

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	1
2. Análisis univariado	6
3. Análisis bivariado	6
4. Estimación del modelo	11
5. Evaluación	12
6. Predicciones	13
7. Conclusiones	15
8. Anexos - Repositorio Código fuente	16

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" "Frecuentement
## $ 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 ...
```

Con base en la estructura del dataframe acerca de la rotación laboral se muestra la información de 1470 empleados. Se detallan aspectos clave como la edad promedio de los empleados (36.92 años) y su experiencia laboral (11.28 años), junto con una diversidad de características demográficas y laborales como género, estado civil, departamento y cargo. Además, se exploran variables relacionadas con la satisfacción y el rendimiento laboral, así como factores potencialmente influyentes en la rotación, como el equilibrio entre trabajo y vida personal, horas extras y antigüedad en la empresa.

Teniendo en cuenta la información que contiene el dataframe, se seleccionan las siguientes variables

Variables categóricas:

Viaje de Negocios: Esta variable representa la frecuencia de los viajes de negocios de los empleados. Podría estar relacionada con la rotación, ya que los viajes constantes podrían afectar el equilibrio entre el trabajo y la vida personal de los empleados.

Departamento: El departamento en el que trabaja un empleado puede influir en su satisfacción laboral y en sus oportunidades de crecimiento. Esto podría estar relacionado con la rotación.

Horas Extra: La cantidad de horas extra trabajadas por los empleados podría estar relacionada con la rotación, ya que el exceso de trabajo puede llevar al agotamiento y a la insatisfacción laboral.

Variables cuantitativas:

Edad: La edad de los empleados podría estar relacionada con la rotación, ya que las personas en diferentes etapas de su vida pueden tener diferentes prioridades y necesidades laborales.

Ingreso Mensual: El salario de los empleados puede influir en su decisión de quedarse o irse de una empresa. Una remuneración más alta podría estar relacionada con una menor rotación.<7li>

Antigüedad en el Cargo: Cuánto tiempo ha estado un empleado en su puesto actual puede estar relacionado con su rotación. Los empleados con menos tiempo en el cargo podrían ser más propensos a rotar en busca de nuevas oportunidades.

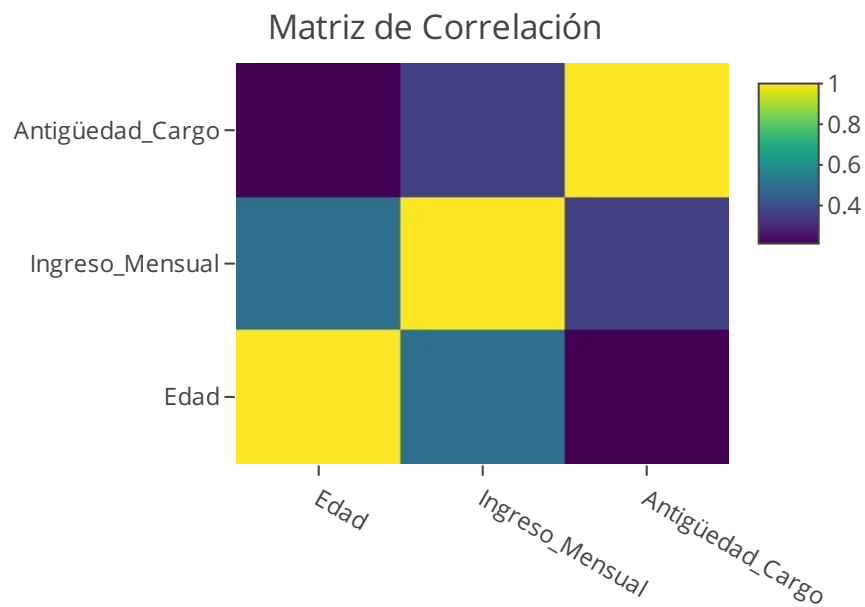
```
## tibble [1,470 x 6] (S3: tbl_df/tbl/data.frame)
## $ Viaje de Negocios: chr [1:1470] "Raramente" "Frecuentemente" "Raramente" "Frecuentemente" ...
## $ Departamento    : chr [1:1470] "Ventas" "IyD" "IyD" "IyD" ...
## $ Horas_Extra      : chr [1:1470] "Si" "No" "Si" "Si" ...
## $ Edad             : 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
```

