

## EMPRESA TELECOM X

### Informe de Análisis de Baja de Clientes - TelecomX

#### 1. Introducción

El objetivo de este análisis es identificar patrones de clientes que se dan de baja y aquellos que permanecen, de acuerdo a la información obtenida por parte del cliente, se reporta una alza en la cancelación de servicios, se solicita el acceso a la base de datos del cliente para poder verificar la información, entender su situación y poder ofrecer estrategias de retención.

#### 2. Limpieza y Tratamiento de Datos

Los datos obtenidos por parte del cliente se encuentran en el repositorio de GitHub ( [https://github.com/ingridcrith/challenge2-data-science-LATAM/blob/main/TelecomX\\_Data.json](https://github.com/ingridcrith/challenge2-data-science-LATAM/blob/main/TelecomX_Data.json))

Los datos se encuentran en un archivo .JSON, los cuales se importaron mediante la librería pandas y se establecieron como un DataFrame denominado datos, estos datos de inicio nos entregaron 6 columnas, de las cuales 4 tenían columnas con diccionarios/JSON anidados, los cuales se procedieron a transformar en columnas individuales, se contatenaron ( unieron) al dataframe datos para trabajarlo todo junto.

verificamos que no existan valores vacíos, eliminamos símbolos en las columnas pertinentes, cambiamos el nombre de churn a Baja para facilitar el análisis, eliminamos las columnas Cliente ID, y Total, ya que estos no generan valor para este análisis.

Se inicia el análisis exploratorio para determinar % de retención, y encontramos que hay datos que no sabemos si el cliente se encuentra o no activo, pero representa menos del 5% de los datos, por lo cual procedemos a segmentar solamente y trabajar con clientes activos y no activos, siendo ahora este nuestro 100% del universo. Verificamos que los datos tuvieran la correlación lógica correspondiente, esto para la parte de servicio telefónico, si no tenían contratado, que coincidiera con los valores de no o no servicio telefónico, de igual manera con las correlaciones del servicio de internet.

Separamos los datos de manera categórica y numérica para proceder con el análisis. Transformando los datos de género: Male (1)/Female (0) , yes(1)/No(0) y creamos variables tipo dummy, separándolas por variable para tratarlas de manera booleana, y después de manera entera True = 1 / False =0.

#### 3. Análisis exploratorio de datos

De los datos suministrados por el cliente obtuvimos un total de 7267 registros

distribuidos de la siguiente manera:

- Total de clientes: 7043
- Porcentaje de clientes dados de baja: 26.54 %

para facilitar el análisis se clasificará de diferentes maneras, lo cual nos permitiera tener una visión más clara de como están distribuidos los datos, quedando las clasificaciones de la siguiente manera:

- Segmentos demográficos
  - Gender
  - SeniorCitizen
  - Partner
  - Dependents

Donde el 50.48% de los clientes son hombres, 48.30% tienen pareja, 29.96 tienen dependientes, y el 16.21 son señores mayores a 65 años

La base de clientes está compuesta por hombres y mujeres en proporciones prácticamente iguales (50/50). Cerca de la mitad tiene pareja (48%), pero sólo el 30% tiene dependientes. Los adultos mayores representan 16%, lo cual indica que la base está dominada por clientes de edad media.

- Servicios contratados:
  - PhoneService
  - MultipleLines
  - OnlineSecurity

	Porcentaje
<b>gender</b>	<b>50.48 %</b>
<b>Partner</b>	<b>48.30 %</b>
<b>Dependents</b>	<b>29.96 %</b>
<b>SeniorCitizen</b>	<b>16.21 %</b>

Figure 1: Segmentos demográficos

- OnlineBackup
- DeviceProtection
- TechSupport
- StreamingTV
- StreamingMovies
- PaperlessBilling
- InternetService\_DSL
- InternetService\_Fiber optic
- No InternetService
- Contract\_Month-to-month
- Contract\_One year
- Contract\_Two year
- PaymentMethod\_Bank transfer (automatic)
- PaymentMethod\_Credit card (automatic)
- PaymentMethod\_Electronic check
- PaymentMethod\_Mailed check

