

DESCRIPCION DEL DATASET

Datasets utilizados como fuentes de datos

Para este proyecto se utilizaron 3 datasets, los cuales se detallan a continuación:

1) 16_1_04_Visitas-al-Parque-Nacional-TDF.xlsx

El archivo muestra la cantidad de entradas al parque nacional Tierra del Fuego con cualquier finalidad (ocio, negocios u otro motivo personal) que no sea ser empleado por dicho parque, teniendo en cuenta también su condición de residencia. Cada vez que se cruza la frontera del área protegida, se genera una visita. Los datos comprenden el periodo entre Enero del 2015 y Agosto de 2024.

Los datos que muestra el dataset son:

- **Periodo:** Muestra el periodo al cual corresponde el dato. Se muestra en meses y años.
- **Total Visitas:** Muestra la sumatoria de las visitas del periodo correspondiente.
- **Visitas por condición de residencia:** Muestra la cantidad de visitantes según su condición de residente o no residente. Esto es:
 - **Residentes:** visitas realizadas por personas que residen en el país.
 - **No residentes:** visitas realizadas por personas extranjeras.

Origen del dataset: <https://ipiec.tierradelfuego.gob.ar/>.

Link de descarga: https://ipiec.tierradelfuego.gob.ar/wp-content/uploads/2024/09/16_1_04_Visitas-al-Parque-Nacional-TDF.xlsx.

Las fuentes de datos para la confección del mismo son: Ministerio del Interior. Subsecretaría de Turismo. Dirección Nacional de Mercados y Estadísticas con base en datos de la Dirección de la Administración de Parques Nacionales.

2) 16_1_01_Pernoctaciones_ingresos_estadia_promedio.xlsx

El archivo muestra la cantidad de pernoctaciones, ingreso de viajeros y estadía promedio, según condición de residencia, en Ushuaia y Río Grande. Los datos comprenden el periodo entre Enero del 2004 y Julio de 2024.

Los datos que muestra el dataset se repiten por cada una de las ciudades, y los mismos son:

- **Periodo:** Muestra el periodo al cual corresponde el dato. Se muestra en meses y años.
- **Total Pernoctaciones:** Muestra la sumatoria de las pernoctaciones del periodo correspondiente, donde una pernoctación es una noche que el turista se aloja en el establecimiento hotelero. También se muestra el total según su condición de residente o no residente. Esto es:
 - **Residentes:** visitas realizadas por personas que residen en el país.
 - **No residentes:** visitas realizadas por personas extranjeras.
- **Total Viajeros:** Muestra la sumatoria de los viajeros del periodo correspondiente, donde un viajero es toda persona que se ha trasladado de su lugar de residencia habitual por razones de diversa índole, tales como el ocio, los negocios, la visita a familiares o amigos, etc, y que realiza una o más pernoctaciones seguidas en el mismo establecimiento hotelero o parahotelero y que abona por tal servicio. También se muestra el total según su condición de residente o no residente. Esto es:
 - **Residentes:** visitas realizadas por personas que residen en el país.
 - **No residentes:** visitas realizadas por personas extranjeras.
- **Total Estadía Promedio:** Duración promedio en días en que se alojaron los viajeros. Esta medida en días y se calcula como Pernoctes / Viajeros. También se muestra según su condición de residente o no residente. Esto es:
 - **Residentes:** visitas realizadas por personas que residen en el país.
 - **No residentes:** visitas realizadas por personas extranjeras.

Origen del dataset: <https://ipiec.tierradelfuego.gob.ar/>.

Link de descarga: https://ipiec.tierradelfuego.gob.ar/wp-content/uploads/2024/09/16_1_01_Pernoctaciones_ingresos_estadia_promedio.xlsx.

Las fuentes de datos para la confección del mismo son: INDEC, Encuesta de Ocupación Hotelera (EOH) y Observatorio Estadístico, Secretaría de Modernización e Innovación, Municipio de Río Grande.

3) 14_5_08_Evolucion-de-Pasajeros-por-tipo.xlsx

El archivo muestra la cantidad de recaladas y pasajeros ingresados al puerto de Ushuaia, correspondientes a lo declarado en los manifiestos de los buques. Los datos comprenden el periodo entre 2008 y 2023.

Los datos que muestra el dataset son:

- **Temporada:** Muestra el periodo al cual corresponde el dato. Se muestra en años.

- **Recaladas:** Muestra la sumatoria de recaladas del periodo correspondiente, donde una recalada es la llegada de un buque/nave, después de una navegación, a la vista de un punto de la costa, como fin de viaje o para, después de reconocido, continuar su navegación.
- **Total Tipo de Pasajeros:** Muestra la sumatoria en el periodo correspondiente de las personas que viajan en el buque/nave. A estos se los puede clasificar como:
 - **Tripulantes:** es el conjunto de personas que van en una embarcación, dedicadas a su maniobra y servicio. Se trata de una estructura jerarquizada, organizada en departamentos según su función.
 - **Pasajeros:** son los que realizan recorridos turísticos.

