



BIENVENIDOS AL CURSO DE “DATA ANALYTICS”



Te recomiendo la lectura completa de este documento (antes de iniciar el curso o una vez que hayas asistido a la primera clase).

Te ayudará a comprender los temas que serán abordados en cada una de las clases, los programas que deberás tener instalados para poder hacer las prácticas junto con el profesor y el desarrollo de tu Proyecto Final.

Programas a utilizar durante la cursada:

SQL SERVER MANAGEMENT STUDIO Y POWER BI

INSTALACIÓN DE SQL SERVER MANAGEMENT STUDIO

El primer programa a utilizar durante la cursada (a partir de la 2ª clase) será **SQL Server Management Studio**.

Este programa puede presentar algunas complicaciones durante su instalación, por lo que se aconseja que lo hagas lo antes posible, para poder analizar las distintas situaciones que se presenten durante su configuración.

Para acceder a los instructivos de instalación, ingresa en el siguiente link:

https://drive.google.com/drive/folders/1iHUu_flx_KSn-p7JA98MQ5kS8kUgm1zB

Dentro de este Drive, encontrarás distintos instructivos para poder instalar el programa.

Encontrarás un archivo con el nombre **INSTALACIÓN DE SQL SERVER DEVELOPER EDITION (PROFESOR)**. Te aconsejo que utilices este instructivo para llevar a cabo la instalación.

Ante cualquier problema que tengas con la instalación, recuerda que tienes tutoriales en video para la instalación del programa dentro de la misma carpeta.

Si trabajas con MAC, no podrás instalar este programa directamente en MacOS, por lo que deberás generar una máquina virtual que emule al sistema operativo Windows. También encontrarás un instructivo para lograrlo.

INSTALACIÓN DE POWER BI

El segundo programa a utilizar durante la cursada (a partir de la **CLASE Nº 10**) será **Power BI**.

Para acceder a los instructivos de instalación, ingresa en el siguiente link:

https://drive.google.com/drive/folders/1iHUu_flx_KSn-p7JA98MQ5kS8kUgm1zB

Dentro de este Drive, encontrarás distintos instructivos para poder instalar el programa.

Encontrarás un archivo con el nombre **INSTALACIÓN DE POWER BI (PROFESOR)**.

Te aconsejo que utilices este instructivo para llevar a cabo la instalación.

Al igual que con **SQL Server Management Studio**, si trabajas con MAC, no podrás instalar este programa directamente en MacOS, por lo que deberás generar una máquina virtual que emule al sistema operativo Windows. También encontrarás un instructivo para lograrlo.

Si instalas estos programas lo antes posible y pruebas su funcionamiento, te asegurarás una cursada mucho más fluida y sin complicaciones.

CLASE Nº 1:

Se explicará el desarrollo del curso, los programas a utilizar durante la cursada, las fechas en las que se deberán entregar los 3 trabajos prácticos y el Proyecto Final (todos éstos están relacionados) y una introducción al Análisis de Datos.

Por otro lado, se explicará el **PRIMER TRABAJO PRACTICO**.

La presentación se habilitará el día de la **CLASE Nº 3** y durante un período de 7 días corridos (es decir, a partir del día en que se dicte la **CLASE Nº 3**, tendrás 7 días corridos para hacer la presentación, antes de que el botón que permite la presentación, se deshabilite).

Para este trabajo práctico, deberás seleccionar el Dataset con el que trabajarás a partir de ahora y hasta la presentación del Proyecto Final. Si bien puedes utilizar un Dataset propio, se aconseja utilizar alguno de los Datasets que se incluyen en el siguiente link:

<https://drive.google.com/drive/folders/1BPloFA4Z5YPu0tVZV2dOIXAFgzQvwl3d>

Estos Datasets están configurados para que puedas cumplir con todos los requerimientos que exige el **PROYECTO FINAL** de CoderHouse. Estos requerimientos serán detallados durante la primera clase del curso.

En el siguiente link encontrarás ejemplos de los distintos trabajos prácticos a entregar, como así también un ejemplo del Proyecto Final:

<https://drive.google.com/drive/u/0/folders/1ha0UjKWhDMdoN1eZhWoLdz3DuXOOkCCU>

¿CÓMO DEBERÁS HACER LAS ENTREGAS?

