Observatorio de Evaluación de la Información Geoespacial

Lic. Horacio Castellaro – Lic. Lautaro Simontacchi

Instituto Geográfico Nacional

Universidad Nacional de La Plata

IDERA

Infraestructura de Datos Espaciales de la República Argentina

IDERA
Infraestructura de Datos Espaciales
de la República Argentina

OBJETIVOS

- Establecer definiciones relativas a la calidad de productos geoespaciales
- Describir procedimiento para determinar la calidad
- Realizar análisis prototípicos de datos publicados en IDERA
- Definir como debe ser reportada esta información (metadatos)

¿Para qué conocer la calidad?

La información sobre la calidad de datos geográficos permite a un productor validar qué tan bien **cumple** un conjunto de datos los criterios previstos en las **especificaciones de producto** y ayuda a un usuario de datos a determinar la capacidad del producto para **satisfacer los requisitos de su aplicación** en particular.

Fuentes de errores

Proceso	Motivo
Modelización conceptual	Errores en el modelo conceptual
Recolección de datos	Error en los trabajos de campo
	Error en las fuentes de información utilizadas
Captura de datos	Inexactitud en la digitalización
	Inexactitud inherente a los elementos geográficos
Almacenamiento	Insuficiente precisión numérica y/o espacial
	Errores de procesamiento
Manipulación	Intervalos de clase inapropiados
	Errores de superposición
	Propagación de errores
	Errores en la operación de coordenadas
Salidas cartográficas	Inexactitud de escala
	Inexactitud del dispositivo de salida
	Deformaciones en el soporte
Uso de los resultados	Entendimiento incorrecto
	Uso inapropiado

Normas de Información Geográfica

19157 Calidad del dato

19113 Principios de Calidad

19114 Evaluación de la Calidad.

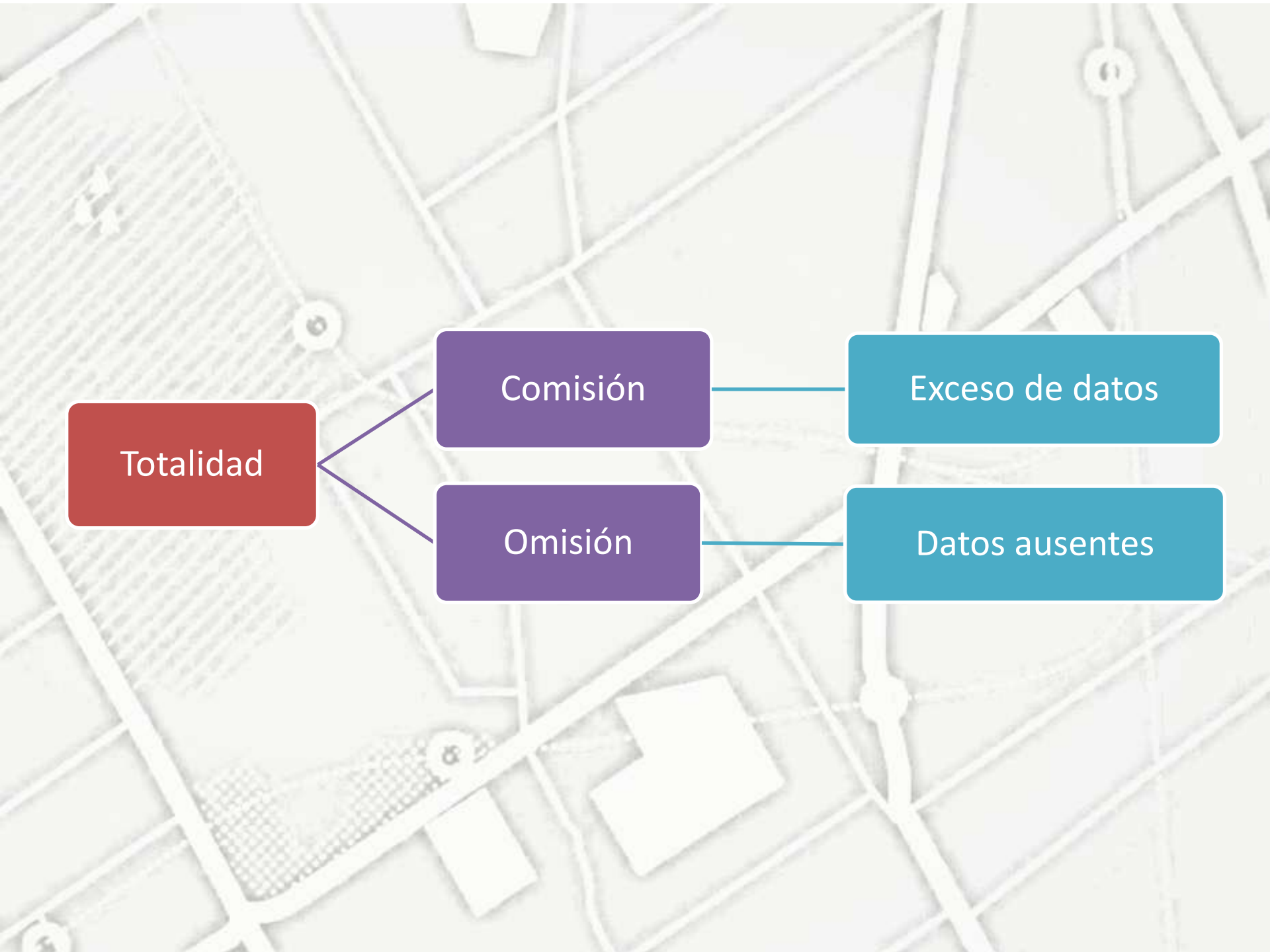
19138 Medidas de calidad de los datos

La calidad no es una única dimensión

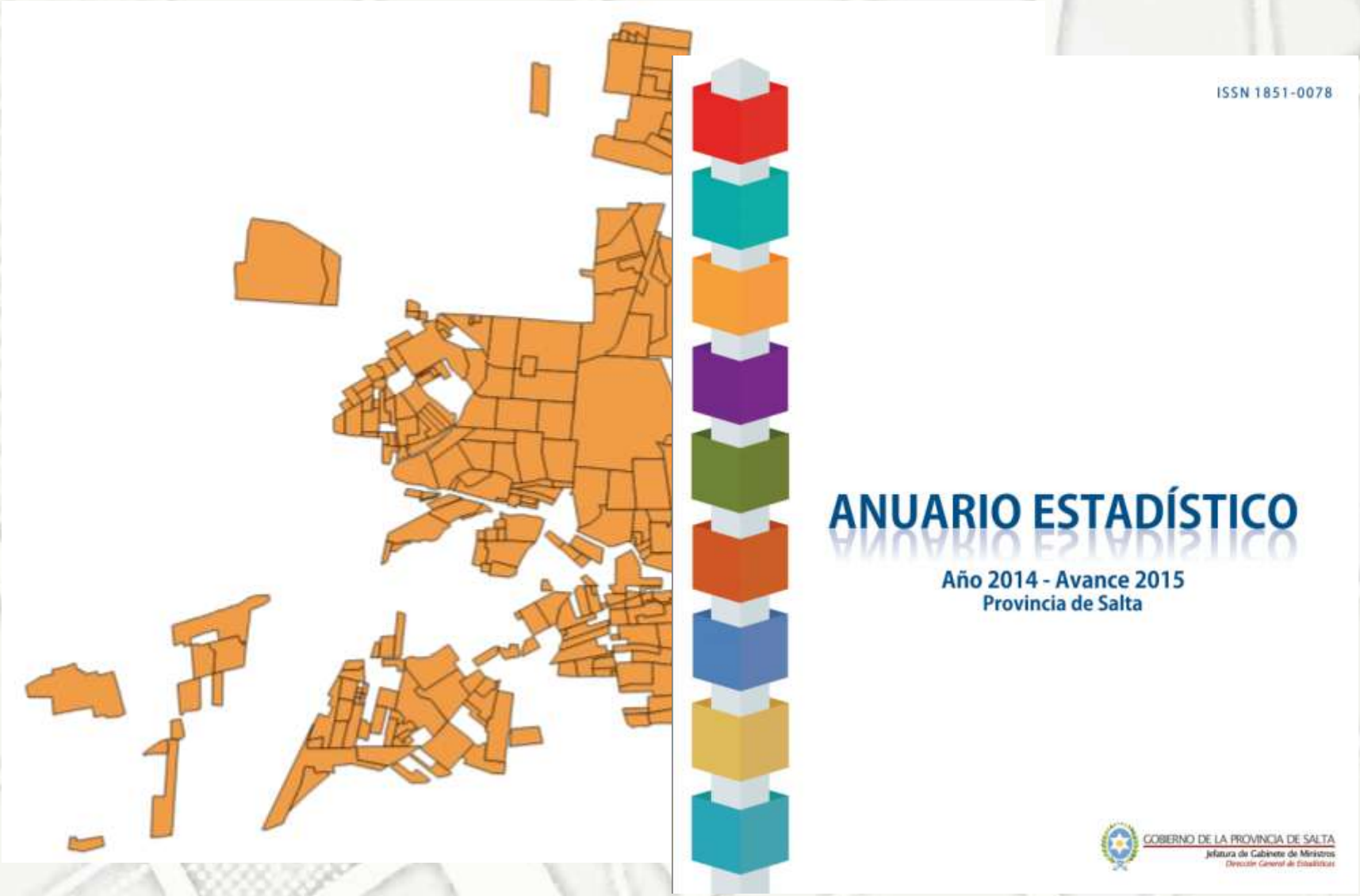
Existen elementos cualitativos y elementos cuantitativos que permiten la evaluación de la calidad.

