

Actividad3 Unidad3: Rotación de Cargo

Juan José Restrepo Rosero y Miller Eduardo Perdomo

2025-03-31

Contents

Rotación de cargo	1
1. Selección de variables	2
Variables Categóricas:	3
Variables Cuantitativas:	4
2. Análisis univariado	6
3. Análisis bivariado	7
Comparación de los Coeficientes con las Hipótesis:	9
Comparación de Coeficientes con Hipótesis a través de Odds Ratios:	11
Discusión:	11
Discusión:	11
4. Estimación del modelo	12
5. Evaluación	13
6. Predicciones	14
7. Conclusiones	16
8. Anexos - Repositorio del Código	17

Rotación de cargo

En una organización, se busca comprender y prever los factores que influyen en la **rotación de empleados** entre distintos cargos. La empresa ha recopilado datos históricos sobre el empleo de sus trabajadores, incluyendo variables como la antigüedad en el cargo actual, el nivel de satisfacción laboral, el salario actual, edad y otros factores relevantes. La gerencia planea desarrollar un modelo de regresión logística que permita estimar la probabilidad de que un empleado cambie de cargo en el próximo período y determinar cuales factores indiquen en mayor proporción a estos cambios.

Con esta información, la empresa podrá tomar medidas proactivas para retener a su talento clave, identificar áreas de mejora en la gestión de recursos humanos y fomentar un ambiente laboral más estable y tranquilo. La predicción de la probabilidad de rotación de empleados ayudará a la empresa a tomar decisiones estratégicas informadas y a mantener un equipo de trabajo comprometido y satisfecho en sus roles actuales.

A continuación se describen los pasos que la gerencia ha propuesto para el análisis:

1. Selección de variables

Seleccione 3 variables categóricas (distintas de rotación) y 3 variables cuantitativas, que se consideren estén relacionadas con la rotación.

```
##      Rotación          Edad      Viaje de Negocios  Departamento
## Length:1470      Min.    :18.00      Length:1470      Length:1470
## Class :character  1st Qu.:30.00      Class :character  Class :character
## Mode  :character  Median :36.00      Mode  :character  Mode  :character
##                               Mean  :36.92
##                               3rd Qu.:43.00
##                               Max.   :60.00
## Distancia_Casa      Educación      Campo_Educación      Satisfacción_Ambiental
## Min.    : 1.000      Min.    :1.000      Length:1470      Min.    :1.000
## 1st Qu.: 2.000      1st Qu.:2.000      Class :character  1st Qu.:2.000
## Median : 7.000      Median :3.000      Mode  :character  Median :3.000
## Mean   : 9.193      Mean   :2.913      Mean   :2.722
## 3rd Qu.:14.000      3rd Qu.:4.000      3rd Qu.:4.000
## Max.   :29.000      Max.   :5.000      Max.   :4.000
## Genero              Cargo              Satisfacción_Laboral  Estado_Civil
## Length:1470      Length:1470      Min.    :1.000      Length:1470
## Class :character  Class :character  1st Qu.:2.000      Class :character
## Mode  :character  Mode  :character  Median :3.000      Mode  :character
##                               Mean   :2.729
##                               3rd Qu.:4.000
##                               Max.   :4.000
## Ingreso_Mensual  Trabajos_Anteriores  Horas_Extra
## Min.    : 1009      Min.    :0.000      Length:1470
## 1st Qu.: 2911      1st Qu.:1.000      Class :character
## Median : 4919      Median :2.000      Mode  :character
## Mean   : 6503      Mean   :2.693
## 3rd Qu.: 8379      3rd Qu.:4.000
## Max.   :19999      Max.   :9.000
## Porcentaje_aumento_salarial  Rendimiento_Laboral  Años_Experiencia
## Min.    :11.00      Min.    :3.000      Min.    : 0.00
## 1st Qu.:12.00      1st Qu.:3.000      1st Qu.: 6.00
## Median :14.00      Median :3.000      Median :10.00
## Mean   :15.21      Mean   :3.154      Mean   :11.28
## 3rd Qu.:18.00      3rd Qu.:3.000      3rd Qu.:15.00
## Max.   :25.00      Max.   :4.000      Max.   :40.00
## Capacitaciones  Equilibrio_Trabajo_Vida      Antigüedad      Antigüedad_Cargo
## Min.    :0.000      Min.    :1.000      Min.    : 0.000      Min.    : 0.000
## 1st Qu.:2.000      1st Qu.:2.000      1st Qu.: 3.000      1st Qu.: 2.000
## Median :3.000      Median :3.000      Median : 5.000      Median : 3.000
## Mean   :2.799      Mean   :2.761      Mean   : 7.008      Mean   : 4.229
## 3rd Qu.:3.000      3rd Qu.:3.000      3rd Qu.: 9.000      3rd Qu.: 7.000
```

```
## Max.      :6.000    Max.      :4.000          Max.      :40.000    Max.      :18.000
## Años_ultima_promoción Años_acargo_con_mismo_jefe
## Min.      : 0.000    Min.      : 0.000
## 1st Qu.: 0.000    1st Qu.: 2.000
## Median : 1.000    Median : 3.000
## Mean   : 2.188    Mean   : 4.123
## 3rd Qu.: 3.000    3rd Qu.: 7.000
## Max.    :15.000    Max.    :17.000

## tibble [1,470 x 24] (S3: tbl_df/tbl/data.frame)
## $ Rotación                : chr [1:1470] "Si" "No" "Si" "No" ...
## $ Edad                    : num [1:1470] 41 49 37 33 27 32 59 30 38 36 ...
## $ Viaje de Negocios       : chr [1:1470] "Raramente" "Frecuentemente" "Raramente" "Frecuentemente" ...
## $ Departamento           : chr [1:1470] "Ventas" "IyD" "IyD" "IyD" ...
## $ Distancia_Casa          : num [1:1470] 1 8 2 3 2 2 3 24 23 27 ...
## $ Educación               : num [1:1470] 2 1 2 4 1 2 3 1 3 3 ...
## $ Campo_Educación         : chr [1:1470] "Ciencias" "Ciencias" "Otra" "Ciencias" ...
## $ Satisfacción_Ambiental  : num [1:1470] 2 3 4 4 1 4 3 4 4 3 ...
## $ Genero                  : chr [1:1470] "F" "M" "M" "F" ...
## $ Cargo                   : chr [1:1470] "Ejecutivo_Ventas" "Investigador_Cientifico" "Tecnico_L..."
## $ Satisfacción_Laboral    : num [1:1470] 4 2 3 3 2 4 1 3 3 3 ...
## $ Estado_Civil            : chr [1:1470] "Soltero" "Casado" "Soltero" "Casado" ...
## $ Ingreso_Mensual         : num [1:1470] 5993 5130 2090 2909 3468 ...
## $ Trabajos_Anteriores     : num [1:1470] 8 1 6 1 9 0 4 1 0 6 ...
## $ Horas_Extra             : chr [1:1470] "Si" "No" "Si" "Si" ...
## $ Porcentaje_aumento_salarial: num [1:1470] 11 23 15 11 12 13 20 22 21 13 ...
## $ Rendimiento_Laboral    : num [1:1470] 3 4 3 3 3 3 4 4 4 3 ...
## $ Años_Experiencia        : num [1:1470] 8 10 7 8 6 8 12 1 10 17 ...
## $ Capacitaciones          : num [1:1470] 0 3 3 3 3 2 3 2 2 3 ...
## $ Equilibrio_Trabajo_Vida : num [1:1470] 1 3 3 3 3 2 2 3 3 2 ...
## $ Antigüedad              : num [1:1470] 6 10 0 8 2 7 1 1 9 7 ...
## $ Antigüedad_Cargo        : num [1:1470] 4 7 0 7 2 7 0 0 7 7 ...
## $ Años_ultima_promoción    : num [1:1470] 0 1 0 3 2 3 0 0 1 7 ...
## $ Años_acargo_con_mismo_jefe : num [1:1470] 5 7 0 0 2 6 0 0 8 7 ...
```