Para poder hacer las entregas, deberás generar una carpeta en tu Drive de Google o Outlook con el nombre **DATA ANALYTICS** y dentro de esa carpeta, generar otras 4 carpetas con los nombres **PRIMERA ENTREGA**, **SEGUNDA ENTREGA**, **TERCERA ENTREGA** y **PROYECTO FINAL**. Todo el material correspondiente a cada entrega, se deberá colocar en la carpeta correspondiente y, al momento de hacer la presentación a través de la plataforma, deberás especificar la URL de la carpeta. Recuerda que **todas las carpetas deben ser públicas** para que los tutores puedan acceder a ellas y a su contenido y poder hacer las devoluciones.

Por último, durante el curso, se utilizarán distintas bases de datos para llevar a cabo ejercitaciones. Se aconseja descargar todos los archivos contenidos entre Drive y almacenarlos en una carpeta con el nombre CURSO DATA ANALYTICS. Todos estos archivos podrás descargarlos desde el siguiente link:

https://drive.google.com/drive/u/0/folders/1b_Z794Dn3G1CMusmEyNG04ELTrG_81x5

CLASE Nº 2:

Durante esta clase, conocerás los conceptos y generalidades de una base de datos, qué objetos la componen, los distintos tipos de bases de datos que existen, los distintos motores de bases de datos existentes en el mercado laboral y generarás una base de datos dentro del servidor SQL.

Nota: en las próximas clases, trabajaremos con una base de datos llamada NEPTUNO. Esta base de datos la podrás descargar desde el siguiente link:

https://drive.google.com/drive/folders/1b_Z794Dn3G1CMusmEyNG04ELTrG_81x5

Ya deberías tenerla en la carpeta del curso si la descargaste al comienzo de la cursada.

Para poder restaurar la base de datos, ingresa en el siguiente link para acceder al archivo RESTAURAR BASE DE DATOS NEPTUNO y sigue el paso a paso del instructivo:

https://drive.google.com/drive/folders/1iHUu_flx_KSn-p7JA98MQ5kS8kUgm1zB

Esto te permitirá restaurar la base de datos con la que trabajaremos a partir de las clases próximas.

Otras bases de datos con las podrás ejercitar todos los temas vistos en clase son CONTOSO.BAK y CODERHOUSE.BAK. Puedes descargarlas también desde el mismo link y restaurarlas, utilizando como instructivo, el utilizado para restaurar la base de datos NEPTUNO.

CLASE Nº 3:

En la tercera clase, abordaremos el concepto de **BASE DE DATOS RELACIONAL**.

Conoceremos la estructura de las tablas, el concepto de tipos de datos, las restricciones aplicables a los campos y las distintas claves (primarias y foráneas).

A través de las tablas y sus correspondientes campos, tipos de datos, restricciones y claves, aprenderás a generar un **MODELO ENTIDAD-RELACIÓN**.

Este tema es sumamente importante por varios motivos: por un lado, SQL Server y todos los motores de bases de datos relacionales trabajan con este concepto, al igual que **Power BI** (programa a abordar a partir de la **CLASE Nº 11**); por otro lado, para que una base de datos relacional funcione adecuadamente, debe partir de un **MODELO E-R** coherente y, por último, este tema será evaluado en el **2º TRABAJO PRACTICO** a entregar en la **8ª CLASE** del curso.

Analizaremos el Dataset a utilizar durante el curso para poder generar su **MODELO ENTIDAD-RELACION** (ten en cuenta que este Dataset no podrá ser utilizado para la presentación del Proyecto Final y los entregables previos; sólo se utilizará durante la cursada para explicar cada uno de los temas que, posteriormente, podrás aplicar en el Dataset seleccionado por vos).



Atención: hoy se habilita la primera entrega de tu Proyecto Final. Para hacer la presentación, deberás pulsar el botón **PRESENTAR** que aparece debajo de la clase de hoy en el cronograma de clases de la plataforma y especificar el link de la carpeta que contiene tu entrega. El botón estará habilitado a partir de esta clase y por 7 días; pasados los 7 días, el botón se inhabilitará automáticamente.

Recuerda que la carpeta de Drive en la que hayas colocado tu Dataset y tu documentación debe ser pública para que los tutores puedan ingresar en ella y acceder a los archivos, para poder hacer la devolución correspondiente. Esta primera entrega constará del Dataset seleccionado para hacer el análisis y el documento de Word o PDF con la correspondiente introducción, la descripción de la temática de los datos contenidos en el Dataset, el alcance de tu análisis y la hipótesis de tu proyecto.

