

## Introducción

### Industria del *ridesharing*

La industria del *ridesharing* es aquella en la que un pasajero viaja en un vehículo privado, manejado por su dueño y usando como intermediario para arreglar este encuentro una app o web site. Ejemplos de servicios de estas plataformas son aplicaciones tales como Uber, DiDi, Cabify, Beat, etc. todas ellas marcas presentes en el mercado nacional e internacional.

Esta industria nació en el año 2009 con UberCab y desde entonces, la industria se ha expandido de forma vertiginosa lo cual demuestra el gran impacto que estas nacientes plataformas han tenido sobre la industria de la movilidad, la cual si bien ha sufrido un gran impacto producto de la pandemia mundial, empresas como DiDi han logrado batir récords en el mes de Julio, alcanzando más de 6 millones de órdenes de viaje fuera de China en un solo día.

Dentro de la industria del *ridesharing*, es necesario resolver problemas, inconvenientes o dudas de los usuarios, es por esto que las principales empresas de *ridesharing* tienen departamentos de servicio al cliente, donde se gestionan y resuelven las problemáticas de los usuarios, ya sean los conductores o los pasajeros.

Sin embargo, no hay estudios sobre el efecto que un departamento de servicio al cliente tiene sobre el usuario pasajero, específicamente si un buen o mal servicio tiene un impacto sobre la frecuencia de uso de la aplicación en el futuro. Confirmar la influencia que un buen servicio al cliente tiene sobre el comportamiento futuro de un pasajero tendría un impacto importante, justificando así la inversión en esta área por parte de las principales empresas de *ridesharing* buscando generar lealtad por parte de los usuarios usando un eficaz y eficiente servicio al cliente.

Para el presente estudio, se obtuvieron los datos de la empresa de *ridesharing* DiDi, esta es una empresa de China fundada en el año 2012, la cual comenzó sus operaciones en Chile en junio del año 2019, actualmente, en términos de viajes, es la segunda mayor empresa de *ridesharing* en el Chile, después de Uber y la mayor del mundo, con presencia en .

## **Satisfacción del usuario**

La satisfacción del usuario es definida como la evaluación de la discrepancia percibida por el usuario entre las expectativas y el real desempeño de un producto o servicio (Tse y Wilton, 1988).

En toda organización, se utilizan Indicadores Claves de Rendimiento o KPIs, por sus siglas en inglés (*Key Performance Indicator*), y el área de servicio al cliente no es ajena a esto. En este estudio, se analizarán 3 KPIs usados por el departamento de servicio al cliente de la aplicación de movilidad, los cuales se utilizan para medir el éxito (o fracaso) de las diferentes medidas que se utilizan para mejorar la experiencia de un usuario con el servicio.

La principal métrica dentro del área del servicio al cliente es la Satisfacción del cliente o CSAT (*Customer Satisfaction*), esta métrica indica qué tan satisfecho está un cliente con el servicio que se le otorgó luego de contactarse con el departamento o área de servicio al cliente de una organización. Esta métrica se recolecta al pedir al usuario evaluar la calidad de la atención recibida luego de terminada la gestión con este, generalmente, se utiliza una escala numérica para esta evaluación, dónde mientras más alta la nota, mejor es la atención recibida.

Existen diversos factores que afectan la satisfacción del usuario, sin embargo, se pueden resumir en dos grandes grupos: las características del producto/servicio en cuestión y la calidad de la atención recibida por parte del área de servicio al cliente (Jahanshahi, Gashti, Mirdamadi, Nawaser, y Khaksar, 2011). En el caso de la industria del ride sharing, no se han efectuado estudios sobre los factores que determinan la satisfacción del usuario.

## **Objetivos**

### **Objetivo General**

Este trabajo tiene como fin determinar los factores que más influyen sobre la satisfacción del usuario pasajero al contactarse con el servicio al cliente de DiDi y cómo esta experiencia puede afectar el futuro nivel de uso del servicio.

## **Objetivos específicos**

1. Identificar los factores principales que influyen en la satisfacción del usuario.
2. Determinar qué factores de una mala experiencia con el Servicio al Cliente hacen que un pasajero deje de utilizar la aplicación en el futuro, y a partir de esto definir procedimientos desde el departamento de servicio al cliente que permitan identificar tempranamente a usuarios que tienen una alta probabilidad de dejar de utilizar el servicio.

