

Proyecto

Introducción

La Encuesta Nacional de Victimización y Percepción sobre Seguridad Pública (ENVIPE) es un instrumento realizado y aplicado por el INEGI con múltiples objetivos. Entre ellos se encuentran medir la victimización personal y del hogar, estimar el número de víctimas y número de delitos ocurridos a lo largo del año y realizar algunas otras mediciones y estimaciones relevantes sobre victimización y percepción de la seguridad de los mexicanos. Todo esto se realiza con el fin de generar información disponible para el público general y que sea de utilidad para la implementación y mejora de políticas públicas en la materia. La ENVIPE arroja como resultado información en forma de tabulados con estimaciones y bases de datos de las respuestas obtenidas.

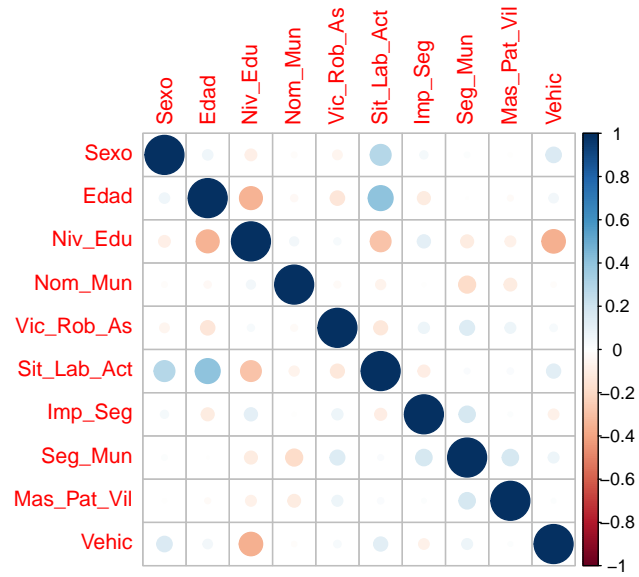
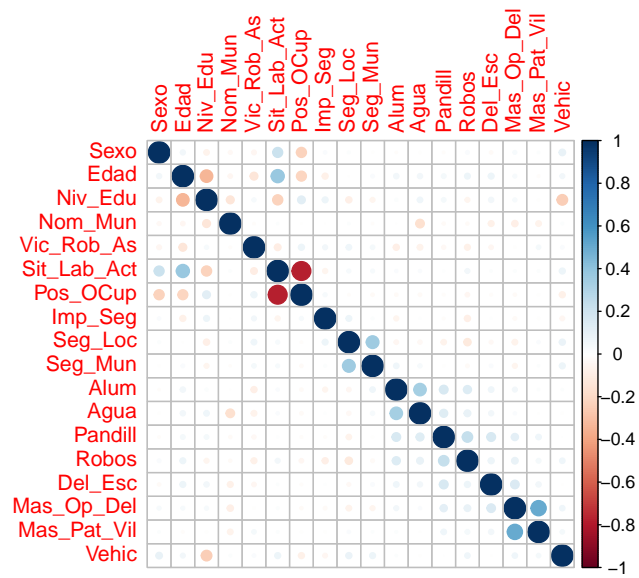
Para la realización de este proyecto, se trabajó con la base de datos de respuestas de la encuesta realizada en 2019, centrándose en estudiar y modelar el número de robos y/o asaltos ocurridos a los encuestados cuya residencia se encuentra en la CDMX. En particular, es de interés contestar a los cuestionamientos: ¿existe alguna relación entre ser víctima de robo y las condiciones y/o características de las personas (de acuerdo con los datos que recopiló la encuesta)? ¿El número de veces que se es víctima a lo largo del año puede tener relación con dichas características?

