Motor de Inteligencia de Mercado para el Sector de Techos en Miami-Dade

Transformando datos públicos dispersos en estrategias de captación accionables

Luis Aurelio Mendoza Michel

El Desafío de la Vulnerabilidad Estructural.

- **Crisis de Seguros:** Aumento de primas y fraude en Florida .
- **Parque Envejecido:** Miles de casas construidas antes del código post-Andrew (1992).
- **Ceguera de Datos:** No existe una base de datos centralizada que diga "cuántos años tiene este techo" .
- **Consecuencia:** El marketing es ineficiente (se vende a ciegas).

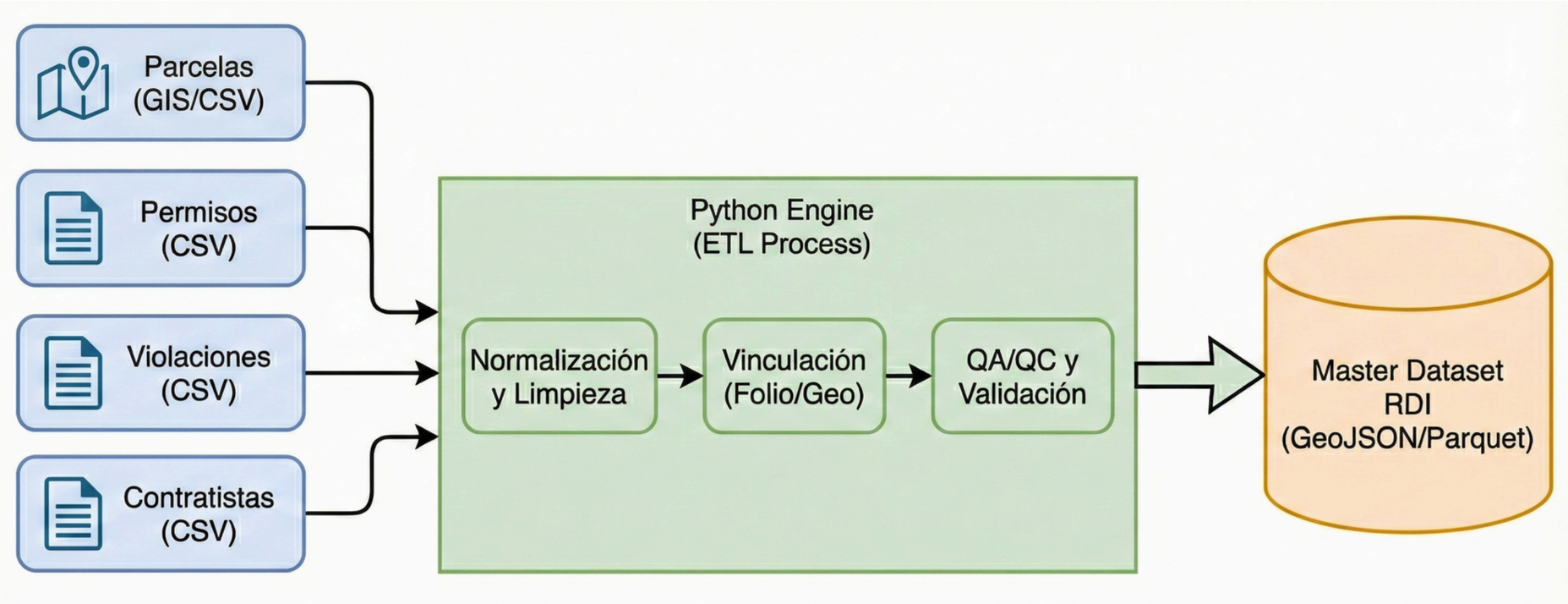
Roofing Demand Index (RDI)

Unificación de 4 fuentes de datos dispersas en un solo indicador de riesgo.

- Catastro (Parcelas).
- Permisos de Construcción.
- Violaciones de Código.
- Licencias de Contratistas.