Lograr los objetivos antes mencionados permitirá a la empresa aumentar las métricas de satisfacción del usuario, y por consiguiente generar una mejor experiencia para los pasajeros, y así disminuir la probabilidad de que estos dejen de usar el servicio por una eventual mala experiencia que puedan haber tenido con el servicio al cliente.

Estos objetivos se lograrán mediante la ejecución de dos modelos:

1. Modelo *Ordered Probit*, para explicar la influencia de distintos factores sobre la satisfacción del usuario.
2. Modelo *Ordered Probit*, para explicar la influencia de distintos factores sobre el cambio en el nivel de uso de la aplicación antes y después del contacto con el servicio al cliente.

## **Estructura**

En este estudio se empezará describiendo el marco teórico, donde se explicarán los estudios realizados sobre la influencia de la satisfacción del usuario sobre su comportamiento futuro y cuales son los principales factores que influyen sobre este. Por otra parte se describirá a fondo el modelo utilizado para explicar la satisfacción del consumidor y el uso futuro de la aplicación.

Luego, pasando a la metodología, se explicará en qué consiste la encuesta de satisfacción al usuario de DiDi, la cual es la base del estudio y el principal medio para

obtener datos, sumado a una descripción de la muestra a utilizar. A su vez, se expondrán las variables dependientes e independientes a ser usadas en el modelo, explicando los criterios para elegir los factores que explicarán las variables dependientes de los dos modelos propuestos. Además dentro de la metodología, también se hará mención de las herramientas utilizadas para el estudio.

En la sección de resultados, se expondrán los resultados de ambos modelos, haciendo énfasis en 3 aspectos de la interpretación de los resultados: La bondad de ajuste del modelo, la significancia de los umbrales del modelo (a explicar más adelante) y un análisis factor a factor: De su significancia y el signo del parámetro.

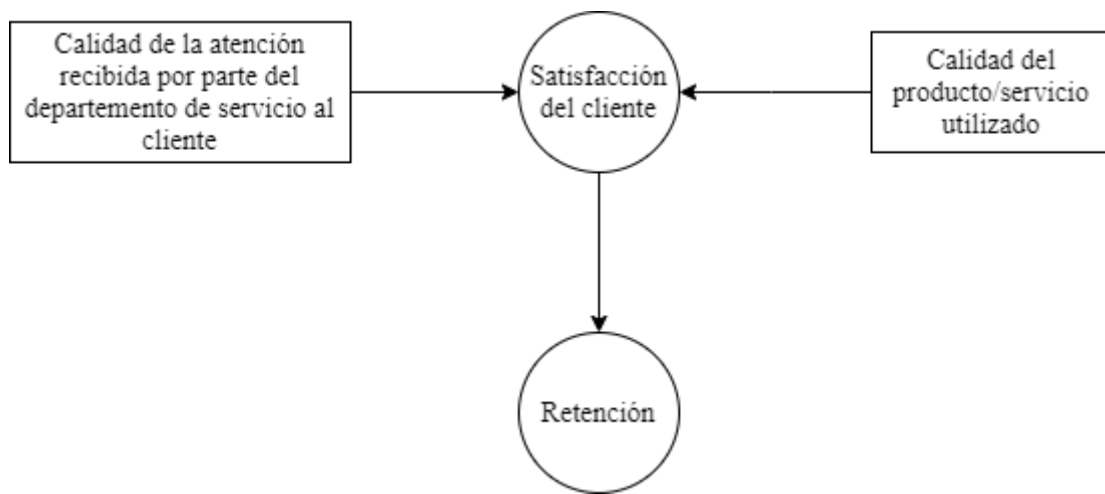
## **Cuerpo de la tesis**

### **Marco teórico**

#### **Antecedentes: Satisfacción del usuario y uso de la app**

La satisfacción del usuario está definida por muchos factores, los cuales tienen una incidencia directa por sobre este. Jahanshahi, Gashti, Mirdamadi, Nawaser, y Khaksar. (2011) demostraron que en la industria automotriz la satisfacción del usuario depende significativamente de la calidad del producto y de la calidad del servicio al cliente .

Por otra parte, también demostraron que existe una relación entre la satisfacción del usuario y la lealtad con una compañía automotriz, esta satisfacción también está fuertemente influenciada por la calidad del servicio al cliente.



