*Telecom X - Informe*

1. **Análisis Descriptivo**

Calculo métricas como **media, mediana, desviación estándar** y otras medidas que ayuden a comprender mejor la distribución y el comportamiento de los clientes.

Numérico





Categórico

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

El contenido generado por IA puede ser incorrecto.

1. **Distribución de evasión**

Distribución variable "churn" (evasión) entre los clientes. Visualización proporción de clientes que permanecieron y los que se dieron de baja.

**Numérico**

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

Tenemos un total de 7267 clientes. En la distribución general de cancelación se observa que 71.2% (5174) de los clientes continúan con el servicio estando Activos, y el 25.7%(1869) corresponde a clientes que se dieron de baja porque desistieron, y existe un pequeño porcentaje 3.08%(224) que tiene un estado de cancelación desconocido.

Se evidencia una alta tasa de cancelación con un 25.7% que nos lleva a realizar un análisis más detallado.

Gráfico, Gráfico circular

El contenido generado por IA puede ser incorrecto.

Entre los clientes con un estado de cancelación definido, los Adultos Mayores (clientes que tienen 65 años o más) tienen una tasa de cancelación considerablemente más baja 22.9% en comparación con los que no son Adultos Mayores (clientes menores de 65 años) con un 40.3% de desistimiento.

En números absolutos, el grupo "No Adulto Mayor" tiene una base de clientes mucho mayor que el grupo "Adulto Mayor". A pesar de esto, el número absoluto de cancelaciones es sustancial en ambos grupos, con 1393 cancelaciones entre los que no son Adultos Mayores y 476 entre los Adultos Mayores (excluyendo el estado indefinido).

Se debe investigar las características y razones detrás de la alta tasa de cancelación entre los No Adultos Mayores en comparación con los que no son Adultos Mayores.

Gráfico, Gráfico circular

El contenido generado por IA puede ser incorrecto.

El número de cancelaciones para los clientes que tienen pareja es de 19%, y para los que no tienen pareja es de 32% con 1200 cancelaciones, siendo mas alto el desistimiento en clientes que no tienen pareja.

Gráfico, Gráfico circular

El contenido generado por IA puede ser incorrecto.

Los clientes que no tiene dependientes corresponde a una base amplia, y tiene mayor desistimiento con 30.3%(1543 clientes) frente a un 14.9%(326) de clientes que tienen dependientes

Gráfico, Gráfico circular

El contenido generado por IA puede ser incorrecto.

Los clientes que prefiere recibir la factura online, reflejan una alto desistimiento 32.5%(1400 clientes) frente a un 15.9%(469 clientes) que reciben factura fisica. Lo que nos lleva a validar si los clientes están teniendo problema al realizar el pago de forma electrónica

Gráfico, Gráfico de barras, Gráfico en cascada

El contenido generado por IA puede ser incorrecto.

Tenemos un base total de 7267 clientes, donde el 25.7%(1869) corresponde a clientes que se dieron de baja porque desistieron. De los cuales tenemos que los No Adultos Mayores (clientes menores a 65) tienen una tasa de cancelación considerablemente más alta con un 40.3% de desistimiento. De igual forma los clientes que no tienen pareja, tienen una alta tasa de cancelación con 32% que corresponde a 1200 cancelaciones de clientes. De igual forma los clientes que no tiene dependientes tiene mayor desistimiento con 30.3%(1543 clientes).

De igual forma se identifica que los clientes que prefiere recibir la factura online, reflejan una alto desistimiento 32.5%(1400 clientes).

Esto nos lleva a detectar que de la base las personas jóvenes, sin dependientes, sin pareja, y que reciben la factura online, son los clientes que mas desisten.

1. **Recuento de evasión por variables categóricas**

Exploración distribución de evasión según variables categóricas, como género, tipo de contrato, método de pago, entre otras.

Este análisis puede revelar patrones interesantes, por ejemplo, si los clientes de ciertos perfiles tienen una mayor tendencia a cancelar el servicio, lo que ayudará a orientar acciones estratégicas.

