

# Parte de R

Kevin David Calderon Martinez. Link del repositorio:  
[https://github.com/QuebimXD/Examen\\_II\\_G1\\_C4D511-KevinCalderonMartinez.git](https://github.com/QuebimXD/Examen_II_G1_C4D511-KevinCalderonMartinez.git)

Primero, leemos la base de datos.

## Ejercicio 2

Tabla 1: Resumen de 5 numeros y promedio.

MEDIDOR	SALARIO	EDAD
MINIMO	10.0	14.00000
MAXIMO	16858742.0	99.00000
CUARTIL_1	297429.0	31.00000
MEDIANA	408495.0	40.00000
CUARTIL_3	700000.0	52.00000
PROMEDIO	657534.3	41.90634

## Ejercicio 3

Tabla 2: Aplicacion del Z-score.

Z_SALARIO	OUTLIER_SALARIO	Z_EDAD	OUTLIER_EDAD
NA	NA	0.0066098	NO
3.2767840	OUTLIER	2.4767464	OUTLIER
-0.4772468	NO	0.2183358	NO
-0.1388214	NO	0.5006371	NO

Z_SALARIO	OUTLIER_SALARIO	Z_EDAD	OUTLIER_EDAD
-0.5518286	NO	0.0066098	NO
-0.3210648	NO	0.1477605	NO
0.2431284	NO	-0.2756915	NO
-0.3163175	NO	-0.4168422	NO
-0.4843763	NO	-1.4754721	NO
4.9563262	OUTLIER	-0.1345408	NO

## Ejercicio 4

Tabla 3: Porcentaje de valores nulos en cada columna.

	COLUMNA	PORCENTAJE
IDENTIFICACION	IDENTIFICACION	0.00
SEXO	SEXO	0.00
SALARIO	SALARIO	0.06
TIPO_ASEGURAMIENTO	TIPO_ASEGURAMIENTO	0.00
SECTOR	SECTOR	0.00
INDICADOR_ACTIVO	INDICADOR_ACTIVO	0.00
INDICADOR_EXTRANJERO	INDICADOR_EXTRANJERO	0.00
INDICADOR_MOROSO	INDICADOR_MOROSO	0.00
EDAD	EDAD	0.06
Z_SALARIO	Z_SALARIO	0.06
OUTLIER_SALARIO	OUTLIER_SALARIO	0.06
Z_EDAD	Z_EDAD	0.06
OUTLIER_EDAD	OUTLIER_EDAD	0.06

## Ejercicio 5

Tabla 4: Datos imputados por las columnas vacias.

SALARIO_IMPUT	EDAD_IMPUT	TIPO_ASEGURAMIENTO_IMPUT
657534.3	42	ASALARIADO
3315014.0	77	PENSIONADOS EN SU PROPIO SISTEMA
270486.0	45	ASALARIADO
544949.8	49	ASALARIADO
210000.0	42	ASALARIADO
397150.0	44	ASALARIADO
854712.0	38	ASALARIADO
401000.0	36	ASALARIADO
264704.0	21	ASALARIADO
4677126.8	40	ASALARIADO

## Ejercicio 6

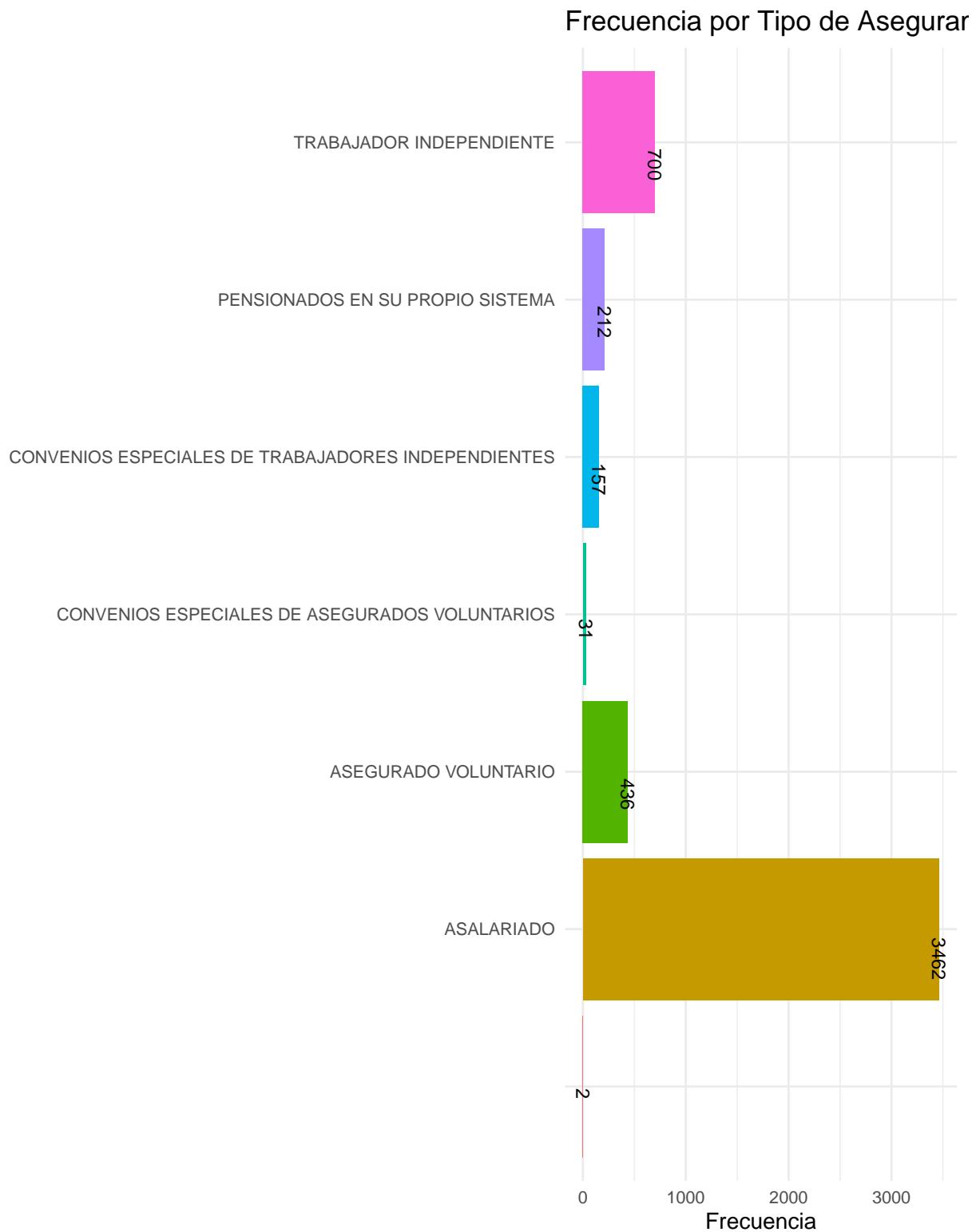


Figura 1: Grafico de la distribucion de asegurados de acuerdo al tipo

Como se puede observar, casi el 80% de los asegurados son asalariados. Por lo que la mayoria se encuentra en situacion de empleado. El de menor frecuencia tiene que ver con el convenio de los asegurados voluntarios. Tambien se aprecia que hay mas trabajadores independientes asegurados, que los asegurados voluntarios.

