

Análisis geoespacial y temporal del valor patrimonial del mobiliario y equipamiento educativo en la Región San Martín mediante procesos ETL y analítica visual

Estudiantes:

Gabriel Quiroz Gómez

Miguel Vega Soldevilla

Universidad Nacional de Ingeniería, Facultad de Ciencias,
e-mail: gabrielquiroz@uni.pe e-mail: miguel.vega.s@uni.pe

Curso:

CC531A - Análisis en Macrodatos
Laboratorio 02

Resumen

Este trabajo analiza el “Registro de Mobiliario y Equipamiento de Instituciones Educativas de la Región San Martín (GRSM)” publicado en la Plataforma Nacional de Datos Abiertos (PNDA). Se implementó un flujo ETL y un conjunto de consultas en Python (pandas, numpy, matplotlib) para explorar la distribución, evolución temporal y estado patrimonial de los bienes educativos, y se diseñó un dashboard en Power BI para análisis interactivo, incluyendo un mapa geográfico basado en coordenadas LATITUD/LONGITUD. Los resultados muestran la concentración del valor patrimonial por provincia y tipo de gestión, tendencias de inversión a lo largo del tiempo y señales de depreciación que orientan decisiones de mantenimiento o renovación.

Keywords: Educación; inventarios patrimoniales; PNDA; ETL; pandas; matplotlib; Power BI; mapas; depreciación.

Índice

1	Introducción	2
2	Marco Teórico	2
3	Metodología	2
3.1	Dataset y fuente	2
3.2	ETL (Extract, Transform, Load)	3
3.2.1	Extract	3
3.2.2	Transform	3
3.2.3	Load	3
3.2.4	ETL en Power BI	3
3.3	Consultas y visualizaciones	3
3.3.1	Python	3
3.3.2	PowerBI	4
4	Discusiones y resultados	4

1. Introducción

La gestión patrimonial educativa requiere información confiable y trazable sobre bienes inventariados: su adquisición, depreciación, ubicación y administración. El dataset “Registro de Mobiliario y Equipamiento de Instituciones Educativas de la Región San Martín – GRSM” permite analizar estas dimensiones con granularidad a nivel de sede educativa. Este informe integra analítica con Python y visualización en Power BI para responder a la pregunta: ¿Cómo se distribuye y evoluciona el valor patrimonial de los bienes educativos en la Región San Martín según la gestión y el estado de conservación, y qué patrones geográficos se observan?

2. Marco Teórico

Los datos representan objetos físicos (bienes patrimoniales) que se adquieren, deprecian y asignan a sedes educativas. Conceptualmente:

- Entidades: bienes (COD_PATRIMONIAL), sedes (NOMBRE_SEDE, COD_MODULAR), unidades territoriales (PROVINCIA, DISTRITO), gestión (D_GESTION), y tiempos (FECHA_COMPRA, FECHA_ALTA, FECHA_MOV).
- Atributos cuantitativos: HVALOR_INICIAL, HDEPR_INICIAL, HDEPR_EJERCICIO, VALOR_NETO.
- Dinámica: depreciación contable (valor en el tiempo), movimientos (TIPO_MOV, TIPO_TRAN), estado físico (ESTADO_CONSER).
- Geografía: LATITUD, LONGITUD permiten análisis espacial y mapas.

Reglas y supuestos típicos:

- Depreciación reduce el VALOR_NETO desde HVALOR_INICIAL ($\text{VALOR_NETO} \leq \text{HVALOR_INICIAL}$).
- Fechas coherentes ($\text{FECHA_COMPRA} \leq \text{FECHA_ALTA}$; FECHA_MOV dentro del ciclo de vida).
- Ubicaciones válidas (LATITUD $\in [-90, 90]$, LONGITUD $\in [-180, 180]$).
- La gestión (D_GESTION) influye en patrones de adquisición y conservación.