Objetivo: Priorizar parcelas por urgencia técnica y oportunidad comercial.

Ingeniería del Datos: Del Caos al Master Dataset



Normalización y Limpieza

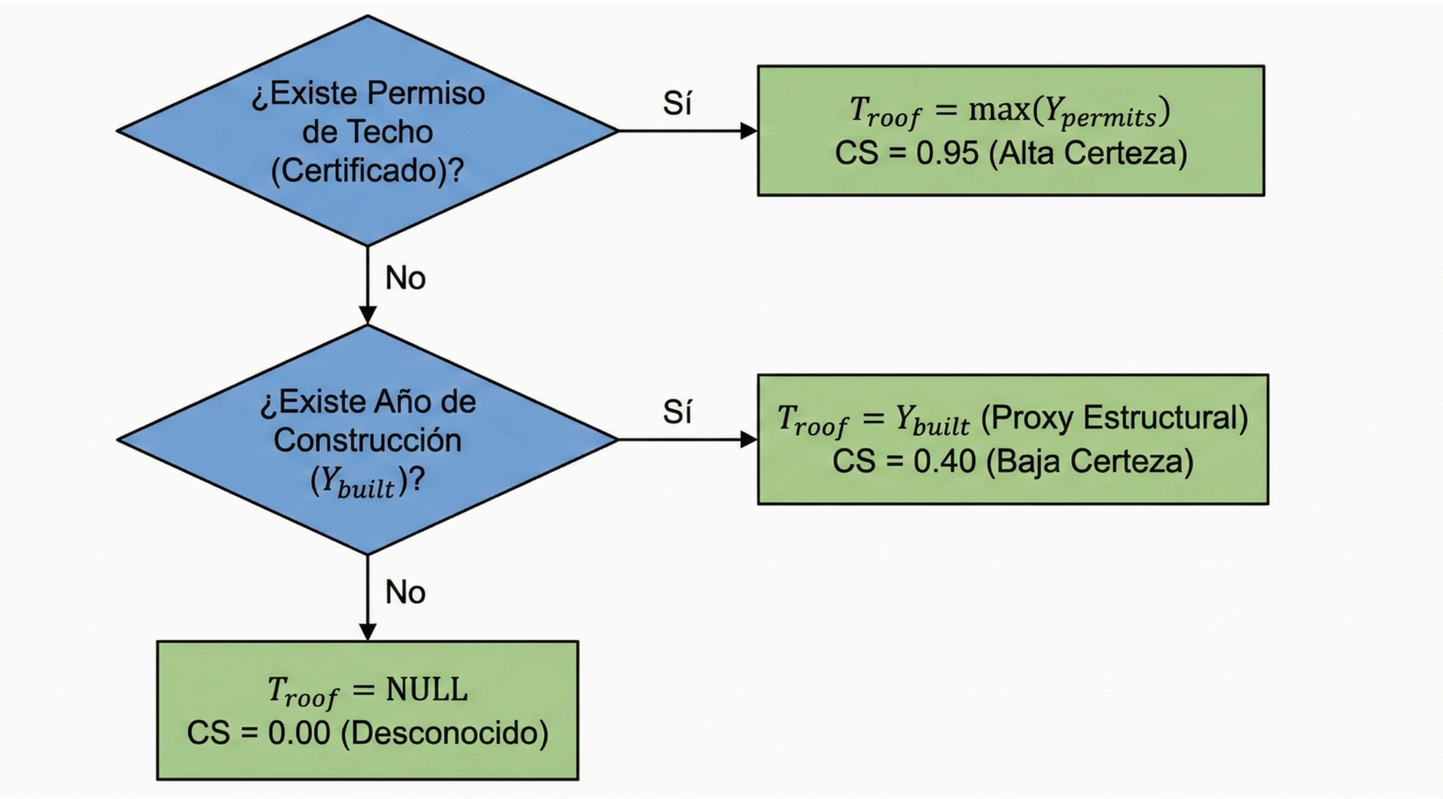
Estandarización de 4 fuentes heterogéneas. Se limpian inconsistencias sintácticas y se homologan las llaves primarias (FOLIO/ZIP) para eliminar la fricción entre formatos.