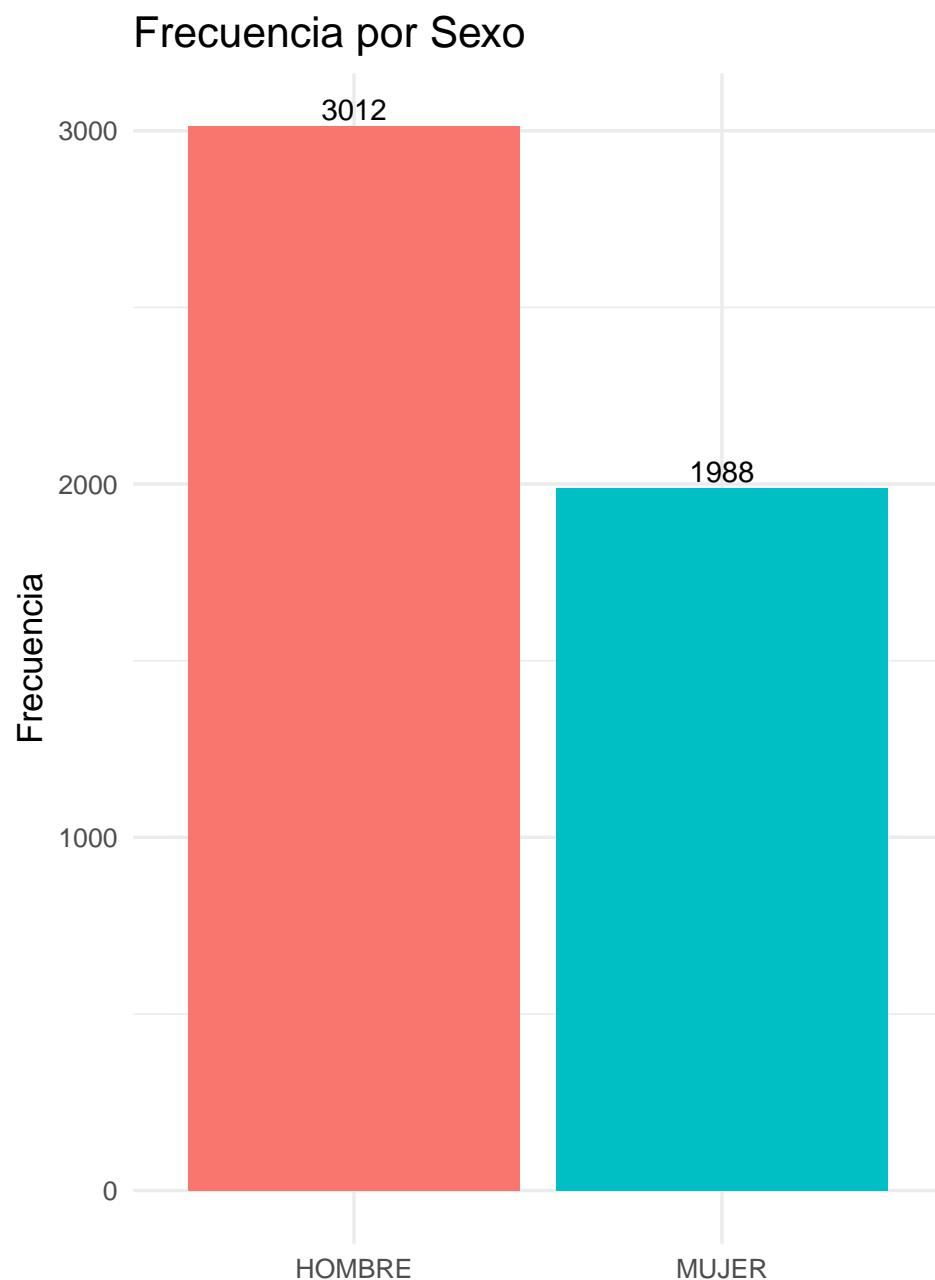


Figura 2: Grafico de la distribucion de asegurados de acuerdo al sexo

Note que poco mas del 60% de los asegurados son hombres, lo cual contrasta en una diferencia de mas de un 20% con las mujeres aseguradas.