	Porcentaje
<b>PhoneService</b>	<b>90.32 %</b>
<b>PaperlessBilling</b>	<b>59.22 %</b>
<b>Contract_Month-to-month</b>	<b>55.02 %</b>
<b>InternetService_Fiber optic</b>	<b>43.96 %</b>
<b>MultipleLines</b>	<b>42.18 %</b>
<b>StreamingMovies</b>	<b>38.79 %</b>
<b>StreamingTV</b>	<b>38.44 %</b>
<b>OnlineBackup</b>	<b>34.49 %</b>
<b>DeviceProtection</b>	<b>34.39 %</b>
<b>InternetService_DSL</b>	<b>34.37 %</b>
<b>PaymentMethod_Electronic check</b>	<b>33.58 %</b>
<b>TechSupport</b>	<b>29.02 %</b>
<b>OnlineSecurity</b>	<b>28.67 %</b>
<b>Contract_Two year</b>	<b>24.07 %</b>
<b>PaymentMethod_Mailed check</b>	<b>22.89 %</b>
<b>PaymentMethod_Bank transfer (automatic)</b>	<b>21.92 %</b>
<b>InternetService_No</b>	<b>21.67 %</b>
<b>PaymentMethod_Credit card (automatic)</b>	<b>21.61 %</b>
<b>Contract_One year</b>	<b>20.91 %</b>

Figure 2: Factores

Como podemos observar, la empresa es principalmente proveedora del servicio de telefonía, siendo este un servicio requerido por el 90% de sus clientes, mas de la mitad utiliza el sistema PaperlesBilling, aproximadamente la mitad (55.02%) de sus usuarios tienen un contrato mes a mes (posible riesgo) y el segundo servicio mas contratado por sus clientes es el internet de fibra óptica con un 43.96% del total de sus clientes (Factor importante a analizar).

- Variables continuas:
  - tenure
  - Monthly

las cuales nos permitiran evaluar el tiempo promedio de retencion del cliente, ver cuales son los periodos de mayor estadia como cliente, el costo mensual que representa a la media del los clientes, y el segmento de clientes con mayor riesgo de baja.

### 3.1 Análisis de Variables Continuas

Variable	Media Baja	Media No Baja	Mediana Baja	Mediana No Baja
tenure	17.98	37.57	10.00	38.00
Monthly	74.44	61.27	79.65	64.43

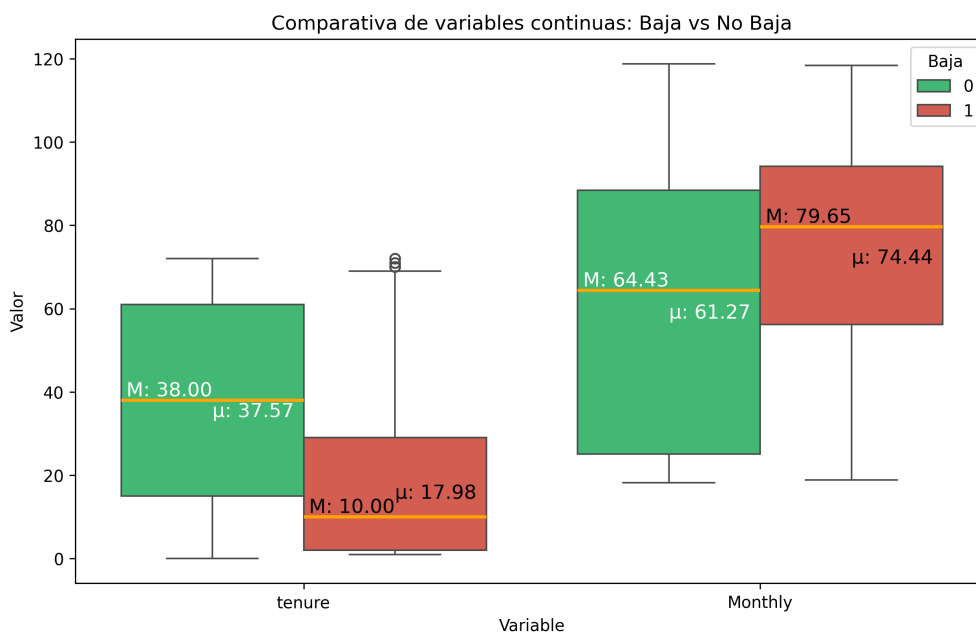


Figure 3: Boxplot Baja vs No

aunque el promedio de estancia de los clientes es de 18 meses, al calcular la mediana obtendremos que la mayor parte de los clientes que se van, tienen 10 meses o menos, de igual manera, visualizamos que los clientes que se dan de baja el costo mensual del servicio que tienen contratado es alto.

