



REPÚBLICA DEL ECUADOR

Escuela Politécnica Nacional

" E SCIENTIA HOMINIS SALUS "

La versión digital de esta tesis está protegida por la Ley de Derechos de Autor del Ecuador.

Los derechos de autor han sido entregados a la "ESCUELA POLITÉCNICA NACIONAL" bajo el libre consentimiento del (los) autor(es).

Al consultar esta tesis deberá acatar con las disposiciones de la Ley y las siguientes condiciones de uso:

- Cualquier uso que haga de estos documentos o imágenes deben ser sólo para efectos de investigación o estudio académico, y usted no puede ponerlos a disposición de otra persona.
- Usted deberá reconocer el derecho del autor a ser identificado y citado como el autor de esta tesis.
- No se podrá obtener ningún beneficio comercial y las obras derivadas tienen que estar bajo los mismos términos de licencia que el trabajo original.

El Libre Acceso a la información, promueve el reconocimiento de la originalidad de las ideas de los demás, respetando las normas de presentación y de citación de autores con el fin de no incurrir en actos ilegítimos de copiar y hacer pasar como propias las creaciones de terceras personas.

***Respeto hacia sí mismo y hacia los demás.***

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE INGENIERÍA DE SISTEMAS**

**APLICANDO LA ANALÍTICA DE DATOS PARA ANALIZAR Y  
VISUALIZAR EL FENÓMENO DE GENTRIFICACIÓN COMO  
HERRAMIENTA PARA LA TOMA DE DECISIONES  
ESTRATÉGICAS EN CIUDADES INTERMEDIAS DE ECUADOR.**

**PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO DE  
MAESTRÍA EN SISTEMAS DE INFORMACIÓN MENCIÓN  
INTELIGENCIA DE NEGOCIOS Y ANALÍTICA DE DATOS MASIVOS**

**LUIS ENRIQUE YULÁN MENDOZA**  
**luis.yulan@epn.edu.ec**

**DIRECTOR: Ing. Tania Calle, PhD.**  
**tania.calle@epn.edu.ec**

**Quito, 2023**

## DECLARACIÓN

Yo, Luis Enrique Yulán Mendoza, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

---

Luis Enrique Yulán Mendoza

## **CERTIFICACIÓN**

Certifico que el presente trabajo fue desarrollado por Luis Enrique Yulán Mendoza, bajo mi supervisión.

---

**Ing. Tania Calle, PhD**  
**DIRECTOR DE PROYECTO**

## **AGRADECIMIENTO**

Mi agradecimiento primeramente a Dios, por llenarme de bendiciones, permitiéndome superar cada uno de los obstáculos que se me han cruzado a lo largo de mi vida.

A mi mamita, por todo su amor para criarme como padre y madre a la vez, apoyándome constantemente para que pudiera culminar cada una de las etapas de mi vida con éxito.

A mí novia Shirley que me ha apoyado constantemente a lo largo de mi carrera dándome fuerzas para poder seguir semestres de mi maestría, gracias a su amor y apoyo incondicional.

Agradezco también Tania Calle por haberme brindado su apoyo, confianza y amistad para que pudiera desarrollar y culminar este proyecto.

Finalmente, a la Escuela Politécnica Nacional y sus profesores por todo el tiempo que dedicaron para mi formación profesional.

## **DEDICATORIA**

Todo esfuerzo y sacrificio realizado en este proyecto de titulación va dedicado a Dios y mi familia, por entregarme todo su apoyo y amor necesarios para que pueda superar los obstáculos que se presentaron a lo largo de mi del desarrollo de mi Maestría.

## ÍNDICE DE CONTENIDO

<b>DECLARACIÓN.....</b>	<b>II</b>
<b>CERTIFICACIÓN .....</b>	<b>III</b>
<b>AGRADECIMIENTO .....</b>	<b>IV</b>
<b>DEDICATORIA .....</b>	<b>V</b>
<b>ÍNDICE DE CONTENIDO.....</b>	<b>VI</b>
<b>RESUMEN .....</b>	<b>XI</b>
<b>CAPÍTULO I.....</b>	<b>1</b>
<b>1. INTRODUCCIÓN.....</b>	<b>1</b>
1.1 Problema. ....	4
1.2 Justificación. ....	5
1.3 Objetivo General. ....	6
1.4 Objetivos Específicos. ....	6
1.5 Marco Teórico. ....	7
1.5.1 Análisis de Datos. ....	7
1.5.2 Conjunto de Datos. ....	7
1.5.3 Aprendizaje de Máquina. ....	8
1.5.4 Web Scraping. ....	10
1.5.5 Sistemas de Información Geográfica (SIG). ....	11
1.5.6 Ciudad Intermedia. ....	11
1.6 Estado del Arte. ....	11
1.6.1 Gentrificación en Ecuador.....	11
1.6.2 Modelos de Gentrificación.....	14
1.6.2.1 Modelo De Hammel Y Wyly.....	15
1.6.2.2 Modelo De Freeman.....	15
1.6.2.3 Modelo De Nesbitt. ....	15
1.6.2.4 Modelo de Chapple, Waddell, Chatman, Loukaitou-Siders.....	16
1.6.3 Herramientas SIG. ....	18
1.6.3.1 Elementos de un Sistema SIG. ....	19
1.6.3.2 Análisis De Datos En Sistemas SIG. ....	21
1.6.4 Herramientas de SIG. ....	22
<b>CAPÍTULO II.....</b>	<b>24</b>

<b>2. Implementación De Modelo.....</b>	<b>24</b>
2.1 Fase 1: Extracción y Obtención de Datos. ....	24
2.1.1 Extracción de datos mediante Web scraping. ....	28
2.1.2 Extracción de datos mediante Urbanpy.....	31
2.2 Fase 2: Limpieza de Datos. ....	33
2.3 Fase3: Visualización de Datos.....	34
2.3.1 Creación De Zonas. ....	38
2.4 Fase 4: Desarrollo de Método Automático. ....	39
2.5 Fase 5: Validación y Evaluación del Modelo.....	48
2.5.1 Coeficiente de Determinación $R^2$ .....	48
2.5.2 Error Cuadrático Medio RMSE. ....	49
2.5.3 Coeficiente de Variación de Pearson (CV). ....	49
2.6 Fase 6: Presentación de Resultados por Ciudad.....	50
<b>CAPÍTULO III.....</b>	<b>52</b>
<b>3. Diseño e Implementación del Dashboard .....</b>	<b>52</b>
3.1 Registro. ....	52
3.2 Creación de Mapa Base. ....	53
3.3 Carga de Capas de ArcGIS Desktop a ArcGIS Online.....	54
3.4 Generación de Mapas Temáticos por Ciudad. ....	57
3.5 Análisis de Resultados .....	59
3.6 Discusión .....	61
3.7 Diagnósticos de la problemática para gobiernos autónomos. ....	62
<b>CAPÍTULO IV.....</b>	<b>65</b>
<b>4. Conclusiones y Recomendaciones .....</b>	<b>65</b>
4.1 Conclusiones.....	65
4.2 Recomendaciones .....	66
<b>ANEXOS.....</b>	<b>A-1</b>

## LISTA DE FIGURAS

Figura 1.1 Etapas de la gentrificación. Fuente: ARQUITASA [2].....	2
Figura 1.2 Conjunto de entrenamiento en aprendizaje supervisado. Fuente: Diego Calvo [18]. .....	8
Figura 1.3 Ejemplo de uso de un modelo de clasificación utilizando aprendizaje supervisado. Fuente: Diego Calvo [18].....	9
Figura 1.4 Ejemplo de aprendizaje no supervisado. Fuente: Nilton Guerra [19].....	9
Figura 1.5 Red de redes neuronales en Deep learning. Fuente: Nilton Guerra [19].....	10
Figura 1.6 Esquema de Web scraping. Fuente: Autoría propia.....	10
Figura 1.7 Esquema de un SIG con sus tres subsistemas fundamentales. Fuente: Autoría Propia.....	18
Figura 1.8 Formato raster en sistemas SIG. Fuente: Autoría propia. ....	20
Figura 1.9 Representación vectorial de puntos, líneas y polígonos. Fuente: Autoría propia.....	20
Figura 1.10 Análisis de proximidad. Fuente: Esri [21]. ....	22
Figura 1.11 Análisis de superposición. Fuente: Esri [21]. ....	22
Figura 2.1 Fases de procesos a implementar para medir la gentrificación en una ciudad intermedia. Fuente: Autoría propia. ....	24
Figura 2.2 Descarga de data empataada nacional censo 2010. Fuente: Autoría propia. ....	25
Figura 2.3 Software Redatam Instalado. Fuente: Autoría propia .....	25
Figura 2.4 Visualización jerárquica de censo en Redatam. Fuente: Autoría propia. ....	26
Figura 2.5 Selección de cantón analizar. Fuente: Autoría propia. ....	26
Figura 2.6 Categoría ocupacional. Fuente: Autoría propia.....	27
Figura 2.7 Tabla generada en software Redatam para análisis de estado económicos. Fuente: Autoría propia .....	27
Figura 2.8 Exploración de estructura HTML de la página. Fuente: Autoría propia. ....	28
Figura 2.9 Geocodificar direcciones en latitud y longitud. Fuente: Autoría propia. ....	30
Figura 2.10 Concatenación de nombre de ciudad en columna de dirección. Fuente: Autoría propia.....	30
Figura 2.11 Archivo obtenido de Web scraping. Fuente: Autoría propia. ....	31
Figura 2.12 Límite urbano de una ciudad con urbanpy. Fuente: Autoría propia. ....	32
Figura 2.13 Tabla de acceso alimentos de una ciudad con urbanpy. Fuente: Autoría propia.....	33
Figura 2.14 Filtro por lugares de interés dentro de una ciudad utilizando urbanpy. Fuente: Autoría propia.....	33

<b>Figura 2.15 Selección de témpate de mapa en ArcGIS pro. Fuente: Autoría propia. ....</b>	<b>34</b>
<b>Figura 2.16 Ejecución de la herramienta XY a punto. Fuente: Autoría propia. ....</b>	<b>35</b>
<b>Figura 2.17 Ingreso de tabla dentro de mapa en ArcGIS Pro. Fuente: Autoría propia. ....</b>	<b>35</b>
<b>Figura 2.18 Tabla con visualización de puntos ArcGIS Pro. Fuente: Autoría propia. ....</b>	<b>36</b>
<b>Figura 2.19 Relación de variables en base al coeficiente de Pearson. Fuente: Autoría propia.....</b>	<b>37</b>
<b>Figura 2.20 Creación de zona limítrofe de una ciudad intermedia. Fuente: Autoría propia.....</b>	<b>38</b>
<b>Figura 2.21 Tabla relacional generada por herramienta Spatial Join. Fuente: Autoría propia.....</b>	<b>39</b>
<b>Figura 2.22 Creación de clases para variable precio. Fuente: Autoría propia. ....</b>	<b>40</b>
<b>Figura 2.23 Creación de modelo de regresión. Fuente: Autoría propia. ....</b>	<b>41</b>
<b>Figura 2.24 Herramienta de análisis para realizar regresión lineal en los datos de tabla. Fuente: Autoría propia .....</b>	<b>41</b>
<b>Figura 2.25 Ingreso de variables dependientes e independientes en el modelo. Fuente: Autoría propia. ....</b>	<b>42</b>
<b>Figura 2.26 Model builder de regresión lineal en base a características de modelo de gentrificación. Fuente: Autoría propia.....</b>	<b>42</b>
<b>Figura 2.27 identificación de zonas en base a variables de gentrificación. Fuente: Autoría propia.....</b>	<b>43</b>
<b>Figura 2.28 Ingreso de variables a clúster. Fuente: Autoría propia.....</b>	<b>43</b>
<b>Figura 2.29 Resultado de ejecución de script. Fuente: Autoría propia. ....</b>	<b>44</b>
<b>Figura 2.30 Creación de notebook en ArcGIS PRO. Fuente: Autoría propia. ....</b>	<b>45</b>
<b>Figura 2.31 Relación entre predicción y valor real. Fuente: Autoría propia..</b>	<b>48</b>
<b>Figura 3.1 Asignación de rol publicador. Fuente: Autoría propia.....</b>	<b>53</b>
<b>Figura 3.2 Selección de mapa base ArcGIS Online. Fuente: Autoría propia. ....</b>	<b>53</b>
<b>Figura 3.3 Ajustes de permisos mapa base. Fuente: Autoría propia.....</b>	<b>54</b>
<b>Figura 3.4 Sincronización de ArcGIS Desktop con ArcGIS Online. Fuente: Autoría propia.....</b>	<b>55</b>
<b>Figura 3.5 Compartición de Capa en ArcGIS Desktop. Fuente: Autoría propia. ....</b>	<b>55</b>
<b>Figura 3.6 Publicar capa en ArcGIS ONLINE. Fuente: Autoría propia. ....</b>	<b>56</b>
<b>Figura 3.7 Agregar capa a mapa base. Fuente: Autoría propia.....</b>	<b>56</b>
<b>Figura 3.8 Galería de ciudades Intermedias ArcGIS Online. Fuente: Autoría propia. ....</b>	<b>57</b>
<b>Figura 3.9 Resultados de zonas gentrificadas para la ciudad de Ambato. Fuente: Autoría propia.....</b>	<b>58</b>
<b>Figura 3.10 Turistificación en la ciudad de Manta. Fuente: Autoría propia... ..</b>	<b>59</b>

<b>Figura 3.11 Gentrificación en ciudades de la Sierra. Fuente: Autoría propia.</b>	<b>61</b>
<b>Figura 3.12 Despliegue de puntos de interés en zonas de gentrificación de una ciudad.</b>	<b>63</b>

## LISTA DE TABLAS

<b>Tabla 1 Precio de metro cuadrado en la parroquia de Cumbayá. Fuente: Revista Bienes Raíces 2011.</b>	<b>12</b>
<b>Tabla 2 Comparativa de los modelos de análisis del fenómeno de gentrificación.</b>	<b>17</b>
<b>Tabla 3 Diferencias entre Formato Raster y Vectorial.</b>	<b>21</b>
<b>Tabla 4 Comparativa de aplicaciones de escritorio para uso de SIG.</b>	<b>23</b>
<b>Tabla 5 Librerías cargadas en ArcGIS.</b>	<b>45</b>
<b>Tabla 6 Resultados obtenido por modelos ML.</b>	<b>50</b>

## LISTA DE CÓDIGOS

<b>Código 2.1 Importación de campo nombre a través de Web scraping</b>	<b>29</b>
<b>Código 2.2 Extracción múltiple variable</b>	<b>29</b>
<b>Código 2.3 Nombre de campos</b>	<b>29</b>
<b>Código 2.4 Librerías para acceder a datos urbanpy.</b>	<b>31</b>
<b>Código 2.5 Limitación de librería a ciudad</b>	<b>31</b>
<b>Código 2.6 Ejecución de K_means en Python de ArcGIS pro.</b>	<b>44</b>
<b>Código 2.7 Librería para código Deep learning.</b>	<b>46</b>
<b>Código 2.8 Búsqueda y eliminación de datos vacíos.</b>	<b>46</b>
<b>Código 2.9 Modelo de capas.</b>	<b>47</b>
<b>Código 2.10 Validación de modelo.</b>	<b>47</b>
<b>código 2.11 Verificación de modelo utilizando Deep learning.</b>	<b>47</b>

## RESUMEN

La presente investigación plantea visualizar zonas con gentrificación que es un proceso de regeneración urbanística cuya finalidad aprovecha del alza de precios debido a la especulación; lo cual provoca el desplazamiento de la población original de un barrio por una de un mayor poder adquisitivo. Se realizó un enfoque especial en ciudades intermedias que son aquellas que tienen un aproximado de 50.000 a 1 millón de habitantes con una actividad económica única y representativa de Ecuador que para el año 2030 tendrán un protagonismo mayor ya que las poblaciones se van incrementando y tiene movimientos mercantiles representativas como: sitios turísticos, actividades industriales, exportación de frutas, flores, ganadería o ensamblaje de algún tipo de electrodoméstico. Estas ciudades en su totalidad no se encuentran gentrificadas evidenciando únicamente la gentrificación en cascos comerciales como son el centro de la ciudad a diferencia de las metrópolis como Quito y Guayaquil que tiene zonas gentrificadas de norte a sur en varios de sus barrios.

Para ello se planteó un análisis de datos sobre las ciudades intermedias donde se represente las zonas gentrificadas para que los gobiernos centrales eviten tener gastos excesivos en la creación de nuevos barrios y centros de salud dando oportunidad de crear una política de buen vivir que todavía se estaría a tiempo cosa que no paso en las metrópolis. Con el fin de brindar una solución a la toma de decisión por parte de los gobiernos centralizados a esta problemática de la gentrificación se extrajeron datos censales junto con variables de sitios web para obtener el valor de la plusvalía de una zona y permitir predecir qué pasaría en dicha zona utilizando herramientas de visualización geográfica mediante el uso de mapas donde se ha sintetizado la acumulación de locales comerciales, sitios turísticos y hoteles de lujo donde se centra el fenómeno de la gentrificación dando el poder de decisión a las autoridades mediante la visualización de la zona de frenar el aumento desmedido de la plusvalía en unos pocos años.

**Palabras claves:** Gentrificación, Sistemas de Información Geográfica, Analítica de datos y machine learning

## ABSTRACT

This research proposes to visualize areas with gentrification, which is a process of urban regeneration whose purpose is to take advantage of rising prices due to speculation, which causes the displacement of the original population of a neighborhood for one with a higher purchasing power. A special focus was made on intermediate cities, which are those with an approximate population of 50,000 to 1 million inhabitants with a unique and representative economic activity in Ecuador that by the year 2030 will have a greater role as populations are increasing and have representative mercantile movements such as: tourist sites, industrial activities, export of fruits, flowers, livestock, or assembly of some type of household appliances. These cities are not gentrified, evidencing gentrification only in commercial areas such as the city center, unlike metropolis such as Quito and Guayaquil, which have gentrified areas from north to south in several of their neighborhoods.

To this end, an analysis of data on the intermediate cities was proposed where the gentrified areas are represented so that the central governments avoid having excessive expenses in the creation of new neighborhoods and health centers, giving the opportunity to create a policy of good living that would still be on time, something that did not happen in the metropolises. In order to provide a solution to the decision making by centralized governments to this problem of gentrification, census data was extracted along with variables from websites to obtain the value of the surplus value of an area and allow predicting what would happen in that area using geographic visualization tools through the use of maps where the accumulation of commercial premises has been synthesized, tourist sites and luxury hotels where the phenomenon of gentrification is centered giving the power of decision to the authorities through the visualization of the area to stop the excessive increase of the capital gain in a few years.

**Keywords:** Gentrification, Geographic Information Systems, Data analytics and machine learning.

# CAPÍTULO I

## 1. INTRODUCCIÓN.

El termino gentrificación fue acuñado por primera vez en el año de 1964 por Ruth Glass para identificar el desplazamiento de la clase obrera británica hacia sectores aledaños provocado por la regeneración urbana del sector que lo había convertido en un sitio de alta demanda que se resulta atractivo para turistas y clases sociales aburguesadas provocando que sus antiguos moradores tuvieran que migran por el alza de renta en los arriendos y locales comerciales [1].

La gentrificación es un fenómeno que no resulta de la espontaneidad suele ser un proceso gradual que consta de cuatro etapas debidamente planeadas por grandes grupos inmobiliarios:

- **Degradación:** Parte con el reconocimiento de un barrio humilde que se convierte en atractivo para inversores debido al desarrollo urbano de la ciudad. En donde crean campañas negativas sobre el barrio sobre delincuencia y venta de drogas provocando la carencia de dicho barrio; esto provoca que varios dueños de casas quieran vender sus propiedades a precios módicos.
- **Encarecimiento:** Una vez adquiridos lo bienes inmuebles, se da la segunda etapa de la gentrificación en donde se plantea edificar locales con nuevos elementos vanguardistas.
- **Expulsión:** En esta tercera etapa, se da a conocer por prensa y televisión los nuevos proyectos vanguardistas cosa que convierte a este barrio en algo novedoso debido a los nuevos proyectos inmobiliarios. Además, se da la apertura de locales comerciales donde antes eran viviendas creando una zona comercial. Esto provoca el aumento de precios de vivienda dejando a los antiguos arrendatarios del barrio sin poder costear los nuevos precios creando la migración hacia zonas más económicas.

- **Comercialización:** Finalmente, en esta etapa ya se encuentra todo debido estructurado en una zona comercial que seguirá creciendo de forma gradual a lo largo de la zona [1].

En la figura 1.6 se puede visualizar cómo funciona el proceso de la gentrificación a lo largo de sus cuatro etapas.

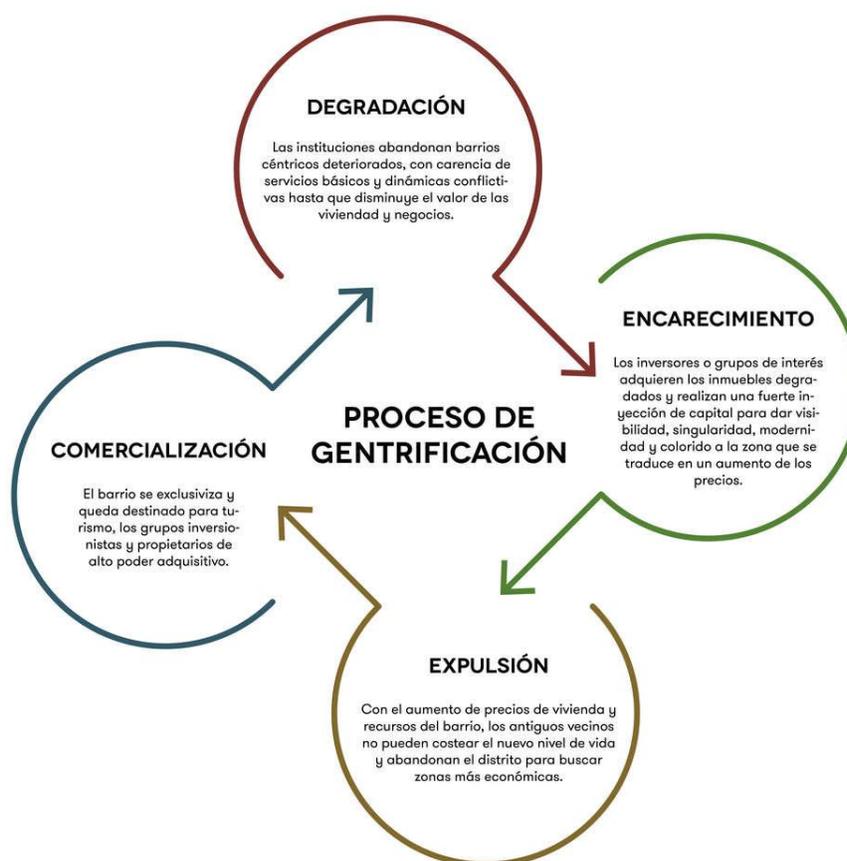


Figura 1.1 Etapas de la gentrificación. Fuente: ARQUITASA [2]

La gentrificación junto con sus etapas se ha ido adaptando a los cambios creando nuevas formas de realizarlas. En el siglo XXI se potencia a través de la creación de lugares turísticos, dando lugar el concepto de turistificación donde de lleno se plantea convertir a un barrio en un lugar dedicado de lleno al turismo. Esto es muy común en

zonas costeras donde se crean grandes atractivos hoteleros junto con centros comerciales.

Entre algunos ejemplos de gentrificación y turistificación se pueden citar a varias ciudades:

- **Paris (Francia):** Dentro de esta ciudad cada año migran alrededor de 12.000 habitantes a zonas totalmente aledañas porque los precios de arriendo en el casco principal se han duplicado en los últimos diez años. Una de las principales causas son los museos y tiendas de moda internacional lo que convierte a sus sectores en zonas llamadas “instagrameables” (sector digno de postear en redes sociales). Varias ciudades europeas copian mucho la situación que aplica sucede en Francia.
- **Nueva York:** Situado en el continente americano es un claro ejemplo de gentrificación porque los arriendo en manhattan varían entre un 10% y 20 % de acuerdo con la ubicación de lugares turísticos que se van regenerando.
- **Colonia Condesa (México):** Barrio ubicado en la ciudad de México se convirtió en un epicentro de diversión en la vida nocturna obligando a sus habitantes a buscar un lugar más tranquilo. La gentrificación se fue extendiendo hasta otra colonia llamada Roma creando una zona rosa denominada “corredor Roma-Condesa”.
- **Cartagena de Indias (Colombia):** Antes de ser una zona turística para Sudamérica era puerto de pescadores. Hoy en día existe una gran cadena de hoteles y se ofertan muchos paquetes turísticos a nivel Internacional [3].
- **Santiago de Chile:** Dentro de esta ciudad se dio un Oligopolio inmobiliario que provoco el desplazamiento de miles de personas en la década de los 90, esto fue provocado con complicidad el gobierno. Dando lugar a muchos actos de corrupción [3].
- **Buenos Aires (Argentina):** Es una de las ciudades más grandes de Latinoamérica, siempre presenta significativos cambios en la zona urbana a

partir de la reestructuración económica neoliberal, se evidencian casos de inversión inmobiliaria grandes dentro del área central de la ciudad, con

- maniobras distintas por parte del gobierno y el municipio local como son:
  - 1) Las transformaciones sucedidas en el antiguo mercado “El Abasto”.
  - 2) Las diversas inversiones en el barrio de “Palermo”.

En base a estos ejemplos se puede evidenciar que la teoría de la gentrificación ofrece un potente lente analítico para vislumbrar la economía urbana en un contexto de predominio de mercado en base a ganancias por medio de la plusvalía [4]. La gentrificación en sí es un problema de extracción y comercialización desigual de la ganancia obtenida por el uso del suelo, y una distribución desigual en el acceso a los bienes públicos urbanos localizados en la periferia central o casco colonial; todo ello posibilitado, y a veces conducido, por el estado y municipios locales.