En el siguiente link (ya compartido en las primeras páginas de este documento) encontrarás un ejemplo del documento a presentar en esta primera entrega:

https://drive.google.com/drive/folders/1Zlxd-M8m-EDyhoyvOZIP8JStlHmUpW_W

Con todos los temas vistos en esta clase, ya podrás empezar a diseñar el **MODELO E-R** de tu Dataset y la correspondiente documentación para ser entregados en la **CLASE Nº 8**.

En el siguiente link encontrarás el ejemplo del trabajo práctico a entregar en la **CLASE Nº 8**:

<https://drive.google.com/drive/folders/1-TIT1BNgZU65MP-bouD8FOWUNoSG4kPo>

CLASE Nº 4:

En esta clase, aprenderás a utilizar los distintos comandos SQL para generar el **MODELO ENTIDAD-RELACION** creado en la clase anterior, dentro de tu base de datos, a través de un grupo de comandos denominado **DDL: LENGUAJE DE DEFINICION DE DATOS**.

Dentro de este grupo, existe un comando llamado **CREATE**, el cual permite crear distintos tipos de objetos dentro de una base de datos. En este caso, utilizaremos este comando para crear tablas. Anteriormente, habíamos utilizado el comando **CREATE** para crear la base de datos; ahora es momento de crear tablas dentro de ella, dado que son las tablas las que contienen la información.

Recuerda que, para la próxima clase, deberás tener restaurada la base de datos NEPTUNO.BAK en tu SQL Server Management Studio.

CLASE Nº 5:

En esta clase, continuaremos viendo comandos pertenecientes al grupo **DDL: LENGUAJE DE DEFINICION DE DATOS**. De este grupo, aplicaremos los comandos **ALTER** y **DROP**.

Posteriormente, nos adentraremos en el segundo grupo de comandos, denominado **DML: LENGUAJE DE MANIPULACION DE DATOS**. Uno de los comandos pertenecientes a este grupo, es el comando **INSERT**, el cual permite cargar registros (datos) en las tablas generadas en la clase anterior. Veremos 2 formas distintas de cargar registros en una tabla: la forma completa y la forma simplificada.

Por otro lado, también utilizaremos el comando **UPDATE** para corregir algún dato ya cargado en una tabla y el comando **DELETE** para eliminar registros de una tabla que cumplan con una determinada condición.

CLASE Nº 6:

En esta clase, comenzaremos a utilizar uno de los comandos más utilizados por todos los empleados de una empresa que acceden a las bases de datos del servidor: el comando **SELECT**. A partir de esta clase, aprenderás a utilizar este comando y sus distintas cláusulas: **FROM**, **DISTINCT**, **TOP**, **JOIN**, **AS**, **WHERE**, **GROUP BY**, **HAVING** y **ORDER BY**.

Conocerás los distintos tipos de operadores (**LÓGICOS**, **MATEMÁTICOS** y **COMPARATIVOS**) para poder hacer consultas.

También utilizaremos algunas funciones de texto dentro de las instrucciones SQL.

CLASE Nº 7:

En esta clase, veremos las cláusulas restantes para conocer a fondo el uso del comando **SELECT**: **GROUP BY** y **HAVING**. Aprenderás a utilizar las funciones de agregado: **SUM**, **MAX**, **MIN**, **AVG** y **COUNT**, las cuales te permitirán desarrollar consultas agrupadas. Empezaremos a ver cómo, a través de instrucciones SQL, traer información de distintas tablas a través del uso de la cláusula **JOIN**. Aplicaremos funciones de fecha dentro de las instrucciones SQL.

CLASE Nº 8:

En esta clase, aplicarás otros tipos de funciones (funciones de conversión y fórmulas matemáticas). Veremos cómo utilizar el condicional **CASE** en las consultas.



Atención: hoy se habilita la segunda entrega de tu Proyecto Final. Para hacer la presentación, deberás pulsar el botón PRESENTAR que aparece debajo de la clase de hoy en el cronograma de clases de la plataforma y especificar el link de la carpeta que contiene tu entrega. El botón estará habilitado a partir de esta clase y por 7 días; pasados los 7 días, el botón se inhabilitará automáticamente.

Recuerda que la carpeta de Drive en la que hayas colocado tu Dataset y tu documentación debe ser pública para que los tutores puedan ingresar en ella y acceder a los archivos, para poder hacer la devolución correspondiente.