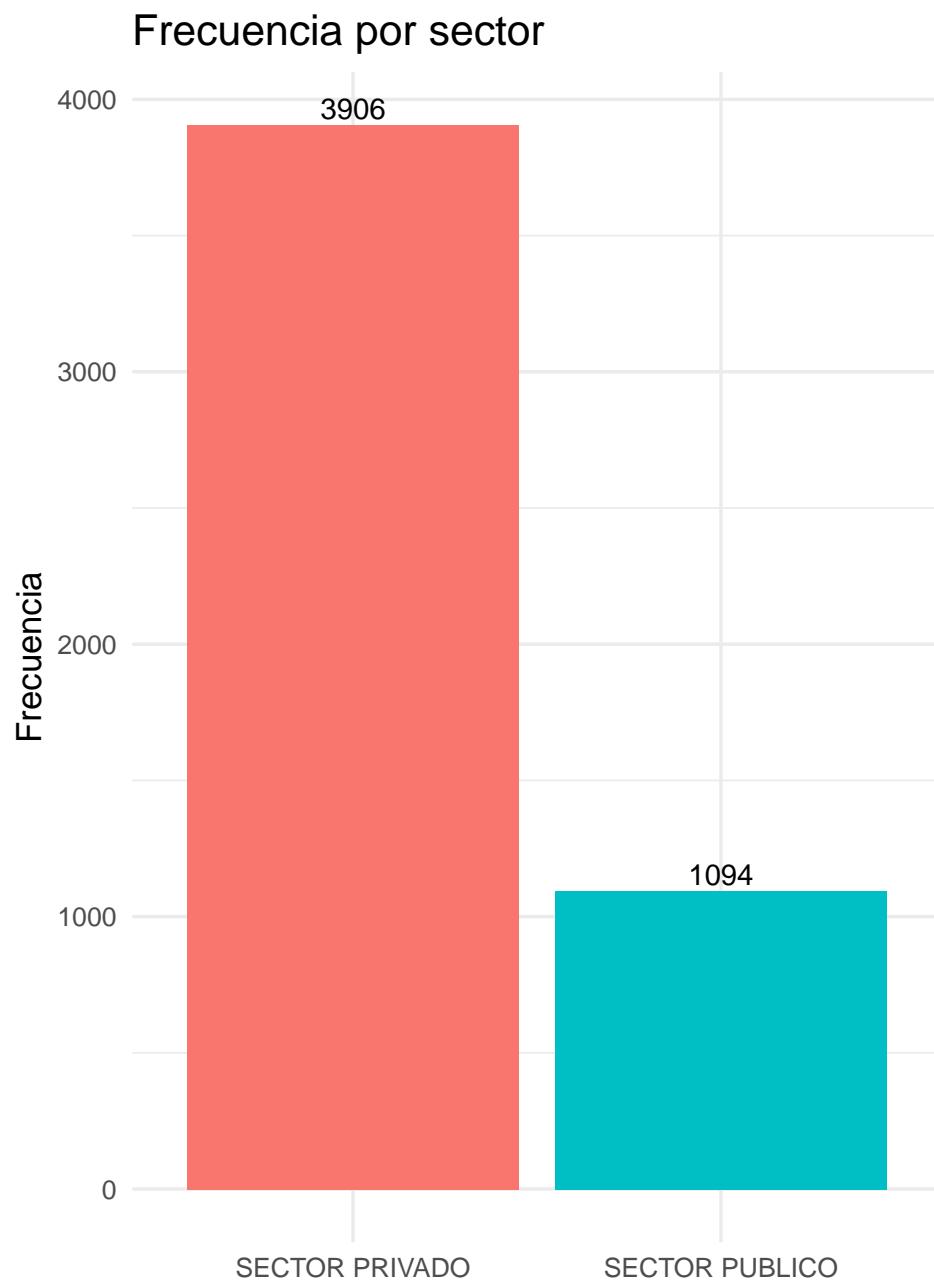


Figura 3: Grafico de la distribucion de asegurados de acuerdo al sector del asegurado

Poco mas de un 20% de empleados son del sector publico, por lo que hay una gran brecha entre la cantidad de empleados del sector publico y privado.

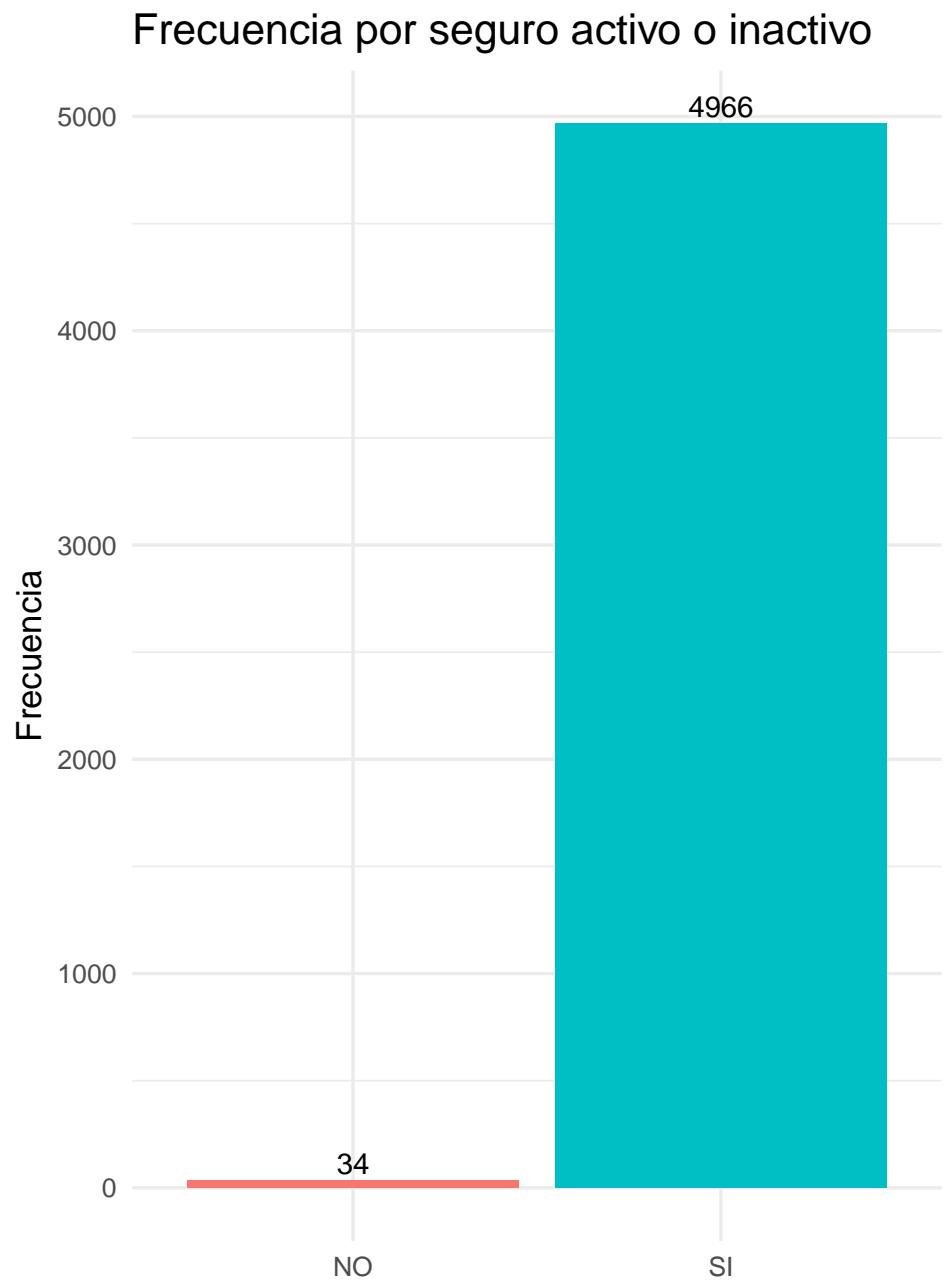


Figura 4: Grafico de la distribucion de asegurados de acuerdo a si el seguro se encuentra activo o no

Aquí casi la totalidad de los asegurados poseen su paquete en estado Activo. Solo 34 personas no tienen su seguro activo. Esto es positivo, debido a que hay poco margen de asegurados sin forma de acceder a su seguro, pues se encuentra inactivo.