Elementos cualitativos:

- **Linaje o Genealogía**
- **Uso**
- **Propósito**

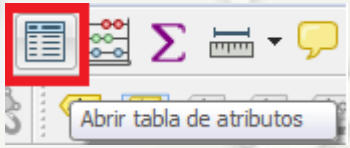

Totalidad
Comisión
Omisión
Exceso de datos
Datos ausentes

Totalidad: Barrios de la ciudad de Salta 2010

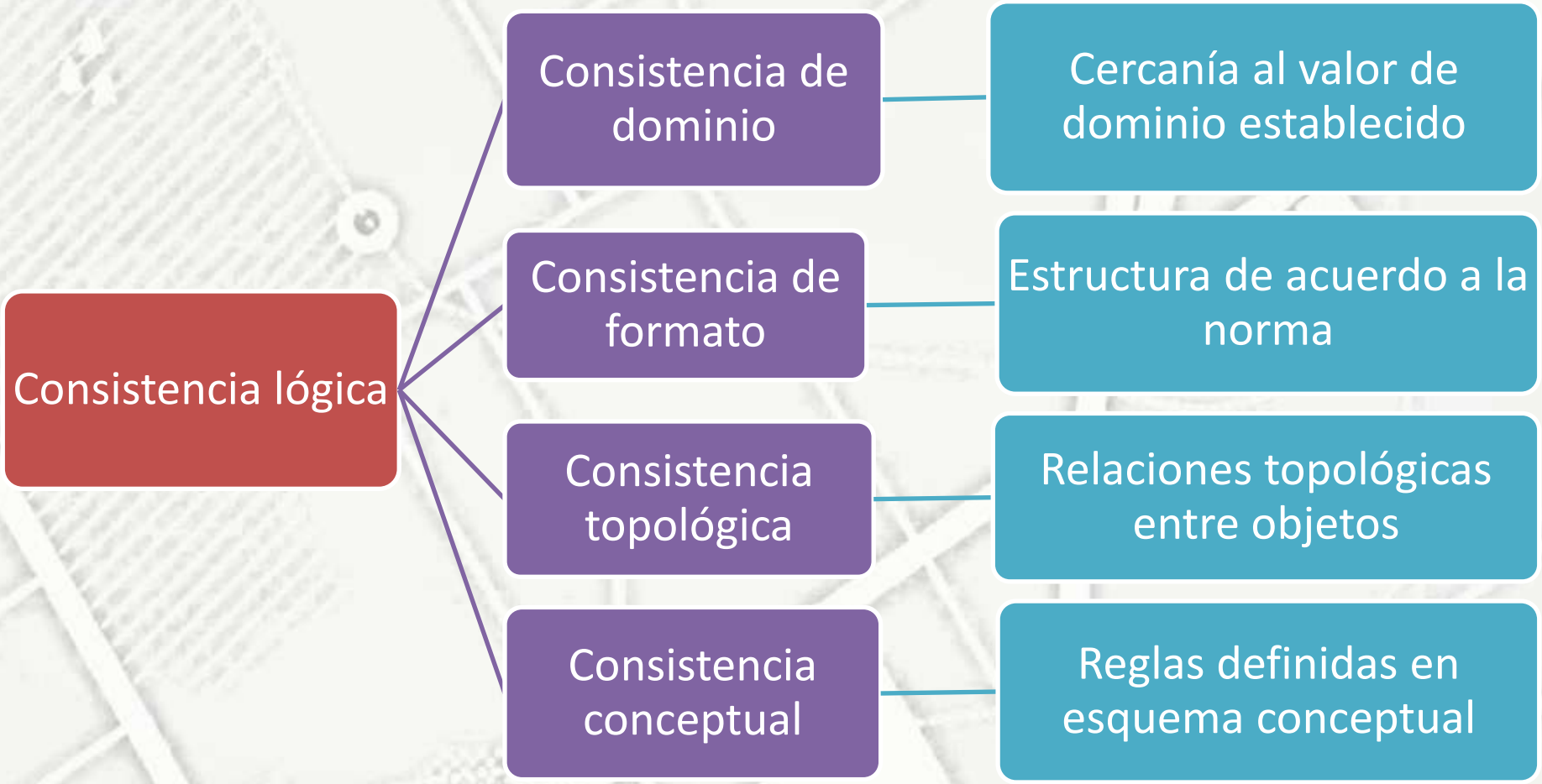


3.5.4_Hogares y Población: Total y con Necesidades Básicas Insatisfechas (NBI) según Barrios y Villas de la Ciudad de Salta.

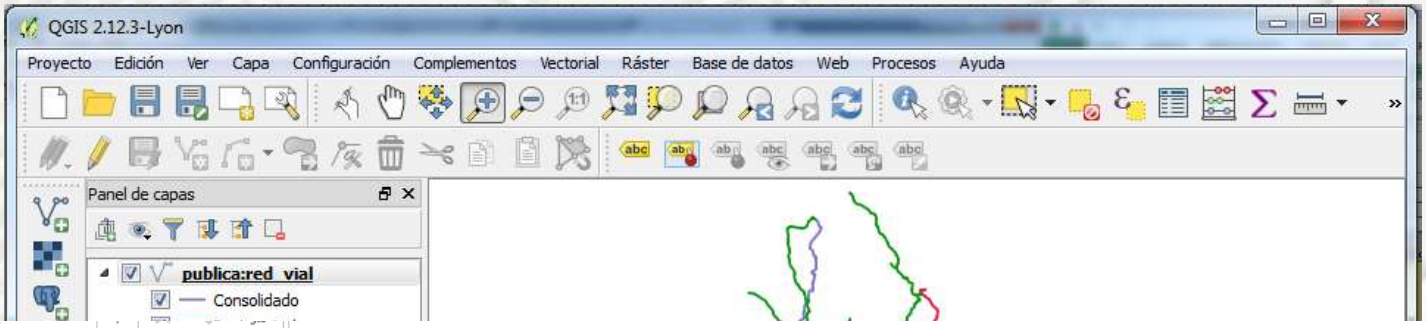
Código	Nombre del barrio y villa	Hogares (1)			Población (1)		
		Total	Con NBI (2)	% (3)	Total	Con NBI (2)	% (4)
	Total	**106.946**	**19.644**	**18,4**	**457.427**	**96.863**	**21,2**
1	15 de Setiembre	326	169	51,8	1.407	809	57,5
2	Juan Manuel de Rosas	635	280	44,1	2.918	1.486	50,9
3	La Tradición	126	56	44,4	518	221	42,7
4	Patricia Heitman	136	59	43,4	563	289	51,3
5	Ciudad del Milagro	1.836	77	4,2	6.[illegible]		
6	17 de Octubre	780	271	34,7	3.[illegible]		
7	La Unión	395	129	32,7	2.[illegible]		
8	1º de Mayo	406	105	25,9	2.[illegible]		
9	Juan Pablo II	347	120	34,6	1.[illegible]		
10	Leopoldo Lugones	418	146	34,9	1.[illegible]		
11	Castañares	2.494	185	7,4	12.027	1.158	9,6
12	Manuel J. Castilla	474	49	10,3	2.345	313	13,3
13	15 de Febrero	310	165	53,2	1.427	832	58,3
14	Loteo Barrio Policial	50	10	20,0	188	44	23,4
15	Asentamiento Universitario	74	25	33,8	397	159	40,1
16	Universitario	414		1,9	1.656	40	2,4



