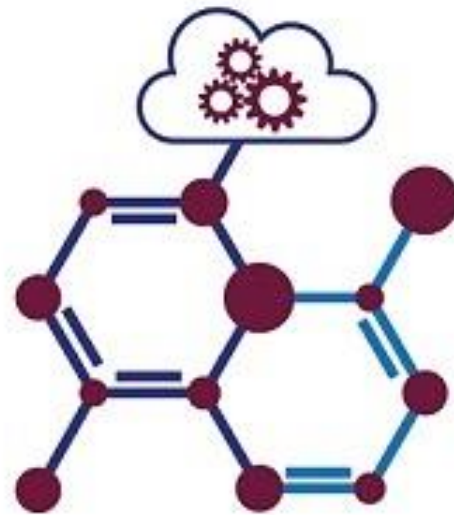




TECNOLÓGICO
NACIONAL DE MÉXICO®



CIENCIA DE
LOS DATOS

Diplomado en:

CIENCIA DE LOS DATOS E INTERNET DE LAS COSAS

VIRTUAL





MTI. Jose Gabriel Rodríguez Rivas, LI



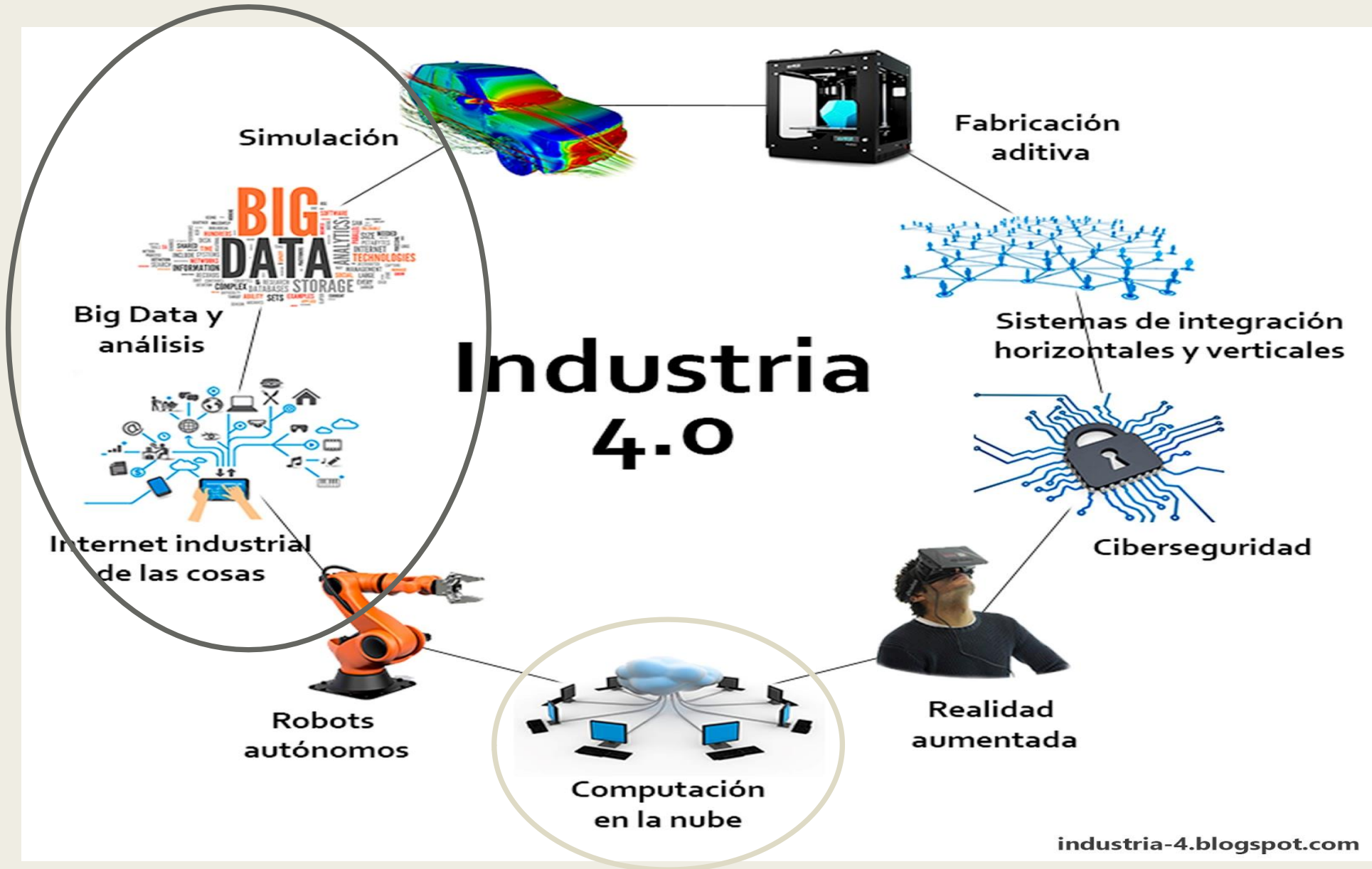
MTI. Marzo Antonio Rodríguez Zúñiga, ISC