### **1.1 Problema.**

El desarrollo y expansión de las ciudades en Ecuador, ha provocado un reordenamiento de sectores por clases creando cada vez más una disparidad en las ciudades donde la clase obrera tiende a vivir en las periferias [5]. Esta forma de reordenamiento por clases trae consigo un fenómeno conocido como gentrificación, cuyo significado es una adaptación adecuada al español del término inglés *gentrification*, que consiste en el desplazamiento de una población original de un sector o barrio con características populares de clase media a baja por uno de mayor nivel adquisitivo desplazando a los habitantes originales del barrio a un sector marginal o limítrofe de la ciudad [6]. Este desplazamiento es provocado por la falta de capacidad adquisitiva que no tendría las clases bajas debido al alza de arriendos de casas y locales comerciales de manera desmesurada provocados por la creación de conjuntos habitacionales por parte de la inversión inmobiliaria sin control de los alcaldía y gobiernos provinciales [7].

El fenómeno de la gentrificación debería ser controlado para que las ciudades no se extiendan en áreas rurales limítrofes y creen un mayor tiempo de desplazamiento de dichos sectores al casco financiero donde se sitúa la demanda de fuerza laboral

generando mayores gastos para la ciudadanía y los gobiernos provinciales porque deben crear nuevas rutas de buses urbanos, expansión de sistemas de alcantarillado, nuevas líneas eléctricas, y centros de salud [8].

Las tres principales ciudades de Ecuador Quito, Guayaquil y Cuenca la mayoría de sus barrios han sido gentrificadas debido a que entre la clase media y baja debe recorrer para llegar a su lugar de trabajo por medio de línea de bus más de una hora, y se observa que existe una gran diferencia social entre el norte y sur de dichas ciudades [9]. Por ello, las entidades de gobierno deberían tomar cartas en el asunto para reducir el gasto público en nuevas obras para el sector rural [10].

La poca experiencia y conocimiento sobre el tema de gentrificación ha provocado que grupos capitalistas llamados gentrificadores puedan aprovechar de espacios que son denominados vulnerables creando una oportunidad de inversión, pues se compra bienes inmuebles o terrenos a un precio muy económico para luego con una pequeña inversión crear un sitio gentrificado que les da ganancias por medio de la plusvalía [11]. Es por ello, que dentro de este trabajo se propone realizar un análisis de datos aplicados a las plusvalía de los terrenos cuyo enfoque principal este basado en el fenómeno de la gentrificación donde se pueda evidenciar zonas afectas a través de un sistema de información geográfica (SIG) con ARGIS aplicando herramientas de machine learning y analítica de datos obtenidas de fuentes como: google maps, airbnb, y trivago, las cuales serán plasmadas en un artículo científico de alto nivel que detalle el fenómeno de la gentrificación en ciudades intermedias de Ecuador para que se conozca sobre el fenómeno y se pueda considerar tomar alguna medida por parte de los gobiernos locales [12].

## **1.2 Justificación.**

En la actualidad, las ciudades no cuentan con un plan de crecimiento urbano-rural. Por lo general, el nacimiento de un barrio se da por medio de invasiones, la cuales son provocadas por la gentrificación esto a su vez causa problemas a los gobiernos provinciales debido a que tienen que extender sus servicios hacia zonas más aledañas

de la ciudad en donde se han creado sectores vulnerables en cuanto a delincuencia, salud y escolaridad [11].

Bajo esta perspectiva, se debería controlar la gentrificación a través de la creación leyes que permitan tener una equidad de clases sociales, y mejor distribución de tierras evitando en lo posible las invasiones de estas dando lugar a un esquema cantonal y provincial planificado para su crecimiento donde se plantee la creación de barrios con acceso a transporte público, escuelas, UPC (Unidades de Policía Comunitaria) y centros de salud [13].

Para evidenciar la gentrificación en ciudades intermedias de Ecuador se plantea utilizar herramientas actuales que permitan predecir los sectores que van a ser gentrificados siendo necesario recurrir a técnicas machine learning con Python que permita aplicar la analítica de datos sobre fuentes de sitios turísticos mencionadas en el punto anterior [14].

Para tener una mayor especificación e identificación de las zonas que se están gentrificando se plantea elaborar un dashboard que despliegue sectores de una ciudad evidenciando por colores la gentrificación, creando de esta manera una iniciativa para evidenciar el fenómeno y fomentar el desarrollo de una geografía urbana crítica [15].

### **1.3 Objetivo General.**

Aplicar la analítica de datos en SIG para visualizar y analizar el fenómeno de gentrificación en ciudades intermedias de Ecuador para la toma de decisiones estratégicas.

### **1.4 Objetivos Específicos.**

- Realizar una revisión de literatura, enfocada en el estudio de contribuciones relacionadas con el fenómeno de gentrificación mediante analítica de datos para ciudades intermedias del Ecuador.
- Realizar una evaluación sobre la situación actual de los modelos de gentrificación en diferentes ciudades del mundo.

- Elaborar un marco comparativo entre herramientas de analítica de datos y SIG y seleccionar la herramienta más idónea para representar el fenómeno de gentrificación.
- Diseñar un modelo para obtener datos sobre plusvalía con Python para clasificar de acuerdo con zonas de mayor cotización con referencia a las más accesibles en ciudades intermedias de Ecuador.
- Desplegar un dashboard del fenómeno de gentrificación de las Ciudades Intermedias de Ecuador incorporando datos actuales.
- Diagnosticar la problemática que genera la gentrificación para la toma de decisiones estratégicas a los gobiernos autónomos descentralizados (GAD).
- Elaborar artículo científico con los resultados obtenidos.

## **1.5 Marco Teórico.**

Para tener un mejor entendimiento sobre el tema desarrollado, se describe a continuación algunos de los términos utilizados durante el desarrollo del presente proyecto.

### **1.5.1 Análisis de Datos.**

El análisis de datos en la última década ha asumido un rol importante al momento de tomar decisiones, analizar fenómenos y diagnosticar enfermedades. Debido a que el análisis de datos consiste en aplicar operaciones estadísticas de tipo inferenciales, cuyo propósito es obtener conclusiones precisas que permitan alcanzar objetivos. Para ello se debe recolectar la mayor información posible sobre la variable analizar en un conjunto de datos [16].

### **1.5.2 Conjunto de Datos.**

Un conjunto de datos (del inglés dataset) hace relación al contenido de una tabla en específico de una base de datos. En donde, cada columna representa a una variable en particular y cada fila es el contenido de las variables.

### 1.5.3 Aprendizaje de Máquina.

Según Ethem Alpaydin [17], machine learning es una parte de la inteligencia artificial que se encarga de reproducir sistemas de manera automática para comprender el contexto de un problema a través de la analítica de datos. Es decir, machine learning crear un algoritmo que revisa datos y es capaz de predecir comportamientos a futuro para la toma de decisiones sin intervención de una persona razón por la cual machine learning se encuentra dividido en tres grandes categorías:

- **Aprendizaje Supervisado.**

Este tipo de aprendizaje resuelve problemas conocidos a partir de un conjunto de datos etiquetados en donde ya se ha identificado la variable a analizar. En base a la variable el aprendizaje supervisado crea un modelo mediante un conjunto de datos de entrenamiento para luego al pasar a un conjunto de datos donde se pueda predecir. Por ejemplo, en la figura 1.2, crea modelos que predice: "¿Cuál es el color de la fruta?", "¿Cuántas frutas diferentes hay en la data?"

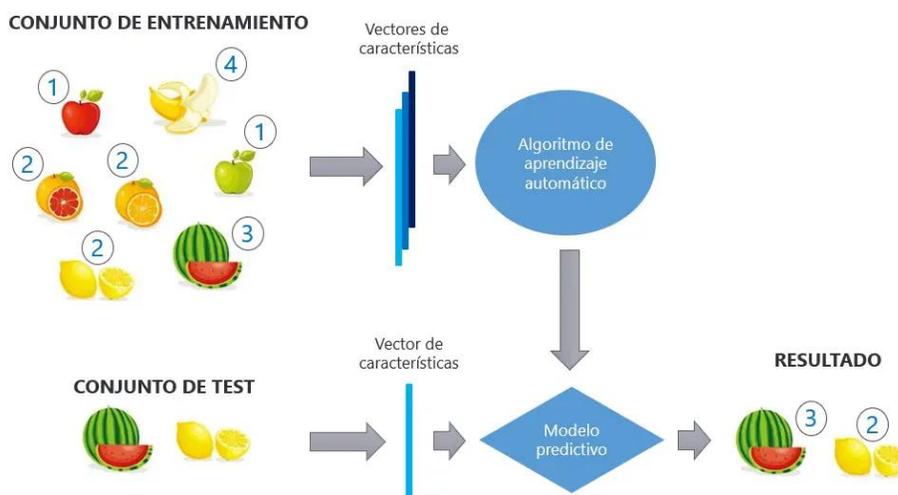


Figura 1.2 Conjunto de entrenamiento en aprendizaje supervisado. Fuente: Diego Calvo [18].

Por lo general, se suele utilizar este tipo de aprendizaje en problemas regresión o clasificación. En el primer caso, se utiliza para identificar un valor único como en el ejemplo de las frutas de la figura anterior. En cambio, la clasificación se aplica a un

grupo o clase en específico. Por ejemplo, en la figura 1.3, identificar las frutas del ejemplo anterior de acuerdo con su color o tipo.

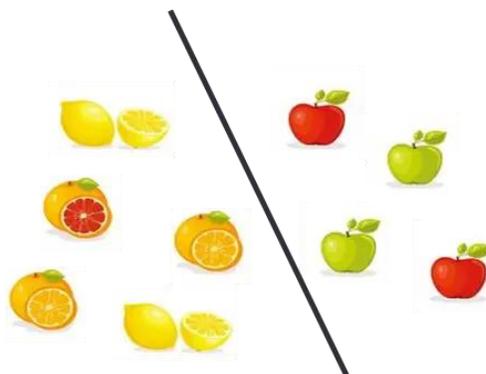


Figura 1.3 Ejemplo de uso de un modelo de clasificación utilizando aprendizaje supervisado. Fuente: Diego Calvo [18].

- **Aprendizaje no supervisado.**

Los métodos utilizados por el aprendizaje no supervisado hacen referencia a un proceso de entrenamiento en donde no se define las variables a analizar por medio de etiquetas o clases. Es decir, se dedica analizar un grupo de datos de entrenamiento sin ningún objetivo a priori para luego obtener similitudes (vector de características) y agruparlos en base a dicha característica. Se utiliza métodos como agrupamiento jerárquico y K-means.

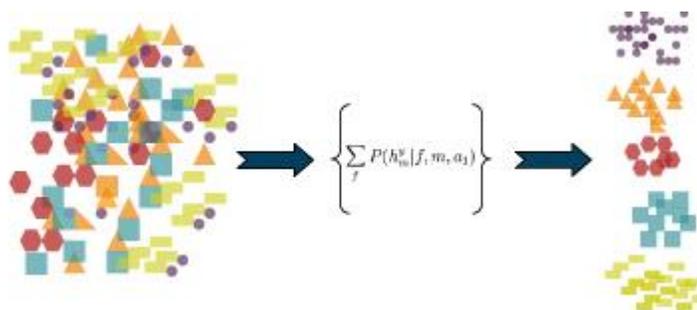


Figura 1.4 Ejemplo de aprendizaje no supervisado. Fuente: Nilton Guerra [19].

- **Aprendizaje profundo.**

Aprendizaje profundo (del inglés Deep learning) realiza procesos más complejos pues utiliza redes neuronales artificial que se asimilan con las neuronas del ser humano,

pero estas se componen en base a un nivel jerárquico. Este nivel consiste en realizar un análisis simple y luego en el siguiente nivel se va mejorando la predicción del modelo hasta reducir la función de error y sea lo más preciso posible evitando el overfitting [17]. Por ejemplo, en la figura 1.5: al realizar el análisis de una foto se quiere identificar si es un perro, se lo va realizando pixel por pixel hasta llegar a una capa final que arroje un resultado.

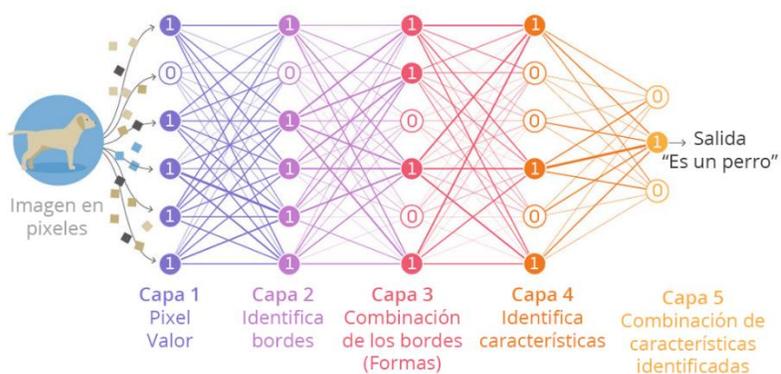


Figura 1.5 Red de redes neuronales en Deep learning. Fuente: Nilton Guerra [19].

### 1.5.4 Web Scraping.

Web scraping o raspado web en español, es una técnica diseñada para extraer datos de páginas web mediante software y convertirlos en datos estructurados que sirva para su análisis. Es decir, convertir de la estructura HTML a tablas para ser almacenados en un base de datos. Mediante es técnica, se puede analizar datos de algún tipo de interés que no sean de carácter libre sin incursionar en algún tema legal [20].

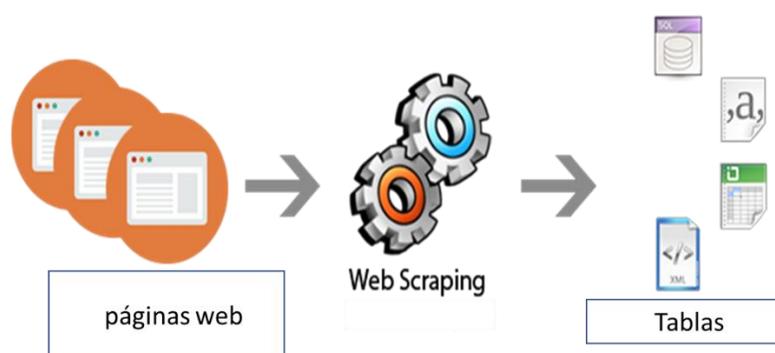


Figura 1.6 Esquema de Web scraping. Fuente: Autoría propia.

### **1.5.5 Sistemas de Información Geográfica (SIG).**

Un sistema de información geográfica (SIG), es un software de trabajo específico que permite a los usuarios realizar consultas interactivas, integradoras capaces de analizar y representar de una forma eficiente cualquier tipo de información geográfica referenciada y asociada a un territorio, que permita conectar mapas con bases de datos. Es decir, se puede visualizar un fenómeno de estudio a través de mapas para identificar zonas vulnerables [21] .

### **1.5.6 Ciudad Intermedia.**

En Ecuador son ciudades con poblaciones aproximadas a 1 millón de habitantes o emplean un alto desarrollo en funciones comerciales e industriales [22]. Dentro de esta clasificación se encuentran 11 ciudades según la secretaria nacional de Planificación y Desarrollo. Estas ciudades son: Cuenca, Manta, Santo Domingo, Loja, Ambato, Esmeraldas, Quevedo, Riobamba, Ibarra, Latacunga y Tulcán.

Dentro de la conferencia Hábitat III realizada en Quito en el año 2016, se planteó una serie de preguntas sobre el desarrollo urbano que tendrían las ciudades intermedias como acceso a la educación de calidad, restauración de estas, pero con un sentido de vivir digno para todos. Razón por la cual es necesario conocer si eso se está cumpliendo o copiando el modelo neoliberalista de las grandes ciudades para ello se plantea dentro de este trabajo de titulación evaluar el nivel de gentrificación de dichas ciudades. Para ello, se debe seleccionar un modelo adecuado de gentrificación y aplicarlo a dichas ciudades.

## **1.6 Estado del Arte.**

### **1.6.1 Gentrificación en Ecuador.**

De acuerdo a la literatura revisada, se han encontrado artículos de investigación y documentos que registran el proceso de gentrificación en Ecuador evaluando principalmente a las metrópolis que son Quito, Guayaquil y Cuenca fueran de esas existe documentación para la ciudad de salinas ubicada en la provincia de Santa

Elena. Para el resto de las ciudades no se han desarrollado investigaciones parecidas, solo se hace una simple mención sobre que debería estar pasando.

El artículo presentado por Gustavo Duran [23] de la universidad FLACSO, Ecuador detalla el proceso de gentrificación para las parroquias de Cumbayá y Calderón que en los años 70 eran zonas agrícolas y pasaron a formar un eje policéntrico que generaba conjuntos habitacionales cerrados para la clase media rompiendo el eje centro-norte replicando el modelo de suburbios estadounidense. También, se detalla que debido al clima en la parroquia de Cumbayá y la obra de vía como la general Rumiñahui y la ubicación de centros educativos de elite incrementaron el valor del metro cuadrado de suelo de construcción es así como ese valor se ha disparado desde el año 2005. En la tabla 1 de dicho artículo se puede visualizar el incremento de metro cuadrado hasta el año 2011. Dentro de este estudio se detalla la gentrificación aplicada a las zonas periurbanas que para el caso de esas dos parroquias más de 50% de suelo se encontraba vacante. Esto evidencia claramente la falta de planificación y control del GAD municipal que permitió la masificación de construcción de conjuntos habitacionales sujetas a una gran especulación por parte de inmobiliarias que aprovecharon el proceso de dolarización para aumentar sus beneficios

**Tabla 1.1 Precio de metro cuadrado en la parroquia de Cumbayá. Fuente: Revista Bienes Raíces 2011.**

Año	Precio del metro cuadrado de la tierra (dólares)
1995	13
1992	30
1993	50
1998	70
2000	110
2011	160
2011 (vía Pillagua)	+ de 200

La autora Adela Vargas Murillo [24] evidencia cambios y regeneración en la ciudad de Guayaquil que han sido a través de un proceso de gentrificación debido a la

segregación de lugares en donde se sostiene la idea de inseguridad, delincuencia, prostitución para luego a través de medios y alcaldía se priorice la remodelación del sector. De hecho, muchos lugares, se han gentrificado en esta ciudad como la avenida 9 de octubre y la calleas aledañas que son actualmente lugares con locales comerciales junto con esculturas relucientes. En esta ciudad es donde más fácil resulta gentrificar debido a la inseguridad que se promociona por radio y televisión.

La Universidad Estatal Península de Santa Elena en el año 2020 publicó un artículo "*El proceso de Gentrificación, intervención urbana arquitectónica en la ciudad de Salinas-Ecuador*" [10]. Dentro del mismo se detalla cómo se ha dado el proceso de turistificación en la parte costera del malecón y sus alrededores promovidos en la década de los 90 por el sector privada que cambio la imagen a una zona de atractivo turístico que luego es apoyada por el gobierno local, ya que la ciudad se convierte en un eje turístico y depende de esa actividad. En la actualidad, se visualiza más del 85 % de sobrepoblación de edificios que son ocupados en la mayoría de los casos fines de semana o feriados por turistas. Debido a esto los habitantes de salinas han tenido que desplazarse hacia las afuera de la ciudad porque los precios del suelo, vivienda y locales comerciales se han duplicado en una década. Esta población, invaden ilegalmente sectores aledaños para su habitabilidad, creando condiciones de vida deplorado que es exhaustivamente notorio y obliga al gobierno a invertir en proyectos para mejorar esas condiciones.

Los autores como Steel y Klaufus [25], mostraron evidencias empíricas de gentrificación en la ciudad de Cuenca debido a la llegada de población jubilada norteamericana. Este tipo de movilidad y migración entre los países americanos hacia destinos latinoamericanos es debido al bajo costo que tiene la canasta básica en comparación de dichas ciudades o también se debe a motivos por cuestiones de ocio y estilo de vida. Han generado proyectos inmobiliarios en sectores de la avenida Ordoñez Lasso, avenida primero de mayo y otros sectores cerca del hotel oro verde.

Dentro de lo analizado en la mayoría de los artículos se puede evidenciar claramente que el proceso de gentrificación en Ecuador obedece a tres variables que son:

1. Regeneración urbana.
2. Aumento de valor de suelo a través de la plusvalía.
3. Boom cultural a través de la creación de sitios turísticos o lugares regenerados como: Bares, discotecas, locales de comida, centros comerciales, etc.

Por tanto, la gentrificación en Ecuador se evidencia que se da por falta del control inmobiliario y en algunos casos por políticas neoliberales en las ciudades que son metrópolis, pero se desconoce que sucede en ciudades intermedias que como los cita José María Llop en su libro titulado “*EL DERECHO A LA CIUDAD EN EL CONTEXTO DE LA AGENDA URBANA PARA CIUDADES INTERMEDIAS EN EL ECUADOR*” [22]. Estas ciudades debido al proceso de globalización en la actualidad toman un rol importante porque realizan exportaciones a nivel internacional, tienen industrias manufactureras, puertos de desembarque, tienen zonas privilegias para la producción de algún tipo de alimento y esto genera que se mueva la economía del país. Es decir, dentro de estas ciudades se está generando empleo y a su vez se evita la migración a ciudades grandes [26]. Para el año 2030, se estima que el 50% de la población urbana vivirá en ciudades intermedias (I-Cities o Intermediary Cities). En 2050, ese porcentaje aumentará hasta el 70%. Razón por la cual estas ciudades deben de planificarse y estructurar un desarrollo urbano coherente evitando que sucedan los mismo que las metrópolis, que en su mayoría de barrios ya se encuentran gentrificados. Para ello se plantea evaluar el estado de gentrificación que presenta las ciudades intermedias actualmente.

### **1.6.2 Modelos de Gentrificación.**

Dentro de la literatura sobre modelos que evalúan el grado de gentrificación en la mayoría de ellos se utiliza variables que miden la transformación sociodemográfica de la ciudad que incorporan diferentes grados de análisis que suelen ser extraídos de datos censales, tales como: nivel de estudios del individuo, estado de viviendas, número de habitaciones, área de la vivienda y etnia a la que pertenece una familia entre otras. Como los censos se realizan cada cierto tiempo algunos autores de

modelo han incluido otras variables para que se pueda ver la evolución de la gentrificación en un tiempo mucho más corto.

#### **1.6.2.1 Modelo De Hammel Y Wyly.**

Como es sabido en la década de los 80 y 90 se produjeron grandes crecimientos urbanísticos los autores de este modelo se basaron en datos de los censos debido a que en el año 1993 no existía medios de difusión virtual como los es hoy en día el internet [13]. Dentro de su modelo analizaron la edad en comparación de grado académico para saber si en dicho distrito predominaban personas con nivel académico de tercer nivel. Esto indicaría que esas personas tienen una situación laboral estable, cosa que se comprueba con otra pregunta del censo a través del acceso a seguridad privada. También considera el índice de materialidad de la casa pues una casa con acabados de lujo debe estar presente en una zona gentrificable. Al ser el primer modelo de estudio de gentrificación ha sido aplicado en un sin número de estudios.

#### **1.6.2.2 Modelo De Freeman.**

Para el año 2005 este autor proporciona nuevas formas de medir la gentrificación en una ciudad esta vez incluye un PSID (Panel de estudio de dinámica de ingreso, del inglés *Panel Study of Income Dynamics*) este panel es muy común en países norteamericanos y europeos que es el equivalente a la central de riesgos en Ecuador solo que en nuestro país esos datos no son abiertos al público [11]. Además, introdujo el proceso de Densificación que miden la cantidad de personas que viven en un hogar alegando que las familias que no tienen altos ingresos tienen que compartir habitaciones excluyendo a las parejas de casados.

Este modelo también considera la altura máxima de las edificaciones junto con la cercanía a centros de abastecimiento de comida.

#### **1.6.2.3 Modelo De Nesbitt.**

Este modelo fue introducido en el año 2007 hace uso de herramientas de SIG para representar los datos en donde asocia al modelo de Hammel en un conjunto de variables llamadas AGEB (Área geoestadística básica) y las contrasta con el aumento

del valor de m<sup>2</sup> por arriba del promedio de la ciudad. Para su utilización se debe empatar con la data censal del lugar donde se va a evaluar la gentrificación. Se debe utilizar muchas herramientas de extracción de datos para aplicarlo como en el caso de este trabajo de titulación que su uso Web scraping que empata con la geodatabase del INEC. Nesbitt interpreta a las variables AGB de la siguiente manera:

- Si existe un porcentaje de población con educación universitaria por arriba del promedio de la ciudad.
- Si existe un porcentaje en que aumentó el valor unitario del metro cuadrado del suelo por arriba del promedio de la ciudad.
- Si el porcentaje de viviendas habitadas rentadas es inferior o superior a la media de la ciudad.

#### **1.6.2.4 Modelo de Chapple, Waddell, Chatman, Loukaitou-Siders**

Es el modelo más actual de gentrificación datado en el año 2017, pero a su vez es el más complejo pues mide niveles de entropía en base a la construcción, accesibilidad al transporte público y DCN (Distancias a centros de negocios) como centros comerciales. Al igual que el modelo de Nesbitt analiza variables geoestadísticas básicas AGEB como dependientes y como variables independientes selecciona la cercanía a hoteles, restaurantes, bares discotecas, boutiques elegantes. Para aplicar dicho modelo se necesita tener una gran cantidad de datos de cada una de las variables independientes. No fue considerado debido a que utiliza los datos de accesibilidad al transporte público no son libres y también se debería scrapear mucha data lo cual no permitiría analizar a todas las ciudades intermedias.

En la tabla 2 ha manera de resumen se muestra la comparativa de modelos de gentrificación junto con su año de implementación detallando las ventajas y desventajas que tendría el implementar dichos modelos.

Tabla 1.2 Comparativa de los modelos de análisis del fenómeno de gentrificación.