Vinculación

Cruce relacional y espacial. Se utiliza la parcela como 'Entidad Ancla' para integrar atributos físicos (Permisos/Violaciones) y métricas de mercado (Contratistas por ZIP).

QA/QC y Validación

Control de calidad automatizado. Se valida la integridad referencial y la cobertura geométrica (75%) antes de exportar el Master Dataset final en formato GeoJSON/Parquet.



¿Cómo sabemos la edad del techo?

Problema: Falta de permisos explícitos.
Solución: Algoritmo de decisión en cascada.
Si hay permiso = Alta Certeza (0.95).
Si no, usamos Año de Construcción = Proxy Estructural (0.40).

Resultado: Variable roof_age auditable para todo el dataset.

Formulación del Índice de Riesgo.

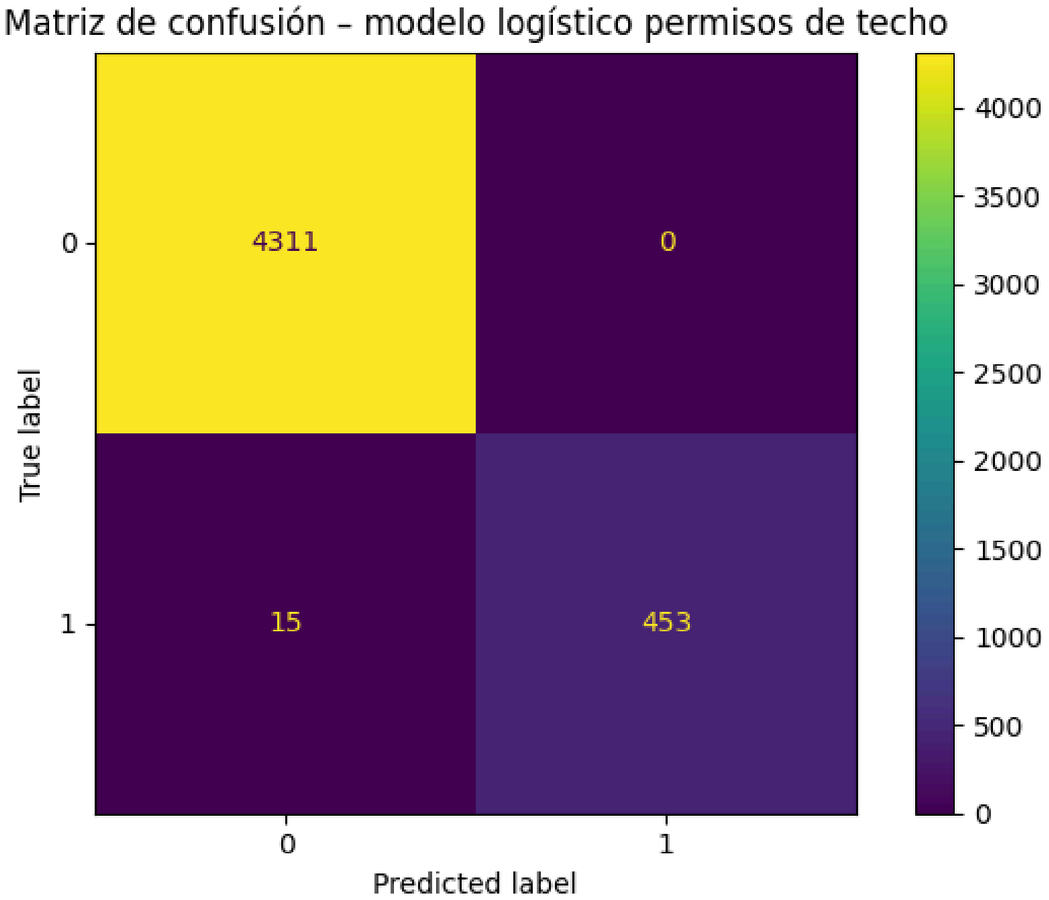
RDI = Edad + Violaciones + Seguro + Oferta

40% Desgaste: Edad del techo (Driver principal).

