurements used to define it

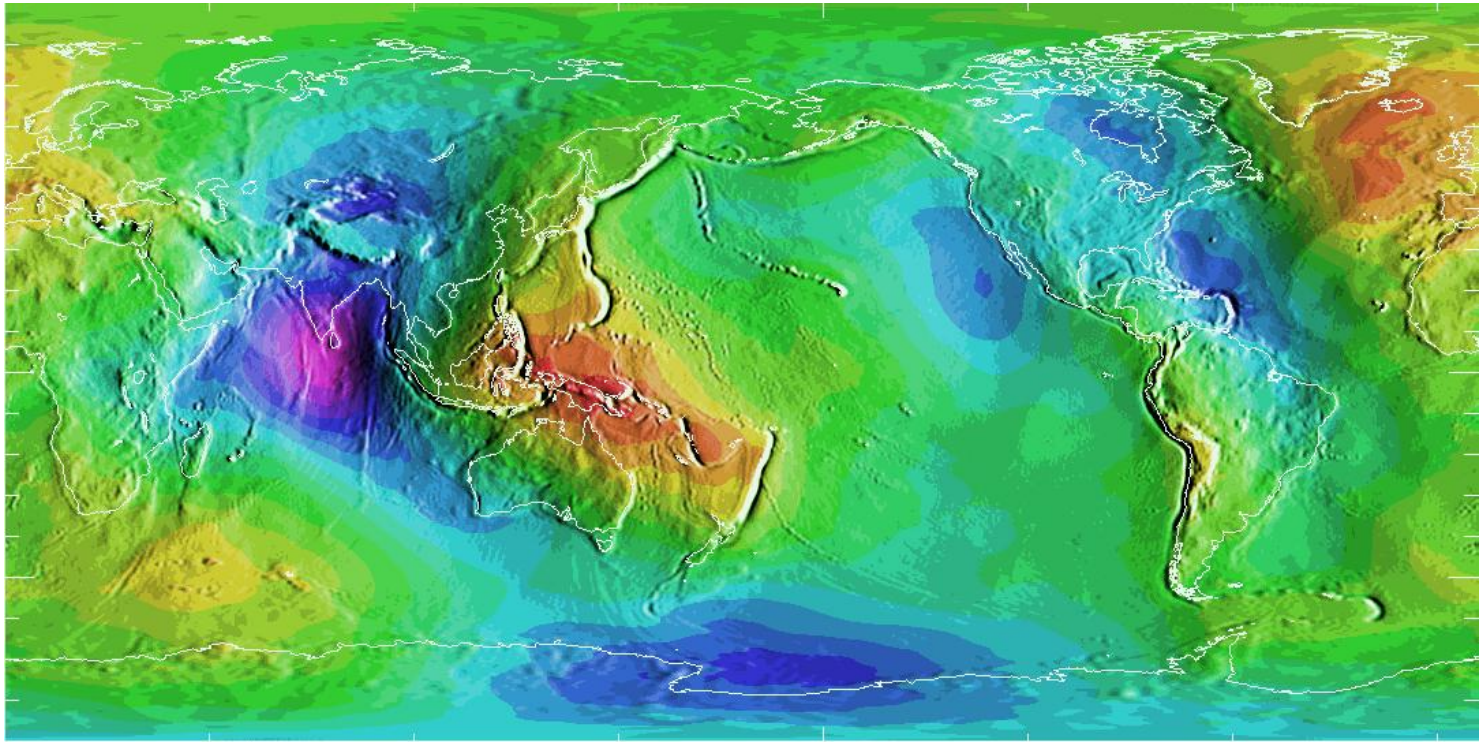
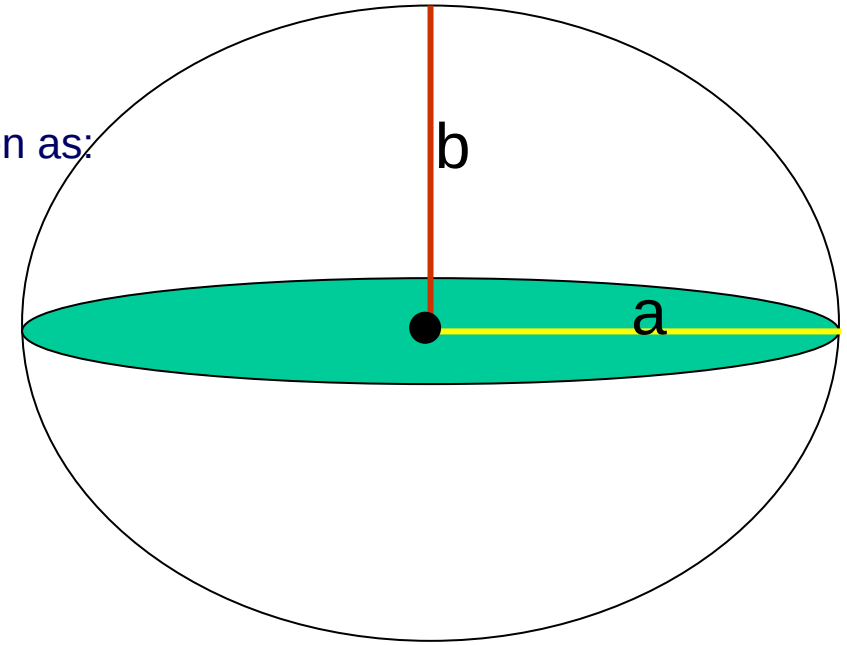