Modelos de identificación y análisis del fenómeno de gentrificación		
Modelo	Ventajas	Desventajas
<b>Hammel y Wyly 1993.</b>	<ul style="list-style-type: none"> <li>•Uso de datos censales que son accesibles a todo público.</li> <li>•Modelo muy difundido y probado en múltiples estudios.</li> <li>•Uso de variables coherentes y de fácil acceso.</li> </ul>	<ul style="list-style-type: none"> <li>•No se trabaja con valores actuales.</li> <li>•Depende un lapso de tiempo entre censos realizados a la misma población civil.</li> </ul>
<b>Freeman 2005.</b>	<ul style="list-style-type: none"> <li>•Uso de datos censales.</li> <li>•Datos geo-codificados.</li> <li>•Modelo basado a nivel de barrios</li> </ul>	<ul style="list-style-type: none"> <li>•Depende mucho de la accesibilidad a los datos financieros y obras de construcción representativas (obras de altura).</li> <li>•No todos los países tienen sus datos Geo-codificados por código postal de forma coherente.</li> </ul>
<b>Nesbitt 2007.</b>	<ul style="list-style-type: none"> <li>•Uso de herramientas SIG.</li> <li>•Aplica datos censales a través de variables AGEb.</li> <li>•Mide el valor de m2 por zona para identificar la gentrificación.</li> </ul>	<ul style="list-style-type: none"> <li>•Comparación de variables censales con sistemas geográficos existentes.</li> <li>•Uso de técnicas más avanzadas de SIG como machine learning.</li> </ul>
<b>Chapple, Waddell, Chatman, Loukaitou-Siders 2017</b>	<ul style="list-style-type: none"> <li>•Uso de SIG por sectores.</li> <li>•Análisis multivariantes por cada sector mejorando la precisión.</li> <li>•Granularidad en la data si se da por código postal.</li> </ul>	<ul style="list-style-type: none"> <li>•Uso de indicadores complejos.</li> <li>•Muchas variables para analizar (Cajeros ATM, centros comerciales, estrellas por parte de clientes).</li> <li>•Datos de transporte públicos no muy claros.</li> </ul>

Para representar y poder visualizar cualquiera de estos modelos sobre la gentrificación de manera gráfica es necesario utilizar una herramienta que utilice SIG que pueda aplicarse a sistemas informáticos y permita realizar la gestión de datos espaciales en base a las variables AGEB.

### 1.6.3 Herramientas SIG.

Las herramientas SIG permiten procesar geoinformación para la gestión y desarrollo de planes territoriales, hidrográficos, riesgos naturales, análisis de suelo entre muchas cosas más que se pueda evidenciar a través de mapa. Pero antes de su representación a nivel de mapa sigue una serie de subsistemas, lo cuales suelen ser tres:

- **Subsistema de datos:** Es la parte encargada de la entrada y salida de información a través de conexiones remotas a bases de datos o carga de archivos en formato CSV que contengan coordenadas e información de la variable a analizar [21].
- **Subsistema de visualización:** Es el encargado de la creación cartográfica a partir de representaciones de datos en mapas, leyendas, áreas, etc. Permite la inserción de datos dentro del mapa en otras palabras funciona como un editor para la visualización de coordenadas o áreas.
- **Subsistema de análisis:** Realiza la interpretación a través de métodos numéricos aplicados al subsistema de datos que luego va a ser representado.

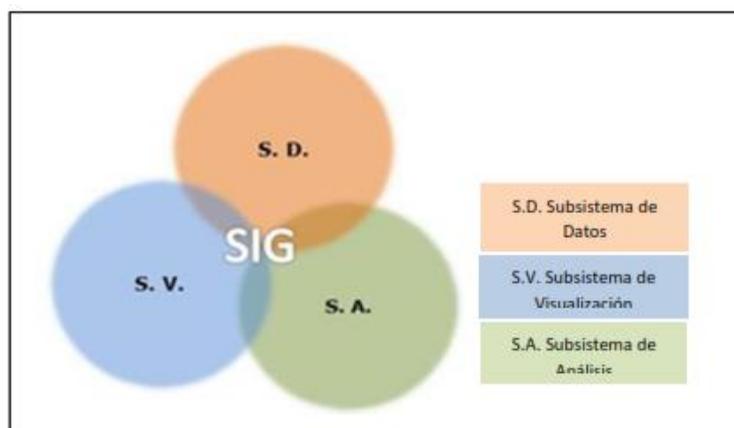


Figura 1.7 Esquema de un SIG con sus tres subsistemas fundamentales. Fuente: Autoría Propia.

### 1.6.3.1 Elementos de un Sistema SIG.

Estos subsistemas están repartidos en base a su procesamiento y obedecen a cinco elementos principales que contemplan toda la metodología del uso de SIG. Cada uno de ellos cumple su respectiva función:

- **Datos:** Son como la materia prima para una herramienta SIG, pues contiene información y ubicación del fenómeno analizar.
- **Métodos:** Son herramientas que facilitan al usuario el análisis de los datos, esto puede ser desde un formulario hasta metodología más complejas.
- **Software:** Es la parte lógica que maneja el usuario a través de la aplicación para implementar los métodos anteriores.
- **Hardware:** Es el equipo físico que se necesita para ejecutar el software cada vez se requiere mejor procesador, memoria RAM y el uso de tarjeta gráfica dependiendo del análisis a realizar.
- **Recursos humanos:** Es el motor de sistema SIG porque se encarga de modelar y ejecutar el software para ver representaciones del fenómeno a analizar.

Si bien hasta ahora solo se hablado de la manera en que interactúan los sistemas SIG entre software y persona existen formatos para representar y analizar su información a través de modelos. Estos modelos son el formato raster y el formato vectorial que dan lugar a la formación de capas que luego pueden ser sobrepuestas para mejora el análisis de un fenómeno.

- **Formato Raster.**

Este formato divide al espacio analizar en un conjunto regular de celdillas, donde cada uno de ellos representa un objeto a manera de variable para representar un único valor. Hace uso de pixeles de imágenes como png, jpeg, tiff, ortofotos entre otros.

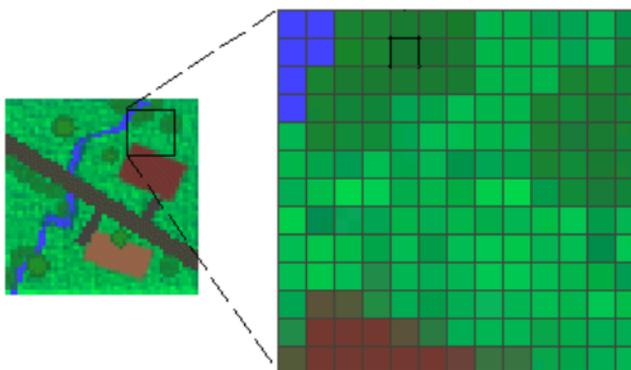


Figura 1.8 Formato raster en sistemas SIG. Fuente: Autoría propia.

- **Formato vectorial.**

Es el formato más común que se utiliza en sistema SIG debido a que representa las variables a analizar en formas de puntos, líneas o polígonos. Sin embargo, al momento de crear los elementos o cargarlos se debe comprobar que estas formas contengan información que será procesada en base a métodos numéricos.

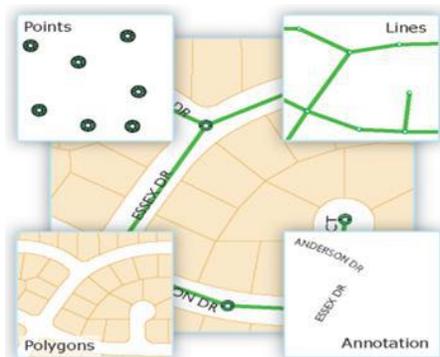


Figura 1.9 Representación vectorial de puntos, líneas y polígonos. Fuente: Autoría propia.

Entre estos dos formatos existen una gran diferencia a la hora de procesar lo datos y realizar un comparativa del fenómeno a analizar. En la tabla 3 se resumen las principales diferencias entre ambos formatos.

Tabla 1.3 Diferencias entre Formato Raster y Vectorial.

VENTAJAS RASTER	VENTAJAS VECTORIAL
Estructura de datos muy simple	Elevación precisión
Facilidad para la representación de entidades espaciales continuas	Facilidad de integración con CAD vectoriales
Elevada capacidad para superposición y combinación de capas	Mapas de elevada calidad
Capacidad para integrar datos de satélite	Estructura de datos con topología
Capacidad de incorporación de imágenes	Ficheros de poco tamaño

### 1.6.3.2 Análisis De Datos En Sistemas SIG.

Los procesos de análisis geográficos frecuentemente llamados de análisis espaciales o geoprosesamiento suelen buscar patrones, tendencias, escenarios similares o características que lo lleven a una conclusión coherente en base a procesos con analítica de datos. Para ello se basan en dos formas de análisis:

- **Análisis de proximidad:** En toda herramienta SIG en la actualidad cuenta con motores de búsqueda que miden la distancia entre un objeto y otro. Suele ser utilizada para contestar preguntas, tales como.
  - I. ¿Cuántas casas se encuentran dentro de una manzana?
  - II. ¿Cuántas clientes existen dentro de un radio de 1km?
  - III. ¿El promedio casas con piscina en un sector?

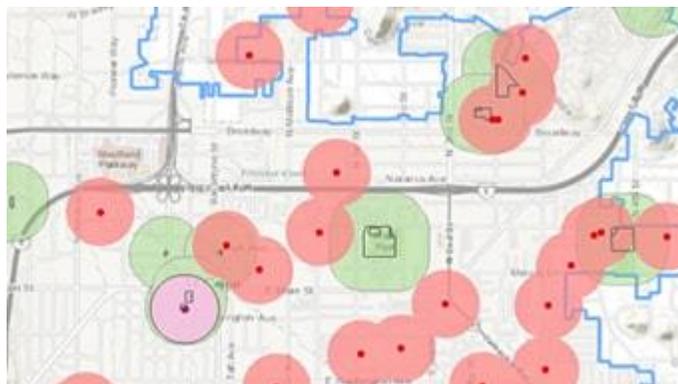


Figura 1.10 Análisis de proximidad. Fuente: Esri [21].

- **Análisis de superposición:** Para realizar este análisis se tiene que ir segregando por capas para luego realizar una comparativa y ver cuan diferente es la una de la otra utilizando extrapolación. Se suele utilizar para medir contaminación del suelo, pendientes, zonas con vegetación o algún cambio por catástrofes naturales.

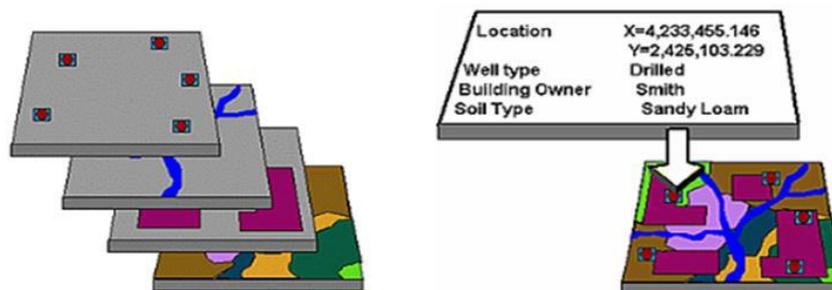


Figura 1.11 Análisis de superposición. Fuente: Esri [21].

#### 1.6.4 Herramientas de SIG.

Durante los últimos años se vienen realizando varios estudios, tesis y publicaciones que comparan varios clientes SIG de escritorio. Entre los que más destacan al momento son: ArcGIS, QGIS y gvSIG.

Los tres tienen sus ventajas y sus inconvenientes, en si depende de la finalidad en que se los requiera. ArcGIS Pro destaca por su facilidad de uso y en su rica documentación generada por la comunidad de Esri. Además, siempre ha sido la herramienta perfecta para realizar estudios topográficos o trabajar con redes, pero muchas de las tareas

que se pueden hacer en ARGIS también se las puede realizar tanto con QGIS como con gvSIG. En cuanto a programación, los tres softwares permiten automatizar tareas de análisis mediante scripting y notebooks. Como dato curioso y a la vez reconfortante, los tres han elegido el lenguaje de Python para sus desarrollos debido a la afinidad que presenta con la analítica de datos. En la tabla 4 se puede evidenciar las principales ventajas y desventajas que presentan estos tres softwares.

Tabla 1.4 Comparativa de aplicaciones de escritorio para uso de SIG.

<b>Comparativa entre diferentes softwares SIG</b>		
<b>Programa</b>	<b>Ventajas</b>	<b>Desventajas</b>
<b>ArcGIS</b>	<ul style="list-style-type: none"> <li>• Interfaz muy amigable.</li> <li>• Acceso a geoprocursos sin necesidad de plugins.</li> <li>• Composición y generación de mapas.</li> <li>• Simbología y etiquetado.</li> <li>• Capacidad raster.</li> <li>• Scripting con ArcPy muy sencillo.</li> </ul>	<ul style="list-style-type: none"> <li>• Licencias.</li> <li>• No permite el despliegue de proyectos entre diferentes versiones.</li> </ul>
<b>QGIS</b>	<ul style="list-style-type: none"> <li>• Software libre y gratuito con actualizaciones cada cuatro meses.</li> <li>• Interoperabilidad.</li> <li>• Conexión a base de datos, en particular PostgreSQL y PostGIS.</li> </ul>	<ul style="list-style-type: none"> <li>• Se requiere un alto índice de conocimiento de Python porque no contiene tantas herramientas amigables.</li> <li>• Todavía sigue en desarrollo.</li> </ul>
<b>gvSIG</b>	<ul style="list-style-type: none"> <li>• NavTable.</li> <li>• Rendimiento con grandes volúmenes de información.</li> <li>• Integración de algoritmos de SEXTANTE.</li> </ul>	<ul style="list-style-type: none"> <li>• No existe documentación.</li> <li>• Es un sistema reciente.</li> </ul>

De estas aplicaciones previamente detalladas se escogió a ArcGIS debido a su documentación realizada por la comunidad, calidad de gráficos. Además, se contaba con una licencia proporcionada por el CEU (Centro de Estudios Universitarios).

## CAPÍTULO II

### 2. Implementación De Modelo.

La parte literaria ya fue analizada en el capítulo anterior en donde se ha seleccionado la metodología de Nesbitt para medir el impacto de la gentrificación en Ciudades Intermedias de Ecuador. Así como también se seleccionó el software ArcGIS como herramienta de trabajo en la parte de SIG. Ahora lo primero que se debe hacer es la extracción de datos censales del último censo año 2010 para cargar en mapa y contrastar con la data de lugares bienes inmuebles de las páginas web como Airbnb y trivago a través del Web scraping. Para ello se detalla el proceso realizado para presentación del fenómeno en las siguientes 6 fases como se muestra en la figura 2.1.

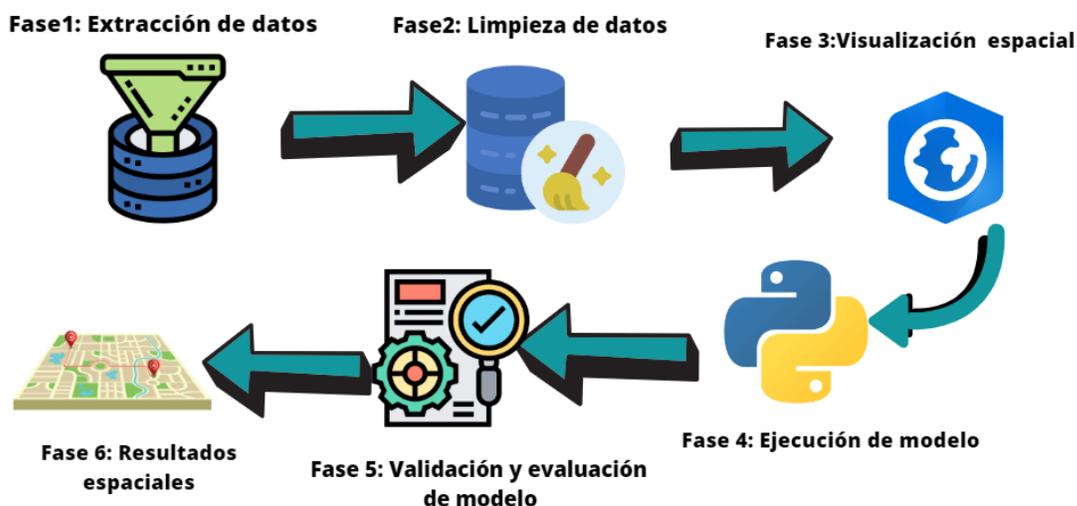


Figura 2.1 Fases de procesos a implementar para medir la gentrificación en una ciudad intermedia.

Fuente: Autoría propia.

#### 2.1 Fase 1: Extracción y Obtención de Datos.

Para acceder a los datos censales que son abiertos es necesario descargar un software que permita leer dichos datos porque como datan del año de 1990 tiene formatos variados mezclados con C y C++ para lo cual se recurre la ayuda del software

Redatam <sup>1</sup>. La data del último censo poblacional 2010 Ecuador se la puede encontrar en la página del INEC<sup>2</sup>.

Donde se debe seleccionar toda la documentación a nivel nacional tal cual como lo muestra la siguiente imagen.



Figura 2.2 Descarga de data empataada nacional censo 2010. Fuente: Autoría propia.

Una vez descargada la data se procede a instalar el software Redatam que no conlleva mayor dificultad como se muestra en siguiente figura. Luego la data del censo se carga en la parte de diccionario.

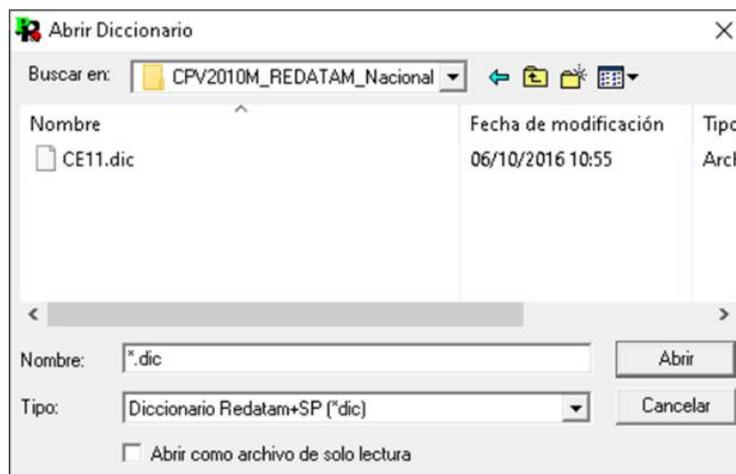


Figura 2.3 Software Redatam Instalado. Fuente: Autoría propia

<sup>1</sup> es un sistema computacional de carácter social amigable e interactivo que facilita el procesamiento, análisis y diseminación web de la información de censos

<sup>2</sup> <https://www.ecuadorencifras.gob.ec/base-de-datos-censo-de-poblacion-y-vivienda-2010-a-nivel-de-manzana/>

Una vez realizado esto se puede visualizar los datos de una manera simple, que contiene entidades jerárquicas dadas por: PROVINCIA, CANTÓN, PARROQUIA, ZONA, SECTOR, MANZANA Y VIVIENDA.

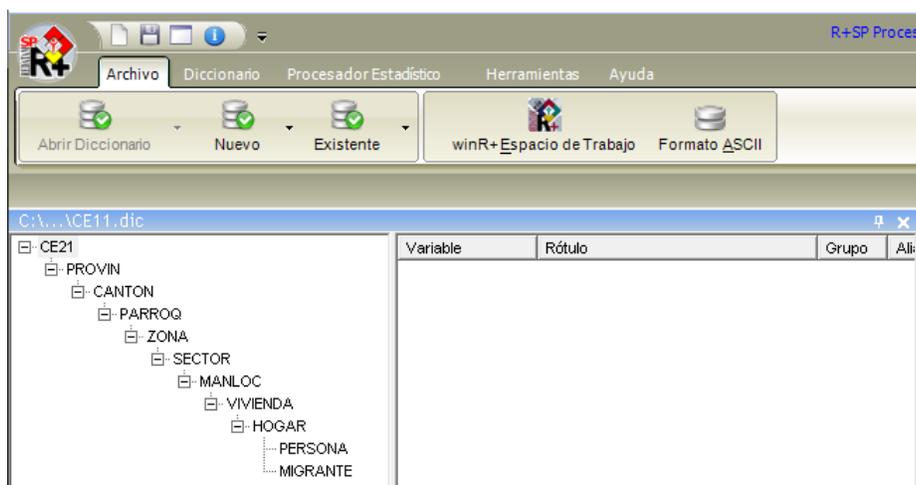


Figura 2.4 Visualización jerárquica de censo en Redatam. Fuente: Autoría propia.

Una vez visualizado el campo jerárquico de la data, se analizan las variables de interés para ser filtradas a través de un archivo Excel. Pero primero se de seleccionar en Nuevo> selección y escoger el cantón analizar en este ejemplo se emplea a Manta que es una de las ciudades Intermedias.

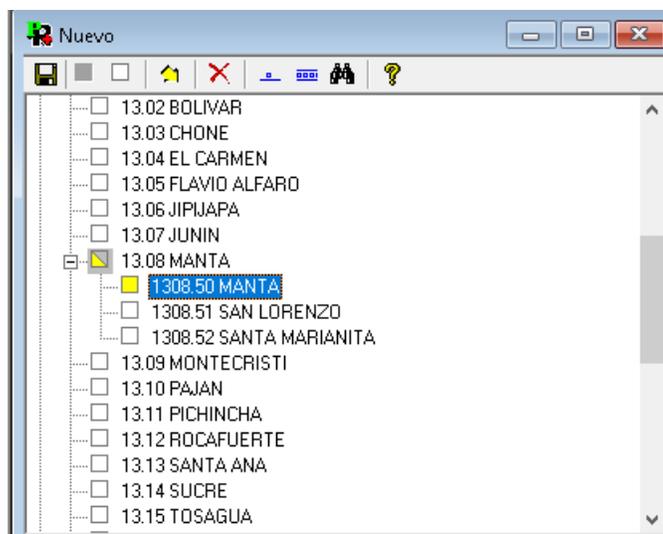


Figura 2.5 Selección de cantón analizar. Fuente: Autoría propia.

Una vez creada la selección se procede a extraer las diferentes variables de Interés en este caso responden a la categoría ocupacional del individuo, que son 10 en total.

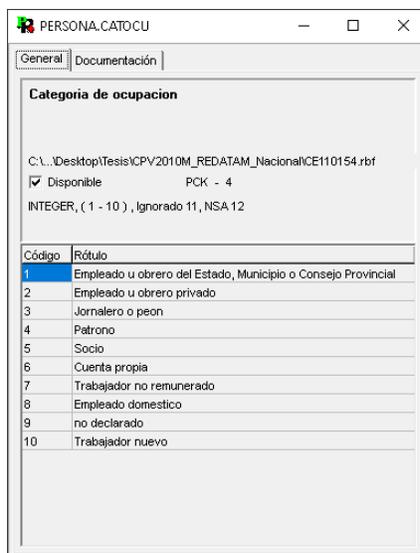


Figura 2.6 Categoría ocupacional. Fuente: Autoría propia.

Finalmente, se debe vincular la variable con la selección de cantón para poder generar la tabla que indica la variable por sector. Para luego generar un proceso estadístico en la plataforma de Redatam y tener la tabla con valores por zonas.

Código	Empleado u obrero del Estado, Municipio o Consejo Provincial	Empleado u obrero privado	Jornalero o peon	Patrono	Socio	Cuent
130850001	153	593	19	122	23	
130850002	265	554	16	120	43	
130850003	343	486	23	64	22	
130850004	246	836	33	134	52	
130850005	267	794	20	70	17	
130850006	264	618	26	74	20	
130850007	178	617	12	69	19	
130850008	231	803	25	91	27	
130850009	286	811	71	66	31	
130850010	173	489	14	35	20	
130850011	158	507	43	31	13	
130850012	159	481	49	30	18	
130850013	133	542	67	21	12	
130850014	160	715	106	31	15	
130850015	181	816	91	27	9	
130850016	269	997	53	49	24	
130850017	203	935	26	83	27	
130850018	172	708	26	76	19	
130850019	256	988	75	68	23	
130850020	127	577	57	21	13	
130850021	128	555	58	13	11	
130850022	536	682	64	18	13	
130850023	203	755	30	48	25	
130850024	415	673	13	53	20	
130850025	190	813	32	82	23	
130850026	142	714	66	73	23	
130850027	169	772	34	91	19	
130850028	227	1.078	40	58	38	

Figura 2.7 Tabla generada en software Redatam para análisis de estado económicos. Fuente: Autoría propia

Con la extracción de este tipo de información se tiene el código de zona para el cantón correspondiente que luego es vinculado con la data obtenida de los sitios de ventas de casas.

### 2.1.1 Extracción de datos mediante Web scraping.

La mayoría de los datos que se encuentra en Internet no son de manera abierta y para un sin número de investigaciones se debe extraer estos por medio de la técnica de Web scraping. Para realizar esta tarea se debe seleccionar un lenguaje y sus bibliotecas en este caso fue Python con dos de sus principales librerías como son:

- **Beautiful Soup:** Permite conocer el formato HTML dentro de la página y así seleccionar el bloque de interés a extraer.
- **Selenium:** Brinda el proceso de paginación ahorrando mucho tiempo, pues en varias páginas se obtienen de 10-15 paginaciones.

Luego, es necesario dirigirse a la página donde se obtendrá la información para explorar su contenido es necesario presionar F12 en Google Chrome y luego explorar los elementos de acuerdo con el elemento div que le corresponda que será el mismo en toda la página.