#### Insight:

- Clientes que se dan de baja son generalmente nuevos y con planes más caros.

### 3.2 Análisis de Segmentos de Clientes

Segmento	Baja=No (%)	Baja=Yes (%)
SeniorCitizen	12.87	25.47
Partner	52.82	35.79
Dependents	34.48	17.44

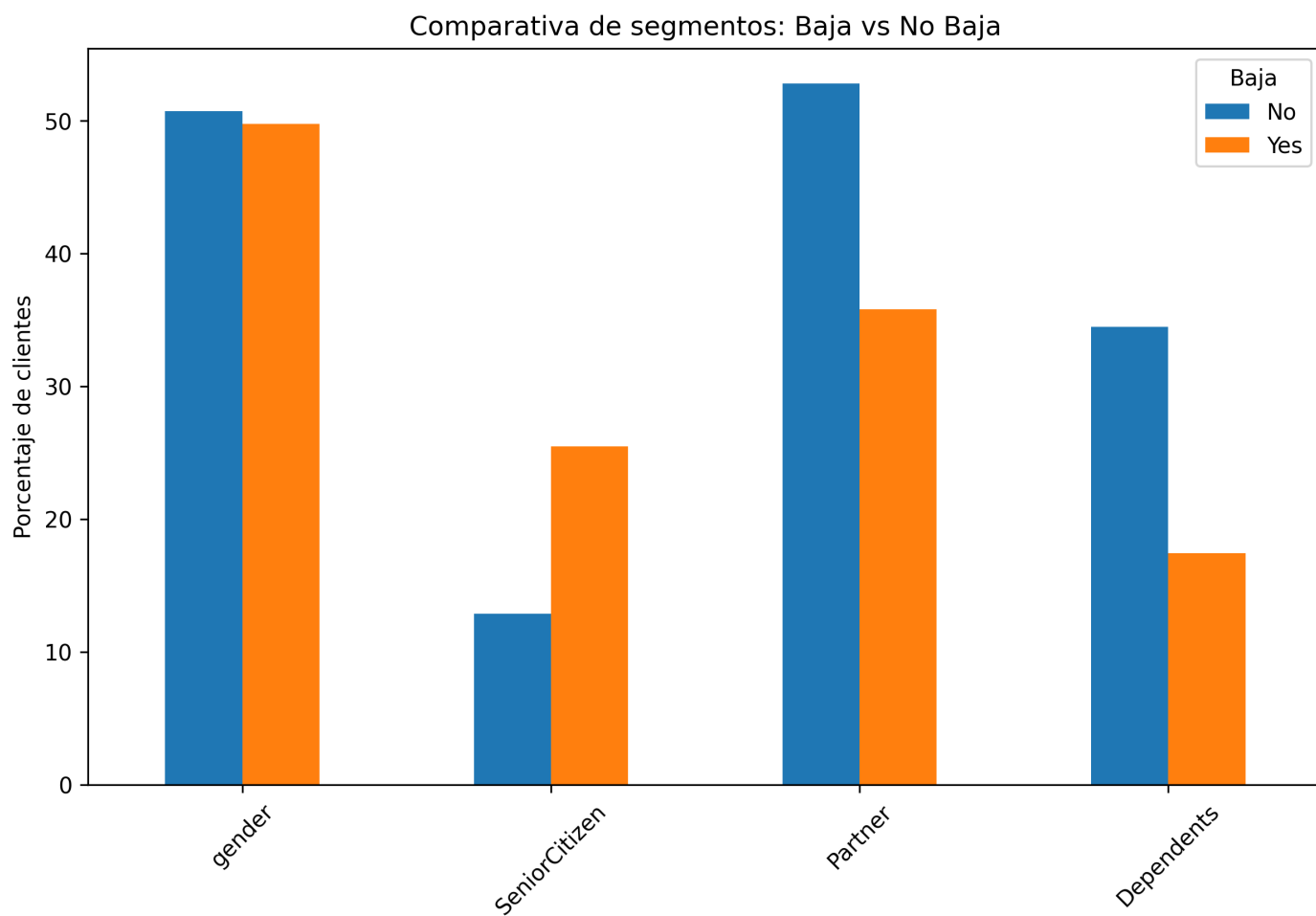


Figure 4: Segmentos comparativa

Podemos observar que el segmento más crítico es: personas mayores a 65, ya que el 25.47 % de estos son los que se van lo que genera retención es: personas con pareja, son 52.82 % de los que conservan el servicio personas con hijos, el 34.48 % de los que tienen hijos se quedan.

