



ANÁLISIS DE DATOS
FACULTAD DE INGENIERÍA Y NEGOCIOS
(ACI777)

Melissa Alegría Arcos, PhD.
malegriaa@udla.cl

Agosto 2021

Descripción del programa

Modalidad : E-SUPPORT

Horario : Martes 19:15 a 22:30

Distribución de horas de dedicación:

- Laboratorio: 54
- Trabajo Personal 126

La asignatura ANÁLISIS DE DATOS tiene por meta formativa que los estudiantes aprendan los fundamentos de análisis basado en datos para resolver problemas organizacionales y apoyar la toma de decisiones en base a la generación de modelos descriptivos y predictivos a partir de datos.

- Esta asignatura cuenta con un aula virtual e-support, la que será fuente oficial de material transversal para la asignatura.
<https://udla.blackboard.com>
- <https://github.com/malegria01/AnalisisDatos>

Descripción del programa

Contenido:

1) Estadística

- Estadística Descriptiva
- Analítica para toma decisiones.

2) Modelos Analíticos Descriptivos

- Análisis Exploratorio de datos

3) Aprendizaje No Supervisado

- Clustering
- Reducción de dimensionalidad
- Análisis de Componentes Principales (PCA)

4) Aprendizaje Supervisado

- Regresión Lineal y Regresión Logística
- Árboles de Decisiones
- Máquinas de Soporte Vectorial (SVM)

