Métodos Multivariados y Datos Categóricos

Proyecto Final

Dante Ruíz
Ana Luisa Masetto
Arantza Ivonne Pineda

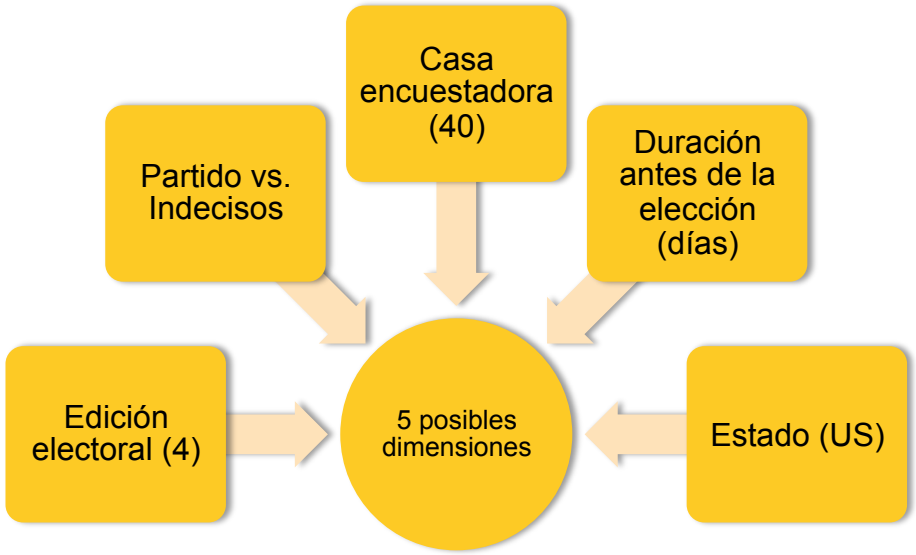
DESARROLLO

- ❑ **Problema**
- ❑ **Tratamiento de los Datos**
- ❑ **Modelos propuestos**
- ❑ **Conclusiones**

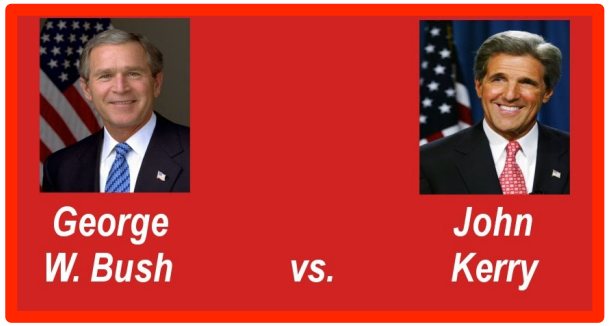
PROBLEMA

- **Encuestas** públicas políticas → **RETO:** **votantes indecisos**
- Altos números de votantes indecisos influyen en **predicciones**
- ¿Cómo tratar este problema de una manera estadísticamente adecuada?

SOLUCIÓN: **Tablas de Contingencia**



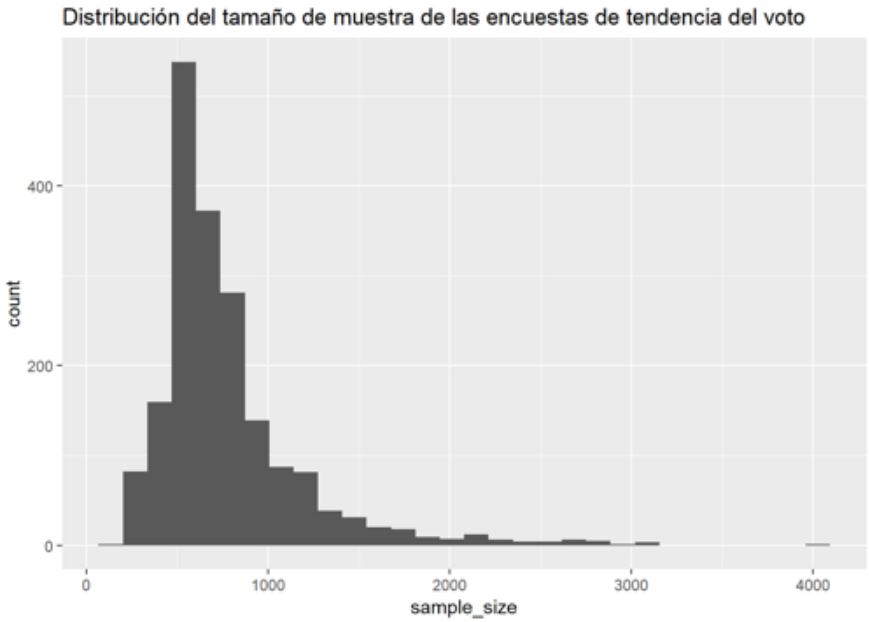
PROBLEMA



OBJETIVO 1:
Evaluar si la intención de voto en el polling por estado es independientemente inducida por el tiempo de duración para el día de la elección

