**Introducción/Resumen**

Este documento está diseñado para ayudarle a comprender cómo puede usar una instancia de Azure Database for PostgreSQL para habilitar el escenario empresarial de selección de sitios. Para ello, nos basaremos en las sólidas capacidades geoespaciales dentro de PostgreSQL.

La selección de sitios es una función importante en una variedad de industrias, incluyendo el comercio minorista, las telecomunicaciones y el transporte/logística. En este ejemplo,ejecutaremos el proceso de selección del sitio desde el punto de vista de una empresa de scooters. La compañía de scooters debe decidir dónde colocar su primera instalación de carga de scooters en una ciudad para asegurarse de que los scooters están cargados y listos para ir cada mañana.

Una vez que la compañía de scooters despliega sus scooters, pueden utilizar los datos de ubicación de los scooters para tomar decisiones sobre las ubicaciones posteriores del sitio de carga. Pero fo la selección inicial del sitio, se basarán en información sobre lo siguiente:

* Población (más alta es mejor)
* Distancia desde las estaciones de metro (más alto es mejor)
* Tiempos promedio de viaje (más alto es mejor)
* Tasa de criminalidad (más baja es mejor)

En resumen, la compañía de scooters quiere dar servicio a un mayor número de clientes potenciales con viajes más largos más lejos de las estaciones de metro en áreas con menorcrimen. Les gustaría ubicar sus estaciones de carga en las esquinas/intersecciones no asociadas con las autopistas para que puedan tomar scooters desde dos calles en lugar de una.

Usaremos la ciudad de Nueva York (NYC) para este ejemplo porque tenemos fácil acceso a todos los datos necesarios para ilustrar este escenario para NYC. Tenga en cuenta que este tutorial está fuertemente inspirado y adaptado de los materiales que se encuentran en el [Taller introductorio PostGIS](https://postgis.net/workshops/postgis-intro/) que se han publicado bajo una [licencia](https://creativecommons.org/licenses/by-sa/3.0/us/) [de Atribución-CompartirIgual 3.0 Estados Unidos (CC BY-SA 3.0 US).](https://creativecommons.org/licenses/by-sa/3.0/us/)

**Herramientas y software requeridos**

Para ejecutar este tutorial, necesitará lo siguiente:

* Una suscripción de Azure (opción gratuita [aquí)](https://azure.microsoft.com/free/free-account-faq/) y una instancia de Azure Database for [PostgreSQL](https://ms.portal.azure.com/#create/hub) (tenga en cuenta que puede usar un único nivel de servicio de servidor y basic con 1 o 2 núcleos virtuales y 50 GB de almacenamiento: esta configuración consumirá aproximadamente 2 USD o menos al día a partir de ya sea su suscripción de prueba de pago o gratuita si la deja funcionando las 24 horas del día).
* PostGIS: las herramientas cliente incluyen El archivo de forma PostGIS 2.0 y DBF Loader Exporter que utilizaremos para importar datos geoespaciales relacionados con NYC a nuestra instancia de PostgreSQL
* [QGIS](https://qgis.org/en/site/forusers/download.html) – un excelente conjunto de herramientas para visualizar datos geoespaciales de PostgreSQL y otras bases de datos
* [pgAdmin – una sólida herramienta de](https://www.pgadmin.org/download/) administración para PostgreSQL

**Creación y población de su base de datos de NYC**

Una vez creada la instancia de Azure Database for PostgreSQL, deberá crear una base de datos PostgreSQL en esa instancia. Use pgAdmin para conectarse a la instancia de Azure Database for PostgreSQL mediante la cadena de conexión de Azure y el nombre de usuario y contraseña que especificó durante la instalación de la instancia. Una vez que pgAdmin esté conectado a la instancia de Azure Database for PostgreSQL, haga clic con el botón derecho en el nombre de la instancia, seleccione Crear > Base de datos... y cree una base de datos denominada nyc utilizando los valores predeterminados.

Una vez creada la basede datos nyc, a continuación utilizará pgAdmin para asegurarse de que su base de datos nyc tiene instalada la extensión postgis. Haga clic con el botón derecho en la basede datos nyc recién creada, seleccione Crear > Extensión... y seleccione postgis en el campo Nombre en la ventana emergente "Crear – Extensión".