### Frecuencia por extranjero

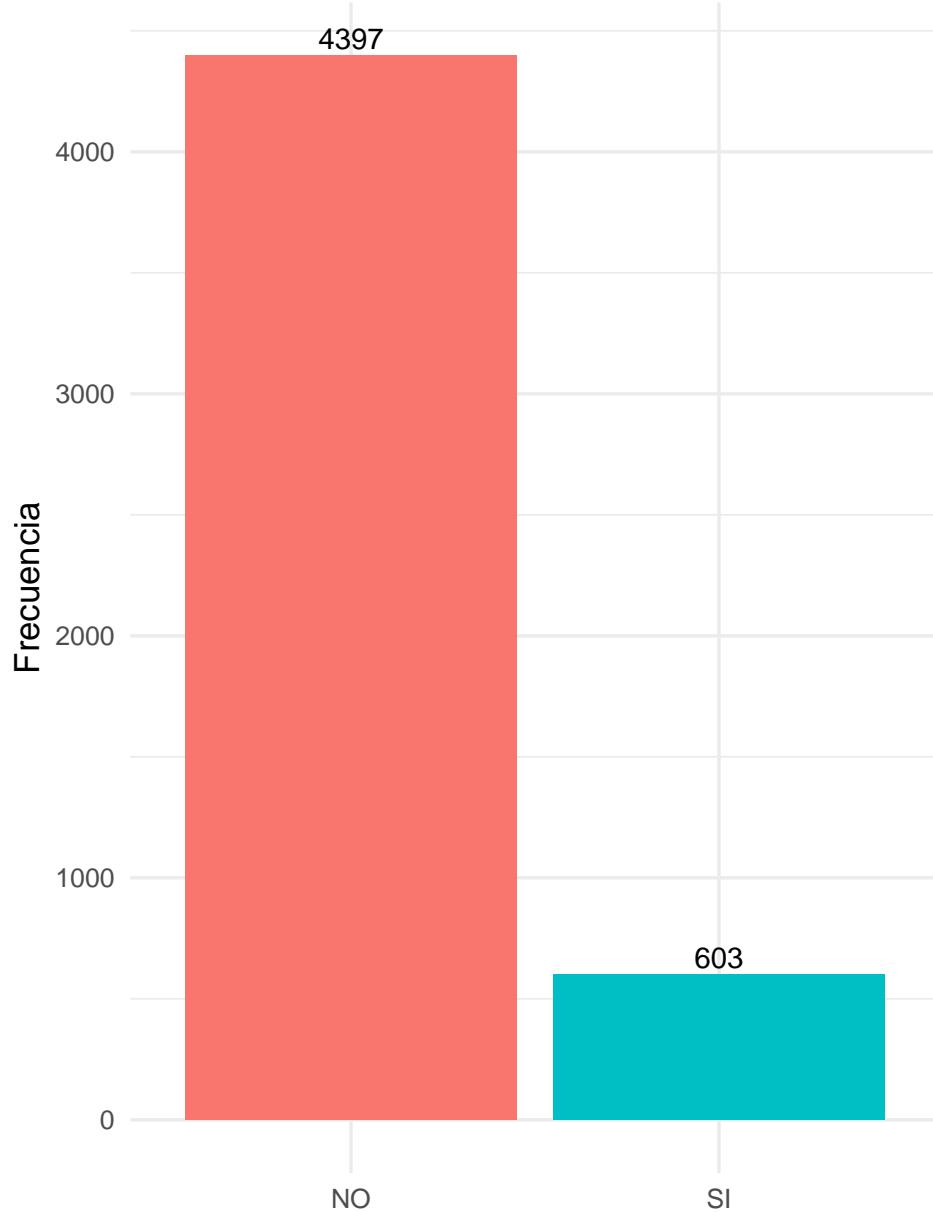


Figura 5: Grafico de la distribucion de asegurados de acuerdo a si el usuario del seguro es extranjero o no

Hay una baja cantidad de asegurados extranjeros. Esto se puede explicar de muchas formas: El seguro esta mas enfocado a servicios de residentes, y no tan enfocados al turismo. Por ende, hay una gran diferencia entre los asegurados que son y no son extranjeros.

## Frecuencia por morosidad

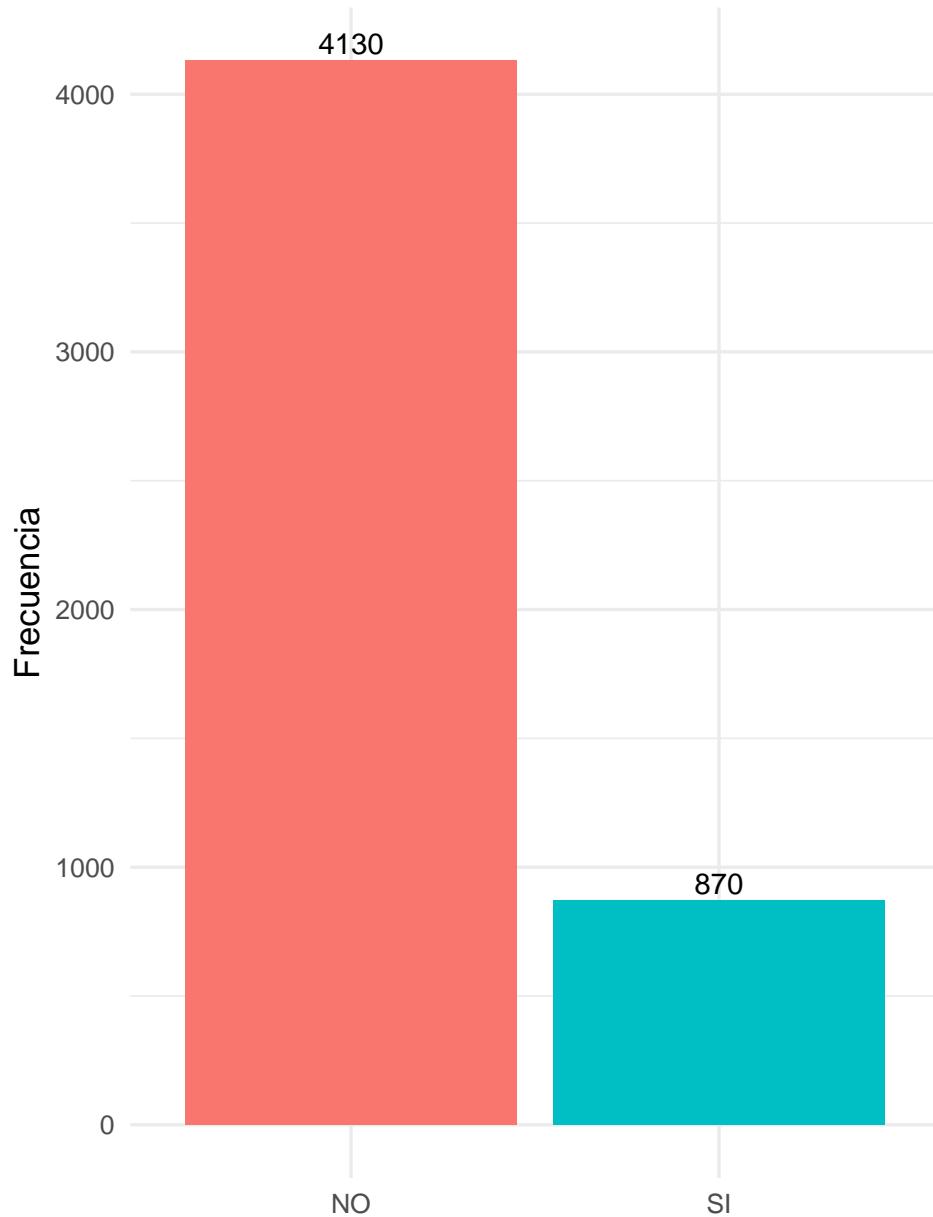


Figura 6: Grafico de la distribucion de asegurados de acuerdo a si el usuario del seguro es moroso o no

Mas del 80% de los asegurados han realizado los pagos de sus obligaciones de manera responsable. Hay 870 de asegurados que han impagado sus seguros. Si bien esta cifra es considerable, no comprende mas del 20% del total de asegurados, lo cual es un dato alentador para el factor incertidumbre de la aseguradora.

