

# Análisis del precio de las viviendas en la ciudad de Manhattan (2024-2025)

---

Proyecto IMT2118  
Entrega Final



## Contexto de la situación

- El distrito de Manhattan es conocido por ser la zona residencial más costosa de toda la ciudad de New York.
- Cuenta con un aproximado de 8,6 millones de habitantes.
- Según la página *Tinsa* en 2013, también es conocido por ser una de las zonas de tiendas más caras y exclusivas de todo el mundo.



# ¿Por qué Manhattan es tan popular?



Puente Brooklyn



Central Park



World Trade Center

## Parámetros a tener en cuenta



## Problemática a abordar

**¿En qué zonas es más probable encontrar propiedades en venta que sean favorables para un futuro comprador, considerando factores económicos, índices delictivos, cercanía a servicios de transporte, de salud y educación, entre otros?**

A raíz de ello, podemos plantearnos las siguientes preguntas:

- **¿Cómo es la urbanización y vegetación en la zona de Manhattan considerando que es el distrito con mayor población de New York?**
- **¿Afecta en algo el precio de la vivienda dado los parámetros anteriores?**
- **¿Cómo se distribuyen espacialmente las zonas que tuvieron viviendas vendidas en Manhattan? ¿Existe un lugar más caro que el resto?**

# Datos y librerías a utilizar

## Datos:

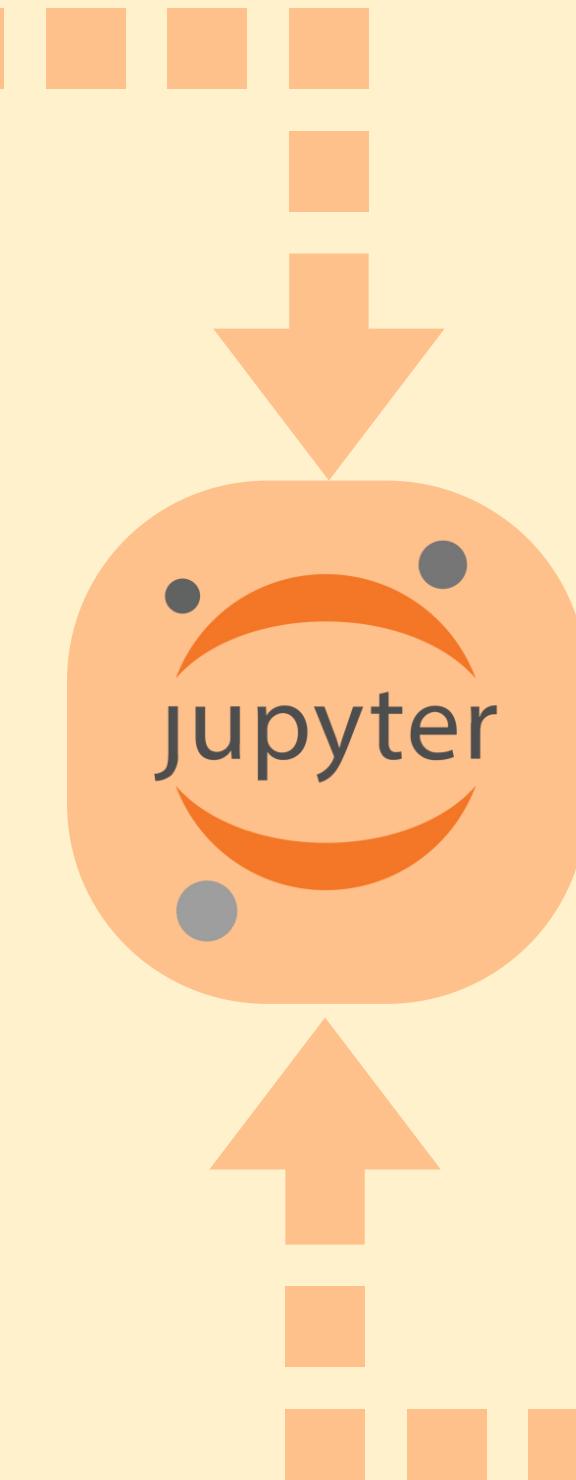
### Datos vectoriales

- property rolling sales data
- NYPD Shooting Incident Data (Historic)
- MTA Subway Entrances and Exits
- hospital
- COLLEGE\_UNIVERSITY
- Shapefile USA