Ahora que las capacidades geoespaciales completas están habilitadas, utilizará la aplicación PostGIS 2.0 Shapefile y DBF Loader Exporter para cargar los cuatro elementos de datos geoespaciales de este ejemplo en la base de datos nyc: calles, estaciones de metro, distritos censales y homicidios. Los datos de homicidios son nuestro apoderado de la tasa de criminalidad y usaremos los distritos censales como un área geográfica relativamente pequeña como nuestra granularidad de alto nivel para la selección del sitio. Tenga en cuenta que tendrá que descomprimir estos datos geoespaciales antes de poder importarlos a PostgreSQL. PostGIS tiene una excelente explicación de cómo utilizar postGIS 2.0 Shapefile y DBF Loader Exporter  [aquí](https://postgis.net/workshops/postgis-intro/loading_data.html).

Como vio en el vínculo anterior, deberá proporcionar sus credenciales de inicio de sesión para la instancia de Azure Database for PostgreSQL en PostGIS 2.0 Shapefile y DBF Loader Exporter. Su nombre de usuario y contraseña serán los que especificó al crear la instancia de Azure Database for PostgreSQL. El host del servidor será el punto de conexión público asociado a la instancia de Azure Database for PostgreSQL, por ejemplo, your\_instance\_name.postgres.database.azure.com. El puerto será 5432 y el nombre de la base de datos será nyc.

Una vez que haya especificado sus credenciales de inicio de sesión, tendrá que agregar el archivo . shp de nuestro dataset geoespacial que extrajo del archivo zip que contiene estos datos. Haga clic en el botón "Agregar archivo", agregue los cuatro . shp y configure el importador para utilizar un valor de 26918 para el campo SRID de cada una de las entidades geoespaciales.

Sabrás que lo has hecho correctamente cuando la ventana de tu archivo de forma PostGIS 2.0 y DBF Loader Exporter tenga este aspecto:

Una captura de pantalla de un teléfono celular

Descripción generada automáticamente

Cuando haya configurado esto correctamente, haga clic en el botón "Importar" para importar estas cuatro entidades geoespaciales a su base de datos nyc.

Cada una de estas cuatro entidades/tablas tiene elementos geoespaciales en la columna geom y elementos relacionales "tradicionales" en otras columnas. El carácter geoespacialde cada entidad puede ser un punto (en el caso de estaciones de metro y homicidios), una/línea múltiple s (en el caso de calles), o uno/múltiples polígonos (en el caso de distritos censales).

Los detalles sobre estas entidades/tablas y algunas de las columnas útiles en ellas se describen en el diccionario de datos a continuación.

## Diccionario de datos geoespaciales

Columnas útiles en la tabla "nyc\_streets":

|  |  |
| --- | --- |
| Nalma | Nombre común de la calle |
| Type | ¿Qué clase de calle es ésta? |
| Geom | Geometría MultiLinestring de la calle |

Columnas útiles en la tabla "nyc\_subway\_stations":

|  |  |
| --- | --- |
| Nombre | Nombre común de la estación |
| Geom | Geometría de puntos de la calle |

Columnas útiles en la tabla "nyc\_homicides":

|  |  |
| --- | --- |
| Arma | Arma usada en el homicidio |
| Geom | Geometría puntual del homicidio |

Columnas útiles en la tabla "nyc\_census\_tracts": MVA

|  |  |
| --- | --- |
| id de distrito | Uncódigo n 1 de1dígito que identifica de forma única cada distritocensal. Por ejemplo:36005000100 |
| Geom | Límite MultiPolygon del tracto |

**Enriqueciendo nuestroconjunto** de datos geoespaciales **con datos socioeconómicos**

Nuestros datos geoespaciales ya tienen algunos elementos relacionales, pero necesitamos combinar estos datos con datos relacionales adicionales para lograr nuestro objetivo de negocio.

Para estos datos adicionales, usaremos el archivo denominado nyc\_census\_sociodata.sql para importar datos a nuestra base de datos nyc. Dentro de pgAdmin, haga clic con el botón derecho en la base de datos nyc y seleccione la herramienta de consulta. Pegue el contenido de nuestro archivo nyc\_census\_sociodata.sql en esta nueva ventana de consulta y, a continuación, ejecute la consulta presionando F5 en el teclado o haciendo clic en el pequeño botón de rayo en la barra de herramientas.

Este script crea una nueva tabla relacional con atributos socioeconómicos para cada distrito censal e incluye una columna denominada tractid que actúa como clave externa para la tabla nyc\_census\_tracts. Además, este script crea tres vistas que juntos forman un algoritmo que califica cada distrito censal por la mayoría de las características que nos importan: población, tiempo medio de tránsito y tasa de criminalidad. Vamos a reservar estas características y puntuaciones por un tiempo para tratar con la cuarta característica (distancia de las estaciones de metro) primero.