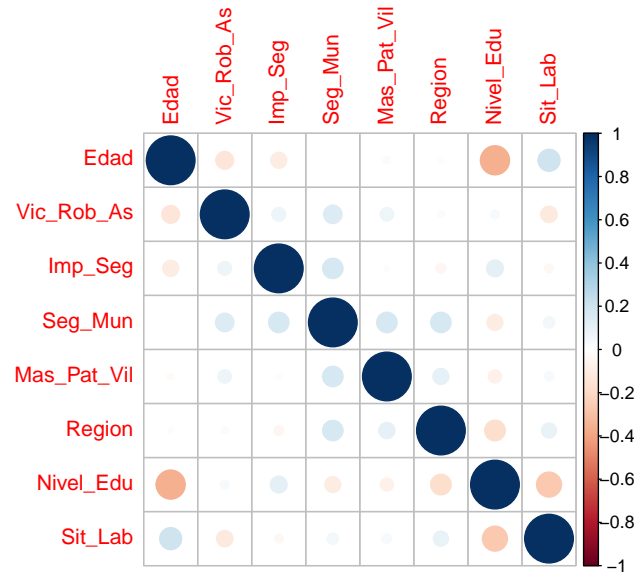
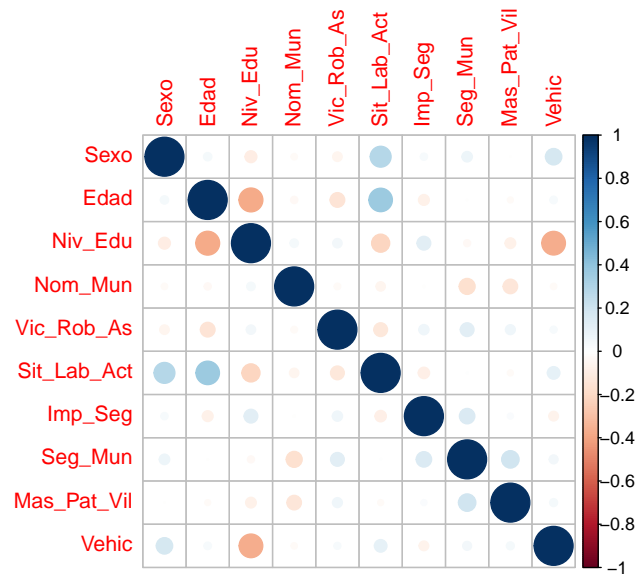
Para contestar estas preguntas, se realizó el análisis y posterior modelación tanto de la probabilidad de sufrir algún robo o asalto como del número de veces que se fue víctima a lo largo del año de alguno de estos delitos. Para ambas variables se consideraron diversas características de los encuestados.

Limpieza de la Base

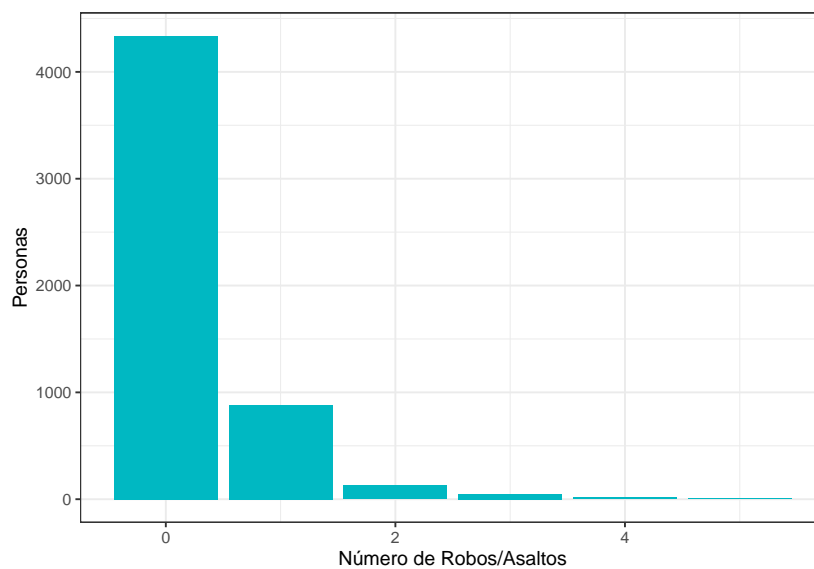
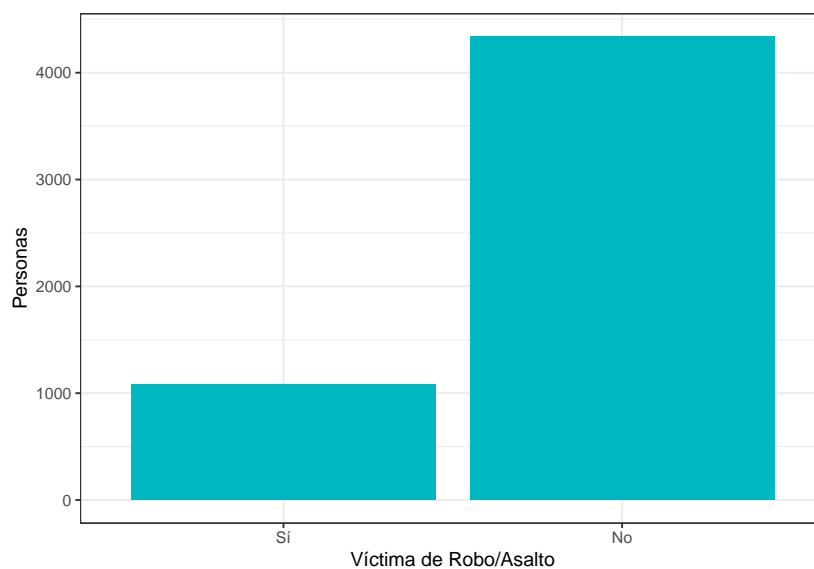
En un principio, tomamos a consideración aquellas preguntas realizadas que a primera instancia pareciesen influir en el valor de nuestras variables respuesta. Así, consideramos como opción las siguientes covariables:

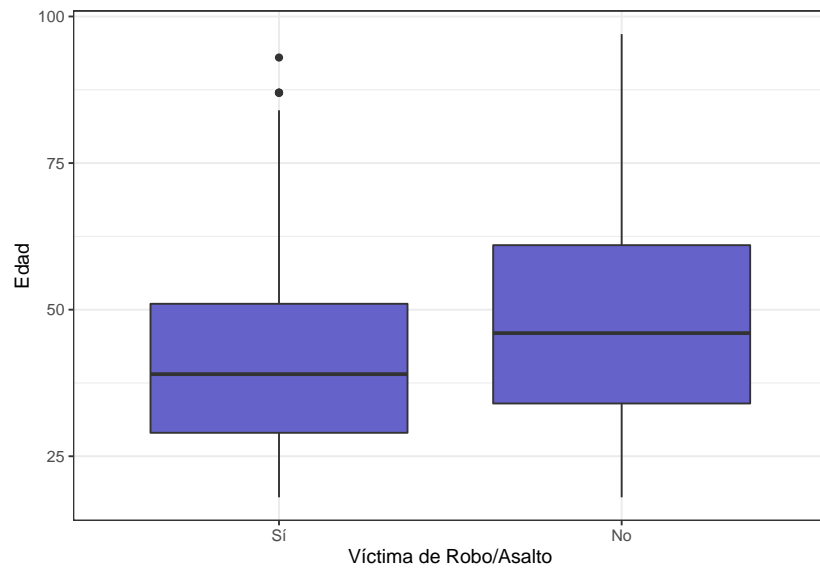
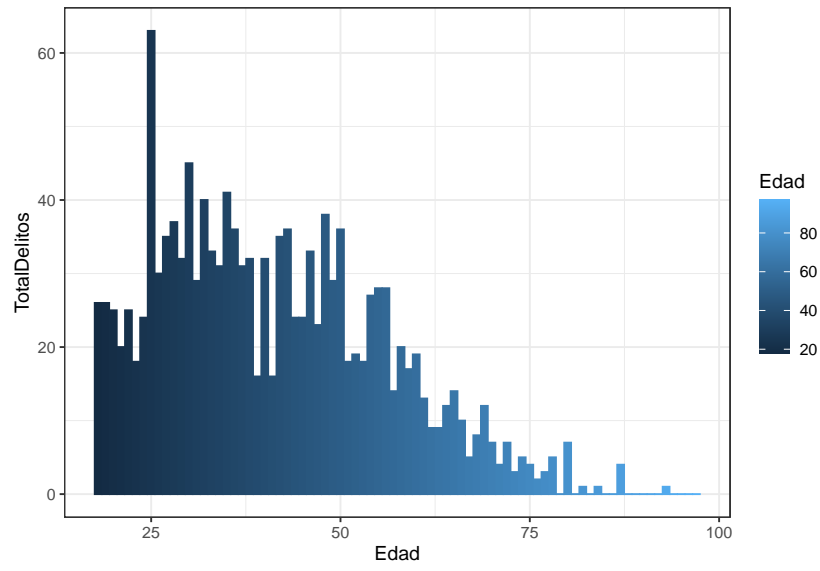
1. Sexo
2. Edad
3. Nivel educativo
4. Situación
- 5.
- 6.
- 7.
- 8.
- 9.
- 10.