A partir de la estructura del dataframe sobre la *rotación laboral*, se presenta la información correspondiente a 1470 empleados. En esta, se destacan aspectos fundamentales como la **edad promedio de los trabajadores (36,92 años)** y su **experiencia laboral media (11,28 años)**, complementados por una amplia gama de características demográficas y profesionales, tales como **género, estado civil, departamento y cargo**. Asimismo, se analizan variables relacionadas con la **satisfacción** y el **desempeño en el trabajo**, además de examinar factores potencialmente determinantes de la rotación, tales como el **equilibrio entre la vida laboral y personal**, la **realización de horas extras** y la **antigüedad en la empresa**.

Teniendo en cuenta la información contenida en el dataframe, se procede a la selección de las siguientes variables:

Variables Categóricas:

- **Viaje de Negocios:** Esta variable indica la frecuencia con la que los empleados realizan desplazamientos por motivos laborales. Una alta frecuencia de viajes podría perturbar el equilibrio entre la vida profesional y personal, lo que, a su vez, podría contribuir a mayores tasas de rotación.
- **Departamento:** El departamento o área en la que se desempeña un empleado influye directamente en su satisfacción laboral y en las oportunidades de crecimiento profesional. La pertenencia a ciertos de-

partamentos puede estar asociada a condiciones laborales que favorecen o desincentivan la permanencia en la organización.

- **Horas Extra:** El número de horas adicionales trabajadas por los empleados puede reflejar una carga laboral excesiva, la cual está vinculada al agotamiento y a la insatisfacción, incrementando el riesgo de que los empleados opten por abandonar la empresa.

Variables Cuantitativas:

- **Edad:** La edad de los empleados es un factor relevante en el estudio de la rotación, ya que las diferentes etapas de la vida profesional y personal implican prioridades y necesidades laborales que pueden influir en la decisión de permanecer o cambiar de empleo.
- **Ingreso Mensual:** El salario mensual es un determinante clave en la retención de personal; una remuneración competitiva puede favorecer la estabilidad laboral, mientras que una compensación inadecuada podría aumentar la tendencia a la rotación.
- **Antigüedad en el Cargo:** El tiempo que un empleado ha ocupado su posición actual refleja su nivel de adaptación y compromiso. Aquellos con menor antigüedad en el cargo pueden estar más inclinados a buscar nuevas oportunidades, mientras que una mayor permanencia podría asociarse a una rotación reducida.

