

Reporte de Revisión de Caracteres en Base de Datos de Estados Financieros SEPS

Pedro Galindo Vera

2023-05-03

Importación de la base de datos

```
library(readr)
```

```
## Warning: package 'readr' was built under R version 4.2.3
```

```
ruta <- "D:\\INNOVACION\\PASANTE\\DESARROLLO\\BalanceFinacieroSFN\\data\\Base de Datos\\SEPS Bases de D  
SEPS <- read_csv(ruta, col_types = cols(FECHA = col_date(), CODIGO = col_integer()))
```

Nombres encontrados en el campo RAZON_SOCIAL

```
nombres <- sort(unique(SEPS$RAZON_SOCIAL))
```

Identificación de caracteres a cambiar

Excluyendo los caracteres alfabéticos y de separación se tiene

```
nombres[grepl("[^[:alpha:] -./]", nombres)]
```

```
## [1] "COOPERATIVA DE AHORRO Y CRÉDITO PROVIDA LTDA"  
## [2] "COOPERATIVA DE AHORRO Y CRÉDITO SAN ANTONIO LTDA LOS RÍOS"  
## [3] "COOPERATIVA DE AHORRO Y CRÉDITO SOLIDARIDAD, EMPRENDIMIENTO Y COOPERACIÓN"  
## [4] "COOPERATIVA DE AHORRO Y CRÉDITO YUYAY LTDA"  
## [5] "COOPERATIVA DE AHORRO Y CREDITO 1 DE JULIO"  
## [6] "COOPERATIVA DE AHORRO Y CREDITO 10 DE SEPTIEMBRE EN LIQUIDACION"  
## [7] "COOPERATIVA DE AHORRO Y CREDITO 13 DE ABRIL"  
## [8] "COOPERATIVA DE AHORRO Y CREDITO 14 DE MAYO"  
## [9] "COOPERATIVA DE AHORRO Y CREDITO 15 DE ABRIL LTDA"  
## [10] "COOPERATIVA DE AHORRO Y CREDITO 15 DE AGOSTO DE PILACOTO"  
## [11] "COOPERATIVA DE AHORRO Y CREDITO 15 DE AGOSTO LTDA"  
## [12] "COOPERATIVA DE AHORRO Y CREDITO 15 DE JUNIO LTDA"  
## [13] "COOPERATIVA DE AHORRO Y CREDITO 15 DE MAYO LTDA"
```

```
## [14] "COOPERATIVA DE AHORRO Y CREDITO 16 DE JULIO LTDA"
## [15] "COOPERATIVA DE AHORRO Y CREDITO 16 DE JUNIO"
## [16] "COOPERATIVA DE AHORRO Y CREDITO 17 DE MARZO LIMITADA"
## [17] "COOPERATIVA DE AHORRO Y CREDITO 18 DE NOVIEMBRE"
## [18] "COOPERATIVA DE AHORRO Y CREDITO 20 DE FEBRERO LTDA"
## [19] "COOPERATIVA DE AHORRO Y CREDITO 21 DE NOVIEMBRE LTDA EN LIQUIDACION"
## [20] "COOPERATIVA DE AHORRO Y CREDITO 22 DE JUNIO"
## [21] "COOPERATIVA DE AHORRO Y CREDITO 23 DE ENERO"
## [22] "COOPERATIVA DE AHORRO Y CREDITO 23 DE JULIO LIMITADA"
## [23] "COOPERATIVA DE AHORRO Y CREDITO 23 DE MAYO LTDA EN LIQUIDACION"
## [24] "COOPERATIVA DE AHORRO Y CREDITO 26 DE JULIO LTDA"
## [25] "COOPERATIVA DE AHORRO Y CREDITO 26 DE SEPTIEMBRE LAZARO CONDO"
## [26] "COOPERATIVA DE AHORRO Y CRÉDITO 26 DE SEPTIEMBRE LAZARO CONDO LTDA"
## [27] "COOPERATIVA DE AHORRO Y CREDITO 27 DE NOVIEMBRE"
## [28] "COOPERATIVA DE AHORRO Y CREDITO 29 DE ENERO CHAGUARPAMBA"
## [29] "COOPERATIVA DE AHORRO Y CREDITO 29 DE ENERO CHAGUARPAMBA EN LIQUIDACION"
## [30] "COOPERATIVA DE AHORRO Y CREDITO 29 DE JUNIO LTDA"
## [31] "COOPERATIVA DE AHORRO Y CREDITO 29 DE OCTUBRE LTDA"
## [32] "COOPERATIVA DE AHORRO Y CREDITO 4 DE OCTUBRE"
## [33] "COOPERATIVA DE AHORRO Y CREDITO 4 DE OCTUBRE SAN FRANCISCO DE CHAMBO"
## [34] "COOPERATIVA DE AHORRO Y CREDITO 5 DE ENERO DE INTENDENCIA LTDA"
## [35] "COOPERATIVA DE AHORRO Y CREDITO 9 DE OCTUBRE LTDA"
## [36] "COOPERATIVA DE AHORRO Y CREDITO BAÃ'OS LTDA"
## [37] "COOPERATIVA DE AHORRO Y CREDITO CAÃ'AR LTDA"
## [38] "COOPERATIVA DE AHORRO Y CREDITO CRECER WIÃ'ARI LTDA"
## [39] "COOPERATIVA DE AHORRO Y CREDITO CREDICOOP 10 DE FEBRERO"
## [40] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA BIBLIAN LIMITADA"
## [41] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA CACPE LOJA LTDA"
## [42] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA CACPE YANTZAZA LTDA"
## [43] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA CACPE ZAMORA CHINCHIPE LTDA."
## [44] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA DE COTOPAXI LIMITADA"
## [45] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA DE PASTAZA LIMITADA"
## [46] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA GUALAQUIZA"
## [47] "COOPERATIVA DE AHORRO Y CREDITO DE LOS EMPLEADOS DE SERVIENTREGA SA; TERCERIZADORAS CONTRATISTAS"
## [48] "COOPERATIVA DE AHORRO Y CREDITO FASAYÃ'AN LTDA"
## [49] "COOPERATIVA DE AHORRO Y CREDITO GAÃ'ANSOL LTDA"
## [50] "COOPERATIVA DE AHORRO Y CREDITO LA LIBERTAD 3"
## [51] "COOPERATIVA DE AHORRO Y CREDITO PABLO MUÃ'OZ VEGA LIMITADA"
## [52] "COOPERATIVA DE AHORRO Y CREDITO SEÃ'OR DE GIRON"
## [53] "COOPERATIVA DE AHORRO Y CREDITO TEXTIL 14 DE MARZO"
```

