



**UNIVERSIDAD
CATÓLICA**
SEDES SAPIENTIAE

FACULTAD DE CIENCIAS ECONÓMICAS Y COMERCIALES



Ciencia de datos para la toma de decisiones



PONENTE PRINCIPAL:
**Mgtr. LUIS ENRIQUE
CAYATOPA RIVERA**

Economista e Investigador experto
en analítica y gestión pública.



MODERACIÓN:
MODESTA GONZÁLEZ SAÉNZ

Coordinadora del Centro de
Investigación y Desarrollo Innovador
para la Regionalización (CIDIR)





Objetivos y ruta de la charla

1. Comprender la disciplina

Distinguir ciencia de datos, big data, analítica, IA y visualización con un lenguaje útil para la gestión.

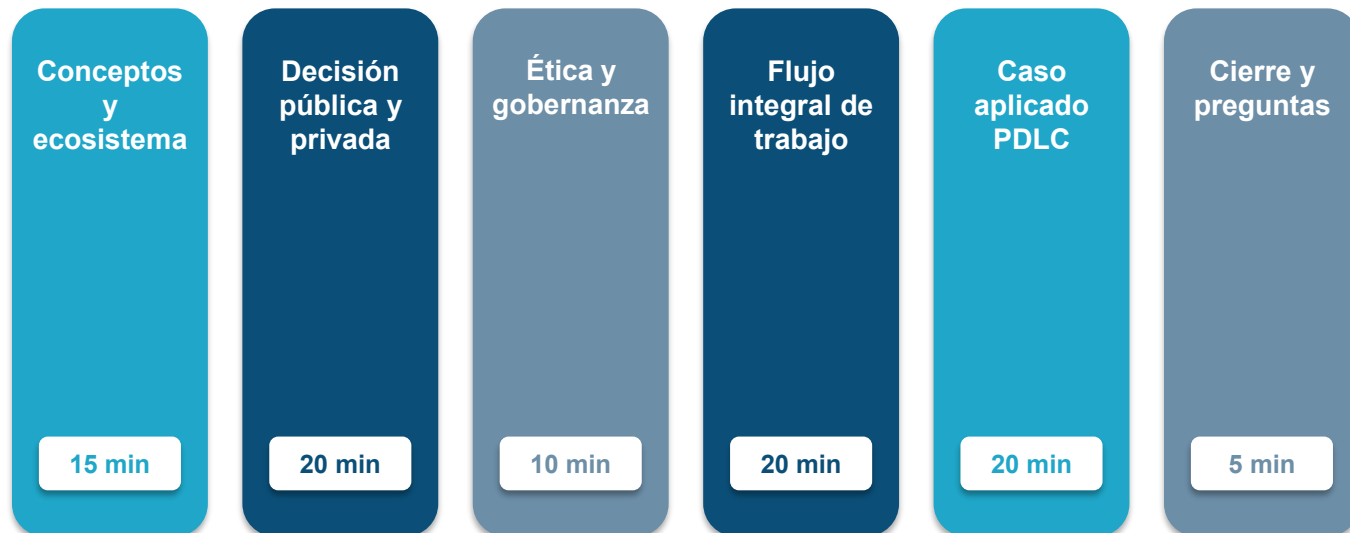
2. Conectar evidencia y decisión

Entender por qué los datos mejoran decisiones públicas y privadas, y cuáles son sus límites.

3. Revisar un flujo aplicable

Pasar del problema al dato, del dato a la evidencia y de la evidencia a la decisión y publicación.

90 minutos





¿Por qué este tema importa hoy?

Más huellas digitales 1

Operaciones, trámites, pagos, sensores, textos, geolocalización y registros administrativos.

Recursos escasos 2

Presupuestos ajustados obligan a priorizar mejor y a justificar decisiones.

La ventaja competitiva es decidir con evidencia

Mayor exigencia de resultados 3

Ciudadanos, directorios y reguladores demandan evidencia, trazabilidad y evaluación.

Riesgos más complejos 4

Fraude, desastres, congestión, abandono de clientes o focalización deficiente requieren anticipación.



I. Conceptualización de la ciencia de datos

Qué es, qué no es, y cómo se relaciona con otras disciplinas y con el big data.



¿Qué es ciencia de datos?

Definición operativa

Es la disciplina que convierte datos en **evidencia accionable** mediante estadística, programación, conocimiento del dominio, visualización y juicio crítico.

Su objetivo no es solo predecir. También describe, explica, monitorea y orienta decisiones.