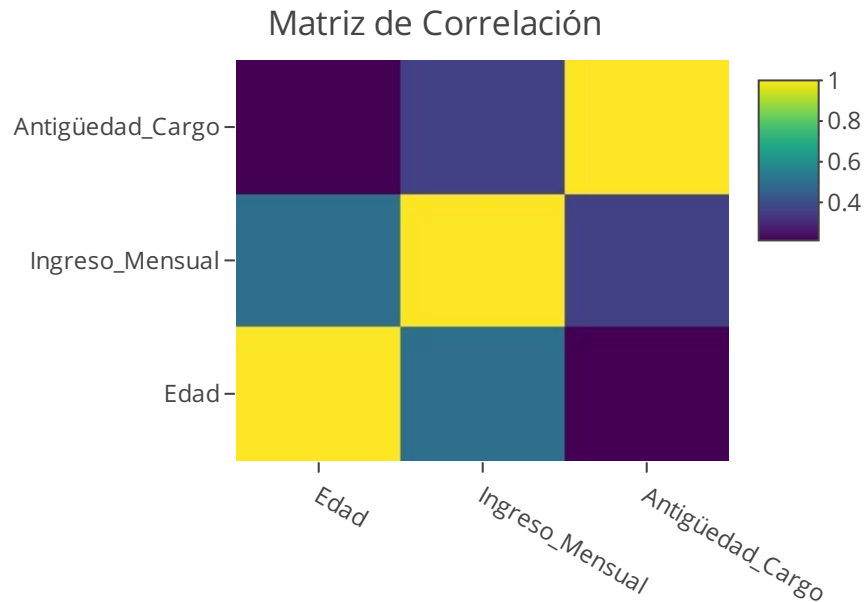
```
## tibble [1,470 x 6] (S3: tbl_df/tbl/data.frame)
## $ Viaje de Negocios: chr [1:1470] "Raramente" "Frecuentemente" "Raramente" "Frecuentemente" ...
## $ Departamento    : chr [1:1470] "Ventas" "IyD" "IyD" "IyD" ...
## $ Horas_Extra     : chr [1:1470] "Si" "No" "Si" "Si" ...
## $ Edad            : num [1:1470] 41 49 37 33 27 32 59 30 38 36 ...
## $ Ingreso_Mensual : num [1:1470] 5993 5130 2090 2909 3468 ...
## $ Antigüedad_Cargo : num [1:1470] 4 7 0 7 2 7 0 0 7 7 ...

## Viaje de Negocios Departamento Horas_Extra Edad
## Length:1470 Length:1470 Length:1470 Min. :18.00
## Class :character Class :character Class :character 1st Qu.:30.00
## Mode :character Mode :character Mode :character Median :36.00
## Mean :36.92
## 3rd Qu.:43.00
## Max. :60.00
## Ingreso_Mensual Antigüedad_Cargo
## Min. : 1009 Min. : 0.000
## 1st Qu.: 2911 1st Qu.: 2.000
## Median : 4919 Median : 3.000
## Mean : 6503 Mean : 4.229
## 3rd Qu.: 8379 3rd Qu.: 7.000
## Max. :19999 Max. :18.000

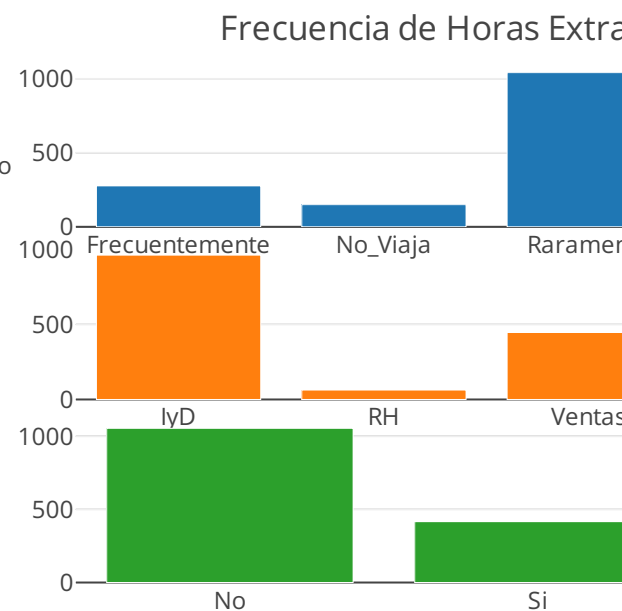
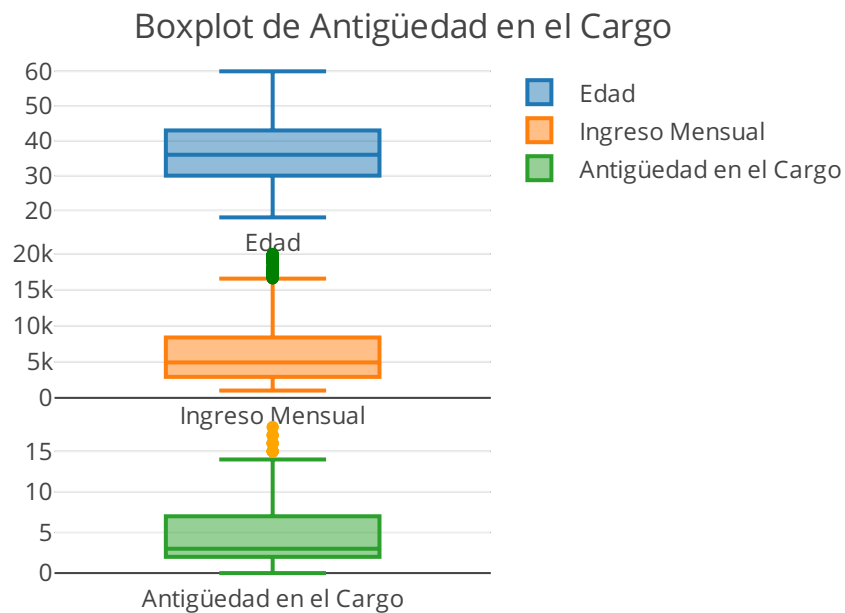
## # A tibble: 6 x 6
## 'Viaje de Negocios' Departamento Horas_Extra Edad Ingreso_Mensual
## <chr> <chr> <chr> <dbl> <dbl>
## 1 Raramente Ventas Si 41 5993
## 2 Frecuentemente IyD No 49 5130
## 3 Raramente IyD Si 37 2090
## 4 Frecuentemente IyD Si 33 2909
## 5 Raramente IyD No 27 3468
## 6 Frecuentemente IyD No 32 3068
## # i 1 more variable: Antigüedad_Cargo <dbl>
```

Clase de las variables seleccionadas:

```
## Viaje de Negocios      Departamento      Horas_Extra      Edad
##      "character"      "character"      "character"      "numeric"
##      Ingreso_Mensual  Antigüedad_Cargo
##      "numeric"      "numeric"
```



```
##      Edad  Ingreso_Mensual  Antigüedad_Cargo
## Edad      1.0000000      0.4977532      0.2128421
## Ingreso_Mensual 0.4977532      1.0000000      0.3638177
## Antigüedad_Cargo 0.2128421      0.3638177      1.0000000
```



Hipótesis:

1. Viaje de Negocios:

Se plantea que la frecuencia de los viajes de negocios se asocia con la rotación de empleados. Es decir, se espera que aquellos empleados que viajan con mayor regularidad tengan una mayor probabilidad de abandonar la empresa en comparación con quienes realizan estos viajes con menor frecuencia.

2. Departamento:

Se hipotetiza que el área o departamento en el que trabaja un empleado influye en su tasa de rotación. La idea es que ciertos departamentos podrían presentar tasas de rotación más elevadas debido a diferencias en la cultura organizacional, la carga de trabajo o las oportunidades de desarrollo profesional.

3. Horas Extra:

La hipótesis sostiene que trabajar horas adicionales se relaciona con un incremento en la rotación. En otras palabras, se espera que los empleados que realizan horas extra tengan una mayor tendencia a abandonar la empresa, posiblemente por el desgaste y la insatisfacción generados por la sobrecarga laboral.

4. Edad:

Se plantea que la edad de los empleados afecta su probabilidad de rotación. Específicamente, se anticipa que los trabajadores más jóvenes, en búsqueda de oportunidades de crecimiento y desarrollo, pueden tener una mayor tendencia a cambiar de empleo, mientras que los empleados de mayor edad podrían mostrar mayor estabilidad debido a compromisos laborales y personales consolidados.

5. Ingreso Mensual:

Se espera que el nivel de ingresos tenga un impacto en la rotación. La hipótesis es que los empleados con salarios más bajos son más propensos a buscar mejores oportunidades de remuneración, mientras que aquellos que reciben ingresos más altos podrían estar más inclinados a permanecer en la empresa debido a la estabilidad financiera que les proporciona su salario.

