

## Módulo 1

### Entendiendo la data

↳ Base de datos: descripción de los datos

↳ Primera fila es la etiqueta de cada dato

### Librerías para análisis

- Pandas
- Numpy
- Scipy

### Librerías para visualización

- Matplotlib
- Seaborn

### Librerías para algoritmos

- Scikit-learn
- Stats model

### ⇒ Importar y exportar Datos

- Formato
- Path ("lugar donde se encuentra la base")

→ Comando .read → si es .csv asume que hay un encabezado

↳ añadir encabezado con creación de lista

→ Comando .to - → guarda la base de datos en el formato especificado

### ⇒ Análisis de data

→ Entender los datos antes de iniciar análisis

→ verificar formato de datos



object  
int  
float  
datetime

- ayuda a evitar confusiones
- debe ser compatible con métodos y funciones de pandas

→ `df.describe()` ⇒ evaluación estadística del dataframe (min, max, etc...)

→ `df.info()` ⇒ Top 30 filas / Bottom 30 filas

→ Nos indica la presencia de valores NA en la base de datos

## Resultado de Review Questions (Módulo 1)

Progreso del curso para el estudiante 'diegoku' con el correo (a01793499@TEC.MX)