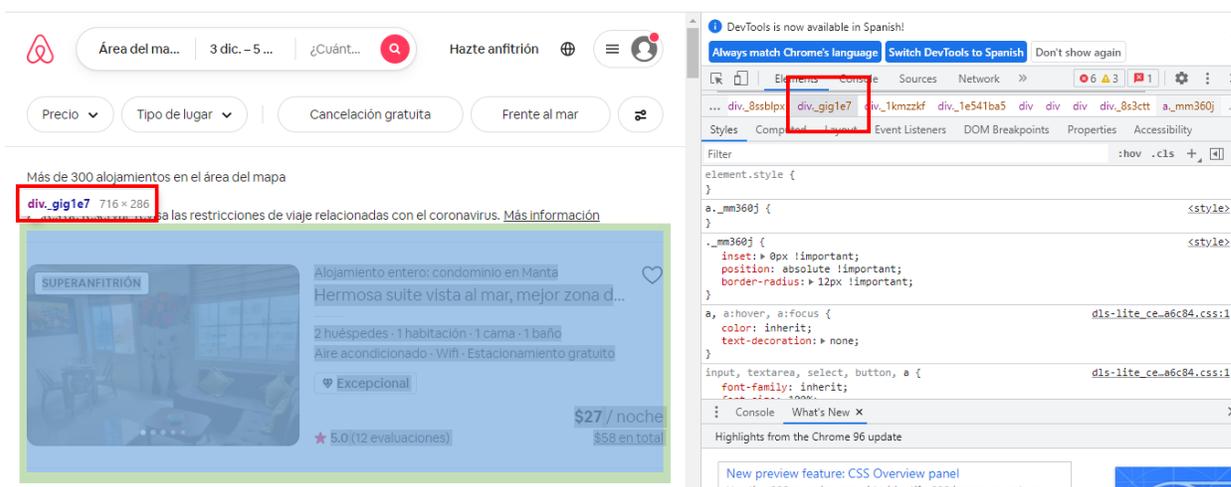


Figura 2.8 Exploración de estructura HTML de la página. Fuente: Autoría propia.

Con la ayuda de la librería Beautiful Soup ingresamos el campo div que se ha visualizado y que es necesario extraer. A continuación, se detalla cómo debería ir el llamado de las librerías junto con la URL de la página en Python.

```
import request
from bs4 import BeautifulSoup
URL = "url de pagina"
pagina = solicitudes . obtener (URL)
sopa = BeautifulSoup(page . content, "html.parser" )
list = page_soup . findAll( "div" , { "clase" : "nombre de clase" })
```

**Código 2.1 Importación de campo nombre a través de Web scraping.**

Para el uso de más variable y obtención de más datos es necesario crear un bucle controlado dentro de Python.

```
for n_element in job_elements:
    direccion_element = job_element.find("div1", class_="")
    precio_element = job_element.find("div2", class_="precio")
    localizacion_element = job_element.find("div3", class_="localizacion")
    print(title_element)
    print(company_element)
    print(location_element)
    print()
```

**Código 2.2 Extracción múltiple variable.**

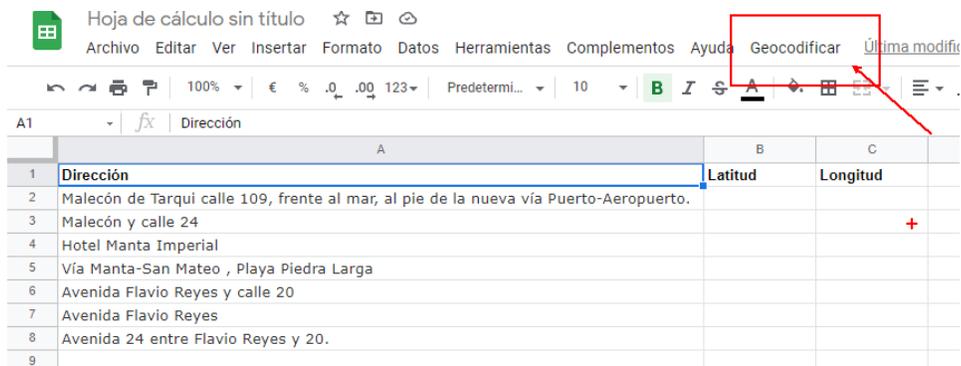
La data deseada será extraída en formato CSV que debe contener las siguientes variables en el caso ideal sería:

```
'Nombre',
'Área construida m2',
'Precio', 'Estrato',
'Antigüedad', 'Sector',
'Numero de Piso.',
' Número de Habitaciones',
' Número de Baños'
```

**Código 2.3 Nombre de campos.**

Como no en todas las páginas funciona la extracción de la latitud y longitud, es necesario ejecutar un script en un sheet de Google Driver para encontrar las

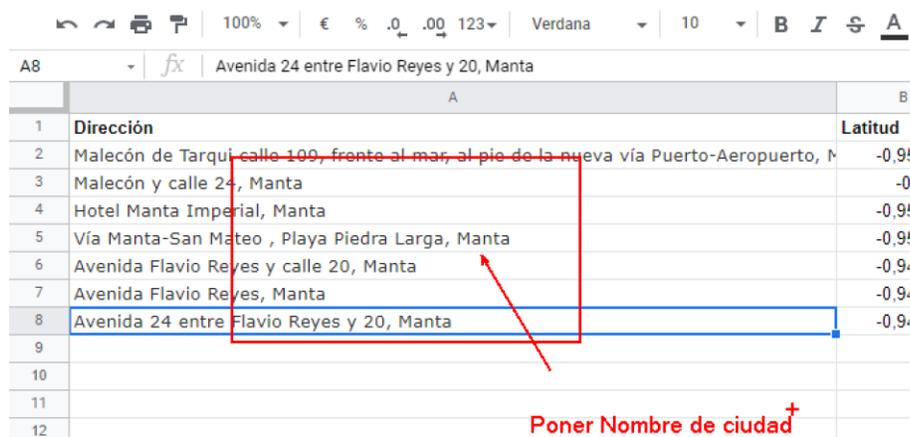
coordenadas espaciales. Para ellos se debe copiar el script dentro de la opción de Geocodificar<sup>3</sup>.



	A	B	C
1	Dirección	Latitud	Longitud
2	Malecón de Tarqui calle 109, frente al mar, al pie de la nueva vía Puerto-Aeropuerto.		
3	Malecón y calle 24		
4	Hotel Manta Imperial		
5	Vía Manta-San Mateo , Playa Piedra Larga		
6	Avenida Flavio Reyes y calle 20		
7	Avenida Flavio Reyes		
8	Avenida 24 entre Flavio Reyes y 20.		
9			

**Figura 2.9 Geocodificar direcciones en latitud y longitud. Fuente: Autoría propia.**

Cabe recalcar que para mejora la exactitud es necesario agregar en la celda de dirección el nombre de la ciudad.



	A	B
1	Dirección	Latitud
2	Malecón de Tarqui calle 109, frente al mar, al pie de la nueva vía Puerto-Aeropuerto, Manta	-0,94
3	Malecón y calle 24, Manta	-0
4	Hotel Manta Imperial, Manta	-0,94
5	Vía Manta-San Mateo , Playa Piedra Larga, Manta	-0,94
6	Avenida Flavio Reyes y calle 20, Manta	-0,94
7	Avenida Flavio Reyes, Manta	-0,94
8	Avenida 24 entre Flavio Reyes y 20, Manta	-0,94
9		
10		
11		
12		

Poner Nombre de ciudad<sup>+</sup>

**Figura 2.10 Concatenación de nombre de ciudad en columna de dirección. Fuente: Autoría propia.**

Una vez obtenida toda la data y ejecutado el proceso de geocodificación se obtiene un archivo csv que no está concluido del todo porque es necesario realizar la limpieza de datos como se puede visualizar en la imagen siguiente.

<sup>3</sup><https://tesel.mx/convertir-direcciones-a-coordenadas-de-latitud-y-longitud-con-google-sheets-y-google-maps-geocoding-7928/>

	A	B	C	D	E	F	G
1	Precio	Descripcion	Direccion	Lugar	Latitud	Longitud	
2	\$ 65	2 Dor. - 2 Bñ. - 85 m2	Departamento vacacional / renta tempo	24 de Mayo, Manabí	-0,9676533	-80,7089101	
3	\$ 1,650	2 Dor. - 2 Bñ. - 114 m2	departamento de alquiler en manta zona	Manta, Manabí	-0,9476708	-80,7229347	
4	\$ 1,450	2 Dor. - 2 Bñ. - 114 m2	departamento de alquiler en manta zona	Manta, Manabí	-0,9476708	-80,7229347	
5	\$ 150	1 Dor. - 1 Bñ. - 25 m2	Se alquila departamento en Manta secto	Manta, Manabí	-0,9447141	-80,739241	
6	\$ 750	2 Dor. - 2 Bñ. - 98 m2	ALQUILO DEPARTAMENTO EDF VISTA AL	Manta, Manabí	34,0036263	-118,4227185	
7	\$ 200	Estudio Dor. - 0 Bñ. - 36 m2	Alquilo habitacion en Manta amoblada c	24 de Mayo, Manabí	-0,9447141	-80,739241	
8	\$ 270	1 Dor. - 1 Bñ. - 80 m2	Alquilo departamento en Manta amoblada	24 de Mayo, Manabí	-0,9870138	-80,7211125	
9	\$ 850	2 Dor. - 2 Bñ. - 170 m2	Alquilo departamento en Manta semiam	24 de Mayo, Manabí	-0,9676533	-80,7089101	
10	\$ 320	1 Dor. - 1 Bñ. - 40 m2	Alquilo departamento en Manta amoblada	Manta, Manabí	-0,9536786	-80,7644591	
11	\$ 1,300	3 Dor. - 3 Bñ. - 180 m2	Alquilo departamento en Manta amoblada	24 de Mayo, Manabí	-0,9476708	-80,7229347	
12	\$ 350	3 Dor. - 3 Bñ. - 107 m2	Alquilo casa en Manta en urbanización	Manta, Manabí	-0,9870138	-80,7211125	
13	\$ 350	3 Dor. - 3 Bñ. - 109 m2	Alquilo casa en Manta en urbanización	Manta, Manabí	-0,9870138	-80,7211125	
14	\$ 350	3 Dor. - 3 Bñ. - 100 m2	Alquilo casa en Manta amoblada en urba	Manta, Manabí	-0,9870138	-80,7211125	
15	\$ 280	2 Dor. - 1 Bñ. - 100 m2	Alquilo casa en Manta en urbanización	Manta, Manabí	-0,963088	-80,753395	
16	\$ 550	3 Dor. - 2 Bñ. - 130 m2	Alquilo departamento en Manta cerca de	24 de Mayo, Manabí	-0,9476708	-80,7229347	
17	\$ 320	1 Dor. - 1 Bñ. - 60 m2	Alquilo departamento en Manta amoblada	24 de Mayo, Manabí	-0,9447141	-80,739241	
18	\$ 350	3 Dor. - 3 Bñ. - 108 m2	Alquilo casa en Manta en urbanización	Manta, Manabí	-0,9870138	-80,7211125	
19	\$ 400	3 Dor. - 3 Bñ. - 100 m2	Alquilo casa en Duran en Urbanización	Manta, Manabí	-2,186493	-79,8190093	

Figura 2.11 Archivo obtenido de Web scraping. Fuente: Autoría propia.

## 2.1.2 Extracción de datos mediante Urbanpy

La librería permite tener acceso a diferentes tipos de datos que incurren en una ciudad como son: escuelas, bancos, centros comerciales, sitios turísticos, hospitales, etc. Que son ubicados en base a latitud y longitud permitiendo calcular el tiempo de viaje hacia ese destino a través de los medios de transportes urbanos. Para ello se debe instalar la librería en un entorno de ejecución Python como se muestra a continuación:

```
pip install urbanpy
```

A veces es necesario reiniciar el entorno para que la librería surta efecto. Una vez instalada se procede a cargar el entorno de ejecución junto con las librerías a utilizar.

```
importar urbanpy como up
importar pandas como pd
importar matplotlib.pyplot como plt
importar plotly.express como px
desde tqdm.auto importar tqdm
```

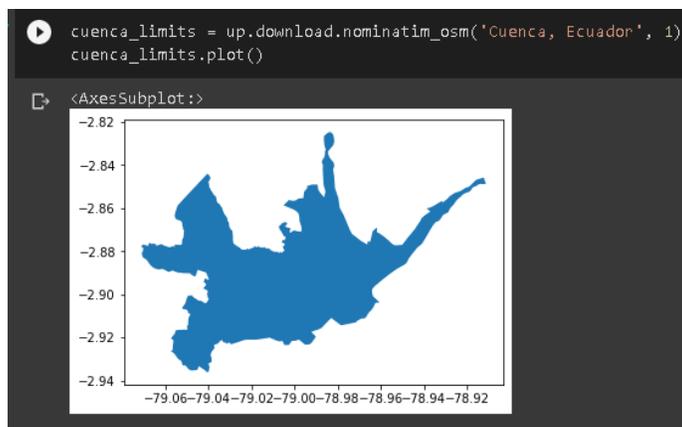
Código 2.4 Librerías para acceder a datos urbanpy.

Es necesario establecer los límites de la ciudad que en este caso entraría la lista de ciudades intermedias para ello ejecuto el siguiente código:

```
cuenca_limits = up.download.nominatim_osm('Cuenca, Ecuador', 1)
cuenca_limits.plot()
```

Código 2.5 Limitación de librería a ciudad.

Una vez ejecutado el código se procederá a tener una gráfica de los límites de la ciudad donde se obtendrá los datos comerciales como se muestra en la figura.



**Figura 2.12** Límite urbano de una ciudad con urbanpy. Fuente: Autoría propia.

Para obtener los datos de la base de urbanpy de igual manera se filtra de acuerdo con la ciudad.

```
cuenca_pop = up.geom.filter_population(ecuador_pop, cuenca_limits)
```

Dentro de la librería existen 4 grupos que constantemente se están actualizando como son: alimentos, salud, finanzas y centros de estudios que son ubicados en base a latitud y longitud. En este ejemplo se lo realizó la ubicación de centros que expender comida a lo largo de la ciudad.

```

facilidad_alimentos=up.download.overpass_pois(bounds=cuenca_hexs.total_bounds,
facilities='food')

```

Mediante la ejecución de este último código se genera una tabla con todos los lugares de acceso alimentos como son: centros comerciales, supermercados, minimarkets, etc. Para observar los primeros elementos de la tabla generada se ejecuta un head que permite observa los primeros elementos de una tabla.

```
[19] facilidad_alimentos = up.download.overpass_pois(bounds=cuenca_hexs.total_bounds, facilities='food')

facilidad_alimentos.head()
```

	type	id	lat	lon	tags	geometry	poi_type
0	node	617188415	-2.895971	-79.005697	{'name': 'La Prensa', 'shop': 'mall'}	POINT (-79.00570 -2.89597)	mall
1	node	617188417	-2.896458	-79.002622	{'name': 'El Correo', 'shop': 'mall'}	POINT (-79.00262 -2.89646)	mall
2	node	620528594	-2.900412	-79.022410	{'addr:housenumber': '4-104', 'addr:street': '...'}	POINT (-79.02241 -2.90041)	convenience
3	node	660174842	-2.929040	-78.971116	{'name': 'T. SN', 'shop': 'convenience'}	POINT (-78.97112 -2.92904)	convenience
4	node	690205085	-2.889667	-79.024255	{'name': 'Supermaxi Las Américas', 'opening_ho...'}	POINT (-79.02425 -2.88967)	supermarket

Figura 2.13 Tabla de acceso alimentos de una ciudad con urbanpy. Fuente: Autoría propia.

De igual manera se puede ubicar estos puntos dentro de un mapa para saber la distribución de estos. Se puede cargar de igual manera para centros de salud y centros financieros y mediante un filtro crear una tabla con lugares de interés como cajeros atm, hospitales y centros comerciales, tal como se muestra en la figura 2.14. Esta tabla puede ser llevada a cualquier herramienta GIS al transformarla en csv mediante el código:

```
data.to_csv('data_verticescuenca.csv')
```

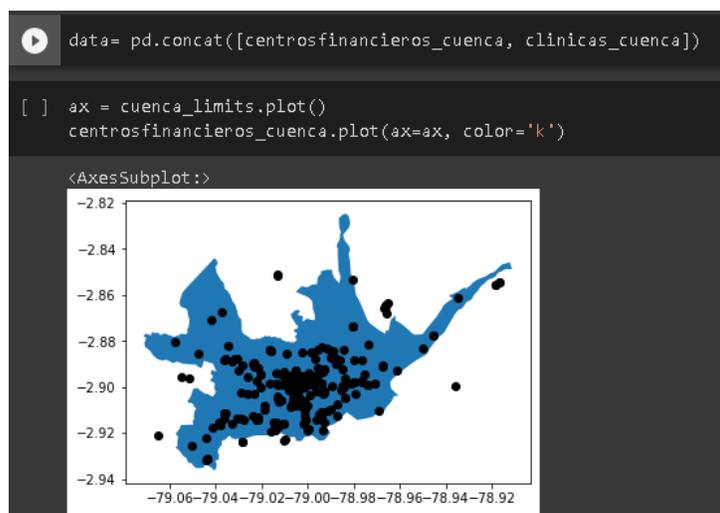


Figura 2.14 Filtro por lugares de interés dentro de una ciudad utilizando urbanpy. Fuente: Autoría propia.

## 2.2 Fase 2: Limpieza de Datos.

En este punto se realiza el análisis exploratorio datos para detectar datos inválidos, que permita realizar las respectivas correcciones o eliminación de datos nulos si

existieran. Siendo necesario modificar columnas, cambiar tipos de variables, eliminar elementos basura o no deseados como: m<sup>2</sup>, \$, %, -, \* Dor, Baños, etc. Que infieran al momento de ejecutar el análisis estadístico del conjunto de datos. Para realizar la limpieza respectiva se utilizó las siguientes librerías en Python:

- **Pandas:** Librería de código abierto que permite la manipulación de estructura de datos de una manera potente, fácil y flexible.
- **Missingno:** Permite la visualización de valores nulos a través de diagrama de barras dentro de un marco de datos con la ayuda de pandas.
- **Recordlinkage:** Realiza la unión de dos o más valores que pertenecen a una misma entidad.
- **Duplicated:** Permite la validación de uno o más datos duplicados dentro de una tabla.
- **Replace:** Se utiliza mucho para sustituir símbolos indeseados dentro de una columna en una tabla.

### 2.3 Fase3: Visualización de Datos.

Una vez limpiada la data se procede a cargar la misma al sistema de georreferenciación que en este caso es ARCGIS para ello se debe crear un nuevo proyecto y seleccionar mapa como se ve en la figura 2.15.

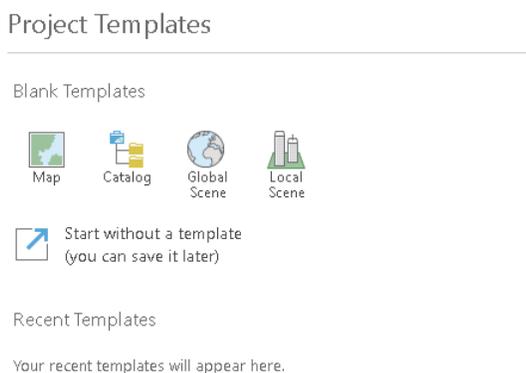
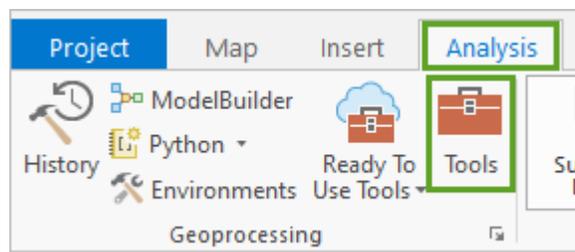


Figura 2.15 Selección de template de mapa en ArcGIS pro. Fuente: Autoría propia.

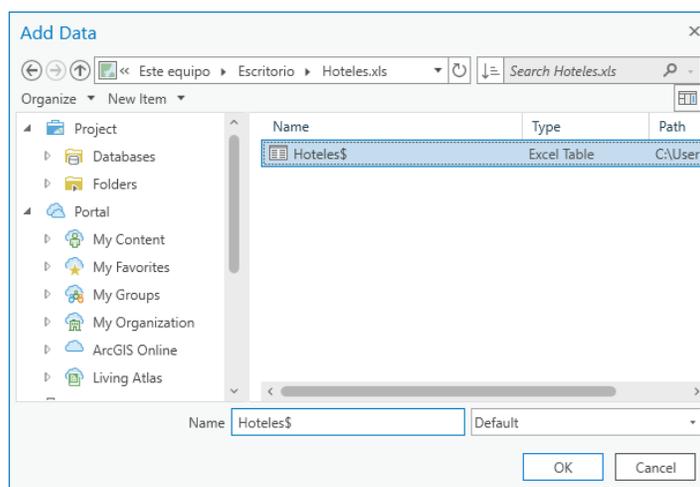
Una vez creado el mapa para cargar los datos a analizar se debe ir a la opción de geoprocesamiento> herramientas y escribir tabla XY.



**Figura 2.16 Ejecución de la herramienta XY a punto. Fuente: Autoría propia.**

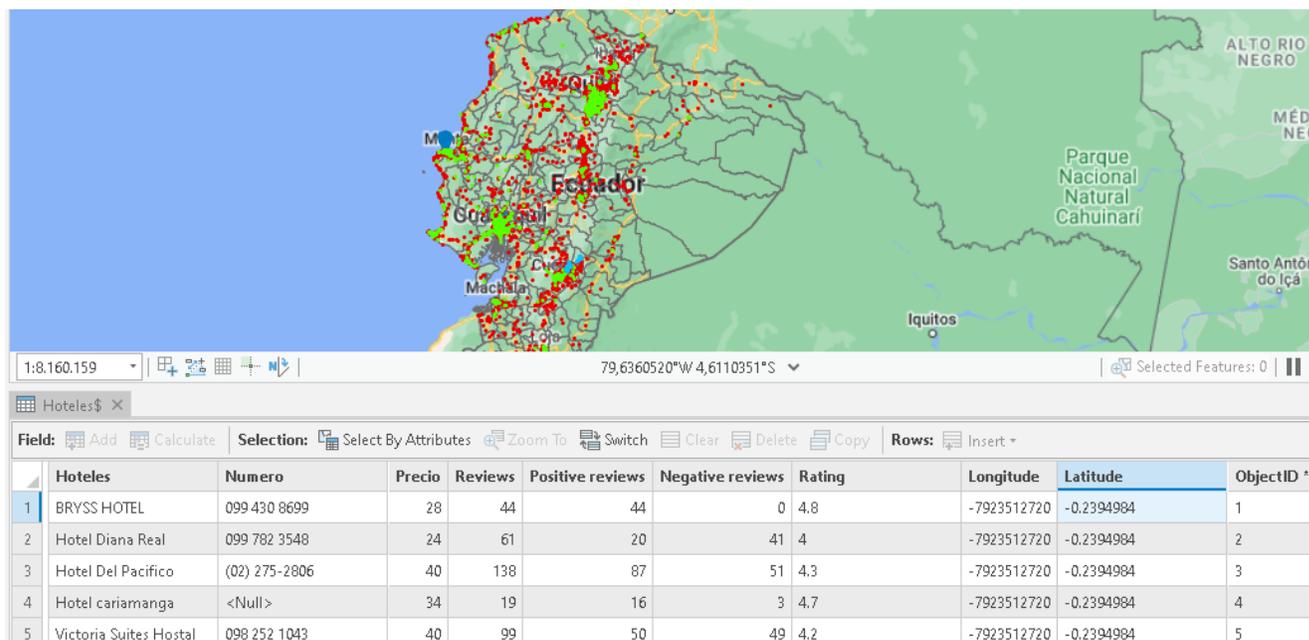
Esta opción permite el despliegue de punto de datos a lo largo del mapa creado para ello dentro de los requerimientos, se encuentra llenar los campos XY que son en sí la correspondencia de la latitud y longitud respectivamente. Se debe almacenar la data en un lugar de referencia específico de preferencia en un servidor de base de datos donde se pueda llamar a los diferentes tipos de tablas creados para el respectivo análisis en base a la zona de referencia de Ecuador como geodatabase, tablas obtenidas de páginas web por medio del Web scraping, tablas de variables socioeconómicas de Redatam del censo 2010.

Para añadir tablas y shape se deber ir a la parte de *add data>add data to map* como se puede visualizar en la figura 2.17.