### Datos raster

- Sentinels 2
- Sentinels 1



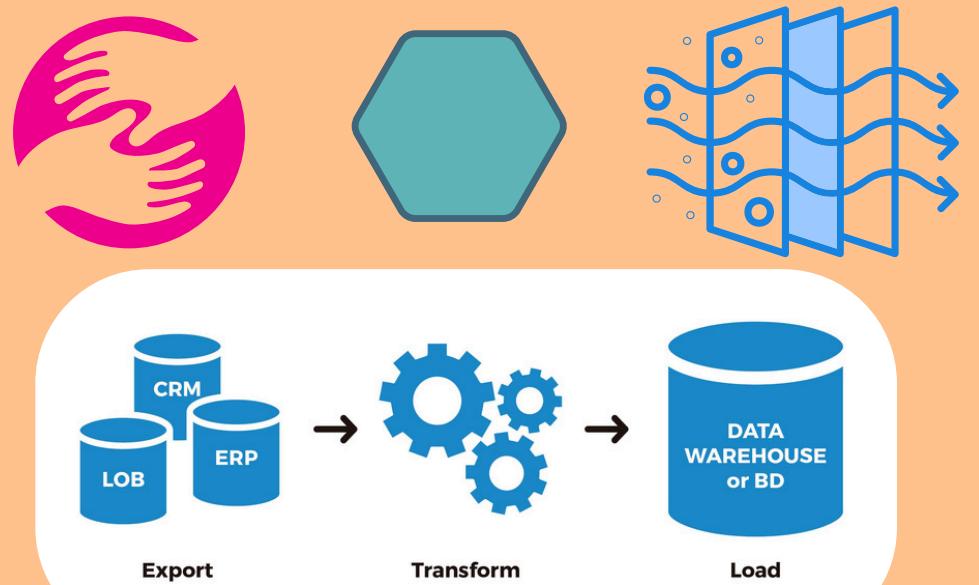
## Librerías:



# ¿Cómo se desarrolla el proyecto?

## Procesamientos

- ETL
- Filtramientos por condición
- Uniones
- Segmentación hexagonal



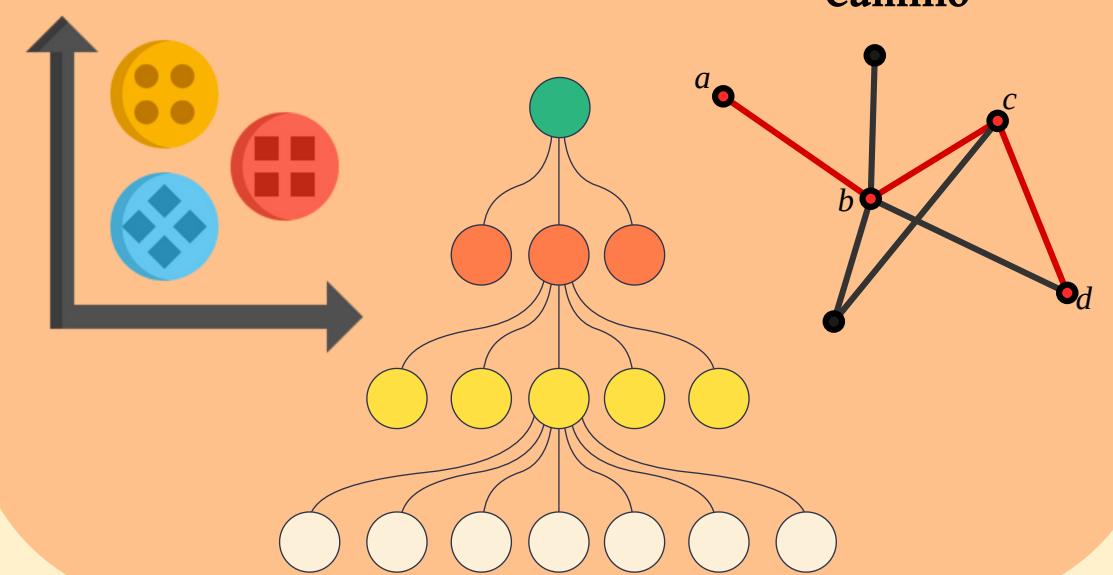
## Análisis

- Espacial
- Temporal
- Estadístico
- Predictivo