## Ejercicio 7

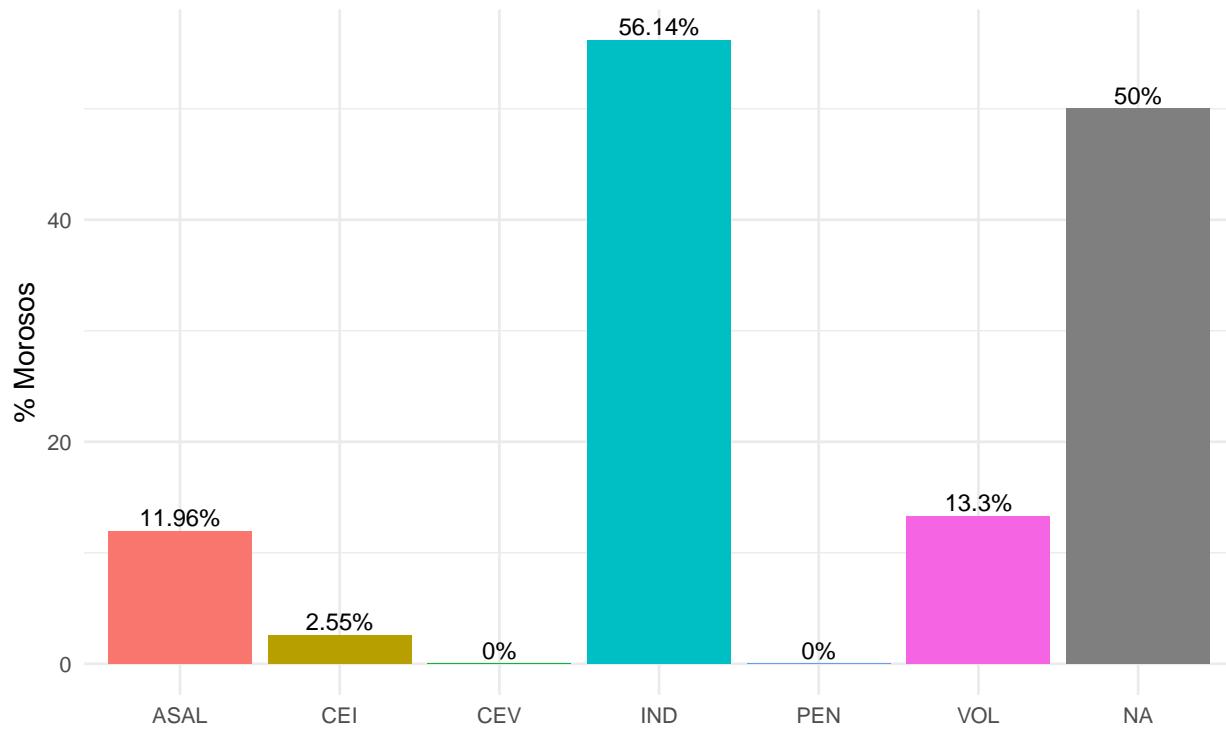


Figura 7: Grafico de la porcentaje de asegurados morosos de acuerdo al tipo de seguro