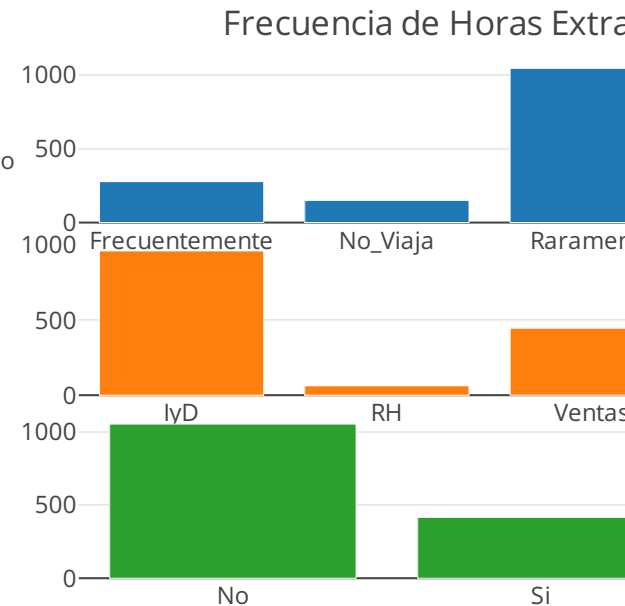
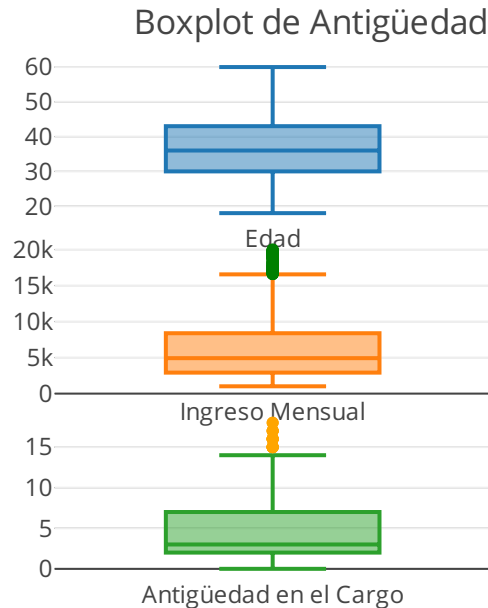
```
## # A tibble: 6 x 6
## 'Viaje de Negocios' Departamento Horas_Extra Edad Ingreso_Mensual
## <chr> <chr> <chr> <dbl> <dbl>
## 1 Raramente Ventas Si 41 5993
## 2 Frecuentemente IyD No 49 5130
## 3 Raramente IyD Si 37 2090
## 4 Frecuentemente IyD Si 33 2909
## 5 Raramente IyD No 27 3468
## 6 Frecuentemente 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 espera que exista una relación entre la frecuencia de los viajes de negocios y la rotación de empleados. La hipótesis es que los empleados que viajan frecuentemente por negocios tienen una mayor probabilidad de rotar que aquellos que viajan raramente.
2. **Departamento:** Se espera que exista una relación entre el departamento en el que trabaja un empleado y su rotación. La hipótesis es que ciertos departamentos pueden tener tasas de rotación más altas que otros debido a diferencias en la cultura organizacional, la carga de trabajo o las oportunidades de crecimiento.
3. **Horas Extra:** Se espera que exista una relación entre trabajar horas extra y la rotación de empleados. La hipótesis es que los empleados que trabajan horas extra tienen una mayor probabilidad de rotar que aquellos que no lo hacen.
4. **Edad:** Se espera que exista una relación entre la edad de los empleados y su rotación. La hipótesis es que los empleados más jóvenes pueden tener una mayor probabilidad de rotar debido a la búsqueda de oportunidades de crecimiento, mientras que los empleados mayores pueden tener una menor probabilidad de rotar debido a la estabilidad laboral y personal.
5. **Ingreso Mensual:** Se espera que exista una relación entre el nivel de ingresos de los empleados y su rotación. La hipótesis es que los empleados con salarios más bajos pueden tener una mayor probabilidad de rotar en busca de oportunidades de mejor remuneración, mientras que aquellos con salarios más altos pueden tener una menor probabilidad de rotar debido a la estabilidad financiera.
6. **Antigüedad en el Cargo:** Se espera que exista una relación entre la antigüedad en el cargo de los empleados y su rotación. La hipótesis es que los empleados con menos antigüedad en el cargo pueden tener una mayor probabilidad de rotar en busca de mejores oportunidades de carrera, mientras que aquellos con más antigüedad pueden tener una menor probabilidad de rotar debido a la estabilidad laboral y las 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

```

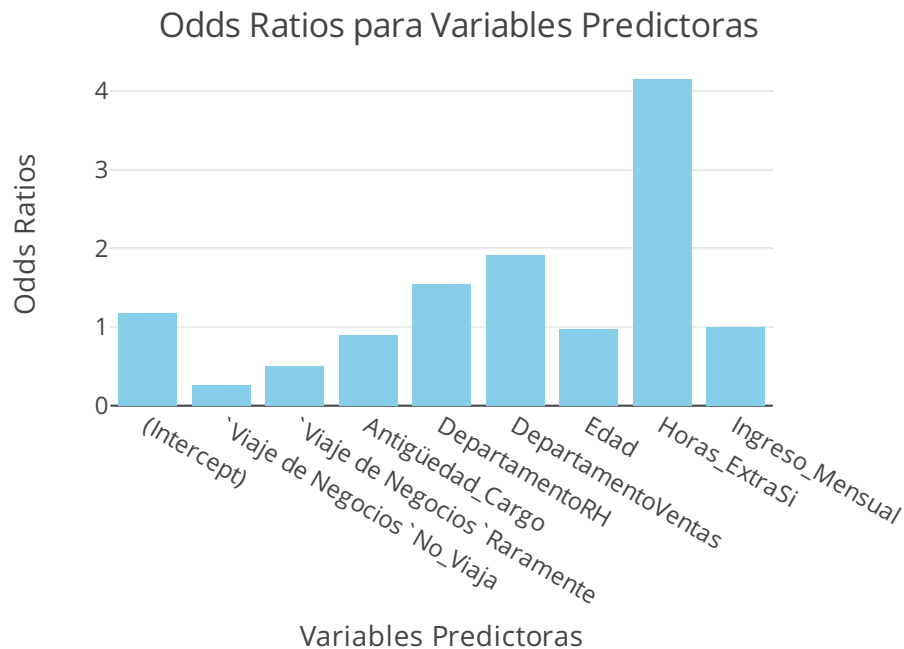
```