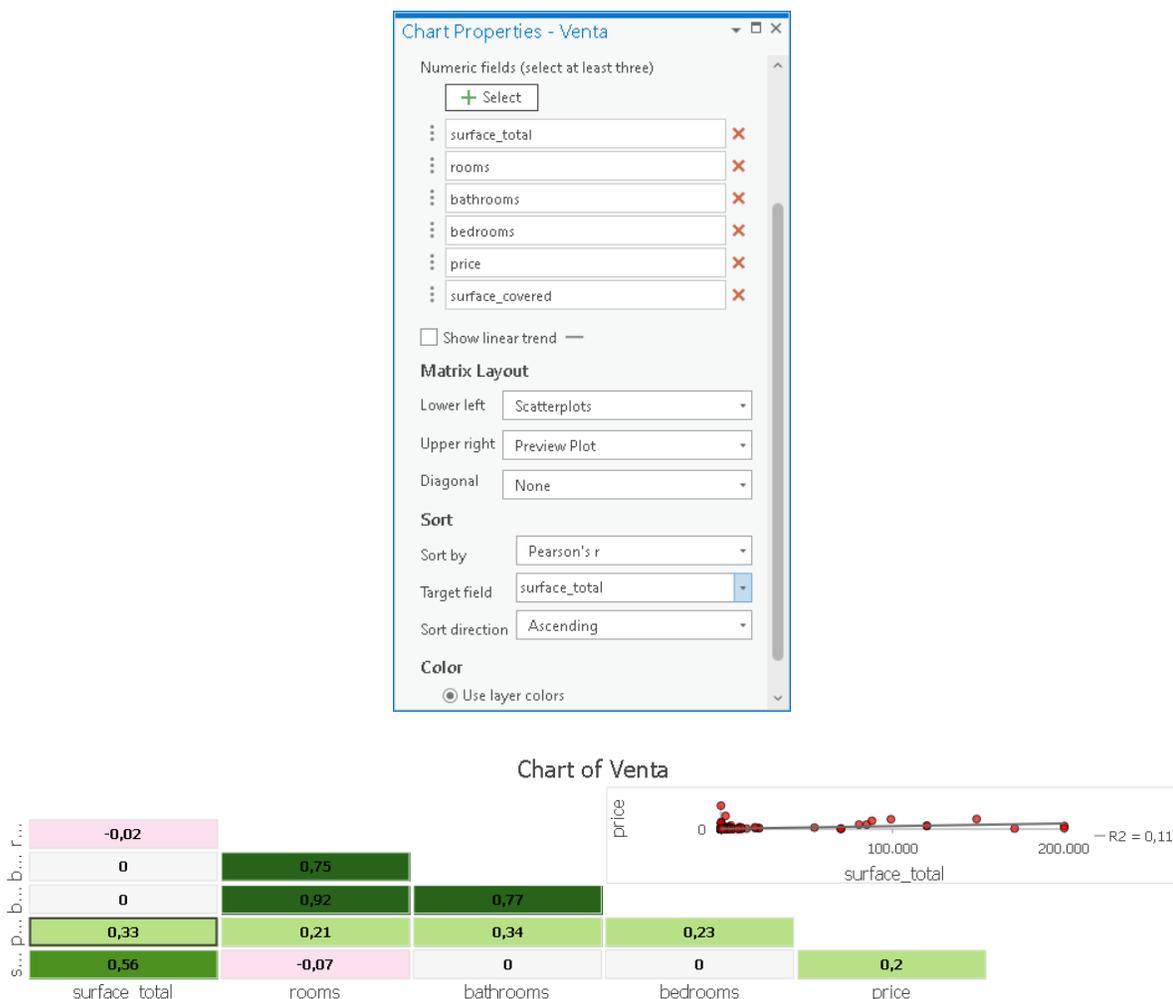
**Figura 2.17 Ingreso de tabla dentro de mapa en ArcGIS Pro. Fuente: Autoría propia.**

El anterior se lo realiza por cada una de las tablas obtenidas para poder visualizar el contenido de las misma se debe dar clic derecho > abrir tabla obteniendo el resultado de la figura 2.18 que se encontraría listo para empezar analizar dentro de la herramienta.



**Figura 2.18** Tabla con visualización de puntos ArcGIS Pro. Fuente: Autoría propia.

A continuación, se creará y explorará una matriz de gráficos de correlación de las variables en cuanto a venta de casas seleccionando los datos de tipo numéricos. Seleccionando como principal la matriz de coeficientes de Pearson como se muestra en la figura 2.19.



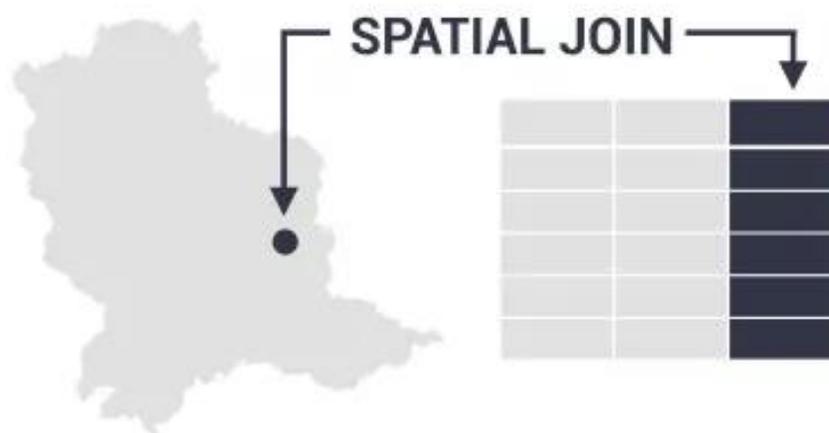
**Figura 2.19** Relación de variables en base al coeficiente de Pearson. Fuente: Autoría propia.

Se puede evidenciar una fuerte relación entre la superficie de la vivienda con el precio que es lo que determina su valor, pues en la mayoría de las páginas no detalla los acabados. Con esto se puede elaborar un modelo que identifique zonas cuyo precio difiera a una vivienda con las mismas características, tales como: precio, número de cuartos, número de baños y metros cuadrados. Para relacionar con los demás datos se debe especificar las zonas a evaluar para ello se procedió a crear nuevos shapefiles<sup>4</sup> de acuerdo con la ciudad Intermedia.

<sup>4</sup> Un shapefile es un formato sencillo y no topológico que se utiliza para almacenar la ubicación geométrica y la información de atributos de las entidades geográficas

### 2.3.1 Creación De Zonas.

Para crear cada una de las zonas que pertenece a cada ciudad intermedia se deber utilizar la herramienta selección de entidades que permite crear un área limítrofe en el mapa que permite la unión de atributos de varios shapefiles. Se utilizó la Geodatabase del INEC relacionando a través de un JOIN (unión), a cada una de las tablas para realizar el análisis de acuerdo a parroquias de la ciudad en base al código INEC o a través de la ubicación en latitud y longitud, creando una relación punto espacio como se muestra en la figura 2.20. En ArcGIS se debe ir a Analysis Tools > Overlay > Spatial Join.



**Figura 2.20 Creación de zona limítrofe de una ciudad intermedia. Fuente: Autoría propia.**

Una vez empastada la data y encapsulada dentro de una zona se crea una super tabla que contiene las variables de censo, hoteles en la región, data de plataformas de internet mediante Web scraping. Para comprobar la correcta asignación se debe ir a propiedades > tabla de atributos.

Current Layer: Zonas_cuenca_SpatialJoin1									
Visible	Read Only	Field Name	Alias	Data Type	Allow NULL	Highlight	Number Format	Domain	Default
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	OBJECTID_1	OBJECTID_1	Object ID	<input type="checkbox"/>	<input type="checkbox"/>	Numeric		
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Shape	Shape	Geometry	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Join_Count	Join_Count	Long	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Numeric		
<input checked="" type="checkbox"/>	<input type="checkbox"/>	TARGET_FID	TARGET_FID	Long	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Numeric		
<input checked="" type="checkbox"/>	<input type="checkbox"/>	JOIN_FID	JOIN_FID	Long	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Numeric		
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Id	Id	Long	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Numeric		
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_ZONA	DPA_ZONA	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_VALOR	DPA_VALOR	Short	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Numeric		
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_ANIO	DPA_ANIO	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_PARRROQ	DPA_PARRROQ	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_NOMBRE	DPA_NOMBRE	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_TIPO	DPA_TIPO	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_LAM	DPA_LAM	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_CATEG	DPA_CATEG	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	DPA_ESCALA	DPA_ESCALA	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Codigo	Codigo	Text	<input checked="" type="checkbox"/>	<input type="checkbox"/>			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Que_es_reconocido_por_el_CONESU	Que es reconocido por el CONESUP	Double	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Numeric		

Figura 2.21 Tabla relacional generada por herramienta Spatial Join. Fuente: Autoría propia.

## 2.4 Fase 4: Desarrollo de Método Automático.

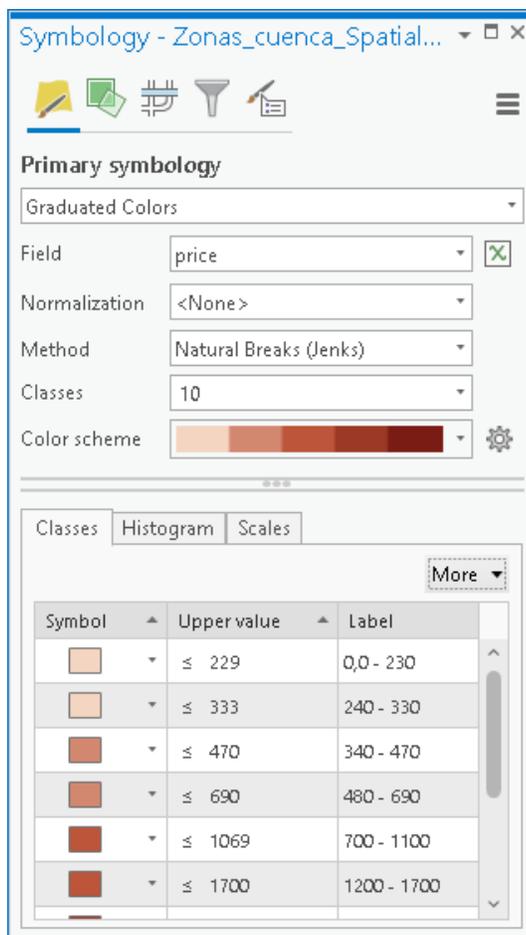
Una vez obtenida la super tabla se explora las diferentes características numéricas para indicar una tendencia a través de una regresión exploratoria en donde se obtenga un modelo que identifique las diferencias áreas de acuerdo con el precio de vivienda, m<sup>2</sup>, nivel de educación superior de acuerdo con el censo 2010, tendencia del número de estrella y precios de hoteles en la zona.

Para ello se debe cambiar la simbología de la zona en ArcGIS para tener una clasificación en base a las parroquias creadas dentro de la ciudad Intermedia. Donde se determinar su impacto de acuerdo con las variables antes mencionadas. Dar clic derecho sobre la zona e ir a simbología y escoger lo siguientes parámetros:

- Simbología principal, graduada por colores.
- Campo principal, variable dependiente el precio de vivienda.
- Clases, seleccionar un numero razonable en este caso 10.
- Método seleccionar jenks<sup>5</sup>.

<sup>5</sup> es un método de agrupación de datos diseñado para determinar la mejor disposición de valores en diferentes clases

En la figura 2.22, se puede visualizar la selección de este método junto con la creación de zonas de acuerdo con varios rangos.



**Figura 2.22 Creación de clases para variable precio. Fuente: Autoría propia.**

Luego de crear las zonas se procede a elaborar un modelo donde se pueda caracterizar las distintas variables, siendo necesario dirigirse a model builder y seleccionar un lienzo en blanco donde se arrastrará la zona con la super tabla creada para juntarla a un iterado que permita realizar una regresión multivariable, tal como se muestra en la siguiente figura 2.23, sin olvidar dar clic en Skip null values para ignorar valores nulos o erróneos.

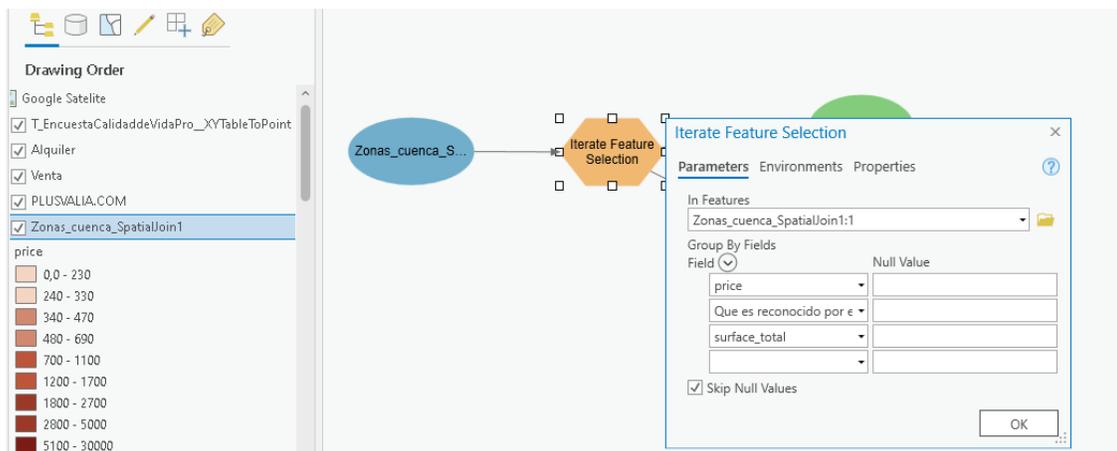


Figura 2.23 Creación de modelo de regresión. Fuente: Autoría propia.

En ArcGIS Pro existe la herramienta de geoprocésamiento de Estadísticas espaciales. Una de las primeras a evaluar es la regresión generalizada (GLR) que se encuentra dentro de la caja de herramientas, tal como se muestra en la figura.

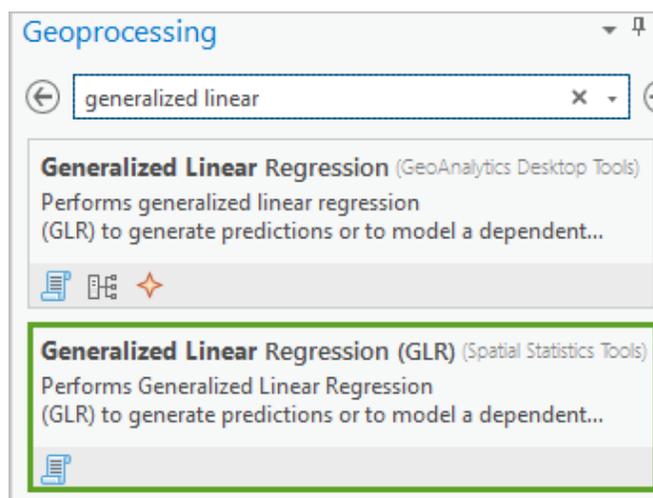
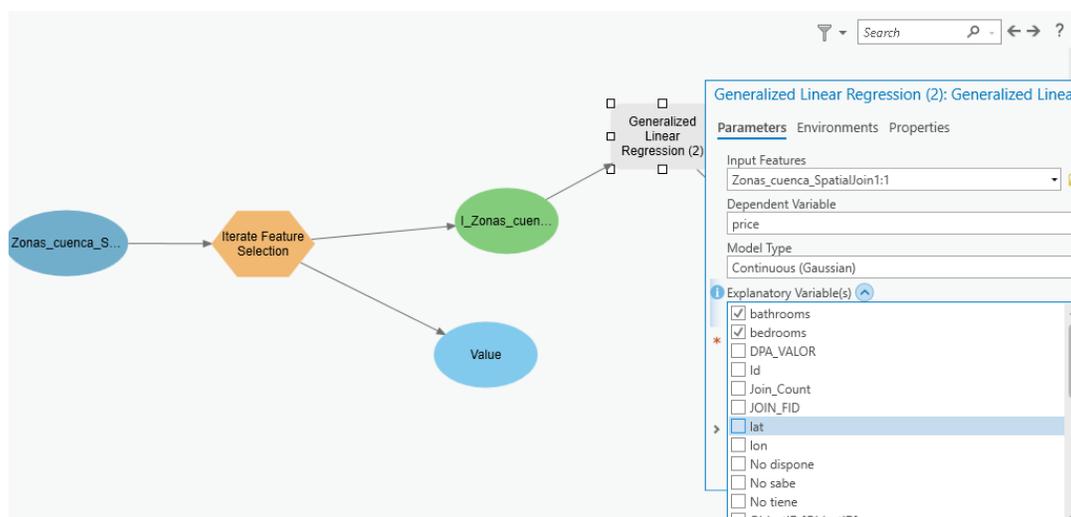


Figura 2.24 Herramienta de análisis para realizar regresión lineal en los datos de tabla. Fuente: Autoría propia

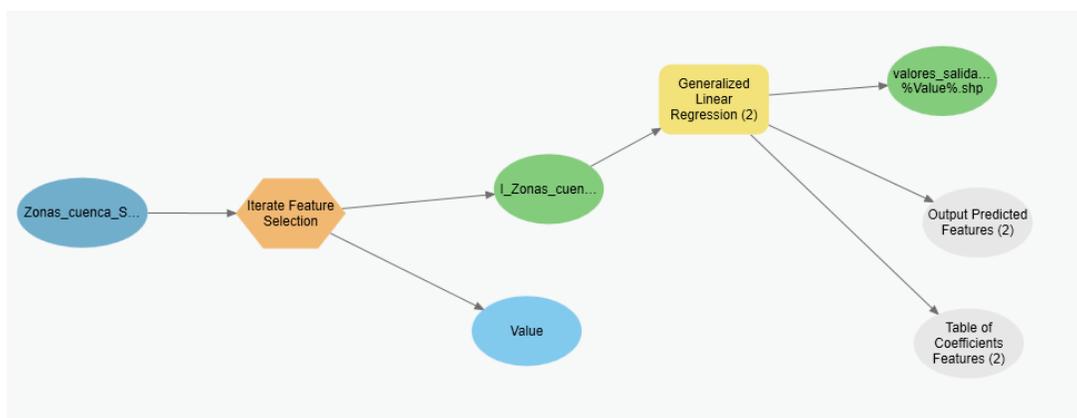
Para ejecutar la herramienta Regresión lineal generalizada (GLR) y predecir distintos tipos de variables dependientes en base al modelo de gentrificación. El modelo correcto para utilizar depende del tipo de variables dependientes e independientes. Dado que predice una variable continua (precio de venta), utilizará un modelo

gaussiano para predecir el precio de venta de la vivienda en promedio en base a una zona de la ciudad Intermedia.



**Figura 2.25 Ingreso de variables dependientes e independientes en el modelo. Fuente: Autoría propia.**

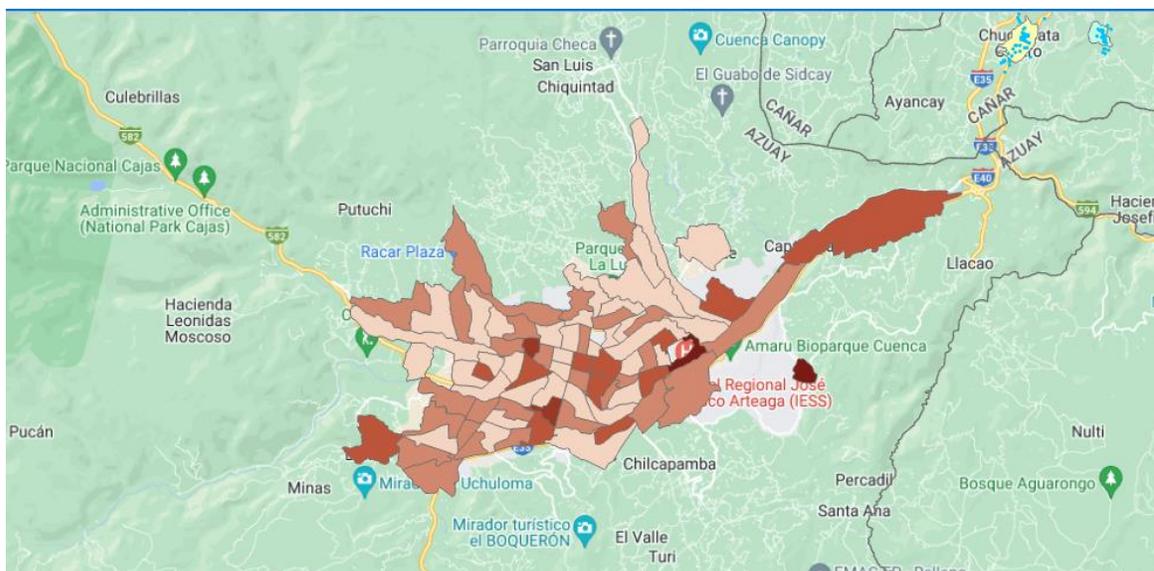
Una vez creado todo el modelo generará una capa que evidenciará un resultado en base a las variables ingresadas con la regresión lineal múltiple.



**Figura 2.26 Model builder de regresión lineal en base a características de modelo de gentrificación.**

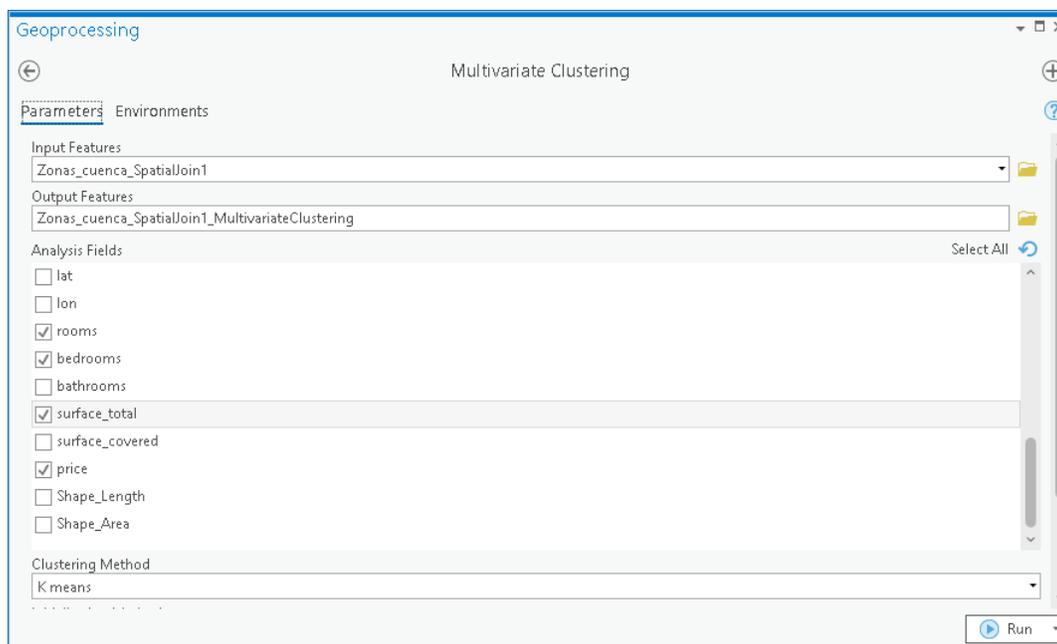
**Fuente: Autoría propia.**

Al realizar un clic en ejecutar dicho modelo probado en una de las principales ciudades intermedia como lo es la ciudad de Cuenca arroja un área sectorizada por precios de arriendos que es la variable dependiente en este caso.



**Figura 2.27** identificación de zonas en base a variables de gentrificación. Fuente: Autoría propia.

De igual manera se puede realizar el estudio de las variables dependientes mediante la herramienta Clustering multivariante que permite clasificar las zonas en base a la característica que se especifique.



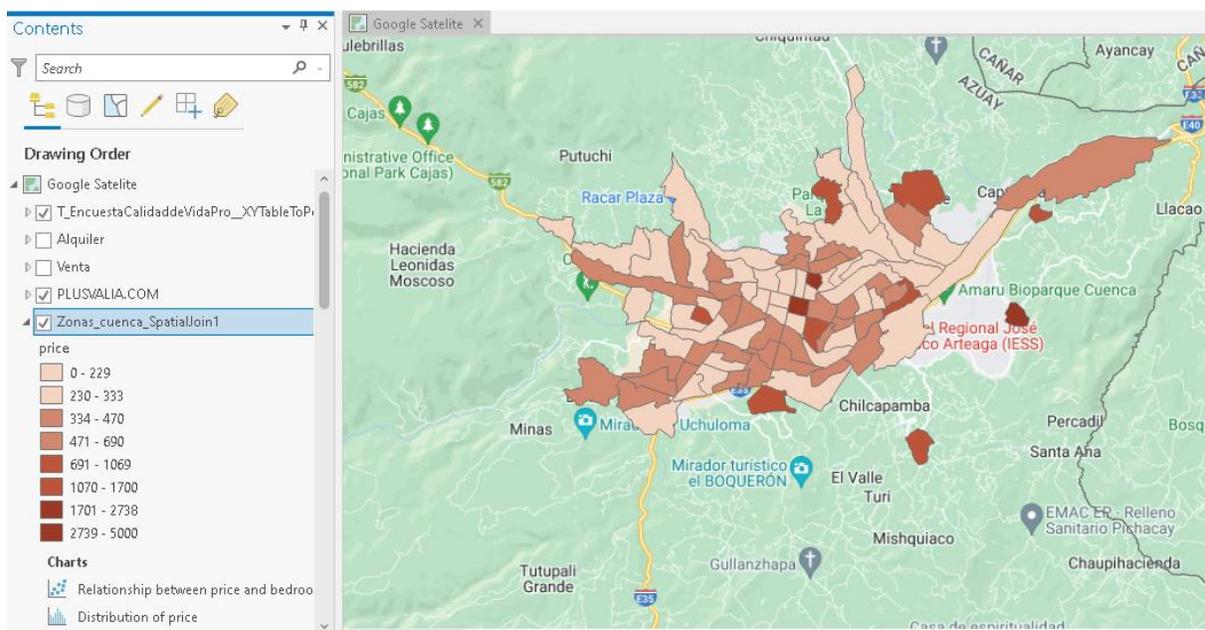
**Figura 2.28** Ingreso de variables a clúster. Fuente: Autoría propia.

También se puede utilizar el algoritmo de clasificación k-meas por medio del Python de ArcGIS. Para ello, se ejecuta un pequeño script que permita clasificar las zonas de acuerdo con las variables dependientes que se ejecutó en la parte gráfica.

```
1 importar arcpy
2 arcpy.env.workspace = r"C:\Datos\LPA\mapagoogole\Default.gdb"
3 arcpy.MultivariateClustering_stats( "price" ,
4     [ "Que es reconocido por el CONESUP", "rating",
5     "precio_hotel"], "K_MEANS", "OPTIMIZED_SEED_LOCATIONS", None, "8")
```

**Código 2.6 Ejecución de K\_meas en Python de ArcGIS pro.**

EL resultado de la ejecución de este código es igual que el de manera gráfica, su resultado es mostrado en la figura 2.29.