Infraestructura de Datos Espaciales de la República Argentina						
Instrumento de evaluación de la calidad de la información geoespacial						
Elemento: Totalidad						
Subelementos: Comisión y omisión						
Objeto	Elementos en el Universo Abstracto	Elementos en el Producto	Error de Comisión	(%)	Error de Omisión	(%)
Barrio	288	188	100	53,2	21	11,2
Totales	288	188	100	**53,2**	21	**11,2**


Consistencia lógica
Consistencia de dominio
Cercanía al valor de dominio establecido
Consistencia de formato
Estructura de acuerdo a la norma
Consistencia topológica
Relaciones topológicas entre objetos
Consistencia conceptual
Reglas definidas en esquema conceptual

Consistencia de dominio: Red vial de la provincia del Neuquén



A ATRIBUTOS

A LISTA DE ATRIBUTOS

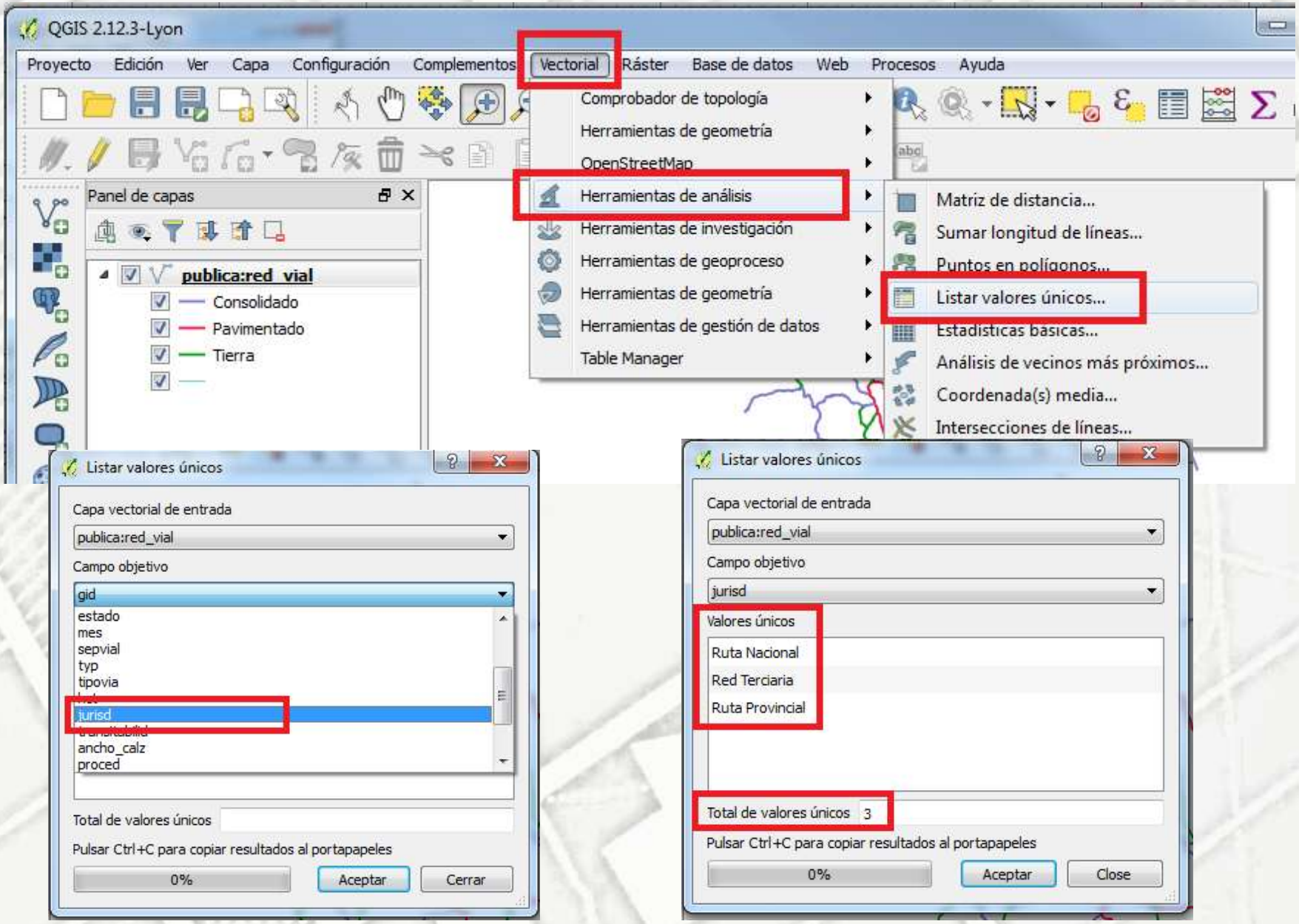
IDERA
Infraestructura de Datos Espaciales de la República Argentina

CATÁLOGO DE OBJETOS GEOGRÁFICOS DE LA INFRAESTRUCTURA DE DATOS ESPACIALES DE LA REPÚBLICA ARGENTINA

VALORES DE DOMINIOS			OBSERVACIONES
Código	Etiqueta	Definición	
HCT	Jurisdicción de vía de transporte		
-1	Información no disponible	Cuando no se dispone o no se tiene certeza de los datos.	
1	Ruta Nacional	Vía de jurisdicción nacional. Forma parte de la red vial troncal del país.	
2	Ruta Provincial	Vía de jurisdicción provincial.	
4	Camino terciario	Vía de jurisdicción local. Incluyendo municipal, vecinal, privada.	

