Introducción

El propósito de este documento es exponer y documentar los avances realizados en el proyecto de ‘Índice de Transitabilidad’. Para ello se utilizará como títulos las etapas propuestas en el diagrama de Gantt de trabajo[[1]](#footnote-0), de manera que este documento resulte complementario al mismo.

Herramientas

Para llevar adelante las distintas etapas del proyecto se utilizaron diferentes herramientas que permiten almacenar las fuentes de datos , procesarla y realizar las transformaciones pertinentes, como también así almacenar y visualizar los resultados.

Se utilizaron las siguientes herramientas:

* Notebook en Databricks: [001.transitabilidad](https://eastus2.azuredatabricks.net/?o=188860186850355#notebook/999827518039874/command/770964731218770)
* Servicio Postgres *godatos* en Azure
* Python
* QGIS

Tablas

Las tablas que intervinieron en los diferentes procesos de actualización del Índice se encuentran almacenadas en la DDBB *transitabilidad* del Postgres *godatos* del equipo de CDD. Éstas son:

* **Etapas de la cuarentena:** esta tabla contiene las diferentes etapas de la cuarentena desde su inicio en Marzo 2020.
* **Franjas horarias:** esta tabla contiene las franjas horarias que serán tenidas en cuenta para la construcción del Índice RUS
* **Transitabilidad\_test\_2020:** esta tabla se genera desde el notebook en Databricks, y contiene para cada eje de calle, dentro de un periodo determinado el indicador sube.
* **calendario\_actividades\_ponderadores:** esta tabla contiene para cada día, para cada franja horaria, para cada actividad comercial definida anteriormente un peso de ponderación asignado. Estos Ponderadores pueden ser actualizados mediante la ejecución de la función ”actualizar\_ponderador” disponible en Postgres.
* **calendario:** el mismo contiene los días desde enero a Diciembre 2020 teniendo en cuenta los días feriados.
* **rus**: esta tabla contiene el resultado del Relevamiento de Usos de Suelo 2017. La misma está disponible también en BA Data.
* **comercios\_clasificados\_rus**: esta tabla se genera mediante un script de Python. Contiene la clasificación de los diferentes usos de suelo, teniendo en cuenta un diccionario creado para tal fin.

Repositorio

El repositorio está disponible en [Github](https://github.com/Florpa/Rus), bajo el usuario Florpa. La carpeta se encuentra clonada en la VM de Windows en Azure, en el directorio *F:\repos\_flor\Rus*.

Etapas

Análisis de oportunidades y mejoras del indicador 2019

Durante esta etapa se actualizó la metodología de cómo se generaba el indicador sube. Esta actualización consistió principalmente en :

* Replicar la metodología en Databricks
* Desagregar el ponderador sube en Inicio y Trasbordo
* Calcular las distancias dentro de la red peatonal de calles, y no simplemente la distancia euclídea o “distancia a vuelo de pájaro”

Guardar los resultados del procesamiento en una tabla dentro de Postgres (**Transitabilidad\_test\_2020**) .

Revisión de la “Nueva realidad” de los usos de suelo para poder adaptar el indicador. Propuestas

Esta etapa consistió en:

* Revisar las categorías del Relevamiento de Usos de Suelos que se realizó en 2017.
* Revisar las franjas horarias que se habían utilizado anteriormente y evaluar su continuidad.
* Revisar las franjas horarias
* Revisar la consistencia de los grupos nomencladores de actividades que habían sido utilizados anteriormente (Residencial,Industrial,etc)

Resultados

En esta etapa se puso énfasis en lograr construir un flujo de datos que permitiera mayor flexibilidad a la hora de

* clasificar los distintos usos de suelos en grandes categorías,
* Modificar el valores de los ponderadores según se crea conveniente
* generar una tabla que contenga para cada día del año, y franja horaria un valor del ponderador según la actividad comercial.