OBJETIVO 2:
Evaluar si la casa encuestadora es independiente de los resultados de la intención de voto

MÉTODO DE AGREGACIÓN



1. **Tamaños de muestra** (cuantiles)

0%	25%	50%	75%	100%
200	545	649	869	4092

- Necesidad de **reescalar** las encuestas para ser **comparables**

```
base_reescalada<- polling %>%
  mutate(
    Undecided = ifelse(is.na(Undecided),0,Undecided),
    muestra_efectiva = Dem_poll + Rep_poll + Undecided, #Sacar la muestra efectiva (#Democrata
s + #Republicanos + #Undecided )
    #NA_voters = sample_size - muestra_efectiva, #Las personas que no decidieron compartir sus
votos
    #Undecided = Undecided + NA_voters, #Según el artículo los que no compartieron su voto se
agregan a los indecisos
    proporcion_Dem = Dem_poll/muestra_efectiva, #Se saca la proporción por categoría de votant
es: #Dem_poll/muestra_efectiva
    proporcion_Rep = Rep_poll/muestra_efectiva, #Se saca la proporción por categoría de votant
es: #Rep_poll/muestra_efectiva
    proporcion_Und = Undecided/muestra_efectiva, #Se saca la proporción por categoría de votan
tes: #(Undecided)/muestra_efectiva
    reescalar_Dem = round(proporcion_Dem*100,0), #Reescalar a 100 votantes (personas por cada
100 votantes)
    reescalar_Rep = round(proporcion_Rep*100,0), #Reescalar a 100 votantes
    reescalar_Und = round(proporcion_Und*100,0)) #Reescalar a 100 votantes
head(base_reescalada)
```

```
## # A tibble: 6 x 26
##   Dem_poll Rep_poll Undecided sample_size mean_days_to_el~ start_days_to_e~
##      <dbl>    <dbl>     <dbl>       <dbl>            <dbl>            <dbl>
## 1       85        7         2         310              4                7
## 2       92        6         3         312              5.5              9
## 3       88        4         1         332             11               14
## 4       81       13         6         310             13               16
## 5       88        5         2         504             18               21
## 6       29       56         5         472              4                7
## # ... with 20 more variables: end_days_to_election <dbl>, state <chr>,
## #   year <dbl>, state_year <chr>, pollster <chr>, state_year_id <dbl>,
## #   pollster2 <chr>, year_id <dbl>, result_margin6 <chr>,
## #   rmargin_year <chr>, rmargin_year_id <dbl>, pollster_grp <chr>,
## #   pollster_id <dbl>, muestra_efectiva <dbl>, proporcion_Dem <dbl>,
## #   proporcion_Rep <dbl>, proporcion_Und <dbl>, reescalar_Dem <dbl>,
## #   reescalar_Rep <dbl>, reescalar_Und <dbl>
```