image source: National Geodetic Survey: http://www.ngs.noaa.gov/images/ngs/jpeg-geo/ww15mgh.jpg

Datum Ellipsoids

- ‘Datum’, as generally used in GIS applications, refers to both the location of the centre of the fitted ellipsoid relative to the Earth (datum) and to the shape of the ellipsoid
- Ellipsoid (spheroid) parameters:
  - a = semi-major axis (equator)
  - b = semi-minor axis (polar)
  - f= flattening = (a-b)/a

  (as f is very small it is generally given as:
  1/f = inverse flattening = a/(a-b)



- Example: WGS84 ellipsoid
  - a = 6378137.0 metres
  - b = 6356752.3142 metres
  - 1/f = 298.257223563

Global geocentric datum

- Based on centre-of-mass of Earth
  - but centre-of-mass varies with Plate motion
  - therefore is date-dependent
- **ITRF**
  - centre-of-mass datum computed annually
  - uses GRS80 ellipsoid shape
    - (Inverse flattening = 298.257222101)
- **WGS84**
  - centre-of-mass datum defined for GPS by NIMA
    - (US National Imagery and Mapping Agency)
  - slight difference in ellipsoid flattening to GRS80
    - (Inverse flattening = 298.257223563)
  - centre revised periodically
    - last updated 2004 (valid until 2010)
  - now close to ITRF

Note that WGS84 uses the zero meridian as defined by the Bureau International de l'Heure in Paris, based on a mean of compiled star observations from different countries.

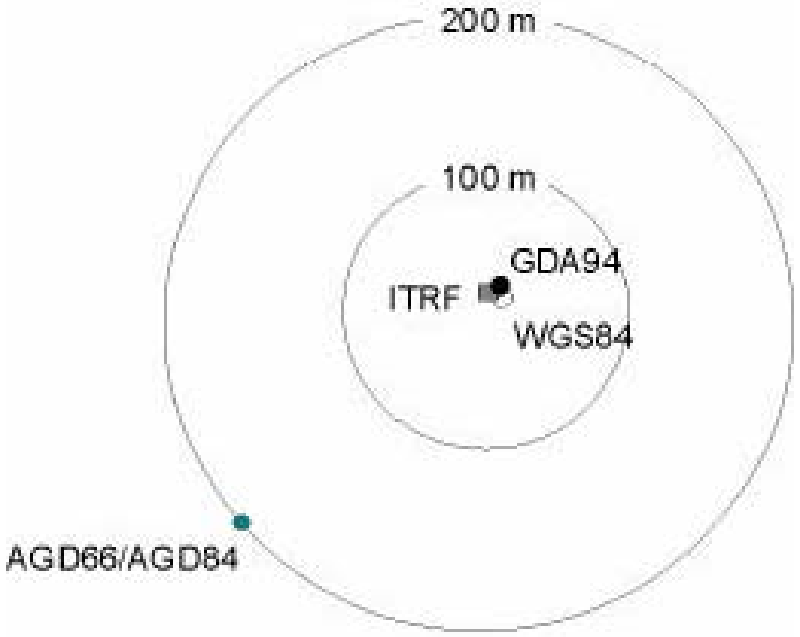
The WGS84 zero meridian is 102.5 metres east of the Prime Meridian that passes through Greenwich Observatory!

