

**Instituto Tecnológico y de Estudios Superiores de Monterrey
Campus Monterrey**

Posgrados



**Tecnológico
de Monterrey**

Avance 1. Análisis exploratorio de datos

Integrantes del equipo de trabajo

Carlos Daniel Villena Santiago A01795127

Edwin David Hernandez Alejandro A01794692

Gustavo Andres Garcia Anguiano A01795493

28 de septiembre del 2025

1. Introducción

El presente documento muestra un análisis exploratorio de los **screenshots** generados a partir del PDF del proyecto **Edificio Tello (E-4034-21-05-2025)**.

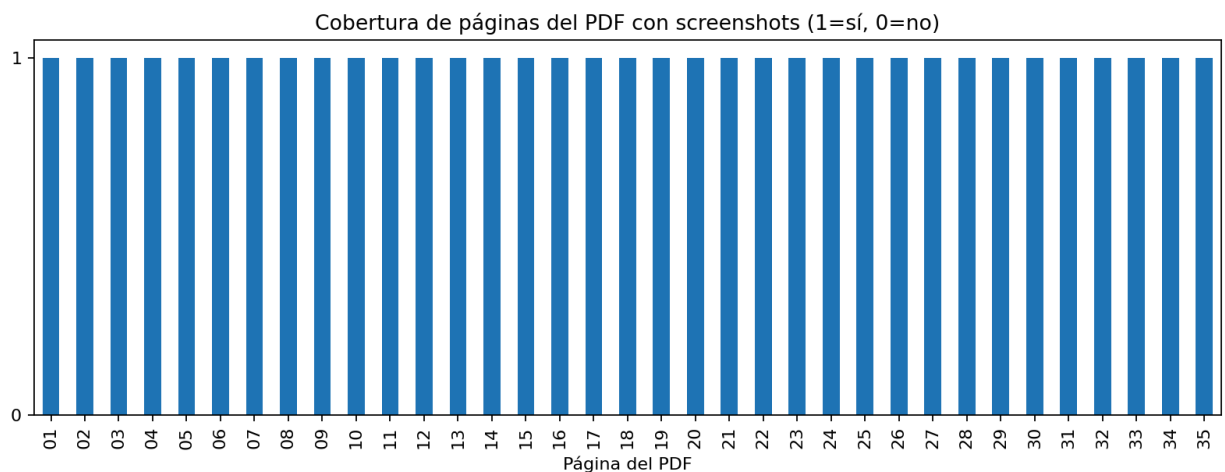
Cada carpeta (**01**, **02**, **03**, ..., **35**) representa las imágenes asociadas a una página del PDF original.

El objetivo es **identificar patrones** en cantidad de imágenes, resolución (ancho y alto), peso en disco y formatos utilizados, para tener una visión global de la documentación generada.

2. Cobertura de páginas

Se verificó si todas las páginas del PDF cuentan con screenshots asociados.

Figura 1. Cobertura de páginas del PDF con screenshots



Hallazgos:

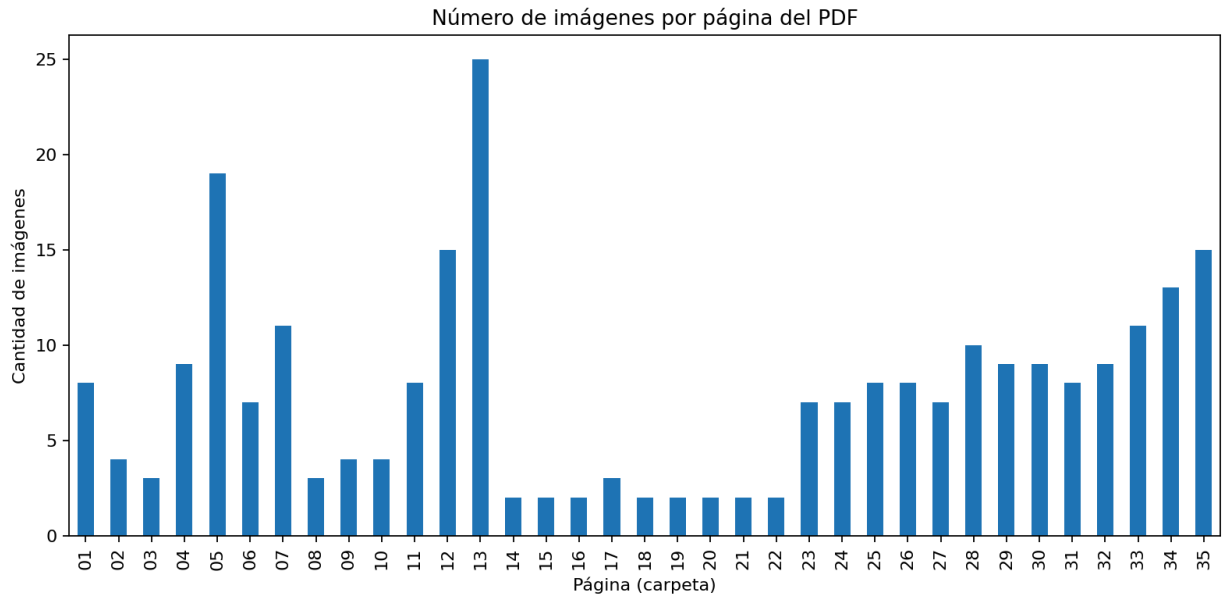
- El 100% de las páginas del PDF cuentan con al menos un screenshot.
- Esto garantiza que no hay pérdida de información en la conversión o captura.

