



Comisión Económica para América  
Latina y el Caribe  
(CEPAL)



"Programa de Cooperación CEPAL-BMZ/giz  
2020-2022: Ciudades Inclusivas, sostenibles e  
inteligentes en el marco de la Agenda 2030".

# Taller: Análisis y Visualización de las Encuestas Origen Destino (EODs) de diferentes ciudades de Latinoamérica

**Paula Vásquez-Henríquez**  
**Ramiro Oliverio Fernández**

**Ciudades Inclusivas, Sostenibles e Inteligentes en el Marco de la Agenda 2030:**  
Caracterización de la Movilidad Urbana de América Latina en 6 ciudades piloto  
Bogotá, Buenos Aires, México, Montevideo, Santiago y Sao Paulo

## ¿Quiénes somos?

### **Paula Vásquez-Henríquez**

Ingeniera Civil Industrial, Magister en Ciencias de la Ingeniería, Universidad del Desarrollo.

Data Governance Engineer, Cornershop by Uber.

Docente Facultad de Ingeniería, Universidad del Desarrollo.



### **Ramiro Oliverio Fernández**

Sociólogo (UBA). Maestría en Data Mining & KDD (UBA).

Analytics Engineer en Addi

Docente UNAB



# Cuadernos Encuesta Origen-Destino

Cuadernos que permiten el análisis descriptivo de los campos incluidos en las encuestas de Uruguay, Argentina, México, Colombia, Chile y Brasil.

Constan de cuatro partes:

- Inicialización e instalación de paquetes
- Preparación de los datos
- Análisis descriptivos
- Mapas

Basados en *Python*



Taller: Análisis y Visualización de las Encuestas Origen Destino (EODs) - Demo y Actividades.ipynb

Archivo Editar Ver Insertar Entorno de ejecución Herramientas Ayuda Se han guardado todos los cambios

+ Código + Texto

▼ Taller: Análisis y Visualización de las Encuestas Origen Destino (EODs)

Por Paula Vásquez-Henríquez, Ramiro Oliverio Fernández y Ariel López

En esta demo podremos observar como se pueden realizar variados análisis y cálculo de indicadores relevantes utilizando la data de la Encuesta Origen-Destino de .

▼ Importando funciones y datos

Lo primero que debemos hacer, es conectarnos a la carpeta donde se encuentran almacenados todos los archivos. Para esto, nos conectaremos con nuestro Google Drive. Debemos permitir el acceso a nuestro Google Drive para poder cargar los datos necesarios. El proceso será exitoso si debajo de la primera celda recibimos el mensaje: Mounted at [/content/gdrive](#)

```
[1] from google.colab import drive
drive.mount('/content/gdrive', force_remount=True)

Mounted at /content/gdrive
```

Luego, debemos asegurarnos de que tenemos los paquetes necesarios para poder ejecutar correctamente todos los análisis. Este paso puede tomar un buen tiempo.

```
[2] %%capture
# La línea anterior oculta el texto de instalación para que sea más fácil de leer el cuaderno, sugerimos
# en primera instancia, comentarla ya que así podremos fijarnos si salta algún error.
!pip3 uninstall matplotlib -y
!pip install -q condacolab
```

# Repositorio

Tantos estos cuadernos, como la data de origen de las EODs y algunos outputs generados por las notebooks están disponibilizados en [este repositorio de github](#), el cual tiene instrucciones para correr los cuadernos tanto de manera local como en Google Colab.

Al ser un repositorio público, cualquiera puede colaborar reportando issues o generando mejoras en el código.

## enmodo : Análisis y visualización de las Encuestas Origen Destino (EODs) de diferentes ciudades de Latinoamérica

### Estructura del repositorio

Hay una carpeta por ciudad. Si se incluyeron diferentes años de las ENMODOS para 1 ciudad, va a haber una subcarpeta por cada año para la misma ciudad (ej. Bogotá).

#### Source

Cada ciudad tiene subcarpetas llamadas `source-{file_extension}` que varían dependiendo el tipo de archivo que haya publicado la ciudad con los datos de la ENMODOS.

#### Output

Cada ciudad tiene carpetas con nombre de extensión ej. `csv` donde hay un output por diferentes tipos de extensión.