E.g Australian and International Datums

- Australian
  - Older imperial map datum
    - Clarke ellipsoid
  - ‘Recent’ metric map datums
    - AGD66
    - AGD84
  - Modern metric map datum
    - GDA94
      - based on ITRF1992
      - fixed on 1/1/1994
- International
  - >50 datums in local use
  - e.g. USA
    - NAD27

    South America
    - SAD69

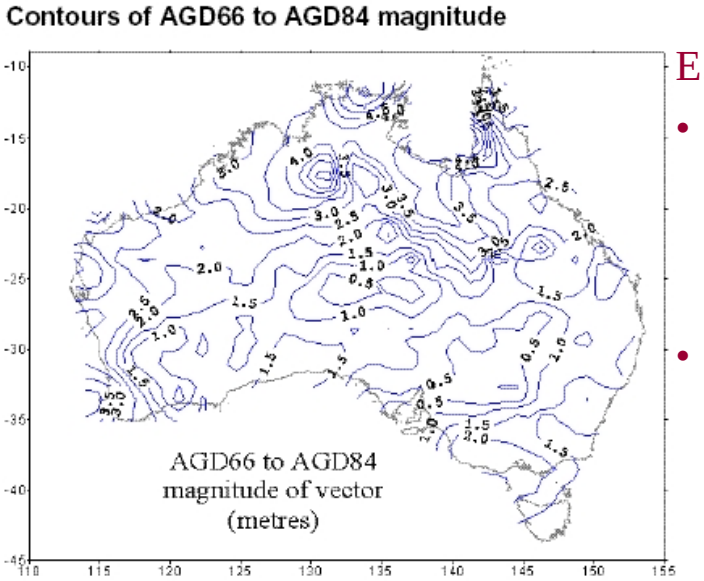
GDA94 ≅ WGS84



Latitude/longitude are not unique

Latitude/longitude values of a point on the Earth depend on the datum used to define latitude/longitude

The datum always needs to be specified



E.g:

- AGD66 to AGD84 difference: 2-3 metres
  - important to geophysics
- AGD84 to GDA94 difference: ~200 metres
  - important to everyone

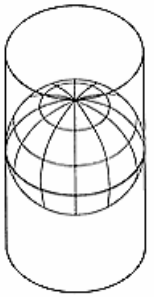
Projections

- The algorithm used to project map data that has been projected from the terrain onto the datum ellipsoid onto a 2D flat surface
  - e.g polar stereographic projections
    - e.g. as used for continental wander path reconstructions

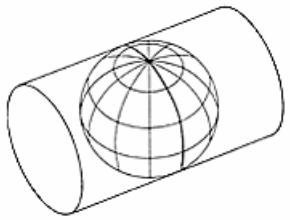
- Most common map projection:
  - Universal transverse mercator (UTM)
- Others
  - conic, polyconic, gauss-kruger (similar to UTM), etc...

Mercator projection

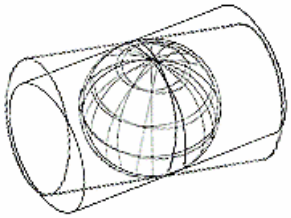
- Mercator Projection
  - Cylinder tangent to equator