Gráfico, Gráfico circular

El contenido generado por IA puede ser incorrecto.

La tasa de cancelación es similar entre géneros: 26.1% para mujeres y 25.3% para hombres.

Dado que las tasas de cancelación son muy similares entre géneros, el género en sí mismo puede no ser un factor significativo que impulse la rotación de clientes (churn). Un análisis adicional debería centrarse en otros factores que podrían influir en las decisiones de cancelación.

Un pequeño porcentaje de los datos de cancelación se categoriza como "Cancelación Desconocida" (2.9% para mujeres y 3.27% para hombres).

Gráfico, Gráfico circular

El contenido generado por IA puede ser incorrecto.

Los clientes con suscripción a fibra óptica están presentando una alta tasa de desistimiento con un 40.6% (1297), respecto a DSL 18.4% (459 clientes) y

Gráfico, Gráfico circular

El contenido generado por IA puede ser incorrecto.

El tipo de contrato Mensual tiene una alta tasa de desistimiento con un 41.3% (1655 clientes), frente al contrato a 1 año con un 3%(166 clientes) y de 2 años (48 clientes)

Gráfico

El contenido generado por IA puede ser incorrecto.

El método de pago 'Electronic check' genera un alto desistimiento con un 43.8% (1071 clientes) respecto a 'Mailed check' (18.8%-308 clientes), 'Credit card (automatic)' (232) 'Bank transfer (258) (automatic)

Interfaz de usuario gráfica, Aplicación, Teams, Gráfico de barras

El contenido generado por IA puede ser incorrecto.

Al analizar estas variables categóricas se observa que el género en sí mismo no es un factor significativo que impulse la rotación de clientes (churn).

La suscripción del servicio de internet con fibra óptica están presentando una alta tasa de desistimiento con un 40.6% (1297), indicador fuerte de que ese tipo de servicio está asociado con la cancelación.

El tipo de contrato Mensual tiene una alta tasa de desistimiento con un 41.3% (1655 clientes), es muy común que los clientes con contratos mes a mes ('Month-to-month') tengan una tasa de cancelación mucho más alta que aquellos con contratos a largo plazo (uno o dos años). La grafica nos confirma esta tendencia y sugiere que los contratos a largo plazo son un factor de retención importante.

El método de pago 'Electronic check' genera un alto desistimiento con un 43.8% (1071 clientes), es un área a investigar.

1. **Conteo de evasión por variables numéricas**

Exploración variables numéricas, como "total gastado" o "tiempo de contrato", como se distribuyen entre los clientes que cancelaron (evasión) y los que no cancelaron.

Este análisis ayuda a entender si ciertos valores numéricos están más asociados con la evasión, proporcionando insights sobre los factores que influyen en el comportamiento de los clientes.

1. Suscripcion a diferentes servicios vs evasión

Gráfico, Gráfico de barras

El contenido generado por IA puede ser incorrecto.

Gráfico, Gráfico de barras

El contenido generado por IA puede ser incorrecto.

Al analizar las suscripciones de los clientes, bajo el análisis de activos y desistimientos, se observa que 6361 clientes tienen el servicio Telefonico, que es la mayor base de suscripciones por servicio, pero de igual forma es el servicio que tienen mas desistimiento con 1699 clientes. Sigue luego la suscripción al servicio de Internet con una base de 6930 clientes, donde la opción de Fibra Optica tiene alto desistimiento con 1297 clientes.

1. Validación columnas Meses\_Contrato, Factura\_Mensual, Cargos\_Totales y Cuentas\_diarias. Para lo cual use un boxplot.

Gráfico, Gráfico en cascada

El contenido generado por IA puede ser incorrecto.

Gráfico, Gráfico de cajas y bigotes

El contenido generado por IA puede ser incorrecto.

Al analizar los **Meses de Contrato**, se observa una clara relación con la cancelación. El boxplot muestra que los clientes que se dieron de baja tienden a tener una antigüedad significativamente menor (mediana de X meses) en comparación con los clientes activos (mediana de Y meses). El gráfico de líneas refuerza esta observación, indicando que el pico de cancelaciones ocurre en los primeros meses de servicio y disminuye gradualmente a medida que los clientes permanecen más tiempo con la empresa.

