

# Geodatabase, Geo-Statistics, Machine Learning and Big Data (SQL – R – Python – Qgis)

*Desarrolladores:* Ing. A. Otiniano , Ing. J.Andrade & Ing. R.Perez

*Instituto de Investigación FIGMM - UNI : Dr. Jimmy Rosales Huamani*

AUSPICIADO POR:

Grupo: Censado Multidisciplinario,  
Accesibilidad Universal y Machine  
Learning





# ***PARTE II***

# PARTE II - A

1. Rstudio
2. Paquetes y Librerías
3. Mi primer Script
  - ☐ R como calculadora científica (funciones de calculadora).
  - ☐ Tipos de variable y Operadores en R.
4. Definición de Vectores en R
  - ☐ Generación de vectores usando números aleatorios.
  - ☐ Funciones Básicas.
  - ☐ Datos Especiales (NA's, Inf & NaN's).
5. Matrices
6. Otras Funciones Básicas (sort(), seq(), rep())
7. DataFrame
8. Estructuras de decisión y repetición:
  - ☐ If, else, else if.
  - ☐ For (loop).
  - ☐ While, next, break.
9. Factores en R.
10. Listas en R.
11. Paquetería de Funciones
  - ☐ Definición de Funciones
  - ☐ Aplicación de Funciones
12. Instalación de paquetes de CRAN, GitHub, Bioconductor u otros.

## ***PARTE II - B***

1. Conectarse desde internet o base de datos -Importar y Exportar Data en R.
  2. Tidyverse y ggplot en el IDE Rstudio.
  3. Limpieza de Data – Filtros, Transformación (Data Wrangle, Data Tidy, Data Transformation).
  4. Introducción a la estadística descriptiva en R.
  5. Estadística Descriptiva en R.
  6. Análisis Exploratorio de Datos (AED):
    - ☐ Análisis e Imputación de Data mediante métodos Robustos (Datos Nulos, debajo o encima del límite de detección, etc.)
    - ☐ Estadística Descriptivas Univariantes – Multivariantes.
    - ☐ Gráficos Estadísticos Descriptivos Univariantes – Multivariantes.
  7. Pruebas Estadísticas - Inferencial
- \* Ejercicios que Contiene esta parte (Parte II – C y D)
- ☐ Revisión de Base de Datos Geológica.
  - ☐ Tratamiento estadístico de una base de datos Geológica.
  - ☐ Aplicaciones generales y Análisis Exploratorio de Datos Geológicos en Recursos Hídricos – Suelos y Roca.

# ***PARTE II - C 1***

Todo estos análisis han tenido una técnica de muestreo representativo de campo y control de calidad en las fases de campo y laboratorio.

1. **Análisis Exploratorio-Geoquímica de Sedimentos de Quebrada (Proyecto Minero Colca y Jala Oca). A.Otiniano & J.Andrade**
  1. Análisis de Base de Datos (estructura y revisión general).
  2. Imputación de Datos debajo del límite de detección (<L.D.).
  3. Análisis Exploratorio mediante Estadísticos Descriptivos.
  4. Correlaciones : metal base - pathfinder
  5. Determinación del Background y Threshold – detección de outliers.
  6. Cluster por poblaciones (según factor).
  7. Índices Mineralógicos.
  8. Imputación de valores sobre el límite de detección.

**Continuará ...**

## PARTE II – C3

Todo estos análisis han tenido una técnica de muestreo representativo de campo y control de calidad en las fases de campo y laboratorio. *Reseña histórica y paper*

### 2. Análisis Exploratorio en Suelos (Corredor Honoria Tournavista). A.Otiniano & J.Andrade

1. Análisis de Base de Datos (estructura y revisión general).
2. Análisis Exploratorio mediante estadísticos descriptivos.
3. Análisis Gráfico Univariante y Bivariante.
4. Estándares de Calidad Ambiental para Suelo (ECAs).
5. Análisis Bivariantes Numéricos (Boxplot Multivariante por Clasificación, Corrplot)
6. Análisis Multivariable (Biplot – PCA).
7. Mapa Geoestadístico de distribución Catiónica.
  1. Crear Grilla con tamaño de celda.
  2. Modelado usando Inverso a la distancia (idw).
  3. Variograma.
  4. Modelado usando Krigging.

Continuará ...

**Los Análisis Multivariantes Avanzados Aplicando Métodos Estadísticos Inferenciales y Machine Learning además del análisis Geoespacial – Geoestadístico detallado se verá al final de la Parte III.**

FIN  
DE LA SEGUNDA  
PARTE

