



# Proyecto Final

**Profesor: José A. Cabezas Jaikel, MTBD.**

## Aspectos Administrativos

A continuación, se citan aspectos administrativos importantes para la realización del proyecto. Además, luego se detallan los contenidos de cada una de las diferentes etapas o entregables del proyecto:

- El proyecto puede ser realizado en grupo (máximo 2 personas), o de forma individual.
- Los entregables se deben realizar de la siguiente forma:
  - Entregable 1: Semana 3.
  - Entregable 2: Semana 6.
  - Entregable 3 (Exposición): Semana 10.
- Toda la documentación y archivos del proyecto pueden ser enviados al correo [jcabezas@ucenfotec.ac.cr](mailto:jcabezas@ucenfotec.ac.cr) o compartidos mediante algún servicio de almacenamiento en la nube a la misma dirección de correo.

## Etapas 1: Recolección de datos

Para el primer entregable del proyecto se requiere lo siguiente:

- Buscar una base de datos, archivos de Excel, archivos de texto o algún otro tipo de fuente de datos estructurados que pueda servir para un análisis de información.
- Extraer información de Facebook de al menos 4 paginas relacionadas con el tipo de información que encontró en el primer ítem, para así poder mezclar ambos tipos de orígenes de datos.
- Crear un repositorio destino, donde se puedan almacenar todas las fuentes de datos, el cual permita luego manipular los datos y visualizarlos.
- Documentación:
  - Portada.
  - Introducción.
  - Descripción de las fuentes de datos:
    - Diccionario de datos.
  - Descripción destino de datos:
    - Repositorio destino.
  - Bibliografía.

## Etapas 2: Limpieza y transformación de datos

Para el segundo entregable del proyecto se requiere lo siguiente:

- Transformación de Datos
  - Tomar columnas de los orígenes de datos que tengan sentido y realizar al menos 10 transformaciones.
  - Ejemplos:
    - Columnas de nombres y apellidos, concatenarlas en una sola columna.
    - Edad, crear una nueva columna de rangos de edades.



- Salarios, crear una nueva columna de rangos salariales.
- Crear una columna de costo total de un producto que sea la suma de todos los costos asociados.
- De una fecha, crear columnas para año fiscal, cuatrimestre, trimestre, semestre, etc.
- Limpieza de Datos
  - Revisar todos los orígenes de datos en busca de alguno de los siguientes aspectos, documentarlos y explicar la corrección aplicada en caso de poder hacerlo:
    - Datos faltantes
    - Datos inconsistentes
    - Datos no útiles para análisis
- Documentación (añadir los siguientes aspectos a la documentación de la etapa 1)
  - Descripción de las transformaciones realizadas
  - Descripción de limpieza realizada

### Etapa 3: Visualización de información

Para el tercer entregable del proyecto se requiere lo siguiente:

- Creación de reportes
  - Seleccionar al menos una herramienta de visualización de datos diferente a las utilizadas en el curso.
  - Cantidad de reportes por presentar:
    - Si se está trabajando individualmente realizar 8 reportes diferentes.
    - Si se está trabajando en grupo, realizar 16 reportes diferentes.
- Documentación (añadir los siguientes aspectos a la documentación de la etapa 2)
  - Documentación de la herramienta de visualización elegida.
    - Nombre de la herramienta.
    - Licenciamiento / costo.
    - Historia.
    - Versiones.
    - Ventajas de la herramienta.
    - Desventajas con herramientas similares.
  - Información de reportes
    - *Screenshots* de cada reporte.
    - Descripción del reporte
    - Orígenes de datos del reporte.
  - Conclusiones del trabajo.
  - Bibliografía actualizada.

----- FIN DE PROYECTO -----