1. Datos
Transaccionales, administrativos, encuestas, texto, imágenes, sensores.



2. Métodos
Limpieza, análisis exploratorio, modelos, experimentos, inferencia.



3. Productos
Indicadores, tableros, alertas, segmentación, pronósticos, recomendaciones.



4. Decisiones
Asignación de recursos, rediseño de procesos, evaluación y aprendizaje.



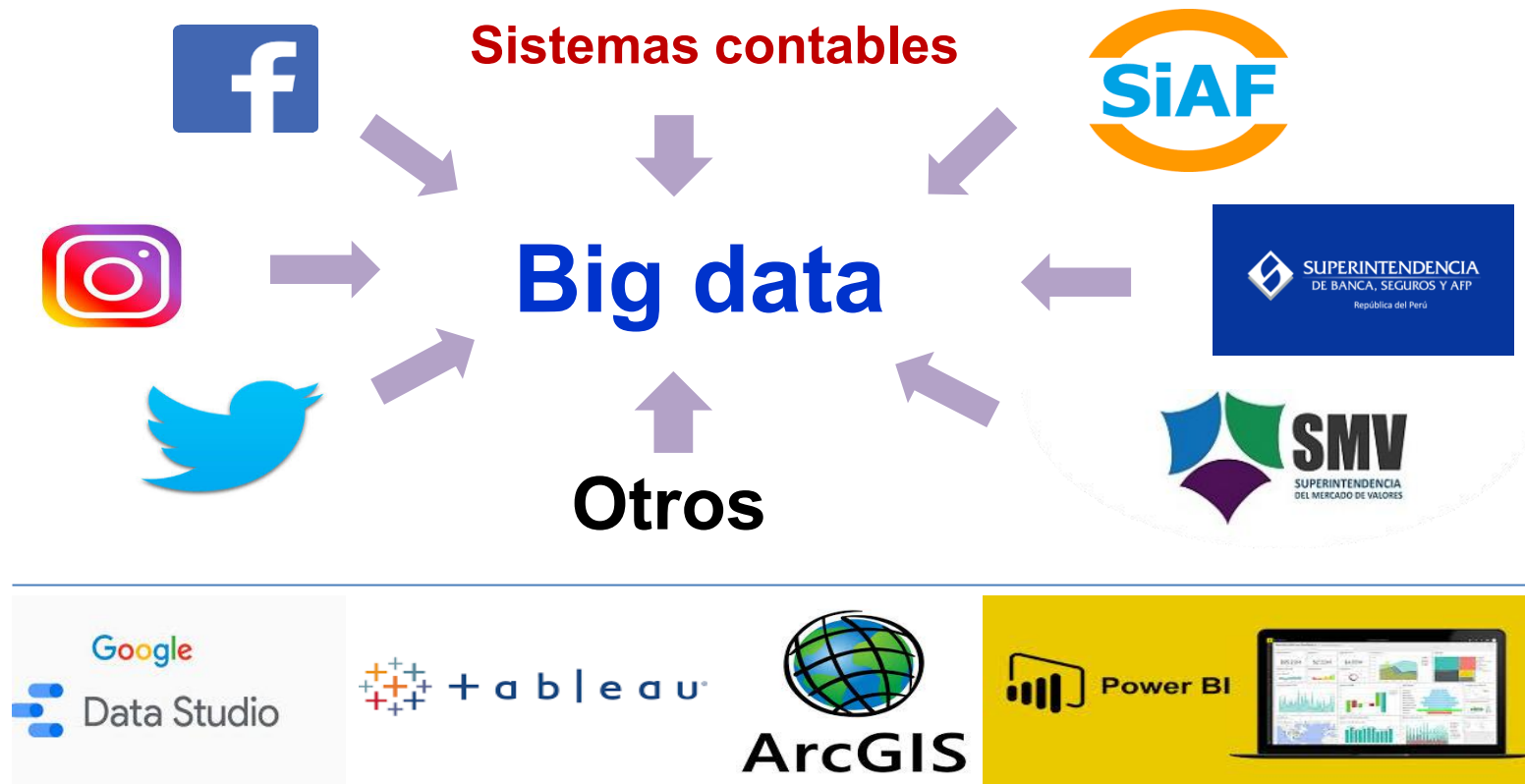
Relación con otras disciplinas y con big data



Big data es el contexto de volumen, velocidad y variedad. No es sinónimo de ciencia de datos.



Diversidad de datos





Big data: cuándo agrega valor y cuándo no

Volumen

Cantidad de registros que supera el manejo manual o local.

Velocidad

Datos que llegan en tiempo real o casi real.

Variedad

Tablas, textos, imágenes, GPS, formularios, audio.

Veracidad

Calidad, sesgos, faltantes y consistencia.

Valor

Capacidad de traducir datos en decisión útil.

Más datos no equivalen a mejores decisiones. Sin pregunta clara, calidad mínima y gobernanza, el proyecto acumula ruido.