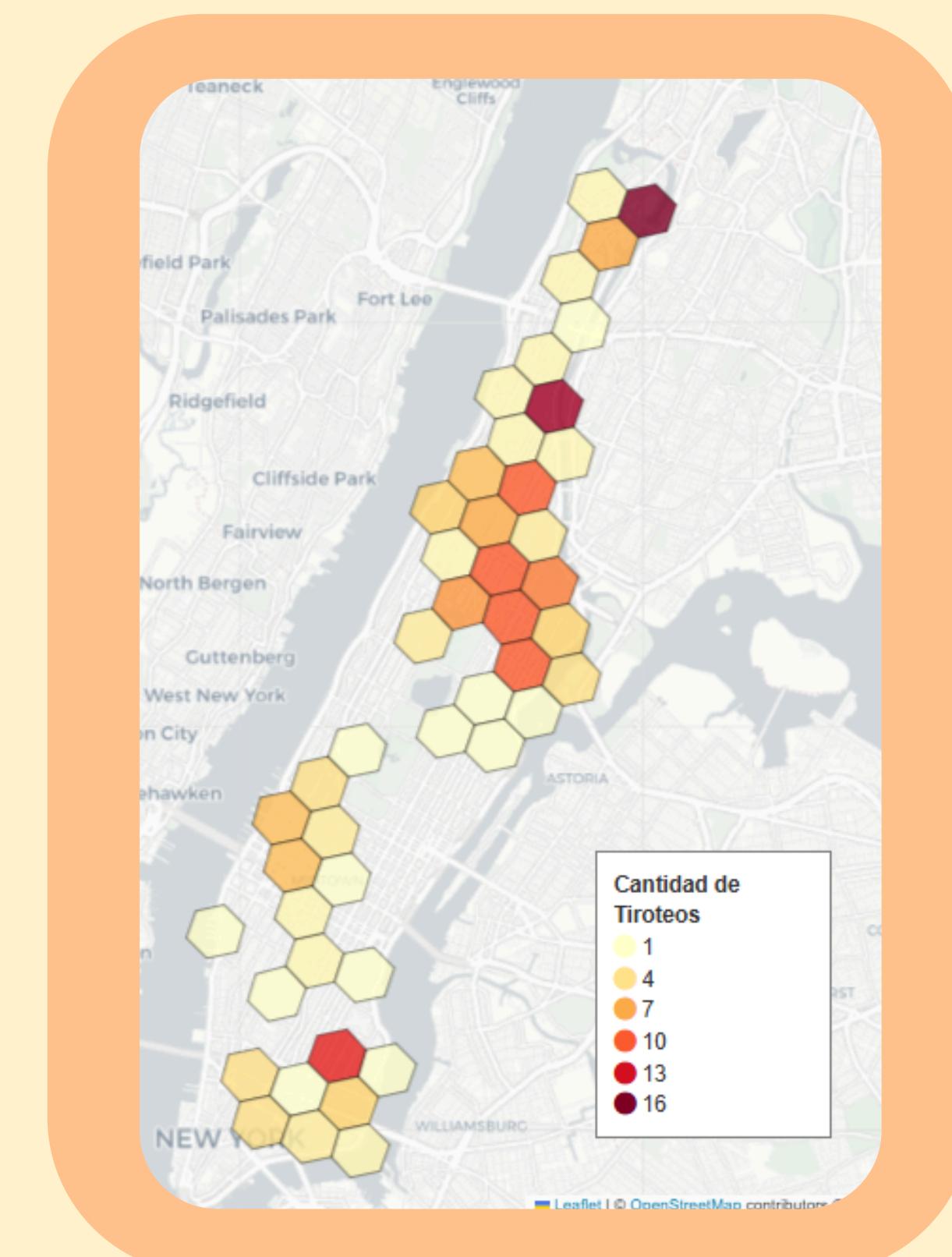
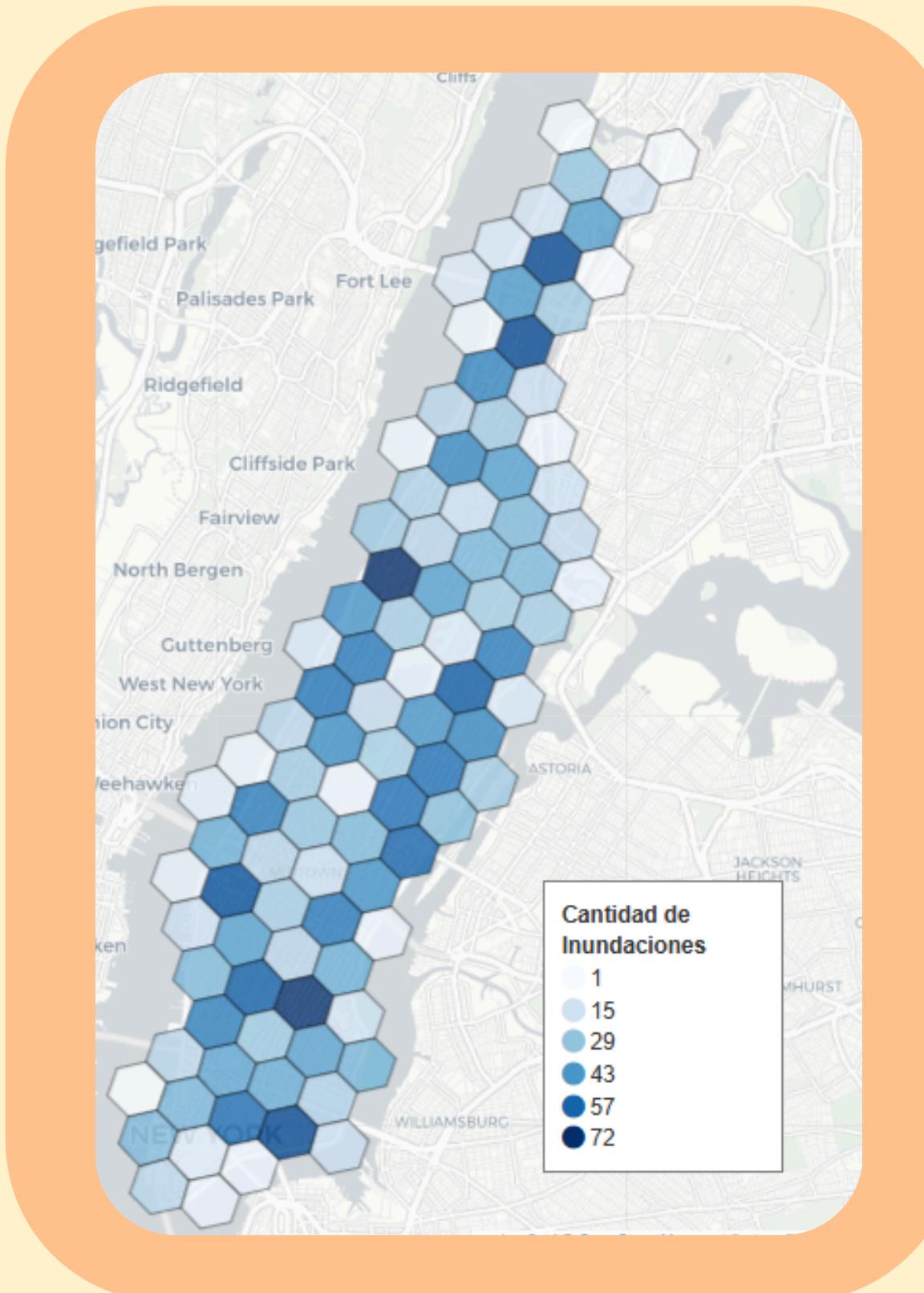
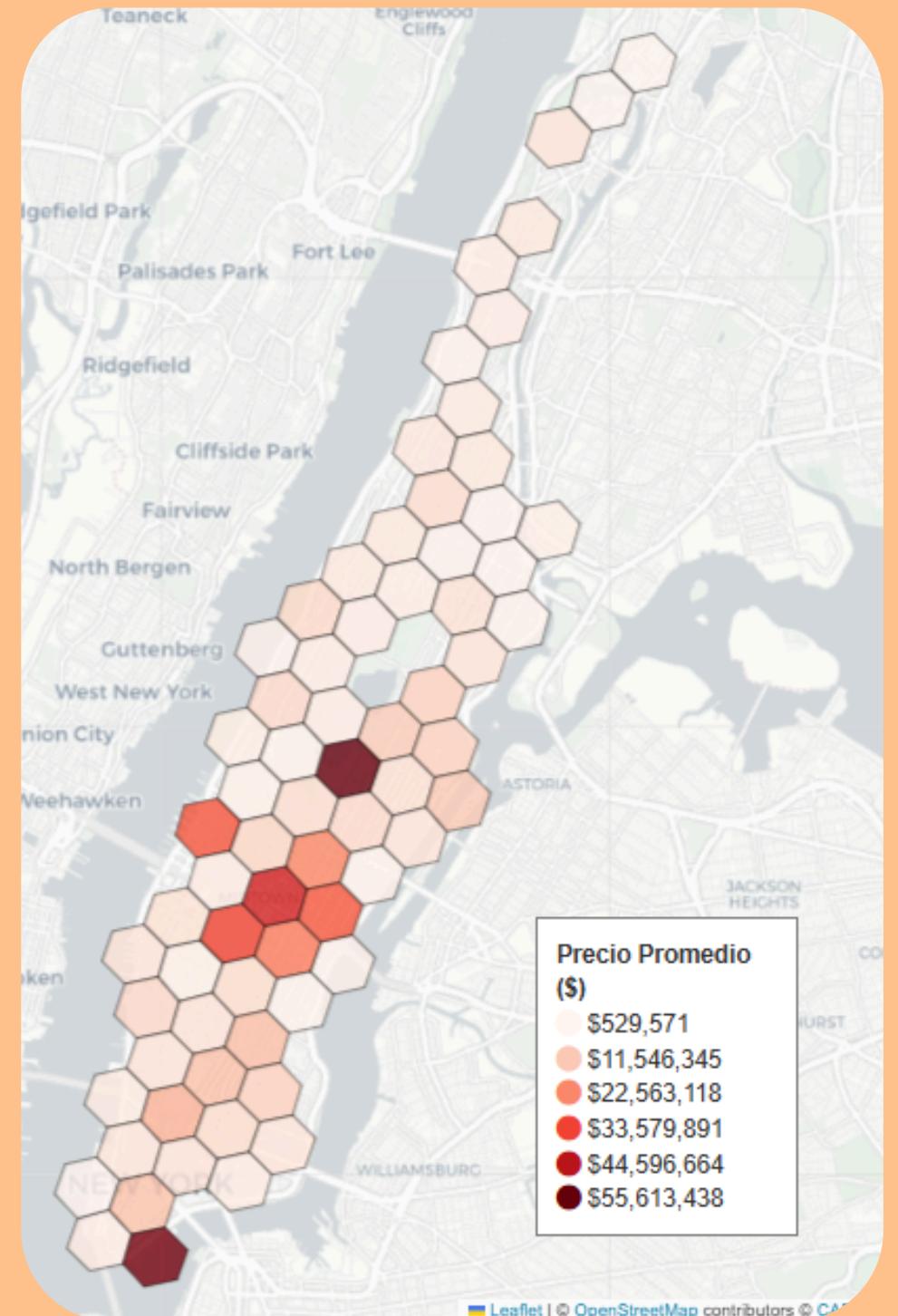


