

PEC2 - Visualización de datos

María Isabel Moreno Carreño

Contents

1. Introducción	1
1 Cargar los archivos de datos	1
2. Extracción de indicadores de los datos para la elaboración.	2
3. Extracción de los datos para España y la UE.	3
4. Creación del ranking del porcentaje de uso de la IA.	5
5. Creación de tabla de datos	5
Creación Bullet Graph.	5
6. Creación de Voronoi	7

1. Introducción

Para realizar la **Infografía** y el **Bullet Graph** se va a utilizar el mismo conjunto de datos, por una parte los indicadores que mide el rendimiento de una de las cuatro dimensiones del Indice de Economía y Sociedad Digital (DESI) relacionadas con la Inteligencia Artificial (IA) en Europa, y por otra parte el porcentaje de adopción de la IA en las empresas.

Son la adopción de la IA por todas las empresas (con 10 o más empleados), excluyendo el sector financiero.

El objetivo de la infografía será presentar el indicador, lo que es y el porcentaje de cumplimiento de las empresas en la Unión Europea en 2025.

El gráfico **Bullet Graph** se usará para comparar el objetivo de adopción de la IA en las empresas españolas comparado con la media de la Unión Europea (UE) para el año 2025 y el objetivo final para 2030.

Este gráfico es ideal para mostrar el cumplimiento de objetivos, se asemeja a un termómetro.

El indicador de barras horizontales se utiliza para comparar una medida con un valor preestablecido, mostrando además el rango cualitativo en el que se encuentra el valor medido. También indica si el valor es alto, medio o bajo.

Para finalizar el tercer gráfico con la técnica **Voronoi** será un mapa de España que muestre los centros más cercanos de innovación digital (EDIH) que ayudan a las empresas a adoptar la Inteligencia Artificial.

1 Cargar los archivos de datos

Se cargan los datos para realizar el gráfico. Por un lado tenemos los valores de los países, años, etc. en el fichero desi_ai-data.csv y por otro lado los datos de los indicadores en dds_ai-data.csv.

```

# Librerías
library(dplyr)

## 
## Adjuntando el paquete: 'dplyr'

## The following objects are masked from 'package:stats':
## 
##     filter, lag

## The following objects are masked from 'package:base':
## 
##     intersect, setdiff, setequal, union

library(readr)
library(ggplot2)

# Carga de Datos
# Indicadores
dd_df <- read_csv("dd_ai-data.csv")

## Rows: 67 Columns: 9

## -- Column specification -----
## Delimiter: ","
## chr (4): country, indicator, breakdown, unit
## dbl (2): period, value
## lgl (3): flags, reference_period, remarks
##
## 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.

# Valores medidos y líneas base
desi_df <- read_csv("desi_ai-data.csv")

## Rows: 560 Columns: 9
## -- Column specification -----
## Delimiter: ","
## chr (6): period, country, indicator, breakdown, unit, flags
## dbl (2): value, reference_period
## lgl (1): remarks
##
## 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. Extracción de indicadores de los datos para la elaboración.

Se muestran las variables y sus tipos.

```

# Objetivo Ideal 2025 del archivo de proyecciones -39.107152%
obj_ideal_2025 <- dd_df %>%
  filter(period == 2025,
         breakdown == "dd_ideal_2025",
         country == "EU") %>%
  pull(value)

print(obj_ideal_2025)

## [1] 39.10715

# Objetivo Final 2030 del archivo de proyecciones - 75%
obj_final_2030 <- dd_df %>%
  filter(period == 2030,
         breakdown == "dd_ideal_2025",
         country == "EU") %>%
  pull(value)

print(obj_final_2030)

## [1] 75

# Línea Base (Valor de la UE 2025) del archivo de países 16.28215
obj_base_2025 <- desi_df %>%
  filter(period == "desi_2025",
         country == "EU",
         breakdown == "ent_all_xfin") %>%
  pull(value)

print(obj_base_2025)

