

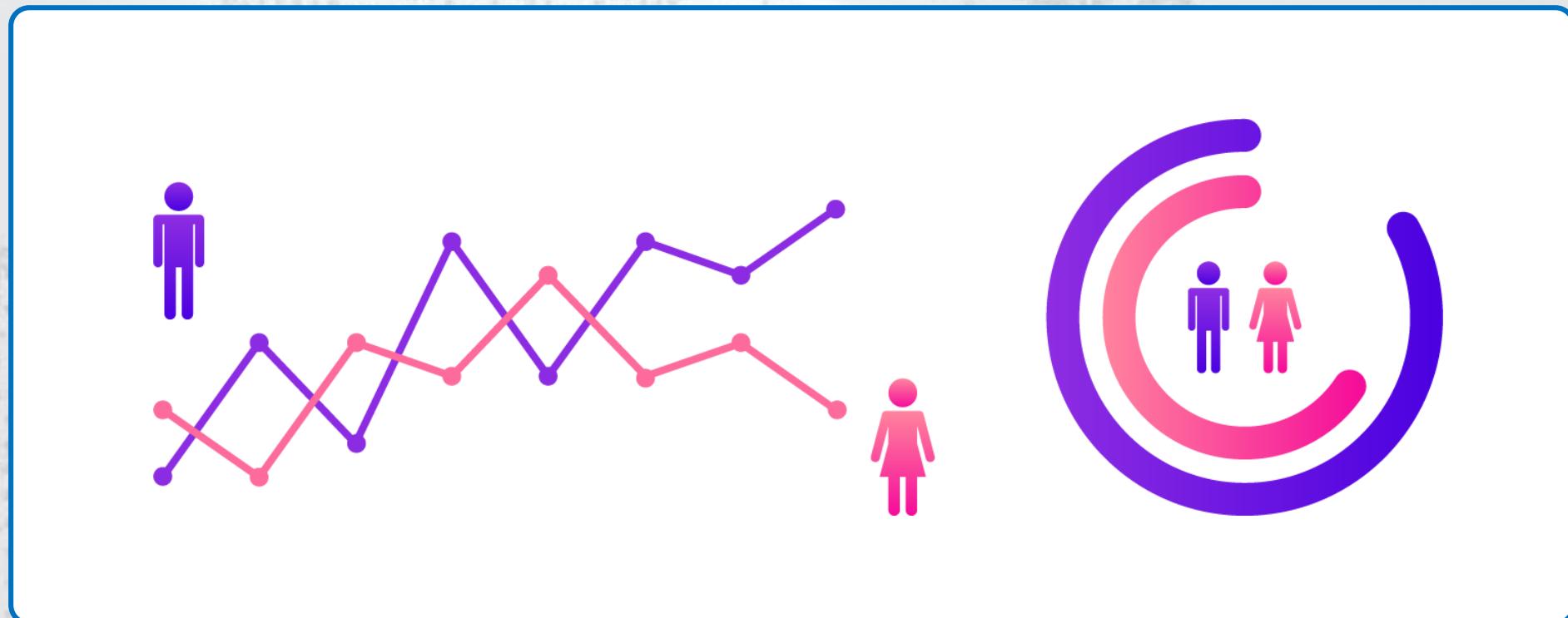


Módulo 2 – Obtención y Preparación de Datos

Actividad: Caso API UN comtrade

Ciencia de Datos

Actividad: Caso Limpieza Datos UN Comtrade



The screenshot shows the UN Comtrade Database homepage. At the top left is the logo 'UN Comtrade' with a green 'C' icon. Below it is the word 'Database'. A central orange box contains the text: 'UN Comtrade Notice: Upgrade plan 2020'. To the right of the notice is a preview message: 'The preview of new data extraction is accessible for public testing at main data extraction page. Please visit [this link](#) to see all the details.' Below the notice is a green 'DASHBOARD' button with several checked items: 'Data Availability', 'Data Availability+' (BETA), 'UN Comtrade Analytics', 'Metadata Survey', and 'Monitor Impact of COVID-19'. At the bottom is a blue 'NEWS' section with a 'Tweets by @UNTradeStats' link. On the left side, there's a 'RECENT DATA RELEASE' section with filters for 'Released' (ultimo 30 días), 'Category' (Todas), 'Period' (Todas), and 'Country' (Todas). It shows statistics: 419 mill. Total Records, 326 Number of Dataset, 49 Number of Country, and 2021-10 Latest Period. Below these are two line charts: one for 'NEW' data (blue line) and one for 'REVISED' data (red line). The x-axis for both charts shows dates from 21 nov to 12 dic. The y-axis for the revised chart ranges from 0 to 30.