## 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 Coeficientes con Hipótesis:

- **Viaje de Negocios:**

- **Coeficiente:** Viaje de Negocios No_Viaja: -1.3389, Raramente: -0.6346
- **Hipótesis:** Se esperaba que los empleados que viajan frecuentemente por negocios tuvieran una mayor probabilidad de rotación que aquellos que viajan raramente.
- **Conclusión:** Consistente con la hipótesis, ya que los coeficientes son negativos, indicando que aquellos que viajan menos tienen menores probabilidades de rotación.

- **Departamento:**

- **Coeficiente:** Departamento RH: 0.38175, Ventas: 0.48116
- **Hipótesis:** Se esperaba que ciertos departamentos tuvieran tasas de rotación más altas que otros debido a diferencias en la cultura organizacional, la carga de trabajo o las oportunidades de crecimiento.
- **Conclusión:** Consistente con la hipótesis, ya que algunos departamentos tienen coeficientes positivos, indicando tasas de rotación más altas.

- **Horas Extra:**

- **Coeficiente:** Horas Extra Si: 1.3274
- **Hipótesis:** Se esperaba que los empleados que trabajan horas extra tuvieran una mayor probabilidad de rotación que aquellos que no lo hacen.
- **Conclusión:** Consistente con la hipótesis, ya que el coeficiente es positivo, indicando que aquellos que trabajan horas extra tienen mayores probabilidades de rotación.

- **Edad:**

- **Coeficiente:** Edad: -0.05225
- **Hipótesis:** Se esperaba que los empleados más jóvenes tuvieran una mayor probabilidad de rotación que los mayores.
- **Conclusión:** Inconsistente con la hipótesis, ya que el coeficiente es negativo pero pequeño, indicando una relación débil entre la edad y la rotación.

