

Benemérita Universidad Autónoma de Puebla

Profesor: Jaime Romero Sierra

Trabajo: Reporte

Alumno: Carmona Hernández Oscar

Matricula: 202511455

Fecha: 05 de octubre de 2025

Semestre: 1° Grupo: 101

Análisis de Precios y Dinámica de Mercado en Airbnb NYC basado en puntos turísticos del barrio

Objetivo de Negocio

Proveer recomendaciones accionables de fijación de precios para anfitriones de Airbnb en la ciudad de Nueva York (NYC), optimizando la rentabilidad sin sacrificar competitividad. Se busca identificar los factores que más influyen en el precio por noche, principalmente ubicación (borough y barrio), lugares turísticos, época del año, tipo de propiedad, disponibilidad, reputación (reseñas), así como moderadores externos como la percepción de inseguridad en el vecindario y flujos de trabajos temporales estacionales, traduciéndolos en lineamientos claros de precios y segmentación. En un mercado saturado como el de NYC, donde los puntos turísticos como Times Square en Manhattan o el Brooklyn Bridge en Brooklyn generan picos de demanda estacional, estos lineamientos permiten a los anfitriones ajustar dinámicamente sus tarifas para capturar valor agregado sin alejar a los viajeros sensibles al precio, considerando también cómo la inseguridad percibida puede reducir reservas en un 15-25% en áreas vulnerables, mientras que trabajos temporales en eventos como la NY Fashion Week fomentan ocupación en temporadas bajas. Por ejemplo, en barrios cercanos a Central Park, los precios podrían incrementarse un 20-30% durante la primavera por la afluencia de turistas, equilibrando así la ocupación anual, y en contextos de baja demanda invernal, paquetes para trabajadores temporales podrían mitigar descuentos forzados por alta disponibilidad.

Objetivo Analítico

Construir un marco de análisis exploratorio y modelos explicativos (por ejemplo, regresión) que permitan cuantificar el impacto de cada variable en el precio, comparar escenarios, y detectar patrones espaciales que apoyen la toma de decisiones, incorporando interacciones entre factores estructurales (ubicación, tipo de habitación), operativos (disponibilidad, minimum_nights) y externos (inseguridad y trabajos temporales). Este marco no solo medirá el efecto directo de la proximidad a puntos turísticos —como el Empire State Building o el High Line—, sino que también integrará análisis geoespacial para mapear "zonas calientes" de demanda, facilitando simulaciones de precios bajo diferentes escenarios estacionales, incluyendo moderadores como tasas de crimen del NYPD que atenúan premiums en barrios como el Bronx o flujos laborales que elevan ocupación en Queens durante conferencias. De esta manera, se generarán insights accionables, como umbrales de precios óptimos por barrio, que ayuden a predecir fluctuaciones basadas en eventos turísticos anuales, como el maratón de NYC o las festividades navideñas, ajustados por estrategias para contrarrestar inseguridad (ej. badges de seguridad) y capitalizar trabajos temporales (ej. paquetes "Work & Stay").

Introducción al Mercado

El mercado de estancias de corta duración en New York City es uno de los más grandes y competitivos del mundo, con más de 40,000 listados activos en plataformas como Airbnb, impulsados por la diversidad de atractivos turísticos que convierten a la ciudad en un imán global, aunque moderados por percepciones de inseguridad y oportunidades de trabajos temporales que influyen en la demanda. Para los anfitriones, fijar el precio óptimo implica balancear múltiples factores: ubicación, tipo de espacio, reputación (reseñas), época del año, estacionalidad (disponibilidad), así como el impacto de la inseguridad que puede presionar precios a la baja en un 20% en áreas con incidentes reportados, y el rol de trabajos temporales en industrias como hospitalidad o eventos que fomentan reservas estables en temporadas bajas. Sin una guía basada en datos, es común subestimar o sobreestimar tarifas, lo que deriva en menor ocupación o en ingresos sacrificados. Asimismo, las diferencias sustanciales entre Manhattan, Brooklyn, Queens, Bronx y Staten Island complican extrapolar aprendizajes entre zonas, donde Manhattan soporta premiums por landmarks icónicos como la Estatua de la Libertad (accesible desde Battery Park), pero enfrenta desafíos de percepción de crimen en el subway, mientras Queens ofrece alternativas asequibles cerca de atracciones culturales emergentes impulsadas por empleos estacionales en retail o producción. Este proyecto busca descomponer la formación de precios utilizando el dataset “New York City Airbnb Open Data (2019)”, para entender los determinantes clave y generar recomendaciones aplicables para distintas combinaciones de barrio, época del año, tipo de habitación y demanda, integrando moderadores como inseguridad y trabajos temporales.