Caso UN Comtrade

- **UN Comtrade** es el seudónimo para la base de datos estadística de comercio de las Naciones Unidas, en las cuales más de 170 países o áreas del mundo, llamados **reporters**, proveen información a la división estadística de las Naciones Unidas con su información anual acerca de comercio e intercambio detallada en categorías de commodities o servicios con sus partners de comercio. Más información en la web oficial de UN Comtrade.
- <https://comtrade.un.org>

STATISTICS**UN Comtrade**

Extract data ▾

Data Availability ▾

Metadata ▾

Reference ▾

Knowledge base

API portal

Dear users, our Comtrade API is experiencing some technical difficulties. To avoid a complete disruption in our service we have restricted to 10,000 records per call for our guest users. **Premium users** are not being affected by this measure. Thank you for your understanding.

1. Type of product & Frequency**Type of product** Goods Services**Frequency** Annual Monthly**2. Classification****HS** As reported 92 96 02 07 12 17**SITC** As reported * Rev. 1 Rev. 2 Rev. 3 Rev. 4**BEC** BEC**3. Select desired data****Periods (year)**

x 2020

All or a valid period. Up to 5 may be selected.

Caso UN Comtrade

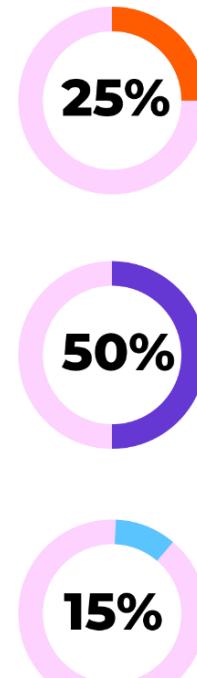
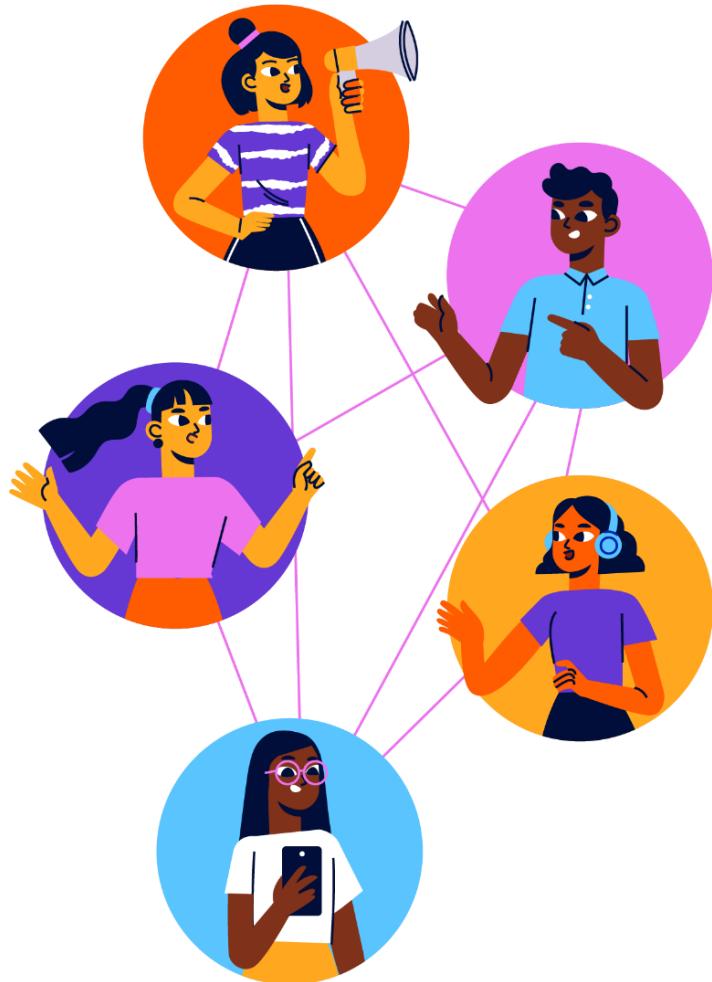
Para la consulta de información, se dispone de una aplicación web de consultas mediante un formulario, así como también una API, la cual tiene acceso público.

El trabajo de obtención ya se ha realizado y se cuenta con un archivo en formato csv con el intercambio comercial de Chile con el resto del mundo entre los años 2016 y 2019.

Caso UN Comtrade

Construya un notebook, ordenado, documentado y reproducible para desarrollar los siguientes requerimientos que se solicitan. Se pide:

- Hacer un análisis exploratorio de los datos, variables numéricas y categóricas.
- Hacer un análisis de valores nulos en el set de datos ¿Qué porcentaje de valores nulos posee las variables numéricas de estudio?.
- Hacer un análisis de outliers ¿Posee outliers las variables NetWeight y TradeValue?.
- Aplicar técnicas de limpieza.
- A partir de los datos limpios, obtener el valor total y peso de las exportaciones de commodities de Chile para el año 2019 y 2020.
- Guardar datos limpios en un archivo Excel.



¡Éxito!