

Alejandro León Líndez¹, Adrian Lizzadro Pla², Marta Medina Muñiz³¹ Máster en Ciencia de Datos; alelin@alumni.uv.es² Máster en Ciencia de Datos; alizpla@alumni.uv.es³ Máster en Ciencia de Datos; memuiz@alumni.uv.es

* Correspondence:

Simple Summary: Resumen.

Abstract: Abstract

Keywords: keyword 1; keyword 2; keyword 3 (list three to ten pertinent keywords specific to the article, yet reasonably common within the subject discipline.).

1. Introducción

2. Carga de librerías e importación del fichero

```
rm(list=ls()) # Borrar todas las variables al principio
```

```
library(readr)
library(ggplot2)
```

```
gastos <- read_delim("data/Gasto_turistas_internacionales_segun_comunidad_paisresi
```

```
## Rows: 3072 Columns: 6
## -- Column specification -----
## Delimiter: ";"
## chr (5): País de residencia, Total Nacional y CCAA, Tipo de dato, Gastos y d...
## dbl (1): Periodo
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message
```

2.1. Transformación a tidy data

```
tipos_datos <- unique(gastos$`Gastos y duración media de los viajes`)
tipos_datos
```

```
## [1] "Gasto total" "Gasto medio por persona"
## [3] "Gasto medio diario por persona" "Duración media de los viajes"
```

```
gastos1 <- subset(gastos,gastos$`Gastos y duración media de los viajes`== tipos_da
colnames(gastos1)[colnames(gastos1)=="Total"]<- "Gasto_total"
gastos1<- subset(gastos1, select = -`Gastos y duración media de los viajes`)
```

Citation: Título. *Journal Not Specified* 2023, 1, 0. <https://doi.org/>

Received:

Revised:

Accepted:

Published:

Copyright: © 2024 by the authors. Submitted to *Journal Not Specified* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

```
gastos2 <- subset(gastos,gastos$`Gastos y duración media de los viajes`== tipos_da
colnames(gastos2)[colnames(gastos2)=="Total"]<- "Gasto_medio_persona"
gastos2<- subset(gastos2, select = -`Gastos y duración media de los viajes`)
```

```
gastos3 <- subset(gastos,gastos$`Gastos y duración media de los viajes`== tipos_da
colnames(gastos3)[colnames(gastos3)=="Total"]<- "Gasto_medio_diario_persona"
gastos3<- subset(gastos3, select = -`Gastos y duración media de los viajes`)
```

```
gastos4 <- subset(gastos,gastos$`Gastos y duración media de los viajes`== tipos_da
colnames(gastos4)[colnames(gastos4)=="Total"]<- "Duracion_media"
gastos4<- subset(gastos4, select = -`Gastos y duración media de los viajes`)
```

```
# Compruebo que en todas las columnas salvo la última todos las filas son iguales
any(gastos1[1:length(nrow(gastos1)-1)] !=gastos2[1:length(nrow(gastos2)-1)])
```

```
## [1] FALSE
```

18

```
any(gastos3[1:length(nrow(gastos1)-1)] !=gastos4[1:length(nrow(gastos2)-1)])
```

```
## [1] FALSE
```

19

```
# Uno los dos datasets en un único dataset con el que trabajar
```

```
gastos12 <- merge(gastos1,gastos2, by = c(colnames(gastos1[1:length(gastos1)-1])))
gastos34 <- merge(gastos3,gastos4, by = c(colnames(gastos3[1:length(gastos3)-1])))
```

```
# Compruebo que en todas las columnas salvo la última todos las filas son iguales
any(gastos12[1:length(nrow(gastos12)-2)] !=gastos34[1:length(nrow(gastos34)-2)])
```

```
## [1] FALSE
```

20

```
# Uno los dos datasets
```

```
datos <- merge(gastos12,gastos34, by = c(colnames(gastos12[1:4])))
```

```
unique(datos$`Tipo de dato`)
```

```
## [1] "Dato base" "Tasa de variación anual"
```

21

```
# Nos quedamos únicamente con los datos base, quitando las tasas de variación
datos <- subset(datos, datos$`Tipo de dato`== unique(datos$`Tipo de dato`)[1])
# Quito la columna irrelevante
datos <- subset(datos, select = -`Tipo de dato`)
```

2.2. Transformación de clases

22

```
lapply(datos,class)
```

```
## $'País de residencia'
```

23

```
## [1] "character"
```

24

```
##
```

25

```
## $'Total Nacional y CCAA'
```

26

```
## [1] "character"
```

27

```

##
## $Periodo
## [1] "numeric"
##
## $Gasto_total
## [1] "character"
##
## $Gasto_medio_persona
## [1] "character"
##
## $Gasto_medio_diario_persona
## [1] "character"
##
## $Duracion_media
## [1] "character"

# Quitar punto de miles
datos[, 4:ncol(datos)] <- lapply(datos[, 4:ncol(datos)], function(x) gsub("\\.", " ", x))
# Sustituir coma decimal por punto decimal
datos[, 4:ncol(datos)] <- lapply(datos[, 4:ncol(datos)], function(x) gsub(",", ".", x))
# Transformar a numerico
datos[, 4:ncol(datos)] <- lapply(datos[, 4:ncol(datos)], function(x) as.numeric(x))
# Comprobar la clase
lapply(datos,class)

## $'País de residencia'
## [1] "character"
##
## $'Total Nacional y CCAA'
## [1] "character"
##
## $Periodo
## [1] "numeric"
##
## $Gasto_total
## [1] "numeric"
##
## $Gasto_medio_persona
## [1] "numeric"
##
## $Gasto_medio_diario_persona
## [1] "numeric"
##
## $Duracion_media
## [1] "numeric"

# Comprobar si hay datos NA
any(is.na(datos))

## [1] FALSE

```

Abbreviations

The following abbreviations are used in this manuscript:

MDPI Multidisciplinary Digital Publishing Institute
DOAJ Directory of open access journals
TLA Three letter acronym
LD linear dichroism

67

Appendix C

68

Appendix C.1

69

The appendix is an optional section that can contain details and data supplemental to the main text. For example, explanations of experimental details that would disrupt the flow of the main text, but nonetheless remain crucial to understanding and reproducing the research shown; figures of replicates for experiments of which representative data is shown in the main text can be added here if brief, or as Supplementary data. Mathematical proofs of results not central to the paper can be added as an appendix.

70

71

72

73

74

75

Appendix D

76

All appendix sections must be cited in the main text. In the appendixes, Figures, Tables, etc. should be labeled starting with ‘A’, e.g., Figure A1, Figure A2, etc.

77

78

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

79

80

81