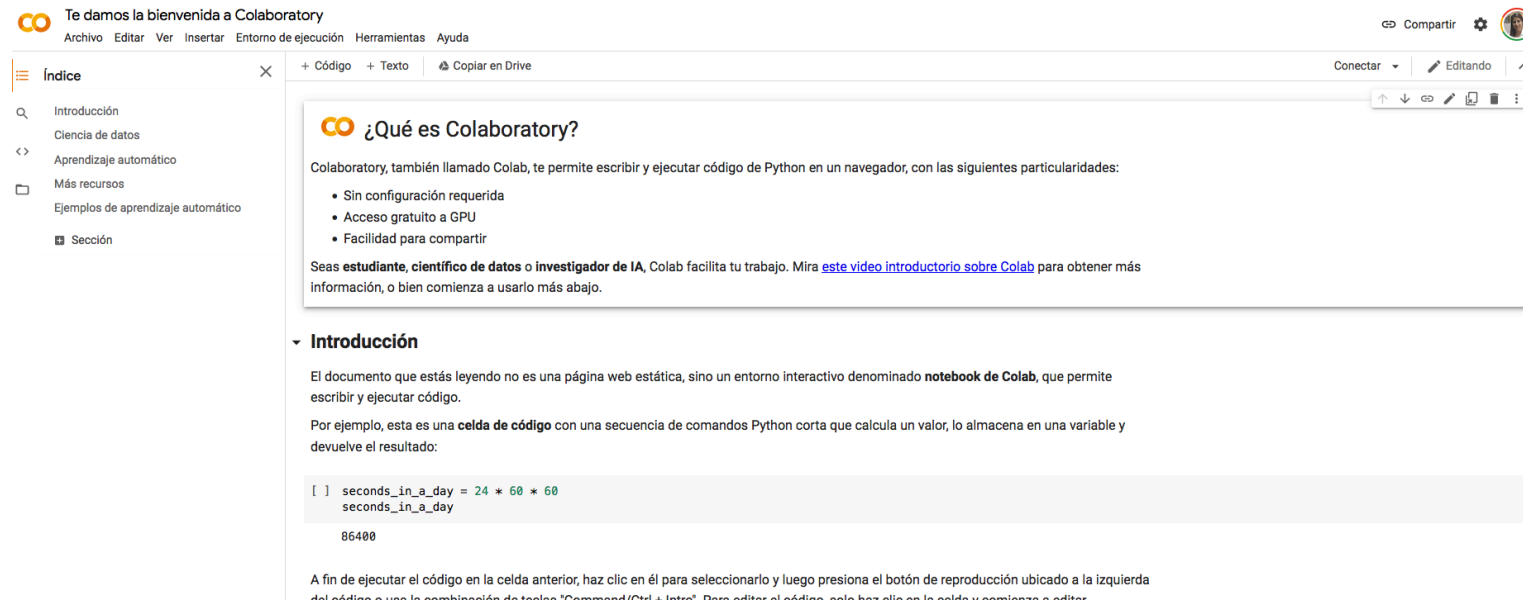
Evaluaciones

Régimen	Ponderación	Componente	% Componente	Subcomponente		% Subcomponente
TODOS	24	EXAMEN	35	EXAMEN	14 o 21 Diciembre	100
		CATEDRA	50	CATEDRA 1	12 Octubre	50
				CATEDRA 2	16 Noviembre	50
		EJERCICIO	15	EJERCICIO 1	07 Septiembre	25
				EJERCICIO 2	05 Octubre	25
				EJERCICIO 3	02 Noviembre	25
				EJERCICIO 4	30 Noviembre	25

Fechas tentativas de catedra y examen

TAREA: Crear una cuenta gratuita en google Colaboratory

<https://colab.research.google.com/notebooks/welcome.ipynb>



The screenshot shows the Google Colaboratory welcome interface. At the top, it says "Te damos la bienvenida a Colaboratory" with a navigation bar including "Archivo", "Editar", "Ver", "Insertar", "Entorno de ejecución", "Herramientas", and "Ayuda". On the left, there is a sidebar with a search icon and a list of items: "Introducción", "Ciencia de datos", "Aprendizaje automático", "Más recursos", "Ejemplos de aprendizaje automático", and "Sección". The main content area has a title "¿Qué es Colaboratory?" and a subtitle "Colaboratory, también llamado Colab, te permite escribir y ejecutar código de Python en un navegador, con las siguientes particularidades:". Below this, there is a bulleted list: "Sin configuración requerida", "Acceso gratuito a GPU", and "Facilidad para compartir". A paragraph follows: "Seas **estudiante, científico de datos o investigador de IA**, Colab facilita tu trabajo. Mira [este video introductorio sobre Colab](#) para obtener más información, o bien comienza a usarlo más abajo." Below this, there is a section titled "Introducción" with a paragraph: "El documento que estás leyendo no es una página web estática, sino un entorno interactivo denominado **notebook de Colab**, que permite escribir y ejecutar código." Another paragraph follows: "Por ejemplo, esta es una **celda de código** con una secuencia de comandos Python corta que calcula un valor, lo almacena en una variable y devuelve el resultado:". Below this, there is a code cell with the following code:

```
[ ] seconds_in_a_day = 24 * 60 * 60
seconds_in_a_day
```

 The output of the code is "86400". At the bottom, there is a paragraph: "A fin de ejecutar el código en la celda anterior, haz clic en él para seleccionarlo y luego presiona el botón de reproducción ubicado a la izquierda del código o usa la combinación de teclas 'Command/Ctrl + Enter'. Para editar el código, solo haz clic en la celda y comienza a editar."

Requerimientos para el curso

Instalacion de R y Rstudio

Ir a : <https://cran.r-project.org/>

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux \(Debian, Fedora/Redhat, Ubuntu\)](#)
- [Download R for macOS](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Al hacer clic sobre "Download R for Windows" seremos dirigidos a la página que se reproduce a continuación. Allí haremos clic sobre **install R for the first time**.



CRAN
Mirror
What's new?
Task Views
Search

About R
R Home page
The R Journal

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed

R for Windows

Subdirectories:

base Binaries for base distribution. This is what you want to **install R for the first time**. There is also information on [this page](#).
contrib Binaries of contributed CRAN packages (for R >= 2.10.0). ~~There is also information on this page.~~ There is also information on [this page](#).
old-contrib Binaries of contributed CRAN packages for outdated versions of R (for R < 2.13.0, managed by Uwe Ligges).
binutils Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

Please do not submit binaries to CRAN. Package developers might want to contact Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables.