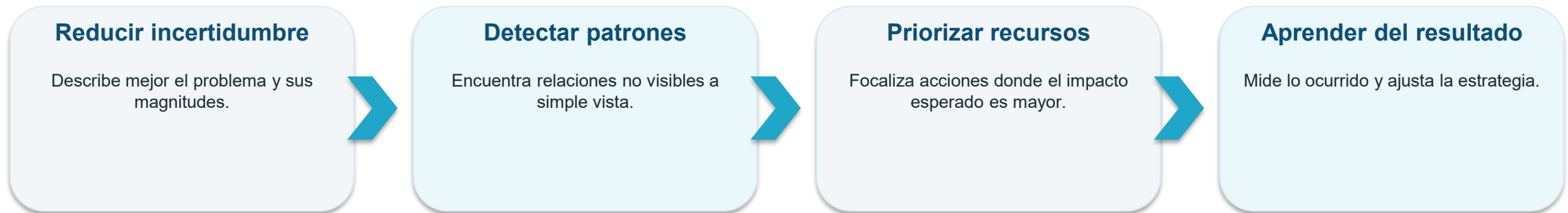


II. ¿Por qué la ciencia de datos mejora la toma de decisiones?

No reemplaza el criterio directivo o político. Lo fortalece con evidencia, trazabilidad y aprendizaje.



Cuatro mecanismos que mejoran una decisión



Idea central

La calidad de la decisión depende menos del algoritmo “más sofisticado” y más de la calidad del problema formulado, los datos y el uso organizacional de la evidencia.



De la analítica descriptiva a la prescriptiva

Descriptiva
¿Qué pasó?
Indicadores, tableros, cortes y mapas.

Diagnóstica
¿Por qué pasó?
Segmentación, correlaciones, análisis de cuellos de botella.

Predictiva
¿Qué podría pasar?
Pronósticos de demanda, riesgo o incumplimiento.

Prescriptiva
¿Qué conviene hacer?
Simulaciones, optimización, reglas y alertas.

La mayoría de organizaciones crea valor cuando consolida bien los niveles descriptivo y diagnóstico antes de saltar al predictivo.



Aplicaciones en el sector público

Focalización y padrones

Identificar hogares, territorios o beneficiarios con mayor probabilidad de vulnerabilidad o exclusión.

Gestión urbana y territorial

Usar datos para movilidad, limpieza, seguridad ciudadana, catastros, riesgos y equipamiento.

Recaudación y control

Detectar morosidad, subdeclaración, evasión o trámites con mayor fricción.

Salud y educación

Monitorear abandono, aprendizaje, cobertura, continuidad y demanda de servicios.

Riesgo de desastres

Anticipar exposición, capacidad de respuesta y prioridades de prevención.

Evaluación y transparencia

Comparar metas, resultados y brechas para corregir políticas.



Aplicaciones en el sector privado

Demanda y ventas

Pronósticos de ventas, surtido, elasticidades y planeamiento comercial.

Precio y rentabilidad

Pricing dinámico, promociones más efectivas y mezcla de productos.

Clientes y churn

Segmentación, valor de vida, abandono y campañas personalizadas.

Fraude y riesgo

Alertas tempranas para anomalías, crédito, lavado o incumplimientos.

Operaciones y logística

Rutas, inventarios, mermas, tiempos de entrega y cuellos de botella.

Talento y productividad

Rotación, ausentismo, clima y diseño de incentivos.



Condiciones habilitantes y límites reales

Sin estas condiciones, el proyecto se debilita

- Problema bien formulado y decisión concreta.
- Datos accesibles, documentados y con calidad mínima.
- Responsables claros del proceso y del uso del resultado.
- Capacidades analíticas y de comunicación.
- Espacio para pilotear, medir y corregir.

La ciencia de datos no resuelve por sí sola

- Falta de objetivos institucionales.
- Incentivos políticos u organizacionales "*mal*" alineados.
- Sesgos de registro y ausencia de datos clave.
- Resistencia al cambio.
- Problemas normativos, éticos o de interoperabilidad.

La pregunta correcta es: qué decisión específica mejorará y bajo qué restricciones.



III. Ética en la ciencia de datos

Toda decisión basada en datos redistribuye riesgos, recursos y oportunidades. Por eso necesita reglas.



Seis principios éticos que no se pueden omitir

Privacidad

Recoger y usar solo los datos necesarios, con base legal y protección adecuada.

Equidad

Revisar sesgos que perjudiquen grupos, territorios o perfiles específicos.

Transparencia

Explicar criterios, supuestos y límites de los modelos o indicadores.

Rendición de cuentas

Definir quién responde por el uso del modelo y sus consecuencias.

Seguridad

Prevenir accesos no autorizados, fugas, manipulación y uso indebido.

Supervisión humana

Reservar revisión humana en decisiones sensibles o de alto impacto.