3. Metodología

3.1. Dataset y fuente

- Nombre: Registro de Mobiliario y Equipamiento de Instituciones Educativas – Región San Martín (GRSM).
- Organización publicadora: Gobierno Regional San Martín (GRSM) en la PNDA.
- Cobertura temática: Educación.
- Estructura de campos (extracto relevante):
 - Identificación y descripción: COD_PATRIMONIAL, DESCRIPCION, NOMBRE.
 - Temporalidad: FECHA_COMPRA, FECHA_MOV, FECHA_ALTA, FECHA_CORTE.
 - Contabilidad: HVALOR_INICIAL, HDEPR_INICIAL, HDEPR_EJERCICIO, VALOR_NETO.
 - Estado/gestión: ESTADO_CONSER, D_GESTION, D_GES_DEP, D_ESTADO.
 - Geografía: DEPARTAMENTO, PROVINCIA, DISTRITO, LATITUD, LONGITUD, UBIGEO.
- Fuente PNDA (URL): <https://datosabiertos.gob.pe/dataset/registro-mobiliario-y-equipamiento-de-instituciones-educativas-de-la-regi>
- Fecha de publicación/actualización: May 29, 2025

3.2. ETL (Extract, Transform, Load)

3.2.1. Extract

- Lectura robusta del CSV con detección de separador y tolerancia a líneas irregulares.
- Normalización de nombres de columnas.

3.2.2. Transform

- Conversión de fechas (FECHA_COMPRA, FECHA_MOV, FECHA_ALTA, FECHA_CORTE) a `datetime`.
- Conversión de numéricos (HVALOR_INICIAL, HDEPR_INICIAL, HDEPR_EJERCICIO, VALOR_NETO, LATITUD, LONGITUD).
- Derivadas: AÑO_COMPRA, MES_COMPRA, TRIM_COMPRA; $\text{PORC_DEPRECIACION} = 1 - \text{VALOR_NETO}/\text{HVALOR_INICIAL}$.
- Validaciones básicas y control de infinitosNaN en ratios.

3.2.3. Load

- Persistencia del DataFrame limpio y export de resultados (tablas de consultas) a CSV para usar en el informe y Power BI.

3.2.4. ETL en Power BI

Durante la importación del dataset en Power BI se presentó un **error de conversión de tipos** al intentar tipar `COD_PATRIMONIAL` como `Int64.Type`. La columna llegaba en *notación científica* y/o como texto (p. ej., `‘‘1,1228E+11’’`), lo que generó filas con error al no poder convertirse a entero. La solución consistió en **cambiar el tipo de dato de COD_PATRIMONIAL a texto**, preservando el código íntegro y eliminando el error de carga.

Transformaciones aplicadas

- **Corrección de tipos**
 - `COD_PATRIMONIAL` → **Texto** (evitar errores por notación científica y mantener el código completo).
 - Columnas de **fecha** (`FECHA_COMPRA`, `FECHA_MOV`, `FECHA_ALTA`, `FECHA_CORTE`) → **Fecha** (habilita agregaciones temporales).
 - `VALOR_NETO`, `HVALOR_INICIAL`, `HDEPR_INICIAL`, `HDEPR_EJERCICIO` → **Número decimal** (necesario para sumas y operaciones).
- **Columnas derivadas para series temporales**
 - `Anio_Compra` y `Mes_Compra` a partir de `FECHA_COMPRA`, para gráficos por año/mes.
- **Formato de visualización**
 - Formato **monetario** (S/) en medidas de valor (p. ej., `VALOR_NETO`, `HVALOR_INICIAL`) para consistencia con la moneda local.

3.3. Consultas y visualizaciones

3.3.1. Python

- Conteo de bienes por provincia : Tabla
- Valor neto total de los bienes por nivel educativo: Tabla
- Valor neto promedio por estado de conservación: Tabla
- Valor total de bienes adquiridos por año: Tabla

- Distribución del estado de conservación): Gráfico de barras.
- Evolución anual del valor neto total: Gráfico de líneas.
- Tendencia acumulativa del Valor total de bienes por tipo de gestión (2000 - 2025) : Gráfico de tendencias acumulativas.
- Distribución de Bienes por Área Censal (Urbano vs Rural): Gráfico circular.
- Distribución de Bienes por Nivel Educativo: Gráfico circular.
- Matriz de correlación (variables: HVALOR.INICIAL, HDEPR.INICIAL, HDEPR.EJERCICIO, VALOR.NETO, AÑO.COMPRA).

3.3.2. PowerBI

- Valor total de bienes por provincia: Columnas apiladas (vertical)
- Cantidad de bienes por nivel educativo: Columnas apiladas (horizontal)
- Tendencia de bienes adquiridos por año: Gráfico de líneas
- Estado de Conservación de los Bienes Educativos: Circular
- Distribución geográfica de inversiones en instituciones educativas de la Región San Martín: Gráfico de mapas
- Distribución geográfica del valor neto del equipamiento educativo: Gráfico de dispersión

4. Discusiones y resultados

Pregunta a responder: ¿Qué provincias concentran la mayor cantidad de bienes inventariados?