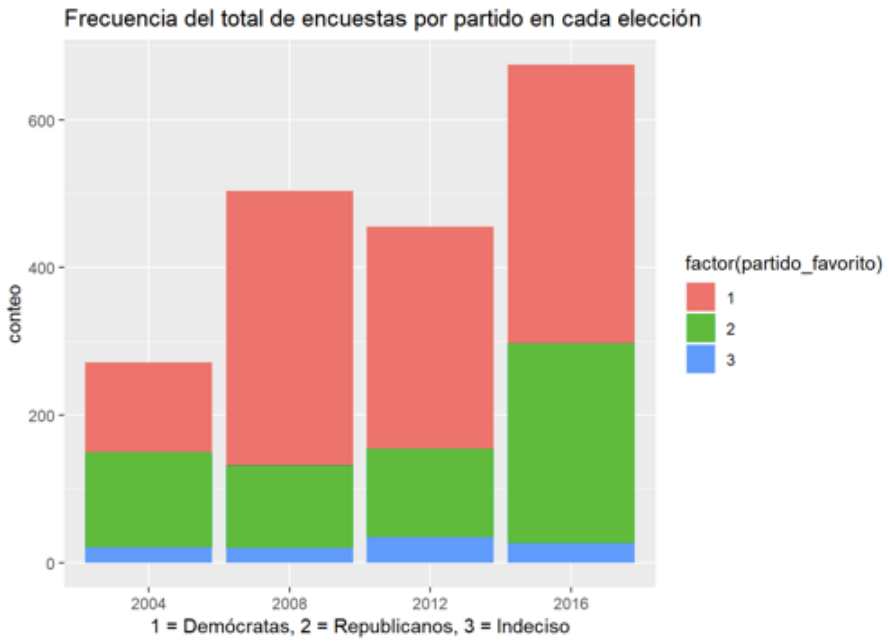
MÉTODO DE AGREGACIÓN

2. Construcción de variable: INTENCIÓN DE VOTO

- toma el valor del partido con mayor proporción de votos al momento de la encuesta
  1. **DEMÓCRATA**
  2. **REPUBLICANO**
  3. **INDECISO**

- Resumen:

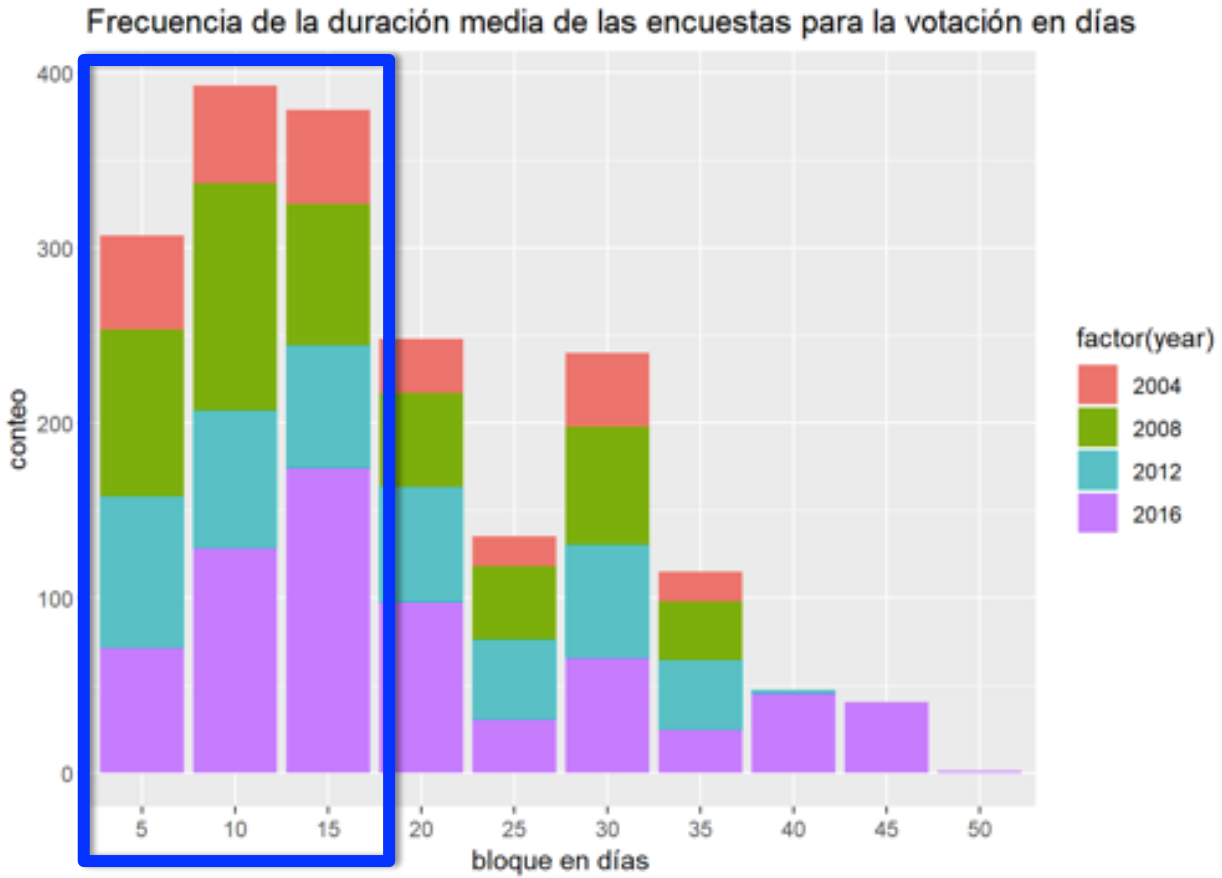
1	2	3
1173	631	101



MÉTODO DE AGREGACIÓN

3. Construcción de variable: TIEMPO MEDIO ANTES DE ELECCIÓN

- Se define en bloques de 5 días



MODELOS

OBJETIVO 1:
Evaluar si la intención de voto en el polling por estado es independientemente inducida por el tiempo de duración para el día de la elección