MAI. Rubén Pizarro Gurrola, LI

Objetivo

- Desarrollar habilidades en el participante en torno al paradigma industrias 4.0 relacionadas con ciencia de los datos e internet de las cosas IoT con el uso de herramientas adecuadas.



Módulo 1. Introducción a Ciencia de los Datos

Objetivo

- Aprender y desarrollar habilidades de programación en lenguaje R, a través de *R Studio* para el análisis de **datos estructurados y semiestructurados** y su correcta interpretación y adecuada comunicación.

Módulo 1. Introducción a Ciencia de los Datos

Objetivos específicos

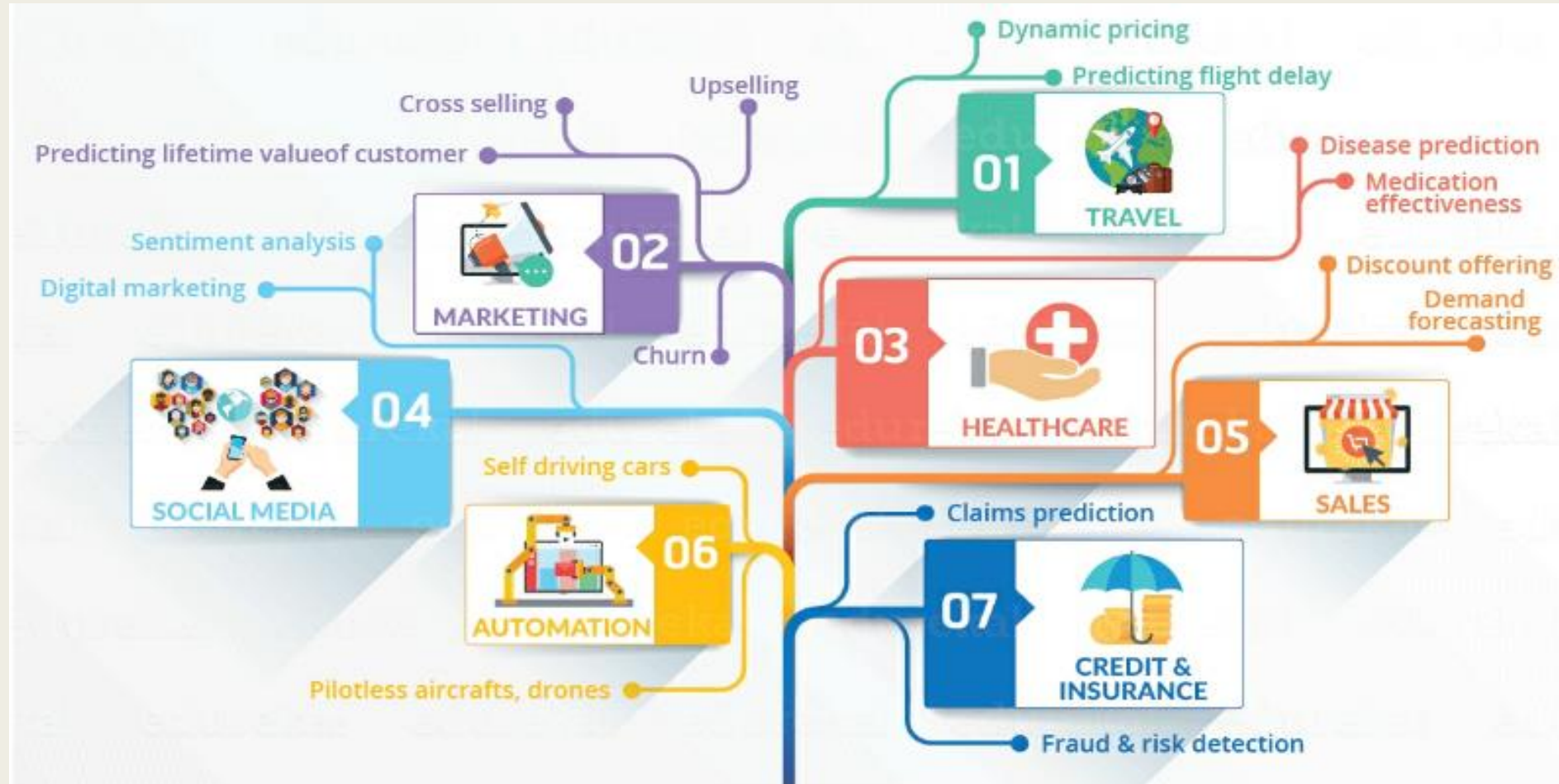
- Estado del arte de ciencia de los datos
- Configurar y organizar entorno de trabajo
- Programar en R
- Analizar datos
- *Construir documentos markdown*
- *Publicar en la nube por medio del servicio RPubs de R Studio.*
- Realizar aplicaciones WEB interactivas a través del servicio R Shiny