**Visualización de nuestros datos geoespaciales**

Para comenzar a visualizar nuestros datos geoespaciales, abra la aplicación QGIS que descargó e instaló anteriormente. En QGIS, inicie un nuevo proyecto,haga clic con elbotón derecho en el nodo PostGIS en el área del navegador y seleccione Nueva conexión...

Una captura de pantalla de una publicación en las redes sociales

Descripción generada automáticamente

Introduzca en la ventana resultante (campos resaltados a continuación) las credenciales e información adecuadas para la instancia de Azure Database for PostgreSQL. Es probable que desee utilizar un nombre descriptivo,como se muestra a continuación. Tenga en cuenta que se requiere SSL para conectarse a la instancia de Azure Database for PostgreSQL.

Una captura de pantalla de un ordenador

Descripción generada automáticamente

Una vez que se hayaautenticado, verá nuestras cuatro entidades/tablas geoespaciales en la base de datos nyc: distritos censales, homicidios, callesy estaciones demetro. Seleccione las cuatro,haga clic con elbotón derecho en cualquiera de ellas y seleccione "Agregar capas seleccionadas al proyecto" como se muestra a continuación.

Una captura de pantalla de una publicación en las redes sociales

Descripción generada automáticamente

Una vez que los datos geoespaciales se cargan en el proyecto QGIS, podemos activar / desactivar capas. Desde el valor predeterminado de ver todas las capas en la visualización resultante, desmarque la casilla junto a homicides para que solo estemos viendo distritos censales, calles y estaciones de metro como se muestra a continuación.

Una captura de pantalla de un teléfono celular

Descripción generada automáticamente

**Manipulación de datos geoespaciales y revisualización de la salida**

Ahora que nuestra configuración inicial en QGIS está completa, realizaremos varias operaciones en pgAdmin. Usted puede recordar que nuestro escenario de negocio implica seleccionar una ubicación en una intersección de calles para que nuestro sitio tenga acceso a más de una calle a la vez. Podemos crear la tabla de intersección ejecutando un script en la herramienta de consulta pgAdmin. En pgAdmin, haga clic con elbotón derecho en su base de datos nyc y seleccione Herramienta de consulta... tener un lugar para pegar/ejecutar el siguiente script:

CREATE TABLE public.nyc\_intersections

(

GID SERIE LLAVE PRIMARIA,

street1 carácter variable(200) COLLATE pg\_catalog." default",

carácter de calle2 variable(200) COLLATE pg\_catalog." default",

geometría geom NO NULL

)

CON (

OIDS - FALSO

)

pg\_default TABLESPACE;

/\* calcular todas las intersecciones en NYC en función de las geometrías de las calles \*/

INSERTAR EN nyc\_intersections (calle1, calle2, geom)

SELECT s1.name, s2.name, ST\_Intersection(s1.geom, s2.geom) AS intersection\_geom

DESDE nyc\_streets COMO s1

Unirse a nyc\_streets AS s2

ST\_INTERSECTS(s1.geom, s2.geom)

Dónde

s1.name NO ES NULL Y

s2.name NO ES NULL Y

s1.name <> s2.name

PEDIDO POR s1.name, s2.name;

/\* Eliminar intersecciones duplicadas \*/

ELIMINAR DE nyc\_intersections

DONDE GID NO IN (

SELECT gid FROM (

SELECT MIN(gid) como gid, geom

DESDE nyc\_intersections

GRUPO POR geom

) COMO intersecciones únicas);

Este script creates punto geografías de las intersecciones de las líneas que son calles. Debido a que también estamos realizando una unión cartesiana entre dos instancias de la tabla de calles, también elimina las intersecciones duplicadas - si tenemos la intersección de1st Avenue y Allen Street, tampoco necesitamos la intersección de Allen Street y 1st Avenue.

Una vez que haya creado sus intersecciones, vuelva a QGIS. Para visualizar las intersecciones recién creadas en QGIS, haga clic con el botón derecho en la conexión PostGIS que creó originalmente y seleccione "Actualizar". Debería ver una nueva entidad geoespacial denominada nyc\_intersections junto con las cuatro entidades originales que vimos anteriormente. Haga clic con el botón derecho en nyc\_intersections y selija "Añadir capa al proyecto". Con casi 43.000 intersecciones no duplicadas,la visualización QGIS debería llenarse de nuevos puntos que cubran casi por completo las otras entidades geoespaciales.