PROVINCIA	
BELLAVISTA	62
EL DORADO	278
HUALLAGA	1121
LAMAS	77638
MARISCAL CACERES	82
MOYOBAMBA	94497
PICOTA	645
RIOJA	36853
SAN MARTIN	212
TOCACHE	933
Name: COD_PATRIMONIAL, dtype: int64	

Figura 1: Tabla - Conteo de bienes por provincia

- Las provincias que concentran la mayor cantidad de bienes inventariados son Moyobamba, Lamas, Rioja.

Pregunta a responder: ¿Qué nivel educativo acumula el mayor valor neto de patrimonio y en qué magnitud?

```
D_NLV_MODULAR
Básica Alternativa-Avanzado      11750.10
Básica Alternativa-Inicial e Intermedio  11707.57
Básica Especial                  8389.95
Básica Especial-Inicial          5759.28
Básica Especial-Primaria        39520.56
Escuela Superior Pedagógica     348853.31
Inicial - Cuna-Jardín           151189.81
Inicial - Jardín                2660875.73
Inicial - Programa no escolarizado  23087.16
Instituto Superior Pedagógico    207127.51
Instituto Superior Tecnológico   665189.50
Primaria                        17767144.73
Secundaria                      17675388.05
Técnico Productiva              66695.09
Name: VALOR_NETO, dtype: float64
```

Figura 2: Tabla - Valor neto total de los bienes por nivel educativo

- El nivel educativo que acumula el mayor valor neto de patrimonio es Primaria con un valor de 17767144,73.

Pregunta a responder: ¿Cuál es el valor neto promedio de los bienes según su estado de conservación (bueno, regular, malo)?

```
ESTADO_CONSER
Bueno      252.179794
Malo       148.040644
Nuevo     229.039950
Regular     67.382282
Name: VALOR_NETO, dtype: float64
```

Figura 3: Tabla - Valor neto promedio por estado de conservación

- El valor neto promedio de bienes que se encuentra en buen estado es de 252,179794.
- El valor neto promedio de bienes que se encuentra en mal estado es de 148,040644.
- El valor neto promedio de bienes que se encuentra en un estado nuevo es de 229,039950.
- El valor neto promedio de bienes que se encuentra en buen estado es de 67,382282.

Pregunta a responder: ¿En qué años se realizaron las mayores inversiones en bienes?

Valor total de bienes adquiridos por año:

Anio_Compra	
1900	1040.00
1905	700.00
1961	5297.00
1965	2699.00
1967	170.00
1968	25.00
1970	5519.00
1974	21500.00
1975	770.00
1976	9466.00
1978	45.00
1979	2443.00
1980	46500.00
1981	10391.00
1982	1310.00
1983	6020.00
1984	10404.00
1985	1910.00
1986	11382.00

Figura 4: Tabla - Valor total de bienes adquiridos por año

- El año con mayor inversión total en bienes fue 2021 con un total de 15277852,34.

Pregunta a responder: ¿Cuál es la distribución de bienes por estado de conservación y qué categoría predomina?

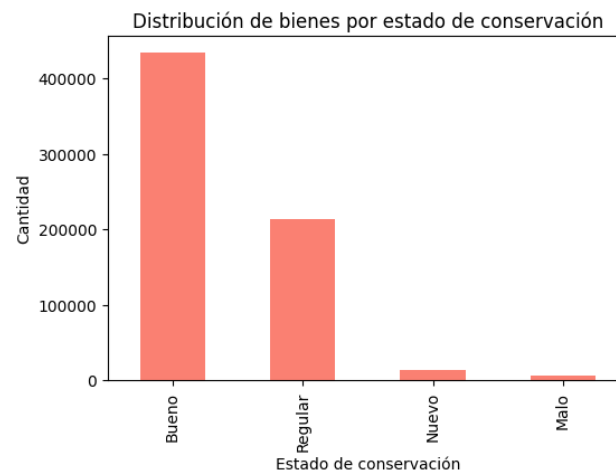


Figura 5: Gráfico de barras - Distribución del estado de conservación

- La categoría de estado de conservación que predomina es 'Bueno'.