6. Antigüedad en el Cargo:

Se supone que la duración que un empleado ocupa su puesto actual influye en su rotación. Se anticipa que los empleados con menor antigüedad en el cargo tienen una mayor probabilidad de buscar nuevas oportunidades, en comparación con aquellos con más años en el puesto, quienes podrían beneficiarse de una mayor estabilidad y oportunidades de promoción.

2. Análisis univariado

Realiza un análisis univariado (caracterización) de la información contenida en la base de datos rotación.

Nota: Los indicadores o gráficos se usan dependiendo del tipo de variable (cuantitativas o cualitativas). Incluir interpretaciones de la variable *rotación*.

Tabla de estadísticas descriptivas para variables cuantitativas:

Variable	Mínimo	X1er.Cuartil	Mediana	Media	X3er.Cuartil	Máximo
Edad	18	30	36	36.924490	43	60
Ingreso Mensual	1009	2911	4919	6502.931293	8379	19999
Antigüedad en el Cargo	0	2	3	4.229252	7	18

##

Tabla de frecuencia de las categorías para variables cualitativas:

Variable	Frecuencia
Viaje de Negocios	3
Departamento	3
Horas Extra	2

3. Análisis bivariado

Realiza un análisis de bivariado en donde la variable respuesta sea *rotacion* codificada de la siguiente manera ($y = 1$ es si rotación, $y = 0$ es no rotación). Con base en estos resultados identifique cuales son las variables determinantes de la rotación e interpretar el signo del coeficiente estimado. Compare estos resultados con la hipótesis planteada en el punto 2.

```
##
## Call:
## glm(formula = Rotación ~ 'Viaje de Negocios', family = "binomial",
##      data = rotacion)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -1.1034     0.1389  -7.943 1.98e-15 ***
## 'Viaje de Negocios'No_Viaja  -1.3389     0.3315  -4.039 5.36e-05 ***
## 'Viaje de Negocios'Raramente  -0.6346     0.1638  -3.873 0.000107 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 1298.6  on 1469  degrees of freedom
## Residual deviance: 1274.8  on 1467  degrees of freedom
## AIC: 1280.8
##
## Number of Fisher Scoring iterations: 5

##
## Call:
## glm(formula = Rotación ~ Departamento, family = "binomial",
##      data = rotacion)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -1.82866     0.09342 -19.576 < 2e-16 ***
## DepartamentoRH      0.38175     0.33417   1.142  0.25330
## DepartamentoVentas  0.48116     0.14974   3.213  0.00131 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 1298.6  on 1469  degrees of freedom
## Residual deviance: 1288.1  on 1467  degrees of freedom
## AIC: 1294.1
```

```

##
## Number of Fisher Scoring iterations: 4

##
## Call:
## glm(formula = Rotación ~ Horas_Extra, family = "binomial", data = rotacion)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.1496     0.1007 -21.338  <2e-16 ***
## Horas_ExtraSi  1.3274     0.1466   9.056  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1298.6  on 1469  degrees of freedom
## Residual deviance: 1217.2  on 1468  degrees of freedom
## AIC: 1221.2
##
## Number of Fisher Scoring iterations: 4

##
## Call:
## glm(formula = Rotación ~ Edad, family = "binomial", data = rotacion)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.20637     0.30597   0.674    0.5
## Edad        -0.05225     0.00870  -6.006  1.9e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1298.6  on 1469  degrees of freedom
## Residual deviance: 1259.1  on 1468  degrees of freedom
## AIC: 1263.1
##
## Number of Fisher Scoring iterations: 4

##
## Call:
## glm(formula = Rotación ~ Ingreso_Mensual, family = "binomial",
##      data = rotacion)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -9.291e-01  1.292e-01  -7.191  6.43e-13 ***
## Ingreso_Mensual -1.271e-04  2.162e-05  -5.879  4.12e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##

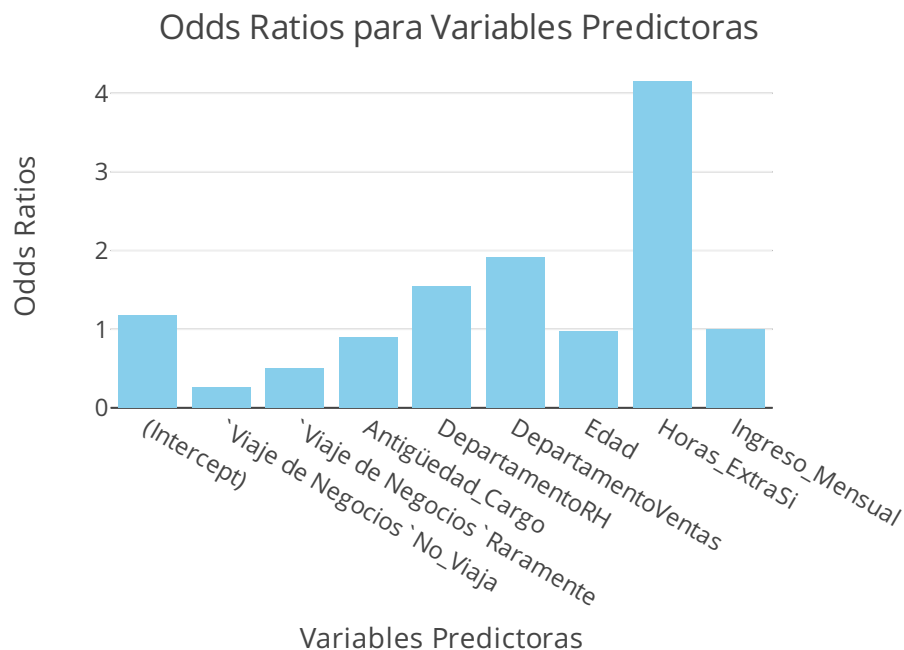
```



```
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1298.6 on 1469 degrees of freedom
## Residual deviance: 1253.1 on 1468 degrees of freedom
## AIC: 1257.1
##
## Number of Fisher Scoring iterations: 5

##
## Call:
## glm(formula = Rotación ~ Antigüedad_Cargo, family = "binomial",
## data = rotacion)
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.11841 0.10380 -10.775 < 2e-16 ***
## Antigüedad_Cargo -0.14628 0.02424 -6.033 1.61e-09 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1298.6 on 1469 degrees of freedom
## Residual deviance: 1255.9 on 1468 degrees of freedom
## AIC: 1259.9
##
## Number of Fisher Scoring iterations: 5
```