- **Ingreso Mensual:**

- **Coeficiente:** Ingreso Mensual: -0.0001271

- **Hipótesis:** Se esperaba que los empleados con salarios más bajos tuvieran una mayor probabilidad de rotación.
- **Conclusión:** Inconsistente con la hipótesis, ya que el coeficiente es negativo pero cercano a cero, indicando una relación débil entre el ingreso mensual y la rotación.
- **Antigüedad en el Cargo:**
 - **Coefficiente:** Antigüedad en el Cargo: -0.14628
 - **Hipótesis:** Se esperaba que los empleados con menos antigüedad en el cargo tuvieran una mayor probabilidad de rotación que los que tienen más antigüedad.
 - **Conclusión:** Inconsistente con la hipótesis, ya que el coeficiente es negativo pero pequeño, indicando una relación débil entre la antigüedad en el cargo y la rotación.

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

La información de los Odds Ratios para las variables predictoras añade valor al informe al proporcionar una medida de la asociación entre cada variable y la probabilidad de rotación de empleados. Los Odds Ratios muestran cuánto aumenta o disminuye la probabilidad de rotación cuando una variable aumenta en una unidad, manteniendo todas las demás variables constantes.

Horas Extra (Horas_ExtraSi): Un Odds Ratio de 4.155 indica que los empleados que trabajan horas extra tienen aproximadamente 4.155 veces más probabilidades de rotar que aquellos que no lo hacen, manteniendo todas las demás variables constantes.

Departamento (DepartamentoVentas y DepartamentoRH): Los Odds Ratios de 1.914 y 1.547 respectivamente sugieren que pertenecer a los departamentos de Ventas o Recursos Humanos aumenta la probabilidad de rotación en aproximadamente 1.914 y 1.547 veces en comparación con otros departamentos, manteniendo todas las demás variables constantes.

Ingreso Mensual (Ingreso_Mensual): Un Odds Ratio de 0.9999181 sugiere que por cada aumento unitario en el ingreso mensual, la probabilidad de rotación disminuye en un 0.0000819%, manteniendo todas las demás variables constantes. Esto indica una asociación muy débil entre el ingreso mensual y la rotación.

Edad (Edad) y Antigüedad en el Cargo (Antigüedad_Cargo): Los Odds Ratios de 0.968 y 0.894 respectivamente indican que por cada aumento unitario en la edad o la antigüedad en el cargo, la probabilidad de rotación disminuye en un 3.2% y un 10.6% respectivamente, manteniendo todas las demás variables constantes.

Viaje de Negocios (Viaje de Negocios Raramente y No Viaja): Los Odds Ratios de 0.5067 y 0.2582 respectivamente sugieren que los empleados que viajan raramente o no viajan tienen una menor probabilidad de rotación en comparación con aquellos que viajan frecuentemente por negocios, manteniendo todas las demás variables constantes.

En resumen, los Odds Ratios proporcionan información adicional sobre la fuerza y la dirección de la asociación entre cada variable y la probabilidad de rotación de empleados, lo que ayuda a comprender mejor la influencia relativa de cada variable en el fenómeno de la rotación.

Discusión:

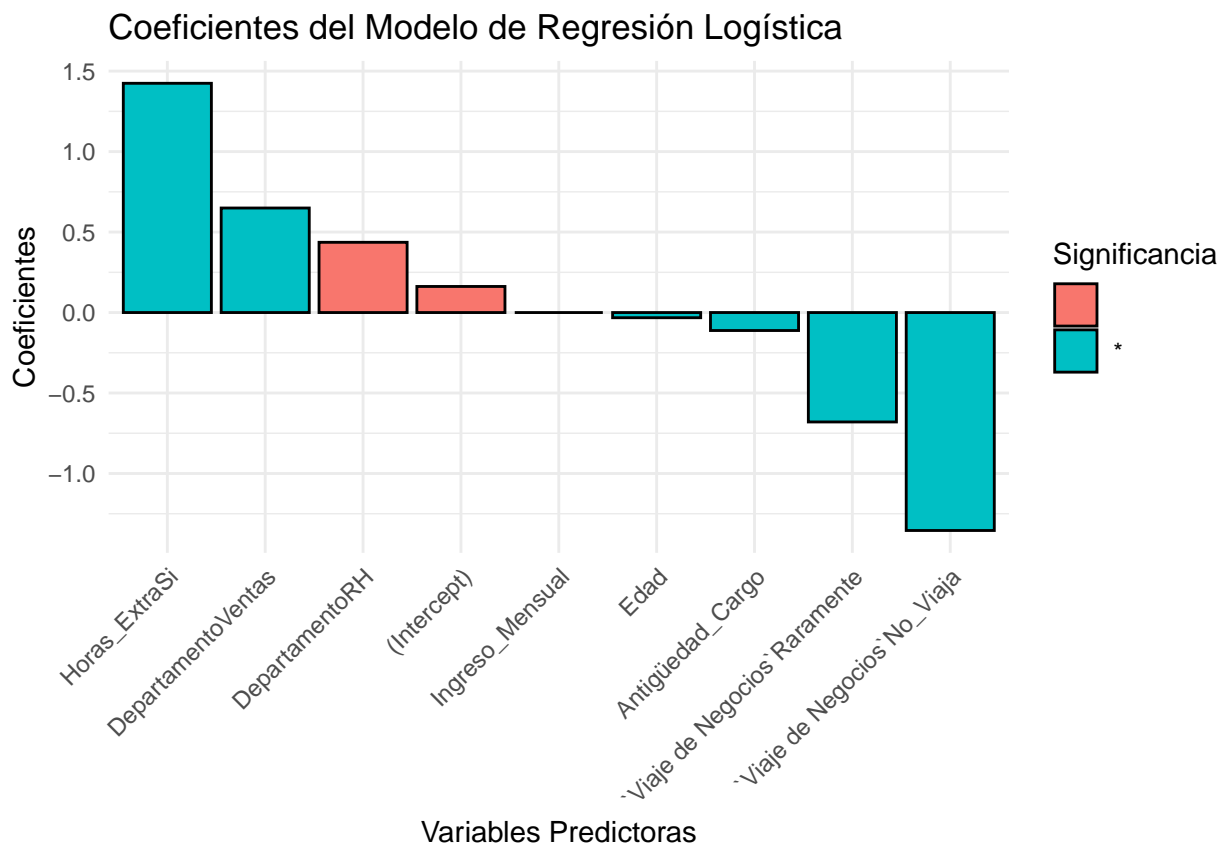
Al interpretar los coeficientes de regresión logística, nos centramos en la dirección y la magnitud del efecto de cada variable en la probabilidad de rotación. Sin embargo, al considerar los Odds Ratios, obtenemos una perspectiva más detallada de esta relación.

Algunas discrepancias entre la interpretación de los coeficientes y los Odds Ratios pueden surgir debido a la transformación logarítmica aplicada en el modelo de regresión logística para calcular los coeficientes. Esta transformación no siempre refleja directamente la asociación entre las variables y la probabilidad de rotación.

Por ejemplo, en el caso del ingreso mensual, el coeficiente estimado podría indicar una relación débil con la rotación, pero el Odds Ratio sugiere que incluso pequeños aumentos en el ingreso mensual pueden estar asociados con reducciones significativas en la probabilidad de rotación. Esta discrepancia resalta la importancia de considerar tanto los coeficientes como los Odds Ratios para obtener una comprensión completa de la relación entre las variables predictoras y la variable de respuesta. Mientras que los coeficientes de regresión logística nos proporcionan información sobre la dirección y la magnitud de la relación entre las variables predictoras y la probabilidad de rotación, los Odds Ratios nos ofrecen una perspectiva más intuitiva de cómo estos cambios afectan la probabilidad de rotación en términos de probabilidades relativas.

4. Estimación del modelo

Realiza la estimación de un modelo de regresión logística en el cual la variable respuesta es rotación ($y=1$ es si rotación, $y=0$ es no rotación) y las covariables 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	