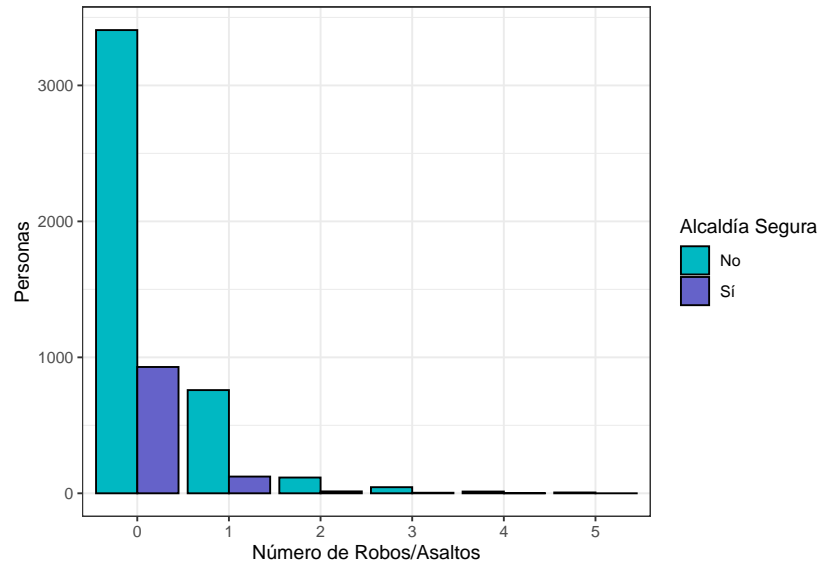
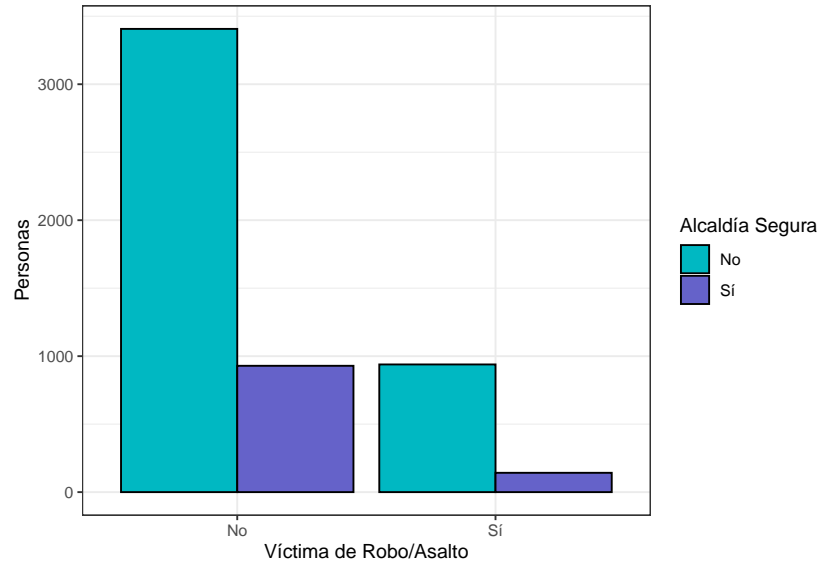


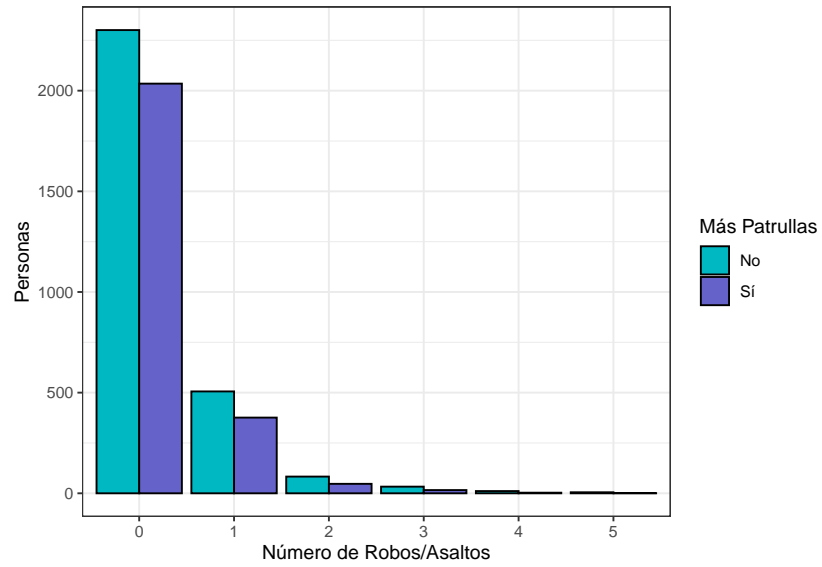
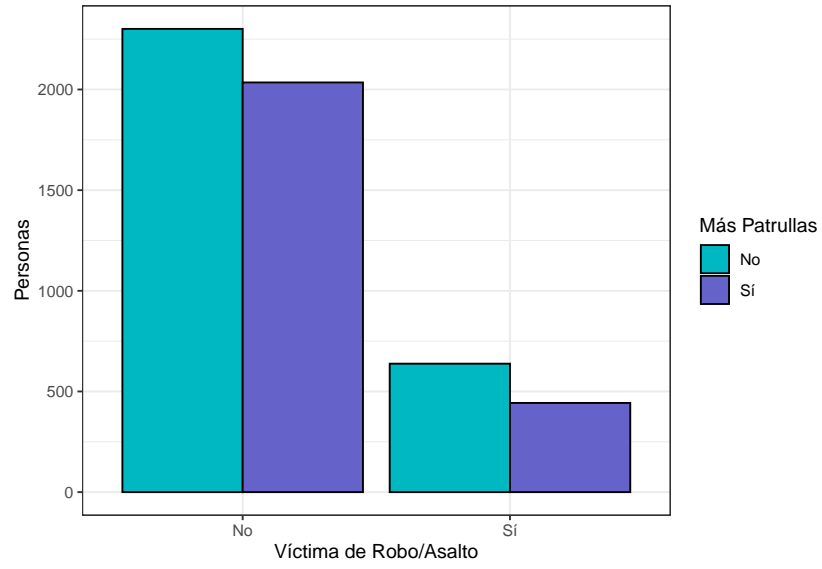