By Brendan Tierney, 2012



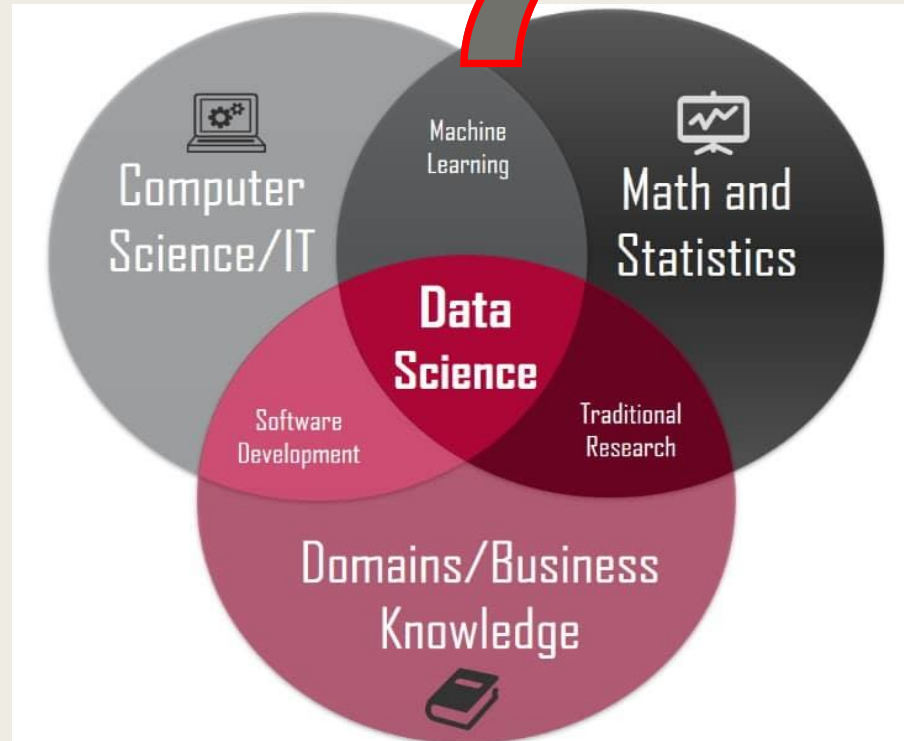
https://www.google.com/search?q=relacion+Big+data+con+BI,+Machine+Learning+y+Data+Mining&sxsrf=ALeKk03UMTsvGPFWnbBg-vNpxVOXzRJMfw:1590674376201&source=lnms&tbm=isch&sa=X&ved=2ahUKEwj-5Xp29bpAhVKS6wKHUARC4IQ_AUoAnoECAwQBA&biw=1600&bih=722#imgsrc=MrQ-5VzNJBf1jM&imgdii=FyF1fRvtIEbs9M

¿En dónde aplicar Ciencia de los Datos?