Además de cuantificar el efecto de ubicación y tipo de propiedad, exploraremos cómo variables operativas se relacionan con señales de demanda y precio, enfatizando el rol integrador de estos factores junto con la inseguridad que disuade turistas en un 10-15% y trabajos temporales que elevan ocupación en un 20-25% durante lows. En particular, se enfatizará el rol de los puntos turísticos del barrio, como la proximidad a museos en el Upper East Side o parques en Williamsburg, que actúan como multiplicadores de precio durante temporadas altas (verano y otoño), pero se atenúan en zonas con mayor crimen percibido, mientras que flujos laborales como crew de cine en Harlem sostienen demanda invernal. El entregable de esta fase son mapas de calor, mostrando el problema, distintas hipótesis medibles y un plan de análisis que facilite la ejecución técnica en fases posteriores. Estos mapas no solo visualizarán densidades de listados, sino que superpondrán capas de concurrencia turística —basadas en datos proxy como reseñas estacionales—, inseguridad (tasas NYPD) y hotspots laborales (ej. venues de eventos), para resaltar cómo la dinámica de mercado se ve influida por flujos de visitantes y trabajadores, permitiendo a los anfitriones anticipar "ventanas de oportunidad" para ajustes de precios en contextos de riesgo o demanda alternativa.

Tecnología y Herramientas

Python (Pandas, NumPy), visualización (Matplotlib, Seaborn), geopandas (cartografía básica), Jupyter Notebook para el flujo exploratorio, y un video del proyecto realizado para la comunicación de resultados. Estas herramientas se seleccionaron por su capacidad para manejar datos geoespaciales, esenciales para mapear puntos turísticos, barrios, zonas de inseguridad (integrando datos NYPD) y áreas de trabajos temporales (ej. cerca de Madison Square Garden); por ejemplo, Geopandas permitirá crear choropleths que ilustren variaciones de precios por proximidad a sitios como el MoMA o el Yankee Stadium, superponiendo capas de crimen y eventos laborales para simular impactos moderadores. Jupyter Notebook facilitará un workflow iterativo, donde se puedan probar visualizaciones interactivas de estacionalidad, como gráficos de series temporales proxy basados en availability_365, para simular impactos de épocas del año en la rentabilidad, incluyendo escenarios de alta inseguridad que reducen ocupación o flujos temporales que la estabilizan.

Datos Disponibles

Los datos disponibles (Kaggle – New York City Airbnb Open Data, archivo AB_NYC_2019.csv). A continuación se incluye un diccionario de variables y su tipo, con la interpretación usada en este proyecto. Aunque el dataset es de 2019, servirá como base proxy para patrones estructurales, complementado con insights cualitativos sobre evolución turística post-pandemia, como el aumento de demanda en barrios periféricos debido a la gentrificación y nuevos hotspots como el Edge en Hudson Yards, así como datos externos sobre inseguridad (NYPD reports) y trabajos temporales (Departamento de Trabajo NYC) para enriquecer análisis de moderadores.

Diccionario de Datos

Columna	Tipo	Descripción (interpretación)
id	Numérico	Identificador único del arrendador
name	Texto	Nombre
host_id	Numérico	Identificador del anfitrión
host_name	Texto	Nombre del anfitrión
neighbourhood_group	Categórico	Borough: Manhattan, Brooklyn, Queens, Bronx,

		Staten Island
neighbourhood	Categórico	Barrio dentro del borough
latitude	Numérico	Coordenadas
longitude	Numérico	Coordenadas
room_type	Categórico	Tipo de espacio: home/apt, private room, shared room, etc
price	Numérico	Precio por noche (USD)
minimum_nights	Numérico	Mínimo de noches requeridas
number_of_reviews	Numérico	Número total de reseñas
last_review	Numérico	Fecha de la última reseña
reviews_per_month	Numérico	Promedio de reseñas por mes
calculated_host_listings_count	Numérico	Cantidad de listados del anfitrión
availability_365	Numérico	Días disponibles al año