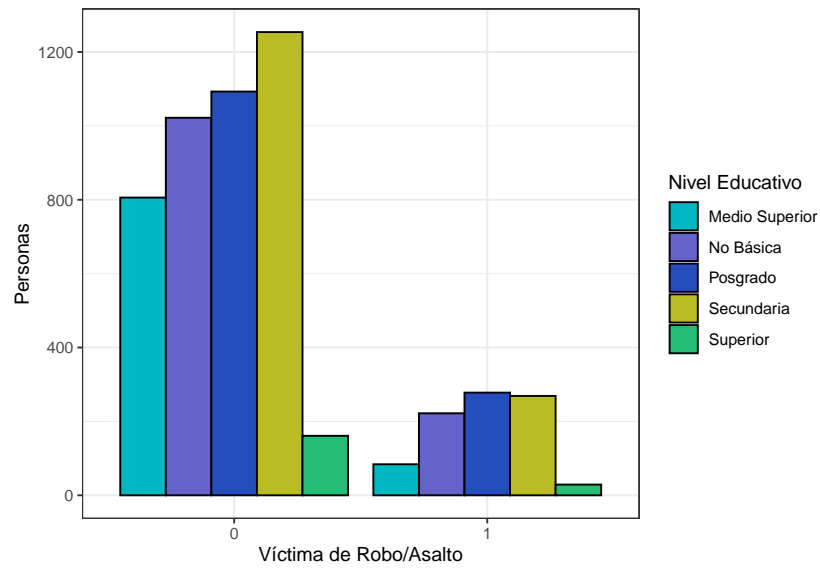
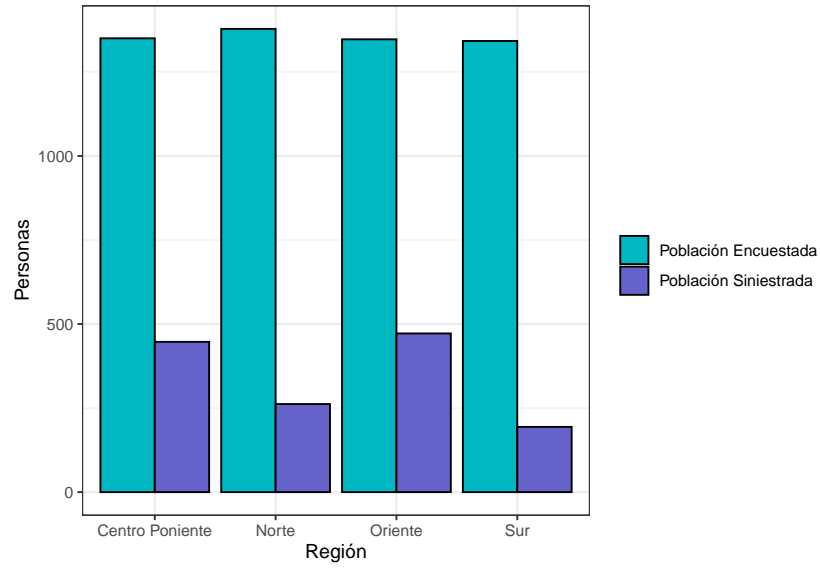
Análisis Descriptivo