**Figura 2.29 Resultado de ejecución de script. Fuente: Autoría propia.**

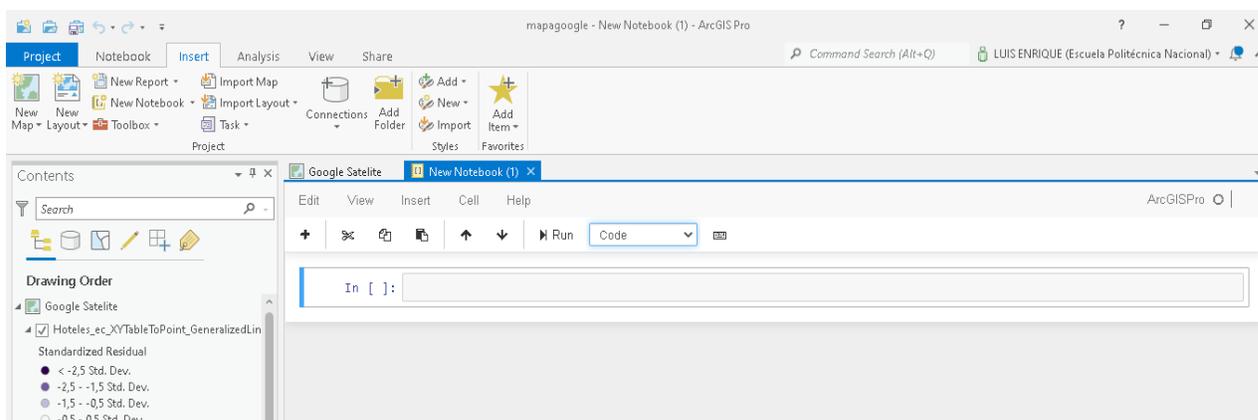
Como se puede visualizar esta herramienta crear mensaje y gráficos para tener una mejor comprensión en cuanto a la zona de análisis cabe recalcar que los campos de análisis siempre deben ser numéricos. Para obtener estos resultados se debe tener en cuenta que se debe particionar el dataset en este caso se utilizó un 70% de datos para el entrenamiento del modelo y el 30% para las pruebas respectivas del mismo.

En base a este entrenamiento se pudo generar un modelo en base a Deep learning utilizando Python de ArcGIS PRO, siendo necesario descargarse las librerías de desde el repositorio<sup>6</sup>. Una vez instalado el ambiente se procedió a emplear las siguientes librerías detalladas en la siguiente tabla:

**Tabla 2.1 Librerías cargadas en ArcGIS.**

<b>LIBRERÍA</b>	<b>UTILIDAD</b>
<b>PANDAS</b>	Se utilizó para manipulación y análisis de datos.
<b>NUMPY</b>	Ejecución de vectores y matrices de datos.
<b>SEABORN</b>	Se empleó para visualización de datos.
<b>SKLEARN</b>	Librería de aprendizaje automático de Python.
<b>TENSORFLOW</b>	Librería para implementación de redes neuronales.

Para la importación de la ejecución de este modelo es necesario crear un nuevo notebook dentro del ambiente, tal como se muestra en la figura 2.30.



**Figura 2.30 Creación de notebook en ArcGIS PRO. Fuente: Autoría propia.**

<sup>6</sup><https://github.com/Esri/deep-learning-frameworks/blob/master/include/install-deep-learning-frameworks-manually-2-8.pdf>

Si es necesario se debe reiniciar el IDLE de Python para tener una conexión con el kernel. La importación de las librerías quedo de la siguiente manera:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler

from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.optimizers import Adam

from sklearn.metrics import
mean_squared_error, mean_absolute_error, explained_variance_score
from sklearn.metrics import classification_report, confusion_matrix
```

**Código 2.7 Librería para código Deep learning.**

Es de vital importancia antes de empezar cual modelo el asegurar tener data limpia, aunque para los anteriores modelos se usó la ayuda que proporcionaba ArcGIS tanto en regresión como em k-means. En este se debe ejecutar lo siguientes comando para cerciorarse que el conjunto de datos quede sin datos nulos.

```
#Realizala busqueda de datos vacios
df . es nulo() . valores _ any()
len (df)
#Elimina datos vacios
df = df . dropna()
len (df)
```

**Código 2.8 Búsqueda y eliminación de datos vacíos.**

Cabe recalcar que los datos desde el scraping fueron limpiados, pero se lo realiza por simple inspección. Luego, se procedió a carga el conjunto de datos junto con el análisis multivariable que se encontró en el mismo para crear el conjunto de entrenamiento con el 70% de la data original de acuerdo con cada ciudad intermedia que resultado del spatial join.

```
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.7,random_state=20)
```

Posteriormente se levanta la arquitectura de la red neuronal que utiliza hiper parámetros mencionados en los puntos anteriores que permiten la constitución de un modelo de capas constituida por 19 capas de entradas, 15 capas ocultas y 1 capa de salida.

```

model = Sequential()
#capa entrada
model.add(Dense(19, activation='relu'))
#capas ocultas
model.add(Dense(15, activation='relu'))
model.add(Dense(5, activation='relu'))
#capa salida
model.add(Dense(1))

model.compile(optimizer='adam', loss='mse')

```

**Código 2.9 Modelo de capas.**

Dentro de este modelo de capas se utilizó el optimizador de Adam que es un algoritmo de optimización de reemplazo para el descenso de gradiente estocástico para entrenar modelos de aprendizaje profundo. Adam combina las mejores propiedades de los algoritmos AdaGrad y RMSProp para proporcionar un algoritmo de optimización que puede manejar gradientes dispersos en problemas ruidosos con es el caso del precio de viviendas y hoteles en el sector de una ciudad. Así mismo se valida el conjunto de entrenamiento en base al batch size y numero de épocas para entrenar,

```

model.fit(x=X_train,y=y_train.values,validation_data=(X_test,y_test.values
),
        batch_size=150,
        epochs=200,
        verbose=1)

```

**Código 2.10 Validación de modelo.**

Finalmente, se comprueba el modelo al predecir un dato al azar con el resto de los datos que no participaron en el entrenamiento.