## Modelos

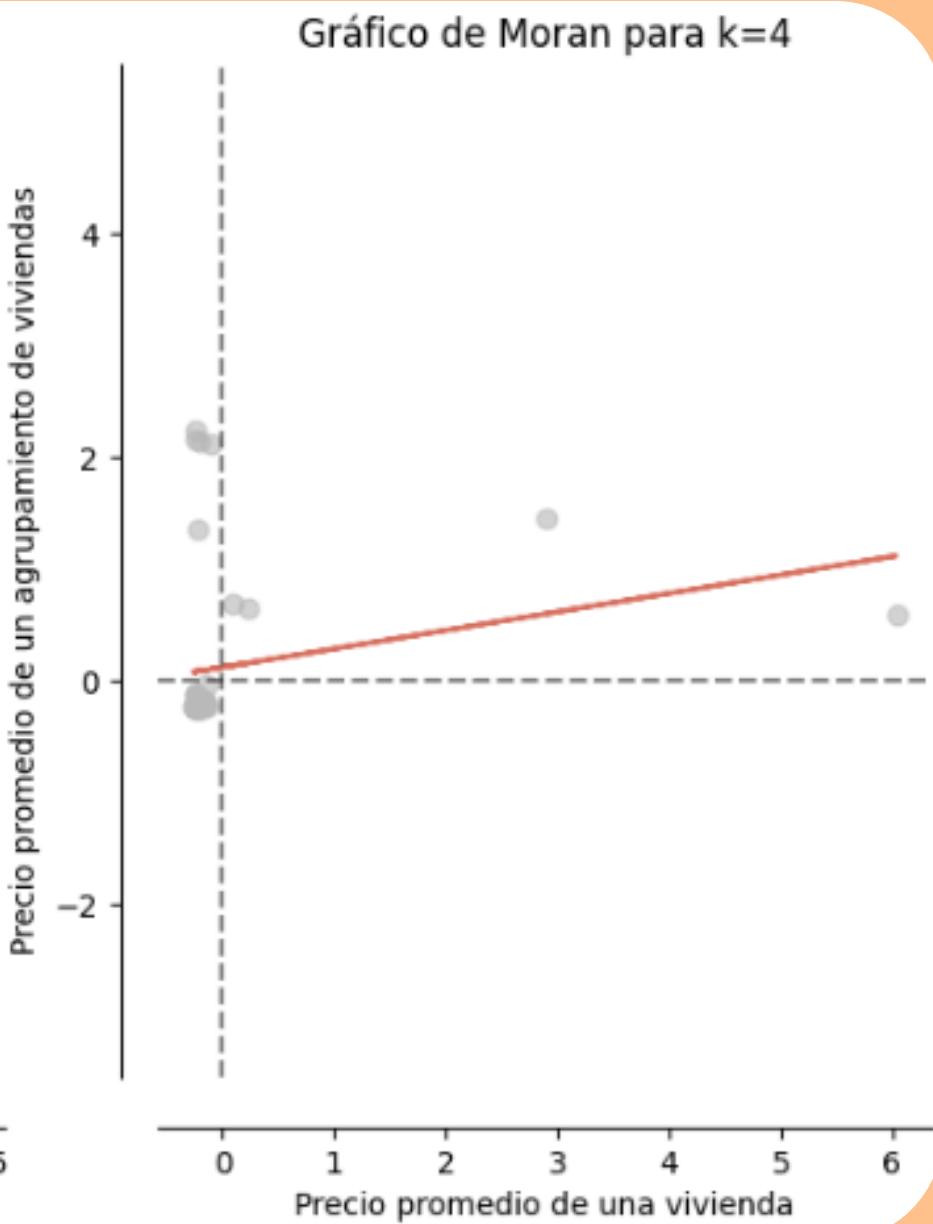
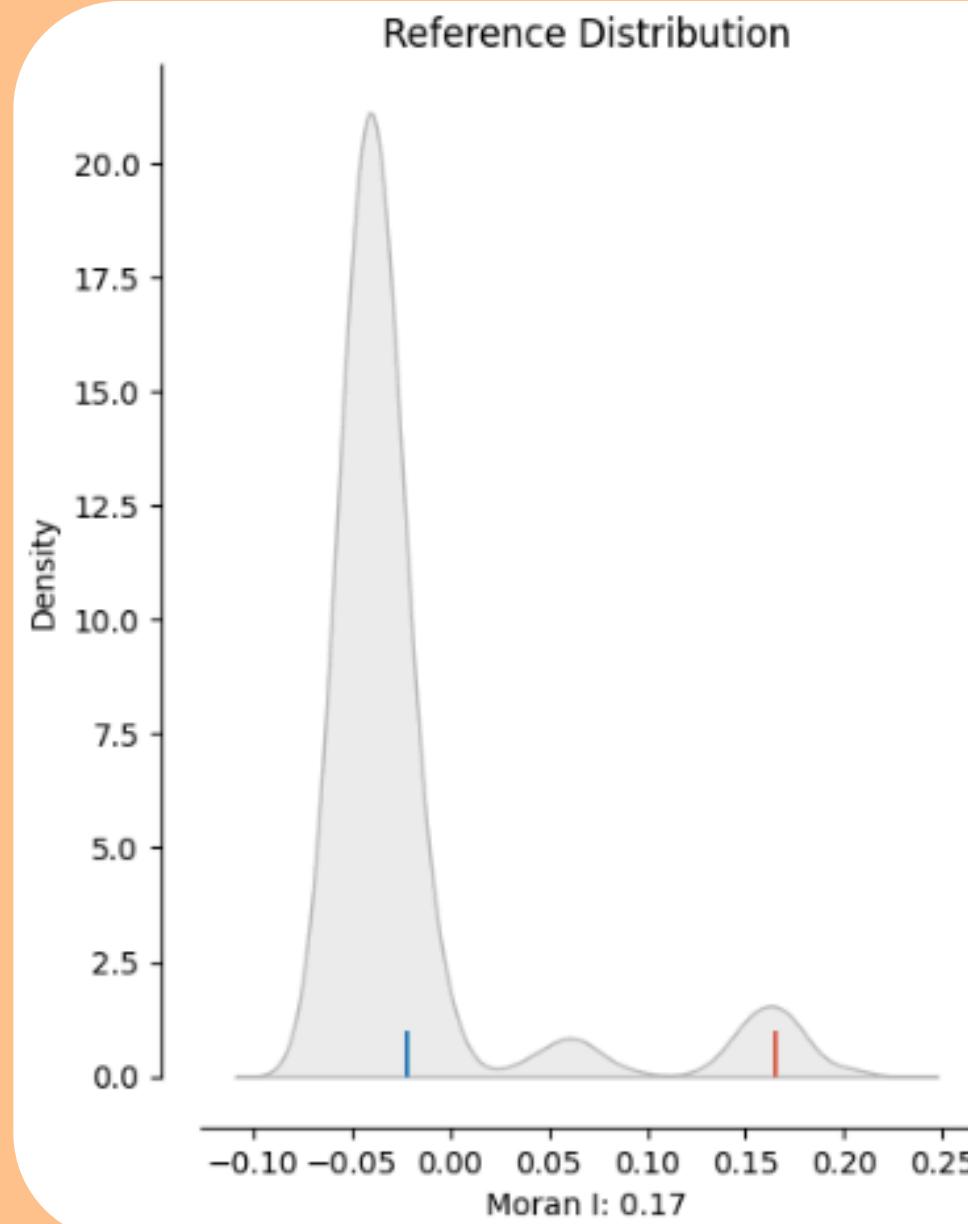
- DBSCAN
- Búsqueda de caminos
- Dijkstra
- Random Forests