## [1] 13.48

```

3. Extracción de los datos para España y la UE.

```

# Datos Actuales para EU y todos los países.

paises_eu <- desi_df %>%
  filter(
    period == "desi_2025" &
      breakdown == "ent_all_xfin"
  ) %>%
  select(country, value) # Ya no creamos la columna 'label_pais'

print(paises_eu)

## # A tibble: 28 x 2
##       country   value
## * <chr>     <dbl>
## 1 ES          13.48
## 2 AT          1.000
## 3 DE          1.000
## 4 PT          1.000
## 5 NL          1.000
## 6 FR          1.000
## 7 IE          1.000
## 8 RO          1.000
## 9 PL          1.000
## 10 HU          1.000
## 11 MT          1.000
## 12 SI          1.000
## 13 CY          1.000
## 14 BG          1.000
## 15 LV          1.000
## 16 GR          1.000
## 17 ES          1.000
## 18 ES          1.000
## 19 ES          1.000
## 20 ES          1.000
## 21 ES          1.000
## 22 ES          1.000
## 23 ES          1.000
## 24 ES          1.000
## 25 ES          1.000
## 26 ES          1.000
## 27 ES          1.000
## 28 ES          1.000

```

```

##      <chr>    <dbl>
## 1 EU        13.5
## 2 BE        24.7
## 3 BG        6.47
## 4 CZ        11.3
## 5 DK        27.6
## 6 DE        19.8
## 7 EE        13.9
## 8 IE        14.9
## 9 EL        9.81
## 10 ES       11.3
## # i 18 more rows

# write.csv(paises_eu, "paises_eu.csv", row.names = TRUE)

# Datos para la presentación 2
datos_filtrados <- desi_df %>%
  filter(
    period == "desi_2025" &
    breakdown == "ent_all_xfin" &
    country %in% c("EU", "ES") # "Todas las empresas (con 10 o más empleados), excluyendo el sector fi
  ) %>%
  select(country, value) %>%
  mutate(
    label_pais = ifelse(country == "ES", "España", "Unión Europea (EU)")
  )

print(datos_filtrados)

## # A tibble: 2 x 3
##   country value label_pais
##   <chr>    <dbl> <chr>
## 1 EU        13.5 Unión Europea (EU)
## 2 ES        11.3 España

datos_España <- desi_df %>%
  filter(
    period == "desi_2025" &
    breakdown == "ent_all_xfin" &
    country %in% c("ES") # "Todas las empresas (con 10 o más empleados), excluyendo el sector financie
  ) %>%
  select(country, value)%>%
  mutate(
    label_pais = ifelse(country == "ES", "España")
  )
print(datos_España)

## # A tibble: 1 x 3
##   country value label_pais
##   <chr>    <dbl> <chr>
## 1 ES        11.3 España

```

4. Creación del ranking del porcentaje de uso de la IA.

```
# Ranking de países por porcentaje de uso de IA en 2025

ranking_ia <- paises_eu %>%
  arrange(desc(value)) %>%
  mutate(ranking = row_number()) %>%
  select(ranking, country, value)
```

5. Creación de tabla de datos

```
# Tomamos los datos filtrados de EU y ES
tabla_final_grafico <- datos_España %>%

# Añadimos los 3 objetivos como nuevas columnas
mutate(
  linea_base_2025      = obj_base_2025,
  objetivo_ideal_2025   = obj_ideal_2025,
  objetivo_final_2030   = obj_final_2030
)

print(tabla_final_grafico)

## # A tibble: 1 x 6
##   country value label_pais linea_base_2025 objetivo_ideal_2025
##   <chr>    <dbl> <chr>           <dbl>            <dbl>
## 1 ES        11.3 España          13.5             39.1
## # i 1 more variable: objetivo_final_2030 <dbl>
```

Creación Bullet Graph.