Para ello se generó el script **001\_clasificacion\_actividades.** El mismo tiene como fin tomar como input la tabla RUS 2017 y para cada id y rama comercial relevada, asignarle una categoría usando un diccionario. Estos resultados se almacenan en **comercios\_clasificados\_rus**. Esta última sumada al calendario con los inputs necesarios para poder construir la tabla **calendario\_actividades\_ponderadores.** (ver imagen adjunta)

La tabla **calendario\_actividades\_ponderadores.** tiene para cada dia de la semana , franja y actividad un valor en el ponderador de RUS. En caso de querer modificar el valor del ponderador, se debe ejecutar la siguiente funcion en Postgres:

**public.actualizar\_ponderador.** La debe recibir los siguientes parámetros:

* el mes a actualizar (1,2,3, etc)
* el nombre del dia (lunes, martes, miércoles),
* la franja horaria (ver tabla **Franjas horarias:** )
* la categoría comercial (comercial, gastronomia\_hoteleria)
* y el valor del ponderador que se desee**.**

En caso de no agregar alguno de estos parámetros la actualización se hará sobre todo el conjunto de datos. (ver imagen adjunta).Al finalizar la ejecución de la función, te muestra cuantos registros se vieron afectados a la actualización.

Para asignar valores de ponderación según las actividades, se deben tener en cuenta las diferentes etapas de la cuarentena y la habilitación de las diferentes actividades comerciales por fecha.

Evaluación de nuevas fuentes de datos a incorporar y Complementar con sube y RUS.

En esta etapa se evaluaron las siguientes fuentes alternativas como complementarias a la información del Indicador Sube y RUS.

**Comercios AGC:** corresponde al censo de comercios realizado en Enero del 2020. Solo incluye rubros comerciales

**Datos Telecom:** este set de datos se recibe semanalmente. Contiene información sobre la cantidad de dispositivos que estaban conectados a las antenas de telefonía móvil, desagregado por hora.

Resultados de la primera evaluación.

**Comercios AGC:** los comercios censados por AGC en ENERO 2020, resultan ser información complementaria al RUS 2017. Si bien sólo se censaron rubros comerciales, esta información se encuentra más actualizada que la relevada en 2017. Por tanto, podría servir de actualización para esta categoría, eliminando y/o agregando todos los comercios que se hayan modificado, de 2017 a 2020.

**Datos Telecom:** sobre este set de datos actualmente se realizan procedimientos que permiten trasladar la información de las distintas antenas disponibles para la CABA, a una unidad espacial diferente (radio, fraccio, grillas).

Resultó del análisis que esta información complementa a la proporcionada por SUBE, ya que capta la cantidad de personas presentes en la CABA, donde no necesariamente hay transacciones. Podría utilizarse en aquellas zonas donde se sabe hay gran tránsito peatonal (Calle Florida Por ejemplo),pero el indicador SUBE no lo refleja. El único inconveniente que se presenta al querer usar los datos de telecom , es que no se encuentra en la misma unidad espacial que el Índice de Transitabilidad.

Existen posibilidades que se pueda trasladar la cantidad de dispositivos a la unidad espacial “calles”, pero para ello hay construir una jerarquía que permita asignar los valores en función del tipo de calle que sea (avenidas, calles,etc). De lo contrario la distribución de la cantidad de dispositivos se dará de manera uniforme.

Actualmente existe una tablero donde se han puesto en dialogo estas diferentes set de datos. El mismo esta disponible en [Transitabilidad- Teste de Fuentes de datos](https://app.powerbi.com/groups/82b62191-0f43-44bb-b6ae-f8c3fab89e50/reports/98382e8b-6d39-427d-a5b4-a4ea7c5f4366?ctid=2377472e-8004-4648-8456-bd9687afa150)

1. Gantt de trabajo: [LINK](https://docs.google.com/spreadsheets/d/15w3G_a6Kn7OFKZ1hVwxYwcYcA2OM5x0UGM8mpoLsW9I/edit?usp=sharing). [↑](#footnote-ref-0)