

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 = 'Análisis 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:

```
[('1', '4', 'Guate Noroccidente', 'Guate Central', 'Intoxicación Alimentaria Pan con Pollo', '135'), ('2', '4', 'Petén Norte', 'Intoxicación Alimentaria', 'envuelto de coliflor', '20'), ('3', '6', 'Quetzaltenango', 'Rotavirus', '17', '4'), ('5', '7', 'Quetzaltenango', 'Intoxicación Alimentaria', 'carne asada', '28'), ('6', '15', 'Alta Verapaz', 'Intoxicación Alimentaria', 'caldo de pollo', '23'), ('7', '15', 'Huehuetenango', 'Enfermedad Diarreica', 'agua contaminada', '39'), ('8', '18', 'Chiquimula', 'Intoxicación Alimentaria', 'Hot Dog', '5'), ('9', '20', 'Huehuetenango', 'Rotavirus', '-', '42'), ('10', '22', 'Totonicapán', 'Fiebre tifoidea', '-', '10'), ('11', '24', 'Suchitepéquez', 'Intoxicación Alimentaria', 'caldo de pollo', '34'), ('12', '25', 'Sololá', 'Intoxicación Alimentaria Hongo Amanita', 'Muscaria', '5'), ('13', '25', 'Huehuetenango', 'Intoxicación Alimentaria', 'frijoles', '10'), ('14', '29', 'Suchitepéquez', 'Intoxicación Alimentaria carne asada, ensalada', 'de coditos', '32'), ('15', '33', 'Chiquimula', 'Intoxicación Alimentaria', 'panes con pollo', '120'), ('16', '34', 'Escuintla', 'Hepatitis A', '-', '10'), ('17', '34', 'Guatemala Sur', 'Intoxicación Alimentaria tortilla con pollo, repollo', 'y crema', '171'), ('18', '41', 'Guatemala Sur', 'Intoxicación Alimentaria', 'ceviche', '16'), ('19', '49', 'Escuintla', 'Fiebre tifoidea', '-', '7'), ('20', '50', 'Chimaltenango', 'Intoxicación Alimentaria', 'Estofado de Pollo', '19'), ('21', '50', 'Escuintla', 'Hepatitis A', '-', '3'), ('22', '52', 'Alta Verapaz', 'Intoxicación Alimentaria', 'Frijoles, café', '8')]
```

No	SE	AreaSalud	ETA	Fuente	Casos
0	1	Guate Noroccidente/	Guate Central	Intoxicación Alimentaria Pan con Pollo	135
1	2	Petén Norte	Intoxicación Alimentaria	envuelto de coliflor	20
2	3	Quetzaltenango	Rotavirus		17
3	5	Quetzaltenango	Intoxicación Alimentaria	carne asada	28
4	6	Alta Verapaz	Intoxicación Alimentaria	caldo de pollo	23
5	7	Huehuetenango	Enfermedad Diarreica	agua contaminada	39
6	8	Chiquimula	Intoxicación Alimentaria	Hot Dog	5
7	9	Huehuetenango	Rotavirus	-	42
8	10	Totonicapán	Fiebre Tifoidea	-	10
9	11	Suchitepéquez	Intoxicación Alimentaria	Caldo de pollo	34
10	12	Sololá	Intoxicación Alimentaria Hongo Amanita	Muscaria	5
11	13	Huehuetenango	Intoxicación Alimentaria	frijoles	10
12	14	Suchitepéquez	Intoxicación Alimentaria carne asada, ensalada	de coditos	32
13	15	Chiquimula	Intoxicación Alimentaria	panes con pollo	120
14	16	Escuintla	Hepatitis A	-	10
15	17	Guatemala Sur	Intoxicación Alimentaria tortilla con pollo, ...	y crema	171
16	18	Guatemala Sur	Intoxicación Alimentaria	ceviche	16
17	19	Escuintla	Fiebre Tifoidea	-	7
18	20	Chimaltenango	Intoxicación Alimentaria	Estofado de Pollo	19
19	21	Escuintla	Hepatitis A	-	3
20	22	Alta Verapaz	Intoxicación Alimentaria	Frijoles, café	8