```
# Bullet Graph.

bullet_graph_ia <- ggplot(datos_España, aes(y = label_pais)) +

  # Capas de rangos
  geom_col(aes(x = obj_final_2030), fill = "#E6E6E6", width = 0.4) +
  geom_col(aes(x = obj_ideal_2025), fill = "#BDBDBD", width = 0.4) +
  geom_col(aes(x = obj_base_2025), fill = "#8C8C8C", width = 0.4) +
  
  # Barra negra (valor actual)
  geom_col(aes(x = value), fill = "black", width = 0.2) +
  
  # Línea del objetivo ideal
  geom_segment(
    # --- AÑADE ESTA LÍNEA PARA FILTRAR ---
    data = . %>% filter(country == "ES"),
```

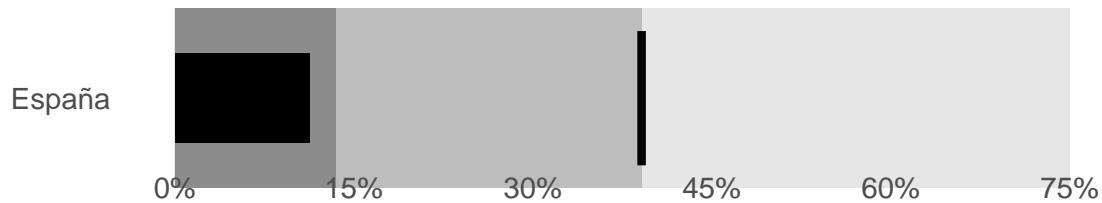
```

aes(
  x = obj_ideal_2025, xend = obj_ideal_2025,
  y = as.numeric(factor(label_pais)) - 0.15,
  yend = as.numeric(factor(label_pais)) + 0.15
), color = "black", linewidth = 1.5
) +
# Escala X
scale_x_continuous(
  limits = c(0, 80),
  breaks = seq(0, 75, by = 15),
  labels = c("0%", "15%", "30%", "45%", "60%", "75%")
) +
labs(
  title = "Transformación Digital en las Empresas Españolas 2025",
  subtitle = "Progreso del porcentaje de Adopción de Inteligencia Artificial frente
  a los Objetivos de la Unión Europea",
  x = "",
  y = "",
  caption = "Fuente: Comisión Europea, Índice de Economía y Sociedad Digital (DESI)"
) +
theme_minimal(base_size = 14) +
theme(
  plot.title = element_text(hjust = 0.5, face = "bold"),
  plot.subtitle = element_text(hjust = 0.5),
  plot.caption = element_text(hjust = 0.5),
  panel.grid.major.y = element_blank(),
  panel.grid.major.x = element_blank(),
  panel.grid.minor.x = element_blank(),
  axis.text.x = element_text(vjust = 33),
  legend.position = "none"
)
print(bullet_graph_ia)

```

Transformación Digital en las Empresas Españolas 2020

Progreso del porcentaje de Adopción de Inteligencia Artificial frente a los Objetivos de la Unión Europea



Fuente: Comisión Europea, Índice de Economía y Sociedad Digital (DESI)

```
# Guardar el gráfico
ggsave(
  "bullet_graph_ia.png",
  plot = bullet_graph_ia,
  width = 10,
  height = 5.625,
  units = "in",
  dpi = 300
)
```

6. Creación de Voronoi

Los Centros Europeos de Innovación Digital (EDIH) son ventanillas únicas que ayudan a las empresas y organizaciones del sector público a responder a los desafíos digitales:

- Proporcionan acceso a conocimientos técnicos y pruebas.
- Proporcionan Servicios de Innovación, como asesoramiento para una transformación digital exitosa.
- ayuda a las empresas en el uso de tecnologías digitales para la sostenibilidad y la circularidad.

En este apartado se extraen los datos de los EDIH, para crear un mapa de Voronoi de los que se encuentran en España.

