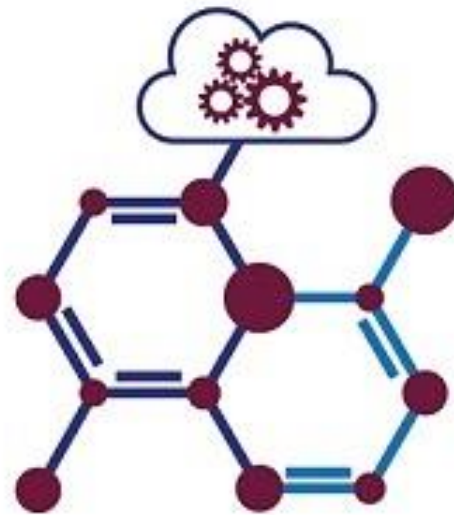




TECNOLÓGICO  
NACIONAL DE MÉXICO®



CIENCIA DE  
LOS DATOS



TECNOLÓGICO  
NACIONAL DE MÉXICO®



El **Centro de Innovación Tecnológica del ITD** en  
conjunto con la **Academia CISCO del Departamento  
de Sistemas y Computación del ITD** te invitan al

# Diplomado en: **CIENCIA DE LOS DATOS E INTERNET DE LAS COSAS**

## VIRTUAL



**INICIO:**  
**8 DE OCTUBRE 2021**

**TÉRMINO:**  
**19 DE FEBRERO 2022**

ALUMNOS DEL ITD PODRÁN OBTENER  
2 CRÉDITOS COMPLEMENTARIO

### **Horario:**

Viernes de 16:00 a 20:00 hrs.  
Sábados de 09:00 a 14:00 hrs.

**Duración:** 120 horas

PRE-REGISTRO:

### **INFORMES:**

Ing. Fátima del Carmen Perez Quiñonez  
Cel. 618 333 5865 / 618 807 0168





MTI. Jose Gabriel Rodríguez Rivas, LI



MTI. Marzo Antonio Rodríguez Zúñiga, ISC

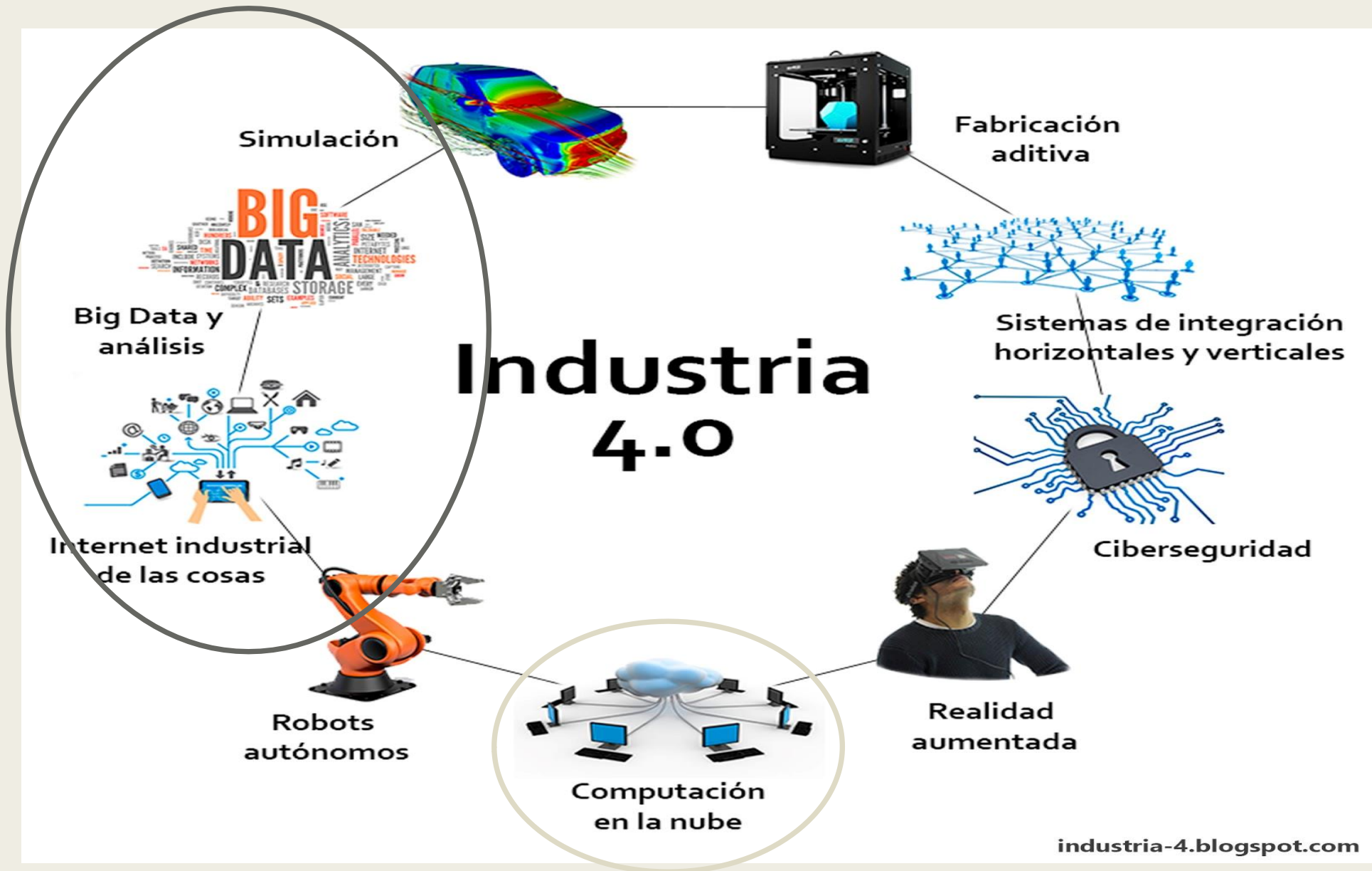


MAI. Rubén Pizarro Gurrola, LI

# 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.

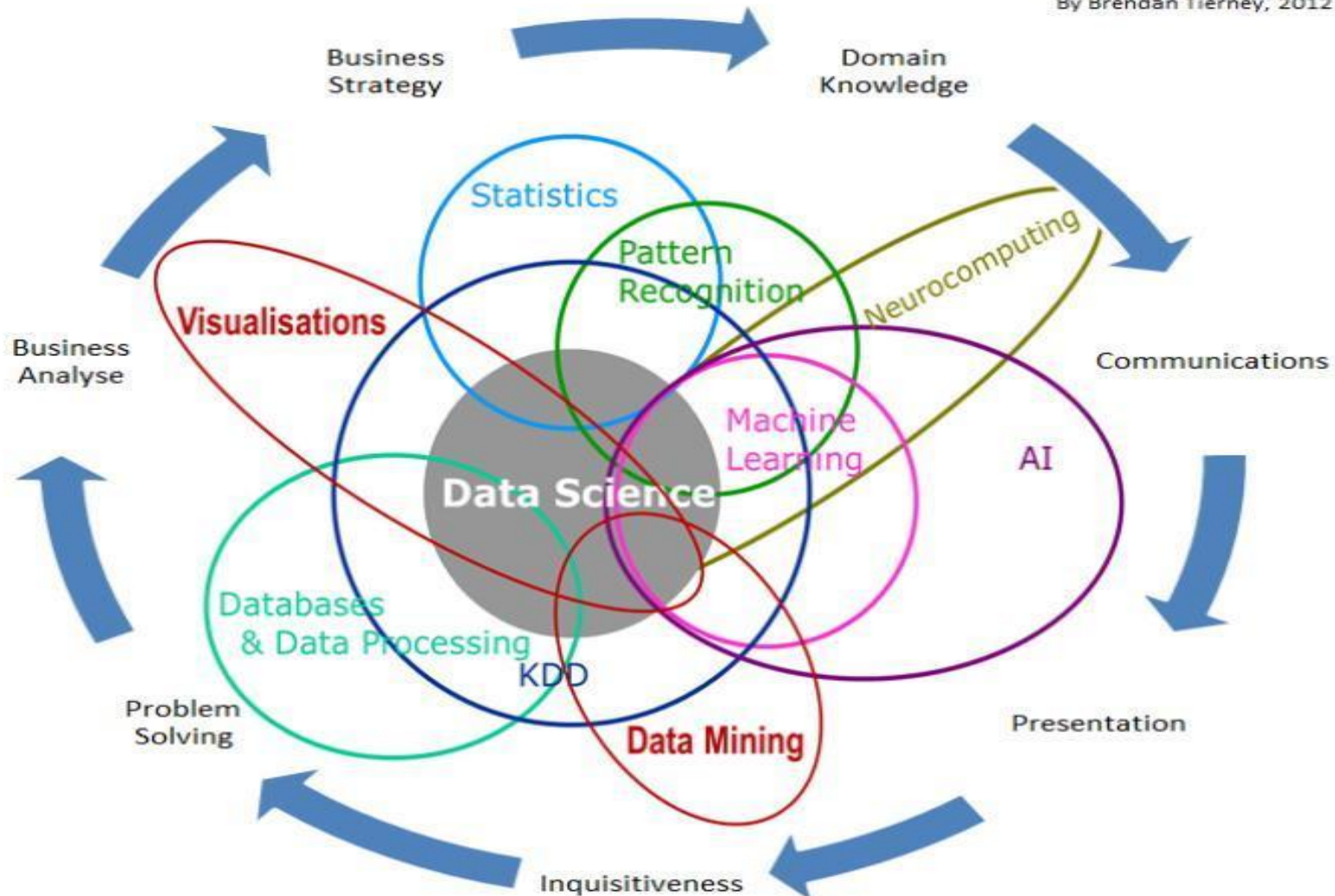
# Módulo 1. Introducción a Ciencia de los Datos