Debido a que sólo queremos considerar intersecciones seguras sin incluir una carretera que están a 500 metros o más de una estacióndemetro, podemos reemplazar el script pgAdmin anterior con el siguiente script.

/\* Eliminar de las intersecciones de consideración que están dentro de 500 metros de una estación de metro \*/

ELIMINAR DE nyc\_intersections

DONDE gid IN(

SELECT nyc\_intersections.gid

DESDE nyc\_intersections

UNIRSE a nyc\_subway\_stations

ST\_DWITHIN(nyc\_intersections.geom, nyc\_subway\_stations.geom, 500)

);

/\* Delete aquellas intersecciones que incluyen una o más autopistas o enlaces de autopistas) \*/

ELIMINAR DE nyc\_intersections

DONDE gid IN (

SELECT nyc\_intersections.gid

DESDE nyc\_intersections

JOIN (SELECT \* FROM nyc\_streets tipo IN ('autopista', 'motorway\_link')) nyc\_streets

ST\_DWITHIN(nyc\_intersections.geom, nyc\_streets.geom, 10)

);

Running este script en pgAdmin eliminará más de 16,000 intersecciones de la consideración. Podemos ver el impacto de esta manipulación volviendo a QGIS y actualizando nuestra vista de las intersecciones. Para ello, haga clic con el botón derecho en la lista de nyc\_intersections en la ventana de capas de QGIS y seleccione "Actualizarcapa SQL... " como se muestra a continuación.

Una captura de pantalla de un mapa

Descripción generada automáticamente

Haga clic en el botón Actualizar en la esquina inferior derecha de la ventana resultante del Administrador de bases de datos para actualizar los datos de intersección,como se muestra a continuación:

Una captura de pantalla de una publicación en las redes sociales

Descripción generada automáticamente

Cuando QGIS termine de actualizar la capa de intersecciones, verá que ahora hay una zona de búfer alrededor de las paradas de metro. En caso de que las estaciones de metro y las intersecciones tengan puntos de color similares asociados con ellos, suzona deamortiguación se hace fácilmente visible si se alterna la visibilidad de las estaciones de metro.

**Manipulación final de datos**

Ahora que hemos eliminado de la consideración de las intersecciones que involucran autopistas o están demasiado cerca de las estaciones de metro, podemos volver a reducir nuestras opciones basadas en la población, los tiempos promedio de viaje diario y la tasa de criminalidad. Usted puede recordar que una de las vistas que creamos previamente puntuaciones distritos censales por estas variables - podemos ver los 10 principales identificadores de distrito censal ejecutando la siguiente consulta en pgAdmin:

SELECT tractid

DESDE vw\_nyc\_site\_selection\_data

LÍMITE 10

Los resultados de esta consulta nos permiten crear una consulta que excluye todas las intersecciones que no están en o cerca de esta lista Top 10 de distritos censales:

/\* eliminar aquellas intersecciones que NO están DENTRO DE 500 METROS de los 10 principales distritos censales \*/

ELIMINAR DE nyc\_intersections

DONDE GID NO IN (

SELECT nyc\_intersections.gid

DESDE nyc\_intersections

UNIRSE a nyc\_census\_tracts

ON ST\_DWithin(nyc\_intersections.geom, nyc\_census\_tracts.geom, 500)

DONDE tractid IN ('36005046201', '36061013600', '36061015400','36061013800', '36061006600', '36061013400', '36061024500', '36061015700', '36061013200', '36081045500')

)

Después de ejecutar esta última consulta en pgAdmin, puede actualizar la capa nyc\_intersections de nuevo en QGIS para ver que hemos reducido nuestras consideraciones significativamente,hasta poco más de350 intersecciones que cumplen nuestros criterios. Puede que no sea posible refinar automáticamente más que esto porque también tendríamos que evaluar la disponibilidad de la propiedad, los precios de alquiler / compra, etc. Sin embargo, hemos eliminado muy rápidamente más del 99% de las intersecciones en la ciudad de Nueva York a través de una combinación de consultas tradicionales y geoespaciales como se muestra a continuación.

Un primer plano de un mapa

Descripción generada automáticamente

Para obtener más detalles sobre los datos geoespaciales y las consultas,consulte el taller [Introducción a PostGIS,](https://postgis.net/workshops/postgis-intro/) que inspiró e informó este tutorial.