3. Número de imágenes por página

Se analizó la cantidad de imágenes almacenadas en cada carpeta/página.



Figura 2. Número de imágenes por página del PDF



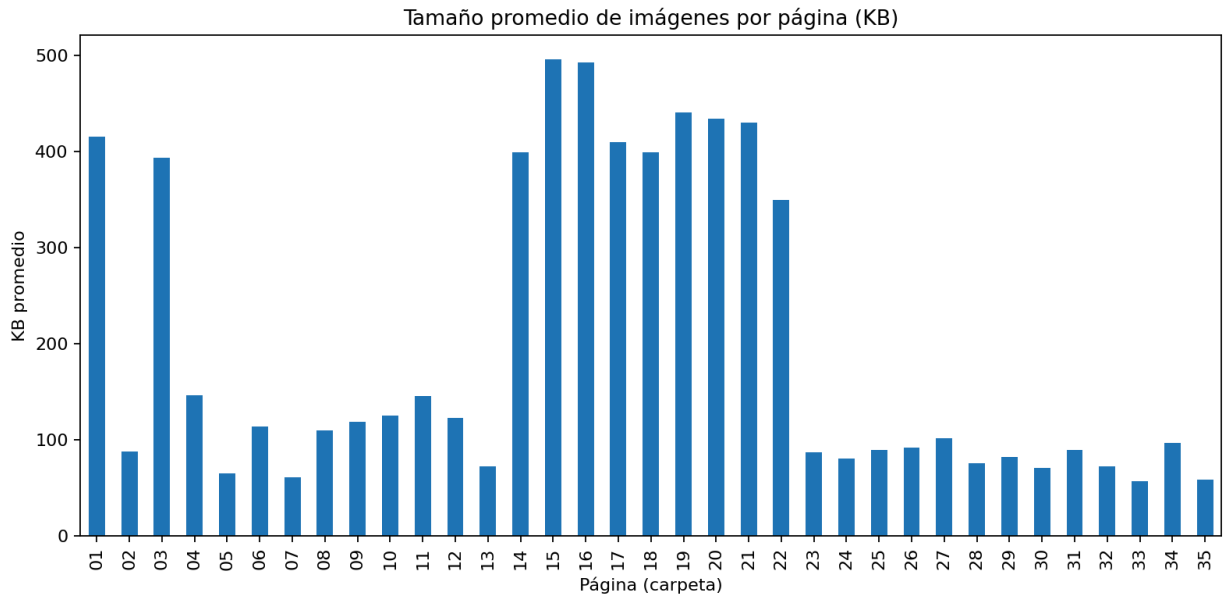
Hallazgos:

- La distribución no es uniforme: algunas páginas cuentan con más de 20 imágenes (páginas 12 y 13), mientras que otras tienen menos de 5.
- Esto indica que hay páginas con mayor nivel de detalle gráfico o subdivisión en múltiples capturas.

4. Tamaño promedio de imágenes

Se midió el tamaño promedio de archivo (en KB) por página.