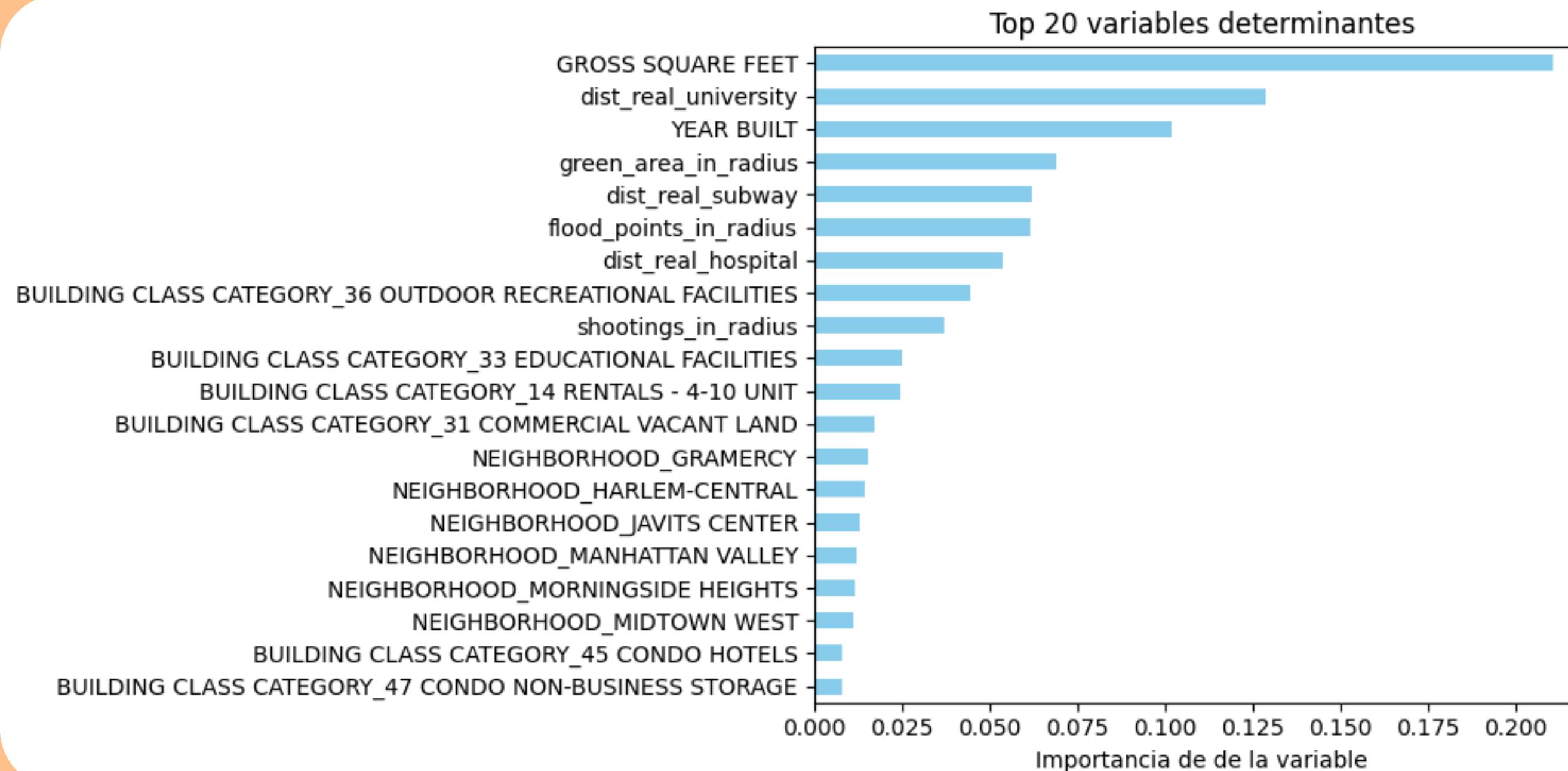
# Resultados obtenidos



# Resultados obtenidos



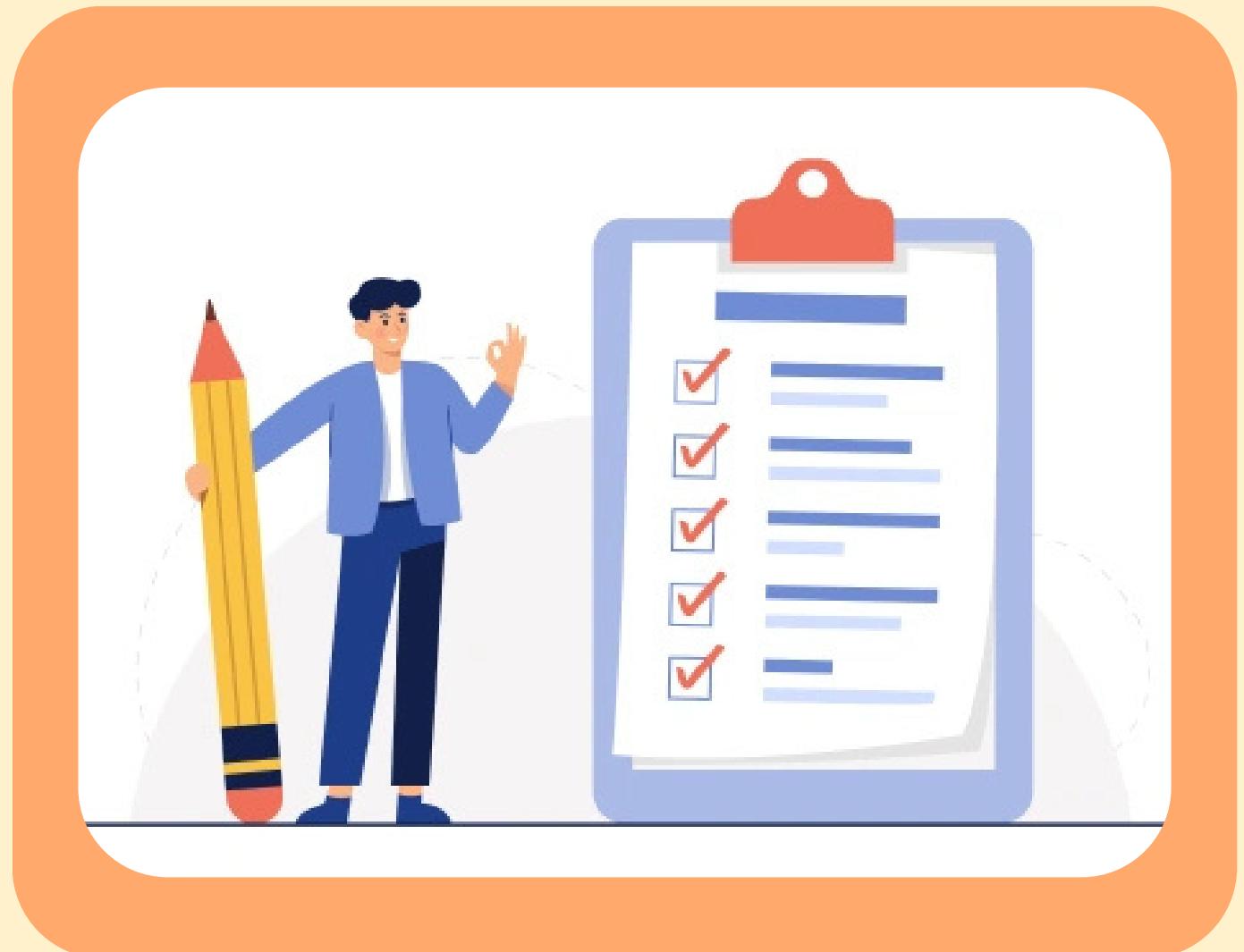
# Resultados obtenidos



## Conclusiones

Si bien los resultados son prometedores, existen diversas áreas en las que el proyecto puede perfeccionarse, entre las que destacan:

- **Ocupar una mayor cantidad de modelos y análisis para realizar estudios más efectivos**
- **Las distancias incluyen calles cerradas por infraestructuras de la ciudad**
- **Para lograr mejores resultados para el proyecto, sería necesario disponer de un conjunto de datos más amplio y equilibrado**