#### Procesamiento

En las carpetas `python` se encuentra 1 notebook por cada ciudad-año con el procesamiento de los indicadores, diferentes visualizaciones y la generación de los archivos de output.

### Ejecución notebooks

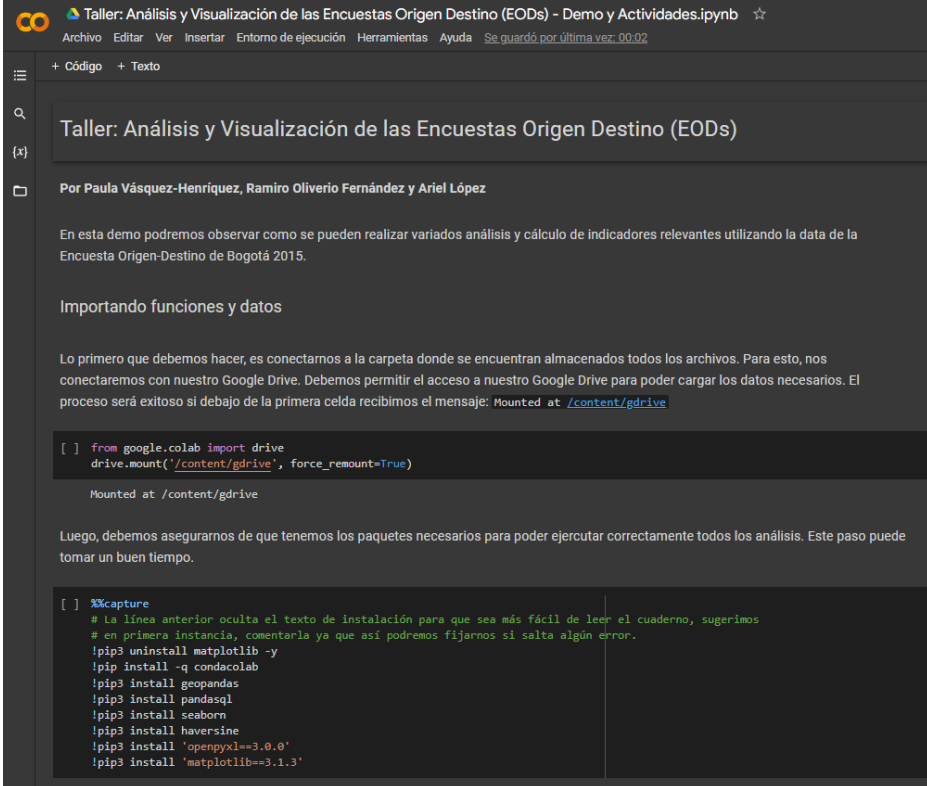
Para ejecutar las notebooks de manera local, luego de clonar el repositorio, en la notebook que se utilice reemplazar la línea donde se define el objeto `data_path` por el path local

```
# Reemplazar el string por el path local correspondiente
data_path = '/home/roliverio/Documentos/ADDI/enmodo/bogota/'
```

# Demo Santiago 2012

Para mostrar la utilización de los cuadernos, hemos preparado una demo sobre una de las EOD para repasar las funcionalidades del mismo. El cuaderno está montado en la herramienta Google Colab, que permite que directamente puedan correr el código sin necesidad de realizar alguna instalación.