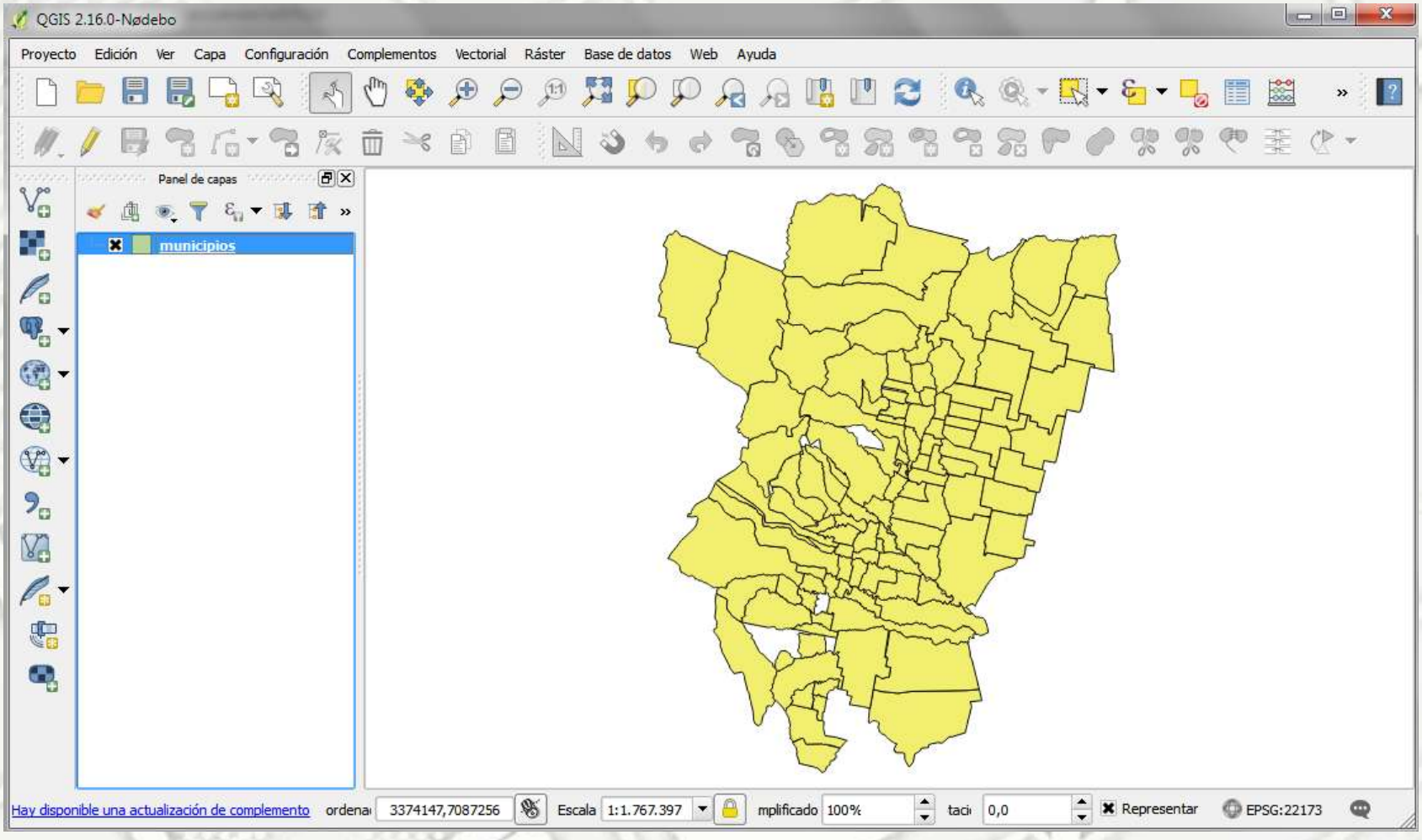
ACD AMG BCP BSC CA1 CAA CAU CEA CEM CER CIF DMA ESI ETD FAA FRI FUC FUN GDD HCT HQC HYP IDA

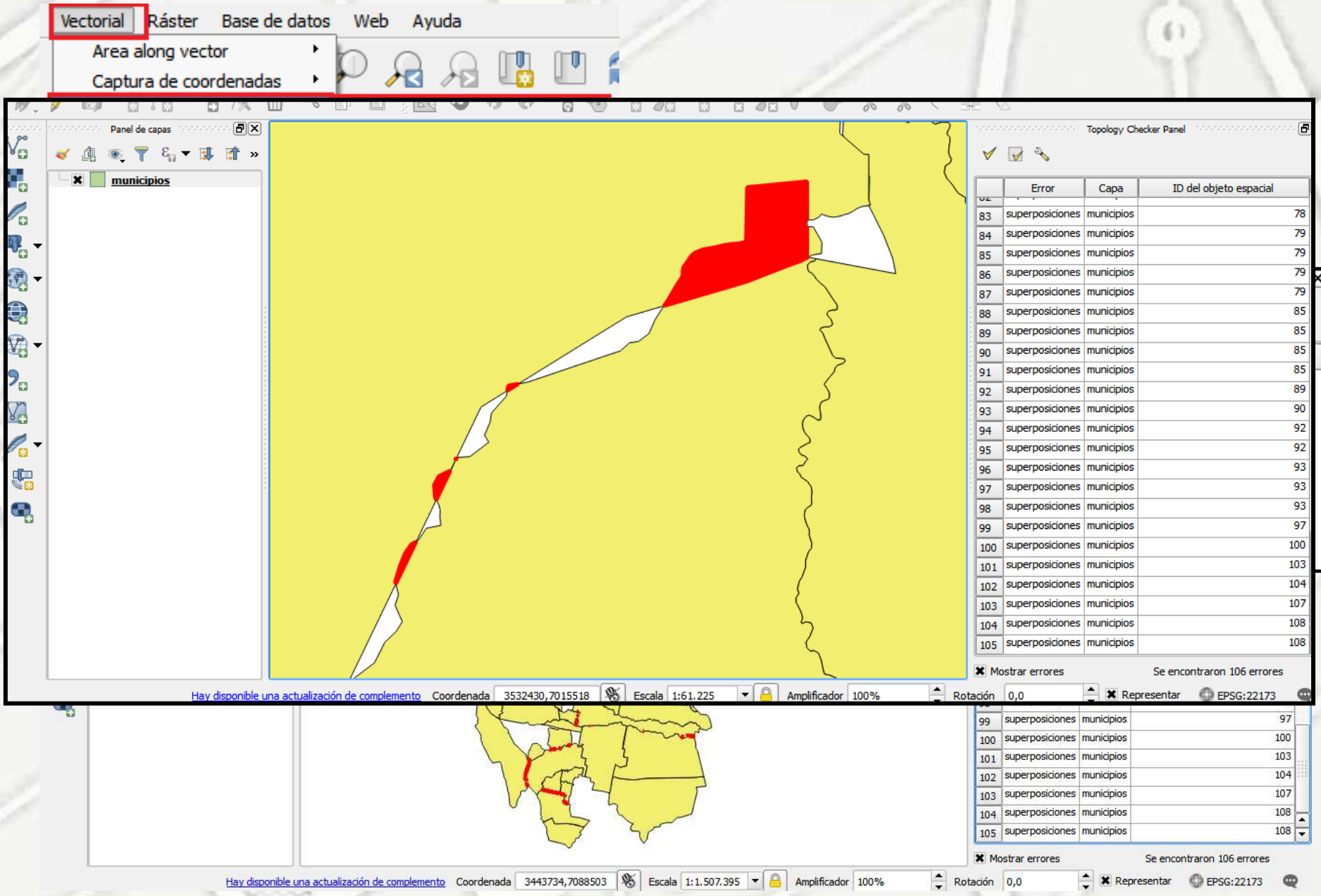
oordenada 2686819,5947055 scala 1:3.396.623 otación 0,0 Representar EPSG:22182 (al vuelo)