Hallazgos

Observando la lista anterior identificamos los siguientes caracteres extraños

```
nombres[grepl("Ã'", nombres)]
```

```
## [1] "COOPERATIVA DE AHORRO Y CREDITO BAÃ'OS LTDA"
## [2] "COOPERATIVA DE AHORRO Y CREDITO CAÃ'AR LTDA"
## [3] "COOPERATIVA DE AHORRO Y CREDITO CRECER WIÃ'ARI LTDA"
## [4] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA BIBLIAN LIMITADA"
## [5] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA CACPE LOJA LTDA"
## [6] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÃ'A EMPRESA CACPE YANTZAZA LTDA"
```

```
## [7] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE ZAMORA CHINCHIPE LTDA."
## [8] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA DE COTOPAXI LIMITADA"
## [9] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA DE PASTAZA LIMITADA"
## [10] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA GUALAQUIZA"
## [11] "COOPERATIVA DE AHORRO Y CREDITO FASAYÁN LTDA"
## [12] "COOPERATIVA DE AHORRO Y CREDITO GAÑANSOL LTDA"
## [13] "COOPERATIVA DE AHORRO Y CREDITO PABLO MUÑOZ VEGA LIMITADA"
## [14] "COOPERATIVA DE AHORRO Y CREDITO SEÑOR DE GIRON"
```

```
nombres[grepl("Ã%", nombres)]
```

```
## [1] "COOPERATIVA DE AHORRO Y CRÃDITO PROVIDA LTDA"
## [2] "COOPERATIVA DE AHORRO Y CRÃDITO SAN ANTONIO LTDA LOS RÃ\u00d0S"
## [3] "COOPERATIVA DE AHORRO Y CRÃDITO SOLIDARIDAD, EMPRENDIMIENTO Y COOPERACIÃ"N"
## [4] "COOPERATIVA DE AHORRO Y CRÃDITO YUYAY LTDA"
```

```
nombres[grepl("Ã\u00d", nombres)]
```

```
## [1] "COOPERATIVA DE AHORRO Y CRÃDITO SAN ANTONIO LTDA LOS RÃ\u00d0S"
```

```
nombres[grepl("Ã", nombres)]
```

```
## [1] "COOPERATIVA DE AHORRO Y CRÃDITO SOLIDARIDAD, EMPRENDIMIENTO Y COOPERACIÃ"N"
```