Pregunta a responder: ¿Cómo ha evolucionado el valor neto total del patrimonio año a año?



Figura 6: Gráfico de líneas - Valor total de bienes adquiridos por año

- Se observa una alta diferencia en el valor neto total del patrimonio entre los años 2020 y 2021.

Pregunta a responder: ¿Qué tipo de gestión acumula más inversión en el periodo 2000–2025 y con qué ritmo?

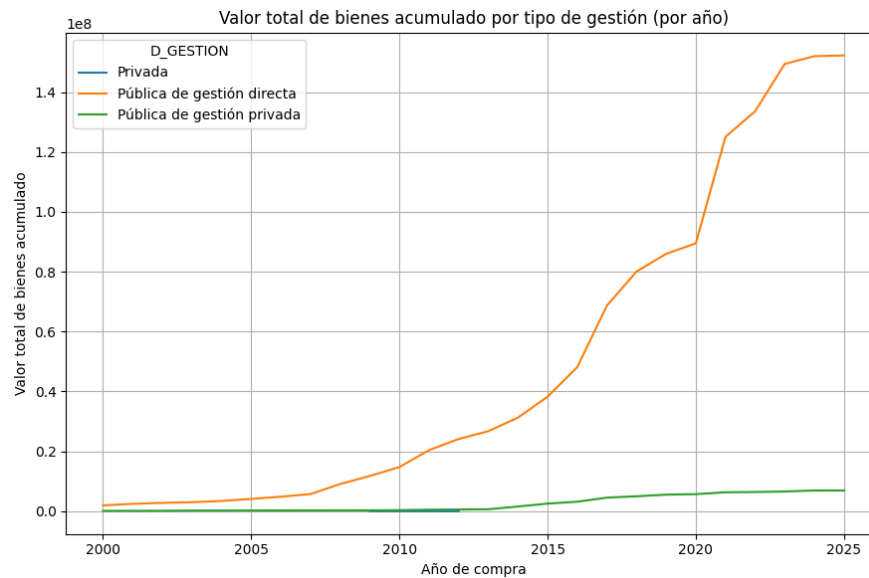


Figura 7: Gráfico de tendencias acumulativas - Valor total de bienes acumulados por tipo de gestión (2000 - 2025)

- El tipo de gestión que acumula más inversión en el periodo 2000–2025 es 'Pública de gestión directa' a un ritmo creciente acelerado.

Pregunta a responder: ¿Qué proporción de bienes está ubicada en zonas urbanas frente a rurales?

Distribución de Bienes por Área Censal (Urbano vs Rural)

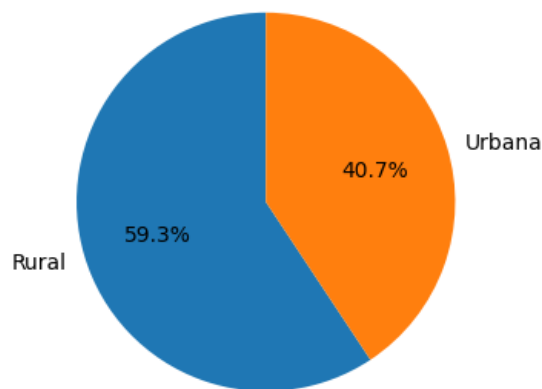


Figura 8: Gráfico circular: Distribución de Bienes por Área Censal

- La proporción de bienes en zonas urbanas es de un 40.7 % frente a un 59.3 % en zonas rurales.

Pregunta a responder: ¿Cómo se reparte el inventario entre los distintos niveles educativos?

Distribución de Bienes por Nivel Educativo ($\geq 3\%$ + Otros)

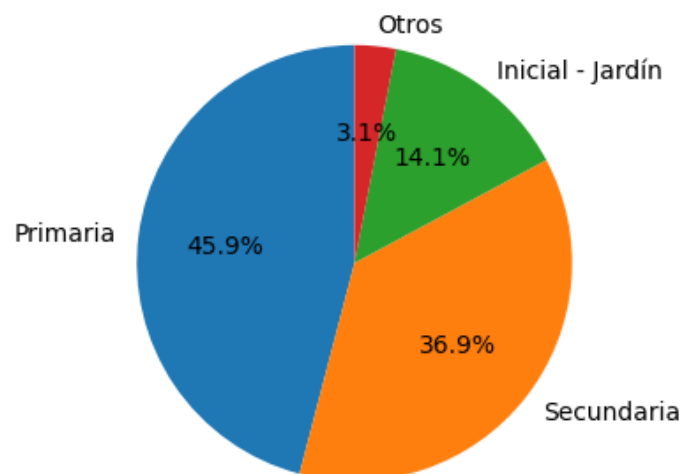


Figura 9: Distribución de Bienes por Nivel Educativo

- El nivel Primaria, Secundaria e Inicial - jardín presentan un 45.9 % , 36.9 % y 14.1 % de los bienes respectivamente. Mientras que otros tipos de niveles educativos presentan solo el 3.1 %.