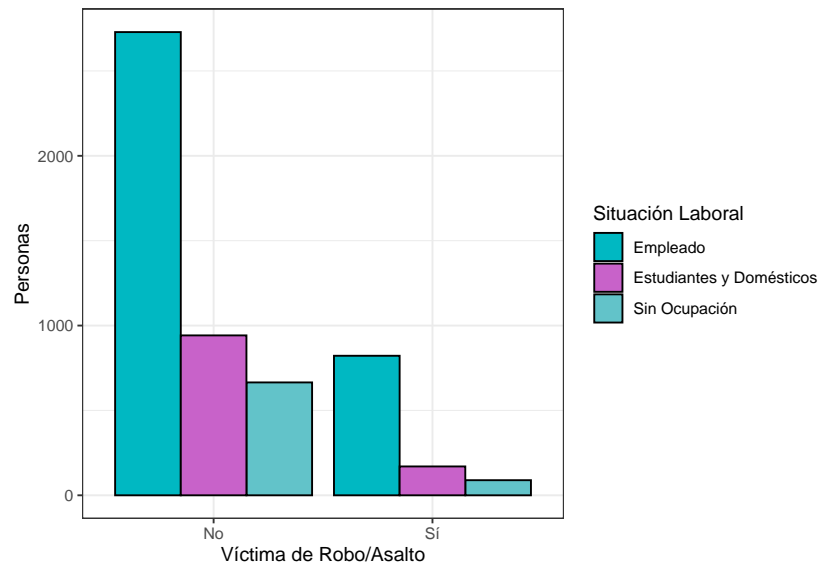
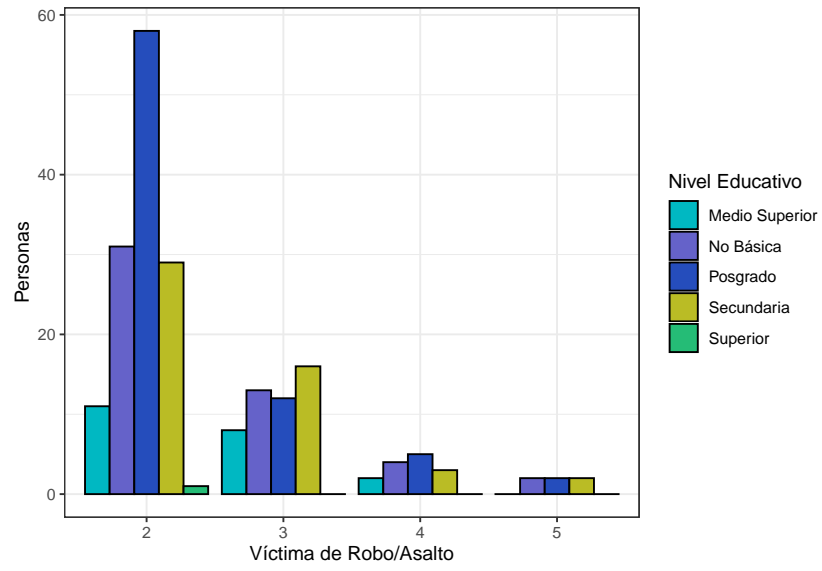


