# 1 DETALLES

* **MAESTRIA EN**: Inteligencia Artificial, Data Science para la transformación de negocios
* **MÓDULO**: Preparación de datos y análisis exploratorio
* **HORAS ACADÉMICAS**: 40 horas en aula virtual
  + Horas Teóricas de Trabajo virtual: 20
  + Horas de Trabajo Colaborativo: 8
  + Horas de interacción Investigativa: 2
  + Horas de Trabajo Independiente Libre en Línea: 4
  + Horas de Evaluación: 6
* **DOCENTE**: MSc. Alvaro Limber Chirino Gutierrez

# 2 ORGANIZACIÓN DE LAS HORAS PROGRAMADAS

Se propone la siguiente organización de las horas programadas:

* Clases teóricas: 40%
* Clases prácticas: 40%
* Atención a consultas: 10%
* Evaluación: 10%

# 3 OBJETIVOS DEL MÓDULO

## 3.1 Objetivo General

Brindar a los estudiantes las habilidades y conocimientos teóricos y prácticos necesarios para dominar las técnicas de preparación de datos y análisis exploratorio en el entorno del lenguaje de programación R. Esto les permitirá manipular, limpiar, transformar, analizar y visualizar conjuntos de datos de manera efectiva, extrayendo información valiosa y preparando los datos para su uso en modelos de aprendizaje automático y otras aplicaciones analíticas.

## 3.2 Objetivos Específicos:

* Comprender los fundamentos del análisis de datos y su importancia en el contexto actual.
* Aplicar técnicas de limpieza y transformación de datos para prepararlos para el análisis.
* Realizar un análisis exploratorio de datos (EDA) para resumir las características de los datos, identificar patrones, detectar anomalías y generar hipótesis.
* Utilizar técnicas de visualización de datos para comunicar información de manera clara y efectiva. Seleccionar las herramientas y técnicas de preparación de datos adecuadas para diferentes tipos de problemas.
* Comunicar los resultados del análisis exploratorio de datos de manera clara, concisa y convincente.

# 4 COMPETENCIAS

* Habilidad para manipular, limpiar y transformar datos utilizando herramientas como R y tidyverse.
* Capacidad para realizar un análisis exploratorio de datos utilizando técnicas estadísticas y de visualización.
* Destreza para comunicar información de manera clara y efectiva a través de visualizaciones de datos.
* Pensamiento crítico y capacidad para resolver problemas relacionados con el análisis de datos.
* Habilidad para trabajar en equipo y colaborar en proyectos de análisis de datos.
* Capacidad para aprender y adaptarse a nuevas herramientas y técnicas de análisis de datos.

# 5 MÉTODOS Y RECURSOS DIDÁCTICOS

Para lograr los objetivos y desarrollar las competencias de la materia, se propone la implementación de una variedad de métodos y recursos didácticos que fomenten el aprendizaje activo, colaborativo y significativo en los estudiantes.

## 5.1 Métodos:

* Clases magistrales
* Clases prácticas con datos reales

## 5.2 Recursos didácticos:

* Github
* Pizarra digital
* R, RStudio

# 6 CONTENIDO ANALÍTICO

1. **Introducción a R:** Sintaxis, objetos y funciones.
2. **Fundamentos del análisis de datos:** Tipos de analíticas, tipos de variables, analítica descriptiva
3. **Fuentes de información y Raspado web :** Enfoque estadístico, fuentes tradicionales, raspado web con R
4. **Gramática de la manipulación de datos:** Filtrado, selección, transformación y resumen.
5. **Introducción a la visualización de datos:** Gramática de figuras, librería ggplot

# 7 ESTRUCTURA DE EVALUACIÓN

* Evaluación teórica 30%
* 2 Evaluaciones prácticas 30% c/u
* Proyecto 10%

# 8 BIBLIOGRAFÍA

* Hadley Wickham, (2024). R for Data Science

Erl, Khattak, y Buhler, Big Data Fundamentals.

Torgo, Data mining with R: Learning with case studies, second edition.

* Claus O. Wilke, (2024). Fundamentals of Data Visualization

# 9 CRONOGRAMA

1. Introducción a R. Semana 1
2. Fundamentos del análisis de datos. Semana 2
3. Fuentes de información y Raspado web. Semana 3
4. Gramática de la manipulación de datos. Semana 4
5. Introducción a la visualización de datos. Semana 4

# 10 VIABILIDAD DEL PLAN DE TRABAJO

* Plataforma virtual
* Asistencia de los participantes en al menos un 80%
* Apuntes virtuales o grabaciones de las sesiones

# 11 METODOLOGÍA DE RELACIONAMIENTO ADMINISTRATIVO

Los canales de comunicación con los estudiantes será mediante:

* La plataforma virtual
* Grupo de whatsapp

M.Sc. Alvaro Limber Chirino Gutierrez

DOCENTE INVITADO

Telf.: 70694453

mail: [achirinog@fcpn.edu.bo](mailto:achirinog@fcpn.edu.bo)