```

prediccion = model.predict(X_test)
print('MAE', mean_absolute_error(y_test,prediccion))
print('VR', explained_variance_score(y_test,prediccion))
MAE 131114.12727507274
VR 0.797185875872096

```

**código 2.11 Verificación de modelo utilizando Deep learning.**

A continuación, se presenta los resultados obtenidos por el modelo aplicando las diferentes capas.

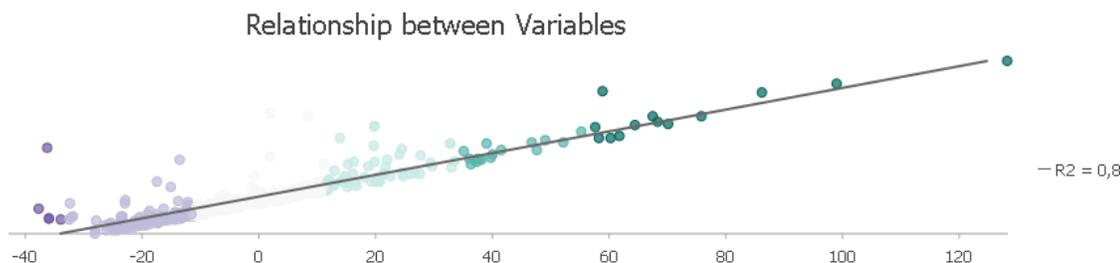


Figura 2.31 Relación entre predicción y valor real. Fuente: Autoría propia.

## 2.5 Fase 5: Validación y Evaluación del Modelo.

En esta fase, fue necesario realizar la validación de cada uno de los modelos antes expuestos, ya que cada ciudad al formar la unión de datos se generarán tablas de análisis con volúmenes diferentes, concentración de zonas resultan en algunos casos saturada y en otros no dependiendo del comercio y nivel de estado de gentrificación. Por lo tanto, se obtuvieron tablas de datos que fueron procesadas con cada modelo para cada una de las ciudades intermedias. Para ello, se efectuará las medidas de criterio de desempeño como son: Coeficiente de determinación, error cuadrático medio y coeficiente de evaluación de Pearson.

### 2.5.1 Coeficiente de Determinación $R^2$ .

Cada uno de los diferentes modelos empleados dentro del entorno de trabajo con el IDLE de ArcGIS fueron regresión lineal, k-meas y redes neuronales. Para conocer si un modelo es preciso es necesario recurrir al coeficiente de determinación que permite ver que tanto se ajusta el modelo a los valores reales. Se parte de la ecuación que permite ver dicho coeficiente:

$$R^2 = \frac{\sum_{t=1}^T (\hat{Y}_t - \bar{Y})^2}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

Donde:

$\hat{Y}_t$ =Precio predicho en el modelo.

$Y_t$  = Precio real de la vivienda.

n= Número total de predicciones.

La ecuación empieza realizando la sumatoria de precios de casas establecidos en el conjunto de datos menos el precio generado por la predicción del modelo; esto también se conoce como la varianza, dividido para la sumatoria de precios que se conoce menos lo predicho por el modelo.

### 2.5.2 Error Cuadrático Medio RMSE.

Es una medida que determina el error que existe en un conjunto de datos tomados como muestra, es decir se realiza la comparación del valor predicho con el real. Mientras más pequeño sea el RMSE, mejor representado está el modelo.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}}$$

Donde:

$P_i$ =Precio calculado por la predicción.

$O_i$ =Precio de vivienda real.

n= Número total de predicciones

### 2.5.3 Coeficiente de Variación de Pearson (CV).

Es una medida estadística empleada para medir la dispersión relativa de un conjunto de datos. Se lo obtiene al dividir la desviación típica con el valor absoluto de la media de precios obtenidas del conjunto de datos.

$$CV = \frac{RMSE}{|\bar{x}|}$$

Donde:

RMSE= Error cuadrático medio.

$|\bar{x}|$ = Es el promedio de la variable que se desea medir.

## 2.6 Fase 6: Presentación de Resultados por Ciudad.

Se realizó el cálculo de cada una de las ciudades obtenido con cada uno de los métodos de machine learning donde se utilizó la medida de R<sup>2</sup>, ya determina que método se ajusta a las variables socioeconómicas del censo 2010 junto con los datos extraídos de viviendas y hoteles de la zona seleccionados los cuales son presentados en la siguiente Tabla:

Tabla 2.2 Resultados obtenido por modelos ML.

CIUDAD	REGRESIÓN LINEAL			K-MEANS			RED NEURONAL		
	R <sup>2</sup>	RMSE	CV	R <sup>2</sup>	RMSE	CV	R <sup>2</sup>	RMSE	CV
ESMERALDAS	0,71	1,42E+08	0,24	0,8	1,56E+08	0,15	0,56	1,4E+09	0,18
MANTA	0,68	6,65E+08	0,24	0,77	4,24E+08	0,18	0,65	3,2E+09	0,27
SANTO DOMINGO	0,53	6,46E+08	0,25	0,71	3,87E+08	0,17	0,72	4,5E+10	0,16
QUEVEDO	0,72	2,57E+08	0,22	0,68	3,79E+08	0,15	0,63	5,9E+10	0,26
LOJA	0,56	7,79E+08	0,25	0,65	4,96E+09	0,19	0,58	7,4E+09	0,27
CUENCA	0,75	6,55E+08	0,17	0,82	2,26E+09	0,15	0,84	7,2E+09	0,14
AMBATO	0,68	1,32E+08	0,19	0,72	4,16E+09	0,18	0,68	3,4E+09	0,19
RIOBAMBA	0,63	5,47E+08	0,23	0,69	2,66E+08	0,17	0,71	4,4E+09	0,22
IBARRA	0,72	2,99E+08	0,16	0,712	1,9E+08	0,18	0,66	5,8E+09	0,23
LATACUNGA	0,61	4,97E+08	0,168	0,66	2,47E+08	0,14	0,58	1,2E+09	0,18
TULCÁN	0,635	1,56E+08	0,15	0,712	3,56E+08	0,17	0,43	1,3E+09	0,27

Como se puede observar en la tabla 6, el modelo que mejor se ajusta a una clasificación multivariable dada por los censos y a través del Web scraping es la de K-MEANS, pues permite analizar grupos en base a un radio específico que son las variables socioeconómicas censo (2010), precios de hoteles de trivago y Airbnb con las características obtenidas por cada página que puede ser: número de estrellas, comodidades, tales como: aire acondicionado, personas por cuarto y ubicación a sitios turísticos. Razón por la cual el error es menor utilizando K-MEANS debido a que la intención es reconocer grupos o zonas donde existe mayor comercio y por ende tiende a ser gentrificado. Cabe recalcar que se puede ir mejorando el modelo con redes neuronales, pero se transforma en un modelo más complicado ya que se está analizando muchas variables y en algunas ciudades no se alcanza un conjunto de

entrenamiento grande debido a que no cuentan con desarrollo comercial a nivel de plataformas de ventas de casas. Es así como al final de predecir y cuantificar las múltiples variables, se obtiene el índice de gentrificación junto con el estado dicho índice. Para luego, ser visualizado en el mapa de acuerdo con un color que se asigne por estado de gentrificación en una zona específica. Los estados que se muestran al final de procesar son gentrificados para un grupo con índice de gentrificación de +7 extremadamente gentrificado, gentrificado entre +4 a +6, +3 en proceso y menos de +2 no gentrificado.

El nivel de gentrificación en una zona se mostrará en el siguiente capítulo, pues es necesario cargar las tablas extraídas de los modelos para que esto se visualice en mapa mediante un dashboard.

## CAPÍTULO III

### 3. Diseño e Implementación del Dashboard

Para el diseño e implementación del dashboard, se hará uso de la cuenta otorgada por la entidad de estudio de ArcGIS Online. La cual permitirá sincronizar los mapas y capas generadas (shapefiles) por cada una de las ciudades intermedias que se ha realizado el análisis de datos. Dentro de la aplicación se clasificarán por ciudad donde se visualizarán los puntos de mayor gentrificación en base al costo de arriendo de habitaciones y hoteles obtenidos de los datos del Web scraping. Los resultados que se obtendrán dentro del dashboard serán de carácter público por lo tanto se creará un enlace de acceso libre en la cuenta de ArcGIS Online. Para configurar y visualizar los resultados se ha dividido en 5 etapas que son:

- a) Registro.
- b) Creación de mapa base.
- c) Carga de capas de ArcGIS desktop a ArcGIS Online.
- d) Generación de mapas temáticos por ciudad.
- e) Despliegue de dashboard por medio de colección.

#### 3.1 Registro.

El presente proyecto se creó una cuenta con el correo institucional de la organización de estudio que tiene convenio con la empresa ESRI propietaria de ArcGIS Online para ello se debe ingresar los datos en la dirección <https://www.arcgis.com/index.html>. Donde se procedió a configurar el perfil que permita publicar mapas y tablas disponibles en ArcGIS desktop. Es necesario cambiar al rol de publicador donde permitirá tener acceso a datos en nube de la organización y contar con un espacio para publicar aplicaciones de tipo web y crear historias para ello es necesario ir a perfil->Mi configuración->roles tal como se muestra en la figura 3.1.

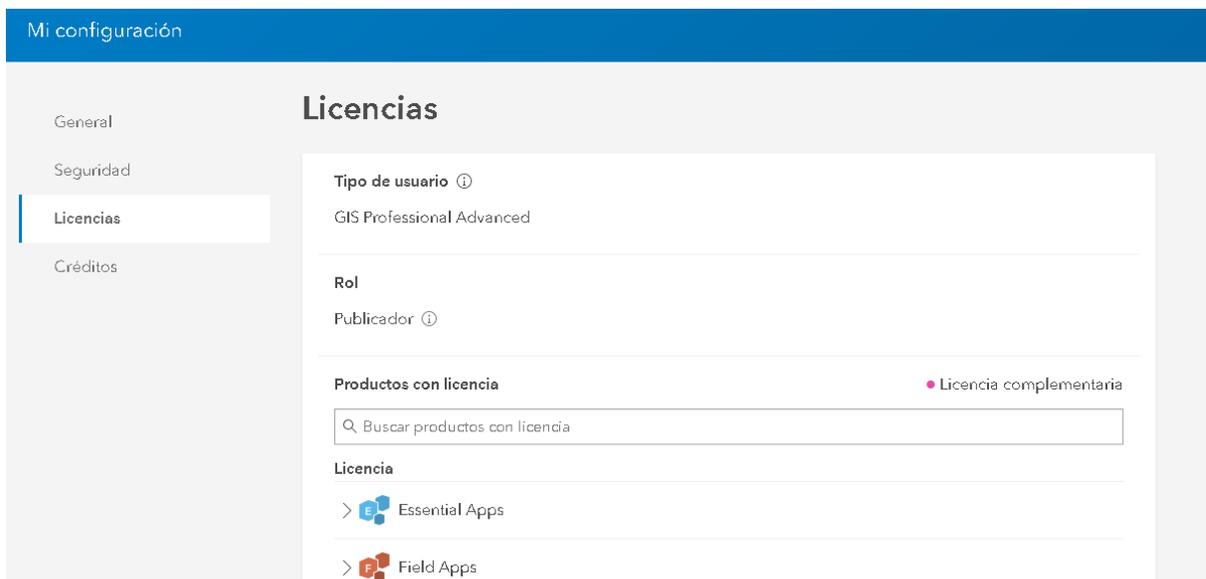


Figura 3.1 Asignación de rol publicador. Fuente: Autoría propia.

## 3.2 Creación de Mapa Base.

Para tener un ente de referencia es necesario crear un mapa base en la cuenta de ArcGIS Online el mismo que servirá para sincronizar con las capas de la herramienta de escritorio. Se debe ir a inicio-> mapa->seleccionar mapa base como se muestra en la figura 3.2.

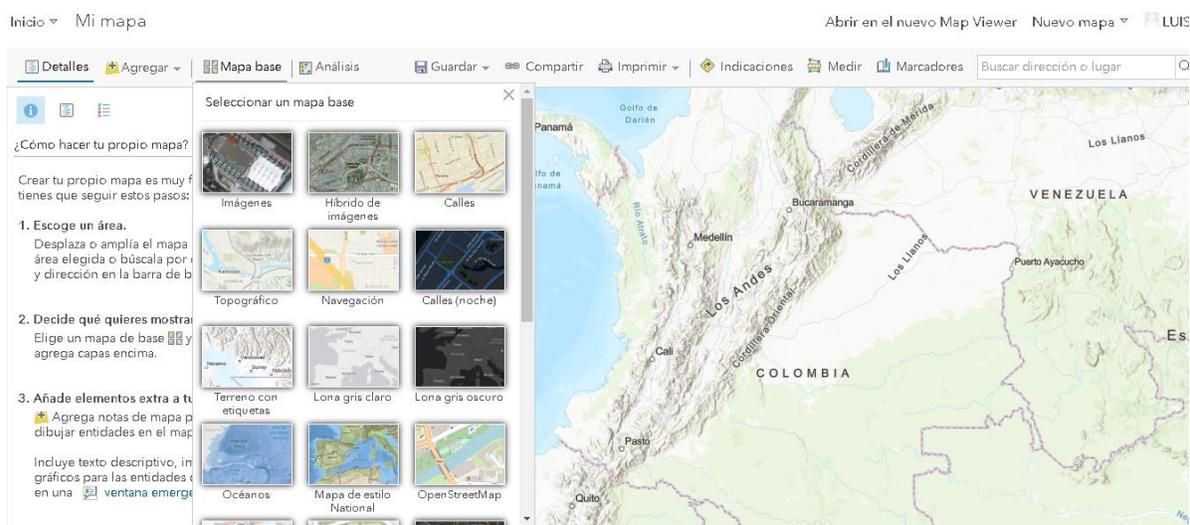


Figura 3.2 Selección de mapa base ArcGIS Online. Fuente: Autoría propia.

Este mapa base creado debe tener los permisos de publicación para que se pueda sincronizar con la cuenta de escritorio.



Figura 3.3 Ajustes de permisos mapa base. Fuente: Autoría propia.

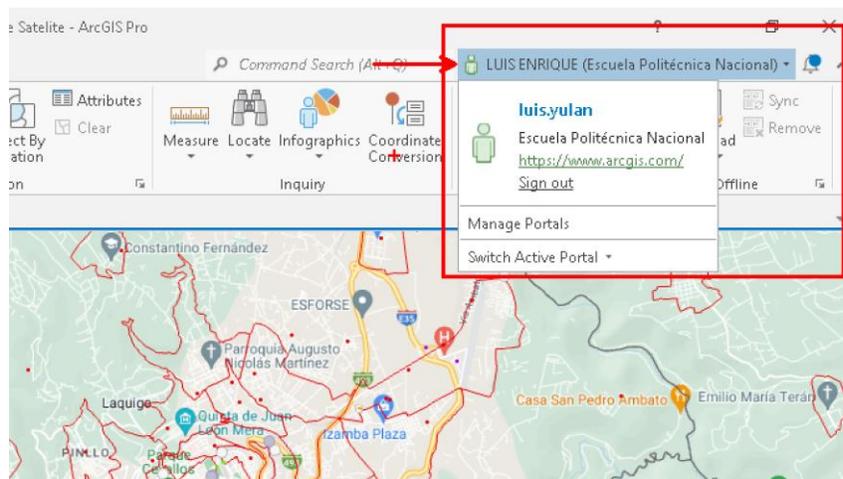
Una vez creado el mapa base, dado los permisos respectivos, ya se puede empezar a cargar contenido.

### 3.3 Carga de Capas de ArcGIS Desktop a ArcGIS Online.

Dentro de la pestaña **Mi contenido** de ArcGIS Online, se crearán varias carpetas que contendrán las zonas de cada una de las ciudades que están divididas en geodatabase, iconos de gentrificación, datos de hoteles, lugares turísticos, entre otros. Una vez ubicado el espacio donde se cargarán los datos es necesario abrir la aplicación de escritorio donde se encuentran todos los datos analizados en el capítulo anterior. ArcGIS Pro se conecta de forma predeterminada con su cuenta Online si es que se utiliza la misma organización, pero en caso de contar con otra se debe agregar el link del portal con los siguientes pasos:

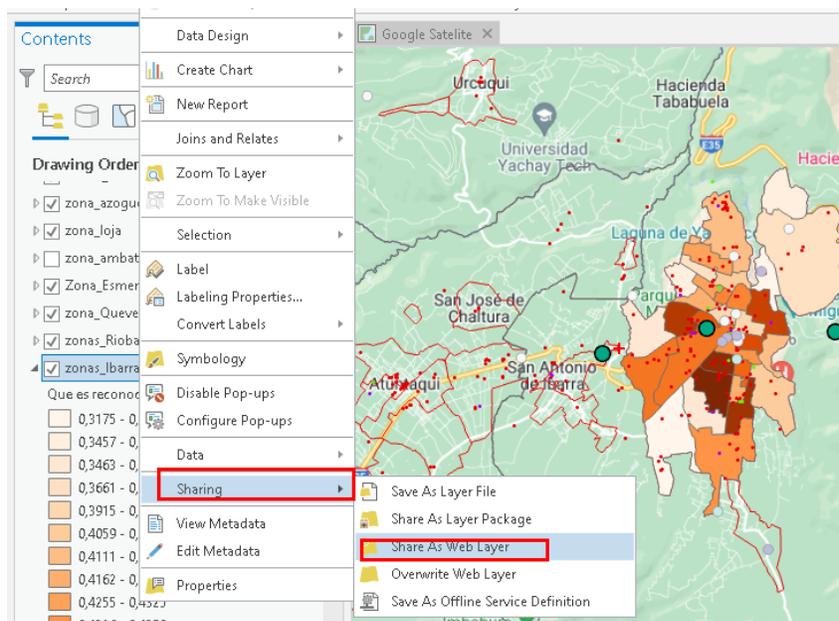
- 1) Hacer clic en el proyecto y dirigirse a portales.
- 2) Dar clic en agregar portal.
- 3) Introducir la URL de la organización en este caso es <https://epn-ec.maps.arcgis.com/>
- 4) Una vez agregado el URL definir como portal activo.

Para comprobar que tuvo éxito la agregación y sincronización visualizar en la parte derecha que aparezca en verde el logo de la cuenta de ArcGIS Online junto con la organización.



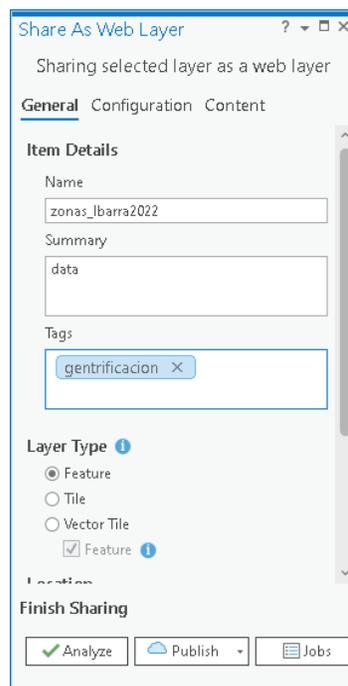
**Figura 3.4 Sincronización de ArcGIS Desktop con ArcGIS Online. Fuente: Autoría propia.**

Para cargar una capa es necesario dirigirse a la misma presionar clic derecho y seleccionar la opción sharing->Share As Web Layer como se muestra en la figura siguiente.



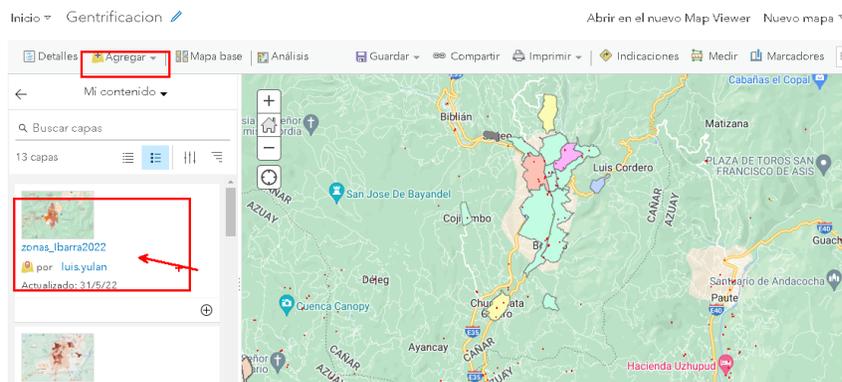
**Figura 3.5 Compartición de Capa en ArcGIS Desktop. Fuente: Autoría propia.**

Una vez marcada esa opción mostrará una ventana emergente donde permitirá cargar la capa hacia los contenidos de ArcGIS Online. Se debe seleccionar siempre el tipo de layer como feature para mantener las características de la capa como se muestra en la figura siguiente. Finalmente se debe presionar el botón publicar.



**Figura 3.6 Publicar capa en ArcGIS ONLINE. Fuente: Autoría propia.**

La capa cargada ya será visible dentro del contenido de ArcGIS online, esto permitirá cargar dicha capa en el mapa base antes creado. Para ello, es necesario dirigirse al mapa y poner en buscar contenido y agregarlo.



**Figura 3.7 Agregar capa a mapa base. Fuente: Autoría propia.**

Esto se debe realizar por cada una de los feature clases que contenga cada ciudad intermedia hasta tener el mapa completo.

### 3.4 Generación de Mapas Temáticos por Ciudad.

La creación de la temática por ciudad es necesario utilizarla herramienta *history map* que cuenta ArcGIS Online.



Figura 3.8 Galería de ciudades Intermedias ArcGIS Online. Fuente: Autoría propia.

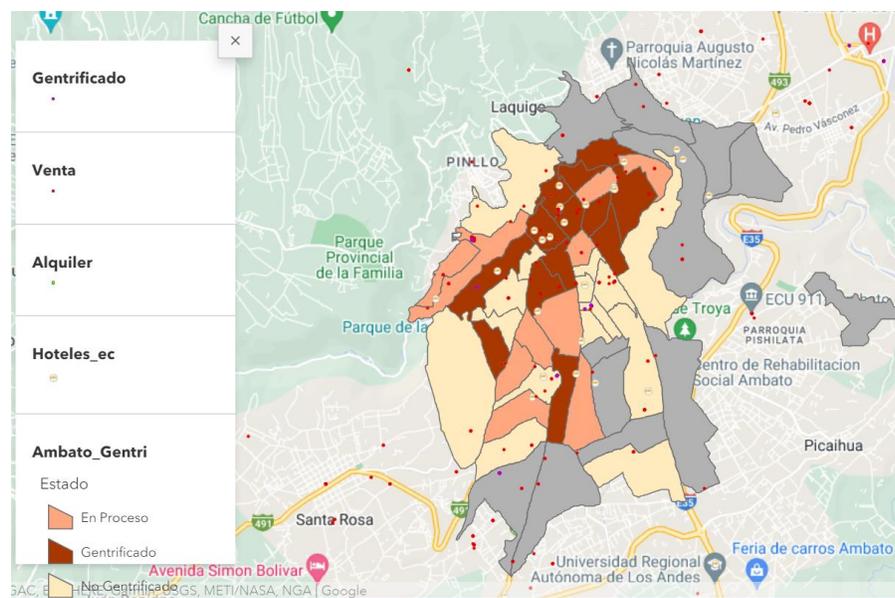
Dentro de cada una de las ciudades se puede visualizar el fenómeno de la gentrificación en base en a las variables seleccionadas socioeconómicas del censo 2010 respecto a hogar y los datos de hoteles y plusvalía de casas por medio del Web scraping creando para ello una tabla que crea un ranking sobre el estado de la gentrificación en un sector de un de las ciudades Intermedias. En la tabla 3.1 se puede visualizar la forma de escoger el nivel de gentrificación de un sector siendo un porcentaje de + 7 para un nivel de gentrificación alto en el sector este valor se lo obtiene de acuerdo con cada variable donde las variables a sumar son: m2 por arriba del promedio, hoteles cinco estrellas, hospitales, centros comerciales, cajeros automáticos, porcentaje por arriba de la media en personas con nivel universitario, porcentaje arriba de la media con afiliación a IESS. En cambio, las variables que restan son tener un nivel alto de personas por habitación, nivel por arriba de la media de personas jubiladas, coste de m2 en un cuartil bajo y variables socioeconómicas por

un nivel bajo de cuartil. De acuerdo con cada variable se irán sumando o restando en una columna independiente que indique los niveles.

**Tabla 3.1 Niveles de Gentrificación en una zona.**

Tipo de Gentrificación en una zona	
Puntuación	Nivel
+7	Extremadamente Gentrificado
+5	Gentrificado
+3 a +5	En proceso
≤ +2	No Gentrificado

Una vez generado el nivel de cada zona se interpola en la aplicación de ArcGIS Online donde se visualiza finalmente dicha zona con puntos críticos donde existe abundancia de locales y hoteles que ayudan a la zona a tener un estatus elevada generando gentrificación en la zona.

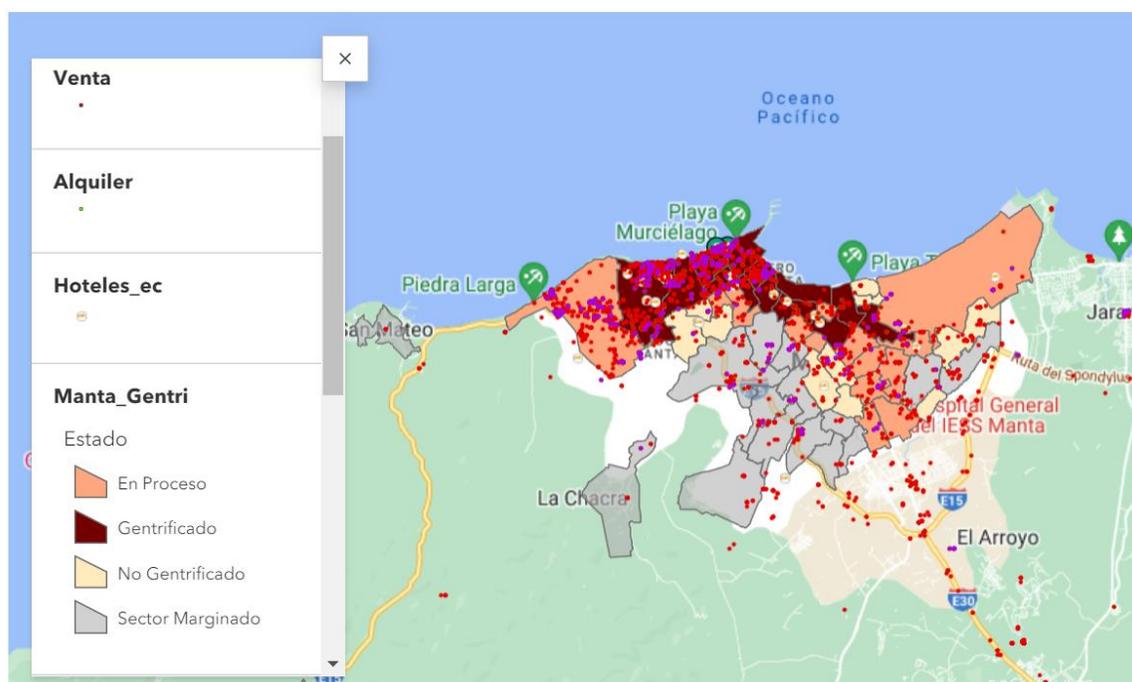


**Figura 3.9 Resultados de zonas gentrificadas para la ciudad de Ambato. Fuente: Autoría propia.**

### 3.5 Análisis de Resultados

Después de haber realizado todas las pruebas para comprobar el correcto funcionamiento del análisis de datos utilizando ArcGIS se obtuvieron los siguientes resultados:

En algunas ciudades de la costa se centran en la turistificación que es una variante de la gentrificación dado el caso de la ciudad de Manta donde se observa que el sector costero está plenamente gentrificado y asociado a variables altas de hoteles 5 estrellas y costos de plusvalías elevados en esa zona. Para el resto de la ciudad no se ve cambios a desde el último censo 2010 concentrando todo en la parte costera como se puede observar en la figura 3.10

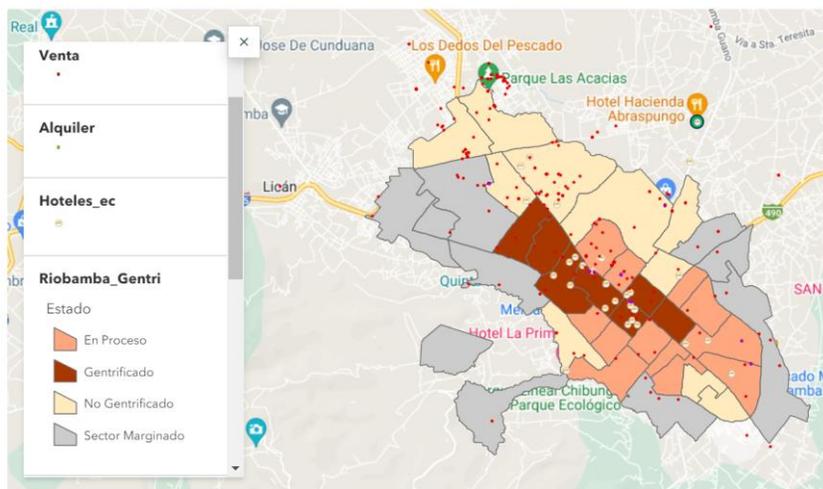


**Figura 3.10 Turistificación en la ciudad de Manta. Fuente: Autoría propia.**

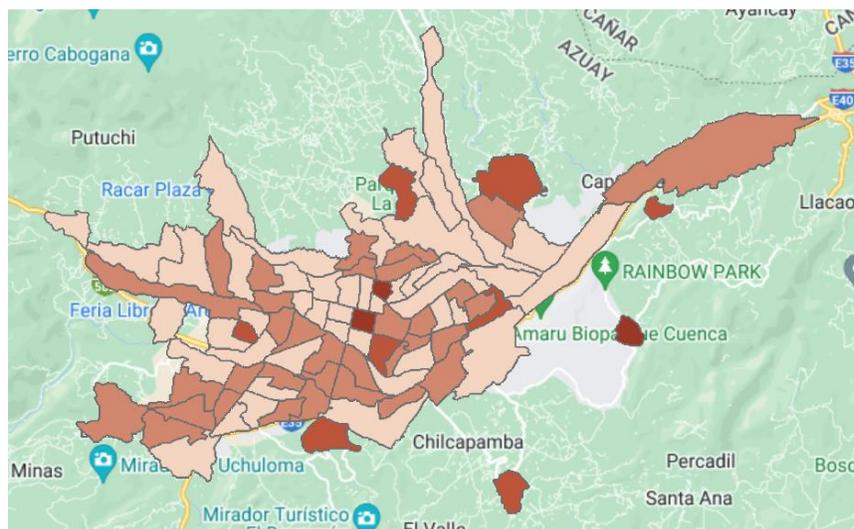
Cabe recalcar que en los sectores marginados ocupa una gran parte en las afueras de la ciudad de Manta y se lo clasifica de esa manera es debido a la delincuencia que afecta la ciudad en la actualidad siendo protegidos y resguardados los sectores conocidos como turísticos que se encuentran marcado en el mapa de color marrón.

En ciudades de la sierra como Riobamba, Cuenca y Loja se observa un alto nivel de gentrificación en los cascos coloniales debido a la ubicación de locales comerciales y hoteles de lujos que dan lugar al alza de plusvalía debido a la concurrencia de turistas que ha provocado la creación de centros comerciales y restaurantes cinco estrellas.

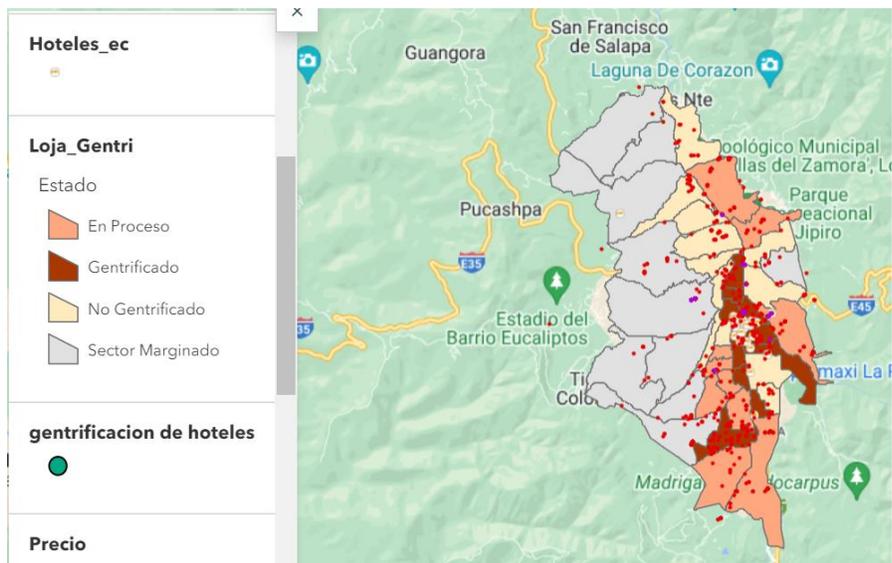
## Riobamba



## Cuenca



## Loja



**Figura 3.11 Gentrificación en ciudades de la Sierra. Fuente: Autoría propia.**

De este modo, las zonas con mayor valor del suelo concentran actividades de comercio y servicios, y mayor accesibilidad a infraestructura de transporte público masivo tienen mayor probabilidad de lograr una transformación sociodemográfica generando gentrificación entre sus habitantes. En cambio, en sectores aledaños de las ciudades tiene mayores distancias y se asocian con menores probabilidades de transformación a menos que se trate de un sector con turistificación donde se tenga algún complejo turístico, zoológico o playa que aporte valor al metro cuadrado en la zona.

### 3.6 Discusión

En la mayoría de los algoritmos, se utilizaron con el mismo propósito el cual fue la identificación de zonas con un alto índice de gentrificación que involucra el precio de arriendo o de ventas de un inmueble por arriba de la media. Para observar, dicho comportamiento se utilizó los datos extraídos del Web scraping junto con los valores de variables socioeconómicas extraídas del censo 2010 para clasificar en base a una tabla de valores que permita el cálculo del índice de gentrificación de una zona en específico de una ciudad intermedia.

El modelo seleccionado predice a complacencia el estado de gentrificación en una época actual pues es reforzado por el método de Nesbitt que involucra variables socioeconómicas de censos con variables extraídas de páginas web del presente año. En base a esto se abre un abanico de nuevos estudios donde se involucren algún tipo de variable numérica en específico que no se encuentre en un censo o sea mucho más actual y se la pueda extrapolar. Para así, ajustar dichos datos a algún tipo de negocio tal como lo hacen muchas plataformas de pago en la actualidad que con la data que obtienen de sus suscriptores para saber el nivel de comercio por zonas o la rentabilidad de venta de una casa.

Es importante tener de base la codificación que utiliza el INEC para las zonas en diferentes ciudades del país, para vincularlas con los datos que mostrarán en base a coropletas el nivel de saturación en unidades de variación espacial siendo más fácil la toma de decisiones.

De esta manera se pudo comprobar que se pueden crear Geodatabase de una forma segura y eficiente para representar algún fenómeno en específico tanto a nivel de variable o multivariable. Por ejemplo, se utiliza en la actualidad para representar el nivel de contagio o vacunación contra la COVID 19.