CRAN
Mirror
What's new?
Task Views
Search

About R
R Home page
The R Journal

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed

R for Mac OS X

This directory contains binaries for a base distribution and packages to run on Mac OS X (release 10.6 and above). Mac OS 8.6 to 9.2 (and Mac OS X 10.1) are no longer supported but you can find the last supported release of R for these systems (which is R 1.7.1) [here](#). Releases for old Mac OS X systems (through Mac OS X 10.5) and PowerPC Macs can be found in the [old-contrib](#) directory.

Note: CRAN does not have Mac OS X systems and cannot check these binaries for viruses. Although we take precautions when assembling binaries, please use the normal precautions with downloaded executables.

Package binaries for R versions older than 3.2.0 are only available from the [CRAN archive](#), so users of such versions should adjust the CRAN mirror setting (<https://cran.r-project.org/>) accordingly.

R 4.0.3 "Bunny-Whisperer Freak Out" released on 2020/10/10

Please check the MD5 checksums of the downloaded image to ensure that it has not been tampered with or corrupted during the mirroring process. For example type:
openssl sha1 -x509 -i 4.0.3.pkg
in the Terminal application to print the SHA1 checksum for the R-4.0.3.pkg image. On Mac OS X 10.7 and later you can also validate the signature using:
pgpkey1 --check-signature R-4.0.3.pkg

Latest releases:

R 4.0.3.pkg (detected and signed by CRAN)

R 4.0.3 binary for macOS 10.11 (High Sierra) and higher, signed and notarized package. Contains R 4.0.3 framework, R app GUI 1.7.9 for 64-bit for Intel Macs, 10179 R 4.0.3 X11 libraries and Tkutils 4.7. The latter two components are optional and can be omitted when choosing "Custom install". They are only needed if you want to use the `utils` R package or build package documentation from sources.

Note: the use of X11 (including `x11`) requires `XQuartz` to be installed since it is no longer part of OS X. Always install XQuartz when upgrading your macOS to a new major version.

Important: this release uses Xcode 10.1 and GNU Fortran 8.2. If you wish to compile R packages from sources, you will need to download and GNU Fortran 8.2 - see the [utils](#) directory.

News features and changes in the R app Mac GUI

NEWS (for Mac GUI)

[Mac GUI 1.7.9 changelog](#)

Sources for the R app GUI 1.7.9 for Mac OS X. This file is only needed if you want to join the development of the GUI. It is not intended for regular users. Read the [INSTALL](#) file for further instructions.

Requerimientos para el curso

Instalar: Rstudio

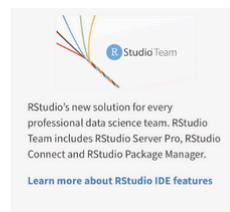
<https://www.rstudio.com/products/rstudio/download/>



Choose Your Version

The RStudio IDE is a set of integrated tools designed to help you be more productive with R. It includes a console, syntax-highlighting editor that supports direct code execution, and a variety of robust tools for plotting, viewing history, debugging and managing your workspace.

[LEARN MORE ABOUT RSTUDIO FEATURES](#)



	RStudio Desktop Open Source License	RStudio Desktop Pro Commercial License	RStudio Server Open Source License	RStudio Server Pro Commercial License
	Free	\$995 /year	Free	\$4,975 /year (5 Named Users)
	DOWNLOAD	BUY	DOWNLOAD	BUY
	Learn more	Learn more	Learn more	Evaluation Learn more
Integrated Tools for R	✓	✓	✓	✓
Priority Support		✓		✓
Access via Web Browser			✓	✓
RStudio Professional Drivers		✓		✓
Connect to RStudio		✓		

Requerimientos para el curso



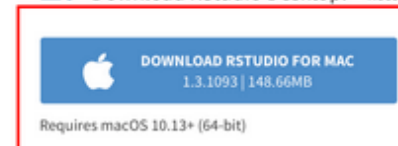
RStudio Desktop 1.3.1093 - [Release Notes](#)