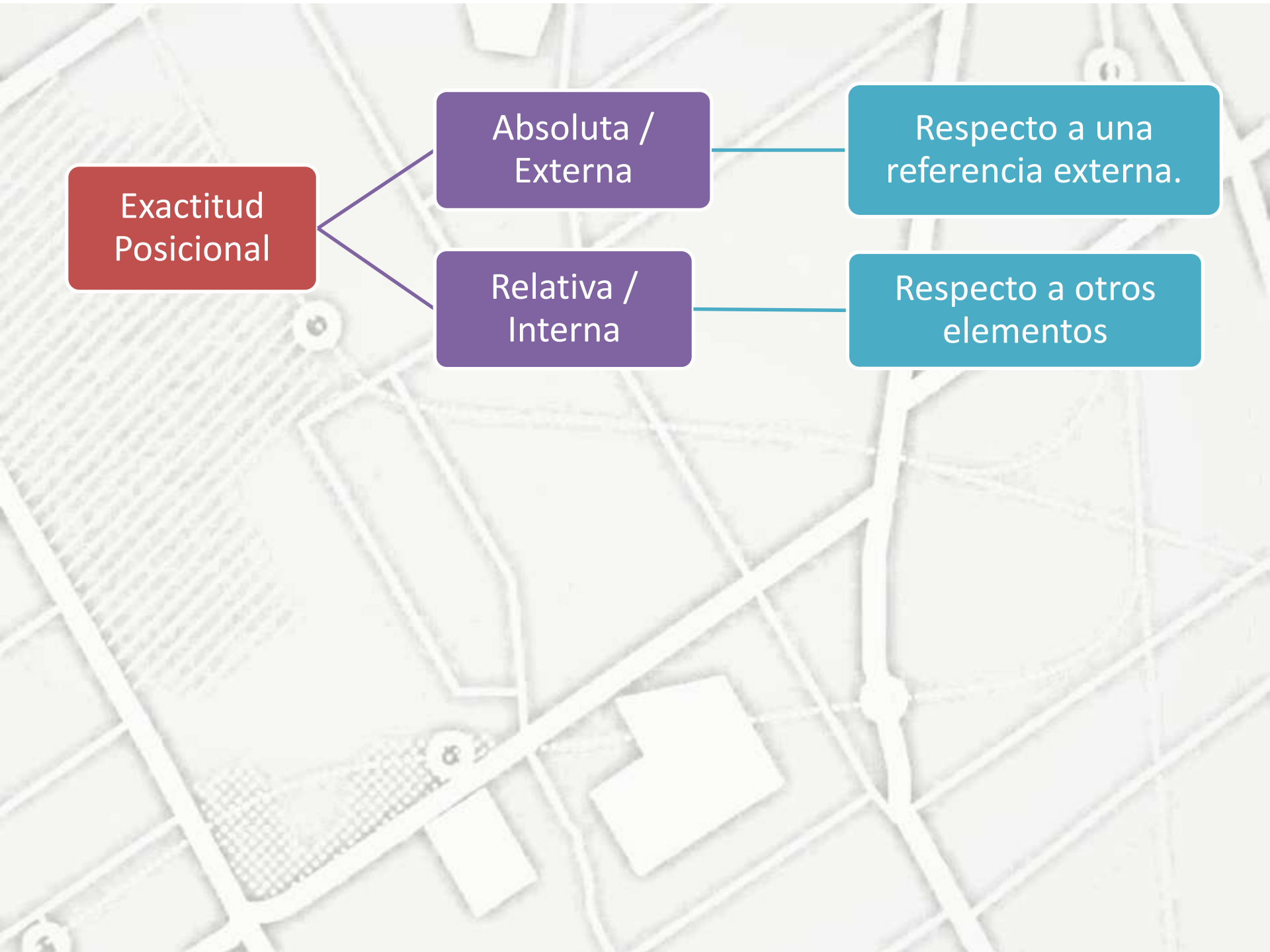

QGIS 2.12.3-Lyon
Proyecto
Edición
Ver
Capa
Configuración
Complementos
Vectorial
Ráster
Base de datos
Web
Procesos
Ayuda
Comprobador de topología
Herramientas de geometría
OpenStreetMap
Herramientas de análisis
Herramientas de investigación
Herramientas de geoproceso
Herramientas de geometría
Herramientas de gestión de datos
Table Manager
Matriz de distancia...
Sumar longitud de líneas...
Puntos en polígonos...
Listar valores únicos...
Estadísticas básicas...
Análisis de vecinos más próximos...
Coordenada(s) media...
Intersecciones de líneas...
Panel de capas
publica:red_vial
Consolidado
Pavimentado
Tierra
Listar valores únicos
Capa vectorial de entrada
publica:red_vial
Campo objetivo
gid
estado
mes
sepvial
typ
tipovia
jurisd
ancho_calz
proced
Total de valores únicos
Pulsar Ctrl+C para copiar resultados al portapapeles
0%
Aceptar
Cerrar
Listar valores únicos
Capa vectorial de entrada
publica:red_vial
Campo objetivo
jurisd
Valores únicos
Ruta Nacional
Red Terciaria
Ruta Provincial
Total de valores únicos 3
Pulsar Ctrl+C para copiar resultados al portapapeles
0%
Aceptar
Close

Infraestructura de Datos Espaciales de la República Argentina			
Instrumento de evaluación de la calidad de la información geoespacial			
Elemento: Consistencia lógica			
Subelementos: Consistencia de dominio			
Atributo	**Rango de dominio en Catálogo de objetos**	**Dominio en la BD**	**Cumplimiento**
Jurisdicción de vía de transporte	Ruta Nacional	Ruta Nacional	SI
Jurisdicción de vía de transporte	Ruta Provincial	Ruta Provincial	SI
Jurisdicción de vía de transporte	Camino terciario	Red Terciaria	NO

Consistencia topológica: Municipios y comunas de Tucumán




Vectorial
Ráster
Base de datos
Web
Ayuda
Area along vector
Captura de coordenadas
Panel de capas
municipios
Topology Checker Panel
Error
Capa
ID del objeto espacial
83 superposiciones municipios 78
84 superposiciones municipios 79
85 superposiciones municipios 79
86 superposiciones municipios 79
87 superposiciones municipios 79
88 superposiciones municipios 85
89 superposiciones municipios 85
90 superposiciones municipios 85
91 superposiciones municipios 85
92 superposiciones municipios 89
93 superposiciones municipios 90
94 superposiciones municipios 92
95 superposiciones municipios 92
96 superposiciones municipios 93
97 superposiciones municipios 93
98 superposiciones municipios 93
99 superposiciones municipios 97
100 superposiciones municipios 100
101 superposiciones municipios 103
102 superposiciones municipios 104
103 superposiciones municipios 107
104 superposiciones municipios 108
105 superposiciones municipios 108
Mostrar errores
Se encontraron 106 errores
Hay disponible una actualización de complemento
Coordenada 3532430,7015518
Escala 1:61.225
Amplificador 100%
Rotación 0,0
Representar
EPSG:22173
Coordenada 3443734,7088503
Escala 1:1.507.395


Exactitud Posicional
Absoluta / Externa
Relativa / Interna
Respecto a una referencia externa.
Respecto a otros elementos

Normas ISO-19157

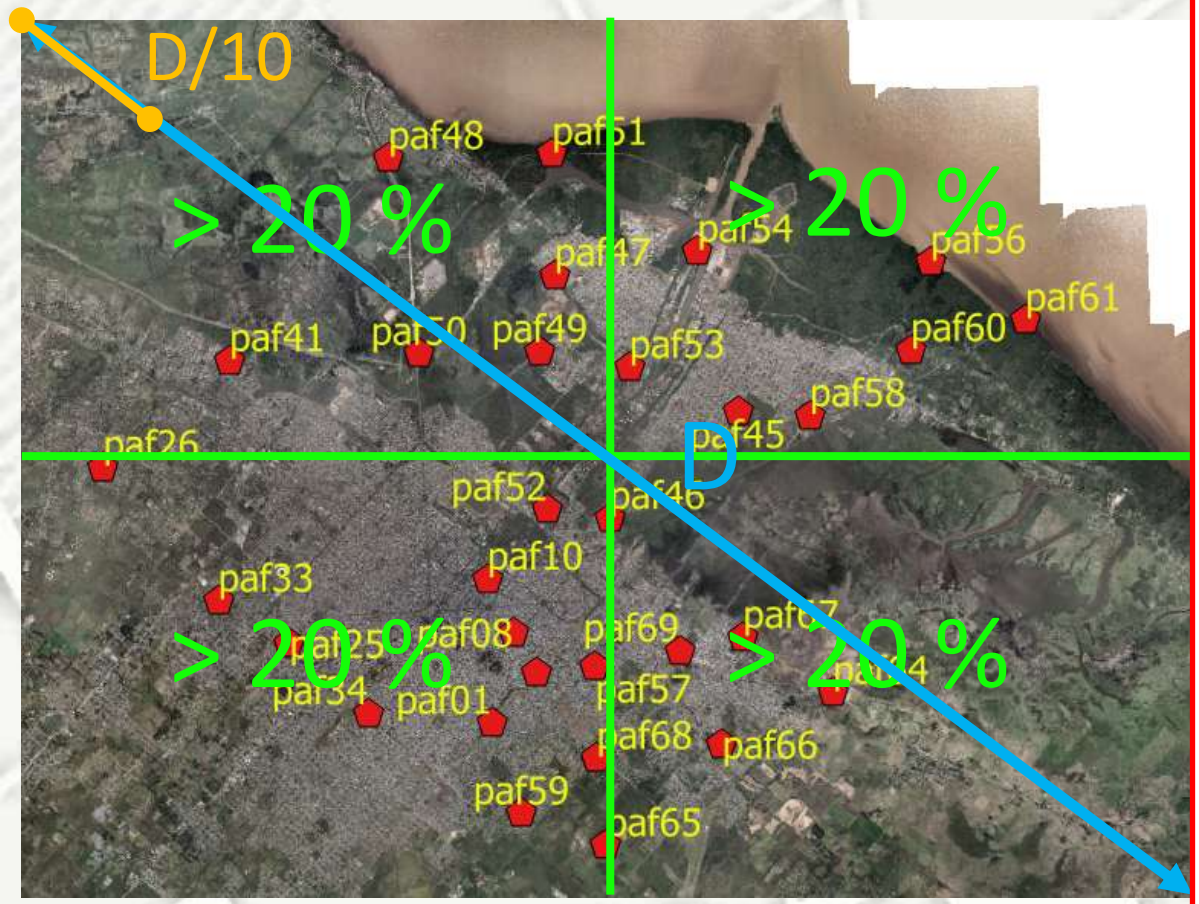
Exactitud Posicional: Algunas definiciones

