

Práctica Dirigida 8

Mg. Sc. J. Eduardo Gamboa U.

En el año 2021 se llevaron a cabo las Elecciones Generales, en la que se eligieron la fórmula presidencial y los congresistas para el periodo 2021-2026. La ONPE disponibiliza los datos de las actas (o mesas) en la [Plataforma Nacional de Datos Abiertos](#). También la data y el diccionario se pueden descargar del [repositorio GitHub](#)

Pregunta 1

Cargar los paquetes necesarios para el desarrollo del caso:

```
library(readr)
library(dplyr)
library(modeest)
library(ggQC)
library(sjstats)
```

Pregunta 2

Leer el archivo de datos. Utilizar **View** para visualizar el data frame leído.

```
datos = read_csv2('PD8 - Resultados_1ra_vuelta_Version_PCM.csv')
```

Pregunta 3

Se realiza el preprocesamiento del conjunto de datos leído mediante los siguientes códigos. ¿Qué acciones están ejecutando?

```
datos |>
  filter(DESCRIP_ESTADO_ACTA %in% c("CONTABILIZADA", "COMPUTADA RESUELTA")) |>
  select(-TIPO_ELECCION, -TIPO_OBSERVACION) -> datos
datos[is.na(datos)] <- 0
```

Acciones que se ejecutan:

-
-
-

Pregunta 4

Interpretar las medidas de tendencia central de la cantidad de ciudadanos que votaron por mesa.

```
datos |> pull(N_CVAS) |> mean()
```

```
[1] 205.9652
```

Interpretación: ...

```
datos |> pull(N_CVAS) |> median()
```

```
[1] 218
```

Interpretación: ...

```
datos |> pull(N_CVAS) |> mfv()
```

```
[1] 229
```

Interpretación: ...

```
datos |>
  summarize(Media = mean(N_CVAS),
            Mediana = median(N_CVAS),
            Moda = mfv(N_CVAS))
```

```
# A tibble: 1 x 3
  Media Mediana  Moda
  <dbl>   <dbl> <dbl>
1  206.    218   229
```

Pregunta 5

Interpretar las medidas de tendencia central de la cantidad de ciudadanos que votaron por mesa en el distrito de Ate.

Pregunta 6

Indicar las acciones que ejecuta el siguiente código, luego interpretar el mayor y menor valor para cada medida de tendencia central.

```
datos |>
  filter(DEPARTAMENTO %in% c("AFRICA", "AMERICA", "ASIA", "EUROPA", "OCEANIA")) |>
  group_by(DEPARTAMENTO) |>
  reframe(Media = mean(N_CVAS),
          Mediana = median(N_CVAS),
          Moda = mfv(N_CVAS))
```

```
# A tibble: 12 x 4
  DEPARTAMENTO Media Mediana  Moda
  <chr>         <dbl>   <dbl> <dbl>
1 AFRICA        7.67      8      6
2 AFRICA        7.67      8      8
3 AFRICA        7.67      8      9
4 AMERICA       66.3     60     40
5 AMERICA       66.3     60     44
6 ASIA          49.7     47     42
7 EUROPA        89.3     90     91
8 OCEANIA       85.8     87     84
9 OCEANIA       85.8     87     85
10 OCEANIA      85.8     87     87
11 OCEANIA      85.8     87     88
12 OCEANIA      85.8     87     94
```

Acciones ejecutadas: ...

Interpretaciones: ...

Pregunta 7

Interpretar el primer cuartil de la cantidad de votos obtenidos por mesa por los partidos políticos Perú Libre, Partido Morado y Fuerza Popular en la provincia de Lima.

```
datos |>
  filter(PROVINCIA=="LIMA") |>
  summarize(Q1_PL = quantile(VOTOS_P16, probs = 0.25),
            Q1_PM = quantile(VOTOS_P3, probs = 0.25),
            Q1_FP = quantile(VOTOS_P11, probs = 0.25))
```

```
# A tibble: 1 x 3
  Q1_PL Q1_PM Q1_FP
  <dbl> <dbl> <dbl>
1      6      3     18
```

Interpretación: ...

Pregunta 8

¿Cuál es la menor cantidad de votos nulos que debe tener una mesa para encontrarse dentro del 15% de las mesas con mayor cantidad de votos nulos?

```
datos |>
  pull(VOTOS_VN) |>
  quantile(probs = ...) # completar con el percentil correcto
```

Pregunta 9

¿Qué provincia de Junín tuvo el valor más bajo en el decil 6 de votos a favor de Juntos por el Perú?

Pregunta 10

Explicar las acciones que se están ejecutando en el siguiente código, luego interpretar el valor obtenido:

```
datos |> filter(PROVINCIA == "HUAMANGA") |> summarize(Rango_JPP = QCrange(VOTOS_P9))

# A tibble: 1 x 1
  Rango_JPP
  <dbl>
1         51
```

Acciones que se ejecutan:

-
-

Interpretación:

Pregunta 11

Explicar las acciones que se están ejecutando en el siguiente código, luego interpretar el valor obtenido:

```
datos |> filter(PROVINCIA == "ILO") |> summarize(RIC_APP = QCrange(VOTOS_P18))

# A tibble: 1 x 1
  RIC_APP
  <dbl>
1       16
```

Acciones que se ejecutan:

-
-

Interpretación:

Pregunta 12

Calcular la varianza de la cantidad de votos por mesa obtenidos por el partido Unión por el Perú

```
datos |>
  pull(VOTOS_P12) |>
  var()
```

```
[1] 2.681974
```

La varianza es igual a

Pregunta 13

Explicar el valor de la desviación estándar de la cantidad de electores hábiles por mesa

```
datos |>
  summarize(media = mean(N_ELEC_HABIL),
            desv  = sd(N_ELEC_HABIL))
```

```
# A tibble: 1 x 2
  media desv
  <dbl> <dbl>
1  292.  26.6
```

La desviación estándar es igual a ... y esto significa que ...

Pregunta 13

Explicar el valor de la desviación estándar de la cantidad de electores hábiles por mesa en cada uno de los siguientes distrits:

- Miraflores
- Comas
- Santa María del Mar

Pregunta 14

¿En qué provincia de Lima, los votos de Fuerza Popular fueron más dispersos?

```
datos |>
  filter(DEPARTAMENTO == "LIMA") |>
  group_by(PROVINCIA) |>
  summarize(CV = cv(VOTOS_P11)*100)
```

```
# A tibble: 10 x 2
  PROVINCIA      CV
  <chr>         <dbl>
1 "BARRANCA"    33.2
2 "CAJATAMBO"  49.1
3 "CANTA"      36.3
4 "CA\xbf\xbfETE" 34.4
5 "HUARAL"     35.3
6 "HUAROCHIRI" 40.0
7 "HUAURA"    41.1
8 "LIMA"       40.6
9 "OYON"       55.0
10 "YAUYOS"    43.2
```

Pregunta 15

Revisar el [caso ENDES](#) disponible en el repositorio GitHub.

Pregunta 16

Revisar el [caso Postulantes](#) disponible en el repositorio GitHub.