



DATA

La naturaleza del dato

Estructura, clasificación
y formalización

Ramírez • Araiza



APRENDA
EDICIONES

¿Realmente entiendes qué es un dato?

En un entorno dominado por la analítica, la inteligencia artificial y la toma de decisiones basada en información, muchos profesionales trabajan con datos sin comprenderlos a profundidad. El resultado: modelos ambiguos, interpretaciones inconsistentes y soluciones débiles.

La naturaleza del dato. Estructura, clasificación y formalización propone un enfoque distinto: comprender los datos antes de analizarlos.

Este libro aborda el dato desde su origen, como representación abstracta de la realidad, y establece fundamentos para su clasificación, modelación relacional y formalización estructurada.

Más allá de herramientas o plataformas, esta obra desarrolla criterio: la capacidad de comprender, estructurar y definir correctamente la información, sentando bases para la automatización inteligente de datos y el uso efectivo de IA generativa.

Dirigido a profesionales, analistas e ingenieros, este libro es una guía esencial para quienes buscan dejar de usar datos... y comenzar a entenderlos.

Aprenderás a:



Describir
fenómenos
usando
datos



Clasificar
datos
mediante
DTXC



Comprender
modelos
relacionales y
semánticos



Preparar
datos para
análisis e IA



Aplicar
mejores
prácticas para
soluciones
consistentes



Forma parte del cuerpo de conocimiento para la certificación **CDE** (**Certified Data Expert**).



Escanea para acceder a
recursos descargables.



APRENDA
EDICIONES

Índice del libro

A continuación encontrarás el índice del libro, para que conozcas a detalle su contenido y estructura.

Ir a librería



Conoce todos los títulos disponibles en APRENDA. Aquí también podrás descargar archivos y recursos complementarios que acompañan a los títulos.

Comprar



Si el libro es de tu interés y deseas comprarlo, aquí puedes hacerlo.

Contenido

1	INTRODUCCIÓN	5
2	¿QUÉ SON LOS DATOS?	9
2.1	Introducción	9
2.2	Fenomenología de los datos	10
2.3	Los datos	11
2.4	Entidades y atributos	12
2.5	Información	13
2.6	Las ideas y los conceptos	14
2.7	El pensamiento	16
2.8	Conclusión	16
3	CLASIFICACIÓN DE LOS DATOS.....	19
3.1	Introducción	19
3.2	Taxonomía de los datos	20
3.3	Categorías Intrínsecas y semánticas en la taxonomía de datos	21
3.4	Clasificación por tipo de medida	23
3.5	Clasificación por tipo de dato (datatype)	26
3.6	Clasificación por tipo de uso	28
3.7	Clasificación por disponibilidad	34
3.8	Clasificación por requerimiento	35

2 La naturaleza del dato

3.9	Clasificación por causalidad	37
3.10	Clasificación por relacionalidad	38
3.11	Conclusiones	40
4	DTXC (DATA TAXONOMÍA CODE)	41
4.1	Introducción.....	41
4.2	Codificación taxonómica de datos (DTXC)	42
4.3	Códigos para el tipo de medida	43
4.4	Códigos para el tipo de dato	44
4.5	Códigos para el tipo de uso.....	44
4.6	Códigos para el tipo de disponibilidad.....	45
4.7	Códigos para el tipo de requerimiento	46
4.8	Códigos para el tipo de causalidad	47
4.9	Códigos para el tipo de relacionalidad	47
4.10	Ejemplo de uso DTXC	48
4.11	Conclusión.....	49
5	ANÁLISIS EXPLORATORIO DE DATOS (EDA)	51
5.1	Introducción.....	51
5.2	Calidad de los datos	52
5.3	Supuestos estadísticos	53
5.4	Etapas y finalidades de aplicar EDA	55
5.5	Herramientas gráficas univariadas.....	58
5.6	Herramientas numéricas univariadas	62
5.7	Herramientas numéricas bivariadas.....	65
5.8	Conclusión.....	66
6	MUESTREO DE DATOS	67
6.1	Introducción.....	67
6.2	Concepto de muestra estadística	68
6.3	Cálculo del tamaño de la muestra estadística.....	69
6.4	Muestra aleatoria simple	70
6.5	Muestra aleatoria estratificada	72
6.6	Conclusión.....	74
7	TRATAMIENTO DE DATOS AUSENTES Y ATÍPICOS	75
7.1	Introducción.....	75
7.2	Identificación de datos ausentes	76
7.3	Tipos de datos ausentes	77
7.4	Estrategias para tratar datos ausentes	81

7.5	Identificación de datos atípicos (outliers)	83
7.6	Percentiles y cuartiles	83
7.7	Regla de Tukey	85
7.8	Estrategias para tratar datos atípicos	85
7.9	Conclusión	88
8	CONCEPTOS DE BASE DE DATOS.....	89
8.1	Introducción	89
8.2	Concepto de bases de datos	90
8.3	Concepto de tablas, registros, campos	91
8.4	Concepto de llaves y relaciones	97
8.5	Concepto de cardinalidad y opcionalidad	100
8.6	Flujo de propagación	103
8.7	Modelo estrella (star) y copo de nieve (snow flake)	105
8.8	Conclusión	107
9	NORMALIZACIÓN DE LA BASE DE DATOS	109
9.1	Introducción	109
9.2	Vicios del modelo relacional	110
9.3	El proceso de normalización	112
9.4	Desnormalización.....	122
9.5	Conclusión	122
10	LOS DATOS Y SU FORMA DE REPRESENTACIÓN EN LOS NEGOCIOS.....	125
10.1	Introducción	125
10.2	Responde la pregunta correctamente	126
10.3	Cantidad de datos a mostrar.....	127
10.4	Segmentación por categorías	128
10.5	Composición de los datos	132
10.6	Secuencia de los datos	133
10.7	Tipos de acumulación.....	134
10.8	Términos de expresión.....	135
10.9	Forma gráfica y textual	136
10.10	Cálculos agregados y valores de referencias.....	139
10.1	Expresión de datos contables.....	139
10.2	Proceso de representación de datos	141
10.3	Conclusión	143
11	MODELOS SEMÁNTICO-RELACIONALES	145
11.1	Integración: el punto de convergencia del dato.....	145

4 La naturaleza del dato

11.2	El modelo relacional y la automatización inteligente de datos.....	146
11.3	Representación tradicional del modelo relacional.....	147
11.4	Estructura y semántica desarticuladas	149
11.5	¿Cómo impacta la falta de integración a la IA Generativa.....	149
11.6	Integración usando modelos semántico-relacionales (SRDM)	150
11.7	Modelo relacional vs. modelo semántico-relacional.....	152
11.8	Conclusión.....	159
12	HERRAMIENTAS DE ANALÍTICA DE DATOS	161
12.1	Introducción.....	161
12.2	Análisis de parámetros	162
12.3	Análisis de proporciones	163
12.4	Análisis de frecuencias	165
12.5	Análisis comparativo	170
12.6	Análisis de tendencias	174
12.7	Análisis de dispersión	176
12.8	Conclusión.....	180
13	ÍNDICE	181
14	REFERENCIAS	187