Checklist mínimo de gobernanza de datos

Antes del proyecto

- Definir propósito y decisión
- Identificar base legal
- Evaluar sensibilidad del dato
- Documentar responsables
- Revisar sesgos ex ante

Durante el desarrollo

- Trazar fuentes y transformaciones
- Medir calidad y faltantes
- Validar con usuarios del dominio
- Registrar supuestos y métricas
- Probar robustez y explicabilidad

Después de desplegar

- Monitorear desempeño y deriva
- Medir impactos no deseados
- Habilitar revisión humana
- Actualizar datos y reglas
- Comunicar límites y uso permitido

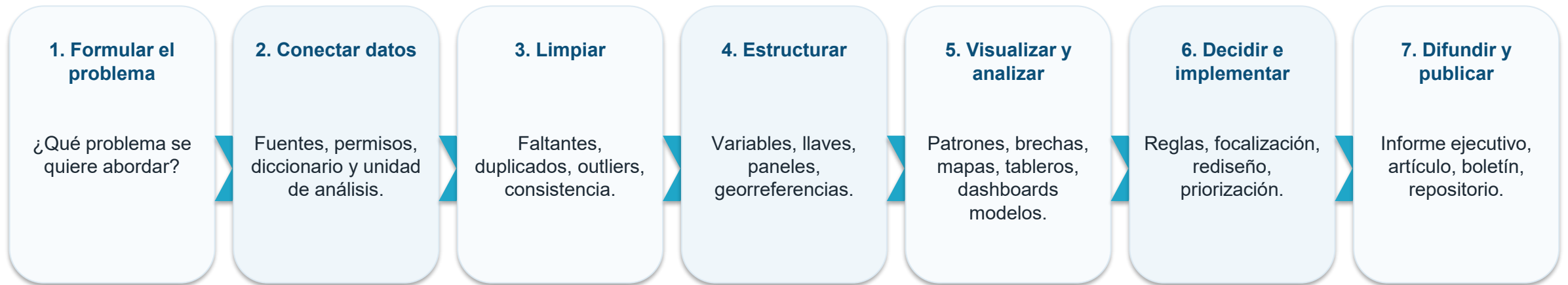


IV. Flujo integral de un proyecto de ciencia de datos

Desde la formulación del problema hasta la decisión, uso institucional y difusión de resultados.



Vista general del flujo

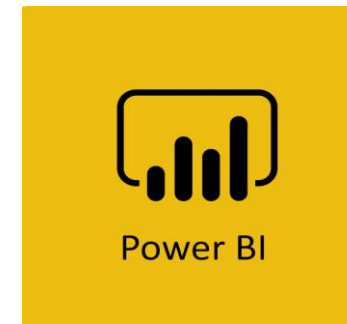


Principio de diseño

Cada etapa debe dejar un entregable verificable: pregunta, base de datos, script o protocolo de limpieza, visualización, recomendación y producto de difusión.



Herramientas Recomendadas





V. Caso aplicado: del problema al resultado publicable

Aplicación del flujo al documento sobre PDLC y brechas de planificación local en el Perú.



Caso aplicado: estudio sobre los Planes de Desarrollo Concertados (PDLC) en el Perú

Municipios que planifican, municipios que esperan: brechas de la Planificación del Desarrollo Local Concertado en el Perú

Luis Enrique Cayatopa-Rivera | Lima
Correo institucional: lcayatopa@ucss.edu.pe

I. Introducción

El Plan de Desarrollo Local Concertado (PDLC), no es un simple requisito administrativo. La Constitución Política del Perú, la Ley de Bases de la Descentralización, la Ley Orgánica de Municipalidades, la Ley Marco del Presupuesto Participativo, el Decreto Legislativo N.º 1088 y la Directiva N.º 001-2024-CEPLAN/PCD lo ubican como un instrumento de gestión territorial que ordena prioridades, orienta la inversión pública y vincula la acción municipal con la sociedad civil. La guía aprobada por la Resolución de Presidencia del Consejo Directivo N.º 094-2024/CEPLAN/PCD precisa, además, que el PDLC tiene un horizonte de diez años y debe articular zonas urbanas, rurales y territorios colindantes.

Por eso su elaboración compromete a varios actores del territorio. El proceso no recae solo en la municipalidad. También exige diálogo con el Consejo de Coordinación Local, organizaciones sociales, empresas, academia, entidades del gobierno nacional, gobiernos regionales y, cuando corresponde, cooperación u organismos internacionales con presencia local. Con base en lo anterior, Máttar y Cuervo (2017), Sandoval, Sanhueza y Williner (2015), Fernández, Fernández y Soloaga (2019), Amar y Erbes (2019), Burjel (2021) y sus coautores, coinciden en que el territorio concentra brechas, intereses y capacidades desiguales; por eso la planificación solo gana sentido cuando combina concertación, evidencia y lectura fina del espacio local.