Correcciones plausibles

```
nombres <- gsub("Ã'", "Ñ", nombres)
nombres <- gsub("Ã%", "É", nombres)
nombres <- gsub("Ã\u00d", "Í", nombres)
nombres <- gsub("Ã", "Ó", nombres)
```

Verificación

```
nombres[grepl("[^[:alpha:] -./]", nombres)]
```

```
## [1] "COOPERATIVA DE AHORRO Y CREDITO 1 DE JULIO"
## [2] "COOPERATIVA DE AHORRO Y CREDITO 10 DE SEPTIEMBRE EN LIQUIDACION"
## [3] "COOPERATIVA DE AHORRO Y CREDITO 13 DE ABRIL"
## [4] "COOPERATIVA DE AHORRO Y CREDITO 14 DE MAYO"
## [5] "COOPERATIVA DE AHORRO Y CREDITO 15 DE ABRIL LTDA"
## [6] "COOPERATIVA DE AHORRO Y CREDITO 15 DE AGOSTO DE PILACOTO"
## [7] "COOPERATIVA DE AHORRO Y CREDITO 15 DE AGOSTO LTDA"
## [8] "COOPERATIVA DE AHORRO Y CREDITO 15 DE JUNIO LTDA"
## [9] "COOPERATIVA DE AHORRO Y CREDITO 15 DE MAYO LTDA"
## [10] "COOPERATIVA DE AHORRO Y CREDITO 16 DE JULIO LTDA"
## [11] "COOPERATIVA DE AHORRO Y CREDITO 16 DE JUNIO"
## [12] "COOPERATIVA DE AHORRO Y CREDITO 17 DE MARZO LIMITADA"
```

```
## [13] "COOPERATIVA DE AHORRO Y CREDITO 18 DE NOVIEMBRE"
## [14] "COOPERATIVA DE AHORRO Y CREDITO 20 DE FEBRERO LTDA"
## [15] "COOPERATIVA DE AHORRO Y CREDITO 21 DE NOVIEMBRE LTDA EN LIQUIDACION"
## [16] "COOPERATIVA DE AHORRO Y CREDITO 22 DE JUNIO"
## [17] "COOPERATIVA DE AHORRO Y CREDITO 23 DE ENERO"
## [18] "COOPERATIVA DE AHORRO Y CREDITO 23 DE JULIO LIMITADA"
## [19] "COOPERATIVA DE AHORRO Y CREDITO 23 DE MAYO LTDA EN LIQUIDACION"
## [20] "COOPERATIVA DE AHORRO Y CREDITO 26 DE JULIO LTDA"
## [21] "COOPERATIVA DE AHORRO Y CREDITO 26 DE SEPTIEMBRE LAZARO CONDO"
## [22] "COOPERATIVA DE AHORRO Y CRÉDITO 26 DE SEPTIEMBRE LAZARO CONDO LTDA"
## [23] "COOPERATIVA DE AHORRO Y CREDITO 27 DE NOVIEMBRE"
## [24] "COOPERATIVA DE AHORRO Y CREDITO 29 DE ENERO CHAGUARPAMBA"
## [25] "COOPERATIVA DE AHORRO Y CREDITO 29 DE ENERO CHAGUARPAMBA EN LIQUIDACION"
## [26] "COOPERATIVA DE AHORRO Y CREDITO 29 DE JUNIO LTDA"
## [27] "COOPERATIVA DE AHORRO Y CREDITO 29 DE OCTUBRE LTDA"
## [28] "COOPERATIVA DE AHORRO Y CREDITO 4 DE OCTUBRE"
## [29] "COOPERATIVA DE AHORRO Y CREDITO 4 DE OCTUBRE SAN FRANCISCO DE CHAMBO"
## [30] "COOPERATIVA DE AHORRO Y CREDITO 5 DE ENERO DE INTENDENCIA LTDA"
## [31] "COOPERATIVA DE AHORRO Y CREDITO 9 DE OCTUBRE LTDA"
## [32] "COOPERATIVA DE AHORRO Y CREDITO CREDICOOP 10 DE FEBRERO"
## [33] "COOPERATIVA DE AHORRO Y CREDITO DE LOS EMPLEADOS DE SERVIENTREGA SA; TERCERIZADORAS CONTRATISTAS"
## [34] "COOPERATIVA DE AHORRO Y CREDITO LA LIBERTAD 3"
## [35] "COOPERATIVA DE AHORRO Y CREDITO TEXTIL 14 DE MARZO"
```