```

library(readxl)
library(dplyr)
library(tidygeocoder)
library(ggplot2)
library(ggforce)
library(maps)
library(sf)

## Linking to GEOS 3.13.1, GDAL 3.11.4, PROJ 9.7.0; sf_use_s2() is TRUE

# Carga el archivo de Excel
edih_raw <- read_excel("edihs_ia_spain.xlsx")

# 1. Filtra por España.
# 2. Selecciona las columnas necesarias para geocodificar y colorear.
edih_spain_seleccionado <- edih_raw %>%
  filter(Country == "Spain") %>%
  select(`EDIH Name`, Provincia)

# Geocodifica la columna 'Provincia' para obtener 'lat' y 'lon'
edih_hubs_coords <- edih_spain_seleccionado %>%
  geocode(address = Provincia, method = "osm")

## Passing 20 addresses to the Nominatim single address geocoder

## Query completed in: 20.3 seconds

print(edih_hubs_coords)

## # A tibble: 25 x 4
##   `EDIH Name`     Provincia      lat    long
##   <chr>          <chr>        <dbl>   <dbl>
## 1 CIDAI          Barcelona     41.4    2.18 
## 2 AGORA DIH      Murcia       38.0   -1.13 
## 3 AgrotechDIH   Sevilla      37.4   -6.00 
## 4 AIR-Andalusia Granada      37.2   -3.60 
## 5 AIR4S          Madrid       40.4   -3.70 
## 6 Aragon EDIH    Zaragoza     41.7   -0.881
## 7 AsDIH          Llanera(Asturias) 43.5   -5.85 
## 8 DATAlife       Santiago de Compostela 42.9   -8.55
## 9 DIGIS3         Salamanca    41.0   -5.66 
## 10 Digital Impulse Hub Terrassa, Barcelona 41.6   2.01
## # i 15 more rows

```

Generación del gráfico Voronoi.

```

# Carga el contorno de España para el mapa (incluye Canarias)
contorno_espana <- map_data("world", region = c("Spain", "Canary Islands"))

# Convierte el contorno a polígono sf
espana_sf <- contorno_espana %>%

```

```

st_as_sf(coords = c("long", "lat"), crs = 4326) %>%
group_by(group) %>%
summarise(geometry = st_combine(geometry), .groups = "drop") %>%
st_cast("POLYGON") %>%
st_union()

# Convierte los puntos EDIH a sf CON LA COLUMNA EDIH Name
puntos_sf <- edih_hubs_coords %>%
  st_as_sf(coords = c("long", "lat"), crs = 4326)

# Genera el diagrama de Voronoi
voronoi.todos <- st_voronoi(st_union(puntos_sf)) %>%
  st_collection_extract("POLYGON") %>%
  st_as_sf() # Convierte a sf

# Recorta al polígono de España
voronoi_recortado <- st_intersection(voronoi.todos, espana_sf) %>%
  st_as_sf()

# Asigna cada polígono Voronoi al punto EDIH más cercano
# Usa st_nearest_feature para asegurar que cada polígono va al punto más cercano
nearest_idx <- st_nearest_feature(voronoi_recortado, puntos_sf)
voronoi_con_datos <- voronoi_recortado %>%
  mutate(`EDIH Name` = puntos_sf$`EDIH Name`[nearest_idx])

# Crea el Gráfico Voronoi PRINCIPAL (Sin Canarias)
grafico_principal <- ggplot() +

  # Dibuja las teselas de Voronoi recortadas con color por EDIH Name
  geom_sf(data = voronoi_con_datos,
    aes(fill = `EDIH Name`),
    colour = "white",
    linewidth = 0.3,
    show.legend = FALSE) +

  # Dibuja el contorno de España
  geom_polygon(
    data = contorno_espana %>% filter(region == "Spain"),
    aes(x = long, y = lat, group = group),
    fill = NA,
    color = "#444444",
    linewidth = 0.5
  ) +

  # Dibuja los puntos de los EDIH
  geom_sf(data = puntos_sf, color = "black", size = 3) +
  coord_sf(xlim = c(-10, 5), ylim = c(35, 44), expand = FALSE) +
  theme_void(base_size = 12) +
  labs(
    title = "Centros IA de Innovación Digital (EDIH) en España",
    subtitle = "Cada celda muestra el Centro de innovación más cercano de Inteligencia Artificial",

```