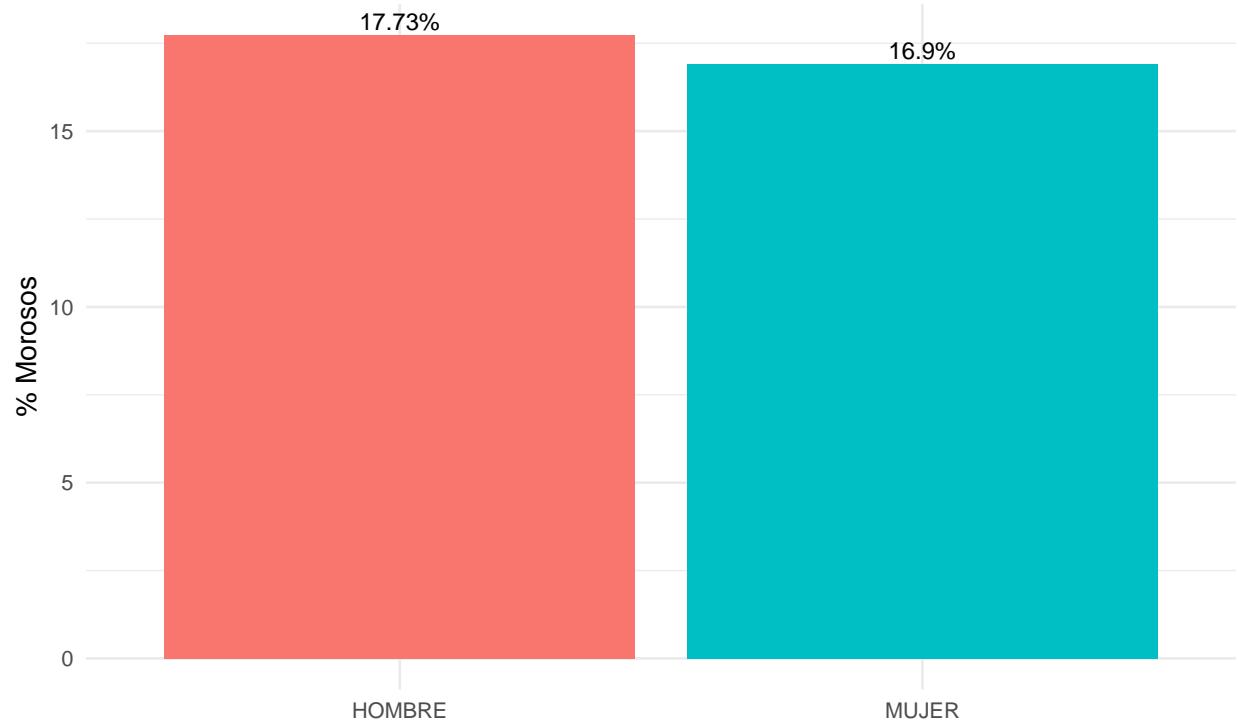


Figura 8: Grafico de la porcentaje de asegurados morosos de acuerdo al sexo

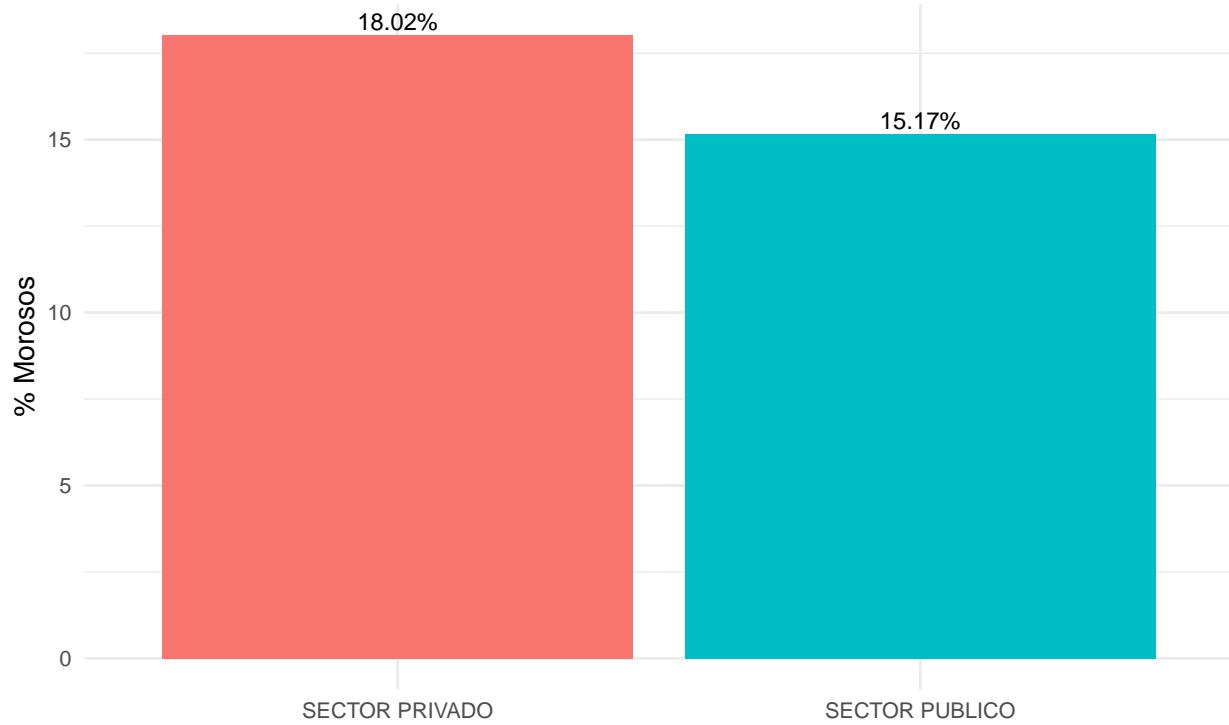


Figura 9: Grafico de la porcentaje de asegurados morosos de acuerdo al sector

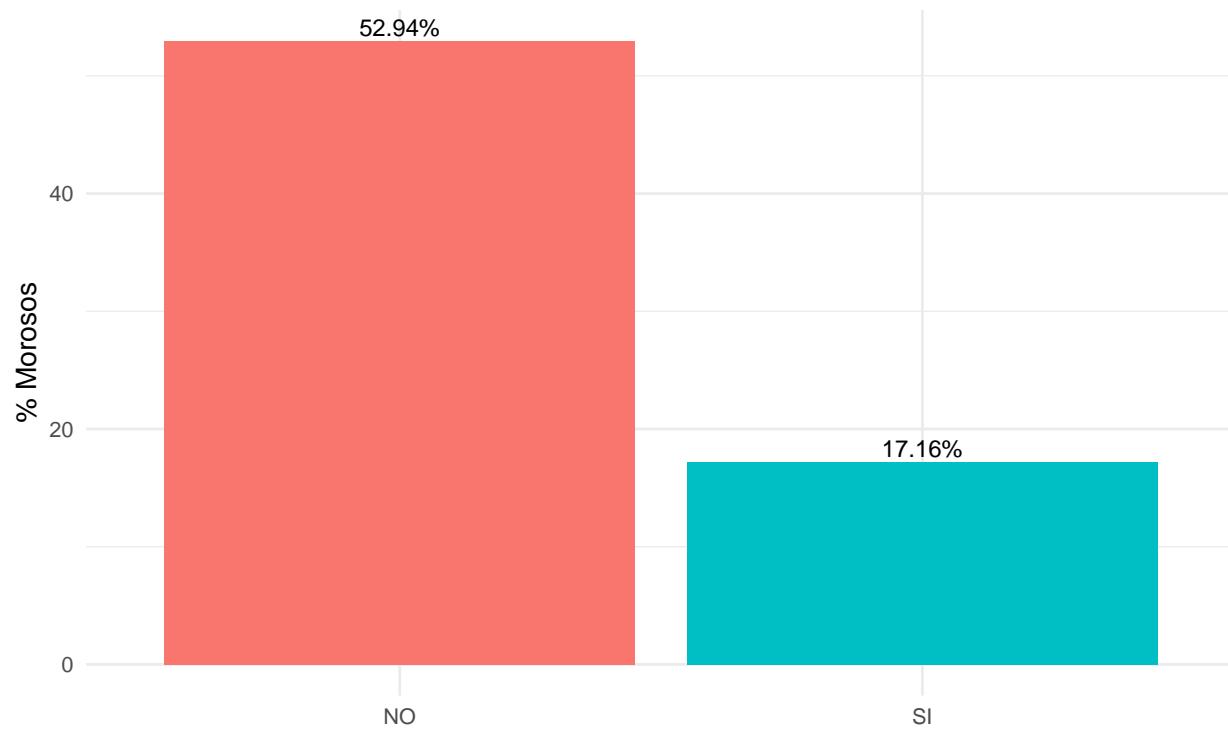


Figura 10: Grafico de la porcentaje de asegurados morosos de acuerdo a si el seguro esta activo o no