Mejora en la expresión regular

Para agregar comillas a la expresión regular, asegúrate de colocar el carácter `-` al final del rango de caracteres dentro de los corchetes `[]`. De lo contrario, R interpretará el carácter `-` como un rango de caracteres y generará un error.

```
nombres <- sort(unique(SEPS$RAZON_SOCIAL))
expresion_regular <- "[^[:alpha:][:digit:] .,;/\\-]"
prueba <- grepl(expresion_regular,nombres)
nombres[prueba]
```

```
## [1] "COOPERATIVA DE AHORRO Y CRÉDITO PROVIDA LTDA"
## [2] "COOPERATIVA DE AHORRO Y CRÉDITO SAN ANTONIO LTDA LOS RÍOS"
## [3] "COOPERATIVA DE AHORRO Y CRÉDITO SOLIDARIDAD, EMPRENDIMIENTO Y COOPERACIÓN"
## [4] "COOPERATIVA DE AHORRO Y CRÉDITO YUYAY LTDA"
## [5] "COOPERATIVA DE AHORRO Y CREDITO BAÑOS LTDA"
## [6] "COOPERATIVA DE AHORRO Y CREDITO CAÑAR LTDA"
## [7] "COOPERATIVA DE AHORRO Y CREDITO CRECER WÍARI LTDA"
## [8] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA BIBLIAN LIMITADA"
## [9] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE LOJA LTDA"
## [10] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE YANTZAZA LTDA"
## [11] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE ZAMORA CHINCHIPE LTDA."
## [12] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA DE COTOPAXI LIMITADA"
## [13] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA DE PASTAZA LIMITADA"
## [14] "COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA GUALAQUIZA"
## [15] "COOPERATIVA DE AHORRO Y CREDITO FASAYÁN LTDA"
## [16] "COOPERATIVA DE AHORRO Y CREDITO GAÑANSOL LTDA"
## [17] "COOPERATIVA DE AHORRO Y CREDITO PABLO MUÑOZ VEGA LIMITADA"
## [18] "COOPERATIVA DE AHORRO Y CREDITO SEÑOR DE GIRON"
```

Observación. Al agregarse en la expresión general las comillas " para que las omitiera `grepl` también dejó de reconocerse a `Ã`.

Identificación de archivos con caracteres mal identificados

Muestreo para identificación por el año, los archivos con caracteres mal identificados

```
tamanho_muestra <- round(0.05*nrow(SEPS))
indices_muestra <- sample(1:nrow(SEPS),tamanho_muestra)
tabla_muestra <- SEPS[indices_muestra,]
expresion_regular <- "[^[:alpha:][:digit:] .,;/-]"
prueba <- grepl(expresion_regular,tabla_muestra$RAZON_SOCIAL)
tabla_muestra_coincidencias <- tabla_muestra[prueba,]
numero_coincidencias <- sum(prueba)
tasa_coincidencias <- 100*sum(prueba)/tamanho_muestra
```

Coincidencias encontradas en la muestra: 88890 de 1.272308×10^6

Tasa de coincidencias: 6.9865158%

Ejemplo de hallazgos

```
library(knitr)
kable(tabla_muestra_coincidencias[1:20,c("FECHA","RAZON_SOCIAL")])
```