1. Install R. RStudio requires R 3.0.1+.
2. Download RStudio Desktop. Recommended for your system:



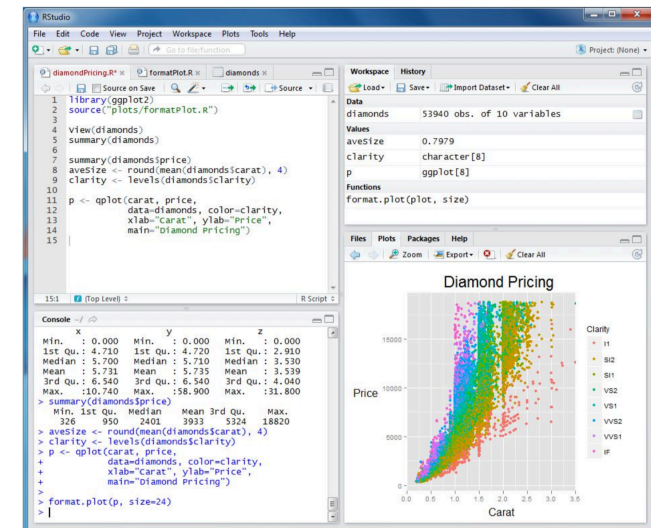
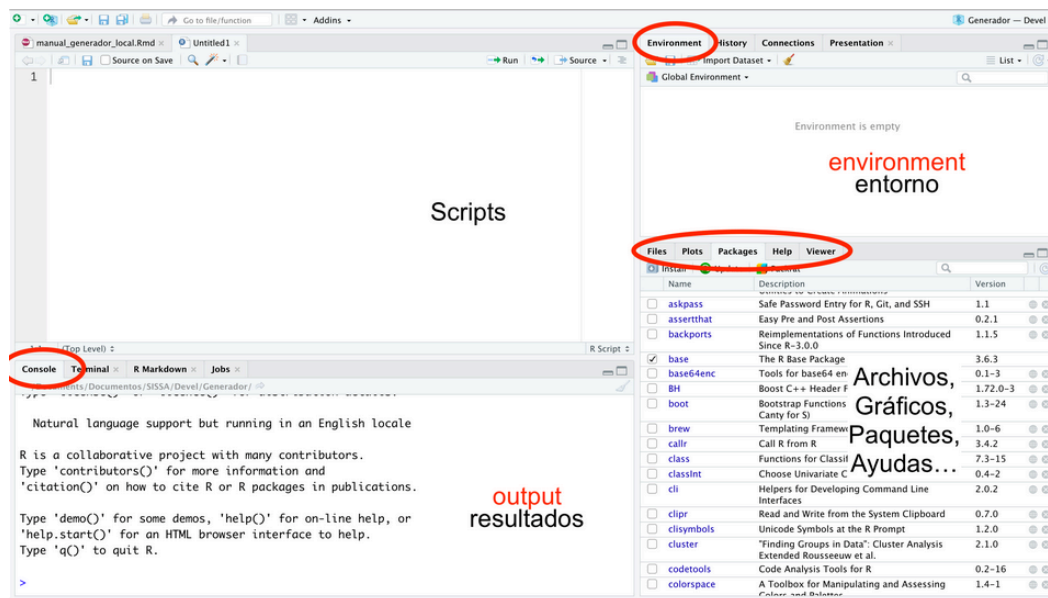
RStudio Desktop 1.3.1093 - [Release Notes](#)

