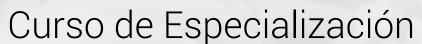


Ayudamos a cambiar historias de vida



# **MUJERES EN TECH II**

















Social N mación



Mejor Programa Educativo para la Inserción Profesional en Programación y Tecnología

# Contenido del curso

### 1. Git, Github & Gitflow

**Objetivo del módulo:** Manejar correctamente git, tomar la costumbre de guardar versiones y aplicar la metodología gitflow durante todos los desarrollos del curso.

**Módulo:** Sesiones a distancia programadas

Duración: 2 sesiones de 4 horas

- · Qué es git y porqué lo necesitamos
- Instalación de git
- Entendiendo git: las tres zonas
- · Guardando el trabajo. git commit, add, rm, log
- Navegando hacia atrás y adelante. git reset, checkout, diff, tags, reflog, referenciando nodos
- · Ramas, merges y conflictos
- · Repositorios remotos
- Presentación de gitflow

#### 2. Data 101

**Objetivo del módulo:** Visión "big picture" del mundo de Big Data: qué problemas resuelve, qué herramientas resuelve y en qué se distingue de los "approaches" anteriores. Casos de uso.

Módulo: Sesión a distancia puntual

Duración: 4 horas

#### 3. Protección de datos

**Objetivo del módulo:** Masterclass sobre el ciclo de vida del dato y la protección del mismo.

Módulo: Sesión a distancia puntual

Duración: 2 horas

- Master Data Management (MDM)
- Procedencia de los datos
- · Gestión del ciclo de vida de los datos
- Seguridad y gestión de acceso
- Privacidad de los datos
- Cuestiones éticas y legales

#### 4. Modelado de Datos & SQL

**Objetivo del módulo:** Aprender y practicar modelado de datos. Aprender SQL.

**Módulo:** Sesiones a distancia programadas

**Duración:** 3 sesiones de 4 horas

#### **Modelado de Datos**

- Modelo Entidad-Relación
- Formas normales

#### Lenguaje SQL

- DDL (lenguaje de definición de datos):
  Créate, alter y drop para tablas, vistas, bbdd, añadir columnas a tablas. Tipos de datos y más.
- DML (lenguaje de manipulación de datos): Insert, Update, Select, Delete
- DCL (lenguaje de control de datos): Otorgar y eliminar permisos (GRANT, REVOKE)



# Contenido del curso

## 5. SQL Avanzado, ETL & Data Warehouse

**Objetivo del módulo:** Diseño y construcción de un Data Warehouse mediante lenguaje SQL en BigQuery. Compresión y diferenciación entre procesos ETL y ELT.

Módulo: Sesiones a distancia programadas

Duración: 6 sesiones de 4 horas

#### **SQL Avanzado**

- Subquerys
- CTEs
- · Funciones condicionales
- · Funciones de agregación
- · Funciones de numeración
- Funciones de conversión
- Funciones matemáticas
- Funciones de navegación
- Funciones de string
- Funciones de fecha
- Funciones de JSON
- Creación de nuevas funciones

#### Introducción a BigQuery

- · Tablas, vistas
- Datasets
- Funciones
- Limpieza de tablas
- Partición y clusterización de tablas
- · Tablas Wildcard
- Tablas externas

#### **Data Warehousing**

- Definición y diferenciación entre Data Warehouse, Data Lake y Data Mart.
- Definición y diferenciación entre ETL y ELT.
- Definición y diferenciación entre bases de datos OLTP y OLAP.
- Herramientas ETL existentes en el mercado.

Construcción de un modelo de datos en BigQuery con un proceso ELT mediante lenguaje SQL

## 6. Exploración y Visualización de datos

**Objetivo del módulo:** Profundizar en el análisis de los datos mediante el *reporting* mediante Tableau. El análisis visual es el objetivo final en un proyecto de Big data ya que a través de este, se puede proporcionar información muy útil para una empresa como respuesta a los problemas de negocio con un aspecto personalizado, ágil e inmediato.

**Módulo:** Sesiones a distancia programadas **Duración:** 6 sesiones de 4 horas

- Visión general del business intelligence.
- Profundización en el análisis de los datos a través de indicadores claves o KPIs.
- Ver los diferentes elementos de una visualización.
- Uso de Tableau Desktop avanzado mediante ejercicios prácticos para aprender a usar las diferentes conexiones, filtrados, tipos de datos, análisis estadísticos, formulación, creación de dashboards, historias, etc.
- Visualización de relaciones entre valores numéricos (diagramas de dispersión, regresión y clustering) de tipos de datos específicos como geográficos o mapas de calor, visualización de distribuciones (histogramas y diagrama de cajas), etc.
- Análisis visual para poder saber cómo construir los gráficos más adecuados para el usuario final y así mejorar la experiencia de este.
- Tableau Desktop y sus módulos
  - Elementos de una visualización
- Conexión a los datos
  - Creación de una conexión
  - · Conexión en vivo vs. creación de extractos
  - Opciones de uso compartido

# Contenido del curso

## 7. Estadística, R, Data Mining & Business Intelligence

**Objetivo del módulo:** El Data Mining es la técnica para extraer el conocimiento a través de grandes volúmenes de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos. Este Módulo se imparte con R.

**Módulo:** Sesiones a distancia programadas **Duración:** 6 sesiones de 4 horas

- Conceptos esenciales de estadística
- Lenguaje R
- Optimización
- Aprendizaje supervisado
- Aprendizaje no supervisado

### 8. Python para Big Data

**Objetivo del módulo:** Herramientas fundamentales para el uso de Python en Big Data con un approach funcional: NumPy, Pandas, Jupyter Labs. Concepto de Data Frames y procesamiento de datos.

**Módulo:** Sesiones a distancia programadas

**Duración:** 1 sesión de 4 horas

- Data Frames
- NumPy
- Pandas

## MasterClass de Desarrollo Ágil

Objetivo del módulo: Masterclass sobre el desarrollo ágil.

**Módulo:** Sesiones a distancia puntual **Duración:** 1 sesión de 4 horas

- ¿Qué es el Agilismo?
- Scrum + Kanban

### **Proyecto Final**

**Objetivo del módulo:** Desarrollar un proyecto final en grupo, ya sea de desarrollo web o big data. Se proporcionará un proyecto base para cada especialización o los alumnos podrán proponer algo que les interese, previa aprobación de KeepCoding.

**Módulo:** Trabajo personal en equipo **Duración:** 40 horas (2 semanas)



\*El contenido de este temario puede sufrir cambios.