Sobre esa base, el análisis de 1,891 municipalidades del país muestra una brecha importante¹. Solo 27.1% reporta contar con este instrumento y 72.9% declara no tenerlo. El rezago no afecta por igual a todos. Las municipalidades provinciales muestran una situación menos crítica que las distritales, mientras que las explicaciones de no tenencia apuntan sobre

¹ **Nota metodológica:** el artículo se elaboró a partir del Registro Nacional de Municipalidades 2025, publicado por el Instituto Nacional de Estadística e Informática (INEI). Se analizaron las variables P23_1 (tenencia del PDLC), P23_1_1 (último año de publicación), P23_1_3 (incorporación de gestión del riesgo), P23_1_4 (motivo de no tenencia), Tipomuni, Ubigeo, Departamento, C96 (ingresos totales) y C97 (gastos totales). Todos los cálculos, tablas, y mapas se elaboraron íntegramente con Python, mediante librerías para lectura de archivos .sav, depuración de datos, análisis estadístico y construcción de cartografía temática.

Pregunta del caso

¿Qué tan extendido está el PDLC en las municipalidades del Perú y qué brechas institucionales y fiscales aparecen?

Datos y unidad de análisis

Registro Nacional de Municipalidades 2025. Unidad de análisis: 1,891 municipalidades.

Métodos y productos

Depuración de variables, análisis descriptivo, tablas, medidas de mediana y cartografía temática >> Visualizaciones e informe de investigación.

Lectura gerencial: Un Dashboard o un boletín técnico puede convertirse en insumo de decisión y publicación.



El flujo del proyecto aterrizado al caso PDLC

Problema

Brecha de planificación territorial y dificultad para sostener el PDLC.

Conexión

Base RNM 2025 + tipo de municipio + ingresos/gastos + cartografía por ubigeo.

Limpieza

Estandarizar códigos, revisar faltantes, clasificar antigüedad del PDLC y motivos de no tenencia.

Estructura

Crear grupos comparables, medianas fiscales y categorías temporales del instrumento.

Visualización

Tablas por tipo de municipio, barras de motivos y mapa territorial.

Decisión

Focalizar asistencia técnica y financiamiento donde la brecha es mayor.

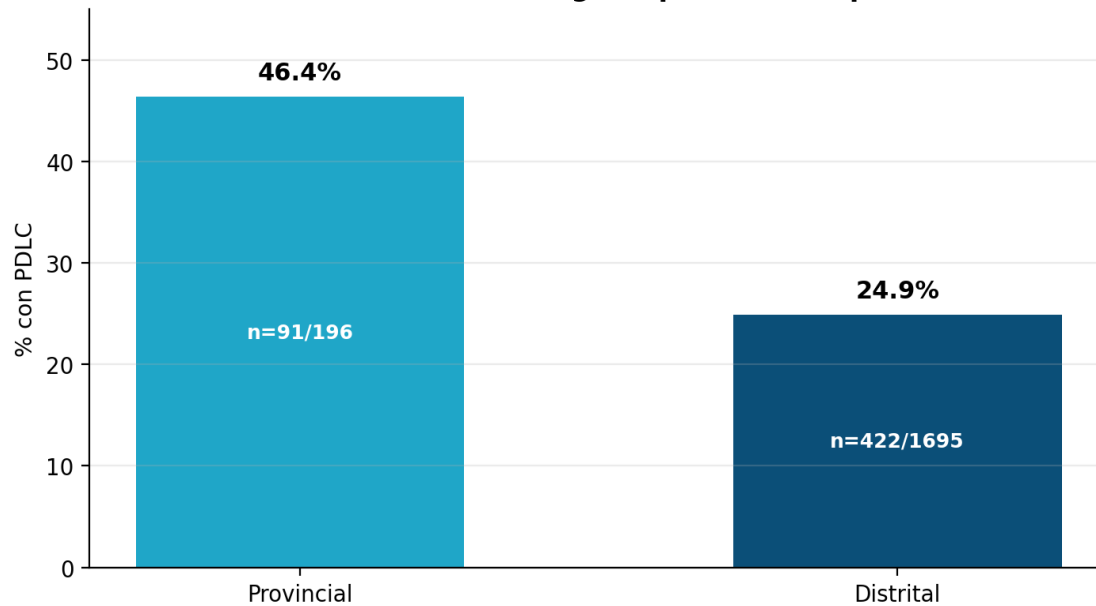
Difusión

Boletín, artículo, repositorio y presentación para actores públicos y académicos.