1. Install R. RStudio requires R 3.0.1+.
2. Download RStudio Desktop. Recommended for your system:



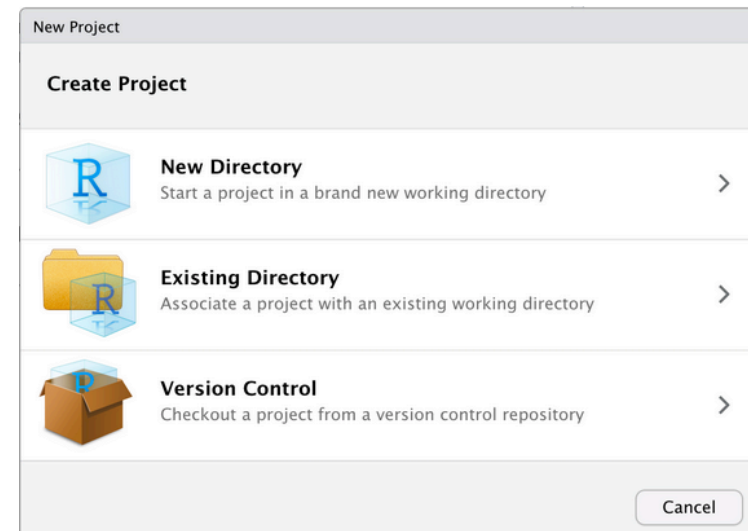
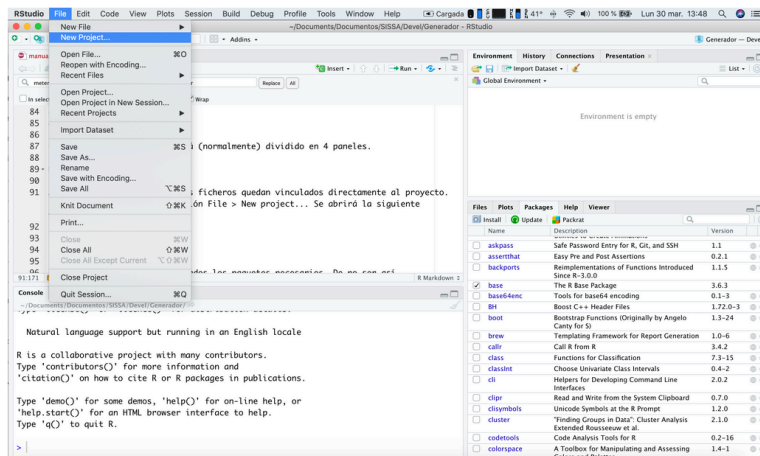
Requerimientos para el curso

Abrir RStudio



Requerimientos para el curso

Crear un proyecto



Requerimientos para el curso

Ejemplo de instalacion de paquetes

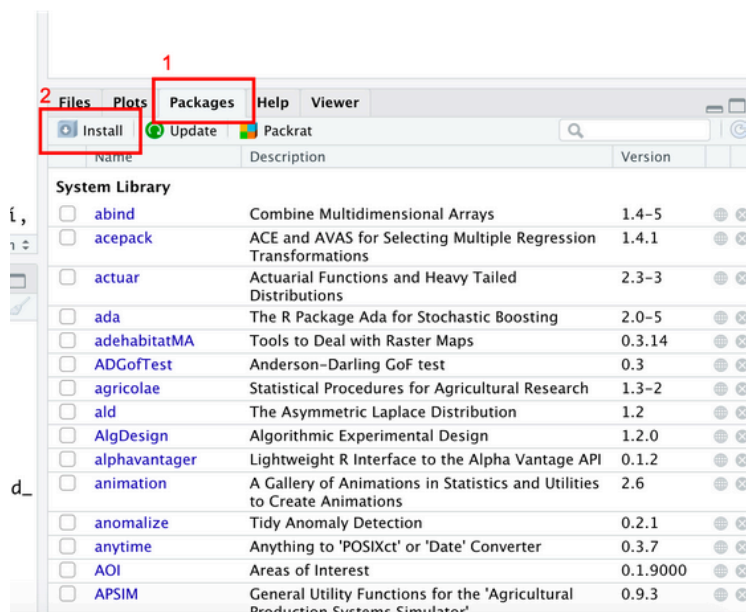
Instalación de paquetes con código

```
#Instalar paquete
install.packages("<nombre del paquete>")
# Ejemplo1
install.packages("ggplot2")
# Instlar mas de un paquete a la vez
install.packages(c("<nombre del paquete 1>", "<nombre del paquete 2>"))
# Ejemplo2
install.packages(c("ggplot2", "dplyr"))
```

Requerimientos para el curso

Instalar paquetes de forma visual en Rstudio

- 1.- hacer click en la pestaña **Packages** y seleccionar el paquete o buscarlo en la lupa
- 2.-Click en el botón **Install** .