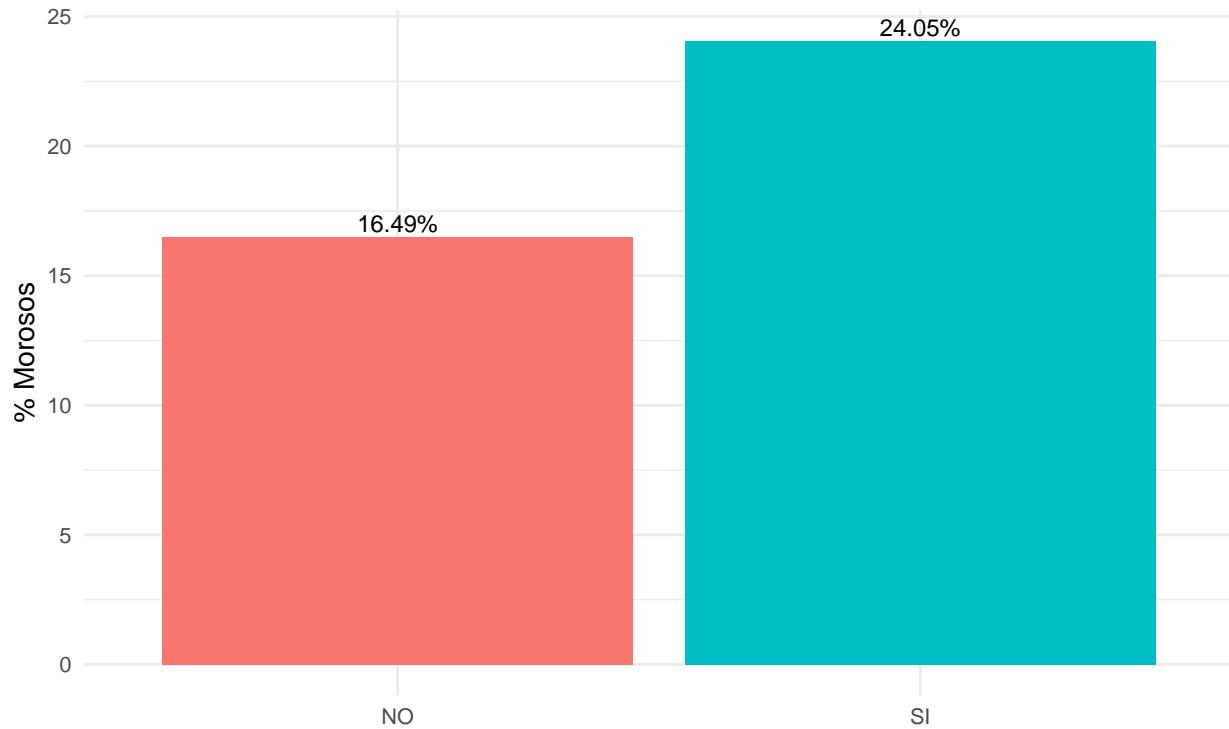


Figura 11: Grafico de la porcentaje de asegurados morosos de acuerdo a si el cliente es extranjero o no

Como puede observar, en los graficos se muestran las distribuciones de cada categoría de acuerdo al porcentaje de morosidad. Los datos a destacar es el porcentaje de morosos de trabajadores independientes, extranjeros, seguros inactivos; son porcentajes altos. Se puede entender en que mas de la mitad de seguros impagados sean inactivos, pero se debe verificar las obligaciones de los clientes morosos extranjeros y los morosos masculinos. Los demás porcentajes no superan el 20%, lo cual es una cifra alentadora.

## Ejercicio 8

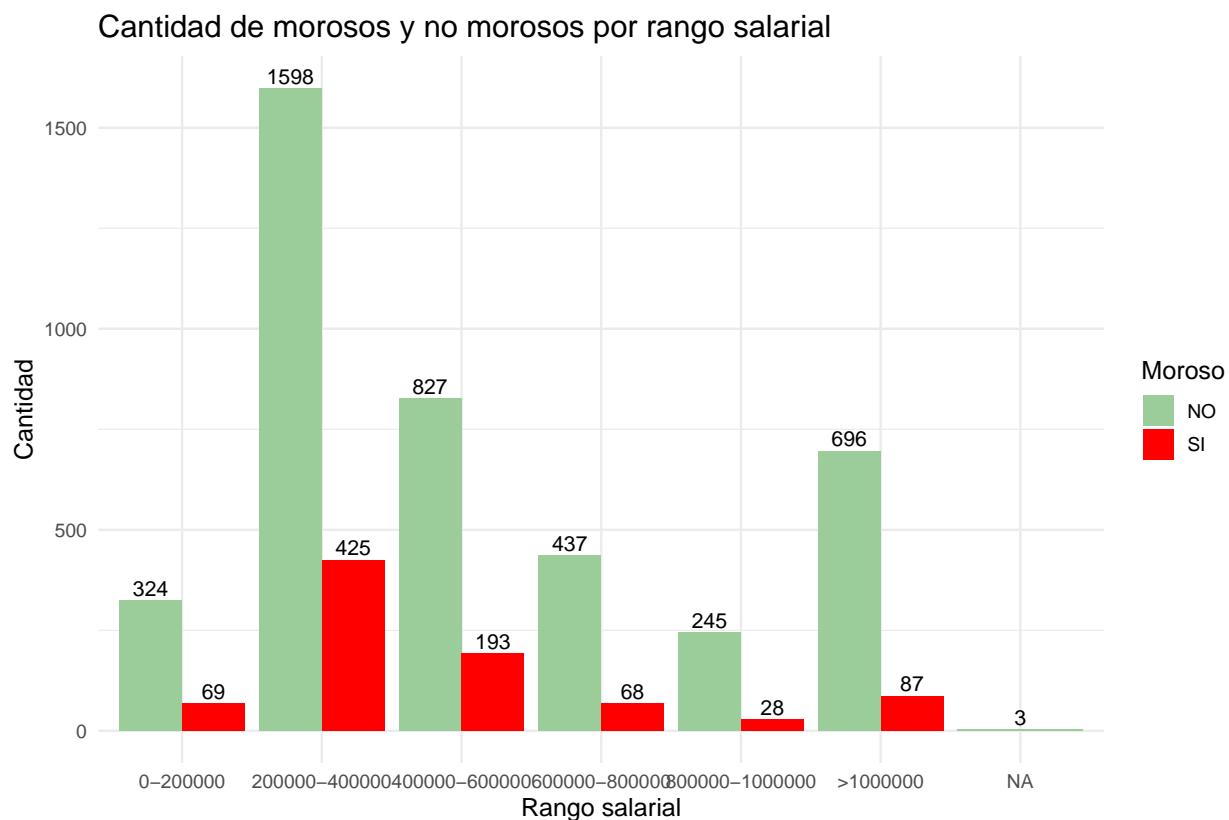


Figura 12: Grafico de cantidad de morosos de acuerdo al salario

Los resultados muestran que la morosidad se concentra principalmente en los niveles salariales intermedios, especialmente entre c200 000 y c600 000, donde se registran los valores más altos de personas morosas. Esto sugiere que quienes reciben ingresos medios pueden enfrentarse a mayores compromisos financieros o a una relación ingreso-gasto más ajustada, lo que incrementa su probabilidad de caer en mora.