25% Violaciones: Historial de daños visibles.

20% Normativa: Si supera los 20 años (Vida útil seguro).

15% Mercado: ¿Cuántos competidores hay en la zona?.



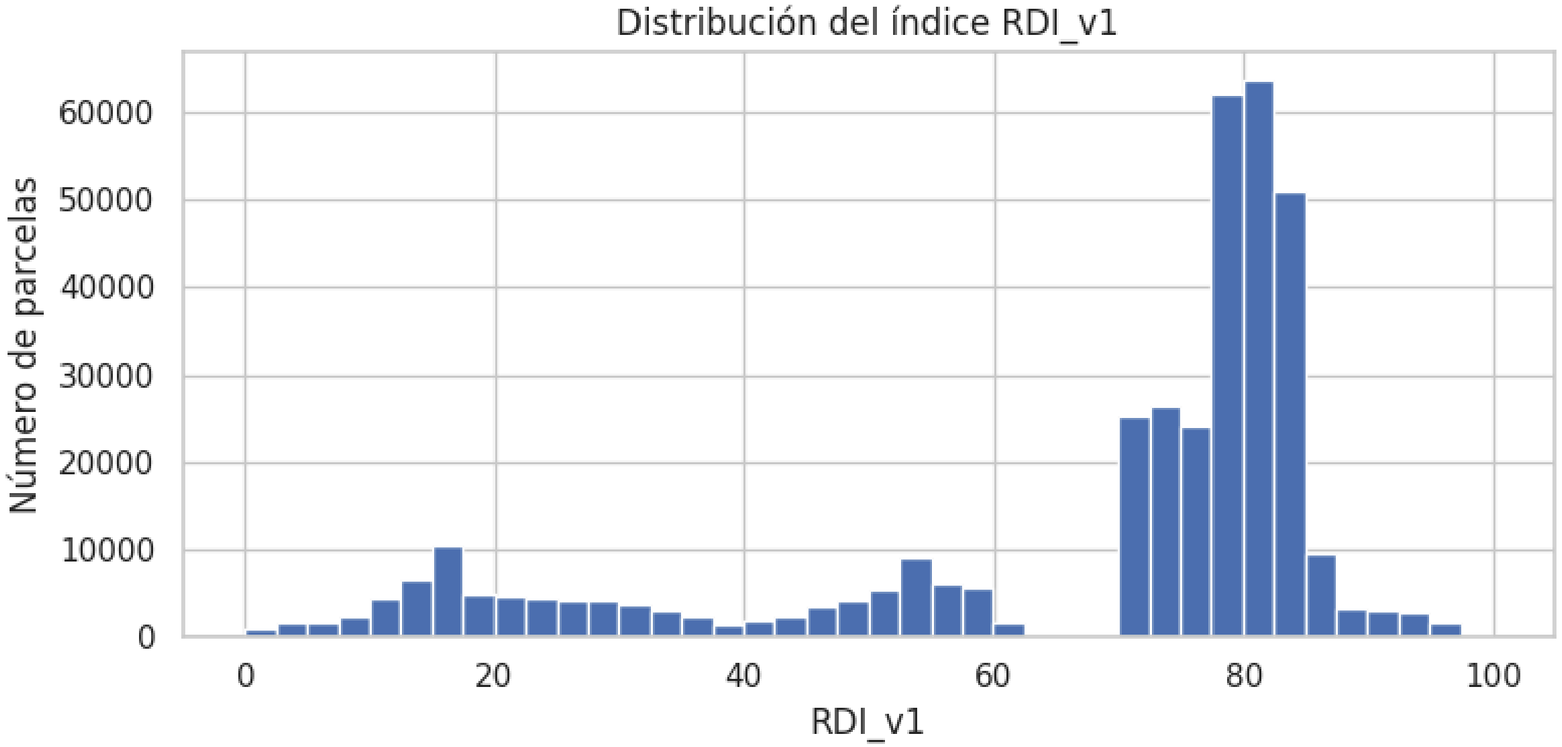
Prueba de Consistencia (Sanity Check)

Validación mediante Regresión Logística.