También se puede tener una cuenta de RStudio Cloud

R Studio Cloud (<https://rstudio.cloud/>) es una plataforma web que permite crear, ejecutar y compartir proyectos implementados en lenguaje R sin necesidad de contar con ninguna infraestructura de hardware o software.

- 1) Ingrese a <https://rstudio.cloud/>
- 2) Hacer click en el link *Sign Up* que aparece en la parte superior derecha de la pantalla
- 3) Seleccione la opción *Cloud Free*
- 4) Haga click en el botón *Sign Up* que aparece en la parte inferior de la pantalla.
- 5) Complete el formulario y siga las instrucciones de registración.

PD1: Si se cuenta con una cuenta de *Google* o *GitHub*, se puede realizar una registración rápida utilizando alguna de estas cuentas (opciones *Sign Up with Google* o *Sign Up with GitHub*)

PD2: Cuentas gratuitas tienen ciertas limitaciones en cuanto a la cantidad de proyectos que puede crear y el tiempo disponible de ejecución sobre una base mensual --> Considerar para tareas y proyectos

Repaso de Python

Se repasarán los aspectos fundamentales de python con el objetivo que todos partan de un nivel parecido.

Recordar que existen conocimientos generales que son necesarios para programar en cualquier lenguaje de programación.

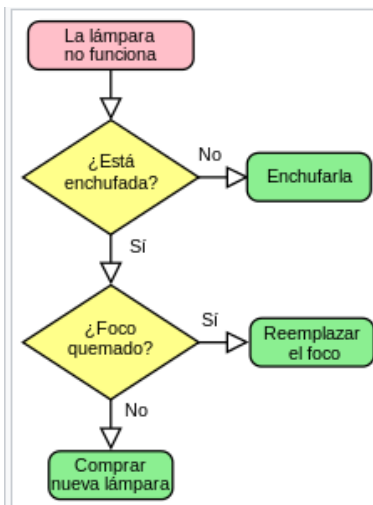
Se puede trabajar de forma local, en jupyter-notebook, jupyter lab, o google colab.

Para entender, debemos recordar que cuando programamos (ya sea una solución o un flujo de trabajo) estamos descomponiendo un problema o flujo en varias partes para luego implementar o desarrollar la solución por medio de código en un lenguaje de programación específico.

Repaso de Python

Algunas definiciones:

Algoritmo: Serie finita de pasos para resolver un problema. Conjunto de instrucciones o reglas definidas de una forma ordenada que permite solucionar un problema.



Los [diagramas de flujo](#) sirven para representar algoritmos de manera gráfica

Imagen extraída de wikipedia

Algunas definiciones:

Variable: Una **variable** es un sitio donde guardamos una determinada información. Es un contenedor de un valor, resultado, expresión, etc. El valor de una variable es modificable, es decir, puede cambiar a lo largo del código.

Terminal: es un programa que interpreta instrucciones o comandos definidos por el usuario. También se le llama consola.

Script: Es un programa simple y que se puede ejecutar desde la terminal.

Bug: Error en el script o programa.

Método/Función: Bloque de código con un nombre asociado que realiza una acción específica. Puede o no recibir argumentos como entrada.

Importar: Añadir código externo.

Pseudocódigo: Descripción de alto nivel. Es una forma de escribir los pasos que va a realizar un programa.