## 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

# Data Science Is Multidisciplinary

By Brendan Tierney, 2012

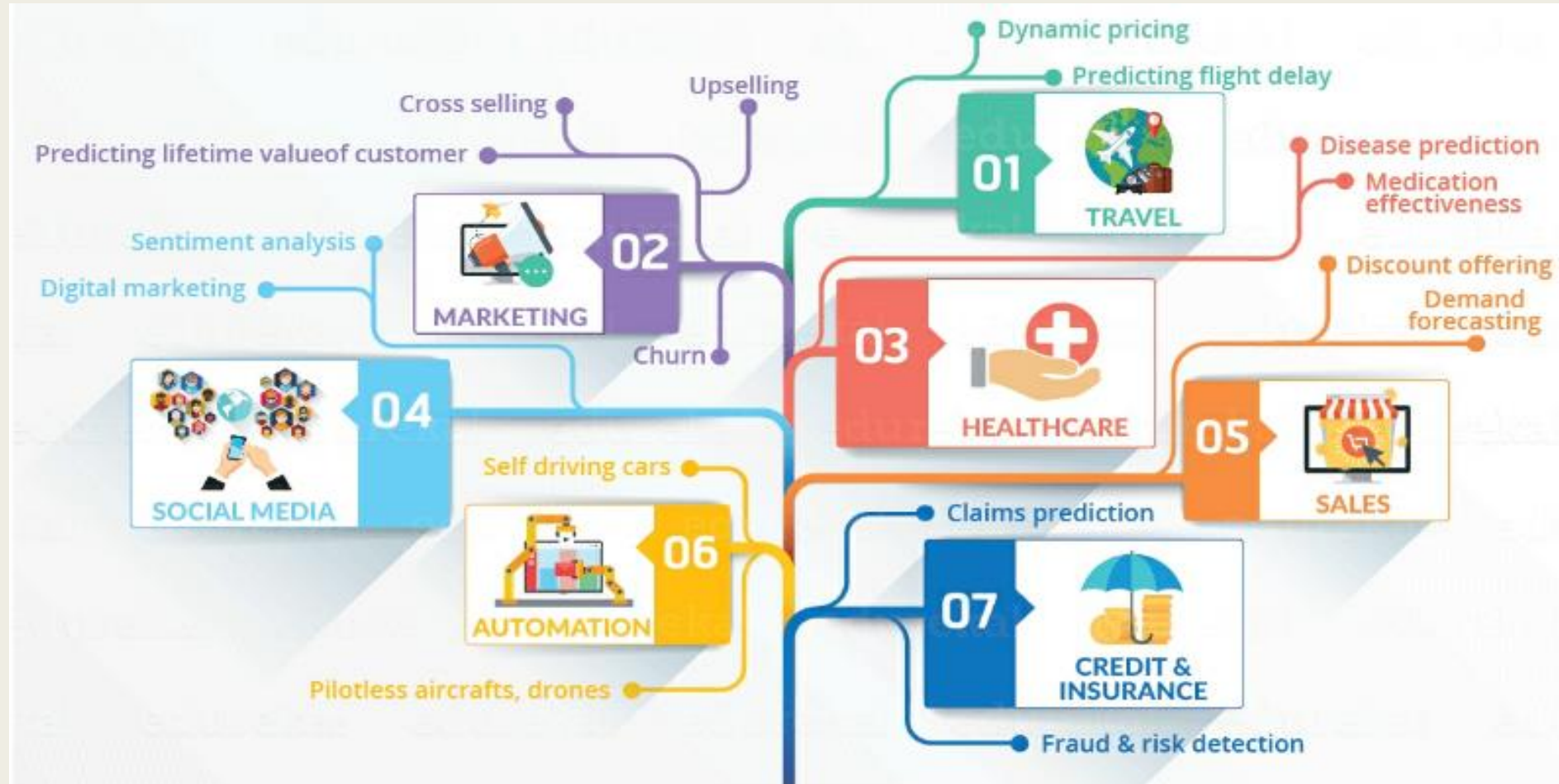


Fuente:

[https://www.google.com/search?q=relacion+Big+data+con+BI,+Machine+Learning+y+Data+Mining&sxsrf=ALeKk03UMTsvGPFWnbBg-vNpxVOXzRJMfw:1590674376201&source=lnms&tbn=isch&sa=X&ved=2ahUKewjr-5Xp29bpAhVKS6wKHUARC4IQ\\_AUoAnoECAwQBA&biw=1600&bih=722#imgsrc=MrQ-5VzNJBf1jM&imgdii=FyF1fRvtIEbs9M](https://www.google.com/search?q=relacion+Big+data+con+BI,+Machine+Learning+y+Data+Mining&sxsrf=ALeKk03UMTsvGPFWnbBg-vNpxVOXzRJMfw:1590674376201&source=lnms&tbn=isch&sa=X&ved=2ahUKewjr-5Xp29bpAhVKS6wKHUARC4IQ_AUoAnoECAwQBA&biw=1600&bih=722#imgsrc=MrQ-5VzNJBf1jM&imgdii=FyF1fRvtIEbs9M)