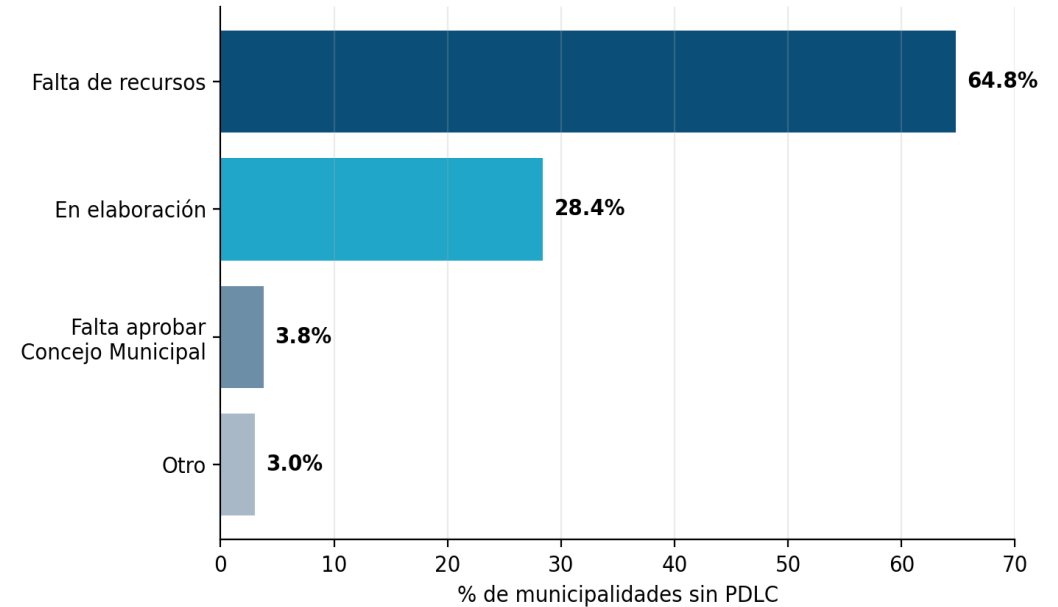


Hallazgos del caso: cobertura y restricciones

Cobertura del PDLC según tipo de municipalidad



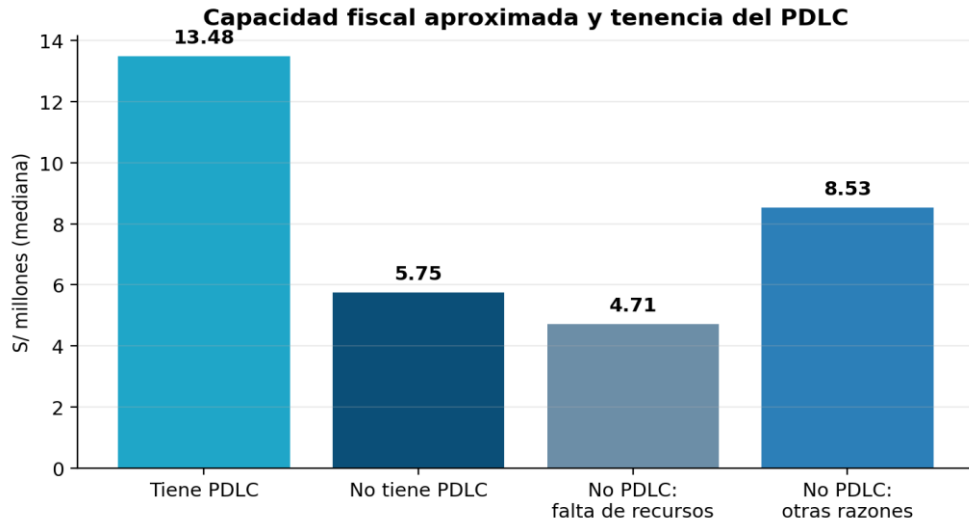
Razones reportadas para no contar con PDLC



