

PROGRAMA DIPLOMADO DATA SCIENCE

Objetivo general: Los estudiantes de este diplomado obtendrán herramientas analíticas y prácticas que permitirán conocer y ejecutar acciones en todo el proceso de Data Science, desde la adquisición y limpieza de datos hasta la creación de modelos predictivos.

Objetivos específicos

1. Entregar competencias que permitan sintetizar problemáticas, abstraer necesidades, comunicar resultados, y entender un conjunto de datos que usualmente pertenece a otra disciplina.
2. Entregar conceptos relevantes en Big Data, Visualización, Cloud Computing, Machine Learning, entre otros.
3. Fortalecer habilidades comunicacionales para relatar de una manera asertiva los hallazgos obtenidos del análisis de los datos, que favorecerá respecto de como las empresas toman ventaja de sus datos.

Contenidos

Data Driven Organization (25 horas teóricas)

Este curso introduce a los alumnos a las llamadas data driven companies que centran su estrategia de toma de decisiones en el uso de datos y Data Science. Mostrando cómo sus técnicas permiten descubrir, enriquecer y responder a las preguntas que enfrentan tanto los tomadores de decisiones como en los distintos niveles de la compañía en una variedad de industrias y organizaciones. Los estudiantes explorarán cómo los individuos y las organizaciones evalúan las opciones, toman decisiones, y cómo estas influyen las decisiones tácticas y estratégicas. El curso incluirá charlas de invitados, lecturas de casos, discusiones para que los alumnos amplíen su forma de ver las empresas y como la nueva economía de datos las impacta, aprender a realizar las preguntas apropiadas, a interpretar los resultados y transmitir los hallazgos a distintas audiencias. Se explorarán casos de áreas como el retail, banca, telecomunicaciones, energía, salud y los medios de comunicación/entretenimiento.

Python Para Análisis de Datos (30 horas teóricas)

Este curso consiste en una introducción al lenguaje de programación Python. El curso instruye en el lenguaje de programación poniendo especial énfasis en los módulos y paquetes de Python para realizar análisis de datos. En este curso los alumnos también aprenderán a usar otras herramientas básicas de Data Science como el uso de la línea de comandos, notebooks de Jupyter y control de versiones con Git y GitHub.

Estadística para Data Science (20 horas teóricas)

El objetivo de este curso es proporcionar a los estudiantes una introducción a métodos de investigación cuantitativa y técnicas estadísticas para el análisis de datos. El contenido se enseña con un enfoque práctico, orientado a usar métodos estadísticos para analizar problemas de datos del mundo real. Entre las temáticas a cubrir se encuentran: estadísticas descriptivas e inferenciales, muestreo, diseño experimental, pruebas paramétricas y no paramétricas de diferencia, regresión de mínimos cuadrados ordinarios y regresión logística.

Análisis y Limpieza de Datos (25 horas teóricas)

Este curso presentará los conceptos básicos de pre-procesamiento y análisis descriptivo de datos. El objetivo principal es poder determinar cuáles datos son susceptibles de ser convertidos en información para apoyar la toma de decisiones, y separar el ruido de la señal. Es el primer paso en un proyecto de ciencia de datos. Los estudiantes aprenderán a identificar las problemáticas que presentan los datos desde el momento de su registro (por ej., error muestral, outliers), así como usar las herramientas necesarias para describirlos (por ej., distribuciones e histogramas), explorarlos (por ej., agrupar o filtrar bajo un criterio específico), y cruzarlos (por ej., utilizando otras fuentes). Asimismo, los estudiantes comprenderán que las etapas de este proceso no son lineales, sino que se benefician del diseño iterativo.

Machine Learning (25 horas teóricas)

Este curso presentará herramientas de aprendizaje automático para generar valor a partir de los datos. Estas herramientas incluyen modelos de clasificación y predicción de fenómenos (por ej, RandomForest), descubrimiento de patrones de manera supervisada (por ej., SVM, redes neuronales) y no supervisada (por ej., Hierarchical Clustering), reducción dimensional, entre otros.

Visualización de la Información (25 horas teóricas)

Este curso introduce una visión global y práctica de las temáticas, herramientas y técnicas principales de Visualización de Información. Los estudiantes, primero, adquirirán capacidades para realizar análisis visual exploratorio, con el fin de responder preguntas de primer orden sobre los datos, y de generar preguntas de segundo orden sobre éstos. Luego, adquirirán capacidades para convertirlos en información interpretable que apoye la toma de decisiones a través de la comunicación visual. Finalmente se fomentará el espíritu crítico, de modo que puedan interpretar, cuestionar, validar y mejorar la comunicación de información por otras entidades.

<i>Duración (hrs.)</i>	150 hrs. académicas – 100 horas Cronologicas
<i>Nº Sesiones TBD</i>	<i>Modalidad Plataforma web y clases presenciales</i>
<i>Lugar:</i>	<i>Centro de Eventos Arrayan / Online – In House</i>
<i>Costos:</i>	El programa de Diplomado Data Science posee un valor de \$2.550.000. Mediante el sistema de Pre Registro y confirma tu asistencia tras pago vía transferencia a la cuenta vista: BANCO ESTADO N° 025-7-196145-1 a nombre de Miningport Consultores Ltda - Rut 76.378.566-1 enviando el comprobante del depósito al mail contacto@miningport.cl
<i>Incluye:</i>	<i>Clase presencial al inicio de cada módulo correspondiente al programa Data Science incluido</i> <i>Almuerzo</i> <i>Coffe Break.</i> <i>Plataformas Webinar – E-learning</i>

Las clases serán semi presenciales impartidas en multi plataformas online, siendo tanto Webinar y E-learning. Cada uno de los módulos en su inicio considera sesiones presenciales en la ciudad de Antofagasta.