# BIG DATA – DATA ANALYTICS (C#23553)

**PROFESOR**: José Luis González

**TUTORA**: Dayana Bernal

# CLASE 0

## CUESTIONES ADMINISTRATIVAS

* Clases serán grabadas.
* Ante cualquier medio de comunicación añadir el número de comisión.
  + Asunto: Nombre – Comisión – Problema

## OBJETIVO

Brindar herramientas que faciliten la inserción laboral en el sector Informática, y en particular, fomentar la participación de las mujeres, con el fin de mejorar su empleabilidad.

## RESUMEN DE CONTENIDOS

* Conceptos Generales. Lógica. Algoritmos.
* Procesos y roles en ciencia de datos.
* Bases de Datos. Relacionales y no relacionales. Trabajando con SQL.
* Planillas de cálculo para análisis de datos.
* Business Intelligence.
* Informes y dashboards.
* Programación en Python. Interfaz gráfica e integración con BBDD.
* Estadística descriptiva básica.
* Librerías de Python orientadas a datos.

## REQUISITOS PARA OBTENER EL DIPLOMA

* Asistir al **65%** de las clases en vivo (sincrónicas).
* Acceder semanalmente al Aula Virtual.
* Realizar los ejercicios obligatorios semanales.
* Cumplir con las fechas de entrega parciales del Proyecto Final Integrador.
* Aprobar el Proyecto Final Integrador. **Fecha límite entrega 10 de noviembre**.
* Aprobar el EFI (Examen Final Integrador).

## OBJETIVO DE ESTE CURSO

Brindar al alumno conocimientos básicos sobre tecnologías, procesos, recursos y equipo de profesionales encargados de convertir grandes volúmenes de información en una herramienta fundamental para la toma de decisiones.

## DATOS, INFORMACIÓN Y CONOCIMIENTO

### DATOS

Son la mínima unidad semántica (que tiene algún sentido), resultado de una medición objetiva, y que de por sí solos, y sin un contexto, carecen de un significado y no son relevantes.

### INFORMACIÓN

Es un conjunto de datos procesados, que tienen un significado (relevancia, propósito y contexto), por lo tanto, son relevantes para la toma de decisiones.

Los datos se transforman en información:

* **Contextualizando**: Se sabe en qué contexto y para qué propósito se generaros.
* **Categorizando**: Se reconocen las unidades de medida que ayudan a interpretarlos.
* **Calculando**: Los datos pueden haber sido procesados matemática o estadísticamente.
* **Corrigiendo**: Se han eliminado errores e inconsistencias de los datos.
* **Condensando**: Los datos se han podido resumir de forma más concisa (Agregación).

#### ¿QUÉ SE BUSCA?

Impactar en la toma de decisiones gracias a la recolección y procesamiento de datos con ayuda de la tecnología.

### CONOCIMIENTO

Es una mezcla de experiencia, valores, información y know-how que sirve como marco para la incorporación de nuevas experiencias e información, y es útil para la acción.

La información se transforma en conocimiento:

* Comparando
* Prediciendo consecuencias
* Buscando conexiones o patrones
* Conversando con otros portadores de conocimiento

## BIG DATA

* Datos masivos, datos a gran escala, inteligencia de datos, macrodatos.
* Enormes volúmenes de datos complejos de múltiples fuentes y en formatos heterogéneos.
* Precisan de aplicaciones informáticas no tradicionales para poder ser procesadas.

## DATA ANALYTICS

Es loa ciencia que obtiene conclusiones por medio del análisis de datos brutos para la **toma de decisiones**.

# CLASE 1

Vemos la utilización de **Deepnote** como herramienta de trabajo a lo largo del cuatrimestre.

* Deepnote se organiza en workspaces.
* Los workspaces se componen de proyectos (Máximo 5 en la versión gratuita).
* Los proyectos contienen archivos.
  + El tipo de archivo más importante es el notebook.

## LÓGICA APLICADA A PROGRAMACIÓN

## ACTIVIDAD LÓGICA PROPOSICIONAL