Esta segunda entrega constará del Dataset seleccionado y ya modificado para hacer el análisis y el documento de Word o PDF con la correspondiente introducción, la descripción de la temática de los datos contenidos en el Dataset, el alcance de tu análisis, la hipótesis de tu proyecto, el Modelo Entidad-Relación y el detalle de cada una de las tablas contenidas en el Dataset. En el siguiente link (ya compartido en las primeras páginas de este documento) encontrarás un ejemplo del documento a presentar en esta segunda entrega:

<https://drive.google.com/drive/folders/1-TIT1BNgZU65MP-bouD8FOWUNoSG4kPo>

CLASE Nº 9:

En esta clase, aprenderemos a generar subconsultas y consultas complejas a través de las cláusulas **UNION** y **JOIN**.

Como tema extra fuera del temario, aprenderemos a generar un **MODELO ENTIDAD-RELACION** dentro del propio programa. Esto te permitirá conocer un poco más cómo están conectadas las tablas entre sí y poder hacer uso de las cláusulas **JOIN** sin ningún tipo de problema.

Recuerda que, para la próxima clase, deberás tener instalado el programa POWER BI.

CLASE Nº 10:

En esta clase, concluiremos con los temas faltantes y aprenderás a conectar una base de datos de **SQL Server** con **Power BI** y sus distintas formas de conexión. Si bien para el Proyecto Final los datos a analizar en Power BI serán tomados desde un Dataset en formato Excel, CSV o TXT (Dataset que vienes trabajando desde la primera entrega), es indispensable saber cómo conectarnos desde **Power BI** con datos provenientes de una base de datos **SQL Server**, dado que la mayoría de las empresas, debido al gran volumen de información con la que trabajan, dicha información se toma desde una base de datos SQL.

CLASE Nº 11:

En esta clase, haremos una introducción a Power BI para conocer sus pantallas y paneles y comenzaremos a utilizar el programa para el análisis de datos a través de tableros gráficos. Conoceremos el mecanismo para poder aplicar fondos a nuestras páginas del proyecto de Power BI y aprenderemos algunas herramientas que permitan aplicar conceptos de **STORYTELLING**. Se explicará la generación de **HIPERVÍNCULOS INTERNOS** y **EXTERNOS** dentro de Power BI para lograr un proyecto más dinámico.

Con estos temas, ya podrás iniciar tu PROYECTO FINAL. Deberás hacer una entrega previa antes de la presentación del PROYECTO FINAL; esta entrega se deberá llevar a cabo en la CLASE Nº 18. Al hacer esta entrega, los tutores verificarán qué tan bien está encaminado tu Proyecto Final y te harán la devolución correspondiente. A partir de allí, podrás continuar desarrollando tu Proyecto Final, para poder entregarlo el día de la última clase del curso (CLASE Nº 23). Con los temas vistos en esta clase, ya estarás en condiciones de generar los fondos para las páginas del proyecto y los hipervínculos, tanto internos como externos, de tu proyecto.

CLASE Nº 12:

En esta clase, haremos unos ejercicios utilizando el **EDITOR DE POWER QUERY**, programa que permite llevar a cabo la transformación de los datos obtenidos desde un Dataset, previo a su carga en Power BI. Aprenderemos a utilizar herramientas de este programa y comenzaremos a generar las relaciones entre las tablas importadas.

Con los temas a desarrollar durante las CLASES Nº 12 y Nº 13, ya podrás importar a Power BI el Dataset seleccionado y analizado durante las entregas anteriores. Recuerda que, todos los procesos de transformación efectuados en el Editor de Power Query, deberás documentarlos. En el siguiente link, encontrarás un ejemplo de distintos Proyectos Finales para que puedas guiarte con su armado:

https://drive.google.com/drive/folders/12DvODVovCxP0qN-zmVZRiM-w4Mp5_MpU

CLASE Nº 13:

En esta clase, continuaremos trabajando con el **EDITOR DE POWER QUERY**, para terminar de importar todas las tablas del Dataset a nuestro proyecto de Power BI. Además, terminaremos de relacionar las tablas importadas para generar un modelo de datos que funcione a la perfección.

Veremos cómo importar datos desde páginas Web a través de tablas y archivos descargables. Por último, conoceremos el mecanismo para generar tablas auxiliares en Power BI, las cuales nos permitirán hacer otros tipos de visualizaciones y análisis con los datos ya importados.

Recuerda que, para cada una de las tablas importadas, deberás documentar todas las transformaciones realizadas en el Editor de Power Query y cómo quedan las mismas después de haberlas relacionado.