**Figura X:** Factores que afectan la satisfacción del usuario (Elaboración propia)

Hasta el momento no existen estudios que hayan investigado la relación entre la satisfacción de un consumidor con el servicio al cliente y su futuro nivel de uso del servicio en la industria del *ridesharing*. Sin embargo, se han realizado estudios para otras industrias, tales como la de las telecomunicaciones, en la cual, el fenómeno estudiado por Kim y Yoon (2004) arrojó que si bien factores tales como la imagen de la marca, calidad del equipo y la calidad de las llamadas son significativas para explicar la lealtad de un cliente a futuro, la satisfacción con el servicio al cliente no es una variable significativa que explique esto.

En este estudio se considerará netamente el aspecto relacionado a la calidad de la atención recibida por el usuario y las características de este usuario, y sus subsecuentes efectos sobre la satisfacción que este usuario pueda tener y su futuro uso sobre la aplicación.

## Modelo

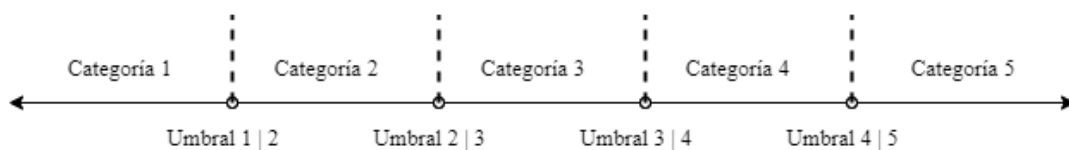
### Ordered Probit

Se utilizará el modelo *Ordered Probit*, propuesto por primera vez por McKelvey y Zavoina (1971), ampliamente utilizado para el análisis de selecciones, respuestas o

resultados que son ordenadas, categóricas y no cuantitativas. (Referencia Libro Modeling Ordered Choices, Greene, Hensher)

Para propósitos de este estudio, este modelo será utilizado para analizar la influencia de un número de factores sobre la satisfacción del usuario y la retención, variables que, como se describió, son ordenadas, categóricas y no cuantitativas.

El modelo es utilizado para el análisis de variables categóricas ordenadas, por ejemplo, para respuestas de encuestas que están en una escala de Likert (Muy malo, malo, regular, bueno, muy bueno). Este modelo utiliza una variable continua, para ser segmentada en diferentes “secciones” que representan cada categoría dentro de la variable categórica ordenada a estudiar. Para definir dónde termina una “sección” y comienza otra, se tienen diferentes umbrales en los cuales se define este límite, como se observa en el ejemplo de la Figura X.



**Figura X:** Recta numérica segmentada representando los valores de la variable dependiente del modelo *Ordered Probit*. (Elaboración propia)

Como se mencionó anteriormente, el modelo a utilizar para este estudio es el modelo *Ordered Probit*, este modelo se ajusta a las características de ambas variables dependientes que se desean estudiar, es decir, ambas son variables categóricas y ordenadas. (Modeling Ordered Choices, Greene, Hensher)

La ecuación del modelo a utilizar es la siguiente:

$$y_i^* = \beta x_i + \varepsilon_i, i = 1, \dots, n, \varepsilon \sim N(0, 1)$$

Donde el vector  $y_i^*$  es la variable continua latente utilizada para ser segmentada posteriormente y a partir de esta obtener la verdadera variable de interés:  $y_i$ .

El vector  $x_i$  representa los valores de las variables, en los modelos a estudiar, en general estas serán binarias, estas variables indican la presencia o no de ciertos factores que pueden influir en lo que se busca estudiar, en este caso la satisfacción del usuario. Un ejemplo de una variable binaria, es la zona geográfica desde la que se contacta un usuario, estas pueden ser: Norte, Centro Norte, Región Metropolitana, Centro Sur, Sur. Cada una de estas zonas toma la forma de una variable independiente en el modelo, sin embargo, una de estas es dejada fuera del modelo sirviendo así como variable “base” a partir de la cual se compara si hay una diferencia estadísticamente significativa en la satisfacción entre cada una de las zonas presentes en el modelo y la zona “base”. Entonces, por ejemplo, para un usuario que se contacta desde la Zona Sur, el valor de las variables para su observación en específico es:

- $x_{Norte} = 0$
- $x_{Centro Norte} = 0$
- $x_{Centro Sur} = 0$
- $x_{Sur} = 1$

Dejando como base (fuera del modelo) la variable  $x_{Región Metropolitana}$

El vector  $\beta$  representa los coeficientes de los diferentes parámetros que se utilizan para determinar el valor de la variable de utilidad continua, por ejemplo, en un modelo para estimar la satisfacción del usuario, se observa que la variable correspondiente a el número de viajes está acompañada de un parámetro  $\beta$  positivo, esto indica que a un mayor número de viajes, más aumenta el valor de  $y_i^*$  (y por consiguiente, la satisfacción), por el contrario, si este parámetro  $\beta$  fuera negativo, esto indicaría que en general la satisfacción es menor a medida que los viajes del usuario en cuestión aumentan.