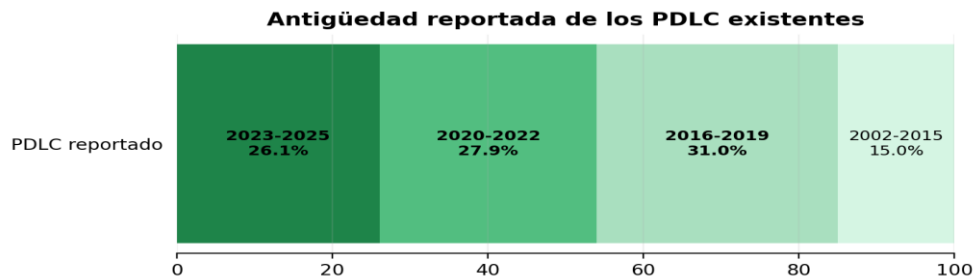
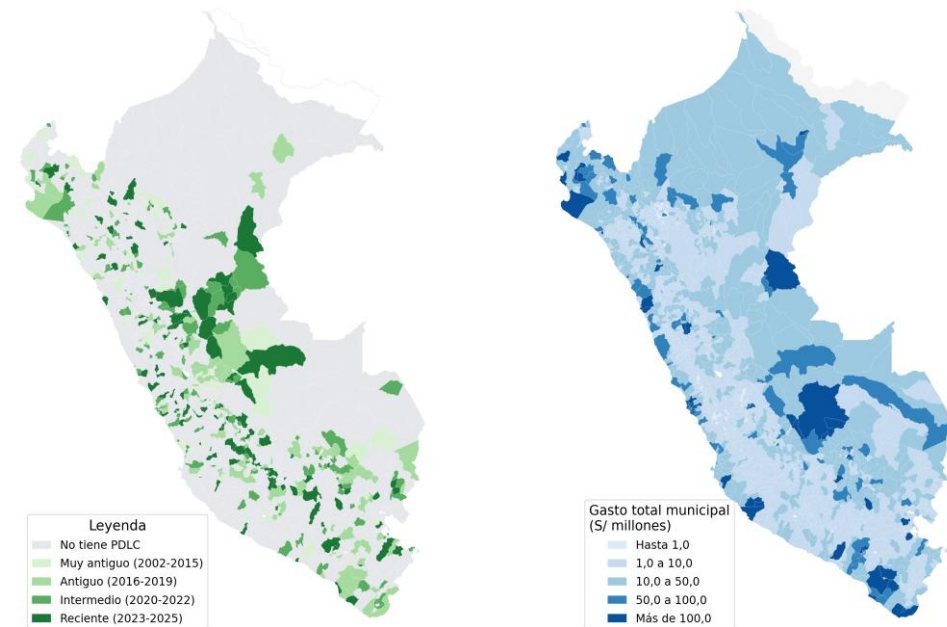
Lectura: la brecha es más severa en municipalidades distritales y la principal restricción declarada es presupuestal.



Hallazgos del caso: capacidad fiscal y territorio



Perú: PDLC, antigüedad reportada y distribución territorial del gasto público municipal, 2025
PDLC reportado y antigüedad del instrumento Distribución territorial del gasto público municipal



Lectura: menor capacidad fiscal se asocia con más dificultad para formular o actualizar el PDLC.



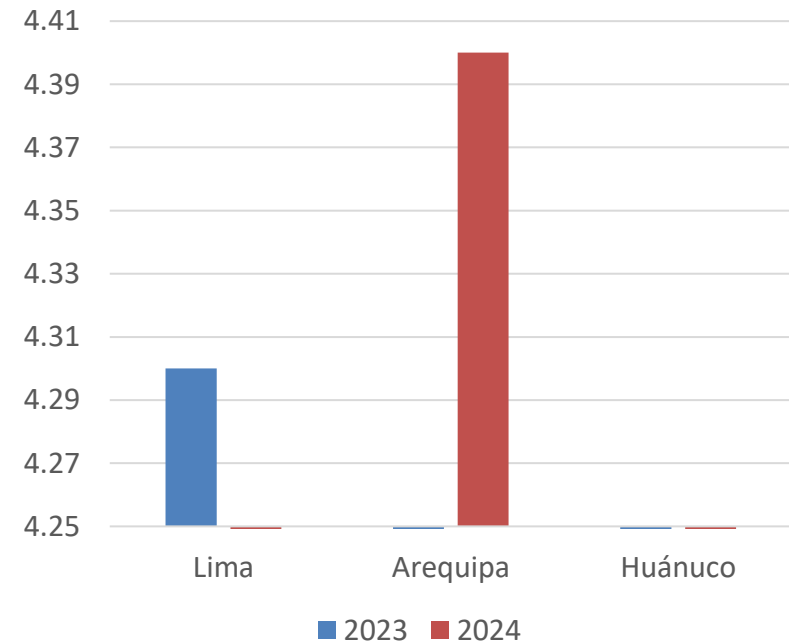
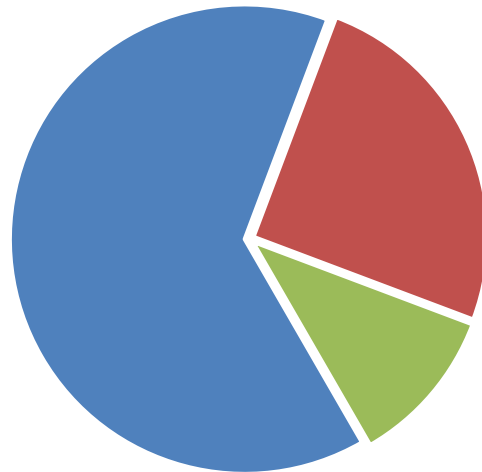
De la evidencia a la decisión

Actor	Decisión posible	Producto
Municipalidad	Priorizar asistencia técnica en distritos con menor capacidad fiscal y sin PDLC vigente.	Plan de trabajo, cronograma y presupuesto.
CEPLAN / nivel rector	Diseñar apoyo diferenciado y lineamientos de actualización según brecha territorial.	Guías, monitoreo y focalización.
Universidad	Acompañar diagnóstico, análisis espacial y evaluación con evidencia replicable.	Dashboard, boletín, paper, policy brief y capacitación.