Pregunta a responder: ¿Qué relaciones (positivas/negativas) existen entre el valor inicial, la depreciación, el valor neto y el año de compra?

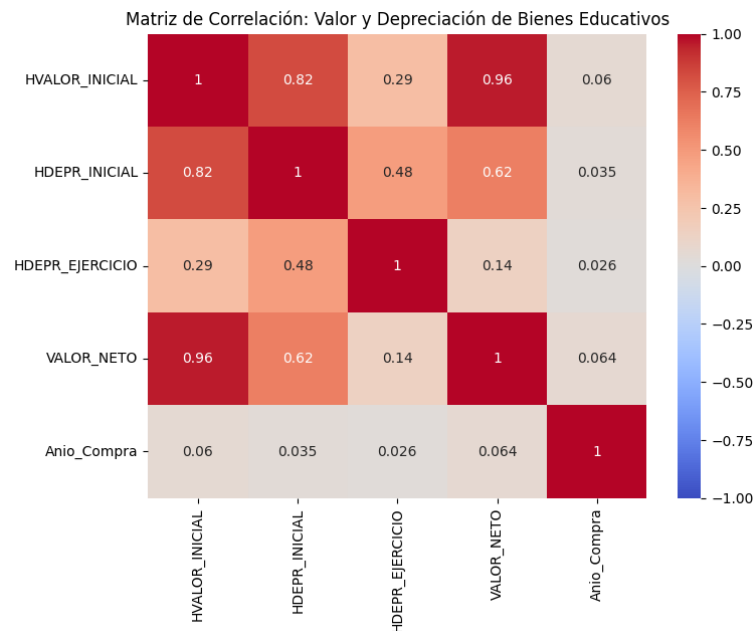


Figura 10: Matriz de correlación (5 variables)

- La matriz muestra relaciones positivas altas entre el valor inicial y el valor neto ($r = 0,956$) y entre valor inicial y depreciación acumulada ($r = 0,825$), lo cual es consistente con el hecho de trabajar con montos absolutos (los bienes de mayor valor inicial tienden a conservar mayor valor neto y acumular mayor depreciación). La depreciación del ejercicio presenta asociaciones más débiles con estos montos, y el año de compra no exhibe correlación lineal relevante con las variables analizadas ($|r| \approx 0,06$).

Pregunta a responder: ¿Cómo se compara el valor total de los bienes entre provincias y qué provincia presenta la mayor inversión total de bienes?



Figura 11: Power BI - Valor total de bienes por provincia

- La provincia que presenta la mayor inversión total de bienes es Lamas.

Pregunta a responder: ¿Qué nivel educativo presenta mayor cantidad de bienes?

