

Universidad Rafael Landívar  
Facultad de Ingeniería  
Licenciatura en Ingeniería en Informática y Sistemas  
Ing. José Abelardo Méndez Pú

# HT01

## Compiladores

Luis David Ixquiac Sac

Carné: 1521223

Quetzaltenango, domingo 18 de enero de 2026

## Código y patrón utilizados:

```

import re

import pandas as pd
from PyPDF2 import PdfReader
import os

rutaArchivo = 'Analisis Anual 2018 ETAS.pdf'
document = PdfReader(rutaArchivo)
pagina = document.pages[4]
texto = pagina.extract_text()
print(texto)

patron = r'\n(\d+)\s+(\d+)\s+(.+?(?:\n.+?))\s{2,}(.)\s{2,}(-|.+?)\s+(\d+)'

datos = re.findall(patron, texto)
print(datos)

etas = pd.DataFrame(datos, columns=['No', 'SE', 'AreaSalud', 'ETA',
'Fuente', 'Casos'])

print(etas)

```

## Salida:

No	SE	AreaSalud	Guate Central	Intoxicación Alimentaria	Pan con Pollo	ETA	Fuente Casos
0	1	4	Guate Noroccidente/				
	2	4	Peten Norte	Intoxicación Alimentaria	envuelto de coliflor	20	
	2	3	6	Quetzaltenango	Rotavirus	17	4
	3	5	7	Quetzaltenango	Intoxicación Alimentaria	carne azada	28
	4	6	15	Alta Verapaz	Intoxicación Alimentaria	caldo de pollo	23
	5	7	15	Huehuetenango	Enfermedad Diarreica	agua contaminada	39
	6	8	18	Chiquimula	Intoxicación Alimentaria	Hot Dog	5
	7	9	20	Huehuetenango	Rotavirus	-	42
	8	10	22	Totonicapan	Fiebre Tifoidea	-	10
	9	11	24	Suchitepequez	Intoxicación Alimentaria	Caldo de pollo	34
	10	12	25	Solola	Intoxicación Alimentaria Hongo Amanita	Muscaria	5
	11	13	25	Huehuetenango	Intoxicación Alimentaria	frijoles	10
	12	14	29	Suchitepequez	Intoxicación Alimentaria carne asada , ensalada	de coditos	32
	13	15	33	Chiquimula	Intoxicación Alimentaria	panes con pollo	120
	14	16	34	Escuintla	Hepatitis A	-	10
	15	17	34	Guatemala Sur	Intoxicación Alimentaria tortilla con pollo , y crema	y crema	171
	16	18	41	Guatemala Sur	Intoxicación Alimentaria	ceviche	16
	17	19	49	Escuintla	Fiebre Tifoidea	-	7
	18	20	50	Chimaltenango	Intoxicación Alimentaria	Estofado de Pollo	39
	19	21	50	Escuintla	Hepatitis A	-	3
	20	22	52	Alta Verapaz	Intoxicación Alimentaria	Frijoles, cafe	8