* Los clientes que cancelan lo hacen al principio de su contrato, por lo que se ve la caja y la mediana de los clientes 'De Baja' concentradas en valores bajos de 'Meses\_Contrato'.
* La dispersión (altura de la caja y bigotes) entre los grupos. Los 'Activos' tienen una distribución de meses de contrato mucho más amplia que los 'De Baja'

Gráfico, Gráfico de cajas y bigotes

El contenido generado por IA puede ser incorrecto.

Gráfico, Gráfico de cajas y bigotes

El contenido generado por IA puede ser incorrecto.

En cuanto a la **Factura Mensual**, los gráficos sugieren que los clientes con facturas mensuales más altas tienen una mayor propensión a cancelar]. El boxplot ilustra que la mediana de la factura mensual para los clientes 'De Baja' es de aproximadamente X, mientras que para los 'Activos' es de Y. El gráfico de líneas muestra que hay un mayor número de cancelaciones en el rango de facturas entre XX y YY USD]."

* + A menudo, los clientes con cargos mensuales más altos tienen una mayor probabilidad de cancelar, especialmente si no perciben un valor acorde.

Imagen que contiene Gráfico

El contenido generado por IA puede ser incorrecto.

Gráfico, Gráfico de cajas y bigotes

El contenido generado por IA puede ser incorrecto.

La variable **Cargos Totales** también presenta una relación con la cancelación, aunque influenciada por la antigüedad del cliente. El boxplot indica que los clientes activos han acumulado, como era de esperar, cargos totales significativamente mayores que los clientes que se dieron de baja. El gráfico de líneas de conteo muestra que la mayoría de las cancelaciones ocurren entre clientes con cargos totales acumulados de hasta X USD, lo que se alinea con la observación de que la cancelación es más frecuente en los primeros meses de contrato.

* + Los clientes con muchos meses de contrato tendrán Cargos Totales altos si no cancelaron. Los 'De Baja' a menudo tendrán Cargos Totales más bajos si cancelaron pronto. La Mediana de los 'De Baja' suele ser mucho más baja que la de los 'Activos'.

Gráfico, Gráfico de cajas y bigotes

El contenido generado por IA puede ser incorrecto.

Gráfico, Gráfico de cajas y bigotes

El contenido generado por IA puede ser incorrecto.

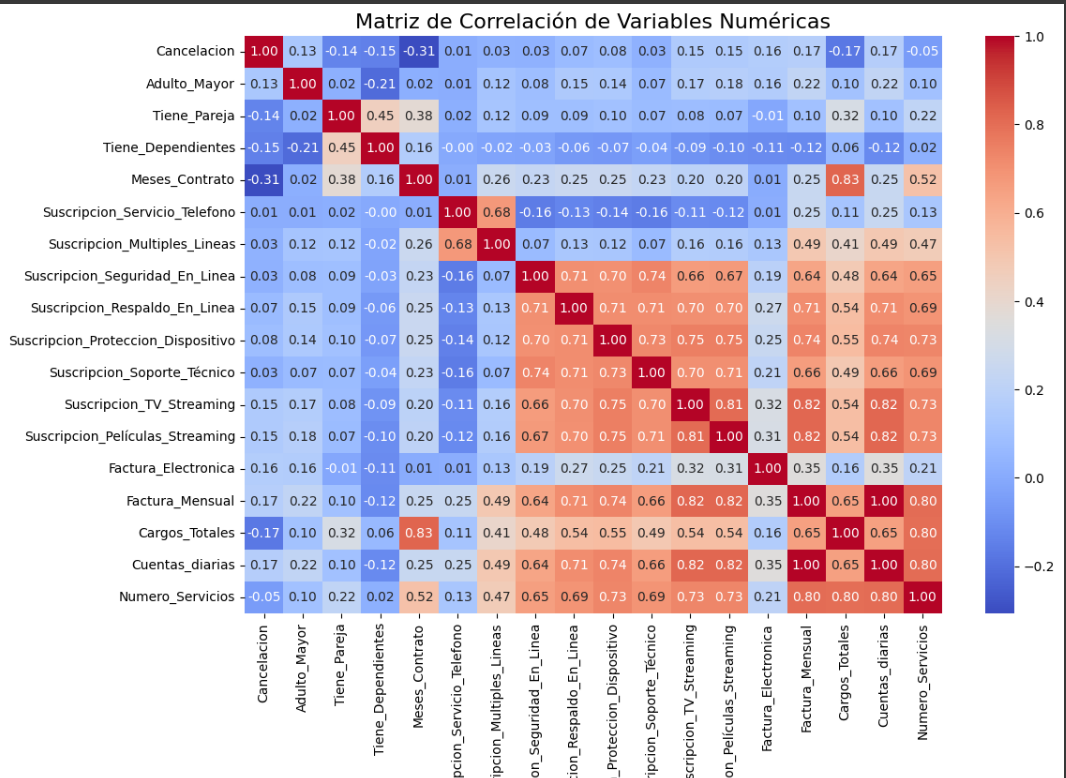
Para comprender mejor cómo el costo diario estimado del servicio influye en la decisión de un cliente de permanecer o cancelar, analizamos la variable **Cuentas Diarias** en relación con el estado de cancelación.