Exactitud: 99.7% en la detección de patrones de renovación.

Conclusión: Las parcelas con RDI alto son matemáticamente idénticas a las que se están renovando hoy

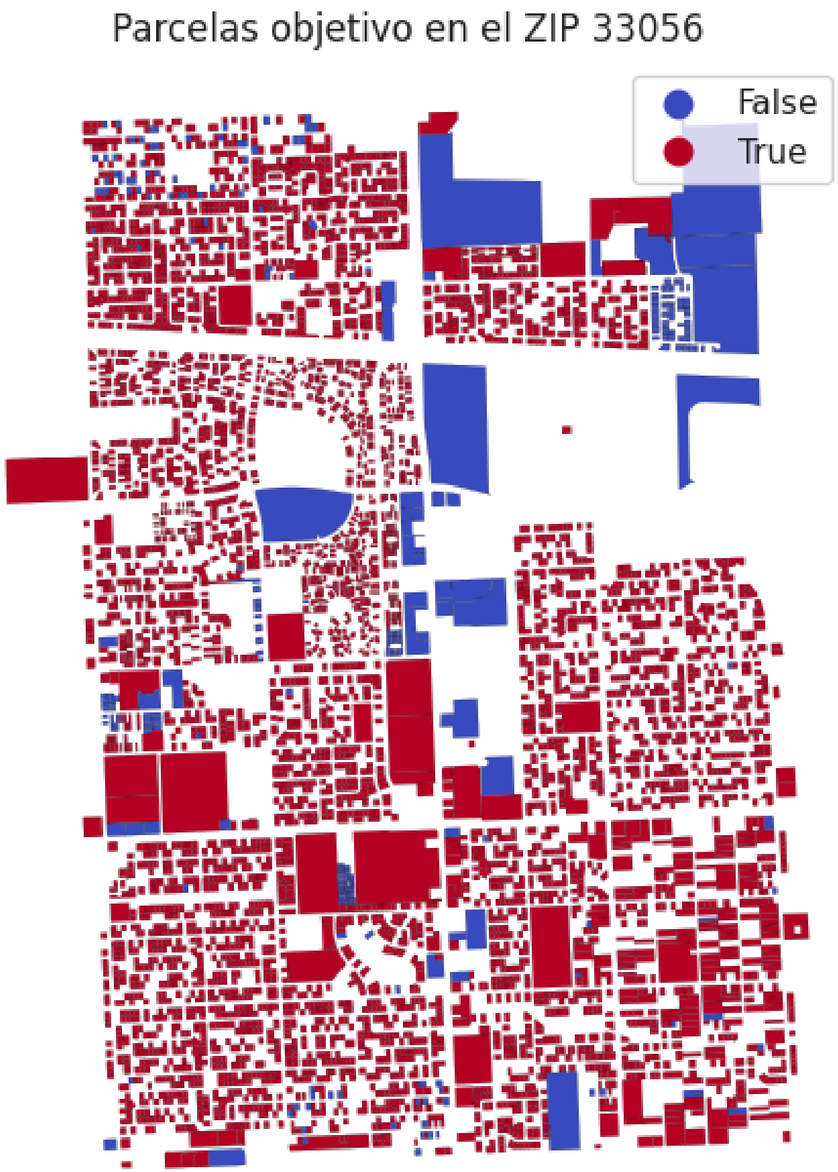
Diagnóstico del Condado Miami-Dade



Distribución Bimodal

Hallazgo: Acumulación masiva en la cola derecha ($RDI > 80$).