Además, para propósitos de este modelo el error  $\varepsilon$  se asume que tiene una distribución normal estándar con media 0 y varianza 1.

Finalmente, la variable  $y_i^*$  es interpretada de la siguiente forma:

$$\begin{aligned}
 y_i &= 1 \text{ si } y_i^* \leq \mu_{1|2} \\
 y_i &= 2 \text{ si } \mu_{1|2} < y_i^* \leq \mu_{2|3} \\
 y_i &= 3 \text{ si } \mu_{2|3} < y_i^* \leq \mu_{3|4} \\
 y_i &= 4 \text{ si } \mu_{3|4} < y_i^* \leq \mu_{4|5} \\
 y_i &= 5 \text{ si } y_i^* > \mu_{4|5}
 \end{aligned}$$

Donde  $\mu_{j|j+1}$ ,  $j = 1, 2, 3, 4$  representa los distintos valores que pueden tomar los umbrales que “segmentan” a la variable continua  $y_i^*$  dependiendo de su valor e  $y_i$  es el valor final observado.

Se observa entonces, que en este modelo se estiman dos cosas: Los parámetros  $\beta$  y los umbrales  $\mu$ , estos parámetros se estiman utilizando la máxima verosimilitud (o *Maximum Likelihood Estimation*) como función objetivo, para desarrollar esta función se necesitan las probabilidades de elección. A partir de las ecuaciones presentadas anteriormente, se puede deducir que la probabilidad de que  $y_i = 1$  será igual a la probabilidad de que  $y_i^*$  sea menor al umbral  $\mu_{1|2}$ , por otra parte, se sabe que  $y_i^*$  es igual al producto entre los vectores de parámetros por el vector de variables más el error ( $\beta x_i + \varepsilon_i$ ), en términos matemáticos esto se ve así:



$$P[y_i = 1] = P[y_i^* \leq \mu_{1|2}]$$

$$P[y_i = 1] = P[\beta x_i + \varepsilon_i \leq \mu_{1|2}]$$

Despejando  $\varepsilon_i$ :

$$P[y_i = 1] = P[\varepsilon_i \leq \mu_{1|2} - \beta x_i]$$

Y ahora, puesto que  $\varepsilon_i$  se distribuye normal:

$$P[y_i = 1] = \Phi[\mu_{1|2} - \beta x_i]$$

En conclusión, se puede decir que la probabilidad de que  $y_i = 1$  es igual a la probabilidad de que  $\varepsilon_i$  (Que se distribuye normal) sea menor a  $\mu_{1|2} - \beta x_i$ .

Ahora, el procedimiento para calcular  $P[y_i = 2]$  (Y por consiguiente  $P[y_i = 3]$  o  $P[y_i = 4]$ ), no es radicalmente diferente al indicado anteriormente:

$$P[y_i = 2] = P[\mu_{1|2} < \beta x_i + \varepsilon_i \leq \mu_{2|3}]$$

$$P[y_i = 2] = P[\mu_{1|2} - \beta x_i < \varepsilon_i \leq \mu_{2|3} - \beta x_i]$$

$$P[y_i = 2] = \Phi[\mu_{2|3} - \beta x_i] - \Phi[\mu_{1|2} - \beta x_i]$$

En palabras simples, se puede concluir, que la probabilidad de que  $y_i = 2$  es igual a la probabilidad de que  $\varepsilon_i$  sea mayor a  $\mu_{1|2} - \beta x_i$  y menor o igual a  $\mu_{2|3} - \beta x_i$ , entonces en general:

$$P[y_i = j] = \Phi[\mu_{j|j+1} - \beta x_i] - \Phi[\mu_{j-1|j} - \beta x_i]$$

Es de observar que la probabilidad de  $y_i$  sea su valor máximo de  $m$  es:

$$P[y_i = m] = 1 - \Phi[\mu_{j-1|j} - \beta x_i]$$

Y la probabilidad de que  $y_i$  sea su valor mínimo:

$$P[y_i = 1] = \Phi[\mu_{j|j+1} - \beta x_i]$$

A partir de esto, se puede definir la función de verosimilitud:

$$LL = \sum_{i=1}^N \sum_{j=1}^m Z_{ij} \ln[\Phi_{ij} - \Phi_{ij-1}]$$

Donde  $Z_{ij}$  es igual 1 si es que  $y_i = j$  y 0 en cualquier otro caso y  $\Phi_{ij} = \Phi[\mu_{j|j+1} - \beta x_i]$ , al maximizar el valor de esta función, se obtiene lo que se conoce como *Maximum Likelihood*, es decir, el punto en el cual la verosimilitud es máxima y donde se encuentran los valores óptimos de los parámetros  $\beta$  y los umbrales  $\mu$ .

Ahora, el efecto que tienen los parámetros sobre las probabilidades de elección de cada uno de los valores de  $y_i$  están definidas por las siguientes ecuaciones, en el caso del modelo de un modelo de 5 niveles, como el ejemplificado en la figura X:

$$\begin{aligned} \frac{\partial P(y_i = 1)}{\partial x_{ij}} &= -\Phi[\mu_{1|2} - \beta x_i] \times \beta_j \\ \frac{\partial P(y_i = 2)}{\partial x_{ij}} &= (\Phi[\mu_{1|2} - \beta x_i] - \Phi[\mu_{2|3} - \beta x_i]) \times \beta_j \\ \frac{\partial P(y_i = 3)}{\partial x_{ij}} &= (\Phi[\mu_{2|3} - \beta x_i] - \Phi[\mu_{3|4} - \beta x_i]) \times \beta_j \\ \frac{\partial P(y_i = 4)}{\partial x_{ij}} &= (\Phi[\mu_{3|4} - \beta x_i] - \Phi[\mu_{4|5} - \beta x_i]) \times \beta_j \\ \frac{\partial P(y_i = 5)}{\partial x_{ij}} &= \Phi[\mu_{4|5} - \beta x_i] \times \beta_j \end{aligned}$$

Al derivar las funciones de probabilidad, se ve que cuando el  $\beta_j$  es positivo, las probabilidades de elección de  $y_i = 1$  disminuyen, dado el signo negativo de la función  $\Phi$ , mientras que las probabilidades de elección de  $y_i = 5$ , aumentan dado el signo positivo de la función antes mencionada.

## Métricas de bondad de ajuste

Para evaluar la bondad de ajuste del modelo propuesto se utilizarán tres métricas:

1. Log Likelihood del modelo estudiado comparado con el modelo nulo: El Log Likelihood es una métrica que indica que tan bien el modelo se ajusta a la realidad, en este caso, los datos (La antes mencionada máxima verosimilitud). Esta métrica por sí sola no dice mucho y es necesario compararla con el Log Likelihood de otros modelos y así tener una referencia de que tanta es la bondad de ajuste de cierto modelo. Por regla general, mientras mayor el valor del Log Likelihood, mejor está ajustado el modelo a los datos.
2. Likelihood Ratio: Esta métrica se utilizará para demostrar que el modelo es significativamente diferente al modelo nulo, a partir de la siguiente fórmula:

$$LR = -2 * [\text{Log Likelihood}(0) - \text{Log Likelihood}(\text{Modelo})]$$

Con este valor de LR, se hará una prueba Chi Cuadrado con el fin de rechazar o comprobar la hipótesis nula, es decir, que el modelo no es significativamente diferente al modelo nulo, al obtener un valor menor a 0,05 en esta prueba se puede rechazar esta hipótesis.

3. Pseudo Rho Cuadrado: Es calculada de la siguiente manera:

$$\rho^2 = 1 - \frac{\text{Log Likelihood}(\text{Modelo})}{\text{Log Likelihood}(0)}$$

En palabras, es 1 menos la división entre el Log Likelihood del modelo estudiado, dividido por el Log Likelihood del modelo nulo.

Esta métrica indica que tanto mejoran los parámetros en el modelo estimado con respecto al modelo nulo, así, mientras mayor sea el pseudo Rho cuadrado, mejor será el modelo.

## Metodología

### Encuesta de satisfacción

Cuando un pasajero decide contactarse con el servicio al cliente de la DiDi, este genera un ticket, el cual al terminar de ser procesado, se cierra, y habiendo sucedido esto se envía una encuesta de servicio al cliente en la que puede evaluar su experiencia con el servicio al cliente de la aplicación de movilidad.