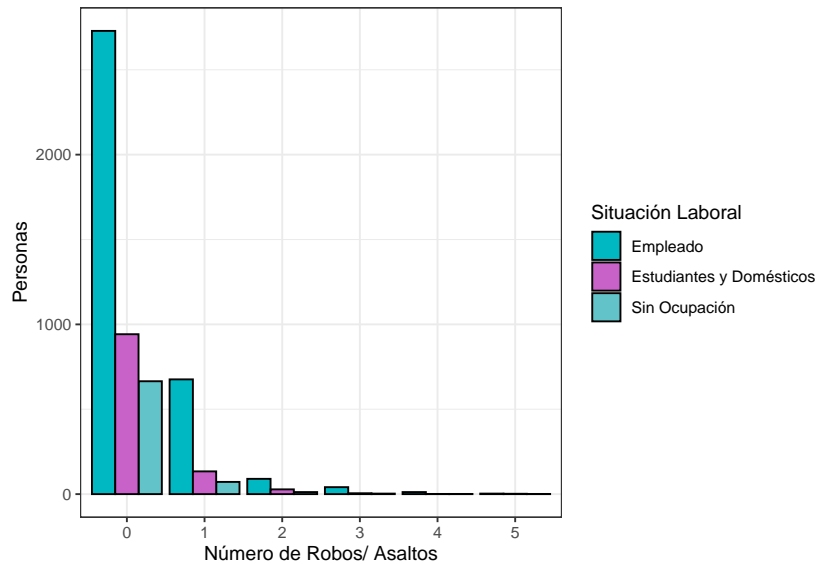










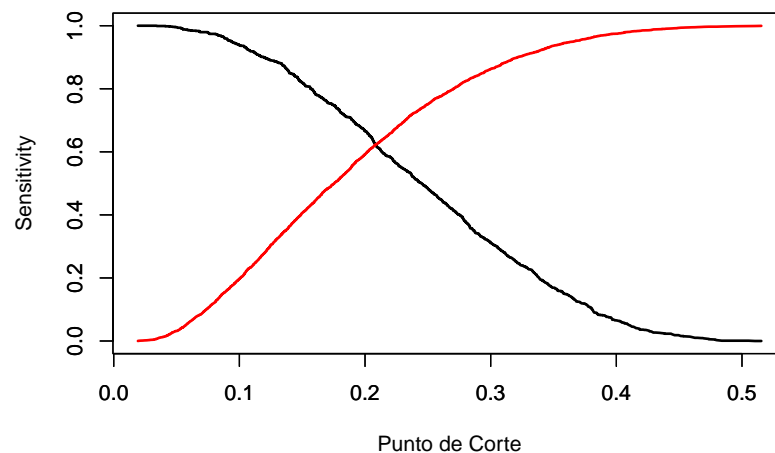


Modelo Logit

Primer modelo

```
##
## Call:
## glm(formula = Vic_Rob_As ~ ., family = "binomial", data = CData_CDMX3)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2037  -0.7203  -0.5547  -0.3551   2.6003
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.59129    0.17915  -3.301 0.000965 ***
## Edad          -0.02033    0.00267  -7.614 2.67e-14 ***
## Imp_Seg1       0.15086    0.09263   1.629 0.103399
## Seg_Mun2       0.49571    0.10222   4.850 1.24e-06 ***
## Mas_Pat_Vil2   0.14669    0.07166   2.047 0.040652 *
## RegionSur     -0.91169    0.10483  -8.696 < 2e-16 ***
## RegionNorte   -0.70081    0.09714  -7.214 5.43e-13 ***
## RegionOriente -0.34078    0.09243  -3.687 0.000227 ***
## Nivel_Edu.L   -0.01029    0.15274  -0.067 0.946269
## Nivel_Edu.Q   -0.34458    0.13307  -2.589 0.009615 **
## Nivel_Edu.C    0.08231    0.09434   0.872 0.382974
## Nivel_Edu^4    0.05765    0.06963   0.828 0.407669
## Sit_LabSin Ocupación -0.28701    0.13220  -2.171 0.029924 *
## Sit_LabEstudiantes y Domésticos -0.52897    0.09572  -5.526 3.27e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##      Null deviance: 5421.1  on 5418  degrees of freedom
## Residual deviance: 5086.2  on 5405  degrees of freedom
## AIC: 5114.2
##
## Number of Fisher Scoring iterations: 4
```



```
##          GVIF Df GVIF^(1/(2*Df))
## Edad      1.360809 1      1.166537
## Imp_Seg    1.016691 1      1.008311
## Seg_Mun    1.053935 1      1.026613
## Mas_Pat_Vil 1.023872 1      1.011865
## Region     1.093541 3      1.015015
## Nivel_Edu  1.254238 4      1.028721
## Sit_Lab    1.204313 2      1.047574
```

```
##
## r1      0      1
## 0 2702 408
## 1 1634 675
```