Oportunidad: ~60,000 propiedades en estado crítico inmediato



De la Estadística a la Acción

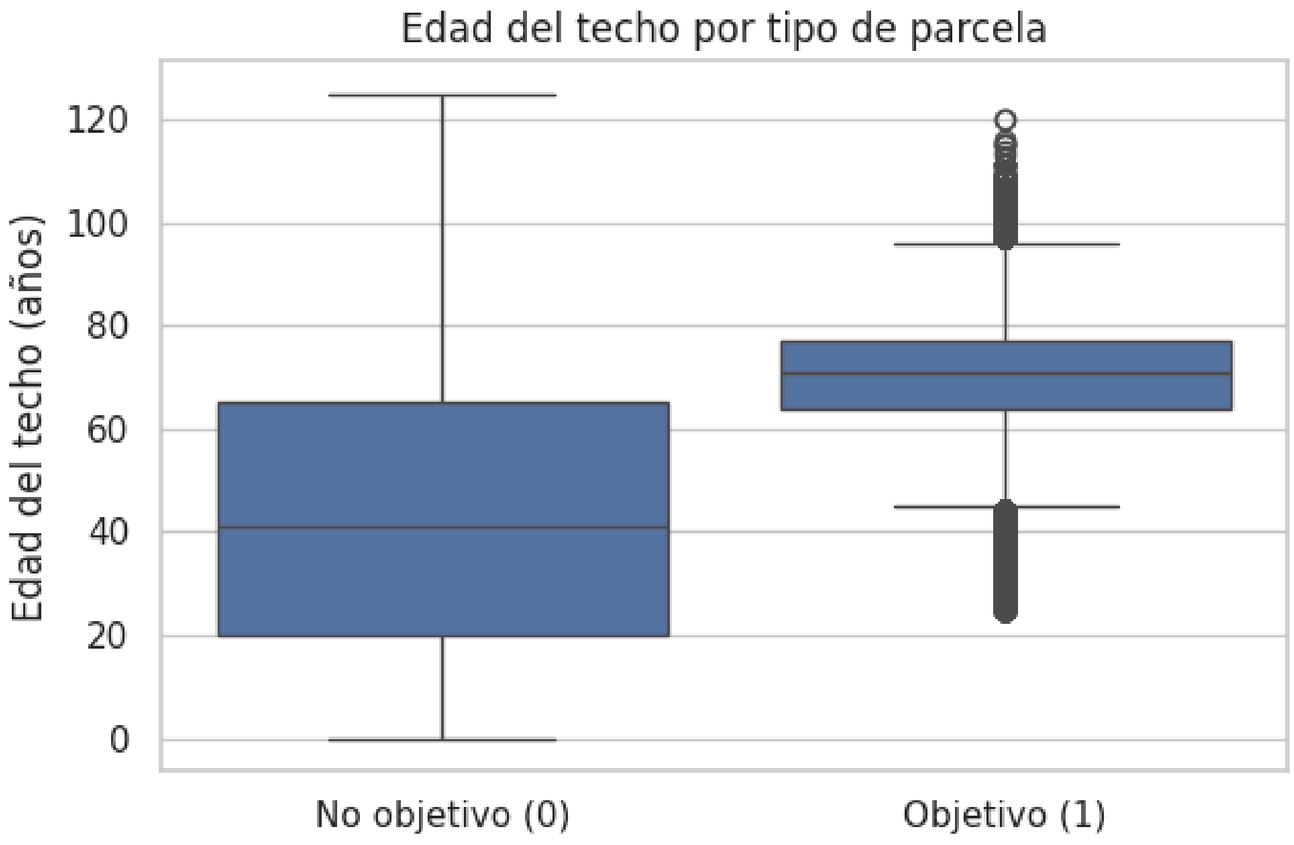
Focalización en parcelas con ≥ 81 (Rojo).