- 4 variables de interés:
  - Intención de voto
  - Estado
  - Tiempo antes de la elección
  - Edición de la elección

- **Tabla de contingencia** (probabilidades muestrales):
  - Por grupos de 5 hasta 50 días antes de la elección
  - Filas: partido político; Columnas: Estados

- Interpretación:

Para un dato `dato` $X_{i,v}$ la `probabilidad` de que adopte la"

$$x_{i,v} = \left(x_{i,1} = \text{republicano } (2), x_{2,1} = \text{alabama } (1), x_{i,3} = 5 \text{ días antes de la elección } (1)\right), x_{i,4} = \text{edición } (4))$$

es

$$\mathbb{P}(X_{i,v} = x_{i,v}) = \theta_{e(1),e(2),e(3)} = 0.0462,$$

con

$$\begin{aligned} e(1) &= \text{republicano } (2), \\ e(2) &= \text{alabama } (1), \\ e(3) &= 5 \text{ días } (1). \\ e(4) &= 2016 \ (4). \end{aligned}$$

MODELOS

OBJETIVO 1:
Evaluar si la intención de voto en el polling por estado es independientemente inducida por el tiempo de duración para el día de la elección

- **Procedimiento de Estimación con Grafos:** encontrar el modelo log-lineal con la mayor probabilidad posterior (máx. verosimilitud)

- 6 Mejores Modelos:

```
##                                                                    formula
## 1 [partido_favorito][state][tiempo_antes_de_la_eleccion][year_id]
## 2  [partido_favorito,state][tiempo_antes_de_la_eleccion][year_id]
## 3  [partido_favorito,year_id][state][tiempo_antes_de_la_eleccion]
## 4  [partido_favorito,tiempo_antes_de_la_eleccion][state][year_id]
## 5  [partido_favorito][state,tiempo_antes_de_la_eleccion][year_id]
## 6  [partido_favorito][state][tiempo_antes_de_la_eleccion,year_id]
##   logPostProb
## 1    1348.988
## 2    1347.031
## 3    1346.242
## 4    1345.665
## 5    1345.369
## 6    1345.283
```

MODELOS

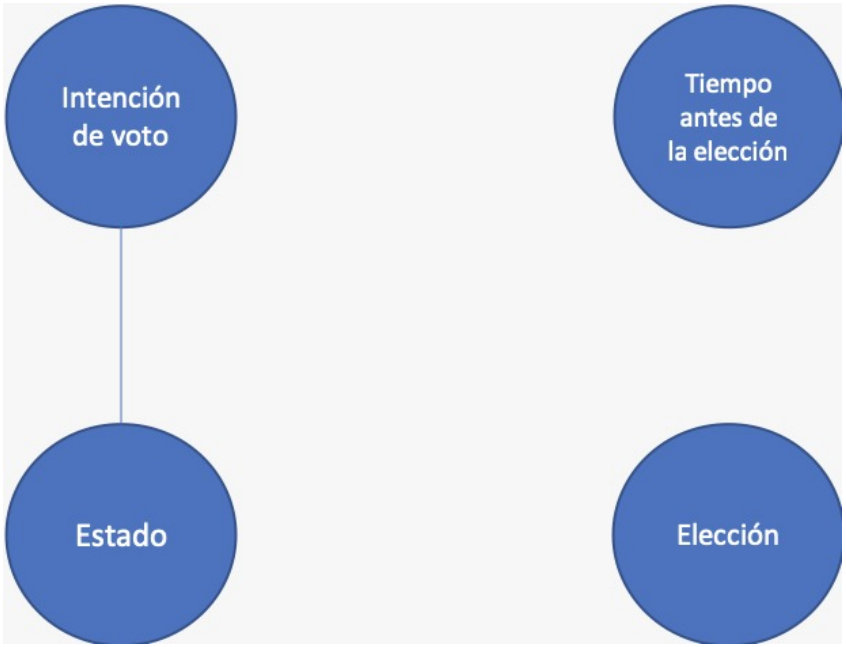
OBJETIVO 1:
Evaluar si la intención de voto en el polling por estado es independientemente inducida por el tiempo de duración para el día de la elección

- Mejor modelo con mayor log-posterior probabilidad:

Modelo Simple

  - Se revisaron las demás configuraciones para identificar modelos con grafos triangulares, pero no se encontró ninguno.

- Considerando el modelo agregado, No se rechaza la Hipótesis Nula, por lo tanto, se prueba independencia, sin embargo, de manera segregada, hay casos donde esto no se respeta.