Cantidad de morosos y no morosos por rango de edad

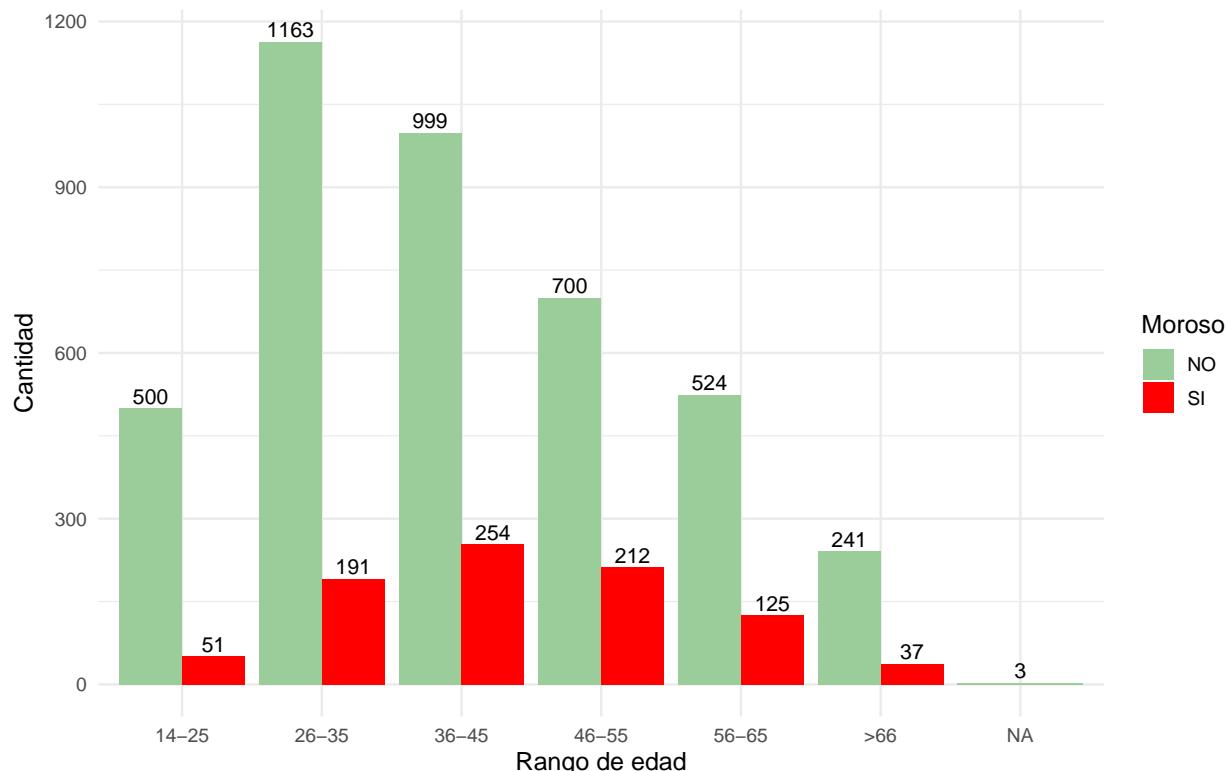
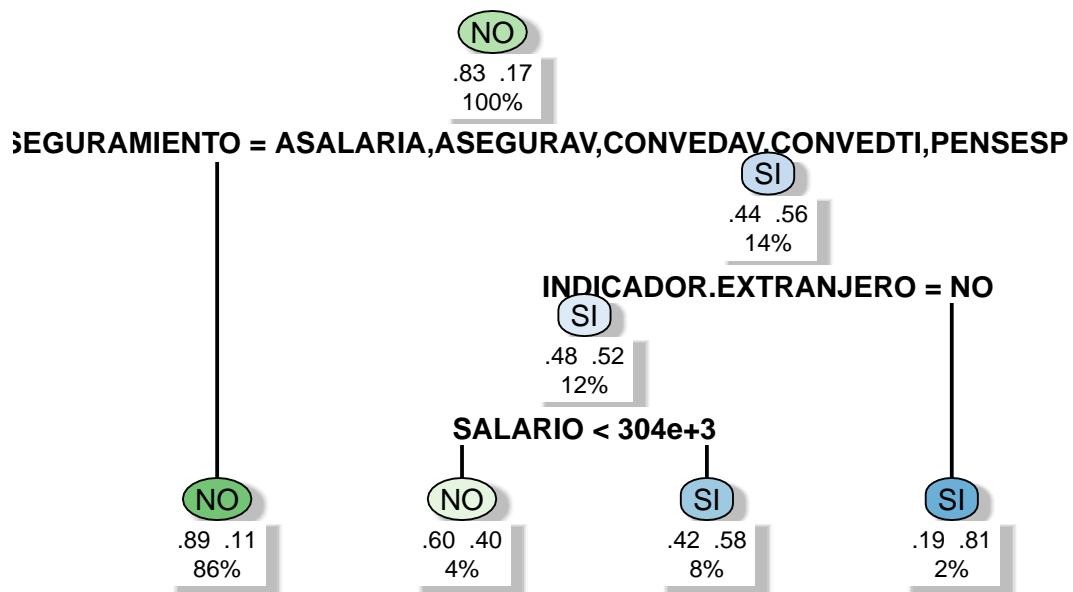


Figura 13: Grafico de cantidad de morosos de acuerdo a la edad

Los resultados muestran que la morosidad es más frecuente en las edades intermedias, especialmente entre los 36 y 55 años. Esto puede deberse a que en esa etapa las personas suelen asumir más compromisos económicos, como préstamos o gastos familiares, lo que aumenta la probabilidad de atrasarse. En cambio, los más jóvenes y los adultos mayores presentan menos casos de morosidad, posiblemente porque tienen menos obligaciones financieras o ingresos más estables.

## Ejercicio 10



El arbol nos afirma que la variable con mayor influencia en la morosidad es el tipo de aseguramiento, pues las categorías tradicionales (asalariado, voluntarios y pensionados) no son tan morosos. Entre quienes no pertenecen a ese grupo, el indicador de extranjero se hace notar como el factor más determinante: las personas extranjeras presentan el mayor nivel de morosidad dentro de toda la base. Finalmente, para los no extranjeros, el salario introduce una división interesante, donde los ingresos mayores a ₡304 000 se asocian con un riesgo más alto de morosidad, posiblemente por mayores compromisos financieros. En conjunto, el modelo sugiere que la morosidad está vinculada principalmente al tipo de aseguramiento, la condición de extranjero y el nivel salarial. Concluya que las características laborales y migratorias junto con el nivel salarial identifica perfiles con mayor chance de morosidad, brindando así una perspectiva mas clara para la toma de decisiones.