

- Entrega 2: 17/06 Descripción del Dataset y Origen

En esta segunda entrega, debe tener acceso al dataset que utilizará en su proyecto de Aprendizaje Automático.

Se informa el link de acceso al dataset: https://github.com/edfedo/Plazas_Disponibles_Hotelero/tree/main/data/external

En dicha ruta, podrá ver dos archivos de Excel XLSX, uno contiene la información tal cual se adquirió "16_3_02_Oferta_Hotelera_por-sector" y el otro "LIMPIO_16_3_02_Oferta_Hotelera_por-sector_LIMPIO" con el cual se procederá a realizar el proyecto, el mismo ya fue tratado y pasado por un proceso de limpieza manual.

Proporcione una descripción completa del dataset, incluyendo la cantidad de instancias, características (columnas), tipos de datos, y cualquier información relevante.

La información que se redacta a continuación, se la puede ver dentro del archivo .md en el repo Git.

Se adjunta link: https://github.com/edfedo/Plazas_Disponibles_Hotelero/blob/main/docs/descripcion_datos.md

Descripción del conjunto de datos:

El dataset contiene datos mensuales sobre la oferta hotelera y parahotelera desde enero de 2008 hasta febrero de 2024. Las variables incluidas son:

Periodo: Mes y año del registro.

Establecimientos_Hotelero: Número de establecimientos hoteleros.

Plazas_Disponibles_Hotelero: Número de plazas disponibles en el sector hotelero.

Habitaciones_Unidades_Disponibles_Hotelero: Número de habitaciones o unidades disponibles en el sector hotelero.

Establecimientos_Parahotelero: Número de establecimientos parahoteleros.

Plazas_Disponibles_Parahotelero: Número de plazas disponibles en el sector parahotelero.

Habitaciones_Unidades_Disponibles_Parahotelero: Número de habitaciones o unidades disponibles en el sector parahotelero.

Este conjunto de datos también contiene información sobre la disponibilidad de establecimientos turísticos, plazas disponibles (camas) y habitaciones o unidades disponibles a lo largo de varios años y meses. Aquí tienes una descripción detallada:

Resumen del Conjunto de Datos

Número de Instancias:

El conjunto de datos tiene datos desde enero de 2008 hasta febrero de 2024, inclusive.

Cada año tiene datos mensuales, excepto algunos valores faltantes debido a circunstancias específicas (como la pandemia de COVID-19).

Hay 192 instancias (filas), por cada sheet.

Características (Columnas) y Tipos de Datos:

Periodo (Año): El año del registro.

Tipo de Datos: Entero

Ejemplo: 2008, 2009, etc.

Mes (Mes): El mes del registro.

Tipo de Datos: Cadena (Categorico)

Ejemplo: "enero", "febrero", etc.

Establecimientos: El número de establecimientos turísticos disponibles.

Tipo de Datos: Entero

Ejemplo: 104, 103, etc.

Plazas disponibles: El número de plazas (camas) disponibles en los establecimientos.

Tipo de Datos: Entero

Ejemplo: 70,649, 64,467, etc.

Habitaciones o Unidades disponibles: El número de habitaciones o unidades disponibles en los establecimientos.

Tipo de Datos: Entero

Ejemplo: 21,452, 19,372, etc.

Características de los Datos:

Rango Temporal: Los datos abarcan desde enero de 2008 hasta febrero de 2024.

Granularidad Mensual: Los datos se proporcionan mensualmente para cada año.

Interrupción en los Datos: Hay una interrupción notable en la recopilación de datos durante 2020 debido a la pandemia de COVID-19, sin datos durante varios meses.

Información faltante:

Algunas entradas están marcadas con "s", Ver Referencias

Los meses en 2023 y 2024 marcados con "(*)" Ver Referencias

Referencia:

-> * **Dato provisorio**

-> s **Dato confidencial por aplicación de las reglas del secreto estadístico**

Ejemplo de Puntos de Datos:

Enero 2008:

Establecimientos: 104

Plazas disponibles: 70,649

Habitaciones o Unidades disponibles: 21,452

Julio 2010:

Establecimientos: 70

Plazas disponibles: 54,188

Habitaciones o Unidades disponibles: 11,625

Agosto 2023 (Datos preliminares):

Establecimientos: 37

Plazas disponibles: 33,635

Habitaciones o Unidades disponibles: 7,936

Observaciones Importantes:

Tendencias a lo Largo del Tiempo: El número de establecimientos, plazas disponibles y habitaciones generalmente disminuye de 2008 a 2024, con fluctuaciones notables posiblemente debido a factores económicos, ambientales o de salud.

Impacto del COVID-19: Significativos vacíos de datos y reducción en los números reportados en 2020 debido a la pandemia. Recuperación y Fluctuaciones: Después de 2020, hay un intento de reanudar la recopilación de datos y las cifras muestran signos de recuperación. Este conjunto de datos proporciona una visión completa del sector de alojamiento turístico durante un período de 16 años, valiosa para el análisis de tendencias y la previsión en la industria del turismo.

Informe sobre el origen del dataset, es decir, de dónde provienen los datos. Esto puede incluir la fuente, la fecha de adquisición y cualquier proceso de recopilación o preprocesamiento que haya realizado.

Si bien el DataSet original es el "16_3_02_Oferta_Hotelera_por-sector", se procedera a usar un DataSet Limpio ya trabajado de nombre "LIMPIO_16_3_02_Oferta_Hotelera_por-sector_LIMPIO"

16_3_02_Oferta hotelera e indicadores seleccionados por sector. Ushuaia y Río Grande, Pcia. de Tierra del Fuego, Antártida e Islas del Atlántico Sur.

Área Temática: Turismo

Tema: Turismo

Subtema: Actividad e Infraestructura de la industria turística

Año Desde: 2008

Año Hasta: 2023

Distribución geográfica: Ushuaia y Río Grande. Tierra del Fuego AeIAS.

Fuente: Instituto Provincial de Análisis e Investigación, Estadísticas y Censos sobre la base de datos de la Encuesta de Ocupación Hotelera del INDEC, y del Observatorio Estadístico, Secretaría de Modernización e Innovación, Municipio de Río Grande.

Formato: Excel XLSX

Link: <https://ipiec.tierradelfuego.gob.ar/>

Link2: <https://ipiec.tierradelfuego.gob.ar/estadisticas-economicas-2/>

Esta entrega debe ser realizada en el repositorio GIT, incluyendo archivos de dataset y documentos.

Link Repo Git: https://github.com/edfedo/Plazas_Disponibles_Hotelero/tree/main

Link de acceso al dataset: https://github.com/edfedo/Plazas_Disponibles_Hotelero/tree/main/data/external

Link descripcion: https://github.com/edfedo/Plazas_Disponibles_Hotelero/blob/main/docs/descripcion_datos.md

Link Origen y descripcion de datos: https://github.com/edfedo/Plazas_Disponibles_Hotelero/blob/main/references/Origen%20y%20descripcion%20datos%20archivos%20Oferta%20Hotelera.md

Link Reporte de resultados: https://github.com/edfedo/Plazas_Disponibles_Hotelero/blob/main/reports/Reporte%20de%20Resultados.md