Notas de Calidad de Datos Esperadas

Valores faltantes en last_review y reviews_per_month; sesgo de precios con cola larga; valores extremos en minimum_nights y price; y concentración geográfica y datos atípicos. Para mitigar estos, se aplicarán imputaciones basadas en mediana por barrio (considerando proximidad turística, inseguridad y proximidad laboral), y winsorización para outliers, asegurando que los análisis de precios reflejen dinámicas reales sin distorsiones, incluyendo ajustes por moderadores como crimen que sesgan percepciones en reseñas. La concentración geográfica, evidente en Manhattan (donde el 40% de listados se ubican cerca de puntos turísticos centrales pero expuestos a subway incidents), requerirá normalizaciones espaciales para evitar sesgos en mapas de calor, incorporando proxies para trabajos temporales en áreas periféricas.

Hipótesis Planteadas

Hipótesis 1°. Efecto Integrado de Factores Estructurales y Operativos en el Precio: Los precios por noche en listados de Airbnb en NYC están positivamente influenciados por una combinación

interactiva de ubicación (borough y barrio, con énfasis en proximidad a zonas turísticas y niveles de seguridad percibida), tipo de habitación, reputación (número de reseñas y reviews_per_month), disponibilidad anual (availability_365) y restricciones operativas (minimum_nights), ajustados por la época del año. Específicamente, listados en Manhattan cerca de atractivos como Broadway o el Rockefeller Center exhiben premiums de hasta \$100/noche (o 50-60% superiores a equivalentes en el Bronx o Staten Island) durante picos turísticos estacionales (verano/otoño), donde "entire home/apt" comandan \$250/noche versus \$150 para "private room" o "shared room", impulsados por preferencias de familias y grupos en áreas de alta concurrencia. Esta superioridad se modera por señales de confianza: un mayor número de reseñas (>50) o reviews_per_month (>2) otorga un " premio por reputación" de 10-15% en precios o mayor ocupación, actuando como endorsement en búsquedas de Airbnb, particularmente en barrios con puntos turísticos estacionales como Central Park. Por el contrario, alta disponibilidad (availability_365 >200 días) y requisitos elevados de minimum_nights (>7) presionan precios a la baja (descuentos del 15-20%) para estimular demanda en épocas bajas (invierno), como en Staten Island lejos de hotspots, mientras que baja disponibilidad en verano cerca de Coney Island soporta precios inflados. En barrios con menor seguridad percibida (ej. partes del Bronx), estos efectos se atenúan en un 10-20%, requiriendo ajustes dinámicos para mantener competitividad sin sacrificar rentabilidad.

Hipótesis 2°. Efecto Moderador de la Inseguridad en la Demanda y Precios de Airbnb: La percepción de inseguridad en barrios de NYC (medida por tasas de crimen reportadas del NYPD y reseñas negativas sobre seguridad en listados) actúa como un moderador negativo en la relación entre proximidad a puntos turísticos y precios/ocupación de Airbnb, reduciendo reservas en un 15-25% en áreas con incidentes elevados (ej. subway en Midtown o calles en Brooklyn), incluso durante temporadas altas, y exacerbando caídas en épocas bajas. Esta hipótesis se basa en la brecha entre realidad (crimen violento bajo un 1.8% en 2023) y percepción mediática (amplificada por CNN o redes sociales), que disuade a turistas internacionales y familias, presionando precios promedio en \$20-50/noche para competir. Por ejemplo, en barrios como el Bronx o partes de Queens con mayor incidencia de delitos menores (robos/vandalismo), listados cerca de atracciones emergentes como el Yankee Stadium ven ocupación 20% inferior comparado con zonas "seguras" como Upper East Side, donde reseñas positivas sobre "entorno tranquilo" mitigan el impacto. Anfitriones pueden contrarrestar esto destacando medidas de seguridad (cerraduras inteligentes, proximidad a patrullas), lo que elevaría la visibilidad en búsquedas y recuperaría hasta un 30% de demanda perdida, alineándose con iniciativas como "Safe Streets" del alcalde Adams para mejorar la percepción turística general.