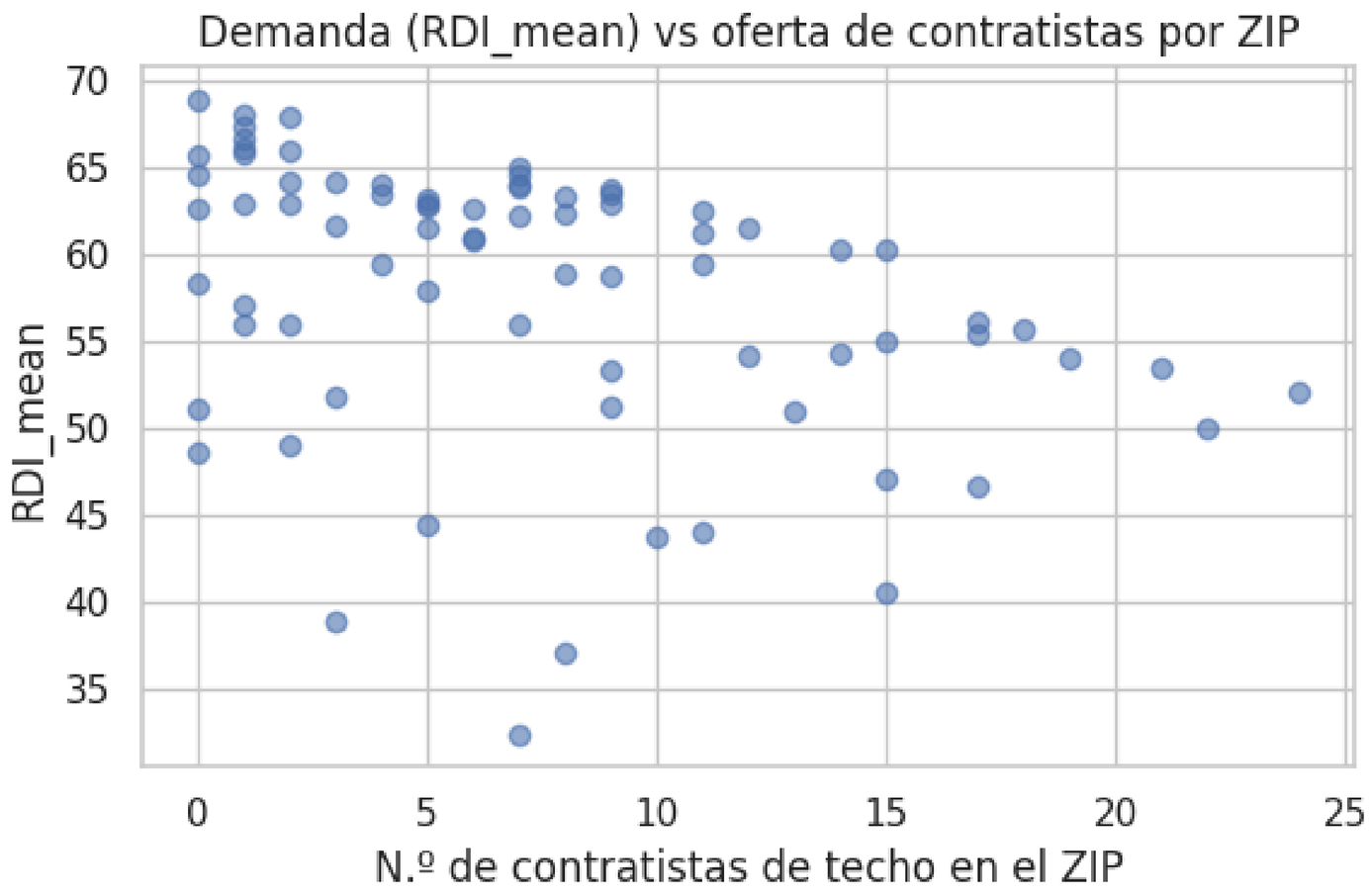
Validación de Clústeres: El riesgo no es aleatorio, afecta manzanas enteras.

Acción: Rutas de venta optimizadas (Barrido sistemático) en lugar de visitas aleatorias.

¿Estamos apuntando a las casas correctas?