# Referencias bibliográficas:

- Tinsa. (2023). Manhattan, la isla más cara del mundo. <https://www.tinsa.es/blog/curiosidades/manhattan-la-isla-mas-cara-del-mundo/>
- Partir a Nueva York. (2023). Manhattan: información del distrito más famoso de Nueva York. <https://www.partir-a-new-york.com/es/distrito-de-nueva-york/manhattan-es>
- GIS and Beers. (2021). Cálculo del índice NDBI para análisis urbanísticos. <https://www.gisandbeers.com/calculo-indice-ndbi-analisis-urbanisticos/>
- J. Leal-Villamil. (2022). Análisis comparativo de asertividad para tres índices de zonas construidas aplicados a ciudades colombianas. Revista de Ingeniería, Universidad Pedagógica y Tecnológica de Colombia (UPTC). [https://revistas.uptc.edu.co/index.php/ingenieria\\_sogamoso/article/view/15018/12232](https://revistas.uptc.edu.co/index.php/ingenieria_sogamoso/article/view/15018/12232)
- Google Earth Engine. (2023). COPERNICUS S2 SR HARMONIZED – Google Earth Engine Dataset Catalog. [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S2\\_SR\\_HARMONIZED#bands](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED#bands)
- EPSG.io. (2023). EPSG Geodetic Parameter Dataset. <https://epsg.io/>
- Esri. (2023). How Spatial Autocorrelation (Moran's I) works. <https://pro.arcgis.com/es/pro-app/latest/tool-reference/spatial-statistics/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm>
- Esri. (2023). What is a z-score? What is a p-value? <https://pro.arcgis.com/es/pro-app/latest/tool-reference/spatial-statistics/what-is-a-z-score-what-is-a-p-value.htm>
- Scikit-learn developers. (2024). Scikit-learn: Machine learning in Python. <https://scikit-learn.org>
- GeoPandas. (2024). GeoDataFrame.dissolve. <https://geopandas.org/en/stable/docs/reference/api/geopandas.GeoDataFrame.dissolve.html>
- PySAL. (2024a). KNN – libpysal.weights.KNN. <https://pysal.org/libpysal/generated/libpysal.weights.KNN.html#libpysal.weights.KNN>
- PySAL. (2024). API Reference – libpysal. <https://pysal.org/libpysal/api.html>
- OSMnx. (2024). User Reference. <https://osmnx.readthedocs.io/en/stable/user-reference.html>
- NetworkX. (2024). multi source dijkstra path length – NetworkX. [https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.shortest\\_paths.weighted.multi\\_source\\_dijkstra\\_path\\_length.html](https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.shortest_paths.weighted.multi_source_dijkstra_path_length.html)
- H3. (2024). H3 Geoindexing System. <https://h3geo.org/>
- GeoPandas. (2024). GeoSeries.buffer. <https://geopandas.org/en/stable/docs/reference/api/geopandas.GeoSeries.buffer.html>
- La Nación. (2023, septiembre 29). Las intensas lluvias provocan inundaciones repentinas en Nueva York e interrupciones de servicios. <https://www.lanacion.com.ar/estados-unidos/las-intensas-lluvias-provocan-inundaciones-repentinas-en-nueva-york-e-interrupciones-de-servicios-nid29092023/>
- Spectrum Noticias NY. (2024, agosto 6). Ha comenzado la lluvia: Advierten de posibles inundaciones. Advertencia de viaje por tormentas eléctricas. <https://spectrumnoticias.com/ny/nyc/noticias/2024/08/06/ha-comenzado-la-lluvia--advierten-de-possibles-inundaciones--advertencia-de-viaje-por-tormentas-electricas>
- Scikit-learn developers. (2024). sklearn.ensemble.RandomForestRegressor. Scikit-learn. <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html>
- GeeksforGeeks. (2023). Random Forest Regression in Python. <https://www.geeksforgeeks.org/machine-learning/random-forest-regression-in-python/>
- Kim, A. (2021, marzo 23). Random Forest Regression - Clearly Explained!. YouTube. <https://www.youtube.com/watch?v=YUsx5ZNIYWc>
- StatQuest with Josh Starmer. (2018, diciembre 4). Random Forests Part 1 – Building, using and evaluating. YouTube. <https://youtu.be/AYIClq5jnhU>
- StatQuest with Josh Starmer. (2018, diciembre 4). Random Forests Part 2 – Out-of-bag error, feature importance and more!. YouTube. <https://youtu.be/kFwe2ZZU7yw?t=718>
- Moonbooks. (n.d.). How to clear the output in a Jupyter Notebook cell after each for loop iteration. <https://en.moonbooks.org/Articles/How-to-clear-the-output-in-a-Jupyter-Notebook-cell-after-each-for-loop-iteration-/>
- IPython. (n.d.). IPython.display — IPython 9.0.2 documentation. <https://ipython.readthedocs.io/en/9.0.2/api/generated/IPython.display.html>
- Akyürek, B. (2020, junio 10). Map Visualization with Folium. <https://medium.com/datascienceearth/map-visualization-with-folium-d1403771717>