- Desviación simple: $e_{xi} = x_{mi} - x_{ti}$ Error: $e_i = |e_{xi}|$
- Error Medio: $e = \frac{1}{\ } \times \Sigma e_i$
- B

Homogéneos
No correlacionados
Tienen una distribución Normal

Error circular: $\frac{1,1774}{\sqrt{2}} \times \sqrt{\sigma_x{}^2 + \sigma_y{}^2}$	Error lineal: 0,6745*σ
	Error standard: 1*σ
CE90% (CMAS): $\frac{2,146}{\sqrt{2}} \times \sqrt{\sigma_x{}^2 + \sigma_y{}^2}$	Exactitud 90 %: 1,645*σ
CE95%: $\frac{2,4477}{\sqrt{2}} \times \sqrt{\sigma_x{}^2 + \sigma_y{}^2}$	Exactitud 95 %: 1,96*σ

Ejemplo de control de exactitud posicional



Elección de los Puntos de Control:

_ Mínimo 20 puntos de control

Table 7. Recommended Number of Check Points Based on Area

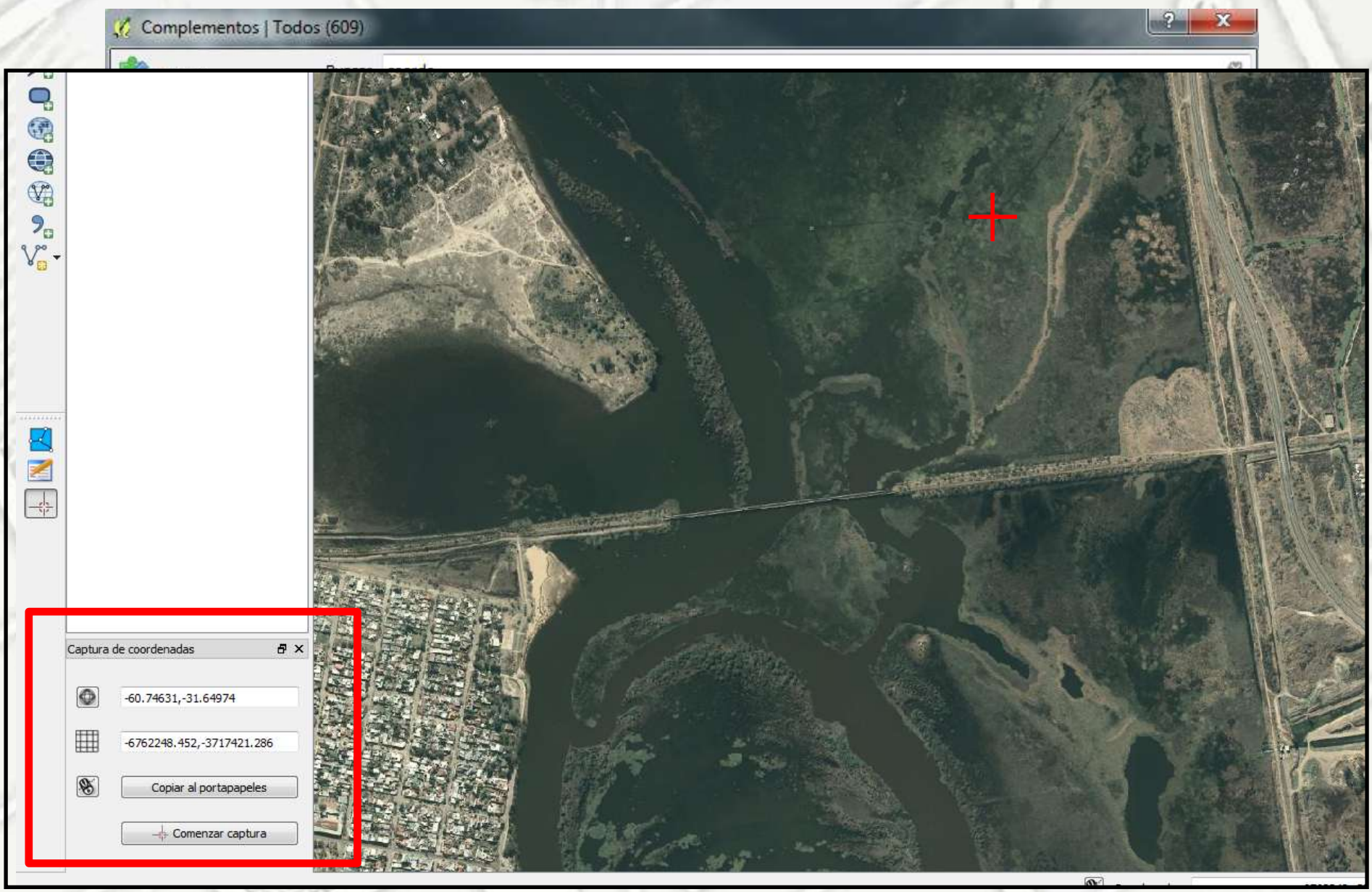
Project Area (Square Kilometers)	Horizontal Testing	Vertical Testing (not clearly-defined points)		
	Total Number of Static Horizontal Check Points (clearly-defined points)	Number of Static Vertical Check Points in NVA	Number of Static Vertical Check Points in VVA	Total Number of Static Vertical Check Points
<500	20	20	0	20
501-750	25	20	10	30
751-1000	30	25	15	40
1001-1250	35	30	20	50
1251-1500	40	35	25	60
1501-1750	45	40	30	70
1751-2000	50	45	35	80
2001-2250	55	50	40	90
2251-2500	60	55	45	100

Puntos de acuerdo al área de la imagen recomendado por el ASPRS

_ Distribuidos en forma uniforme

_ Separados por una distancia equivalente al 10 % de su diagonal

x_{mi} = puntos obtenidos de la imagen obtenida del servicio WMS del IGN
x_{ti} = puntos propios obtenidos mediante medición GPS diferencial


Complementos | Todos (609)
Captura de coordenadas
-60.74631,-31.64974
-6762248.452,-3717421.286
Copiar al portapapeles
Comenzar captura

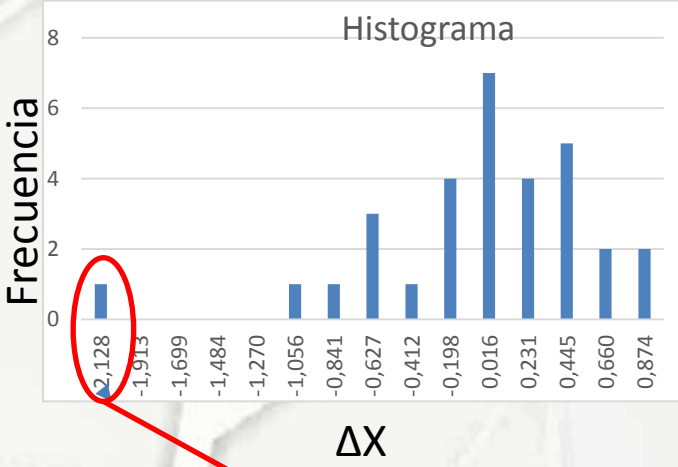
Ejemplo de control de exactitud posicional

Primeros Pasos:

_ Calcular las desviaciones simples: $e_{xi} = x_{mi} - x_{ti}$

_ Determinar los "outliers"

_ Hacer un análisis estadístico



* Media = 0 (test de Student)
  $t = (e_{ixm} \times \sqrt{n-1})/\mathrm{RMSEx}^2$