Para acceder al cuaderno, usar [este link](#). Pueden realizar una copia del mismo en su propio drive para hacer las modificaciones que quieran y seguir el ejercicio propuesto.



```
Taller: Análisis y Visualización de las Encuestas Origen Destino (EODs) - Demo y Actividades.ipynb
+ Código + Texto

Taller: Análisis y Visualización de las Encuestas Origen Destino (EODs)

Por Paula Vásquez-Henríquez, Ramiro Oliverio Fernández y Ariel López

En esta demo podremos observar como se pueden realizar variados análisis y cálculo de indicadores relevantes utilizando la data de la Encuesta Origen-Destino de Bogotá 2015.

Importando funciones y datos

Lo primero que debemos hacer, es conectarnos a la carpeta donde se encuentran almacenados todos los archivos. Para esto, nos conectaremos con nuestro Google Drive. Debemos permitir el acceso a nuestro Google Drive para poder cargar los datos necesarios. El proceso será exitoso si debajo de la primera celda recibimos el mensaje: Mounted at /content/gdrive

[ ] from google.colab import drive
drive.mount('/content/gdrive', force_remount=True)

Mounted at /content/gdrive

Luego, debemos asegurarnos de que tenemos los paquetes necesarios para poder ejecutar correctamente todos los análisis. Este paso puede tomar un buen tiempo.

[ ] %capture
# La línea anterior oculta el texto de instalación para que sea más fácil de leer el cuaderno, sugerimos
# en primera instancia, comentarla ya que así podremos fijarnos si salta algún error.
!pip3 uninstall matplotlib -y
!pip3 install -q condocolab
!pip3 install geopandas
!pip3 install pandasql
!pip3 install seaborn
!pip3 install haversine
!pip3 install 'openpyxl==3.0.0'
!pip3 install 'matplotlib==3.1.3'
```

# Actividades a realizar

Al final de la demo, hay una serie de actividades para modificar parte del código que se ve en los cuadernos para generar ciertos gráficos imputando diferentes variables.

Para que funcionen, anteriormente hay que ejecutar la totalidad del cuaderno hasta esa celda y luego reemplazar los gaps “\_\_\_\_\_” por la variable indicada (ej. “TramoIngreso”)

## ▼ Actividades a realizar

Rellene los espacios con líneas con las variables correspondientes para responder las siguientes preguntas.

### ▼ ACTIVIDAD 1: ¿Cómo es la distribución de personas encuestadas de acuerdo a su actividad principal y género?

```
[72] ax = (personas_df.groupby(['_____', '_____']).size().unstack() # Ocupacion y Sexo
      .pipe(normalize_rows)
      .plot(kind='barh', stacked=True, figsize=(6, 3)))
ax.set_xlabel("% de Encuestados")
ax.set_ylabel(" ")
ax.set_title("Distribución de Personas encuestadas por sexo y actividad principal")
sns.despine()

plt.legend(loc=0, bbox_to_anchor=(1.0, 0.5))
```

### ▼ ACTIVIDAD 2: ¿Cómo es la razón de viajes motorizados por sobre los peatonales por Tramo de Ingresos cuando aperturamos por género?

```
[74] viajes_peaton = viajes_peaton_habiles.groupby(['_____', '_____']).agg(n_peaton= ('PONDERADOR_CALIBRADO_VIAJES', 'sum')).reset_index() # TramoIngreso, Sexo
viajes_privados = viajes_privado_habiles.groupby(['_____', '_____']).agg(n_privados= ('PONDERADOR_CALIBRADO_VIAJES', 'sum')).reset_index() # TramoIngreso, Sexo
n_viajes_modos = pd.merge(viajes_peaton, viajes_privados, on=["TramoIngreso", "Sexo"]) # TramoIngreso, Sexo
n_viajes_modos['motorizado_sobre_peaton'] = n_viajes_modos['n_privados'] / n_viajes_modos['n_peaton']

g = sns.catplot(x="TramoIngreso", y="motorizado_sobre_peaton",
               capsizes=2, height=4, aspect=4,
               kind="point", data=n_viajes_modos, hue='Sexo')

g.fig.suptitle('Razón entre los viajes en transporte motorizado sobre peatonal en un día hábil por género')
g.set_ylabels('Proporción de viajes motorizado sobre peatonal')
g.set_xlabel('Tramo Ingreso')
g.set_xticklabels(rotation=30)
g.set_ylim=(0,15)
```

### ▼ ACTIVIDAD 3: ¿Cuál es la distribución/partición modal de los viajes por género?

## Bibliografía recomendada para partir con Python

- [The Python Handbook](#)
- [Discover the Best Resources to Learn Python for Data Science](#)
- [GitHub - jerry-git/learn-python3: Jupyter notebooks for teaching/learning Python 3](#)
- [GitHub - trekhleb/learn-python: 📖 Playground and cheatsheet for learning Python. Collection of Python scripts that are split by topics and contain code examples with explanations.](#)
- [GitHub - Asabeneh/30-Days-Of-Python](#)
- [A curated list of awesome Python frameworks, libraries, software and resources](#)