En esta encuesta el pasajero puede contestar las siguientes preguntas:

1. **Califica tu experiencia con nuestro servicio al cliente:** Aquí el pasajero puede evaluar el servicio con una nota de 1 a 5. Es desde esta pregunta que se desprende la métrica principal a ser estudiada, el CSAT (*Customer Satisfaction*), el cual está definido como:

$$CSAT = \left( \frac{\text{Número de encuestas con nota 4} + \text{Número de encuestas con nota 5}}{\text{Número de encuestas totales}} \right) \times 100$$

2. **¿Cuántas veces contactaste a servicio al cliente para resolver tu problema?:** En esta pregunta el pasajero tiene 3 opciones: una vez, varias veces, mi problema no se resolvió. Desde aquí se desprenden 2 métricas, muy relacionadas con el CSAT. El FCR (*First Contact Resolution*) y el RR (*Resolution Rate*), los cuales están definidos de la siguiente manera:

$$FCR = \left( \frac{\text{Número de encuestas respondidas con "Una vez"}}{\text{Número de encuestas totales}} \right) \times 100$$

$$RR = \left( \frac{\text{Número de encuestas respondidas con "Una vez"} + \text{Número de encuestas respondidas con "Varias veces"}}{\text{Número de encuestas totales}} \right) \times 100$$

3. **Cuadro de comentario libre:** El pasajero tiene un espacio donde puede escribir libremente comentarios acerca del servicio.

## **Muestra**

A partir de los resultados del modelo propuesto en este estudio, se espera definir lineamientos que permitan a los agentes de servicio al cliente aumentar la satisfacción del usuario, es por esto que uno de los principales factores que se consideraron al seleccionar los factores a considerar (y por consiguiente los datos que se recolectarán), es aquella información a la cual los agentes de servicio al cliente pueden acceder rápidamente durante un contacto con usuario.

Por ejemplo, si el modelo indicara que cuando un usuario tiene más de 0,5 viajes por día este tiene más posibilidades de dar una peor nota de satisfacción y por consiguiente se definen acciones para manejar a usuarios como estos, los agentes de servicio al cliente no pueden acceder fácilmente a esta información, puesto que durante el contacto, estos tendrían que calcular esta métrica (viajes por día) aumentando considerablemente el tiempo de contacto y resolución, y por tanto haciendo difícil definir acciones dirigidas a estos segmentos que no son fácilmente diferenciables.

Por el contrario, si el modelo indicara que cuando un usuario tiene más 100 viajes este tiene más posibilidades de dar una peor nota de satisfacción, para el agente del servicio al cliente es muy fácil ver esta información directamente en el sistema, sin necesidad de hacer ningún tipo de cálculo que pueda aumentar el tiempo de contacto y resolución.

Es por esto que para seleccionar los factores considerados en el modelo estos deben:

1. Estar disponibles.
2. Estar relacionados con la atención recibida o caracterización del usuario.
3. Ser fácilmente accesible para los agentes de servicio al cliente.

La muestra utilizada consiste en 3974 encuestas de satisfacción respondidas por usuarios pasajeros de todo Chile, durante los meses de enero y febrero del año 2021.

## Variables dependientes

En este estudio, se definieron dos modelos, uno con variable dependiente de satisfacción del consumidor y otro con la retención, los cuales se explicarán a continuación:

### Nivel de satisfacción

Como se mencionó anteriormente, la principal variable objetivo de este estudio es la satisfacción del consumidor, en la compañía estudiada, este índice es obtenido como CSAT. Es decir una variable entera ordinal, con valores desde el 1 hasta el 5, donde 1 indica máxima insatisfacción y 5 indica máxima satisfacción.