* Aleatoreos (Chi cuadrado)
  $\chi^2 = [RMSE_x{}^2 \times (n-1)]/\sigma^2$

Datos (n):	30	Student	0,879
Media =	0,039 m		
RMSE =	0,477 m	Chi cuadrado	27,235
σ =	0,5		

	90%	95%
Student	1,31	1,7
Chi cuadrado	39,09	42,56

Ejemplo de control de exactitud posicional

Cálculo de errores:

$\mathrm{RMSE_z} = 0{,}485$ m

Error lineal: 0,6745* $\mathrm{RMSE_z}$ = 0,327 m

Error estándar: 1* $\mathrm{RMSE_z} = 0{,}485$ m

Exactitud 90 %: 1,645* $\mathrm{RMSE_z}$= 0,797 m

Exactitud 95 %: 1,96* $\mathrm{RMSE_z}$= 0,950 m

PEC	EM (m)	EP (m)
A	2,70	1,67
B	5,00	3,33
C	6,00	4,00
D	7,50	5,00

MINISTÉRIO DA DEFESA EXÉRCITO BRASILEIRO DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA
NORMA DA ESPECIFICAÇÃO TÉCNICA PARA CONTROLE DE QUALIDADE DE DADOS GEOESPACIAIS (ET-CQDG)

NSSDA	NMAS
Agencies set thresholds or tolerances for their product specifications	Defined thresholds
95% confidence	90% confidence
Horizontal Accuracy 1.7308 * $\mathrm{RMSE_r}$	Horizontal Accuracy 1.5175 * $\mathrm{RMSE_r}$
Vertical Accuracy 1.9600 * $\mathrm{RMSE_z}$	Vertical Accuracy 1.645 * $\mathrm{RMSE_z}$

National Standard for Spatial Data Accuracy. (Federal Geographic Data Committee)
National Map Accuracy Standard (U.S. BUREAU OF THE BUDGET)

Table 1. Horizontal Accuracy Standards for Orthophotos

Horizontal Data Accuracy Class	RMSEx and RMSEy	Orthophoto Mosaic Seamline Maximum Mismatch	Aerial Triangulation or INS-based RMSEx RMSEy and RMSEz
I	Pixel size x 1.0	Pixel size x 2.0	Pixel size x 0.5
II	Pixel size x 2.0	Pixel size x 4.0	Pixel size x 1.0
III	Pixel size x 3.0	Pixel size x 6.0	Pixel size x 1.5
.....			
N	Pixel size x N	Pixel size x 2N	Pixel size x 0.5N

Table 2. Horizontal Accuracy Standards for Digital Planimetric Data

Horizontal Data Accuracy Class	RMSEx and RMSEy (cm)
I	1.25% of Map Scale Factor (0.0125 x Map Scale Factor)
II	2.0 x Class I Accuracy (0.025 x Map Scale Factor)
III	3.0 x Class I Accuracy (0.0375 x Map Scale Factor)
.....	
N	N x Class I Accuracy

ASPRS (The American Society for Photogrammetry and Remote Sensing) Accuracy Standards for Digital Geospatial Data

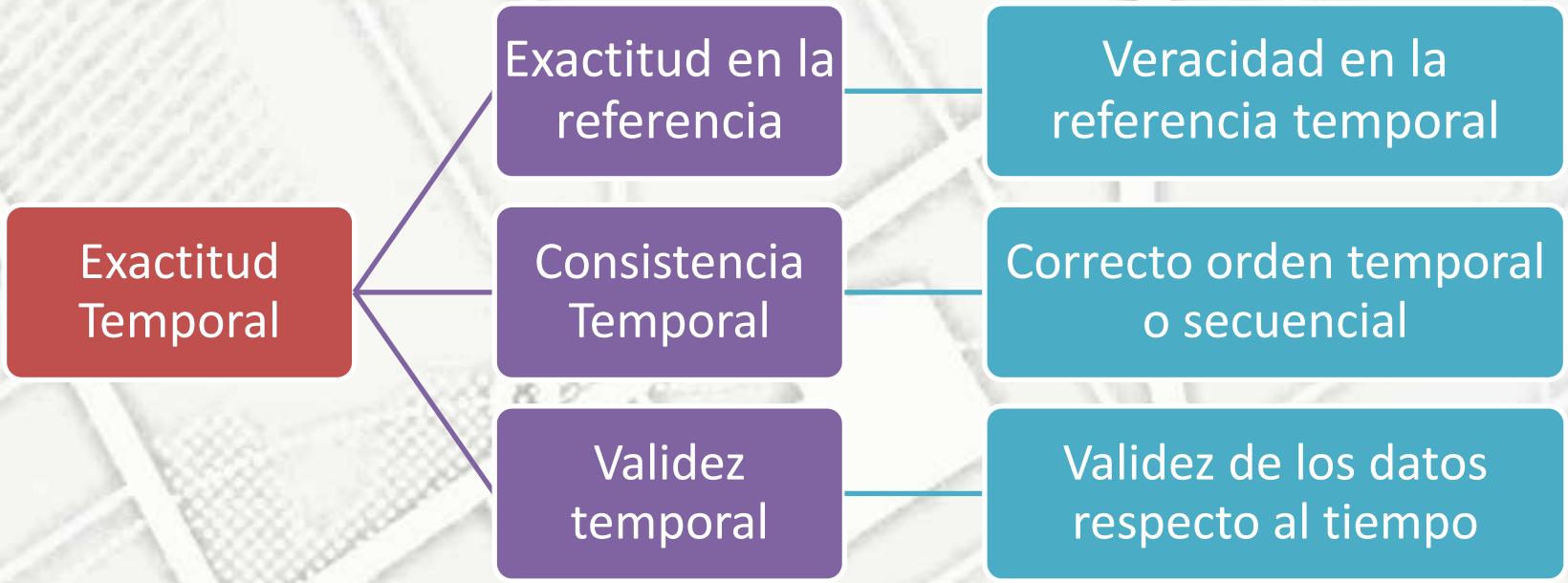
Objetivos:

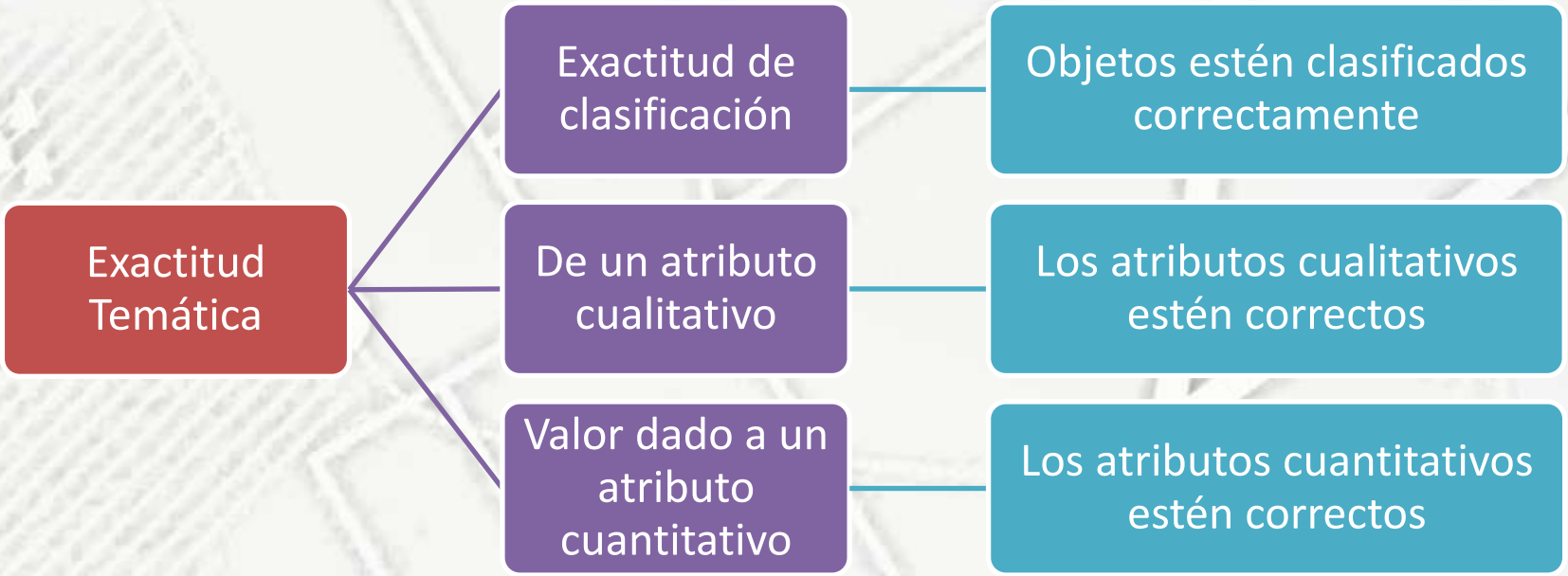
Desarrollar herramientas de cálculo de los datos estadísticos