Hipótesis 3°. Efecto Promotor de Trabajos Temporales en la Demanda de Airbnb durante Temporadas Bajas: Los flujos de trabajadores temporales en industrias estacionales de NYC (turismo, eventos, hospitalidad y producción, con ~500,000 roles anuales según el Departamento

de Trabajo) fomentan un aumento del 20-25% en la ocupación y precios de Airbnb en temporadas bajas (invierno/otoño temprano), particularmente en boroughs periféricos como Queens y Brooklyn, al preferir estancias flexibles y asequibles (20-30% más baratas que hoteles) cerca de venues laborales. Esta hipótesis resalta cómo roles como guías turísticos para el Winter Jazzfest (enero), crew de cine/TV en estudios de Harlem o personal de retail para holidays en Macy's (diciembre) generan "micro-estacionalidad", elevando reservas para "private rooms" o "shared rooms" con minimum_nights bajos (1-7 noches) en áreas subutilizadas, como Astoria cerca del Citi Field o Williamsburg para eventos como NY Fashion Week. Por instancia, durante la UN General Assembly (septiembre, temporada de transición baja), ~50,000 asistentes temporales impulsan demanda en Upper East Side, permitiendo premiums del 10-15% en precios ajustados por disponibilidad alta (availability_365 >250), y reduciendo tiempos muertos en un 25%. Publicidad targeted (ej. en LinkedIn o Indeed con paquetes "Work & Stay") puede maximizar este efecto, segmentando por barrio y tipo de habitación para capturar ingresos estables, contrarrestando la menor afluencia turística general y promoviendo una distribución equitativa de demanda en barrios diversos.

Stakeholders

Anfitriones de Airbnb: Requieren reglas de precio por zona y tipo de habitación, y señales de demanda por temporada, especialmente adaptadas a flujos turísticos locales, moderadores de inseguridad y oportunidades de trabajos temporales para maximizar ingresos en barrios como Chelsea o Harlem, donde paquetes para crew temporal pueden llenar vacíos invernales.

Plataforma (productos de pricing/revenue): Información para ajustar recomendaciones automáticas, incorporando datos de concurrencia turística, tasas de crimen y flujos laborales para algoritmos predictivos que simulen escenarios de riesgo o demanda alternativa.

Analistas e Investigadores Urbanos: Insumos para estudiar patrones de turismo y presión en barrios, evaluando cómo los precios de Airbnb impactan la vivienda local cerca de landmarks, considerando desigualdades exacerbadas por inseguridad en el Bronx o beneficios de empleos estacionales en Queens.

Viajeros: Beneficiarios indirectos a través de mayor transparencia y competencia, permitiendo elecciones informadas basadas en valor por proximidad a atracciones, con énfasis en seguridad y accesibilidad para estancias relacionadas con trabajo temporal.

Preguntas Clave

¿Cuál es el precio medio por borough y por barrio, ajustado por proximidad turística y niveles de inseguridad?

¿Cómo varía el precio por tipo de habitación dentro de cada borough, considerando interacciones con reseñas y disponibilidad?

¿Qué barrios concentran los listados con precios del cuartil superior? (Ej. Midtown en Manhattan por su acceso a Times Square, mitigado por crimen en subway).

¿Cuál es la relación entre disponibilidad_365 y precio, moderada por épocas bajas y flujos de trabajos temporales?

¿Existe relación entre número de reseñas/reviews_per_month y precio u ocupación esperada, especialmente en contextos de inseguridad percibida?

¿Cómo impacta minimum_nights en el precio y la demanda potencial, para estancias de trabajadores temporales?

¿Qué combinación (borough–room_type) maximiza el ingreso estimado por noche?
(Posiblemente Manhattan-entire home/apt en temporada alta, o Queens-private room para eventos laborales).

¿Existen patrones espaciales (clusters) de precios altos/bajos al controlar por tipo de habitación, inseguridad y hotspots laborales?

¿Qué variables explican mejor la variación en el precio (importancia de variables), incluyendo moderadores como crimen y empleos estacionales?

¿Cuáles son los outliers de precio y qué los caracteriza? (Ej. penthouses cerca de atracciones exclusivas, o descuentos en áreas de alta inseguridad).

¿Qué tan concentrada está la oferta por borough/barrio y tipo de habitación, influida por demanda temporal?

¿Qué reglas simples de pricing pueden derivarse (p. ej., recargos por zona/amenidades, descuentos por seguridad o paquetes laborales)?

