

Universidad Europea de Madrid

Escuela de Arquitectura, Ingeniería y Diseño



Actividad 1

Caso de uso de BUSINESS ANALYTICS

Profesor: Christian Vladimir Sucuzhanay Arévalo

Alumno: Laura Teruel Montero

Enunciado

Basados en lo aprendido hasta la fecha y en la práctica realizada en clase, deberá responder a las preguntas que cumpla los siguientes requerimientos:

Un empresario del sector de la hostelería pretende incursionar en el sector de la energía, concretamente en la instalación de gasolineras low-cost, para ello el requiere tomar decisiones de tipo empresarial, que deberá responder usted como en la analista de negocio, este empresario tiene 2 millones de euros y desea invertir en la instalación de gasolineras, pero no quiere pagar franquicias ya que las mismas están sujetas a muchas restricciones, pagos permanente de royalties, no tienen independencia en la toma de decisiones , por lo que se pide:

a. De fuentes veraces, lea los archivos que se indican en el anexo, como

- i. Limpie el dataset (la información debe estar correctamente formateada, por ej. lo que es de tipo texto no debe tener otro tipo que no sea texto) , ponga el formato correcto en los números, etc., etc. genere un informe y explique si encuentra alguna anomalía, en el punto

```
-- Column specification -----
:ols(
  .default = col_double(),
  c_p = col_character(),
  direccion = col_character(),
  horario = col_character(),
  localidad = col_character(),
  margen = col_character(),
  municipio = col_character(),
  precio_hidrogeno = col_logical(),
  provincia = col_character(),
  remision = col_character(),
  rotulo = col_character(),
  tipo_venta = col_character(),
  id_provincia = col_character(),
  idccaa = col_character()
)
i Use spec() for the full column specifications.
```

He tenido que cambiar de forma manual el dato del hidrógeno a double después de llevar a cabo la limpieza.

- ii. cree una columna nueva que deberá llamarse low-cost, y determine cuál es el precio promedio de todos los combustibles a nivel comunidades autónomas tanto para el territorio peninsular e insular, esta columna deberá clasificar las estaciones por lowcost y no lowcost,

```
ds_lowcost <- ds_clean %>%
  group_by(idccaa) %>%
  # Calculo la media por ccaa y clasifico en low cost o no low cost
  mutate(low_cost = (precio_gasoleo_a < mean(precio_gasoleo_a,
na.rm=TRUE))) %>%
  ungroup()
```

low_cost
FALSE
FALSE
FALSE
FALSE
TRUE
FALSE
FALSE
FALSE
FALSE
TRUE
FALSE
FALSE
FALSE

- iii. conseguidos objetivos anteriores, debe guardar este "archivo" en una nueva tabla llamada low-cost_num_expediente y deberá estar disponible también en su repositorio de Github con el mismo nombre y formato csv.

```
write_csv(ds_lowcost, "lowcost_21964596.csv")
```

- b. **Este empresario tiene sus residencias habituales en Madrid y Benalmádena , por lo que, en principio le gustaría implantar en cualquiera de las dos antes citadas, y para ello quiere saber:**

- i. cuántas gasolineras tiene la comunidad de Madrid y la comunidad de Andalucía, cuántas son low-cost, cuantas no lo son

CCAA	Nº total de gasolineras	Nº LowCost	Nº no LowCost
Madrid	793	239	554
Andalucía	2207	954	1253

- ii. Además, necesita saber cuál es el precio promedio, el precio más bajo y el más

Como he ejecutado el código en días distintos hoy la información que obtengo es la siguiente:

CCAA	Gasóleo A (€/L)			Gasóleo premium (€/L)			Gasolina 95 E5 (€/L)		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
MAD	1.21	1.56	1.46	1.27	1.64	1.55	1.34	1.68	1.57
AND	1.08	1.66	1.41	1.26	1.63	1.52	1.169	1.81	1.533

- iii. Conseguido el objetivo, deberá guardar este "archivo" en una nueva tabla llamada informe_CAM_expediente y deberá estar disponible también en su repositorio con el mismo nombre en formato csv

```
write_csv(informeB, "informe_CAM_21964596.csv")
```

- c. **Por sí la comunidad de Madrid no se adapta a sus requerimientos, el empresario también quiere:**

- i. conocer a nivel provincias, cuántas gasolineras son low-cost, cuantas no lo son, cuál es el precio promedio, el precio más bajo y el más caro de los siguientes carburantes: gasóleo A, y gasolina 95 e5 Premium , en todo el TERRITORIO NACIONAL, exceptuando las grandes CIUDADES ESPAÑOLAS ("MADRID", "BARCELONA", "SEVILLA" y "VALENCIA")

Dato	Valor
Nº Gasolineras LowCost	3699

Nº Gasolineras no LowCost	5710
Min Gasóleo A	0.971
Max Gasóleo A	1.659
Mean Gasóleo A	1.408
Min Gasolina 95 E5 Premium	1.149
Max Gasolina 95 E5 Premium	1.799
Mean Gasolina 95 E5 Premium	1.626

- ii.* Conseguido el objetivo, deberá guardar este “archivo” en una nueva tabla llamada **informe_no_grandes_ciudades_expediente** y deberá estar disponible también en su repositorio con el mismo nombre en formato Excel

```
write_excel_csv(informeC, "informe_no_grandes_ciudades_21964596.csv")
```

- d.** Determine :

- i.* que gasolineras se encuentran abiertas las 24 horas exclusivamente, genere una nueva tabla llamada **no_24_horas** sin la variable horario (es decir no debe aparecer esta columna).

```

no_24_horas 6648 obs. of 32 variables
  $ c_p      :
  $ direccion :
  $ latitud  :

```

- ii.* Conseguido el objetivo, deberá guardar este “archivo” en una nueva tabla llamada **no_24_horas** y deberá estar disponible también en su repositorio con el mismo nombre en formato Excel

```
write_excel_csv(no_24_horas, "no_24_horas.xls")
```