

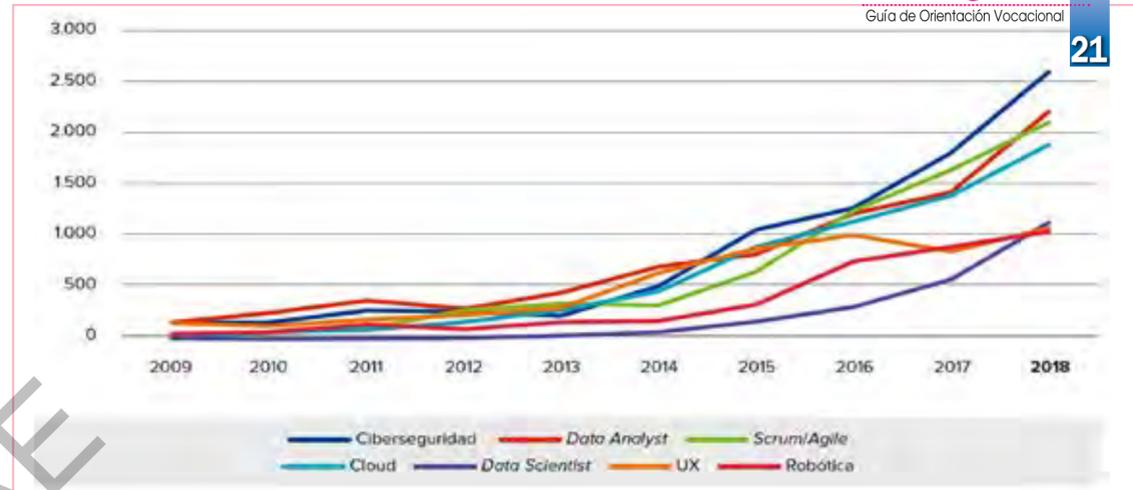
El Big Bang de la Analítica

Cuando el Big Data colisiona con las nuevas tecnologías, la analítica predictiva se hace imprescindible



datos, prácticamente cualquier organización podría obtener beneficios palpables. De hecho, se puede utilizar como una herramienta eficaz en múltiples escenarios: banca y finanzas, comercio, hotelería y turismo, industria, transporte, agricultura, logística, gobierno, educación, entretenimiento, campañas políticas y muchas más actividades humanas.

El actual Gobierno de la República, presentó en el 2018 la 'Estrategia de transformación digital hacia la Costa Rica del Bicentenario 4.0'. Dicha estrategia procura aprovechar las nuevas tecnologías y declara: "Los nuevos adelantos y la penetración de técnicas de inteligencia artificial, ciencia de datos, nanotecnología, biotecnología, y bioingeniería, y desarrollos como internet de las cosas, blockchain, 5G, sistemas de información geográficos y grandes volúmenes de datos, entre otros, están impulsando grandes transformaciones en los procesos productivos y en la forma en que las personas acceden a información y servicios. Esta estrategia considera que la apropiación de estas tecnologías por parte de las empresas, los ciudadanos y las instituciones públicas es central y primordial para el desarrollo del país." (GobiernoCR, 2018) (disponible en <https://micit.go.cr/transformaciondigitalcr/TransfDigitalCR.pdf>)



Puestos emergentes que siguen su senda creciente en España. Fuente: InfoJobs (ESADE), 2019.

res de Analítica de datos, Ciencia de datos e Ingeniería de datos, así como de las tecnologías de Bases de datos e Inteligencia de datos. Empresas como SmartSoft, Predisoft, SimMachines, y Singularities están desarrollando tecnología de clase mundial para la Analítica de datos y la Ciencia de datos. Prácticamente todos los bancos, operadoras de pensiones, aseguradoras y empresas de mediano a gran tamaño, están demandando profesionales en Análisis de datos.