Comparación de los Coeficientes con las Hipótesis:

- Viaje de Negocios:

- **Coefficiente:** No_Viaja: -1.3389, Raramente: -0.6346
- **Hipótesis:** Se esperaba que los empleados que realizan viajes frecuentes tengan una mayor propensión a la rotación en comparación con aquellos que viajan con menos frecuencia.
- **Conclusión:** Los coeficientes negativos confirman la hipótesis, ya que indican que quienes viajan menos presentan menores probabilidades de rotación.
- **Departamento:**
 - **Coefficiente:** Departamento RH: 0.38175, Ventas: 0.48116
 - **Hipótesis:** Se anticipaba que ciertos departamentos mostrarían tasas de rotación más elevadas debido a diferencias en la cultura organizacional, la carga de trabajo y las oportunidades de crecimiento.
 - **Conclusión:** La presencia de coeficientes positivos en algunos departamentos respalda la hipótesis, evidenciando que estos tienen una mayor tasa de rotación.
- **Horas Extra:**
 - **Coefficiente:** Horas Extra Sí: 1.3274
 - **Hipótesis:** Se postulaba que trabajar horas extra incrementa la probabilidad de rotación.
 - **Conclusión:** El coeficiente positivo respalda la hipótesis, ya que sugiere que los empleados que realizan horas adicionales tienen una mayor tendencia a rotar.
- **Edad:**
 - **Coefficiente:** Edad: -0.05225
 - **Hipótesis:** Se esperaba que los empleados más jóvenes tuvieran mayor rotación en comparación con los de mayor edad.
 - **Conclusión:** Aunque el coeficiente es negativo, su magnitud pequeña indica una relación débil, por lo que la hipótesis no se confirma de manera contundente.
- **Ingreso Mensual:**
 - **Coefficiente:** Ingreso Mensual: -0.0001271
 - **Hipótesis:** Se preveía que los empleados con salarios más bajos tendrían mayor probabilidad de rotación.
 - **Conclusión:** Dado que el coeficiente es negativo pero casi insignificante, la relación entre el ingreso y la rotación es débil, lo que contradice la hipótesis planteada.
- **Antigüedad en el Cargo:**
 - **Coefficiente:** Antigüedad en el Cargo: -0.14628
 - **Hipótesis:** Se esperaba que los empleados con menor antigüedad en el cargo mostraran una mayor tendencia a rotar que aquellos con más tiempo en el puesto.
 - **Conclusión:** Aunque el coeficiente es negativo, su pequeño valor sugiere una relación débil entre la antigüedad en el cargo y la rotación, lo que no respalda completamente la hipótesis.

Table 3: Odds Ratios para Variables Predictoras

	Variables	Odds_Ratios
Horas_ExtraSi	Horas_ExtraSi	4.1551317
DepartamentoVentas	DepartamentoVentas	1.9145694
DepartamentoRH	DepartamentoRH	1.5471128
(Intercept)	(Intercept)	1.1758364
Ingreso_Mensual	Ingreso_Mensual	0.9999181
Edad	Edad	0.9683310
Antigüedad_Cargo	Antigüedad_Cargo	0.8942634
Viaje de NegociosRaramente	Viaje de NegociosRaramente	0.5067194
Viaje de NegociosNo_Viaja	Viaje de NegociosNo_Viaja	0.2581596

Variables	Odds_Ratios
-----------	-------------

Comparación de Coeficientes con Hipótesis a través de Odds Ratios:

Los Odds Ratios de las variables predictoras enriquecen el informe al ofrecer una medida cuantitativa de la relación entre cada variable y la probabilidad de rotación de empleados. Estos indicadores cuantifican el cambio en dicha probabilidad cuando una variable aumenta en una unidad, manteniendo constantes las demás variables.

- **Horas Extra (Horas_ExtraSi):**

Un Odds Ratio de 4.155 sugiere que los empleados que realizan horas extra tienen aproximadamente 4.155 veces más probabilidades de rotar que aquellos que no lo hacen, al controlar el efecto de las otras variables.

- **Departamento (DepartamentoVentas y DepartamentoRH):**

Con Odds Ratios de 1.914 y 1.547 respectivamente, se evidencia que pertenecer a los departamentos de Ventas o Recursos Humanos incrementa la probabilidad de rotación en aproximadamente 1.914 y 1.547 veces, en comparación con otros departamentos, manteniendo constantes las demás variables.

- **Ingreso Mensual (Ingreso_Mensual):**

Un Odds Ratio de 0.9999181 indica que, por cada aumento unitario en el ingreso mensual, la probabilidad de rotación disminuye en un 0.0000819%. Esto denota una asociación extremadamente débil entre el nivel de ingreso y la rotación.

- **Edad (Edad) y Antigüedad en el Cargo (Antigüedad_Cargo):**

Los Odds Ratios de 0.968 y 0.894 implican que, por cada incremento unitario en la edad o en la antigüedad en el cargo, la probabilidad de rotación se reduce en un 3.2% y un 10.6% respectivamente, cuando se mantienen las demás variables constantes.

- **Viaje de Negocios (Viaje de Negocios Raramente y No Viaja):**

Con Odds Ratios de 0.5067 y 0.2582, se observa que los empleados que viajan raramente o que no viajan presentan una menor probabilidad de rotación en comparación con aquellos que realizan viajes frecuentes, al mantener constante el efecto de las otras variables.

En síntesis, los Odds Ratios permiten cuantificar tanto la intensidad como la dirección de la relación entre cada variable y la probabilidad de rotación, siendo una métrica que facilita la comprensión detallada de cómo cada factor incide de manera relativa en el fenómeno de la rotación laboral.

Discusión:

Discusión:

Al interpretar los coeficientes de un modelo de regresión logística, se evalúa tanto la dirección como la magnitud del efecto de cada variable en la probabilidad de rotación. No obstante, al analizar los Odds Ratios se obtiene una perspectiva más detallada y cuantificable de dicha relación.

No obstante, las discrepancias que existen entre la *interpretación de los coeficientes* y los *Odds Ratios* pueden surgir debido a la transformación logarítmica que se aplica en el modelo para obtener los coeficientes. Esta transformación no siempre traduce de forma directa la asociación entre las variables y la probabilidad de rotación.