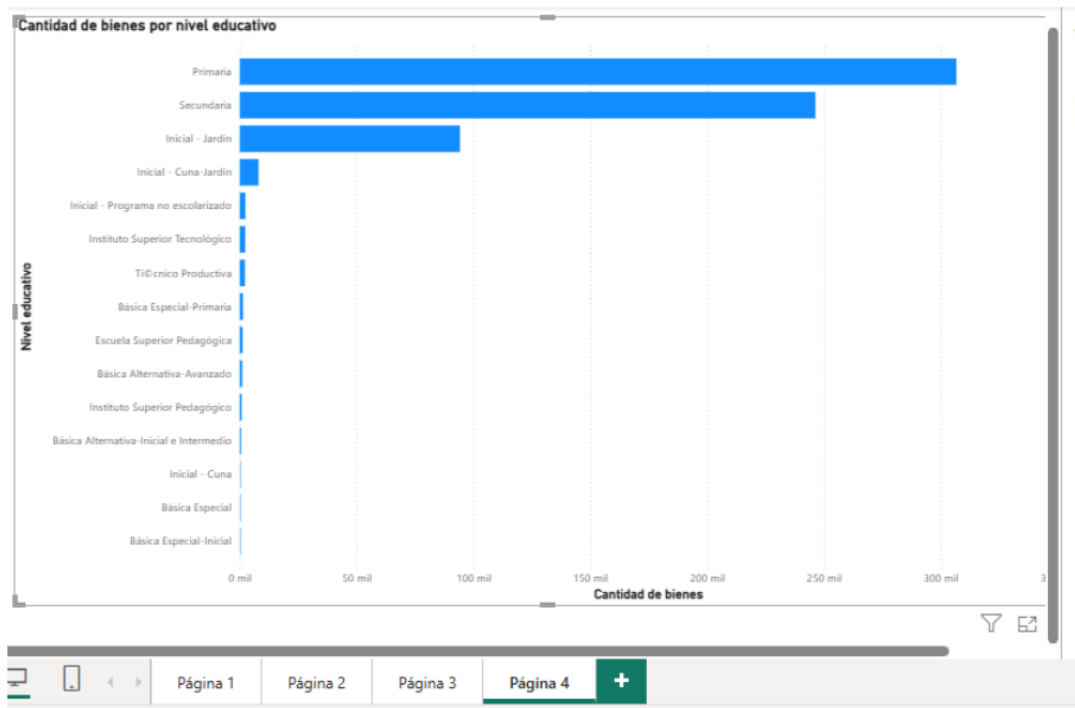


Figura 12: Power BI - Cantidad de bienes por nivel educativo

- El nivel educativo Primaria presenta mayor cantidad de bienes.

Pregunta a responder: ¿Cuál ha sido la tendencia en el número de bienes adquiridos a lo largo del tiempo?

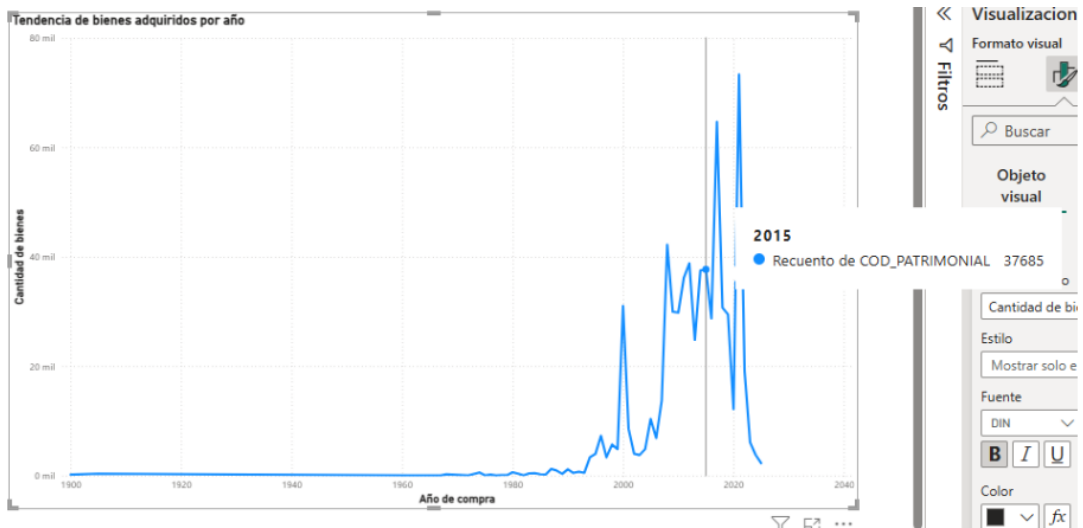


Figura 13: Power BI - Tendencia de bienes adquiridos por año

- Se observa una mayor cantidad de bienes adquiridos en el año 2021.

Pregunta a responder: ¿Qué porcentaje del inventario se encuentra en cada estado de conservación?