- Transverse Mercator Projection
  - Cylinder tangent to line of longitude

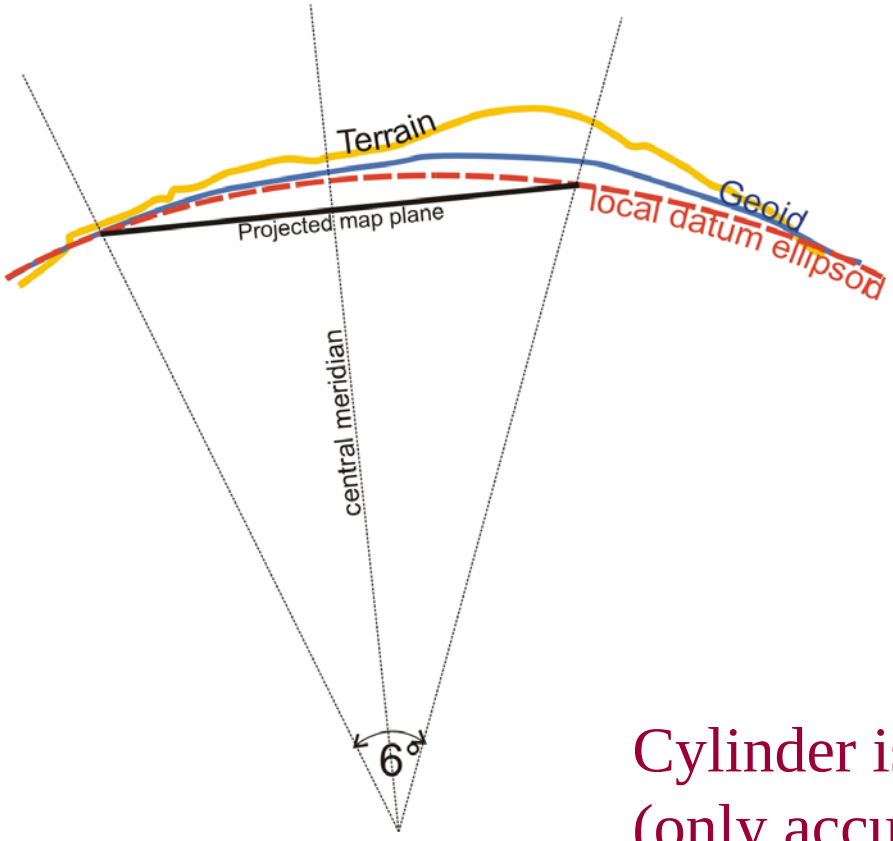


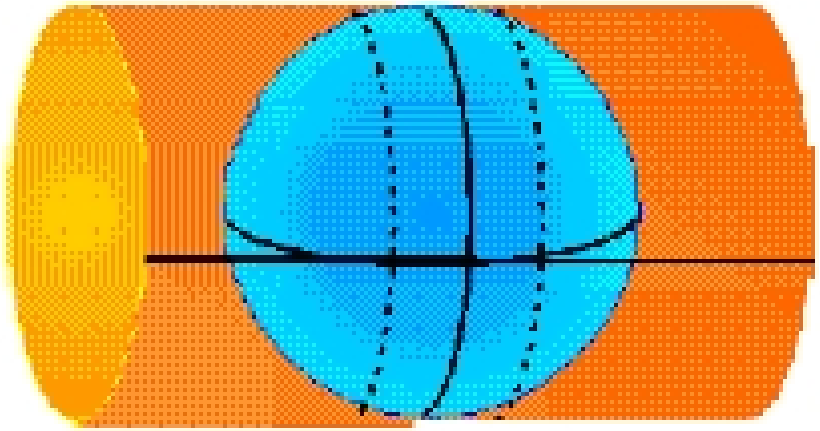
- Universal Transverse Mercator projections
  - standard set of 120 projections around the globe
    - 60 x 6º zones; separate definition of N & S hemisphere



UTM projection

- Projection onto a cylinder tangent to a line of longitude ('Central Meridian' and touching the equator at a point





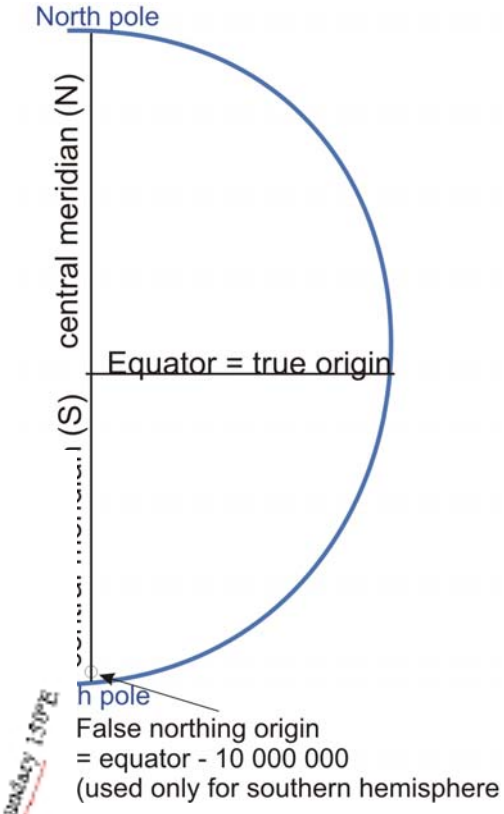
Cylinder is unrolled to give 2D map
(only accurate close to the Central Meridian)