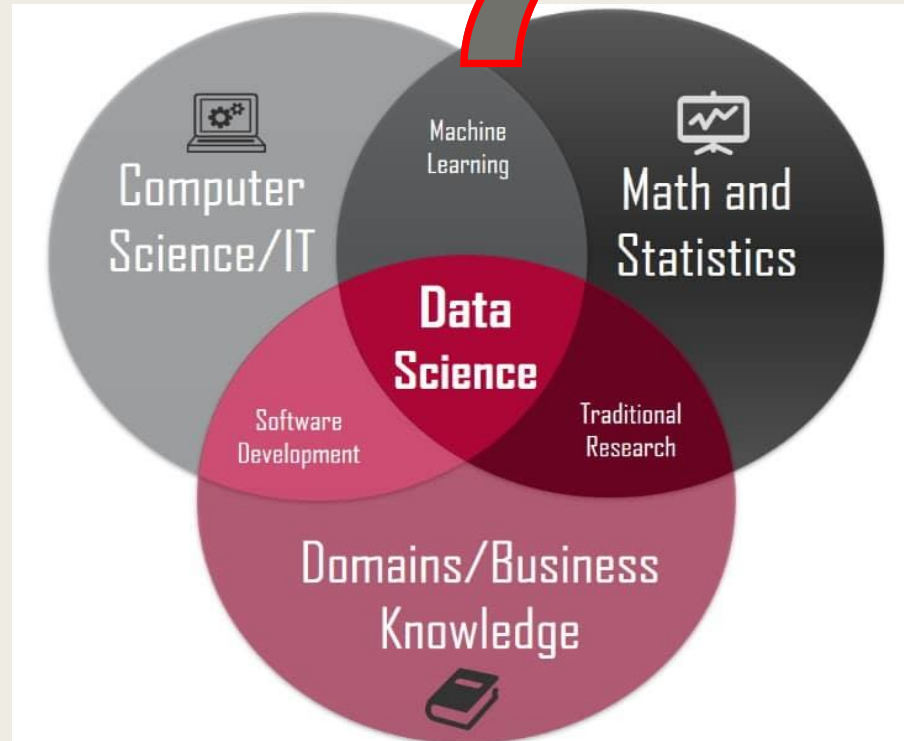


# ¿En dónde aplicar Ciencia de los Datos?



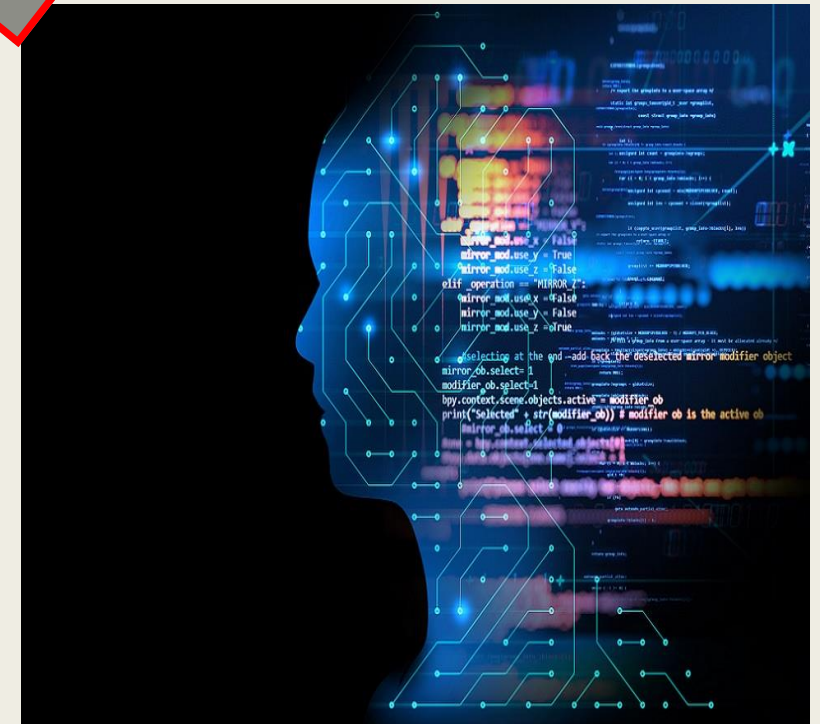
Fuente: Jones, Herbert (2019). Ciencia de los datos: Lo que saben los mejores científicos de datos sobre el análisis de datos, minería de datos, estadísticas, aprendizaje automático ... Data - que usted desconoce (Spanish Edition) .