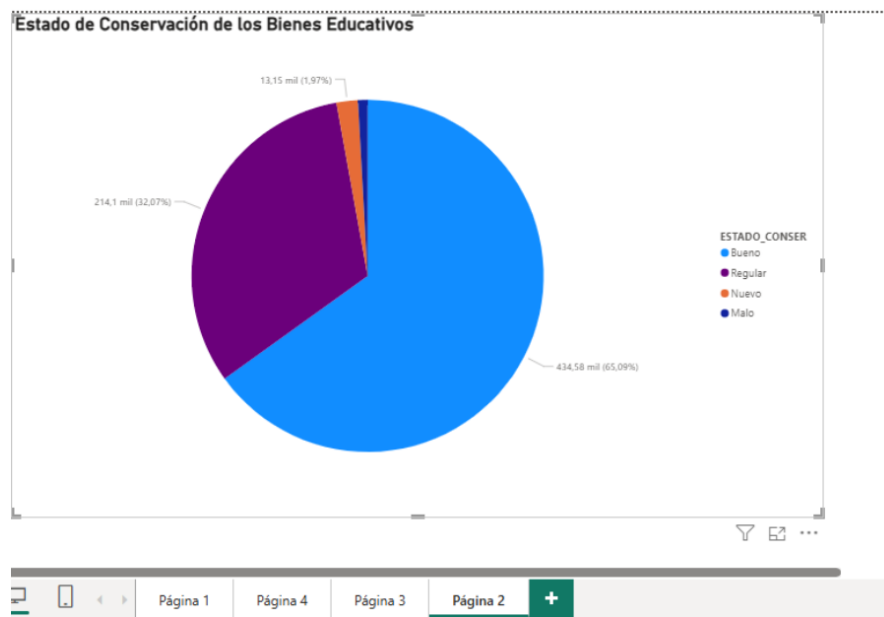


Figura 14: Power BI - Estado de Conservación de los Bienes Educativos

- Se observa que el 65.09 % de bienes se encuentra en buen estado, el 32.07 % se encuentra en estado regular y el 1.97 % se encuentra en estado nuevo.

Pregunta a responder: ¿Dónde se localizan geográficamente la mayor cantidad de inversiones?

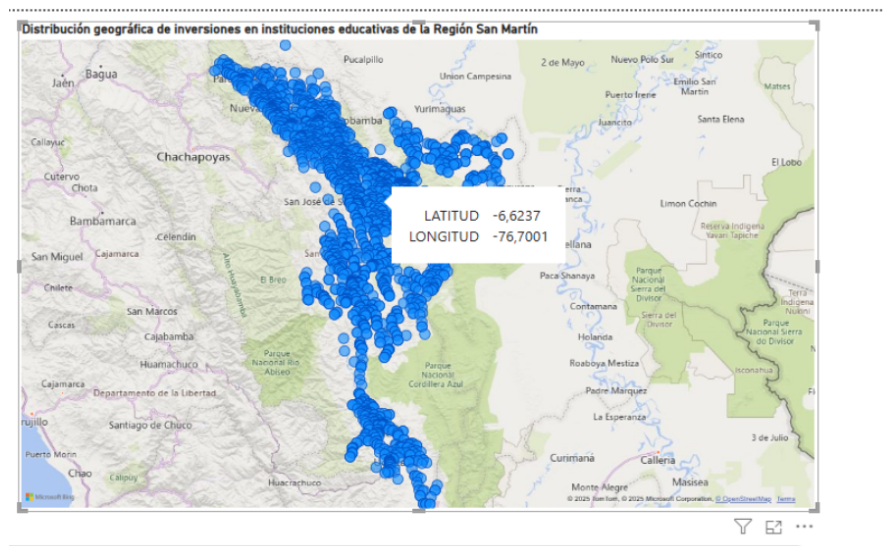


Figura 15: Power BI - Gráfico de mapas

- La mayor cantidad de inversiones se localiza en la región de Moyobamba.

Pregunta a responder: ¿Qué concentración y dispersión presenta el valor neto del equipamiento en el territorio y dónde se hallan los principales focos?