Hasta hace poco tiempo, el perfil del Analista de datos incluía el conocimiento de temáticas como: depósitos de datos (data warehouses), mercados de datos (data marts), bases de datos multidimensionales, herramientas de reporte y visualización de datos, minería de datos, cuadros de mando integral, tableros de mandos (dashboards), sistemas de administración de relaciones con clientes (customer relationship management, CRM), análisis predictivo y administración del desempeño empresarial, entre otras. Las exigencias actuales para el perfil de analista de datos incluyen estudios en Estadística, Modelos matemáticos, aprovechamiento avanzado de SQL, los lenguajes de programación Python y R, y estudios en visualización de datos, entre otros.

De acuerdo con el documento de InfoJobs (parte de ESADE), titulado 'Estado del Mercado Laboral en España 2018' (InfoJobs (ESADE), 2019) presenta una tendencia al crecimiento de nuevos perfiles que precisan aplicar la inteligencia y la tecnología en el análisis y el uso de los datos. En la figura puede apreciarse la demanda creciente de Analistas de datos (Data Analyst).

Los Analistas de datos aportan su conocimiento de negocios, la comprensión de los resultados derivados del análisis y el modelado de los datos, así como la aplicación de los produc-

tos de datos que diseñen y construyan para la organización. El analista de datos puede extraer y analizar conjuntos de datos e identificar tendencias llamativas en ellos, así como profundizar en la indagación de anomalías. El analista debe ser capaz de resumir y presentar sus resultados de una manera clara para habilitar el trabajo de personas de la organización que no son especialistas en tecnología o en ciencias.

Estas son algunas de las tareas que usualmente realizaría un Analista de datos:

- Recopilar, procesar y realizar análisis estadísticos sobre los datos.
- Descubrir posibles usos de los datos para responder preguntas y resolver problemas.
- Limpiar y organizar datos en bruto (aún sin procesar).
- Emplear estadísticas descriptivas para obtener una vista general de sus datos.
- Analizar tendencias llamativas encontradas en los datos.
- Crear visualizaciones y tableros para habilitar la interpretación y la toma de decisiones basadas en datos.
- Presentar los resultados de los análisis a clientes o equipos internos.

A pesar del potencial de la creciente aplicación de los avances científicos y tecnológicos, las oportunidades de aprendizaje y enseñanza no han ido al mismo ritmo del mercado. Son pocas las carreras de Educación Superior que presentan carreras que permitan el estudio profundo del Análisis de datos.

En Costa Rica, la carrera pionera en esta materia es la **Maestría en Tecnología de Bases de Datos de la Universidad Cenfotec**, que contiene 7 cursos alineados con aspectos analíticos y de 'inteligencia de negocios', además de otros 8 cursos avanzados

en Bases de datos. Esta carrera se ofrece en modalidad presencial y virtual. U. Cenfotec también ofrece un **Posgrado denominado 'Analista de Datos'** (presencial), compuesto por cuatro cursos intensivos, así como el **Certificado de Posgrado 'Data Analytics & Big Data' (virtual)** - en alianza con XTOL - con un original método de aprender-haciendo apoyado por mentores en-línea.

La **Licenciatura en Administración de Tecnologías de Información del TEC** incluye cursos sobre Bases de datos e Inteligencia de negocios. **El Bachillerato en Ingeniería de Ciencia de Datos de ULEAD** ofrece cinco cursos relacionados con la Analítica de datos, en el marco de una carrera con buenos fundamentos de Informática y bases en Administración de **un programa de capacitación denominado Ciencias de los Datos**, constituido por cinco cursos avanzados, de nivel de postgrado, que abordan el Aprendizaje automático, el procesamiento de Datos masivos, la Inteligencia de negocios y la Minería de datos, además de cubrir la fundamentación matemática y estadística necesarias.

Es de esperar que pronto veamos ofertas académicas nuevas en el mercado, dada la enorme demanda profesional que hay en los niveles nacional e internacional.