MODELOS

OBJETIVO 2:
Evaluar si la casa encuestadora es independiente de los resultados de la intención de voto

- 4 variables de interés:
  - Partido con mayor intención de voto
  - Casa encuestadora
  - Elección (año)
  - Tiempo antes de la elección

- **Tabla de contingencia**:
  - Filas: Partido político; Columnas: casa encuestadora

- Interpretación:

Para un dato `dato` $X_{i,v}$ la `probabilidad` de que adopte la

$$x_{i,v} = \left(x_{i,1} = \text{democrata } (1), x_{2,1} = \text{survey_monkey } (32), x_{i,3} = 2016\ (4), x_{i,4} = \text{tiempo antes de la elección } (1)\right),$$

es

$$\mathbb{P}(X_{i,v} = x_{i,v}) = \theta_{e(1),e(2),e(3),e(4)} = 0.0462,$$

con

$$\begin{aligned} e(1) &= \text{democrata } (1), \\ e(2) &= \text{survey_monkey } (32), \\ e(3) &= 2016\ (4). \\ e(4) &= 5 \text{ días antes de la elección } (1). \end{aligned}$$

MODELOS

OBJETIVO 2:
Evaluar si la casa encuestadora es independiente de los resultados de la intención de voto

- **Procedimiento de Estimación con Grafos:** encontrar el modelo log-lineal con la mayor probabilidad posterior (máx. Verosimilitud)

- 6 mejores modelos:

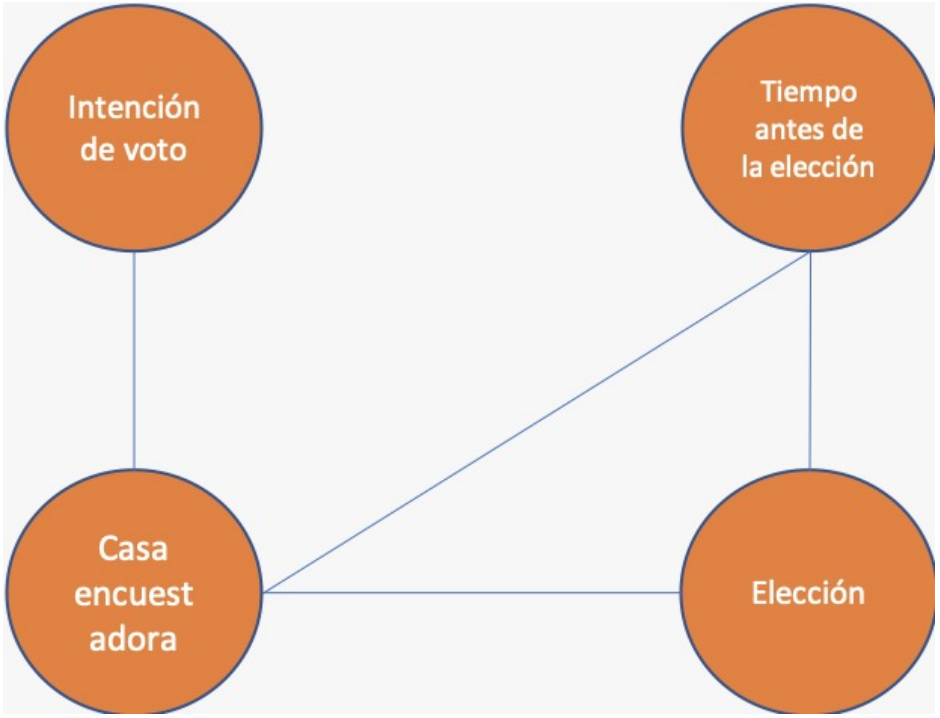
```
##                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                formula
## 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
## 2                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
## 3                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
## 4                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
## 5                                                                                                
## 6 [partido_favorito,pollster_id][partido_favorito,year_id][pollster_id,tiempo_antes_de_la_eleccion]
##   logPostProb
## 1    2287.013
## 2    2286.863
## 3    2286.835
## 4    2286.464
## 5    2284.932
## 6    2284.834
```

MODELOS

OBJETIVO 2:
Evaluar si la casa encuestadora es independiente de los resultados de la intención de voto

- Mejor Modelo:

Modelo donde el partido favorito es condicionalmente independiente de año y tiempo dado la casa encuestadora



- Existe independencia condicional a nivel agregado, sin embargo, una vez que se desgloza por casa encontramos que el resultado del agregado no se sostiene en todos los casos (6 casas de 40 que rechazan la hipótesis nula)

CONCLUSIONES

PAUL J. RICHARDS/AFP/GETTY IMAGES