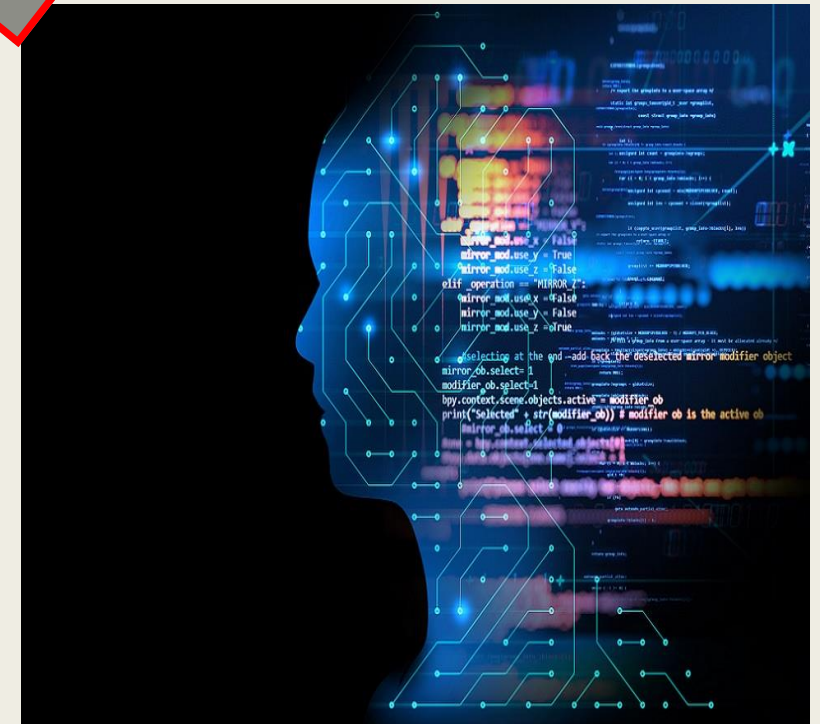
Fuente: Jones, Herbert (2019). Ciencia de los datos: Lo que saben los mejores científicos de datos sobre el análisis de datos, minería de datos, estadísticas, aprendizaje automático ... Data - que usted desconoce (Spanish Edition) .

¿Con qué se relaciona ciencia de los datos?



Fuente: google.com

Inteligencia artificial



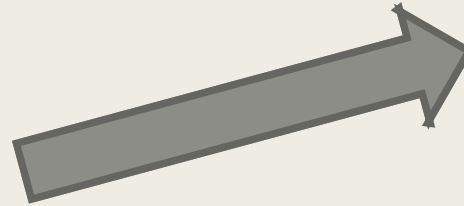
Se puede aplicar en cualquier dominio del conocimiento y/o ramas de ingeniería

*(por supuesto en las ciencias computacionales y TIC),
en cualquier campo laboral, consultoría, de investigación y docencia.*

¿Cómo trabajar?

Construir un documento-libro digital

Instrucciones e indicaciones



Desafíos o retos



Retroalimentación

Temas

Temas	Estimado 20 hrs.
1. Reconocimiento de entorno de trabajo	1
2. Programación R	1
3. Estructuras de datos en R	2
4. Funciones del usuario en R	2
5. Visualización de datos	1
6. Importar y exportar datos	1
7. Librería <i>dplyr</i>	2
8. Archivos markdown y publicación rpubs	1
9. Simulación de datos	1
10. Gráficos avanzados <code>ggplot2()</code> y <code>plotty()</code>	2
11. Strings and Dates en R	1
12. Datos semiestructurados	2
13. Series de tiempo	1
14. R shiny	2

1. Reconocimiento de entorno de trabajo

- Instalación y configuración de entorno
- Organización de carpetas
- Reconocimiento de R y R Studio
- **Desafío:** Mostrar evidencia de instalación de R Studio