Estas preguntas guían un análisis integral, priorizando la integración de factores turísticos, inseguridad y trabajos temporales para derivar insights como recargos del 25% en barrios con alta densidad de visitantes como SoHo durante ferias de arte, o descuentos del 20% en invierno para crew de eventos en Brooklyn.

Fuentes de Datos

Kaggle – New York City Airbnb Open Data (AB_NYC_2019.csv).

Documentación de tipos de espacio de Airbnb (Help Center) para definir “entire place”, “private room” y “shared room”.

Inside Airbnb / Repositorios públicos como referencia secundaria de estructura del dataset. Para enriquecer con datos turísticos, se podrían complementar con APIs abiertas de NYC (ej. puntos de interés de datos.gov), datos de inseguridad (NYPD Crime Reports) y empleos temporales (NYC Department of Labor statistics), aunque el foco principal permanece en el dataset base para patrones estructurales.

Justificación

La ciudad de Nueva York es uno de los destinos turísticos más visitados del mundo, con millones de visitantes anuales que buscan alojamiento flexible, accesible y bien ubicado. Sin embargo, las variaciones en la delincuencia urbana, los cambios en la regulación de alquileres a corto plazo (como la Ley Local 18) y las percepciones de seguridad influyen significativamente en el comportamiento de los turistas y, por tanto, en la demanda de alojamientos tipo Airbnb. Barrios con menor percepción de seguridad tienden a experimentar menor ocupación o precios más bajos, mientras que zonas turísticas consolidadas mantienen tarifas elevadas. Analizar estas diferencias permite identificar oportunidades para ajustar estrategias de precios y promoción.

Por otro lado, la economía local de NYC también se ve impulsada por la presencia de trabajadores temporales, profesionales en proyectos de corto plazo, estudiantes internacionales y personal de eventos o producción audiovisual, quienes utilizan Airbnb fuera de las temporadas turísticas altas. Comprender los patrones de estancia de estos grupos y su relación con la disponibilidad y la oferta puede ayudar a los anfitriones a mantener ingresos estables durante todo el año, reduciendo el impacto de la estacionalidad.

Este proyecto, al analizar los datos públicos de Airbnb, busca ofrecer información útil tanto a anfitriones como a investigadores y responsables de políticas urbanas. Los resultados pueden emplearse para optimizar precios dinámicos, mejorar estrategias de publicidad en temporadas bajas y fomentar un turismo sostenible que beneficie tanto a los residentes como a los visitantes. Además, el análisis servirá como caso práctico de cómo la analítica de datos puede aplicarse para resolver problemas reales de mercado, promoviendo decisiones basadas en evidencia y contribuyendo al desarrollo económico y social de la ciudad.

Plan de Trabajo:

- 1° Limpieza del DataSet: Validación de estructura, auditoría de faltantes, duplicados y outliers; tipificación de variables; diccionario validado.
- 2° Valor descriptivo: estadísticas por vecindario y tipo de cuarto; boxplots/log-precio; mapas de calor por precio mediano, fechas concurridas; cruces con availability_365; análisis de colas .
- 3° Modelado explicativo: regresión lineal con transformaciones (log-precio) y dummies por borough y room_type; experimentos con árboles de decisión para importancia de variables; validación cruzada simple.
- 4° Reglas de negocio: construcción de matriz de recargos/descargos por segmento y guía de precios objetivo; recomendaciones sobre mínimos de noches y ventanas de disponibilidad.
- 5° Comunicación: dashboard estático (tablas y gráficos clave) junto a un reporte ejecutivo con hallazgos y próximos pasos.

Tipo de Datos:

El archivo AB_NYC_2019.csv contiene 48,895 registros y 16 columnas. Los tipos abarcan: categóricas (neighbourhood_group, neighbourhood, room_type), numéricas (price, minimum_nights, number_of_reviews, reviews_per_month, calculated_host_listings_count, availability_365, latitude, longitude) y temporales (last_review).

Fuentes:

- Kaggle – New York City Airbnb Open Data (2019).
- GitHub – kaggledatasets: referencia del archivo AB_NYC_2019.csv.
- Airbnb Help Center – Tipos de lugares (definiciones de “Entire place”, “Private room”, “Shared room”).