Nivel de satisfacción	Número de observaciones	Porcentaje del total
1	1236	31,10%
2	270	6,79%
3	399	10,04%
4	466	11,73%
5	1603	40,34%

**Tabla X:** Composición de la muestra por frecuencia de uso después del contacto con servicio al cliente.

### Variación de la frecuencia de uso

El nivel de uso de la aplicación después del contacto es la cantidad de viajes efectuados por un usuario en un periodo de 30 días después de su contacto con el servicio al cliente. A partir de estos viajes, se han definido 3 grupos de pasajeros:

1. Bajo uso: 0 viajes
2. Medio uso: Entre 1 y 5 viajes

### 3. Alto uso: 6 o más viajes

Estos grupos representan aproximadamente un tercio cada uno:

Nivel de uso después del contacto	Número de observaciones	Porcentaje del total de la muestra
Bajo uso después del contacto	1043	26.25%
Medio uso después del contacto	1597	40.19%
Alto uso después del contacto	1334	33.57%

**Tabla X:** Composición de la muestra por frecuencia de uso después del contacto con servicio al cliente.

El nivel de uso de la aplicación antes del contacto es la cantidad de viajes efectuados por un usuario desde el momento de la creación de su cuenta en la aplicación hasta el contacto con el servicio al cliente. Se han definido 3 grupos de pasajeros según su nivel de uso:

1. Bajo uso: 17 o menos viajes
2. Medio uso: Entre 18 y 59 viajes
3. Alto uso: 60 o más viajes

Estos grupos representan aproximadamente un tercio cada uno:

Nivel de uso antes del contacto	Número de observaciones	Porcentaje del total de la muestra
Bajo uso antes del contacto	1405	35,45%
Medio uso antes del contacto	1319	33,19%
Alto uso antes del contacto	1250	31,45%

**Tabla X:** Composición de la muestra por frecuencia de uso antes del contacto con servicio al cliente.

La variación en la frecuencia de uso está calculada de la siguiente forma.

Variación en la frecuencia de uso = Nivel de uso de la aplicación después - Nivel de uso de la aplicación antes

Por ejemplo: Un usuario con 23 viajes antes del contacto con el servicio al cliente tiene un nivel de uso medio con un valor de 2, si este usuario después del contacto con el servicio al cliente tiene 0 viajes, esto corresponde a un nivel de uso bajo, con un valor de 1. Por lo tanto la variación en la frecuencia de uso para este usuario en particular sería  $1 - 2 = -1$ .

Vemos entonces a partir de este ejemplo, que esta variable puede tomar valores enteros desde el -2 (Máxima variación negativa en la frecuencia de uso) hasta el 2 (Máxima variación positiva en la frecuencia de uso).

## **Variables independientes**

### Canal de contacto

Un usuario se puede contactar con el servicio al cliente mediante diversos canales, estos son:

1. FAQ: Mediante la sección de ayuda dentro de la aplicación
2. Phone: Llamando al número de teléfono del servicio al cliente
3. Email: Escribiendo al correo electrónico del servicio al cliente
4. Self Help: Ayuda automatizada, a la cual el usuario puede ingresar mediante la sección de ayuda de la aplicación.

Canal de contacto	Número de observaciones	Porcentaje del total
-------------------	-------------------------	----------------------



FAQ	1051	26,45%
Phone	335	8,43%
Email	197	4,96%
Self Help	2391	60,17%

**Tabla X:** Composición de la muestra por canal de contacto.

Para los modelos, se utilizó como base aquellas encuestas con tickets gestionados mediante Self Help.

#### Canal de respuesta de la encuesta

Los usuarios pueden responder la encuesta de satisfacción del cliente por dos medios:

1. Dentro de la aplicación
2. *Email*

Canal de respuesta	Número de observaciones	Porcentaje del total
Dentro de la app	2720	31,56%
<i>Email</i>	1254	68,44%

**Tabla X:** Composición de la muestra por canal de respuesta.

Para los modelos, se utilizó como base aquellas encuestas respondidas mediante el canal de *Email*.

#### Razón de contacto

Al contactar los usuarios al servicio al cliente, los agentes categorizan el ticket generado dentro de 4 grandes grupos de problemas:

1. Problemas con un viaje
2. Problemas con el método de pago

3. Problemas con una promoción
4. Información de la aplicación

Razón de contacto	Número de observaciones	Porcentaje del total
Problemas con un viaje	3401	85,58%
Problemas con el método de pago	390	9,81%
Problemas con una promoción	73	1,84%
Información de la aplicación	110	2,77%

**Tabla X:** Composición de la muestra por razón de contacto.

Para los modelos, se utilizó como base aquellas encuestas con tickets con razón de contacto problemas con un viaje.

#### Número de contactos para la resolución

Como se mencionó anteriormente al describir la encuesta, una de las preguntas es “¿Cuántas veces contactaste a servicio al cliente para resolver tu problema?” dónde los usuarios tienen la posibilidad de responder una de 3 opciones.