El **gráfico de caja** (izquierda) muestra la distribución de las Cuentas Diarias para cada grupo de clientes (Activos, De Baja, No definidos). Observamos que los clientes que se dieron de baja tienden a tener un costo diario promedio (mediana) ligeramente superior en comparación con los clientes activos. La dispersión de los datos parece similar entre los grupos Activos y De Baja.

El **gráfico de líneas** (derecha) complementa esta vista al mostrar el conteo de clientes en cada estado de cancelación a lo largo de los diferentes rangos de Cuentas Diarias. Este gráfico nos permite identificar dónde se concentran las cancelaciones en términos de costo diario. Se evidencia que si bien hay clientes activos en todos los rangos de Cuentas Diarias, el número de clientes 'De Baja' presenta un pico notable en el rango de costo diario entre X y Y USD, y disminuye para costos diarios más altos. Los clientes 'No definidos' se distribuyen de manera más uniforme en los rangos de costo diario más bajos"].

**En resumen**, el análisis de las Cuentas Diarias sugiere que, aunque no es el único factor, un costo diario percibido como alto podría estar asociado con una mayor probabilidad de cancelación, especialmente en ciertos rangos de precios. Esta información es valiosa para estrategias de pricing o la identificación de clientes en riesgo basados en su consumo diario estimado.

1. **Análisis de correlación entre variables**



La matriz de correlación y el mapa de calor es fundamental para entender qué variables numéricas tienen la relación más fuerte con la cancelación de clientes.

**Interpretación del Mapa de Calor de Correlación**

El mapa de calor muestra la fuerza y dirección de la relación lineal entre cada par de variables numéricas.

* **Colores:** La intensidad y el color de cada celda indican el coeficiente de correlación (un valor entre -1 y 1).
  + Los colores cálidos (rojo) indican una **correlación positiva**: cuando una variable aumenta, la otra tiende a aumentar también.
  + Los colores fríos (azul) indican una **correlación negativa**: cuando una variable aumenta, la otra tiende a disminuir.
  + Los colores cercanos a cero (generalmente blanco o colores neutros en el centro del rango de color) indican una **correlación débil o nula**.
* **Valores Numéricos:** Los números dentro de cada celda son los coeficientes de correlación.
  + Un valor cercano a 1 (positivo o negativo) indica una correlación fuerte.
  + Un valor cercano a 0 indica una correlación débil.
* **Diagonal Principal:** La diagonal principal siempre tendrá un valor de 1.00 porque representa la correlación de una variable consigo misma.

