

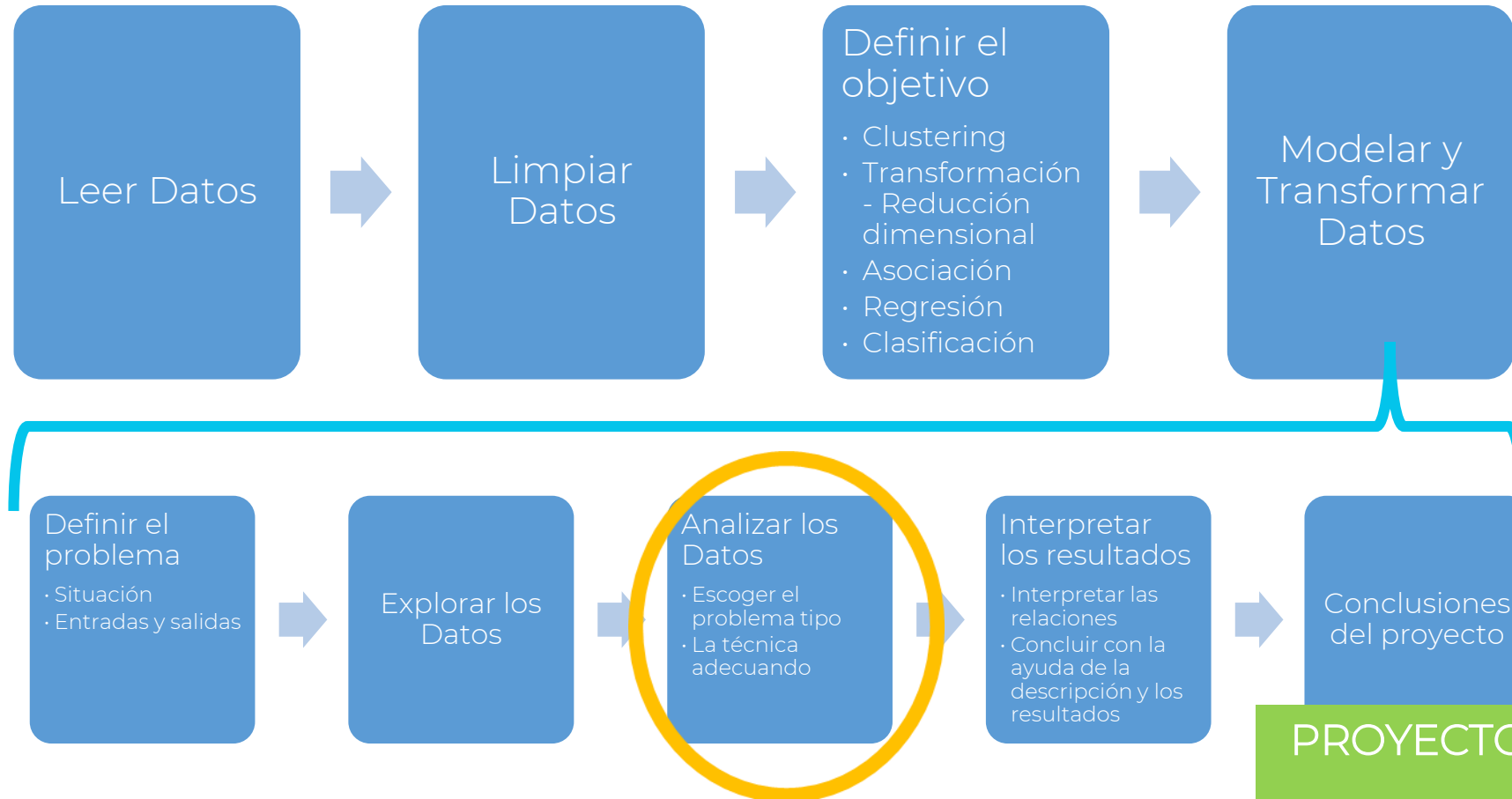
Técnicas de transformación de variables

Conoce las técnicas más relevantes para transformar los datos



Las etapas de un proyecto de ciencia de datos

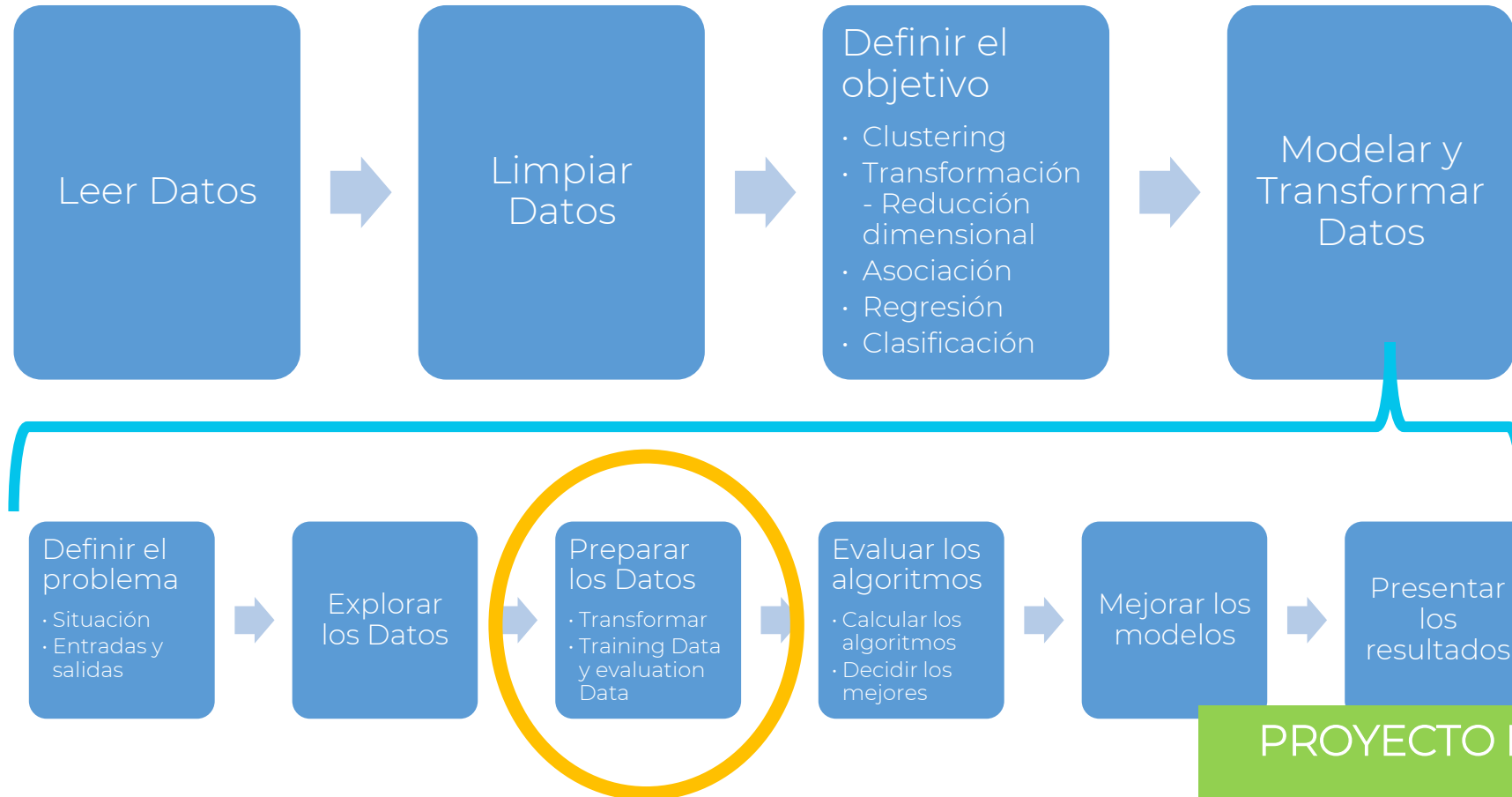
Instalación de R
Interfaz RStudio
Preparación
de Paquetes





Las etapas de un proyecto de ciencia de datos

Instalación de R
Interfaz RStudio
Preparación
de Paquetes



Las distintas transformaciones de datos

- Centrar – “center”
- Escalar – “scale”
- Estandarizar – “center” + “scale”

Mirar las variables desde un mismo origen
Quitar las unidades
Mismo origen y mismas unidades

- Normalizar – “range”

Variables de 0 a 1

- Box-Cox – “BoxCox”
- Yeo Johnson – “YeoJohnson”

Mejorar Normalidad
Mejorar los residuos de modelos
Mejorar correlación entre predictoras (entradas)

- PCA – “center” + “scale” + “pca”
- ICA – “center” + “scale” + “pca”

Cambiar a variables totalmente independientes
Cambiar a variables de una misma fuente

Centrar, escalar y estandarizar

- Centrar – “center”
- Escalar – “scale”
- Estandarizar – “center” + “scale”

Mirar las variables desde un mismo origen

Quitar las unidades

Mismo origen y mismas unidades

Normalizar

- Normalizar – “range”

Variables de 0 a 1

Corregiendo asimetría con Box-Cox y Yeo Johnson

- Box-Cox – “BoxCox”
- Yeo Johnson – “YeoJohnson”

Mejorar Normalidad

Mejorar los residuos de modelos

Mejorar correlación entre predictoras (entradas)

Las distintas transformaciones de datos

- PCA – “center” + “scale” + “pca”
- ICA – “center” + “scale” + “pca”

Cambiar a variables totalmente independientes
Cambiar a variables de una misma fuente

Take away

El resumen de la lección

Lo más importante de la lección

- La transformación de datos nos sirve para afinar el cálculo del modelo
- Nos sirve para mejorar las distribuciones a gaussianas y que puedas trabajar con modelos lineales generalizados
- Si quieres utilizar un modelo de forma interpretativa, con las transformaciones pierden sentido físico

Tú turno

A por las piezas de lego

Tú turno

- Descarga la hoja de trabajo de las transformaciones y practica con ellas. Te aconsejo que explores los datos antes y después para ver cómo han cambiado
- Utiliza las piezas de lego de la exploración que ya conoces 😊
- ¡Adelante!