

```

1  ### SCRIPT 1 - Leer y juntar Excels de OpenPLACSP ###
2  #
3  # Lee los excels de todos los años de OpenPLACSP y los junta en
4  # un unico datafame
5  #
6  #####
7  library(ggplot2)
8  library(dplyr)
9  library(openxlsx)
10 library(RODBC)
11 library(qanteda) #paquete recomendado en todos laos
12 library(qanteda.textmodels) #aux
13 library(qanteda.textstats) # aux
14 library(readtext) # sencilla manera de leer data de texto en R
15 library(spacyr) # NLP usando la libreria spaCy, incluyendo etiquetado part-of-speech,
entity recognition y dependency parsing.
16 library(devtools)
17 library(zoo)
18 library(DBI)
19 library(tidyverse)
20 library(stringi)
21 library(doParallel)
22
23 #setup parallel backend to use many processors
24 cores=detectCores()
25 cl <- makeCluster(cores[1]-1, type = "PSOCK", outfile="log.txt") #not to overload your
computer
26 registerDoParallel(cl)
27
28 #Ruta para PLACSP
29 setwd("C:/Users/guillermo.alonso/Desktop/Tesis Guillermo/_BD actualizada/Datos
atom/PLACSP")
30
31 ## Inicializo la BD
32 BD_PLACSP<-NULL
33
34 ### CARGA DE DATOS PLACSP #####
35
36 ##Cargo datos 2012 PCSP
37 datos2012_PLACSP<-readWorkbook("Licitaciones_PLACSP_2012.xlsx",sheet = 2,colNames = T,
rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
38 ##Aumento la BD
39 BD_PLACSP<-datos2012_PLACSP
40
41 ##Cargo datos 2013 PCSP
42 datos2013_PLACSP<-readWorkbook("Licitaciones_PLACSP_2013.xlsx",sheet = 2,colNames = T,
rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
43 ##Aumento la BD
44 BD_PLACSP<-rbind(BD_PLACSP,datos2013_PLACSP)
45
46 ##Cargo datos 2014 PCSP
47 datos2014_PLACSP<-readWorkbook("Licitaciones_PLACSP_2014.xlsx",sheet = 2,colNames = T,
rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
48 ##Aumento la BD
49 BD_PLACSP<-rbind(BD_PLACSP,datos2014_PLACSP)
50
51 ##Cargo datos 2015 PCSP
52 datos2015_PLACSP<-readWorkbook("Licitaciones_PLACSP_2015.xlsx",sheet = 2,colNames = T,
rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
53 ##Aumento la BD
54 BD_PLACSP<-rbind(BD_PLACSP,datos2015_PLACSP)
55
56 ##Cargo datos 2016 PCSP
57 datos2016_PLACSP<-readWorkbook("Licitaciones_PLACSP_2016.xlsx",sheet = 2,colNames = T,
rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
58 ##Aumento la BD
59 BD_PLACSP<-rbind(BD_PLACSP,datos2016_PLACSP)
60
61 ##Cargo datos 2017 PCSP
62 datos2017_PLACSP<-readWorkbook("Licitaciones_PLACSP_2017.xlsx",sheet = 2,colNames = T,
rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
63 ##Aumento la BD
64 BD_PLACSP<-rbind(BD_PLACSP,datos2017_PLACSP)

```

```

65
66 ##Cargo datos 2018 PCSP
67 datos2018_PLACSP<-readWorkbook("Licitaciones_PLACSP_2018.xlsx",sheet = 2,colNames = T,
68 rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
69 ##Aumento la BD
70 BD_PLACSP<-rbind(BD_PLACSP,datos2018_PLACSP)
71
72 ##Cargo datos 2019 PCSP
73 datos2019_PLACSP<-readWorkbook("Licitaciones_PLACSP_2019.xlsx",sheet = 2,colNames = T,
74 rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
75 ##Aumento la BD
76 BD_PLACSP<-rbind(BD_PLACSP,datos2019_PLACSP)
77
