

## Trabajo Final

A continuación, se presentan las indicaciones para la confección del trabajo final del módulo de Obtención y Preparación de Datos.

### Contexto del Problema

Como parte de su estrategia de transformación digital, el Banco Digital de Chile (BDC) ha formado un selecto grupo de profesionales, del cual usted es parte, para implementar una serie de proyectos de inteligencia artificial y analítica.

Su rol, dentro de la organización, es apoyar a los científicos de datos ya que ellos no tienen desarrolladas sus habilidades de Data Wrangling y necesitan un estrecho apoyo de los ingenieros de datos para llevar a cabo sus modelos de inteligencia artificial.

El proyecto actual en el que se encuentra participando, requiere que se construya una base de datos de información de la competencia en cuanto a las captaciones, colocaciones e inversiones a lo largo del tiempo. Esta tarea era desarrollada de forma manual por los analistas, quienes se conectaban a la página del Banco Central y descargaban las planillas Excel para alimentar los datos, pero se cometían muchos errores de transcripción, razón por la cual le han encomendado a usted que automatice un proceso de obtención y preparación de datos.

A continuación, se presenta el link para la descarga de los archivos que utilizan los analistas:

<https://www.bcentral.cl/contenido/-/detalle/serie-de-datos-bancarios-septiembre-de-2021>

Nótese que este link podría variar durante el módulo, por lo tanto, deberá buscar la información disponible siguiendo la navegación sugerida en el breadcrumb siguiente.

The screenshot shows a web browser window with the URL <https://www.bcentral.cl/contenido/-/detalle/serie-de-datos-bancarios-septiembre-de-2021>. The browser's address bar and tabs are visible at the top. Below the browser window, the website's header is shown, featuring the Banco Central Chile logo and navigation links: 'El Banco', 'Áreas', 'Noticias y Publicaciones', and 'Contacto Central'. A search bar is also present. The main content area displays the breadcrumb path: 'Home / Noticias y publicaciones / Publicaciones / Estadísticas / Dinero y banca / Serie de Datos Bancarios'. Below this, the word 'Publicaciones' is prominently displayed in a large, bold font. Underneath, the title 'Serie de Datos Bancarios septiembre de 2021' is shown. At the bottom, there is a box containing the text 'Series de datos bancarios'.

### Descripción

Las Series de Datos Bancarios, preparadas por la División de Estadísticas del Banco Central de Chile, contiene información por institución bancaria en materia de colocaciones, depósitos e inversiones. Estas series son publicadas los días 23 de cada mes, o el día hábil siguiente con un desfase de 2 meses y 23 días.

#### Serie de Datos Bancarios septiembre de 2021

Recuadros y gráficos



Haca click en el signo (+) y podrá acceder a los archivos:

#### Serie de Datos Bancarios septiembre de 2021

Recuadros y gráficos



Los archivos que usted debe descargar son los siguientes:

- Colocaciones
- Depósitos
- Inversiones

## Objetivo

Construir un programa en Python para la extracción y preparación de datos de colocaciones, depósitos e inversiones disponible en la página web del Banco Central. Recuerde que debe utilizar las técnicas y buenas prácticas aprendidas en clase en lo que respecta a limpieza de datos y data wrangling.

El formato de salida del reporte es el siguiente:

## Entregable

Los entregables de su trabajo son los siguientes:

- Un archivo csv que tenga la información consolidada del resultado
- El notebook de Python donde realiza la obtención y preparación de la data

El archivo de salida debe considerar las siguiente columnas:

- Fecha
- Institución Financiera
- Tipo (colocación, captación, inversión)
- Subtipo
- Valor
- Tipo de moneda

En el caso de los depósitos, se podrá dar cuenta que hay distintos tipos de colocaciones en las distintas hojas de la planilla, por ejemplo, colocaciones comerciales, colocaciones de consumo, de vivienda, de comercio exterior, etc. Utilice esta información para fijar el campo Subtipo solicitado.

También se podrá dar cuenta que algunas colocaciones están expresadas en dólares, con lo cual podrá alimentar el campo Tipo de Moneda. Lo mismo aplica para los otros dos archivos, es decir, de depósito y de inversiones.

Haga todos los supuestos que estime conveniente y déjelos bien documentados en su notebook. Asuma que el formato de los archivos es siempre el mismo y no cambia con el tiempo, por lo tanto usted podría referenciar una hoja particular sabiendo que se trata de un tipo o subtipo específico.

Haga todas las simplificaciones que estime conveniente para lograr el objetivo, lo importante es que los datos que contenga el archivo de salida sean consistentes.

## Plazo de entrega

Lunes 12 de diciembre de 2023 hasta las 08:30 am hrs. Posterior a esta fecha no se recibirán más trabajos ni se harán excepciones.

## Grupos de Trabajo

El trabajo puede ser realizado de forma individual o grupal, con un máximo de 2 integrantes por grupo.