2. Programación R

- Programación en Consola
 - *print()*, *paste()*, *cat()*, *scan()*
- Crear, ejecutar y almacenar scripts en R
- Creación de variables
- Operadores
- Expresiones
- Coerción = cast = conversión
- Condicionales y flujos de trabajo
- **Desafío:** Crear, abrir, ejecutar, almacenar e interpretar scripts en R y R Studio

3. Estructuras de datos en R

■ Vectores

- *Concatenar elemento con c()*
- *Expresiones y operaciones con vectores*
 - `length()`, `class()`, `which()`, `subset()`

■ Datos numéricos y datos categóricos

■ Factorizar datos categóricos

- *factor()*

■ Descripción de datos

- *min()*, *max()*, *range()*, *mean()*, *var()*, *sd*, *quantile()*, *summary()*, *head()* y *tail()*

■ Matrices

■ **Desafío:** Describir vectores y matrices

3. Estructuras de datos en R

- Data frames o conjuntos de datos
- Describir data frames
 - *str()*,
- Acceder a elementos de un data frame
 - *Por posición y por nombre*
- Utilizando \$ para acceder a columnas de un data.frame
- **Desafío:** Crear un script en donde se construye un data frame a partir de vectores y analizar los datos.

4. Funciones del usuario en R

- Funciones propias del lenguaje
- Programación de funciones en consola
- Crear y almacenar funciones como scripts en R
- Llamar funciones
- Almacenar funciones en la nube
- Llamar funciones de la nube
- **Desafío:** Crear, abrir, ejecutar, almacenar e interpretar funciones en R y R Studio

5. Visualización de datos

- Dispersión: `plot()`
 - *Dispersión, lineal*
- Diagrama de barra: `barplot()`
- Histograma: `hist()`
- Pastel: `pie()`
- Tallo y hoja: `Stem()`
- Diagrama de caja o bigotes: `boxplot()`
- **Desafío:** Crear un script en donde se construya data frame de datos de personas con datos categóricos y numéricos y realizar diversos diagramas

6. Importar y exportar datos en R

- Instalar librerías
 - *install.packages()*
- Cargar librerías
 - *library()*
- Importar datos tipo texto
- Importar datos formateados
- Importar datos excel
- Importar datos tipo JSON
- Exportar datos
- **Desafío:** Cargar, describir, analizar y visualizar datos

7. Librería *dplyr*

- Instalar librerías
 - *install.packages("dplyr")*
- Cargar librerías
 - *library(dplyr)*
- *filter()*
- *select()*
- pipes *%>%*
- *arrange()*
- *mutate()*
- *rename()*
- *summarize()* *group_by()*

8. Archivos markdown y publicación Rpubs

- Instalar librerías
 - *install.packages()*
- Cargar librerías
 - *library(dplyr)*
- Creación de archivos markdown
 - *Edición no visual*
 - *Edición visual*
- Ejecución parcial
- Depuración
- Construcción total del archivo markdown
- Publicación de archivo en servicio rpubs

Desafío: Publicar en servicio RPubS

9. Simulación

- `set.seed()` semilla de inicialización
- `rep()` Repeticiones
- `seq()` Secuencias
- `sample()` Muestras
- `rnorm()` Distribución normal

10. Gráficos avanzados `ggplot2()` y `plotty()`

- Instalar librerías
 - *`install.packages()`*
- Cargar librerías
 - *`ggplot()`*
 - *`plotty()`*

11. Strings and Dates en R

- Instalar librerías
 - *install.packages()*
- Cargar librerías
 - *library(dplyr)*
 - *library(lubridate)*
 - *library(tidytext)*
 - *library(tidyverse)*
- Datos strings y funciones
- Datos fecha y funciones
- **Desafío:** Aplicar funciones sobre datos string

12. Datos semiestructurados en R

- Instalar librerías
 - *install.packages(...)*
- Cargar librerías
 - *library(...)*
- Web Scrapig (barrido o recorrido de páginas)
- Nube de palabras
- Text mining
- Análisis de sentimientos
- Desafío

13. Series de tiempo en R

- Función `ts()`
- **Desafío:** Generar, visualizar e interpretar series de tiempo

14. R Shiny

- Instalar librerías
 - *install.packages()*
- Cargar librerías
 - *library()*
- Crear RShiny
- Publicar RShiny
- **Desafío:** Construir y publicar RShiny