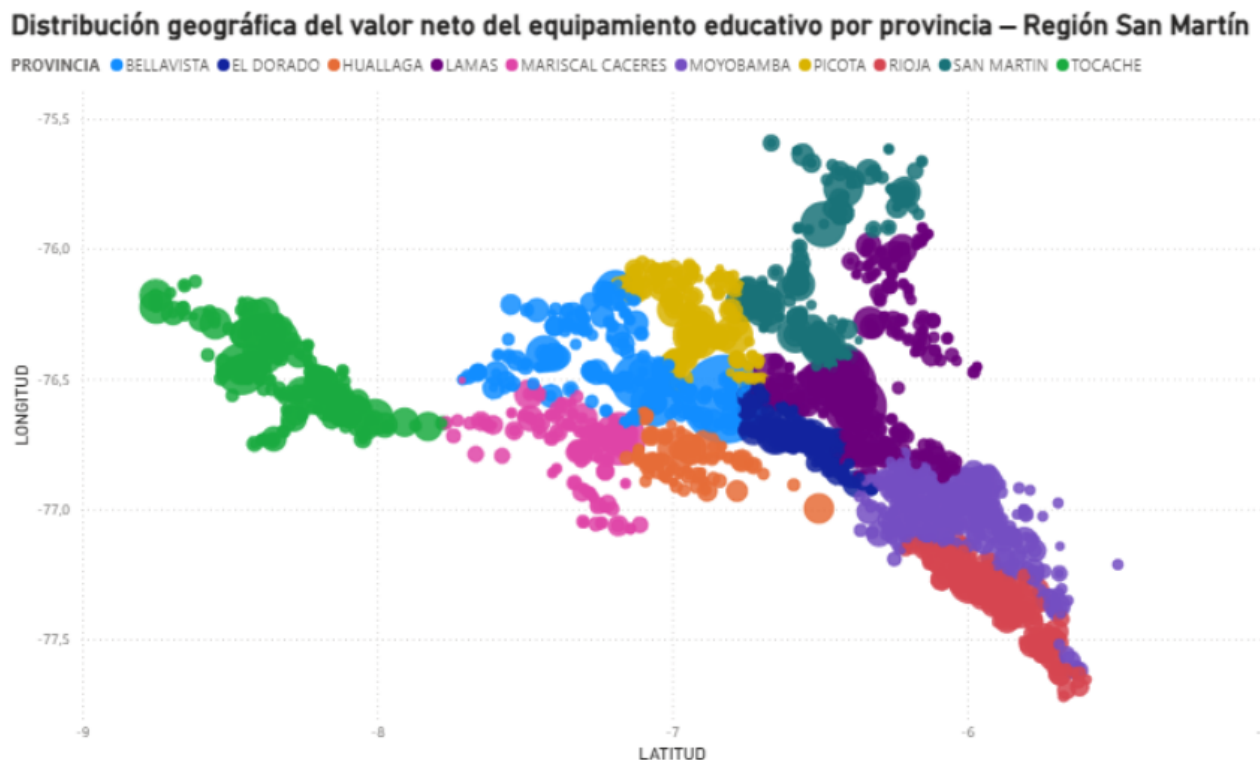


Figura 16: Power BI - Gráfico de dispersión

- Alta concentración en el corredor centro-este de la región (latitudes $\approx -6,8a-6,0$, longitudes $\approx -77,3a-76,5$). Ahí se observan nubes densas de puntos superpuestos y con varias burbujas grandes.
- Focos principales: Núcleos de burbujas grandes y muy próximas en las zonas Moyobamba–Rioja–Lamas. Indican acumulación de mayor valor neto en un radio relativamente pequeño.
- Dispersión más marcada hacia los bordes oeste y sur del mapa. Los puntos aparecen más esparcidos y con burbujas, en promedio, más pequeñas, lo que sugiere menor valor agregado por sede y mayor dispersión territorial.

5. Conclusiones

- El dataset permite una visión integral del patrimonio educativo en GRSM: qué bienes existen, dónde están, cuánto valen y en qué estado se encuentran.
- Se identificaron concentraciones geográficas y patrones temporales de compra que pueden orientar la planificación presupuestal.
- La combinación de ETL + Python + Power BI habilita monitoreo continuo y cuadros de mando reproducibles.

Bibliografía y referencias

- [1] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, et al. Array programming with numpy. *Nature*, 585:357–362, 2020.

- [2] John D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.
- [3] Gobierno Regional San Martín. Registro de mobiliario y equipamiento de instituciones educativas de la región san martín – grsm. Plataforma Nacional de Datos Abiertos (PNDA), 2025. Publicado: 29 May 2025.
- [4] Wes McKinney. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference (SciPy 2010)*, pages 51–56, 2010.
- [5] Microsoft. Create maps in power bi, n.d.
- [6] Microsoft. Power bi documentation, n.d.
- [7] Python Software Foundation. Python language reference, n.d.
- [8] The pandas development team. pandas-dev/pandas: Pandas, 2020. Software.
- [9] Michael L. Waskom. Seaborn: Statistical data visualization. *Journal of Open Source Software*, 6(60):3021, 2021.