78 ##Cargo datos 2020 PCSP
79 datos2020_PLACSP<-readWorkbook("Licitaciones_PLACSP_2020.xlsx",sheet = 2,colNames = T,
80 rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
81 ##Aumento la BD
82 BD_PLACSP<-rbind(BD_PLACSP,datos2020_PLACSP)
83
84 ##Cargo datos 2021 PCSP
85 datos2021_PLACSP<-readWorkbook("Licitaciones_PLACSP_2021.xlsx",sheet = 2,colNames = T,
86 rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
87 ##Aumento la BD
88 BD_PLACSP<-rbind(BD_PLACSP,datos2021_PLACSP)
89
90 ##Cargo datos 2022 PCSP
91 datos2022_PLACSP<-readWorkbook("Licitaciones_PLACSP_2022.xlsx",sheet = 2,colNames = T,
92 rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
93 ##Aumento la BD
94 BD_PLACSP<-rbind(BD_PLACSP,datos2022_PLACSP)
95
96 ##Cargo datos 2023 PCSP
97 datos2023_PLACSP<-readWorkbook("Licitaciones_PLACSP_2023.xlsx",sheet = 2,colNames = T,
98 rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
99 ##Aumento la BD
100 BD_PLACSP<-rbind(BD_PLACSP,datos2023_PLACSP)
101
102 ##Cargo datos 2025 PCSP
103 datos2025_PLACSP<-readWorkbook("Licitaciones_PLACSP_2025.xlsx",sheet = 2,colNames = T,
104 rowNames = F, detectDates = F,skipEmptyRows = F,skipEmptyCols = F, na.strings = "NA")
105 ##Aumento la BD
106 BD_PLACSP<-rbind(BD_PLACSP,datos2025_PLACSP)
107
108 # Paso 1: Eliminar acentos
109 names(BD_PLACSP) <- stringi::stri_trans_general(names(BD_PLACSP), "Latin-ASCII")
110
111 # Paso 2: Reemplazar "/" por "_"
112 names(BD_PLACSP) <- gsub("/", "_", names(BD_PLACSP))
113
114 # Verificar y manejar nombres duplicados
115 names(BD_PLACSP) <- make.unique(names(BD_PLACSP), sep = "_")
116
117 # Reemplazar ";" por ":" en todo el data frame
118 BD_PLACSP <- BD_PLACSP %>%
119   mutate(across(everything(), ~ str_replace_all(., ";", ":")))
120
121 #Corrijo el formato de las fechas
122 # Función para convertir fechas de Excel a formato Date en R
123 convertir_excel_a_fecha <- function(columna) {
124   as.Date(as.numeric(columna), origin = "1899-12-30")
125 }
126
127 BD_PLACSP$Fecha.actualizacion <- convertir_excel_a_fecha(BD_PLACSP$`Fecha.actualizacion`)
128 BD_PLACSP$Fecha.de.presentacion.de.ofertas <- convertir_excel_a_fecha(BD_PLACSP$`Fecha.de.presentacion.de.ofertas`)

```

```
128 BD_PLACSP$Fecha.de.presentacion.de.solicitudes.de.participacion <-  
convertir_excel_a_fecha(BD_PLACSP$  
`Fecha.de.presentacion.de.solicitudes.de.participacion`)  
129 BD_PLACSP$Fecha.del.acuerdo.licitacion_lote <- convertir_excel_a_fecha(BD_PLACSP$  
`Fecha.del.acuerdo.licitacion_lote`)  
130 BD_PLACSP$Fecha.formalizacion.del.contrato.licitacion_lote <- convertir_excel_a_fecha(  
BD_PLACSP$`Fecha.formalizacion.del.contrato.licitacion_lote`)  
131 BD_PLACSP$Fecha.entrada.en.vigor.del.contrato.de.licitacion_lote <-  
convertir_excel_a_fecha(BD_PLACSP$  
`Fecha.entrada.en.vigor.del.contrato.de.licitacion_lote`)  
132  
133  
134 #Guardo Excel de PLACSP  
135 write.csv2(BD_PLACSP,"BD_PLACSP.csv", col.names = T, row.names=F,na="NA",dec=". ",sep=  
";")  
136 #Guardo por separado en RData  
137 save(BD_PLACSP,file="BD_PLACSP.RData")
```