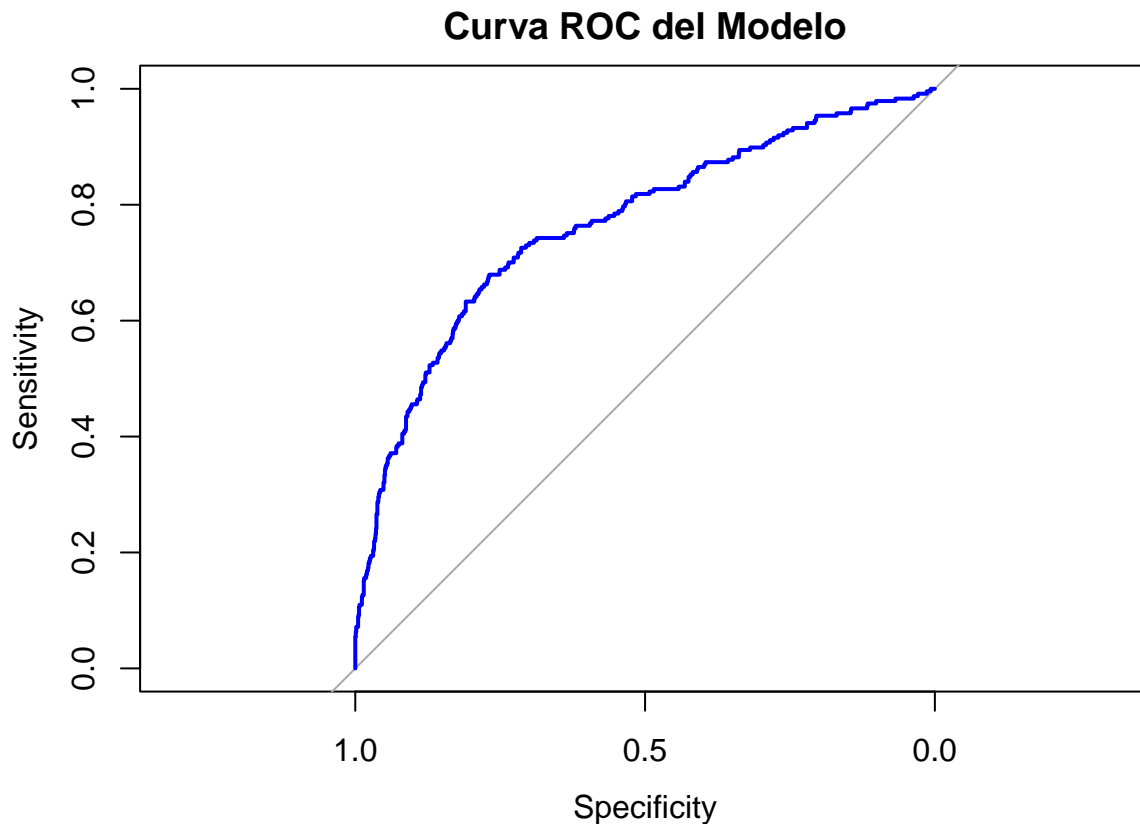
	Variablen	Coefficientes	Significancia
Viaje de NegociosNo_Viaja	Viaje de NegociosNo_Viaja	-1.3541771	*
Viaje de NegociosRaramente	Viaje de NegociosRaramente	-0.6797980	*
DepartamentoRH	DepartamentoRH	0.4363905	
DepartamentoVentas	DepartamentoVentas	0.6494927	*
Horas_ExtraSi	Horas_ExtraSi	1.4243441	*
Edad	Edad	-0.0321813	*
Ingreso_Mensual	Ingreso_Mensual	-0.0000819	*
Antigüedad_Cargo	Antigüedad_Cargo	-0.1117549	*

A continuación se procede con la interpretación de los coeficientes:

1. **Intercept:** Este es el intercepto del modelo. Representa el logaritmo de la razón de probabilidades de rotación cuando todas las variables predictoras son cero. En este caso, es 0.1619797, lo que significa que cuando todas las demás variables son cero, el log-odds de rotación es de aproximadamente 0.16.
2. **Viaje de NegociosNo_Viaja:** Un coeficiente negativo (-1.3541771) indica que los empleados que no viajan por negocios tienen menores log-odds de rotación en comparación con aquellos que viajan por negocios, manteniendo todas las demás variables constantes.
3. **Viaje de NegociosRaramente:** Un coeficiente negativo (-0.6797980) indica que los empleados que viajan raramente por negocios tienen menores log-odds de rotación en comparación con aquellos que viajan frecuentemente por negocios, manteniendo todas las demás variables constantes.
4. **DepartamentoRH y DepartamentoVentas:** Los coeficientes positivos (0.4363905 y 0.6494927 respectivamente) indican que pertenecer al departamento de Recursos Humanos o al departamento de Ventas está asociado con mayores log-odds de rotación, en comparación con el departamento de referencia (probablemente otro departamento no mencionado).
5. **Horas_ExtraSi:** Un coeficiente positivo (1.4243441) indica que trabajar horas extra está asociado con mayores log-odds de rotación.
6. **Edad y Antigüedad_Cargo:** Los coeficientes negativos (-0.0321813 y -0.1117549 respectivamente) indican que el aumento en la edad y en la antigüedad en el cargo están asociados con menores log-odds de rotación.
7. **Ingreso_Mensual:** Un coeficiente negativo (-0.0000819) indica que un aumento en el ingreso mensual está asociado con menores log-odds de rotación, aunque el efecto es muy pequeño y casi despreciable.