	gender	SeniorCitizen	Partner	Dependents
Baja				
0	50.73 %	12.87 %	52.82 %	34.48 %
1	49.76 %	25.47 %	35.79 %	17.44 %

**Insight:** - Mayor riesgo: clientes mayores, sin pareja ni dependientes.

### 3.3 Análisis de Servicios Contratados

mediante el análisis realizado de configuraciones de servicio de la clientela, podemos clasificar los segmentos, para ver cuáles son los que propician la salida o retención del cliente.

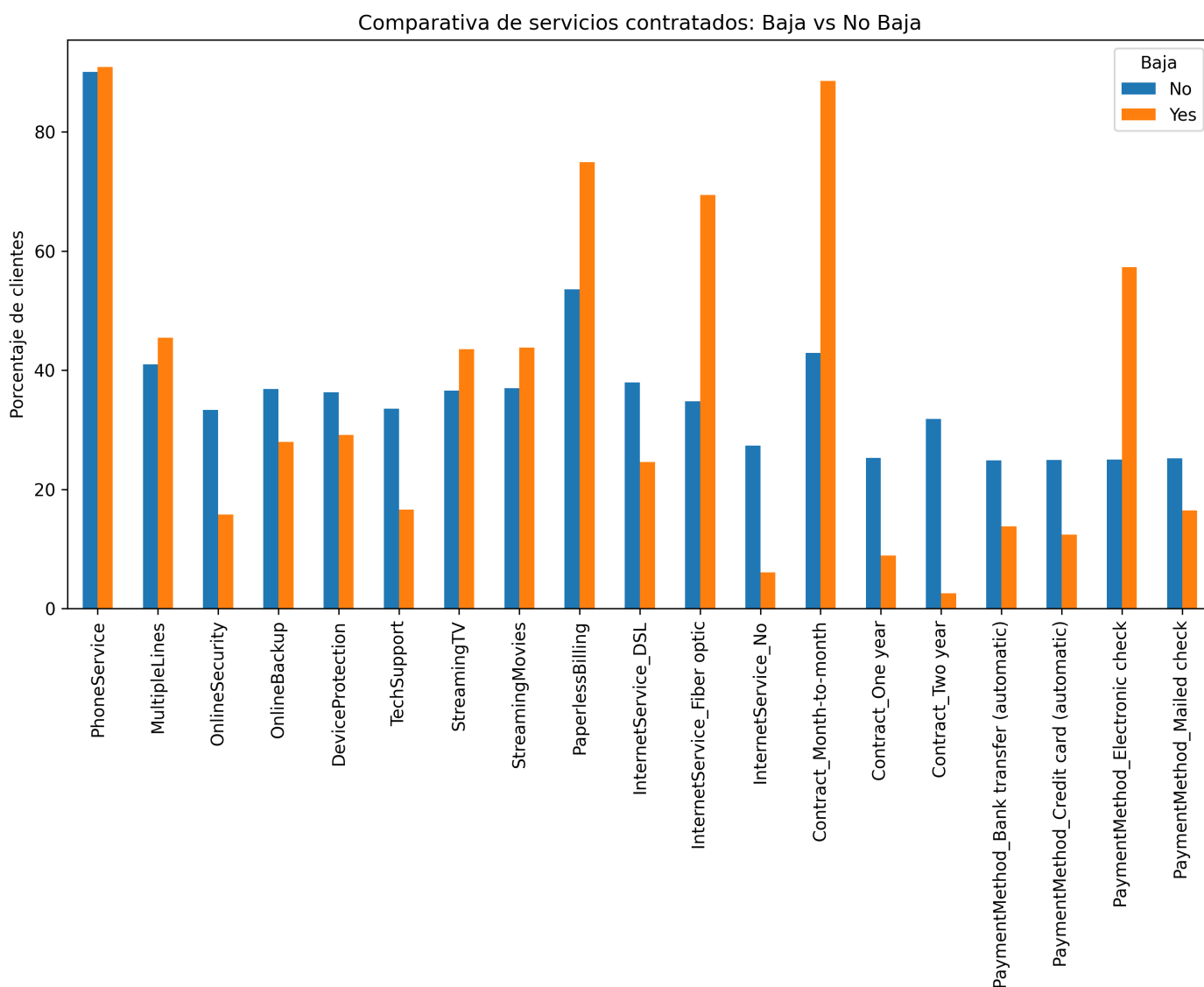


Figure 5: Barras comparativo

Para facilitar el entendimiento de los gráficos, se muestra por medio de una tabla, ya clasificados, como afectan los

servicios que la empresa **TelecomX**, siendo si estos representan un riesgo para la salida o la retencion de clientes.

Baja	0	1
PhoneService	90.10 %	90.90 %
MultipleLines	40.99 %	45.48 %
OnlineSecurity	33.32 %	15.78 %
OnlineBackup	36.84 %	27.98 %
DeviceProtection	36.28 %	29.16 %
TechSupport	33.51 %	16.59 %
StreamingTV	36.59 %	43.55 %
StreamingMovies	36.99 %	43.77 %
PaperlessBilling	53.56 %	74.91 %
InternetService_DSL	37.92 %	24.56 %
InternetService_Fiber optic	34.77 %	69.40 %
InternetService_No	27.31 %	6.05 %
Contract_Month-to-month	42.91 %	88.55 %
Contract_One year	25.26 %	8.88 %
Contract_Two year	31.83 %	2.57 %
PaymentMethod_Bank transfer (automatic)	24.86 %	13.80 %
PaymentMethod_Credit card (automatic)	24.93 %	12.41 %
PaymentMethod_Electronic check	25.01 %	57.30 %
PaymentMethod_Mailed check	25.20 %	16.48 %

Figure 6: Factores style transpuesta

#### Resumen:

Servicio	Baja=No (%)	Baja=Yes (%)	Clasificación
OnlineSecurity	33.32	15.78	Protege
OnlineBackup	36.84	27.98	Protege
DeviceProtection	36.28	29.16	Protege
TechSupport	33.51	16.59	Protege
StreamingTV	36.59	43.55	Riesgo
StreamingMovies	36.99	43.77	Riesgo
Internet Fibra	34.77	69.39	Riesgo
Contract Month-to-month	42.91	88.55	Riesgo
PaymentMethod Electronic Check	25.00	57.30	Riesgo
PaperlessBilling	53.56	74.91	Riesgo