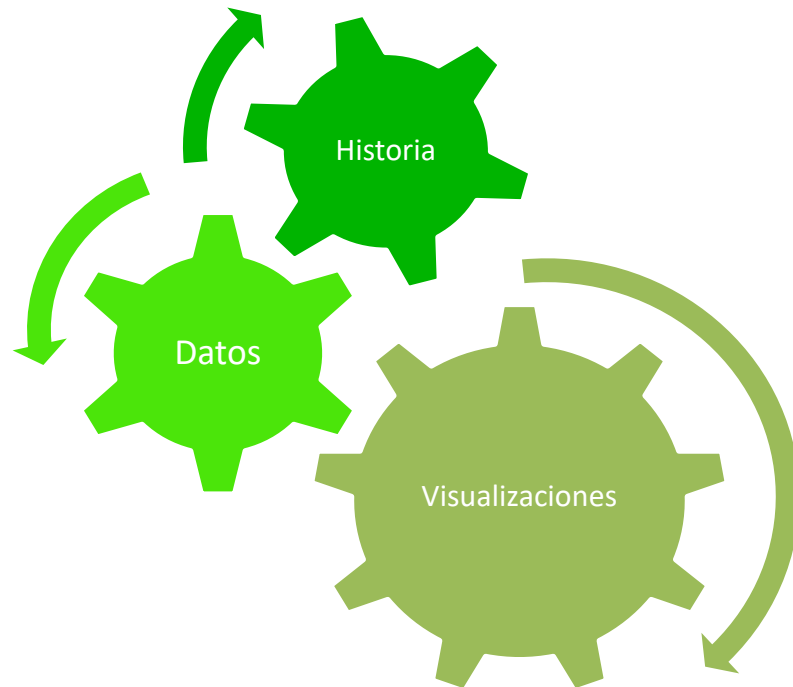
Errores comunes en el análisis de datos

- 1 Ejes mal escalados.
- 2 Falta de etiquetas o leyendas.
- 3 Sobrecarga de información.
- 2 Gráficos que distorsionan la verdad.





¿Qué es el Storytelling en BI?



Es el arte de **comunicar datos** de manera efectiva.

Combina **análisis** con **narrativa** para influenciar decisiones.



Visualización Narrativa

Usar dashboards como
escenarios narrativos.

Combinar visuales con
textos explicativos.

Resaltar lo importante
con color, forma y
jerarquía

Centrado en la mejora
continua.

Herramientas
recomendadas.





Cinco mensajes de cierre

1

La ciencia de datos no es una moda tecnológica. Es una disciplina para mejorar decisiones con evidencia.

2

Su valor aparece cuando conecta problema, dato, método, criterio del dominio y uso institucional.

3

Ética y gobernanza no son anexos. Son condiciones para usar datos de manera legítima y segura.

4

El flujo de trabajo importa tanto como el modelo: formular, conectar, limpiar, estructurar, visualizar, decidir y difundir.



Fuentes clave para la charla

Fuentes conceptuales y metodológicas

- Provost, F., & Fawcett, T. (2013). Data Science for Business.
- Donoho, D. (2017). 50 Years of Data Science.
- Laney, D. (2001). 3D Data Management.
- UNESCO (2021). Recommendation on the Ethics of Artificial Intelligence.
- NIST (2023). AI Risk Management Framework 1.0.

Fuentes del caso aplicado

- Cayatopa-Rivera, L. E. Documento aplicado sobre PDLC.
- Instituto Nacional de Estadística e Informática (2026)



Gracias

Preguntas, comentarios y discusión

Mgtr. Luis Enrique Cayatopa Rivera

Economista | Investigador Renacyt | UCSS



UNIVERSIDAD
CATÓLICA
SEDES SAPIENTIAE

FACULTAD DE
CIENCIAS ECONÓMICAS
Y COMERCIALES

Charla Magistral:

CIENCIA DE DATOS PARA LA TOMA DE DECISIONES



Jueves 23 de abril
19:10 a 20:40 h



Transmitido vía
zoom



PONENTE PRINCIPAL:

**Mgtr. LUIS ENRIQUE
CAYATOPA RIVERA**

Economista e Investigador experto
en analítica y gestión pública.



MODERACIÓN:

MODESTA GONZÁLEZ SAÉNZ

Coordinadora del Centro de
Investigación y Desarrollo Innovador
para la Regionalización (CIDIR)



Escanea el QR e
**INSCRÍBETE
AQUÍ**



Dirigido a: Estudiantes y docentes de la Facultad de Ciencias Económicas y Comerciales (FCEC) de la UCSS.