FECHA	RAZON_SOCIAL
2017-07-31	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA DE COTOPAXI LIMITADA
2017-06-30	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA BIBLIAN LIMITADA
2021-03-31	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE ZAMORA CHINCHIPE LTDA.
2020-08-31	COOPERATIVA DE AHORRO Y CRÉDITO SOLIDARIDAD, EMPRENDIMIENTO Y COOPERACIÓN
2015-10-31	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE ZAMORA CHINCHIPE LTDA.
2020-09-30	COOPERATIVA DE AHORRO Y CREDITO CAÑAR LTDA
2019-08-31	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA BIBLIAN LIMITADA
2016-10-31	COOPERATIVA DE AHORRO Y CREDITO GAÑANSOL LTDA
2017-09-30	COOPERATIVA DE AHORRO Y CREDITO PABLO MUÑOZ VEGA LIMITADA
2017-05-31	COOPERATIVA DE AHORRO Y CREDITO GAÑANSOL LTDA
2019-01-31	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA DE COTOPAXI LIMITADA
2020-11-30	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE LOJA LTDA
2019-04-30	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA BIBLIAN LIMITADA

FECHA	RAZON_SOCIAL
2016-04-30	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE ZAMORA CHINCHIPE LTDA.
2018-01-31	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA DE COTOPAXI LIMITADA
2021-06-30	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA GUALAQUIZA
2016-11-30	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE ZAMORA CHINCHIPE LTDA.
2021-07-31	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE YANTZAZA LTDA
2020-04-30	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA DE PASTAZA LIMITADA
2016-02-29	COOPERATIVA DE AHORRO Y CREDITO DE LA PEQUEÑA EMPRESA CACPE LOJA LTDA

Fechas donde se presentan caracteres mal identificados

```
sort(unique(tabla_muestra_coincidencias$FECHA))
```

```
## [1] "2015-02-28" "2015-03-31" "2015-04-30" "2015-05-31" "2015-06-30"
## [6] "2015-07-31" "2015-08-31" "2015-09-30" "2015-10-31" "2015-11-30"
## [11] "2015-12-31" "2016-01-31" "2016-02-29" "2016-03-31" "2016-04-30"
## [16] "2016-05-31" "2016-06-30" "2016-07-31" "2016-08-31" "2016-09-30"
## [21] "2016-10-31" "2016-11-30" "2016-12-31" "2017-01-31" "2017-02-28"
## [26] "2017-03-31" "2017-04-30" "2017-05-31" "2017-06-30" "2017-07-31"
## [31] "2017-08-31" "2017-09-30" "2017-10-31" "2017-11-30" "2017-12-31"
## [36] "2018-01-31" "2018-02-28" "2018-03-31" "2018-04-30" "2018-05-31"
## [41] "2018-06-30" "2018-07-31" "2018-08-31" "2018-09-30" "2018-10-31"
## [46] "2018-11-30" "2018-12-31" "2019-01-31" "2019-02-28" "2019-03-31"
## [51] "2019-04-30" "2019-05-31" "2019-06-30" "2019-07-31" "2019-08-31"
## [56] "2019-09-30" "2019-10-31" "2019-11-30" "2019-12-31" "2020-01-31"
## [61] "2020-02-29" "2020-03-31" "2020-04-30" "2020-05-31" "2020-06-30"
## [66] "2020-07-31" "2020-08-31" "2020-09-30" "2020-10-31" "2020-11-30"
## [71] "2020-12-31" "2021-01-31" "2021-02-28" "2021-03-31" "2021-04-30"
## [76] "2021-05-31" "2021-06-30" "2021-07-31" "2021-08-31" "2021-09-30"
## [81] "2021-10-31" "2021-11-30" "2021-12-31" "2022-01-31" "2022-02-28"
## [86] "2022-03-31" "2022-04-30" "2022-05-31" "2022-06-30" "2022-07-31"
## [91] "2022-08-31" "2022-09-30" "2022-10-31" "2022-11-30" "2022-12-31"
## [96] "2023-01-31" "2023-02-28" "2023-03-31"
```

Rango de fechas donde se presentan caracteres mal identificados

```
inferior <- min(tabla_muestra_coincidencias$FECHA)
superior <- max(tabla_muestra_coincidencias$FECHA)
```

Inferior: 2015-02-28

Superior: 2023-03-31

Conclusión

Los errores en la codificación de los caracteres se encuentra a lo largo de todas las tablas procesadas.

Recomendación

Implementar un método para aplicar a cada tabla y corregir los caracteres mal identificados