**Insight:** - Servicios de soporte protegen contra la baja. - Contratos mensuales, planes caros y métodos de pago electrónicos aumentan el riesgo.

4.Conclusiones e Insights: Resume los principales hallazgos y cómo estos datos pueden ayudar a reducir la evasión. a lo largo del analisis observamos cuales fueron los sectores demograficos, de riesgo, y la media de cobro y retencion del servicio.

Se emiten las siguientes observaciones:

1. Para contrarrestar la salida de los clientes, se debe de evaluar la **calidad** del servicio que se le brinda a los clientes con respecto al **precio**

Los clientes que adquieren paquetes de internet fibra optica, y algun complemento adicional de streaming, no ve un gran **valor** vs el **gasto promedio** que conlleva dicha suscripcion

2. Los contratos **Month-to-Month**, los tipos de **pago electronico**, son la mayor razon de la no retención de los clientes, no genera un compromiso y al permitir **cancelaciones tempranas**, no deja ver el valor real del servicio al cliente.

si, bien, el servicio de paga electronico, permite al cliente reducir tramites, o movilización para pagar el servicio, al tener servicios **de alto** costo y **poco valor** el cliente optara por culminar su contrato

Se emiten las siguientes observaciones:

1. Focalizar campañas de retención en **clientes nuevos, mayores y sin dependientes**.
2. Promover la contratación de **servicios de soporte** y contratos de mayor duración.
3. Revisar planes de fibra y métodos de pago electrónicos para reducir el riesgo de baja.
4. analisis adicional

se realizaron tablas de correlación para complementar y validar nuestro analisis el cual muestra lo siguiente

	tenure	Monthly
tenure	1.00	0.25
Monthly	0.25	1.00

Figure 7: Correlación estilo

donde nos muestra que individualmente no existe una correlación entre el pago mensual y la baja.

en la siguiente tabla podemos observar cuales son los factores con mayor correlación, posibles causas (cerca de 1) y cuales son los que ayudan como sistema de retención (cerca de -1).