Refactorización: Reestructuración o optimización de código

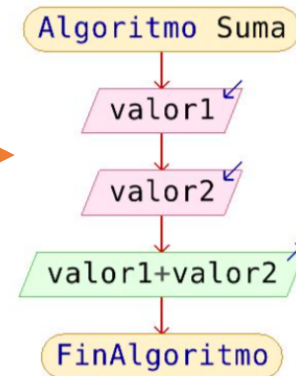
Indentación: Consiste en espacios vacíos que se dan antes de una instrucción para especificar que está dentro de un contexto.

Repaso de Python

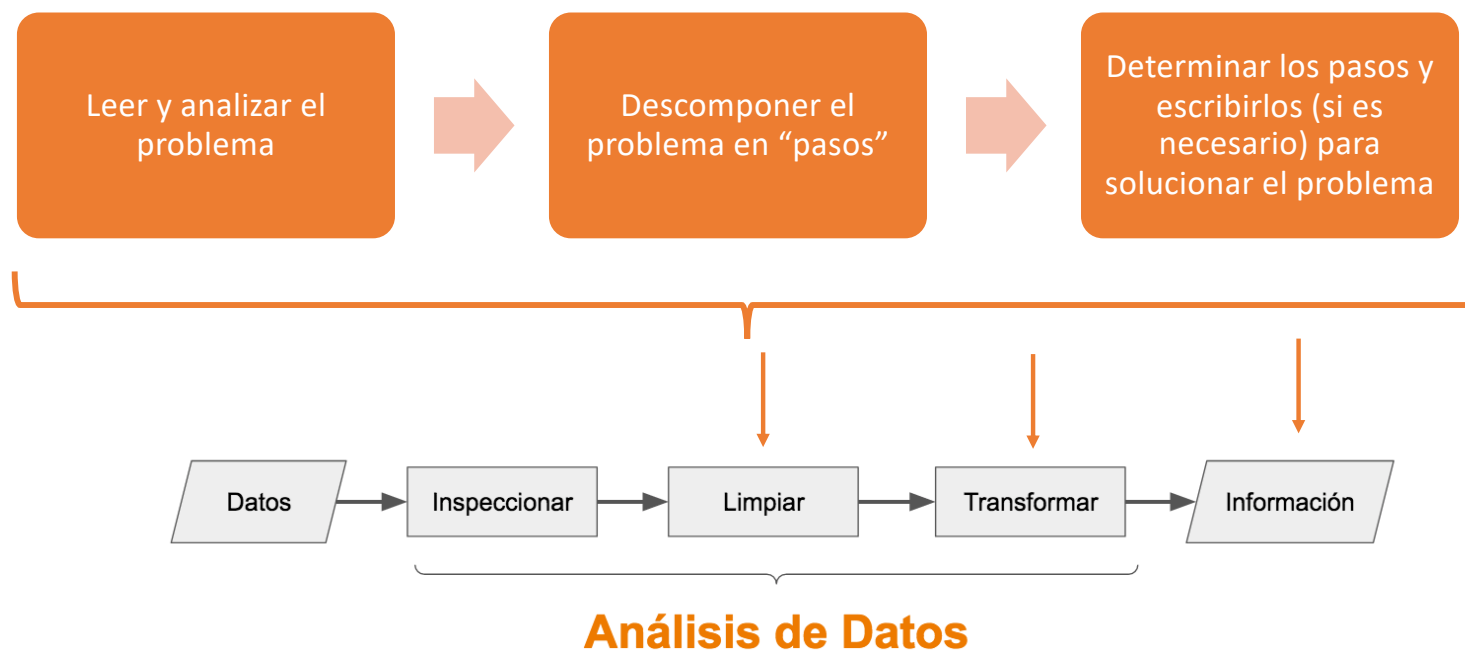
```
Algoritmo Suma  
  Leer valor1  
  Leer Valor2  
  Mostrar valor1 + valor2  
FinAlgoritmo
```

Pseudocódigo

Pseudocódigo a diagrama de Flujo



Pasos a seguir cuando programamos una solución



Gracias por su atención
Mail de contacto malegriaa@udla.cl