

Proyecto Harmony HealthCare

Angel Pineda

2025-03-18

Proyecto Harmony HealthCare

El objetivo de este proyecto es encontrar los variables que mas contribuyen a que un paciente use la sala de emergencia. Los datos tienen 2821 pacientes y 675 variables. El proyecto se desarrolla en R.

Structura del proyecto

El proyecto tiene tres componentes:

1. Limpieza, procesamiento y visualización de datos (eliminación de variables vacías, imputación de información faltante, renombramiento de variables, visualización de relaciones entre variables)
2. Selección de características (estudio de estrategias para identificar las variables más predictivas del uso de la sala de emergencia)
3. Validación (estudio de estrategias para determinar la utilidad de los resultados, por ejemplo, entrenamiento, validación, prueba, validación cruzada, remuestreo)

```
library(readr)
HarmonyHealthcareData <- read_csv("C:/Users/h703294449/OneDrive - Hofstra University/Hofstra/Service/EMALCA
Honduras 2025/Curso de Aprendizaje Estadistico/HarmonyHealthcareData.csv")
```

```
## Rows: 2821 Columns: 139
## — Column specification —————
## Delimiter: ","
## chr  (95): EHR.Sex, Ethnicity, Language, Race, Refugee, Sex.at.Birth, Usual...
## dbl  (41): UDS.Qualifying.Encounter.Count, Active.Medications, Alcohol.Asses...
## lgl   (1): Public.Housing
## date  (2): Age, Obesity
##
## 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.
```

```
View(HarmonyHealthcareData)
head(HarmonyHealthcareData)
```

```
## # A tibble: 6 × 139
##   Age      EHR.Sex Ethnicity      Language Race  Refugee Sex.at.Birth
##   <date>   <chr>   <chr>          <chr>   <chr> <chr>   <chr>
## 1 1900-01-27 female another hispanic, lati... spanish white ignore female
## 2 1900-01-21 female another hispanic, lati... spanish amer... ignore female
## 3 1900-02-17 female another hispanic, lati... spanish unre... ignore female
## 4 1900-02-10 male   not hispanic, latino/a... english blac... ignore male
## 5 1900-02-17 female another hispanic, lati... spanish white ignore female
## 6 1900-02-08 female puerto rican          english blac... ignore female
## # i 132 more variables: Usual.Location <chr>, Usual.Provider <chr>,
## #   Veteran.Status <chr>, Zip <chr>, Most.Recent.Encounter.Type <chr>,
## #   UDS.Qualifying.Encounter.Count <dbl>, UDS.Homelessness.Status <chr>,
## #   Active.Medications <dbl>, ACE.ARB.Med.Name <chr>,
## #   Alcohol.Assessment.Code <chr>, Alcohol.Assessment.Result <dbl>,
## #   Anti.HTN.Med.Name <chr>, Patient.Appointment.No.Show.Count <dbl>,
## #   Patient.Appointment.No.Show.Rate.. <dbl>, AST.Result <chr>, ...
```

```
# Exploración de los datos
```

```
boxplot(HarmonyHealthcareData$Active.Medications~HarmonyHealthcareData$Admission, xlab = 'Admisión al departamento de emergencia', ylab = 'numero de medicaciones')
```