CLASE Nº 14:

En la esta clase, aprenderás a crear una **TABLA DE FECHAS** o **TABLA CALENDARIO** desde Power Query utilizando **Lenguaje M**. Conoceremos sus múltiples usos y aprenderemos a generar columnas calculadas a través de funciones **DAX** para completar la **TABLA DE FECHAS** y darle sus distintos usos.

Por otro lado, comenzaremos a generar visualizaciones para nuestro tablero, conocerás las opciones de formato aplicables a una visualización y aprenderás a descargar visualizaciones no incluidas dentro del programa.

Visualizaciones: **MAPA - SCROLLER**.

En el caso de que no cuentes con usuario de Power BI, podrás descargar las visualizaciones a utilizar en el curso desde el siguiente link:

<https://drive.google.com/drive/folders/1SJXgrTb50TupuaR9QEKd4imml7jYxISb>

Para descargar el código en Lenguaje M que permite la generación de la tabla de fechas, ingresa en el siguiente link:

https://drive.google.com/drive/folders/1oaUxM2pfAko_AyDZ8Jq7nwQLsCqA2fWx

Con los temas desarrollados hasta el momento, ya deberías estar en condiciones de importar todas tus tablas y llevar a cabo todas las transformaciones que éstas requieran. Además, ya deberías generar la TABLA DE FECHAS. Con esta última tabla, tu modelo de datos en Power BI quedará finalizado. No olvides documentar la generación de la tabla de fechas.

CLASE Nº 15:

En esta clase, aprenderemos el concepto de **COLUMNAS CALCULADAS** y **MEDIDAS** (implícitas y explícitas). Si bien las medidas se pueden generar en cualquier tabla del proyecto, un proyecto de Power BI queda mucho más prolijo si todas las medidas explícitas se almacenan en una tabla que almacene específicamente medidas, por lo que conocerás cómo generar una tabla de medidas. Veremos cómo calcular sumatorias, máximos, mínimos, cantidades, promedios, porcentajes y tasas de crecimiento.

Por último, aplicaremos formatos condicionales en las visualizaciones generadas.

Visualizaciones: **MATRIZ - CHICLET SLICER - SEGMENTACIÓN DE DATOS**.

CLASE Nº 16 / CLASE Nº 17:

STORYTELLING: generación de marcadores para el reemplazo de gráficos dentro de una misma página y para la eliminación de filtros aplicados en la página

Visualizaciones: **GRÁFICO DE BARRAS - MEDIDOR.**

MEDIDAS: uso de variables. Funciones **SWITCH, REPT, UNICHAR, CALCULATE.**

CLASE Nº 18:

En esta clase, aprendemos a generar un tooltip personalizado que permita detallar en profundidad los valores mostrados por una de las visualizaciones colocadas en el tablero. Además, conocerás el mecanismo para insertar un archivo del video institucional de la empresa para la que estás generando el proyecto y un archivo de audio para la portada del tablero. **STORYTELLING:** a través de marcadores, podrás definir si el archivo de audio insertado en la portada del proyecto se reproduzca o no.

Por último, aprenderás a clasificar las medidas en distintas carpetas, para lograr un mejor ordenamiento de todas las medidas generadas hasta el momento.

Para utilizar los videos ofrecidos por el profesor, ingresa en el siguiente link:

<https://github.com/lucianojuliangomez/videos>

Para utilizar los audios ofrecidos por el profesor, ingresa en el siguiente link:

<https://github.com/lucianojuliangomez/audios>



Atención: hoy se habilita la tercera entrega de tu Proyecto Final. Para hacer la presentación, deberás pulsar el botón PRESENTAR que aparece debajo de la clase de hoy en el cronograma de clases de la plataforma y especificar el link de la carpeta que contiene tu entrega. El botón estará habilitado a partir de esta clase y por 7 días; pasados los 7 días, el botón se inhabilitará automáticamente.

Recuerda que la carpeta de Drive en la que hayas colocado tu

Dataset, tu documentación y tu archivo de Power BI debe ser pública para que los tutores puedan ingresar en ella y acceder a los archivos, para poder hacer la devolución correspondiente.