Variable	Correlación con Baja
Baja	1.000000
Contract_Month-to-month	0.405103
InternetService_Fiber optic	0.308020
PaymentMethod_Electronic check	0.301919
Monthly	0.193356
PaperlessBilling	0.191825
SeniorCitizen	0.150889
StreamingTV	0.063228
StreamingMovies	0.061382
MultipleLines	0.040102
PhoneService	0.011942
gender	-0.008612
DeviceProtection	-0.066160
OnlineBackup	-0.082255
PaymentMethod_Mailed check	-0.091683
PaymentMethod_Bank transfer (automatic)	-0.117937
InternetService_DSL	-0.124214
PaymentMethod_Credit card (automatic)	-0.134302
Partner	-0.150448
Dependents	-0.164221
TechSupport	-0.164674
OnlineSecurity	-0.171226
Contract_One year	-0.177820
InternetService_No	-0.227890

Variable	Correlación con Baja
Contract_Two year tenure	-0.302253
	-0.352229

Mapa completo de correlacion:

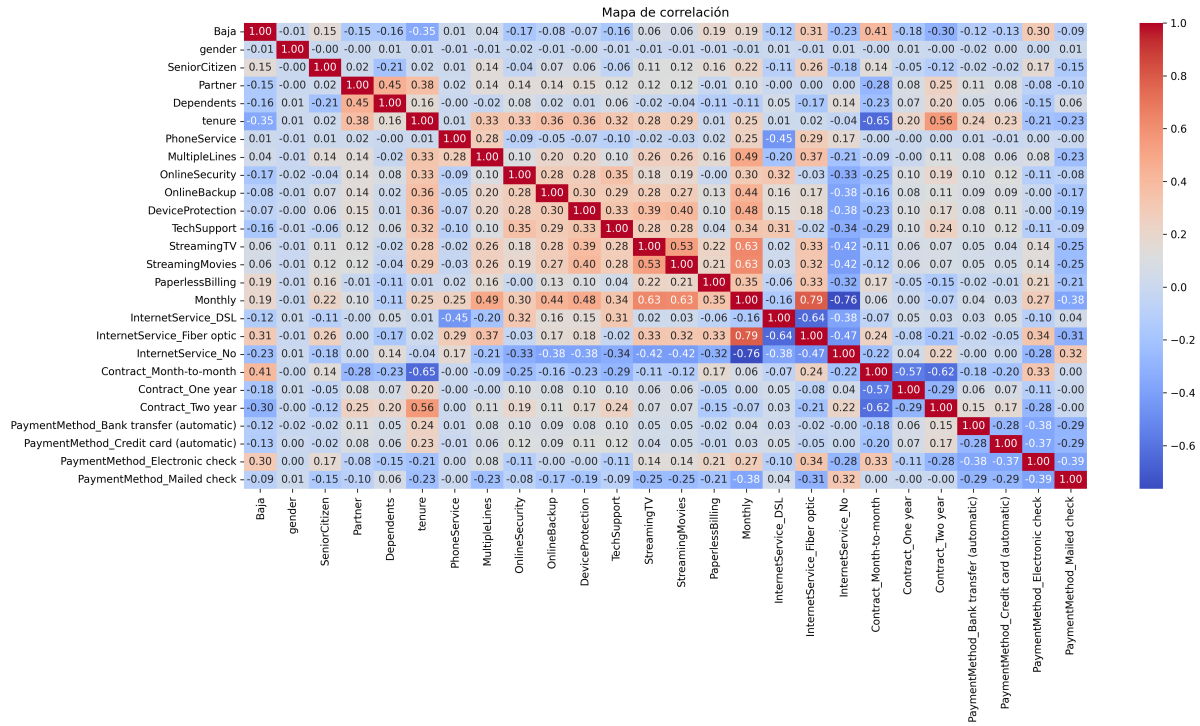


Figure 8: Mapa de correlación