Origen del dataset: <https://ipiec.tierradelfuego.gob.ar/>.

Link de descarga: https://ipiec.tierradelfuego.gob.ar/wp-content/uploads/2024/03/14_5_08_Evolucion-de-Pasajeros-por-tipo.xlsx.

Las fuentes de datos para la confección del mismo son: Dirección Provincial de Puertos.

Dataset versión definitiva

Los dataset mencionados anteriormente no tienen un formato “tidy”. Un formato “tidy”, también conocido como datos ordenados o en formato largo, es un conjunto de datos en el que: cada fila es una observación y cada columna es una variable. Debido a esto, se realizó la extracción de la información útil de cada uno de ellos y se la unificó en un solo archivo CSV.

Los datos en el mismo son los siguientes:

anio: [*tipo numérico*] año al que corresponden los datos de la fila.

mes: [*tipo categórico*] mes al que corresponden los datos de la fila.

visitas_pn_total: [*tipo numérico*] total de visitantes al parque nacional en la fecha especificada. Es la suma de visitas_pn_r y visitas_pn_nr.

visitas_pn_r: [*tipo numérico*] total de visitantes residentes al parque nacional en la fecha especificada.

visitas_pn_nr: [*tipo numérico*] total de visitantes no residentes al parque nacional en la fecha especificada.

pernoctantes_ush: [*tipo numérico*] total de pernoctes en la ciudad de Ushuaia en la fecha especificada.

pernoctantes_ush_r: [*tipo numérico*] total de pernoctes residentes en la ciudad de Ushuaia en la fecha especificada.

pernoctantes_ush_nr: [*tipo numérico*] total de pernoctes no residentes en la ciudad de Ushuaia en la fecha especificada.

viajeros_ush: [*tipo numérico*] total de viajeros en la ciudad de Ushuaia en la fecha especificada.

viajeros_ush_r: [*tipo numérico*] total de viajeros residentes en la ciudad de Ushuaia en la fecha especificada.

viajeros_ush_nr: [*tipo numérico*] total de viajeros no residentes en la ciudad de Ushuaia en la fecha especificada.

estadia_prom_ush: [*tipo numérico*] estadía promedio en días de visitantes en Ushuaia en la fecha especificada.

estadia_prom_ush_r: [*tipo numérico*] estadía promedio en días de visitantes residentes en Ushuaia en la fecha especificada.

estadia_prom_ush_nr: [*tipo numérico*] estadía promedio en días de visitantes no residentes en Ushuaia en la fecha especificada.

pernoctantes_rg: [*tipo numérico*] total de pernoctes en la ciudad de Rio Grande en la fecha especificada.

pernoctantes_rg_r: [*tipo numérico*] total de pernoctes residentes en la ciudad de Rio Grande en la fecha especificada.

pernoctantes_rg_nr: [*tipo numérico*] total de pernoctes no residentes en la ciudad de Rio Grande en la fecha especificada.

pernoctantes_rg_sd: [*tipo numérico*] total de pernoctes sin declarar estado de residencia en la ciudad de Rio Grande en la fecha especificada.

viajeros_rg: [*tipo numérico*] total de viajeros en la ciudad de Rio Grande en la fecha especificada.

viajeros_rg_r: [*tipo numérico*] total de viajeros residentes en la ciudad de Rio Grande en la fecha especificada.

viajeros_rg_nr: [*tipo numérico*] total de viajeros no residentes en la ciudad de Rio Grande en la fecha especificada.

viajeros_rg_sd: [*tipo numérico*] total de viajeros sin declarar estado de residencia en la ciudad de Rio Grande en la fecha especificada.

estadia_prom_rg: [*tipo numérico*] estadía promedio en días de visitantes en Rio Grande en la fecha especificada.

estadia_prom_rg_r: [*tipo numérico*] estadía promedio en días de visitantes residentes en Rio Grande en la fecha especificada.

estadia_prom_rg_nr: [*tipo numérico*] estadía promedio en días de visitantes no residentes en Rio Grande en la fecha especificada.

No todos los datasets cubrían los mismos periodos de tiempo, por lo que se tuvieron que eliminar ciertos sectores de los mismos, dejando solo los registros donde coincidían las fechas. Debido a esto, el dataset quedó constituido por 108 filas y 25 columnas.

Tambien se puede observar que el dataset no incorpora los datos de los turistas que ingresaron a la provincia a través de buques/naves. Esto se debe a que los datos del dataset original están dados en formato anual, lo cual no coincide con el formato de los otros dataset originales y, por lo tanto, no se los pudo utilizar.