Figura 3. Tamaño promedio de imágenes por página (KB)



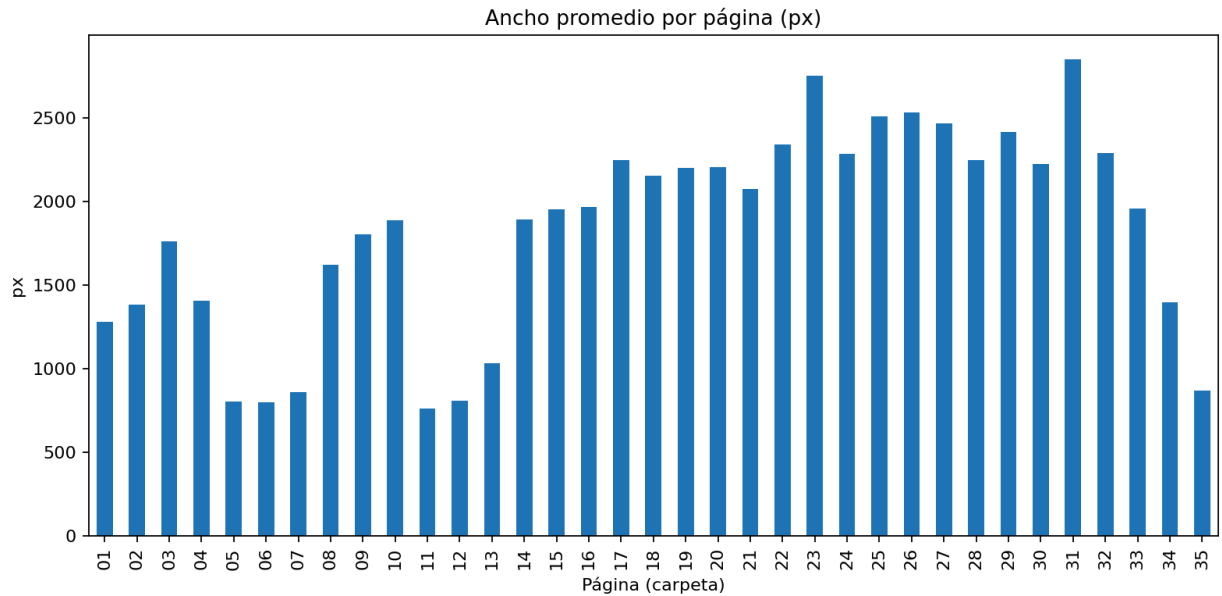
Hallazgos:

- Las páginas 15 a 21 concentran imágenes de mayor tamaño, con promedios cercanos a 500 KB.
- Esto sugiere mayor densidad de información gráfica o resolución más alta.