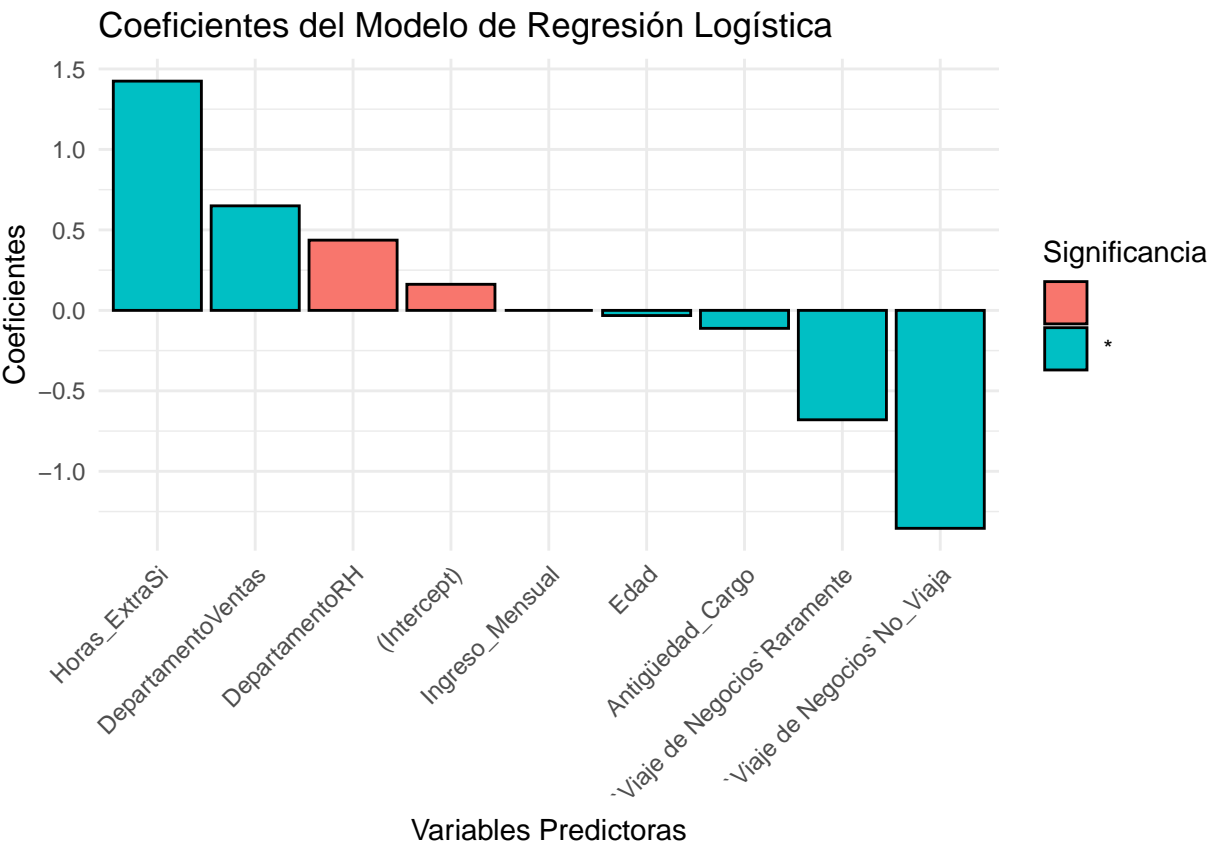
Por ejemplo, en el caso del ingreso mensual, mientras que el coeficiente estimado puede sugerir una relación débil con la rotación, el Odds Ratio revela que incluso pequeños incrementos en el ingreso pueden estar

vinculados a reducciones significativas en la probabilidad de rotación. Esta diferencia resalta la importancia de analizar ambos indicadores para comprender de manera completa cómo las variables predictoras influyen en la variable de respuesta.

En resumen, aunque los coeficientes de la regresión logística proporcionan información sobre la dirección y magnitud de las relaciones, los Odds Ratios ofrecen una interpretación más intuitiva al mostrar el impacto relativo de los cambios en cada variable sobre la probabilidad de rotación.

4. Estimación del modelo

Realiza la estimación de un modelo de regresión logística en el cual la variable respuesta es **rotación** ($y = 1$ indica rotación y $y = 0$ indica no rotación) y las covariables corresponden a las 6 seleccionadas en el punto 1. Interprete los coeficientes del modelo y la significancia de los parámetros.



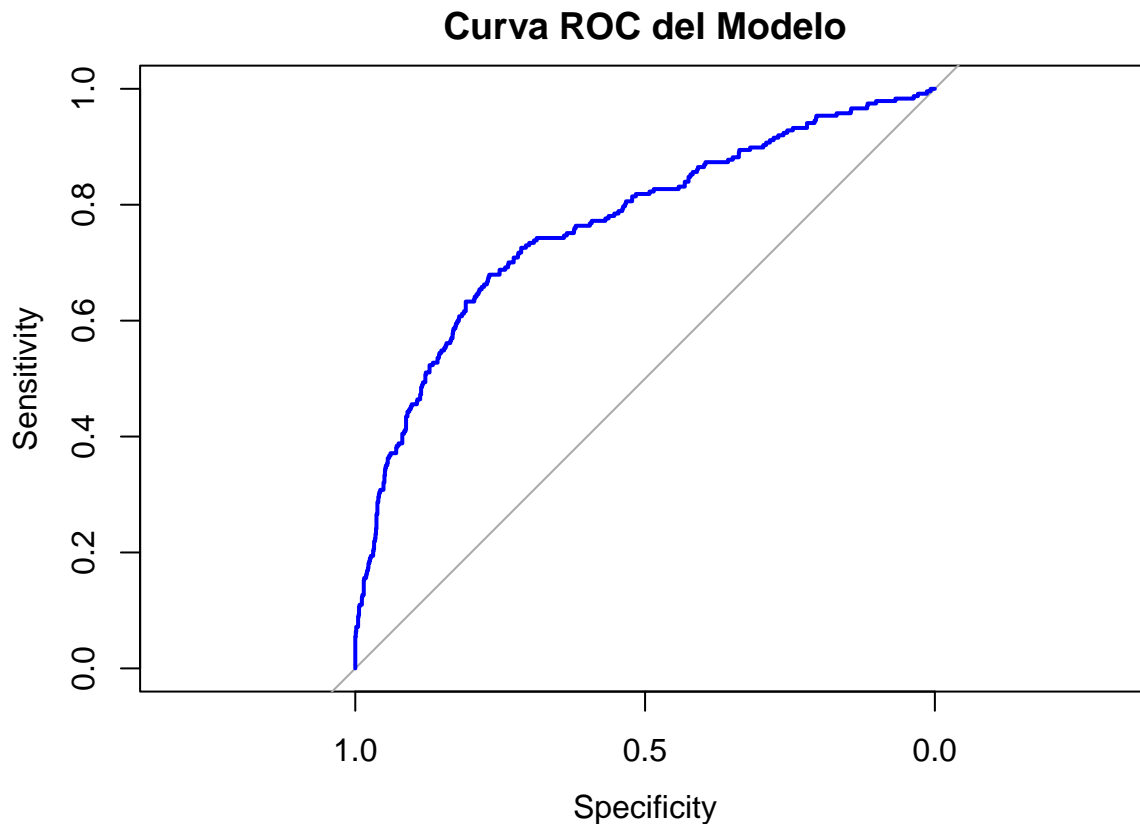
	Variables	Coeficientes	Significancia
(Intercept)	(Intercept)	0.1619797	
Viaje de NegociosNo_Viaja	Viaje de NegociosNo_Viaja	-1.3541771	*
Viaje de NegociosRaramente	Viaje de NegociosRaramente	-0.6797980	*
DepartamentoRH	DepartamentoRH	0.4363905	
DepartamentoVentas	DepartamentoVentas	0.6494927	*
Horas_ExtraSi	Horas_ExtraSi	1.4243441	*
Edad	Edad	-0.0321813	*
Ingreso_Mensual	Ingreso_Mensual	-0.0000819	*
Antigüedad_Cargo	Antigüedad_Cargo	-0.1117549	*