Esta tercera entrega deberá incluir el Dataset en base al cual generaste tu proyecto de Power BI, la documentación en formato Word o PDF y el archivo de Power BI con la portada, la hoja de hipótesis y, por lo menos, una página completa con visualizaciones, filtros y segmentaciones. En el siguiente link (ya compartido en las primeras páginas de este documento) encontrarás 3 ejemplos de la presentación:

https://drive.google.com/drive/folders/11ljXdNBDM17yfSL0kwwQQAXao_XCOr5m

CLASE Nº 19:

En esta clase, denominada **WORKSHOP**, haremos un ejercicio práctico para repasar algunos de los temas vistos durante las clases anteriores. Por otro lado, podrás aprovechar algunos minutos para hacerle todas las consultas que requieras a tu tutor, acerca de tu Proyecto Final, el cual deberás entregar en la **CLASE Nº 23**.

CLASE Nº 20:

En la primera parte de esta clase, recibiremos la visita de un especialista en el uso de bases de datos y generación de tableros en Power BI para que tengas un acercamiento al uso de ambos programas a nivel empresarial.

En la segunda parte de la clase, aprenderás a aplicar un proceso denominado **DRILL THROUGH**, también conocido como **EXPLORACION EN PROFUNDIDAD**. Con este proceso, podrás pasar de una hoja a otra de tu informe de Power BI para analizar más en detalle algún dato de la página principal de tu informe.

CLASE Nº 21:

En esta clase, nos acercaremos al uso de la estadística descriptiva, a través del uso de **MEDIDAS DE TENDENCIA CENTRAL**.

Aprenderás a calcular el promedio, la media y la moda de distintos campos para aplicar dichos cálculos en una visualización.

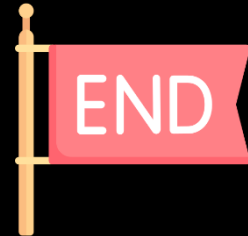
A través de una narrativa, generaremos las conclusiones del proyecto armado en Power BI.

CLASE Nº 22:

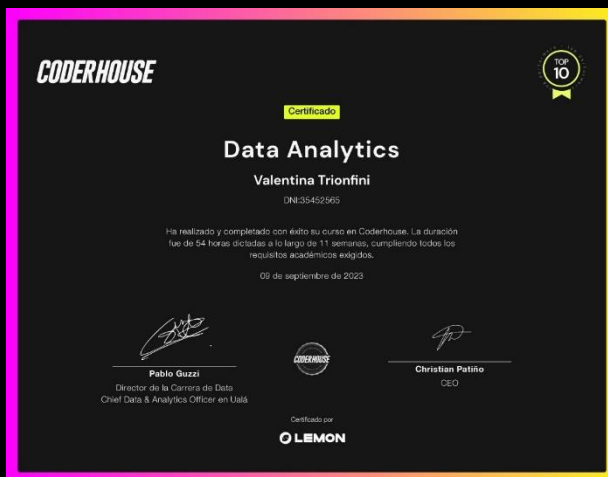
En esta clase, haremos los retoques finales a nuestro proyecto de Power BI y aprenderás a publicarlo para poder compartirlo con compañeros o gerentes de tu empresa a través del **Servicio de Power BI** y, una vez publicado en el servicio, poder incorporarlo a **Microsoft Teams** o **Microsoft SharePoint**.

CLASE Nº 23:

Última clase del curso: en esta clase, los alumnos que tengan su **Proyecto Final** terminado, tendrán la posibilidad de presentarlo ante sus compañeros. Esta presentación no deberá llevar más de 10 minutos por alumno. Esto permitirá a otros alumnos ver los distintos análisis que cada uno de los oradores hayan efectuado a partir de sus Datasets y los tableros generados con tal fin.



Finalizada la presentación de los Proyectos Finales, se utilizará el resto de la clase para evacuar dudas sobre la presentación del Proyecto Final de los alumnos restantes.



Recuerda: en esta clase se habilita la presentación del PROYECTO FINAL. Tendrás 10 días a partir de esta clase para efectuar la presentación. Transcurridos los 10 días a partir de la última clase, el botón que permite hacer la presentación se deshabilitará de tu plataforma automáticamente. Deberás incluir tu Dataset, tu documentación de Word o PDF y tu archivo de Power BI en la carpeta PROYECTO FINAL de tu Drive y, al momento de hacer la presentación, no

olvides especificar la URL de esta carpeta. La carpeta debe ser pública y no privada, para que el profesor pueda corregir tu PROYECTO FINAL.

Una vez que el PROYECTO FINAL quede aprobado, podrás descargar de la plataforma tu diploma.