5. Resolución promedio de imágenes

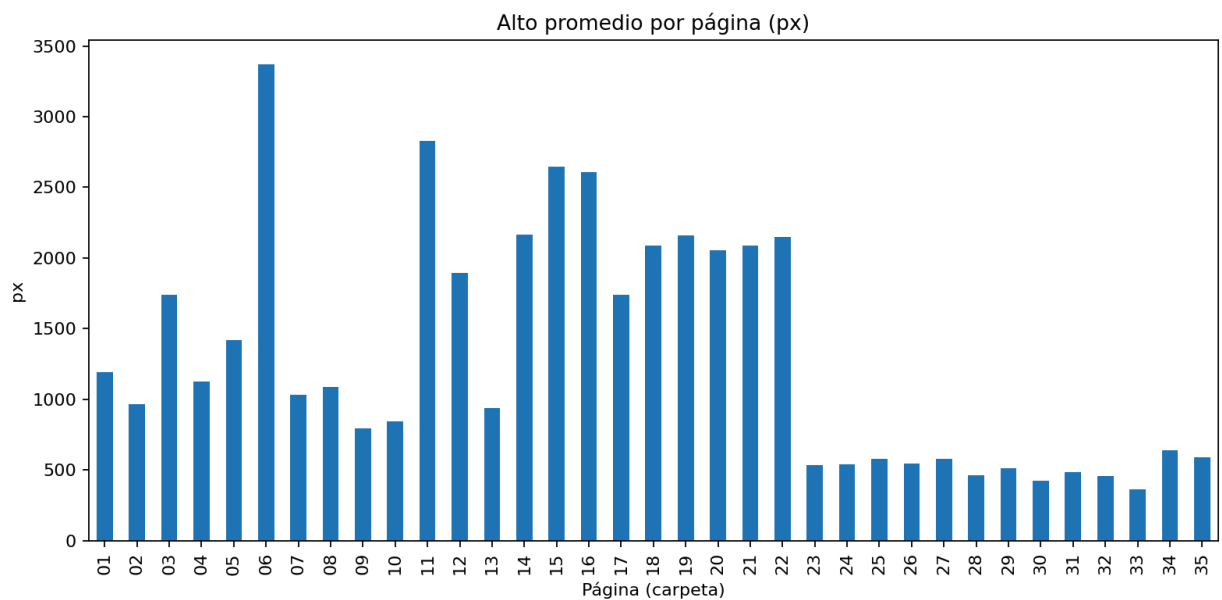
5.1 Ancho promedio

Figura 4. Ancho promedio por página (px)



5.2 Alto promedio

Figura 5. Alto promedio por página (px)



Hallazgos:

- El ancho promedio de las imágenes se mantiene entre **1,500 y 2,500 px**, con picos en páginas como la 30 (~2,800 px).
- El alto promedio presenta mayor variabilidad, con picos de más de **3,000 px en páginas 6 y 11**.
- Esto indica que algunas páginas fueron capturadas como imágenes más “verticales”, posiblemente planos de mayor extensión.

6. Distribución global de dimensiones

Se evaluó la distribución de anchos y altos de todas las imágenes del proyecto.

Figura 6. Distribución de anchos de imagen (px)

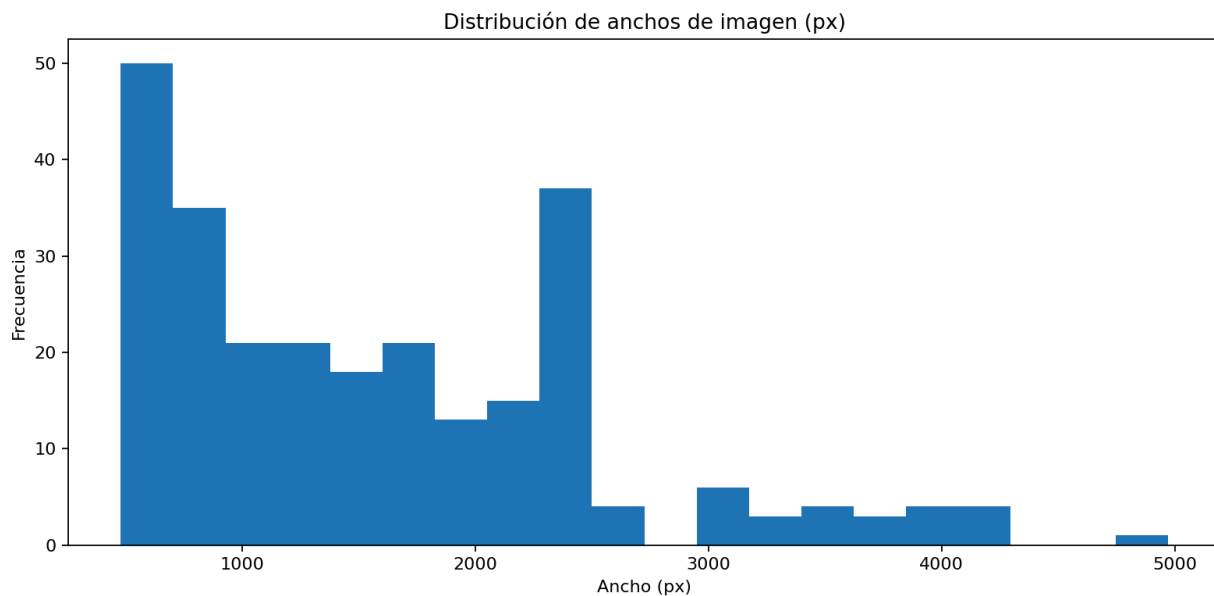
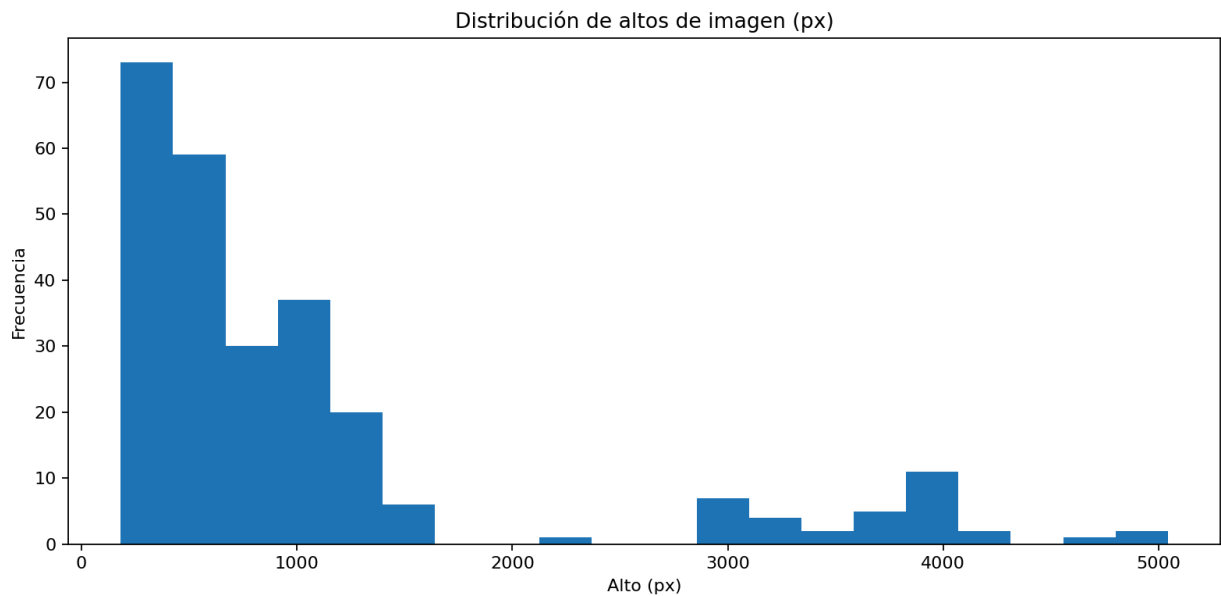