Planilla Excel

Aplicación Online

Establecer estándares de calidad para ser aplicados a IDERA



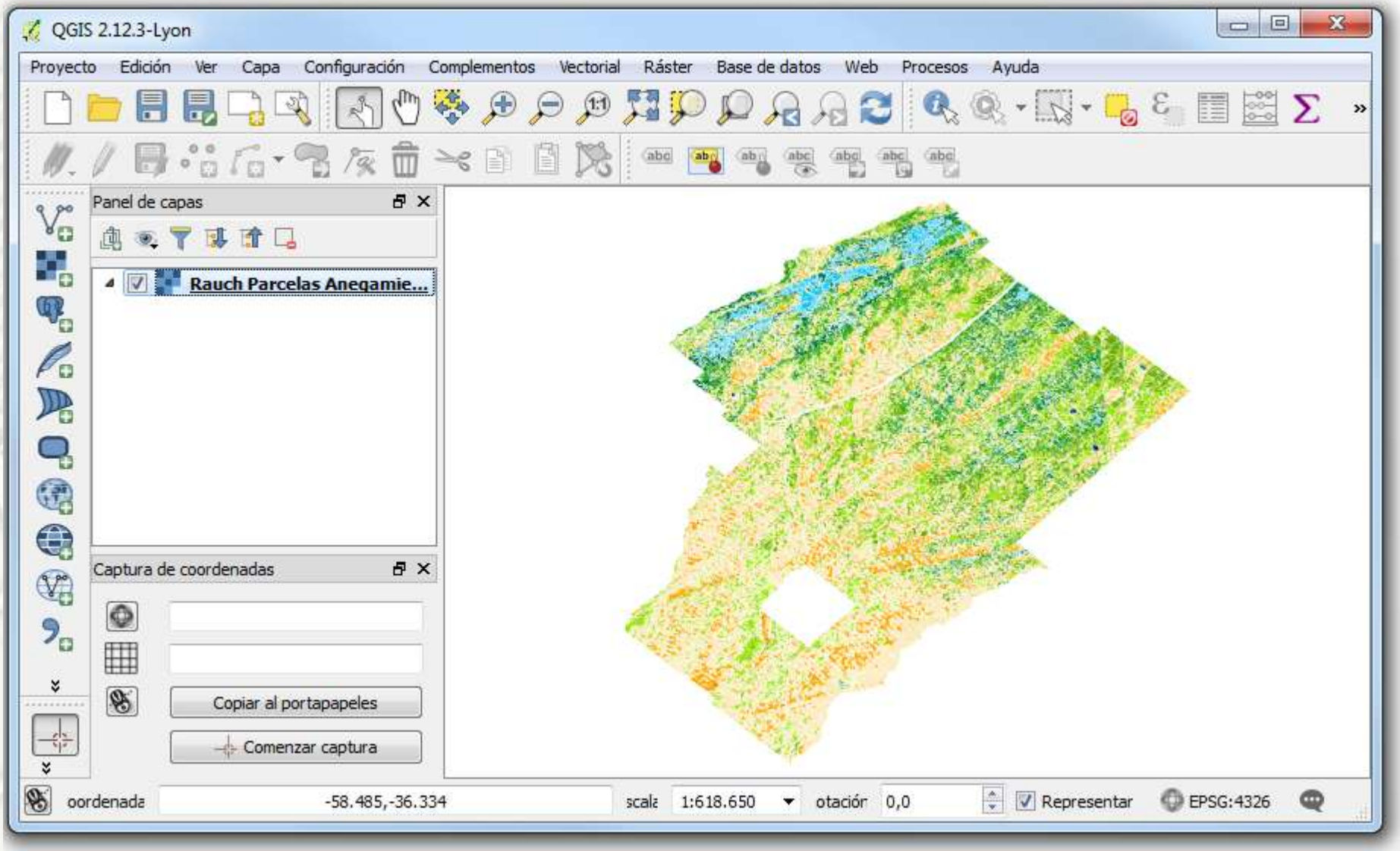


Tamaño de la muestra

$$n = \frac{N \times S^2 \times Z^2}{[N \times E^2 + S^2 \times Z^2]}$$

N = poblacion total
Z = nivel de significancia
Ej: 1,96 para el 95 %
S = proporcion casos favorable (se sugiere S=1/2)
E = Error maximo admitible

Exactitud temática



Componente	Descripción	
Nombre	Nombre de la medida de calidad empleada de acuerdo al elemento y subelemento de calidad a evaluar.	
Nombre alternativo	Nombre alternativo de la medida de calidad a emplea	
Alcance o nivel de medición de la calidad	Definición y características del dato al cual se va a realizar la evaluación de calidad: Conjunto o serie de datos, objeto, tipo de objeto, instancia, atributo, etc.	
Elemento de la calidad	Componente cuantitativo que describe la calidad de un conjunto de datos geográficos, definido por normas nacionales e internacionales que establecen conceptos y parámetros a tener en cuenta.	Elija una Opción
Subelemento de la calidad	Componente de un elemento de calidad de los datos cuantitativos que describe cierto aspecto de dicho elemento.	Elija una Opción
Medida de la calidad		
Medida básica de calidad	Nombre de la medida básica de la cual se deriva la medida de calidad: error, exactitud, conteo de errores, porcentaje de errores, conteo de aciertos, relación de ítems correctos, etc.	
Definición	Definición del método para el cálculo del valor de la medida de calidad.	
Descripción	Descripción de la medida de calidad incluyendo las fórmulas necesarias para obtener resultados de la medida.	
Parámetro	Variable auxiliar usada por la medida.	
Identificador de la medida	Código que identifica la medida de calidad en la NTC 5660 o en el registro de medidas de calidad de la entidad.	
Método de evaluación		
Tipo de método de evaluación	Definición del tipo de método a utilizar para realizar la evaluación de calidad.	
Descripción del método de evaluación	Descripción general de la forma o metodología realizada para hacer la evaluación.	
Resultado de la calidad		
Tipo del valor	Tipo de dato para reportar el resultado de la evaluación de calidad, en función de la medición realizada.	
Fuente de referencia	Es la fuente de la cual se toman los datos para realizar la evaluación de calidad	
Valor	Valor obtenido una vez realizada la medición de calidad.	
Unidad	Para los casos que aplique es la unidad de medida obtenida de la revisión de calidad.	
Fecha	Fecha de cuando se realiza la medición de calidad	
Nivel de conformidad	Umbral de error o resultado que se espera para dar por aprobada la prueba de calidad.	
Interpretación del resultado	Descripción clara de los resultados encontrados en la evaluación de calidad realizada	

Propuestas a futuro para el Observatorio

- Brindar asistencia a los productores de los datos para la determinar sus propias reglas de calidad
- Dar a conocer las normas / estándares y procedimientos.
- Difundir las herramientas para documentar procedimientos
- Promover la evaluación de la calidad en cada paso de los procesos de los datos pertenecientes a IDERA
- Fomentar la documentación y publicación de los resultados de la evaluación para poder hacer una correcta gestión de la calidad
- Desarrollar aplicaciones informáticas en línea y de escritorio para la evaluación de la calidad.

Observatorio de Evaluación de la Información Geoespacial

Muchas gracias

Instituto Geográfico Nacional

Universidad Nacional de La Plata

IDERA

Infraestructura de Datos Espaciales
de la República Argentina