```

    caption = "Fuente: Catálogo EDIH, European Digital Innovation Hubs Network"
) +
theme(
  plot.title = element_text(hjust = 0.5, face = "bold", size = 14),
  plot.subtitle = element_text(hjust = 0.5, size = 12),
  plot.caption = element_text(size = 11, hjust = 1)
)

# Crea el gráfico INSET para Canarias
grafico_canarias <- ggplot() +

# Dibuja las teselas de Voronoi recortadas
geom_sf(data = voronoi_con_datos,
         aes(fill = `EDIH Name`),
         colour = "white",
         linewidth = 0.2,
         show.legend = FALSE) +

# Dibuja el contorno de Canarias
geom_polygon(
  data = contorno_espana %>% filter(region == "Canary Islands"),
  aes(x = long, y = lat, group = group),
  fill = NA,
  color = "#444444",
  linewidth = 0.5
) +

# Dibuja los puntos de los EDIH
geom_sf(data = puntos_sf, color = "black", size = 2) +

coord_sf(xlim = c(-18, -13), ylim = c(27, 29.5), expand = FALSE) +
theme_void(base_size = 10) +
theme(
  plot.background = element_rect(fill = "white", color = "black", linewidth = 1)
)

# Inclusión de Canarias
grafico_voronoi_edih <- grafico_principal +
annotation_custom(
  ggplotGrob(grafico_canarias),
  xmin = -9.8, xmax = -7,
  ymin = 35.2, ymax = 36.8
)

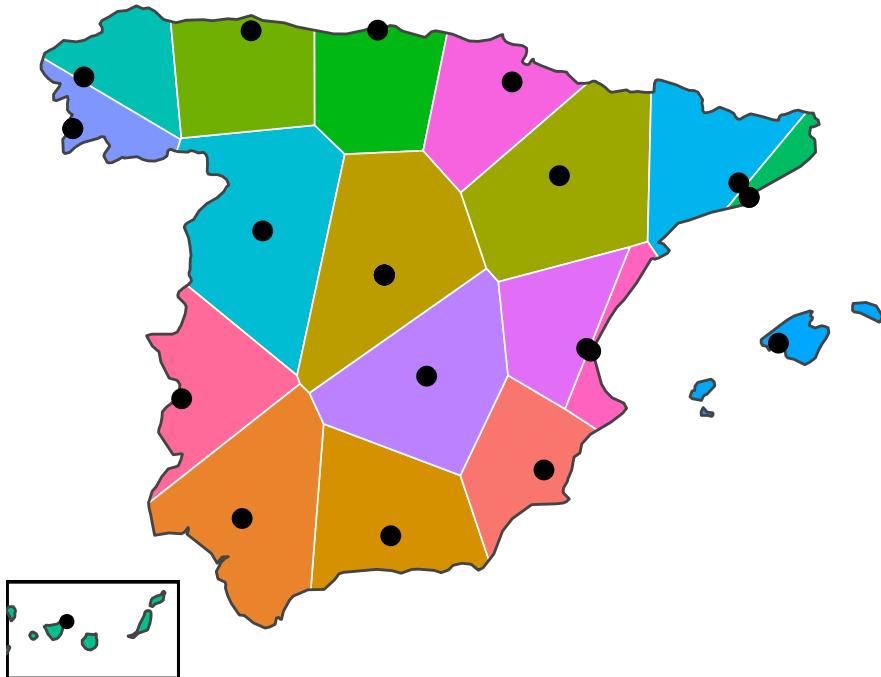
grafico_voronoi_edih <- grafico_voronoi_edih +
theme(
  plot.background = element_rect(fill = "white", color = NA),
  panel.background = element_rect(fill = "white", color = NA)
)

print(grafico_voronoi_edih)

```

Centros IA de Innovación Digital (EDIH) en España

Cada celda muestra el Centro de innovación más cercano de Inteligencia Artificial



Fuente: Catálogo EDIH, European Digital Innovation Hubs Network

```
ggsave(  
  "voronoi_edih_spain.png",  
  plot = grafico_voronoi_edih,  
  width = 14,  
  height = 10,  
  units = "in",  
  dpi = 300,  
  bg = "white"  
)
```