Agradecimiento

Ing. Marco A. Hernández Vásquez, M.Sc., es Arquitecto de Datos en BAC | Credomatic, Profesor de la Maestría en Tecnología de Bases de Datos de la Universidad Cenfotec y Profesor de Ingeniería en Computadores en el Tecnológico de Costa Rica.

Ing. Ignacio Trejos Zelaya, M.Sc., es profesor de Informática en el Tecnológico de Costa Rica y en la Universidad Cenfotec. Su investigación se centra en Lenguajes de programación, Ingeniería del software y Educación en Informática.

Análisis de datos y Analítica de datos

Transformación Digital

Desarrollar los mecanismos que permitan generar, interoperar, analizar, sintetizar y producir conocimiento de valor para la toma de decisiones y la gestión de riesgos a partir de los datos e información pública en manos de las instituciones del sector público costarricense.



Línea de acción para promover la ciencia de datos en la toma de decisiones. Fuente: Gobierno de Costa Rica, 2018.

servicios. El **análisis de datos** procura obtener información a partir de los datos: organizarlos, relacionarlos, ordenarlos, clasificarlos y visualizarlos para sacar sentido de los datos 'crudos'. Las herramientas informáticas ayudan a realizar este trabajo, pero la interpretación es humana.

La Analítica de datos agrupa diversos y numerosos elementos. Se basa en técnicas, tecnologías, métodos y teorías provenientes de varios campos de estudio: Matemática, Estadística, Reconocimiento de patrones, Aprendizaje automático, Bases de datos, Al-

goritmos, Graficación computacional, Interacción humanos-computadoras, Computación avanzada, Visualización de información, Modelaje de la incertidumbre y Simulación, entre otros. Todo lo anterior debe ser conjuntado para extraer un sentido de los datos y, con base ello, crear productos de datos - que facilitarán la obtención de información estratégica y gerencial que pueda ser desplegada y difundida entre usuarios y asociados de negocios.

Los sistemas de Analítica de datos, diseñados por los Analistas de datos, permiten aprovechar los datos

almacenados en los sistemas de información transaccionales mediante la aplicación de herramientas informáticas que permiten explorar, conocer y descubrir información de la empresa, produciéndole conocimiento útil, una mejor comprensión de su contexto de negocios y apoyar sus procesos de toma de decisiones. Los sistemas analíticos proveen vistas históricas, actuales o predictivas sobre las operaciones del negocio, que pueden relacionarse con su entorno competitivo.

Debido a la alta gama de posibilidades para el estudio y el análisis de

En sus orígenes hace poco más de seis décadas, la Informática empresarial era denominada 'Procesamiento de datos'. La tecnología informática ha sido utilizada desde entonces para procesar los datos operativos de las organizaciones. La Internet, las computadoras personales, los dispositivos móviles y la creciente digitalización de los negocios generan enormes volúmenes de datos, que hacen crítica la necesidad de almacenar, administrar, extraer y analizar los datos eficiente y eficazmente.

En la actualidad se reconoce que los datos representan uno de los activos más significativos de cualquier empresa, debido a su criticidad, importancia y aplicabilidad en cualquier campo. Los datos provienen de bases de datos, de la Web, se obtienen a través de millones de teléfonos, por medio de sensores, sistemas de pago, cámaras y una gran variedad de otras fuentes. El valor de los datos está dado por su uso final. Todo lo anterior indica una escalada cambios aún mayor, a medida que técnicas de Inteligencia Artificial alcanzan la madurez (como el aprendizaje automático y el aprendizaje profundo), dando a las computadoras capacidades sin precedentes para emular actividades humanas como resolver problemas, revisar grandes volúmenes de datos, analizar datos complejos y comprender el lenguaje natural.

La **Analítica de datos** es una disciplina que busca convertir los grandes volúmenes de datos de las organizaciones en conocimiento útil y estratégico para tomar decisiones adecuadas y oportunas, que permitan mejorar los procesos, productos y