A continuación se presenta la interpretación de los coeficientes obtenidos por el modelo:

1. **Intercepto:** Representa el logaritmo de la razón de probabilidades de rotación cuando todas las variables predictoras se establecen en cero. En este caso, su valor es 0.1619797, lo que indica que, en ausencia de cualquier influencia de las demás variables, el log-odds de rotación es aproximadamente 0.16.
2. **Viaje de NegociosNo_Viaja:** El coeficiente negativo de -1.3541771 sugiere que, manteniendo constantes las demás variables, los empleados que no viajan por negocios presentan log-odds de rotación inferiores en comparación con aquellos que sí realizan viajes.
3. **Viaje de NegociosRaramente:** Con un coeficiente de -0.6797980, se evidencia que los empleados que viajan raramente por negocios tienen log-odds de rotación menores respecto a quienes viajan frecuentemente, siempre que las demás variables permanezcan constantes.
4. **DepartamentoRH y DepartamentoVentas:** Los coeficientes positivos de 0.4363905 y 0.6494927, respectivamente, indican que pertenecer al departamento de Recursos Humanos o al de Ventas se asocia con mayores log-odds de rotación en comparación con el departamento de referencia, asumiendo que las demás variables se mantienen sin cambios.
5. **Horas_ExtraSi:** Un coeficiente positivo de 1.4243441 muestra que trabajar horas extra está vinculado a un incremento en los log-odds de rotación, lo que implica una mayor probabilidad de rotación, manteniendo constantes las demás variables.
6. **Edad y Antigüedad_Cargo:** Los coeficientes negativos de -0.0321813 para la edad y -0.1117549 para la antigüedad en el cargo indican que, conforme aumenta la edad o la antigüedad, los log-odds de rotación disminuyen, lo que sugiere una menor probabilidad de rotación.
7. **Ingreso_Mensual:** El coeficiente negativo de -0.0000819 implica que cada incremento unitario en el ingreso mensual se asocia con una leve disminución en los log-odds de rotación. Aunque el efecto es casi despreciable, señala una relación inversa entre el ingreso y la probabilidad de rotación.

5. Evaluación

Evaluar el poder predictivo del modelo con base en la curva ROC y el AUC.



El área bajo la curva (AUC) es: 0.7654207

Curva ROC:

La curva ROC representa la relación entre la tasa de verdaderos positivos (TPR) y la tasa de falsos positivos (FPR) a lo largo de diferentes umbrales de clasificación. Se dibuja en un plano bidimensional, donde el eje X corresponde a la FPR y el eje Y a la TPR. Una curva ROC ideal se ubica en la esquina superior izquierda del gráfico, lo que significa que el modelo logra identificar correctamente todos los casos positivos sin incurrir en falsos positivos.

Área bajo la curva (AUC):

El AUC corresponde a la proporción del área total bajo la curva ROC. Este indicador se expresa mediante un valor numérico que oscila entre 0 y 1. Un AUC de 1 implica que el modelo clasifica perfectamente, mientras que un AUC de 0 equivale a un rendimiento similar al de una clasificación aleatoria.

En la imagen presentada, la curva ROC se sitúa por encima de la diagonal que representa la clasificación aleatoria, lo que demuestra que el modelo supera a una simple adivinanza en la clasificación de casos. Con un AUC de 0.7654207, se evidencia un rendimiento sólido del modelo, ya que este valor indica que el modelo detecta de manera eficiente los verdaderos positivos y mantiene bajas las tasas de falsas alarmas.

