

```

1 #####-
2 #' ---
3 #' title: 'Preprocesamiento de Datos de FRE'
4 #' author: 'Daniel S. Parra G.'
5 #' date: '01-01-2021'
6 #' ---
7 ## Propósito del Script: preprocesamiento de datos de FRE
8 ##
9 ##
10 ## Copyright (c) Fondo Nacional de Estupefacientes, 2021
11 ##
12 ## Email: dsparra@minsalud.gov.co
13 #####-
14
15 require(tidyverse)
16 require(readxl)
17 require(lubridate)
18 require(reticulate)
19 require(fuzzywuzzyR)
20
21 #'-----
22 # 1. Lectura de datos base -----
23 #'-----
24 #
25 data <- read_excel(file.path('data', 'intermediate',
26                             '001_Herramienta_Limpieza_Manual.xlsx'),
27                   na = c('', 'NA', 'N.A', 'No aplica'))
28
29 df_DIVIPOLA <-
30   read_csv(file.path('data', 'processed', '798_DANE_DEPARTAMENTO.csv'))
31 df_MUNICIPIO <-
32   read_csv(file.path('data', 'processed', '799_DANE_DIVIPOLA.csv'),
33           locale = locale(encoding = 'latin1'))
34
35 #'-----
36 # 2. Manipulación de fechas -----
37 #'-----
38
39 data$`Fecha creación FRE` %>% supply(function(x) as_date(x))
40
41 #'-----
42 # 3. Cambios de carácter -----
43 #'-----
44 # Obtener valores cercanos
45 data$Departamento_1 <- data$Departamento %>%
46   map_chr(~ str_to_upper(.x) %>%
47     {
48       fuzzywuzzyR::GetCloseMatches(., df_DIVIPOLA$NombreDepartamento,
49                                     cutoff = 0.85)
50     } %>%
51     {
52       `[`(1)
53     })
54 # Unir códigos de departamento
55 data <- data %>%
56   left_join(df_DIVIPOLA, by = c('Departamento_1' = 'NombreDepartamento'))
57
58 # Unir códigos de municipio
59 data$Ciudad1 <- data$Ciudad %>%
60   map_chr(

```

```

61 ~ str_to_upper(.x) %>%
62   fuzzywuzzyR::GetCloseMatches(., df_MUNICIPIO$NOMBRE_MUNICIPIO,
63     cutoff = 0.6) %>%
64   `[`(1)
65 )
66
67 # Cambio en formato de costo de adquisición del Recetario
68 data$`3.06 Costo de adquisición del recetario (COP)` <-
69   data$`3.06 Costo de adquisición del recetario (COP)` %>%
70   str_replace_all(., '(\\$|\\.)', '') %>%
71   str_replace_all(., 'Por definir', '')
72
73 data %>%
74   select(matches('Precio de'))
75
76 data <- data %>%
77   mutate(across(matches('Precio de'),
78     ~.x %>%
79       str_replace_all(., '(\\$)', '') %>%
80       str_replace_all(., '(Por definir|No saben)', '')))
81
82
83
84 #'-----
85 # 2. -----
86 #'-----
87
88 data <- data %>%
89   mutate(across(CodigoDepartamento, ~formatC(.x, width = 2, flag = '0')))
90
91 data1 <- data %>%
92   mutate(across(matches('Nombre del'), ~str_to_title(.x)))
93
94
95 write_csv(data1, file.path('data', 'processed', '001_Herramienta_Procesada.csv'))
96

```