### **3.7 Diagnósticos de la problemática para gobiernos autónomos.**

A nivel de ciudad la gentrificación trae consigo desplazamiento hacia sectores aledaños, debido al encarecimiento del nivel de vida en el sector comúnmente en los cascos comerciales de la ciudad. Esto ocasiona la pérdida de control sobre el precio del metro cuadrados por parte del municipio, razón por la cual siempre las viviendas son compradas a un precio más elevado que el catastro municipal cosa que hoy en día se lo ve muy a menudo. Esta ganancia es aprovechada por las personas que ocasionan la gentrificación en el sector y son muy pocos los que obtienen el beneficio de esto. Para ello deberían hacer un control de la plusvalía de vivienda estableciendo leyes que beneficien a todos en general fomentando el buen vivir, pues el pueblo es el que paga las consecuencias debido a que aumenta el tiempo de traslado y la inversión por parte de gobiernos es cada vez más fuerte al crear nuevas líneas de buses,

sistemas de alcantarillado y eléctricos. Para mejorar el reconocimiento de un área en específico se presentó un diagrama de puntos representados (representación espacial) por colores que permite conocer la agrupación de variables en base a puntos de gentrificación, hoteles y cercanía a centros comerciales. Pues la gentrificación se va desarrollando de manera circular cuadra a cuadra de los puntos gentrificados, esto permitirá conocer y predecir cómo avanza la gentrificación por medio de latitud y longitud. En la figura 3.12 se puede visualizar dichos puntos desplegados en algunas ciudades intermedias.

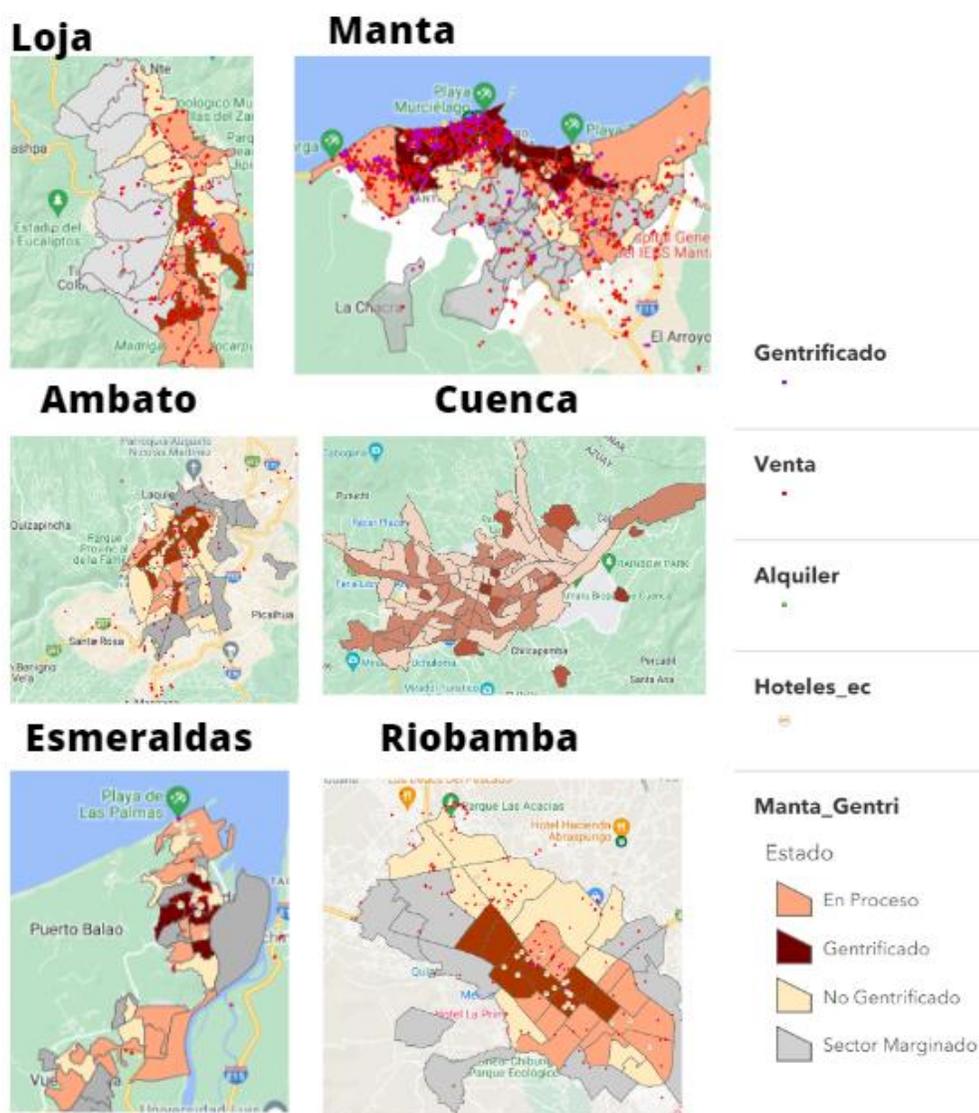


Figura 3.12 Despliegue de puntos de interés en zonas de gentrificación de una ciudad.

En varios países ya se tiene varios controles que frenan los desplazamientos por parte de la gentrificación, llamados “linkage fees” en inglés, son cargos que pagan los urbanizadores que vinculan ciertos beneficios para la ciudad como creación de áreas verdes, construcción de carreteras y hasta beneficios a personas con bajos recursos provocando que se planee bien la construcción de un barrio. Es indudable que, dentro de los poderes del Estado, con la voluntad política suficiente, sería posible realizar una renovación urbana y una revitalización de los barrios históricos de las grandes ciudades sin provocar desplazamiento ni aburguesamiento residencial.

## CAPÍTULO IV

### 4. Conclusiones y Recomendaciones

#### 4.1 Conclusiones

- La gentrificación en Ecuador ha ido tomando fuerza en los últimos 20 años debido a la llegada de tiendas internacionales y crecimiento de la industria inmobiliaria que busca beneficio propio y principalmente por la falta de control de las autoridades de turno.
- Para el año 2023 se prevé que la gentrificación crezca debido a que el sector inmobiliario se está recuperando de la reducción de precios de ventas de casas cosa que se puede visualizar y ponderación en plataformas de venta de casa como plusvalía.com, OLX y Airbnb que no han cambiado desde el año 2019 en donde empezó la pandemia ocasionada por el COVID 19. Para ello, se recomienda que los gobiernos centrales prevean la especulación del valor del m<sup>2</sup> en ciertos sectores de la ciudad.
- Los modelos de gentrificación han ido adoptando cada vez más variables de tipo socioeconómicas, pues el modelo de Hammel es muy bueno pero está basado en censos y estos se dan con un preciosidad regular y un lapso de 10 años principalmente en países Latinoamericanos como Ecuador para ello modelos como Chapple, Waddell, Chatman, Loukaitou-Siders y Ong (2017) proponen índices de precios, cercanía a cajeros y bancos y algunas variables que no son abiertas como accesos a seguros privados datos que no son abiertos en Ecuador razón por la cual se escogió el modelo de Nesbitt que involucra el precio promedio de m<sup>2</sup>, hoteles de lujo y centros comerciales cosas que no son abiertos pero se los extrajo a manera de investigación mediante la técnica de Web scraping para ser unidos con datos del último censo 2010.
- La herramienta GIS facilitó la unión de datos censales con datos extraídos de plataformas web de ventas de casas gracias a su herramienta Inner Join que permitió a datos espaciales (coordenadas de latitud y longitud) unirse con zonas

codificadas establecidas por el INEC (código de ciudad junto con zona) que sirvió para interpolar datos en una super tabla y poder aplicar analítica de datos con Python que viene preinstalada en ArcGIS para poder predecir las diferentes zonas de acuerdo al índice de gentrificación. Esto se realizó por cada una de las ciudades intermedias dentro de Ecuador.

- Se priorizó el uso de variables geoestadísticas básicas como porcentaje de personas con nivel universitario, número de personas por cuarto, porcentaje de personas con seguro social junto con datos extraídos de las zonas como hoteles con estrellas, valor de m<sup>2</sup> y cercanía centros comerciales para crear un ranking en base a pesos que midió el nivel de gentrificación en una zona que contenga todas las variables respecto a otras.
- Finalmente se puede decir que la gentrificación en Ecuador no es un fenómeno tan conocido por la mayoría de sus habitantes permitiendo a ciertos sectores aprovecharse con la especulación de valor de metro cuadrado ocasionando desplazamiento y creación de nuevos barrios en las afueras de las ciudades creando una brecha distante año a año. Por lo cual se debería crear políticas de buen vivir para gozar de los mismos privilegios sin desigualdades de clases.

## **4.2 Recomendaciones**

- Se pudo evidenciar la gentrificación en varias ciudades Intermedias de Ecuador a través de zonas, pero se podría profundizar el análisis en zonas gentrificadas para extraerlas por medio de bloques o manzanas cosa que no se pudo hacer por cuestiones de tiempo y alcances de la investigación dando libertad a otros autores a profundizar en dicho estudio.
- Se debería probar en otras herramientas GIS para ver si se varían los resultados y predicción de zonas en base a la plusvalía de casas y variables socioeconómicas extraídas de los censos.
- Se podría utilizar los datos del próximo censo venidero en noviembre del 2022 para ver si se empatan con las predicciones extraídas mediante la técnica de Web scraping de plataformas de ventas de casas.

- Con las tablas extraídas de sitios web se pueden utilizar para el análisis de negocios y proyecciones de mercado siempre y cuando sea para fines didácticos y no seguir gentrificando en la zona.

## REFERENCIA BIBLIOGRÁFICAS

- [1] L. Loretta, T. Slater y E. Wyly, *Gentrification*, Inglaterra: Taylor and Francis, 2013.
- [2] ARQUITASA, «GENTRIFICACIÓN: PROCESO, CAUSAS Y EFECTOS,» GENTRIFICACIÓN: PROCESO, CAUSAS Y EFECTOS, 15 Julio 2021. [En línea]. Available: <https://arquitasa.com/la-gentrificacion-o-conquistar-barrios-sin-derramar-sangre/>. [Último acceso: 20 Septiembre 2022].
- [3] J. & P. J. L. Rius-Ulldemolins, «ultura, transformación urbana y empoderamiento ciudadano frente a la gentrificación,» *Rius-Ulldemolins*, vol. 42, nº 126, pp. 131-154, 2016.
- [4] E. López, «Gentrificación en Chile: aportes conceptuales y evidencias para una discusión necesaria,» *Scielo*, vol. II, nº 56, pp. 220-256, 2013.
- [5] J. Sequera, «Las políticas de la gentrificación en la ciudad neoliberal nuevas clases medias, producción cultural y gestión del espacio público,» 20 febrero 2013. [En línea]. Available: <https://dialnet.unirioja.es/servlet/tesis?codigo=167038>. [Último acceso: 12 Enero 2022].
- [6] C. Antoine, «Procesos de gentrificación y desplazamiento en América Latina, una perspectiva comparativista,» *Urban*, vol. 1516, nº 160, pp. 200-215, 02 Abril 2013.
- [7] M. Carman, «El proceso de ennoblecimiento y la salida negociada de los innobles en Buenos Aires,» *Metrópole*, vol. 33, nº 25, pp. 257-278, 2011.
- [8] B. Orellana y T. Calle, «Analysis of the Gentrification Phenomenon Using GIS to Support Local Government Decision Making,» *Springer*, vol. 1213, pp. 348-354, 01 Septiembre 2020.
- [9] V. Adler, «Temas que deben cuidar las ciudades para mejorar su calidad de vida,» *Ciudades Sostenibles*, 30 Octubre 2017. [En línea]. Available: <https://blogs.iadb.org/ciudades-sostenibles/es/8-temas-que-deben-cuidar-las-ciudades-para-mejorar-su-calidad-de-vida/>. [Último acceso: 10 Agosto 2021].
- [10] V. DELGADILLO, «Repoblamiento y recuperación del centro histórico de la ciudad de México, una acción pública híbrida,» *Economía, Sociedad y Territorio*, vol. 8, nº 28, pp. 817-845, 2008.
- [11] D. Castro, «El proceso de Gentrificación, intervención urbana arquitectónica en la ciudad de Salinas-Ecuador,» *Universidad de Oriente*, vol. IV, nº 1, pp. 32-44, 2020.

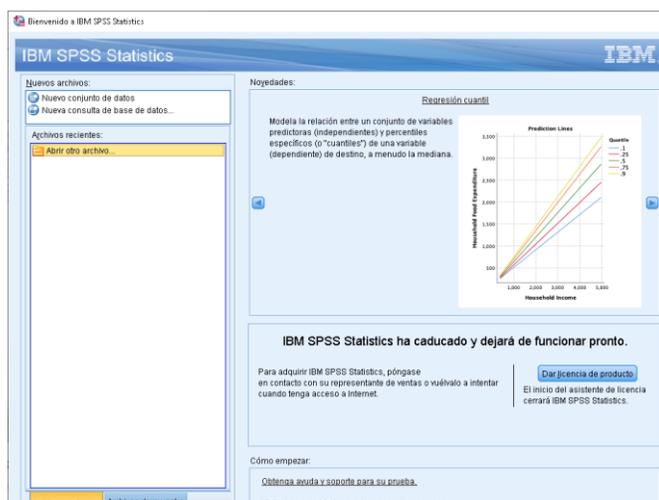
- [12] A. Casgrain, «Gentrificación y resistencia en las ciudades latinoamericanas,» El ejemplo de Santiago de Chile, 12 Agosto 2013. [En línea]. Available: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1870-00632013000200003](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-00632013000200003). [Último acceso: 2021 Agosto 2021].
- [13] P. Zambrano, «Desarrollo de una aplicación web para la visualización y análisis del fenómeno de gentrificación en el Centro Histórico Quito como soporte a la toma de decisiones en el municipio del Distrito Metropolitano de Quito (MDMQ) entre el período de 1990 al 2010,» EPN, Quito, 2020.
- [14] I. Díaz, «La Gentrificación en la cambiante estructura socioespacial de la ciudad,» *INVI*, vol. 88, pp. 27-71, 16 Julio 2016.
- [15] P. Zandbergen, «Introducción a Python en ArcGIS Pro,» learnargis, 17 Febrero 2019. [En línea]. Available: <https://learn.arcgis.com/es/projects/get-started-with-python-in-arcgis-pro/>. [Último acceso: 17 Agosto 2021].
- [16] D. Smith, «Forecasting gentrification in city neighborhoods, with R,» *Revolutions*, 28 Febrero 2017. [En línea]. Available: <https://blog.revolutionanalytics.com/2017/02/forecasting-gentrification.html>. [Último acceso: 17 Agosto 2021].
- [17] J. Serrano-Cobos, «Big data y analítica web. Estudiar las corrientes y pescar en un océano de datos,» *Profesional de la Información*, vol. v. 23, nº 6, pp. 561-565, 2014.
- [18] E. Alpaydin, *Machine Learning*, massachusetts: Cambridge, MIT, 2021.
- [19] D. Calvo, «Definición de aprendizaje supervisado,» [www.diegocalvo.es](http://www.diegocalvo.es), 23 Marzo 2019. [En línea]. Available: <https://www.diegocalvo.es/aprendizaje-supervisado/>. [Último acceso: 02 febrero 2022].
- [20] N. Guerra, *Minería de datos en ofertas laborales*, México: Groppe, 2020.
- [21] F. Polidoro, «Web scraping techniques to collect data on consumer electronics and airfares for Italian HICP compilation,» *IOS PRESS*, vol. 31, nº 2, pp. 165-176, 2015.
- [22] ESRI, *The ArcGIS Book – PDF en Español*, Redlands, California: Esri Press, 2019.
- [23] J. M. Llop y L. V. Cruz, *EL DERECHO A LA CIUDAD EN EL CONTEXTO DE LA AGENDA URBANA PARA CIUDADES INTERMEDIAS*, Cuenca: Universidad de Cuenca, 2017.
- [24] G. Durán, «Crecimiento, segregación y mecanismos de desplazamiento en el periurbano de Quito,» *Revista de Ciencias Sociales*, vol. II, nº 56, pp. 123-146, 2016.

- [25] A. V. Murillo y Y. Ayo, «Gentrificación en Guayaquil: ¿Desde cuándo ‘cambió’ la calle Panamá en la ciudad?,» *UArtes*, 09 Diciembre 2021. [En línea]. Available: <http://ilia.uartes.edu.ec/blog-f-ilia/2021/12/09/gentrificacion-en-guayaquil-desde-cuando-cambio-la-calle-panama-en-la-ciudad/>. [Último acceso: 22 Febrero 2022].
- [26] G. STEEL, «Displacement by/for development in two Andean cities,» *Congress of the Latin American Studies Association*, vol. IV, nº 1, pp. 60-80, 2015.
- [27] J. Palma, T. Calle y B. Orellana, «GIS for Decision-Making on the Gentrification Phenomenon in Strategic Territorial Planning,» *Conference on Information and Communication Technologies of Ecuador*, vol. I, nº 9, pp. 142-156, 2021.
- [28] F. Novoa, «SlideShare,» Grupo Academia Postal, 16 Abril 2014. [En línea]. Available: <https://es.slideshare.net/FranciscoNvoaManuel/ccna-routing-switching-novedades-enrutamiento-ospf-multirea>. [Último acceso: 2017 Julio 5].
- [29] A. Dávila, «UNIVERSIDAD NACIONAL DE CAJAMARCA,» 20 Noviembre 2017. [En línea]. Available: [https://repositorio.unc.edu.pe/bitstream/handle/UNC/1717/T016\\_44166738\\_T.pdf.pdf?sequence=1&isAllowed=y](https://repositorio.unc.edu.pe/bitstream/handle/UNC/1717/T016_44166738_T.pdf.pdf?sequence=1&isAllowed=y). [Último acceso: 25 febrero 2022].
- [30] A. Morales, «mappinggis.com,» 12 Abril 2021. [En línea]. Available: <https://gis.stackexchange.com/questions/13509/comparing-qgis-and-gvsig>. [Último acceso: 28 Febrero 2022].
- [31] I. Díaz Parra y P. Rabasco Pozuelo, «¿Revitalización sin gentrificación? Cooperativas de vivienda por ayuda mutua en los centros de Buenos Aires y Montevideo,» *Cuadernos Geográficos*, vol. 52, nº 2, pp. 99-118, 2020.

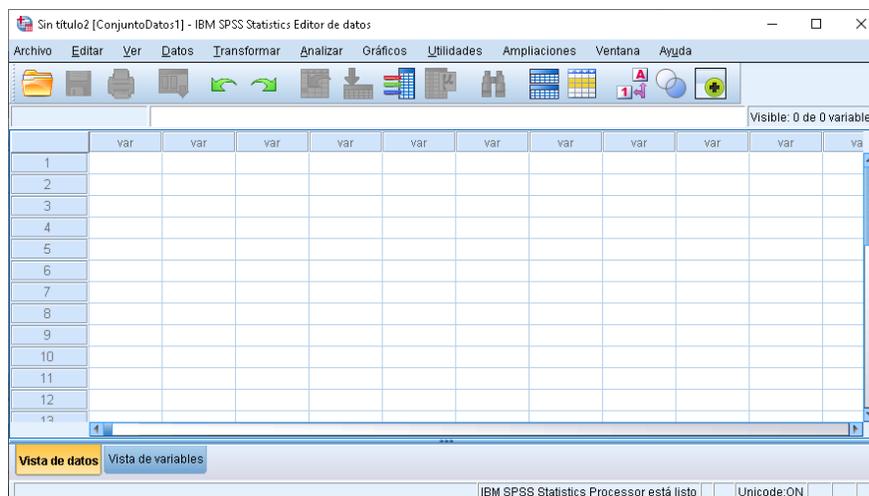
## ANEXOS

### ANEXO I. Guía para extracción de datos censales por ciudades en SPSS.

Para los datos censales del año 2010 se pudo trabajar con la plataforma Redatam sin ningún inconveniente, pero el programa predefinido para el análisis censal en muchos países suele ser SPSS. De hecho, Ecuador cuenta con dicho formato para el año 2010 al igual que censos de años pasados. Para la data registrada en el año 2001, se trabajó con la herramienta SPSS, ya que es el único formato que se contaba para dicho año.



Se ejecuta un nuevo conjunto de datos



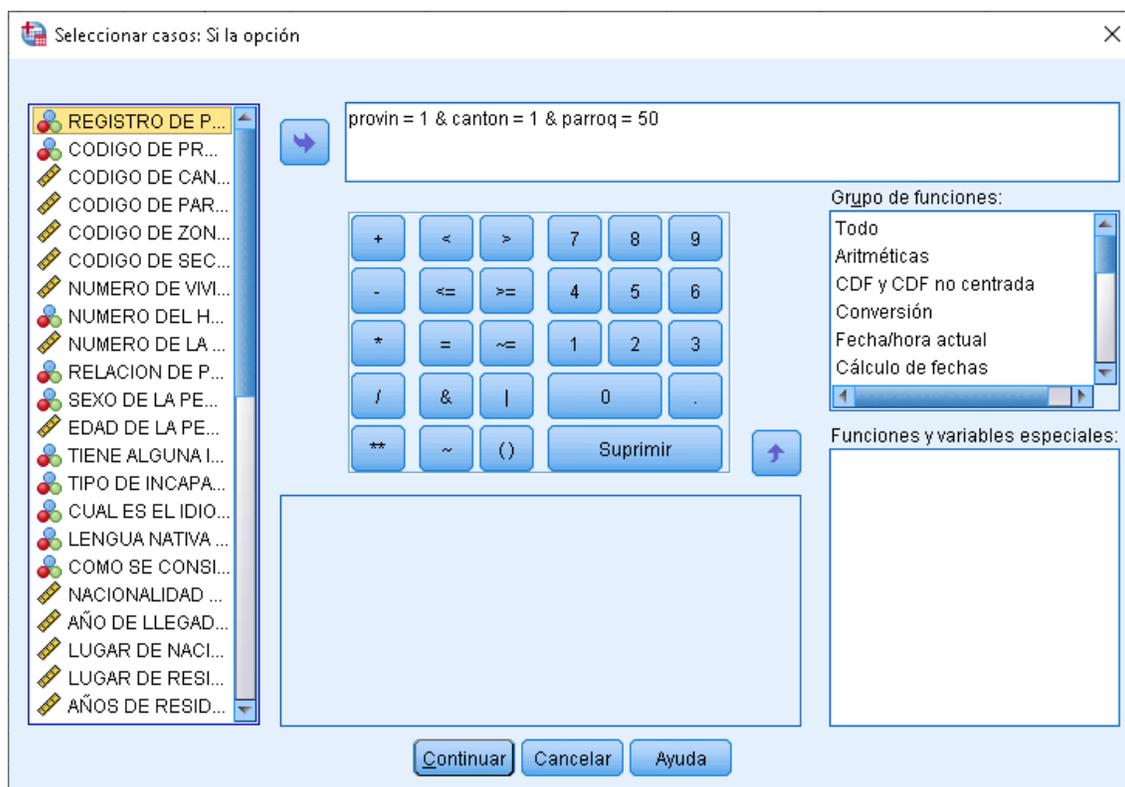
Luego se cargó las variables del censo 2001.

Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rel
1 tipreg	Numérico	1	0	REGISTRO DE...	Ninguna	Ninguna	6	Derecha	Nominal	Entrada
2 provin	Numérico	2	0	CODIGO DE P... (1, Cuenca)...	Ninguna	Ninguna	6	Derecha	Nominal	Entrada
3 canton	Numérico	2	0	CODIGO DE C...	Ninguna	Ninguna	6	Derecha	Escala	Entrada
4 parroq	Numérico	2	0	CODIGO DE P...	Ninguna	Ninguna	6	Derecha	Escala	Entrada
5 zona	Numérico	3	0	CODIGO DE Z...	Ninguna	Ninguna	9	Derecha	Escala	Entrada
6 sector	Numérico	3	0	CODIGO DE S...	Ninguna	Ninguna	6	Derecha	Escala	Entrada
7 nummw	Numérico	3	0	NUMERO DE ...	Ninguna	Ninguna	6	Derecha	Escala	Entrada
8 numhog	Numérico	1	0	NUMERO DEL ...	Ninguna	Ninguna	6	Derecha	Nominal	Entrada
9 ordpob	Numérico	4	0	NUMERO DE L...	Ninguna	Ninguna	6	Derecha	Escala	Entrada
10 parent	Numérico	1	0	RELACION DE ... (0, Jefe)...	Ninguna	Ninguna	6	Derecha	Nominal	Entrada
11 sexo	Numérico	1	0	SEXO DE LA P... (1, Hombre)...	Ninguna	Ninguna	4	Derecha	Nominal	Entrada
12 edad	Numérico	2	0	EDAD DE LA P...	Ninguna	Ninguna	4	Derecha	Escala	Entrada
13 discap1	Numérico	1	0	TIENE ALGUN... (1, Si)...	Ninguna	Ninguna	7	Derecha	Nominal	Entrada
14 discap2	Numérico	1	0	TIPO DE INCA... (1, Para ver)...	Ninguna	Ninguna	7	Derecha	Nominal	Entrada
15 idioma	Numérico	1	0	CUAL ES EL I... (1, Solo esp...)...	Ninguna	Ninguna	6	Derecha	Nominal	Entrada
16 qidioma	Numérico	2	0	LENGUA NATIVA	Ninguna	Ninguna	7	Derecha	Nominal	Entrada
17 raza	Numérico	1	0	COMO SE CO... (1, Indigena)...	Ninguna	Ninguna	4	Derecha	Nominal	Entrada
18 grazza	Numérico	2	0	NACIONALIDA...	Ninguna	Ninguna	5	Derecha	Escala	Entrada
19 anollag	Numérico	4	0	AÑO DE LLEG...	Ninguna	Ninguna	7	Derecha	Escala	Entrada
20 lughac	Numérico	6	0	LUGAR DE NA... (10150, Cue...)...	Ninguna	Ninguna	6	Derecha	Escala	Entrada
21 lughab	Numérico	6	0	LUGAR DE RE... (10150, Cue...)...	Ninguna	Ninguna	6	Derecha	Escala	Entrada
22 aniores	Numérico	2	0	AÑOS DE RES...	Ninguna	Ninguna	7	Derecha	Escala	Entrada
23 mesres	Numérico	2	0	MESES DE RE...	Ninguna	Ninguna	6	Derecha	Nominal	Entrada
24 lugin	Numérico	6	0	DONDE VIVA ... (10150, Cue...)...	Ninguna	Ninguna	6	Derecha	Escala	Entrada

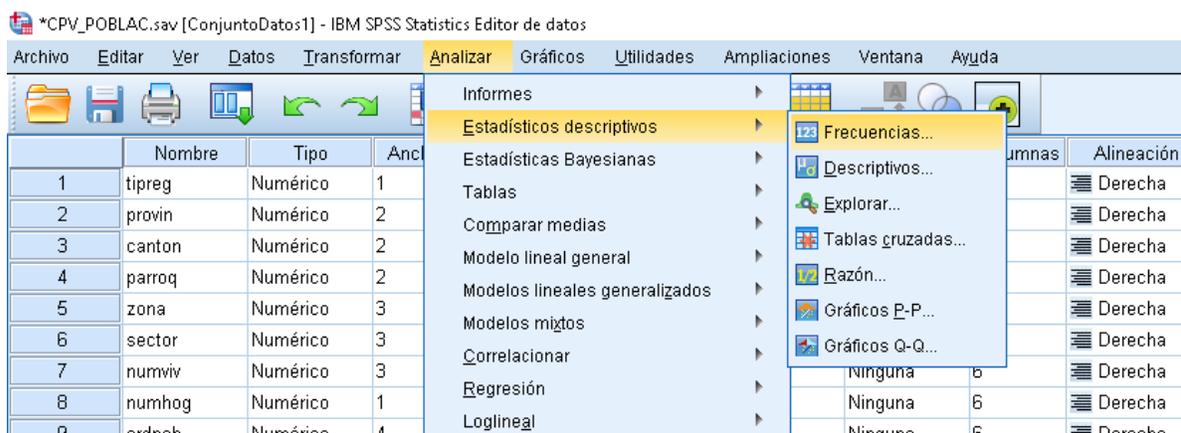
Para realizar filtros para cada ciudad intermedia es necesario ir a datos selección de casos. Por ejemplo: para la ciudad de Cuenca se debe escoger Azuay= 1, cuenca=1 y parroquia=50. La información y codificación por provincia se encuentra en el siguiente enlace:

<http://web.educacion.gob.ec/CNIE/pdf/Anexo%20con%20Codificacion.pdf>

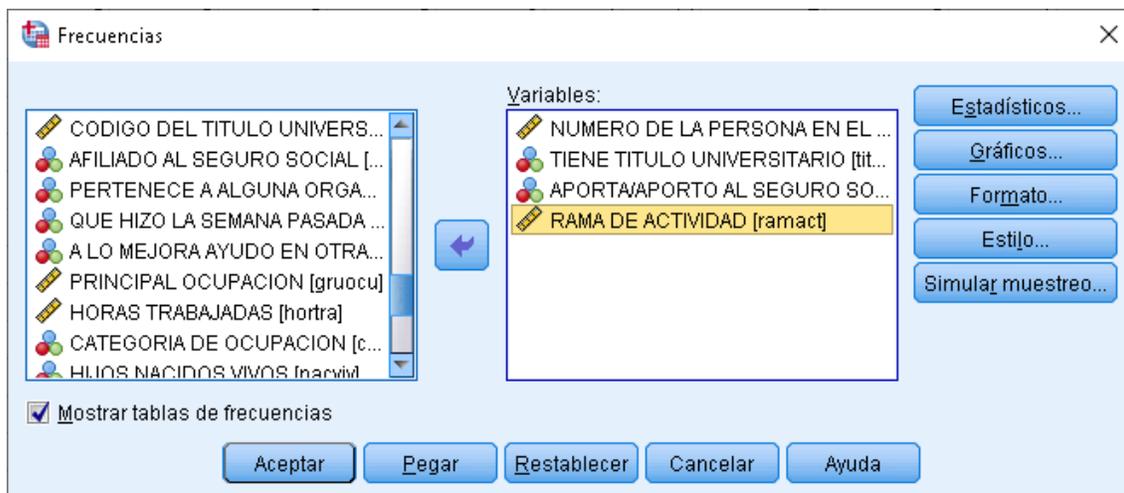
Para filtrar los datos es necesario ir a Analizar Datos y seleccionar Filtrado por casos.



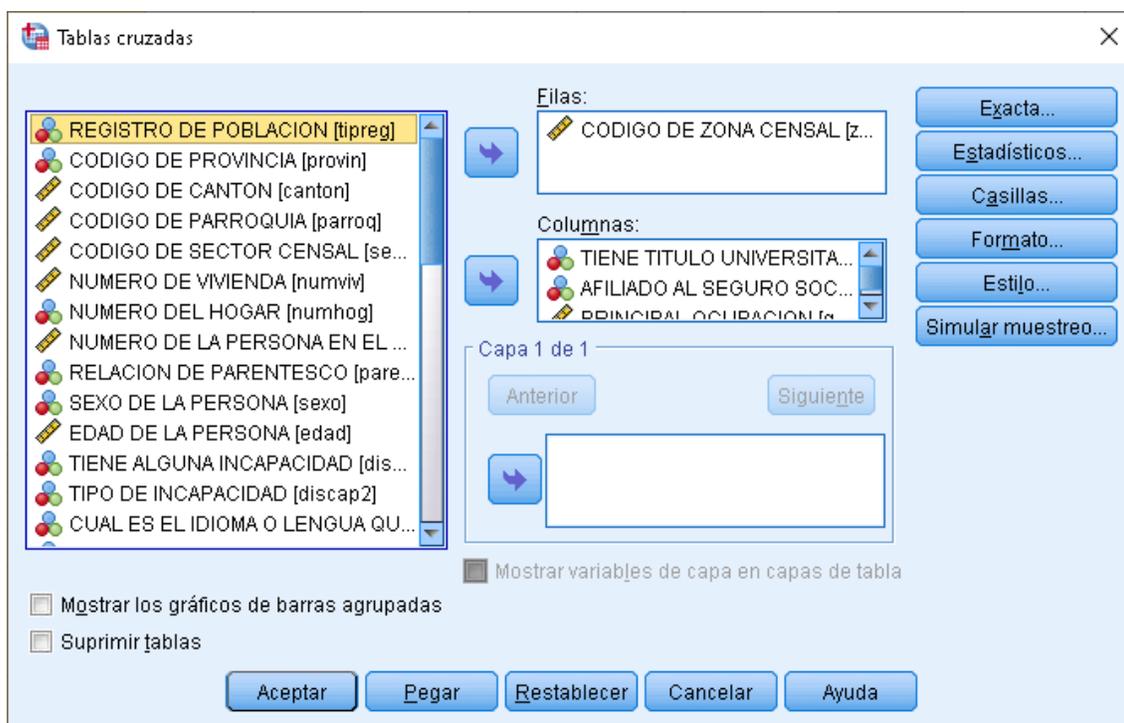
Se puede ejecutar un análisis estadístico aplicando el filtro de la ciudad Intermedia.



Dentro de la tabla de frecuencias es necesario aplicar el filtro junto con el dato de todas las zonas censales de la ciudad Intermedia.



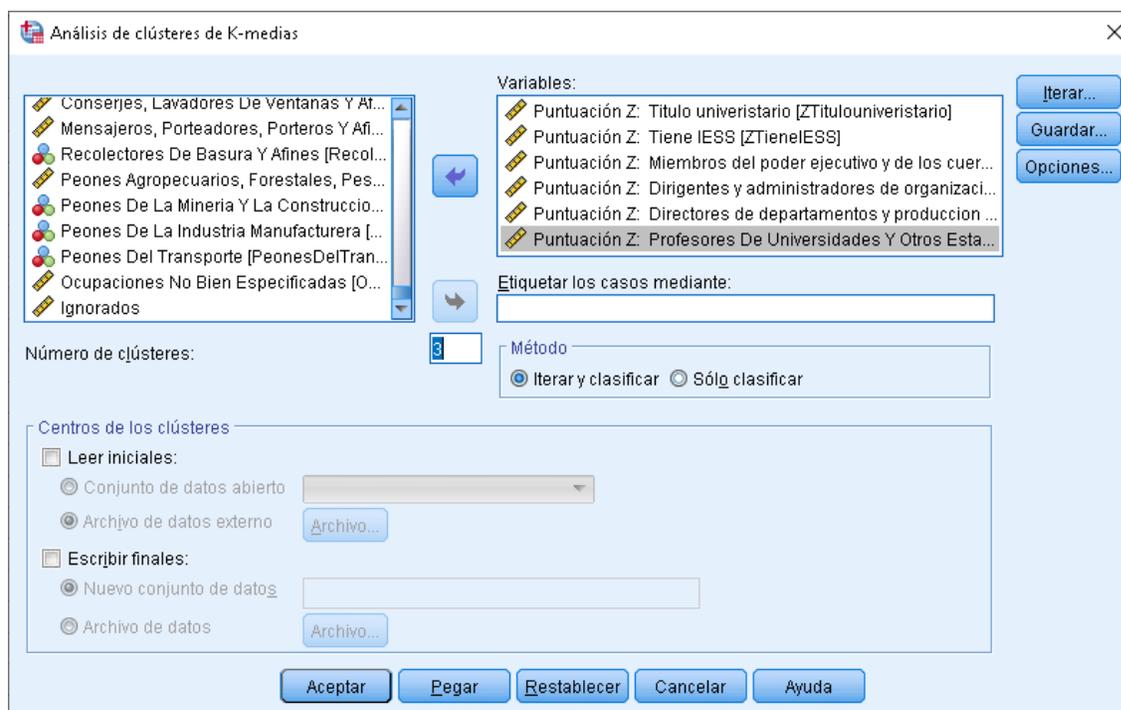
Para obtener la respectiva tabla referente a zonas que abarcan una ciudad del año 2001 se realizó una tabla cruzada. Antes de realizar debe estar activo el filtro de acuerdo a la ciudad.



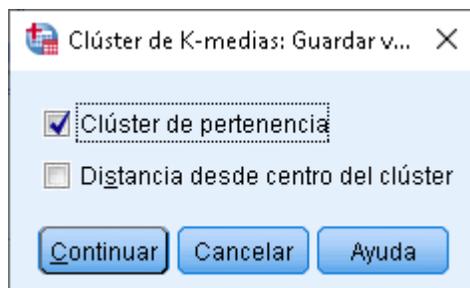
Se debe normaliza los datos antes de clasificarlos y agruparlos. Para ello, es necesario aplicar descriptores con las variables aplicar.



Elegir la z cobers para parametrizar cada una de las variables



Finalmente se aplica al clúster de pertinencia que permita mostrar en una columna de referencia el número de clúster al que pertenece.



## **ANEXO II. Despliegue de Dashboard en ArcGIS Online**

La publicación de las zonas gentrificadas para cada ciudad intermedia se la realizó por medio de la herramienta ArcGIS online que se presenta en el siguiente enlace:

<https://arcg.is/1v8bCa>