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 muestra la probabilidad de detección verdadera (TPR) en función de la probabilidad de falsa alarma (FPR) para diferentes umbrales de clasificación. La curva ROC se traza en un espacio bidimensional, con la FPR en el eje X y la TPR en el eje Y. Una curva ROC ideal es la que se encuentra en la esquina superior izquierda del gráfico, lo que indica que el modelo es capaz de detectar todos los casos positivos sin generar ninguna falsa alarma.

Área bajo la curva (AUC):

El AUC es la proporción de la superficie total del gráfico ROC que se encuentra debajo de la curva. El AUC se expresa como un valor numérico entre 0 y 1. Un AUC de 1 indica que el modelo es perfecto, mientras que un AUC de 0 indica que el modelo no es mejor que una adivinanza aleatoria.

Finalmente, en la imagen que se muestra, la curva ROC se encuentra por encima de la diagonal aleatoria. Esto indica que el modelo es mejor que una adivinanza aleatoria para clasificar los casos. El AUC de la curva ROC es de 0.7654207. Este valor indica que el modelo tiene un buen rendimiento para clasificar los casos. Se puede concluir que el modelo tiene un buen rendimiento para clasificar los casos. La curva ROC se encuentra por encima de la diagonal aleatoria y el AUC es de 0.7654207, lo que indica que el modelo es capaz de detectar los verdaderos positivos con una alta probabilidad y generar pocas 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, una empleada dedicada y comprometida con su trabajo en una empresa de ventas. Lucía ha decidido establecerse en un puesto que no requiere viajes de negocios, ya que prefiere pasar tiempo en casa con su familia. A pesar de su dedicación, últimamente ha tenido que trabajar horas extras para cumplir con las demandas del trabajo. A sus 35 años, Lucía se encuentra en una etapa estable de su vida laboral, con una antigüedad en su cargo de 5 años. Con 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 este caso, teniendo en cuenta la información procesada por el modelo, se determina que no es necesario la intervención a la colaboradora Lucía, ya que, Lucía se siente cómoda y valorada en su posición actual. .

7. Conclusiones

En las conclusiones adicione una discusión sobre cuál sería la estrategia para disminuir la rotación en la empresa (con base en las variables que resultaron significativas en el punto 3).

Teniendo cuenta las variables identificadas en el análisis bivariado del punto 3. se presentan a continuación el diseño de 4 estrategias basadas en: Viaje de Negocios, Departamento y Horas Extra.

Mejorar el ambiente laboral y la conciliación: Se busca reducir la necesidad de viajar frecuentemente por negocios, especialmente en departamentos con altas tasas de rotación, mediante la implementación de políticas que fomenten un equilibrio adecuado entre el trabajo y la vida personal. Esto puede incluir la promoción de flexibilidad laboral, programas de trabajo remoto y beneficios para empleados que favorezcan la conciliación.

Optimizar la distribución de tareas y responsabilidades: Se propone revisar la distribución de carga laboral en departamentos con altas tasas de rotación, como Ventas, para minimizar la necesidad de trabajar horas extras de manera constante. Esto implica una evaluación de la carga de trabajo, la asignación de recursos adecuados y la implementación de sistemas de rotación de turnos que reduzcan la fatiga laboral.

Incentivar el desarrollo profesional y el crecimiento interno: Para contrarrestar la rotación en departamentos específicos, se sugiere implementar programas de desarrollo profesional y oportunidades de crecimiento interno. Estos programas pueden incluir capacitación continua, mentoría y planes de carrera personalizados para motivar a los empleados a permanecer en la organización y avanzar en sus roles.

Revisar y ajustar la política de compensación: Se propone evaluar la estructura de remuneración para garantizar que sea competitiva y equitativa, especialmente en departamentos con altas tasas de rotación. Esto puede incluir ajustes salariales, bonificaciones por desempeño y beneficios adicionales que reconozcan y recompensen el esfuerzo y la dedicación de los empleados.

Estas estrategias buscan abordar las preocupaciones específicas relacionadas con las variables identificadas en el análisis bivariado, con el objetivo de reducir la rotación de empleados y promover un entorno laboral más estable y satisfactorio.

8. Anexos - Repositorio Código fuente

Repositorio de Github