# ¿Con qué se relaciona ciencia de los datos?



Fuente: google.com

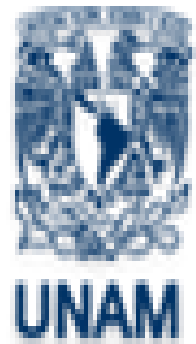
Inteligencia artificial



*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.*

# ¿En donde se estudia Ciencia de los Datos?



## Oferta Académica

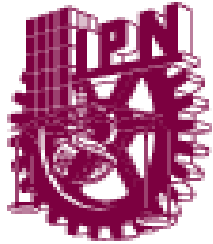
### Licenciatura en Ciencia de Datos

Asignaturas Optativas de Ciencia de Datos	
06 Aprendizaje de Máquina y Minería de Datos Avanzados	06 Temas Selectos de Ciencia de Datos
06 Bioinformática	06 Temas Selectos de Ciencia de Datos en Área Diversa
06 Ciencia Social Computacional	06 Temas Selectos de Ciencia de Datos en Mercadotecnia
06 Ciencia de Datos en Biología	06 Temas Selectos de Estadística
06 Estrategias de Portafolios de Inversión utilizando Ciencia de Datos	06 Temas Selectos de Finanzas Corporativas
06 Introducción a la Investigación Científica	06 Temas Selectos de Procesamiento de Lenguaje Natural
06 Minería de Textos	06 Temas Selectos de Sistemas de Información
06 Seguridad de la Información	06 Temas Selectos de Tópicos Especiales
06 Series de Tiempo	06 Temas Selectos de Visión Computacional
	06 Temas Selectos de Visualización

[http://oferta.unam.mx/planestudios/Ciencia-de-Datos-IIMAS\\_Plan%20de%20estudios19.pdf](http://oferta.unam.mx/planestudios/Ciencia-de-Datos-IIMAS_Plan%20de%20estudios19.pdf)



# ¿En donde se estudia Ciencia de los Datos?



## Licenciatura en Ciencia de Datos

8	Desarrollo de Habilidades sociales para la alta dirección	Gestión empresarial	Trabajo Terminal II	Estancia Profesional		
7	Big Data	Modelos econométricos	Trabajo Terminal I	Administración de proyectos de TI	Optativa C	Optativa D
6	Modelado predictivo	Procesamiento de Lenguaje Natural	Análisis de series de tiempo	Analítica avanzada de datos	Optativa A	Optativa B
5	Minería de datos	Matemáticas avanzadas para ciencia de datos	Procesos estocásticos	Aprendizaje de máquina e Inteligencia Artificial	Analítica y visualización de datos	Metodología de la investigación y divulgación científica
4	Desarrollo de aplicaciones Web	Cómputo de alto desempeño	Estadística	Bases de datos avanzadas	Desarrollo de aplicaciones para análisis de datos	Liderazgo Personal
3	Análisis y Diseño de Algoritmos	Programación para ciencia de datos	Probabilidad	Bases de datos	Métodos numéricos	Finanzas empresariales
2	Algoritmos y Estructuras de Datos	Álgebra Lineal	Cálculo Multivariable	Ética y Legalidad	Fundamentos económicos	
1	Fundamentos de Programación	Matemáticas Discretas	Cálculo	Comunicación Oral y Escrita	Introducción a la Ciencia de Datos	

<https://www.ipn.mx/oferta-educativa/educacion-superior/ver-carrera.html?lg=es&id=69>



# ¿En donde se estudia Ciencia de los Datos?

**Ingeniero en Ciencia de Datos y Matemáticas**



**Especialidad en Ciencia de Datos**

<https://www.ipn.mx/oferta-educativa/educacion-superior/ver-carrera.html?lg=es&id=69>

# Herramientas de programación para ciencia de los datos



# MÓDULO I. Programación R: Temas

Temas	Estimado 20 hrs.
<ol style="list-style-type: none"><li>1. Reconocimiento de entorno de trabajo</li><li>2. Programación R</li><li>3. Estructuras de datos en R</li><li>4. Funciones del usuario en R</li><li>5. Visualización de datos</li><li>6. Importar y exportar datos</li><li>7. Librería <i>dplyr</i></li><li>8. <i>Estadística descriptiva</i></li></ol>	15