- **Grupo Objetivo (1):** Mediana de edad de techo 70 años.
- **Resto (0):** Mediana de edad approx 40 años.
- **El modelo discrimina exitosamente la obsolescencia física.**



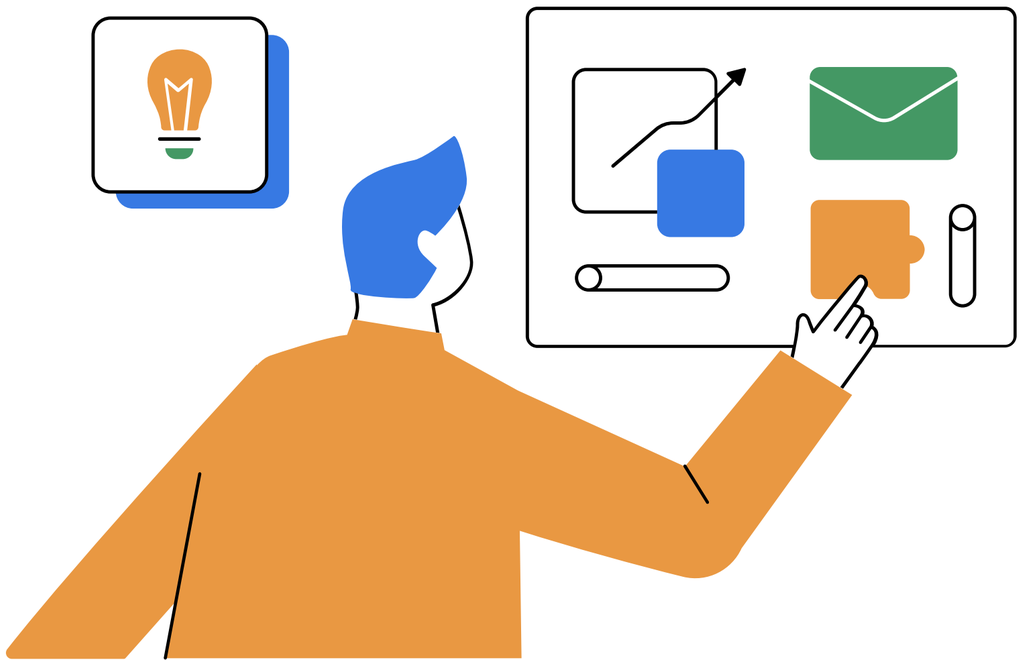


Correlación negativa. Donde más se necesita techo, menos contratistas hay.

Estrategia: Atacar el cuadrante superior izquierdo (Alta Necesidad / Baja Competencia).

Encontrando Océanos Azules

Conclusiones y Valor Entregado



Activo de Datos Unificado

Se consolidó una "fuente de verdad" auditada con 642,000 parcelas, eliminando la dispersión de datos públicos.

Estrategia de "Océanos Azules"

Se entregaron mapas operativos que identifican zonas de alta urgencia técnica y nula competencia comercial.

Tecnología Reproducible

Se implementó un Pipeline ETL automatizado en Python, dejando una infraestructura lista para procesar nuevos datos

Gracias!

Referencias

Florida Building Commission. (2023). Florida building code (8th ed.). International Code Council.

Florida Department of Business & Professional Regulation. (2024). Construction industry licensing board public records [Data set]. State of Florida Open Government. http://www.myfloridalicense.com/dbpr/sto/file_download/

Insurance Information Institute. (2024). Trends and insights: Florida's property insurance crisis.

Kim, W. C., & Mauborgne, R. (2005). Blue ocean strategy: How to create uncontested market space. Harvard Business Review Press.

Miami-Dade County. (2024). Building permits: Historical & current [Data set]. Miami-Dade Open Data Hub. https://opendata.miamidade.gov/

Miami-Dade County. (2024). Code violations: All departments [Data set]. Miami-Dade Open Data Hub. https://opendata.miamidade.gov/

Miami-Dade County GIS. (2024). Parcel boundary & property appraiser tax roll [Data set]. MDC Open Data. https://gis-mdc.opendata.arcgis.com/

Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830.