Figura 7. Distribución de altos de imagen (px)

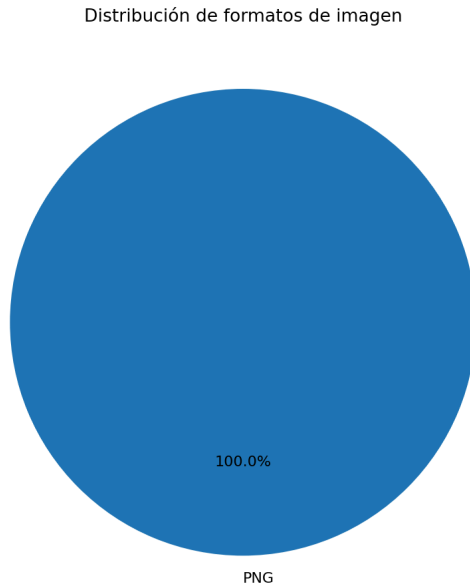


Hallazgos:

- Los anchos más frecuentes se concentran entre **500 y 2,500 px**.
- Los altos se distribuyen de manera bimodal: un grupo en **500–1,000 px** y otro entre **2,500–4,000 px**.
- Esto refleja que existen dos tipos de capturas: imágenes pequeñas (subdetalles) y otras de página completa.

7. Formatos de imagen

Figura 8. Distribución de formatos de imagen



Hallazgos:

- El 100% de las imágenes están en formato **PNG**, lo que asegura calidad sin pérdida pero con archivos más pesados que JPG.
- Esto explica el tamaño elevado en algunas páginas.

8. Conclusiones

- Todas las páginas del PDF tienen screenshots asociados → **cobertura total del proyecto**.
- Existe variabilidad significativa en número de imágenes y tamaños, lo que refleja diferentes niveles de detalle en las páginas.
- La resolución promedio es alta, lo cual favorece la legibilidad pero impacta en el peso total del repositorio.
- El uso exclusivo de PNG asegura fidelidad gráfica, aunque incrementa los KB almacenados.