# MÓDULO I. Programación R: Temas

Temas	Estimado 20 hrs.
<p>9. <i>Probabilidad y distribuciones normales</i></p> <p>10. Archivos markdown y publicación rpubs</p> <p>11. Simulación de datos</p> <p>12. Gráficos avanzados ggplot2() y plotty()</p> <p>13. Strings and Dates en R</p> <p>14. Datos semiestructurados</p> <p>15. Series de tiempo</p> <p>16. R shiny</p>	16



# MÓDULO I. Programación R: Temas

Temas	Estimado 20 hrs.
<p>Sesión 01. Ciencia de los datos, entorno R y R Studio</p> <ul style="list-style-type: none"><li>• Instalar R y R Studio</li><li>• Entorno R Studio</li><li>• Programación R en consola y construcción de scripts</li><li>• Tipos de datos, variables, función c()</li></ul>	2
<p>Sesión 02. Programación R</p> <ul style="list-style-type: none"><li>• Números, operaciones matemáticas y funciones matemáticas, strings, booleanos</li><li>• Operadores de asignación &lt;- y -&gt; =; operadores matemáticos; operadores lógicos</li><li>• Condicionales, Ciclos</li></ul>	2

# MÓDULO I. Programación R: Temas

Temas	Estimado 20 hrs.
<p>Sesión 03. Estructuras de datos en R, funciones</p> <ul style="list-style-type: none"><li>• Creación y llamada de funciones</li><li>• Estructuras de datos: vectores, listas, matrices, arreglos, data frames, factores</li><li>• Visualización de datos, plot(), bar(), ...</li></ul>	2
<p>Sesión 04. Importar y Exportar datos</p> <ul style="list-style-type: none"><li>• Importar y exportar datos read.csv(), writescv()</li><li>• Importar datos Excel</li><li>• Importar datos texto</li></ul>	2

# MÓDULO I. Programación R: Temas

Temas	Estimado 20 hrs.
<p>Sesión 05. Estadística descriptiva en R, comunicación de datos con archivos markdown y publicar en R Pubs</p> <ul style="list-style-type: none"><li>• Variable de interés</li><li>• Muestra con sample()</li><li>• Medidas de tendencia central mean(), median(), quantile(), table(),</li><li>• Medidas de dispersión var(), sd()</li><li>• Función summary()</li><li>• Frecuencias e histograma, funciones which(), subset() e hist()</li><li>• Tablas de frecuencias y su uso; librería fdth()</li><li>• Archivos markdown y publicar en R Pubs</li></ul>	2
<p>Sesión 06 Probabilidad y distribuciones de probabilidad</p> <ul style="list-style-type: none"><li>• Probabilidad</li><li>• Operaciones de conjuntos</li><li>• Variables aleatorias discretas y continuas</li><li>• Distribución Normal</li></ul>	2

# MÓDULO I. Programación R: Temas

Temas	Estimado 20 hrs.
<p>Sesión 07 Simulación y visualización de datos</p> <ul style="list-style-type: none"><li>• Librería dplyr</li><li>• Simulación de datos</li><li>• Gráficos avanzados ggplot2() y plotty()</li></ul>	2
<p>Sesión 08.</p> <ul style="list-style-type: none"><li>• Strings and Dates en R</li><li>• Datos semiestructurados</li></ul>	2



# MÓDULO I. Programación R: Temas

Temas	Estimado 20 hrs.
<p>Sesión 09</p> <ul style="list-style-type: none"><li>Series de tiempo</li></ul>	2
<p>Sesión 10</p> <ul style="list-style-type: none"><li>R Shinny</li></ul>	2

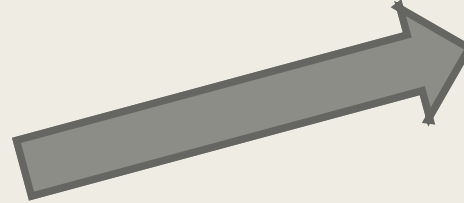
# R y R Studio



# ¿Cómo vamos a trabajar?

Construir evidencia digital

Instrucciones e indicaciones



Ejercicios para comunicar  
datos a través de  
scripts/markdown/rpubs



Retroalimentación