UTM zones

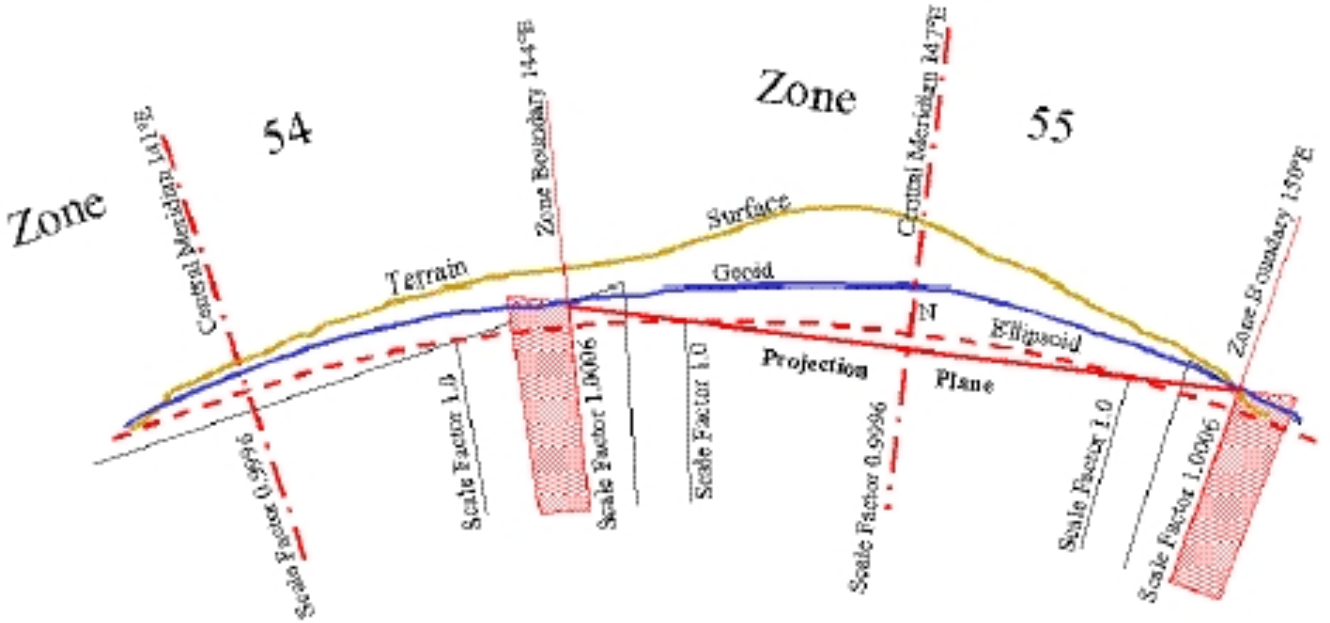
– 60 northern and 60 southern zones

Definition of UTM zones

Longitude of initial central meridian (Zone 1)	177° W
Zone width	6°
Central Scale Factor	0.9996
False easting	500 000 m
False Northing (S hemisphere)	10 000 000 m



Cross section through two UTM zones



Coordinate labels in databases

- Coordinate labels on spreadsheet or database columns should indicate the projection AND the datum
  - Assume that at some point that the data will be used by someone else
  - In the example shown at top right it is a reasonable guess that the projection is UTM – but what is the datum?
    - local alternatives are WGS84, Corrego Alegre, or SAD69
  - The example at lower right at least indicates that, in this case, the UTM guess is correct, and the datum is Corrego Alegre

PNT	X	Y	Lithotype
01	298843	8097407	2
02	298871	8097485	2
03	298909	8097512	2
20	298969	8097808	4
06	298987	8098281	2
07	298995	8098173	2
08	299000	8098096	2

PNT	Xutm_COA	Yutm_COA	Lithotype
01	298843	8097407	2
02	298871	8097485	2
03	298909	8097512	2
20	298969	8097808	4
06	298987	8098281	2
07	298995	8098173	2
08	299000	8098096	2

The UTM zone should also be referenced somewhere