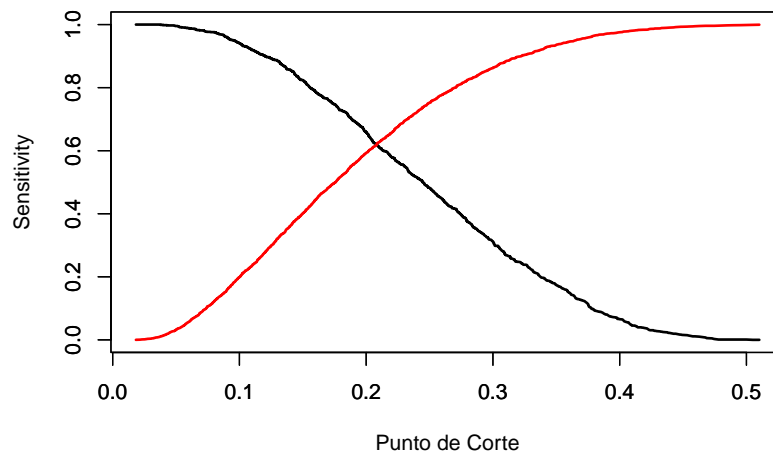
Modelo 2

```
##
## Call:
## glm(formula = Vic_Rob_As ~ Edad + Sit_Lab + Seg_Mun + Mas_Pat_Vil +
##      Nivel_Edu + Region, family = "binomial", data = CData_CDMX3)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1947  -0.7207  -0.5554  -0.3568   2.5602
##
## Coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
```

```

## (Intercept)                -0.475570    0.164138   -2.897 0.003763 **
## Edad                       -0.020384    0.002667   -7.642 2.13e-14 ***
## Sit_LabSin Ocupación       -0.293578    0.132110   -2.222 0.026268 *
## Sit_LabEstudiantes y Domésticos -0.527325    0.095708   -5.510 3.59e-08 ***
## Seg_Mun2                   0.510724    0.101817    5.016 5.27e-07 ***
## Mas_Pat_Vil2               0.146295    0.071641    2.042 0.041146 *
## Nivel_Edu.L                0.002820    0.152506    0.018 0.985247
## Nivel_Edu.Q                -0.346577    0.133041   -2.605 0.009186 **
## Nivel_Edu.C                0.084031    0.094325    0.891 0.372998
## Nivel_Edu^4                0.056999    0.069606    0.819 0.412850
## RegionSur                  -0.916828    0.104767   -8.751 < 2e-16 ***
## RegionNorte                -0.702203    0.097121   -7.230 4.82e-13 ***
## RegionOriente              -0.345291    0.092374   -3.738 0.000185 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 5421.1  on 5418  degrees of freedom
## Residual deviance: 5088.9  on 5406  degrees of freedom
## AIC: 5114.9
##
## Number of Fisher Scoring iterations: 4

```

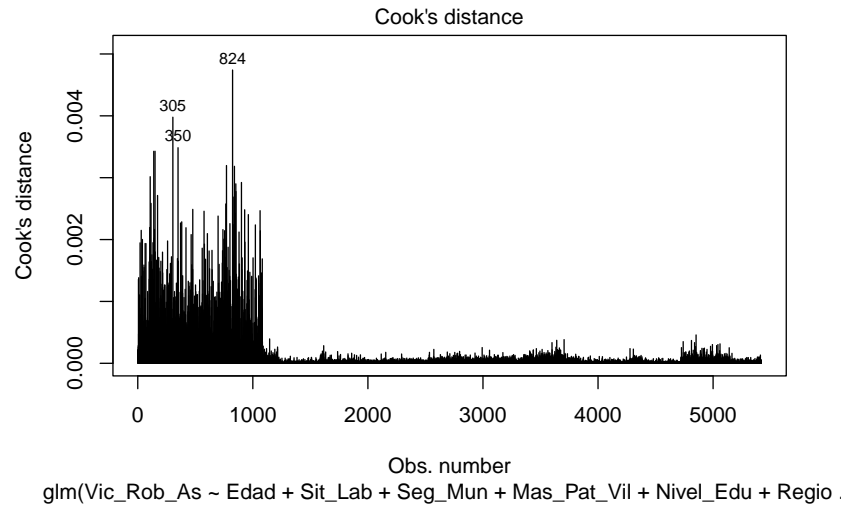


```

##          GVIF Df GVIF^(1/(2*Df))
## Edad      1.359760  1      1.166088
## Sit_Lab   1.203587  2      1.047416
## Seg_Mun   1.045904  1      1.022694
## Mas_Pat_Vil 1.023861  1      1.011860
## Nivel_Edu 1.247328  4      1.028011
## Region    1.092110  3      1.014794
##
##
## r2      0      1

```

```
## 0 2689 411
## 1 1647 672
```



```
## No Studentized residuals with Bonferroni p < 0.05
## Largest |rstudent|:
##      rstudent unadjusted p-value Bonferroni p
## 824 2.572215      0.010105      NA
```

```
##
## vec2    0    1
##    0 3510 815
##    1 826 268
```

Ejemplo