**Enfócate en la Columna 'Cancelacion'**

Para tu informe sobre la evasión de clientes, el enfoque principal estará en la fila o columna correspondiente a la variable **'Cancelacion'**. Observa los coeficientes de correlación en esta fila/columna:

* **Correlaciones Positivas con 'Cancelacion':** Las variables con coeficientes positivos significan que, a medida que aumenta el valor de esa variable, la probabilidad de cancelación también tiende a aumentar. Busca los valores positivos más altos.
  + Por ejemplo, si 'Factura\_Mensual' tiene un coeficiente positivo relativamente alto con 'Cancelacion', sugiere que los clientes con facturas mensuales más altas tienden a cancelar más.
* **Correlaciones Negativas con 'Cancelacion':** Las variables con coeficientes negativos significan que, a medida que aumenta el valor de esa variable, la probabilidad de cancelación tiende a disminuir. Busca los valores negativos más cercanos a -1.
  + Por ejemplo, si 'Meses\_Contrato' tiene un coeficiente negativo relativamente alto con 'Cancelacion', sugiere que los clientes con más antigüedad son menos propensos a cancelar.
* **Correlaciones Débiles con 'Cancelacion':** Las variables con coeficientes cercanos a cero tienen poca o ninguna relación lineal directa con la cancelación.

**Resultados Clave para el Informe**

Basándote en tu mapa de calor, identifica las variables numéricas con las correlaciones más fuertes (positivas o negativas) con 'Cancelacion'.

Aquí tienes una estructura para redactar los resultados en tu informe:

**Análisis de Correlación de Variables Numéricas con la Cancelación**

Realizamos un análisis de correlación para identificar las relaciones lineales entre las variables numéricas de los clientes y su estado de cancelación. El mapa de calor de correlación adjunto visualiza estos resultados, donde los colores indican la fuerza y dirección de la relación (rojo para correlación positiva, azul para negativa).

Los hallazgos clave en relación con la **cancelación de clientes** son los siguientes:

* **Variables con Correlación Negativa Significativa:**
  + La variable **Meses de Contrato** muestra una correlación negativa fuerte con la cancelación (coeficiente: [Valor del coeficiente]). Esto indica que, a mayor antigüedad del cliente, menor es la probabilidad de que cancele el servicio. Este es un hallazgo importante, sugiriendo que la lealtad del cliente aumenta con el tiempo.
  + [Menciona otras variables con correlación negativa notable y su coeficiente, si las hay. Por ejemplo: "Los **Cargos Totales** también presentan una correlación negativa con la cancelación (coeficiente: [Valor]), lo cual está relacionado con la antigüedad del cliente."]
* **Variables con Correlación Positiva Significativa:**
  + La **Factura Mensual** muestra una correlación positiva con la cancelación (coeficiente: [Valor del coeficiente]). Esto sugiere que los clientes con facturas mensuales más altas tienen una mayor tendencia a cancelar.
  + La **Cuenta Diaria** (estimación) presenta una correlación positiva similar con la cancelación (coeficiente: [Valor del coeficiente]), lo cual es coherente con la relación observada en la Factura Mensual.
  + [Menciona otras variables con correlación positiva notable y su coeficiente, si las hay. Por ejemplo: "Ser **Adulto Mayor** muestra una correlación positiva [débil/moderada] con la cancelación (coeficiente: [Valor]), indicando que este segmento puede tener una mayor propensión a irse."]
* **Variables con Correlación Débil o Nula:**
  + [Menciona variables que tengan coeficientes cercanos a cero con 'Cancelacion', indicando que no tienen una relación lineal fuerte, ej: "El **Número de Servicios** contratados muestra una correlación débil con la cancelación (coeficiente: [Valor]), lo que sugiere que la cantidad de servicios por sí sola no es un predictor lineal fuerte de la cancelación."]

**Implicaciones:**

Estos resultados de correlación confirman la importancia de factores como la antigüedad del cliente y el monto de la factura en la probabilidad de cancelación. Estas variables serán candidatas clave para incluir en modelos predictivos de churn.