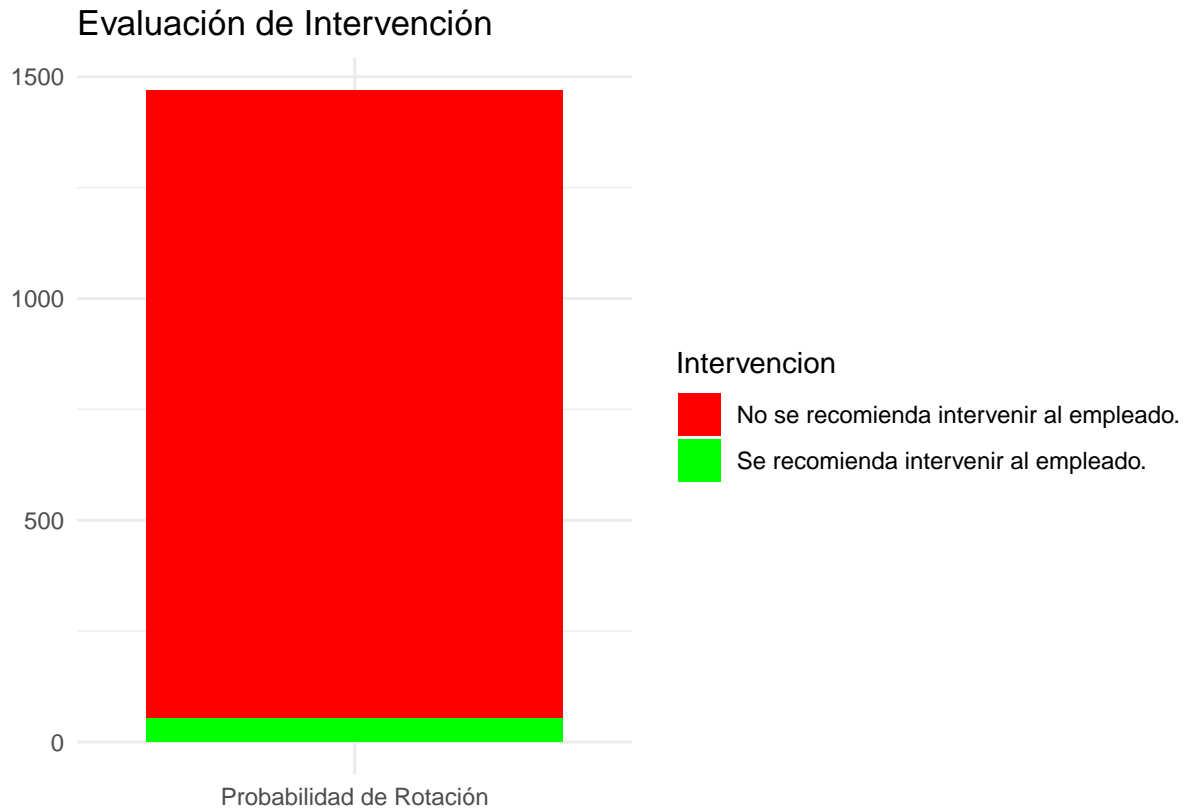
6. Predicciones

Realiza una predicción la probabilidad de que un individuo (hipotético) rote y defina un corte para decidir si se debe intervenir a este empleado o no (posible estrategia para motivar al empleado).

```
##   Viaje.de.Negocios Departamento Horas_Extra Edad Ingreso_Mensual
## 1           No           Ventas           Si   35           3800000
##   Antigüedad_Cargo
## 1           5
```

Lucía es una empleada dedicada y comprometida en una empresa de ventas. Ha optado por un puesto que no requiere viajes de negocios, debido a que prefiere pasar tiempo en casa con su familia. A pesar de su compromiso, últimamente se ha visto obligada a trabajar horas extras para satisfacer las demandas laborales. A sus 35 años, Lucía se encuentra en una etapa estable de su vida profesional, contando con 5 años de antigüedad en su cargo y percibiendo un ingreso mensual de 3,800,000 pesos.

```
##   Probabilidad.de.rotación Recomendación.de.intervención
## 1           0.3316           No (Probabilidad <= 0.5)
## 2           0.0680           No (Probabilidad <= 0.5)
## 3           0.3881           No (Probabilidad <= 0.5)
## 4           0.3785           No (Probabilidad <= 0.5)
## 5           0.1308           No (Probabilidad <= 0.5)
## 6           0.1300           No (Probabilidad <= 0.5)
## 7           0.2296           No (Probabilidad <= 0.5)
## 8           0.1540           No (Probabilidad <= 0.5)
## 9           0.0677           No (Probabilidad <= 0.5)
## 10          0.0528           No (Probabilidad <= 0.5)
## 11          0.0920           No (Probabilidad <= 0.5)
## 12          0.2832           No (Probabilidad <= 0.5)
## 13          0.1216           No (Probabilidad <= 0.5)
## 14          0.1137           No (Probabilidad <= 0.5)
## 15          0.4051           No (Probabilidad <= 0.5)
## 16          0.0365           No (Probabilidad <= 0.5)
## 17          0.3505           No (Probabilidad <= 0.5)
## 18          0.3282           No (Probabilidad <= 0.5)
## 19          0.0234           No (Probabilidad <= 0.5)
## 20          0.2968           No (Probabilidad <= 0.5)
```



En base a la información procesada por el modelo, se concluye que no es necesaria ninguna intervención en el caso de la colaboradora Lucía, dado que se encuentra cómoda y valorada en su posición actual.

7. Conclusiones

En las conclusiones se discute la estrategia para reducir la rotación en la empresa, fundamentada en las variables que resultaron significativas en el análisis bivariado. A partir de los hallazgos del punto 3, se proponen las siguientes cuatro estrategias, basadas en las variables Viaje de Negocios, Departamento y Horas Extra:

- **Mejorar el ambiente laboral y la conciliación:**

Se plantea disminuir la necesidad de realizar viajes frecuentes por negocios, especialmente en aquellos departamentos con elevadas tasas de rotación. Esto se puede lograr mediante la implementación de políticas que promuevan un adecuado equilibrio entre el trabajo y la vida personal, tales como la flexibilidad laboral, la adopción de programas de trabajo remoto y la oferta de beneficios orientados a la conciliación.

- **Optimizar la distribución de tareas y responsabilidades:**

Se recomienda revisar la asignación de la carga de trabajo en departamentos con altas tasas de rotación, como el de Ventas, con el objetivo de reducir la dependencia de horas extras de manera constante. Esto implica evaluar la distribución de tareas, asegurar una asignación óptima de recursos y establecer sistemas de rotación de turnos para disminuir el agotamiento laboral.

- **Incentivar el desarrollo profesional y el crecimiento interno:**

Para contrarrestar la rotación en departamentos específicos, se sugiere implementar programas de desarrollo profesional y de promoción interna. Dichos programas pueden incluir capacitación continua, mentoría y planes de carrera personalizados, que incentiven a los empleados a permanecer en la organización y a desarrollar su potencial.

- **Revisar y ajustar la política de compensación:**

Se propone evaluar la estructura de remuneración para garantizar que sea competitiva y equitativa, en especial en aquellos departamentos donde se observa alta rotación. Esto podría incluir ajustes salariales, la implementación de bonificaciones por desempeño y la incorporación de beneficios adicionales que reconozcan el esfuerzo y compromiso de los empleados.

Estas estrategias están diseñadas para abordar de manera específica las variables identificadas en el análisis bivariado, con el fin de reducir la rotación de empleados y promover un entorno laboral más estable y satisfactorio.

8. Anexos - Repositorio del Código

Repositorio de Github