1. Una vez
2. Varias veces
3. Mi problema no se resolvió

Número de contactos para la resolución	Número de observaciones	Porcentaje del total
Una vez	2218	55,81%
Varias veces	957	20,11%
Mi problema no se resolvió	799	20,11%

**Tabla X:** Composición de la muestra por número de contactos para la resolución.

Para los modelos, se utilizó como base aquellas encuestas mencionando que su problema no fue resuelto.

### Línea de negocio

Dentro de la empresa de *ridesharing* analizada existen 4 líneas de negocio diferentes:

1. Express: Línea de negocio normal.
2. Taxi: Línea de negocio para Taxistas
3. Entrega: Línea de negocio para entrega de paquetes y no para transporte de pasajeros
4. Economy: Línea de negocio económica, con precios reducidos y menores barreras de entrada para los socios conductores.

Línea de negocio	Número de observaciones	Porcentaje del total
Express	2565	64,54%
Taxi	961	24,18%
Economy	204	5,13%
Entrega	244	6,14

**Tabla X:** Composición de la muestra por línea de negocio.

Taxi, Entrega y Economy son reflejadas en el modelo como variables binarias, utilizando Express como opción base.

### Tiempo de resolución de ticket

Cuánto tiempo pasó desde el momento que el usuario ingresó su duda/reclamo y esta fue resuelta, esta variable fue categorizada en 3 variables binarias:

1. Corto tiempo: Menos de 24 horas.
2. Mediano tiempo: Desde 24 horas hasta 50 horas.
3. Largo tiempo: Más de 50 horas.

Tiempo de resolución de ticket	Número de observaciones	Porcentaje del total
Corto tiempo	371	9,34%
Mediano tiempo	2120	53,35%
Largo tiempo	1483	37,32%

**Tabla X:** Composición de la muestra por tiempo de resolución de ticket.

Para el modelo se utilizó como base aquellos usuarios que obtuvieron una respuesta en un mediano tiempo.

#### Etapas a la cual llegó el viaje

Si el usuario se contacta con dudas/reclamos acerca de un viaje, este viaje se adjunta al ticket, este viaje puede no haberse realizado (en caso de que el usuario no llame con razones relacionadas a un viaje), pueden no haberse completado (En caso de que el conductor o el pasajero haya cancelado el viaje antes de finalizar), o puede haber terminado normalmente.

1. Viaje cancelado o sin viaje: El pasajero o el conductor cancelaron el viaje antes de que este terminara de llevarse a cabo o el pasajero no llevó a cabo ningún viaje.
2. Viaje normal: El viaje transcurrió y llegó a su fin normalmente.

Estado del viaje	Número de observaciones	Porcentaje del total
Finalizado	2134	53,70%
Cancelado o sin viaje	1840	46,30%

**Tabla X:** Composición de la muestra por estado del viaje.

Esta es una variable binaria que toma el valor de 1 cuando el viaje finalizó normalmente y 0 cuando fue cancelado o no se realizó el viaje.

Tiempo de respuesta de la encuesta

Cuánto tiempo pasó entre que se cerró el ticket del usuario y este respondió la encuesta, esta variable fue categorizada en 3 variables binarias:

1. Corto tiempo: Menos de una hora.
2. Mediano tiempo: Desde una hora hasta un día.
3. Largo tiempo: Más de un día.

Tiempo de respuesta de la encuesta	Número de observaciones	Porcentaje del total
Corto tiempo	1918	48,26%
Mediano tiempo	1618	40,71%
Largo tiempo	438	11,02%

**Tabla X:** Composición de la muestra por tiempo de respuesta de la encuesta.

Para el modelo se utilizaron aquellos usuarios que respondieron la encuesta en corto tiempo como base.

Presencia de comentario

Si en la encuesta de satisfacción al cliente, el usuario decidió dejar un comentario o no.

Presencia de comentario	Número de observaciones	Porcentaje del total
-------------------------	-------------------------	----------------------

Si	2383	59,96%
No	1591	40,04%

**Tabla X:** Composición de la muestra por presencia de comentario en la encuesta.

Esta es una variable binaria que toma valor de 1 cuando sí hay un comentario y 0 cuando no hay un comentario.

### Zona geográfica del usuario

Zona geográfica en la cual está localizado el usuario al momento de crear un ticket de soporte.

<b>Zona</b>	<b>Número de observaciones</b>	<b>Porcentaje del total</b>
Norte	213	5,36%
Centro norte	602	15,15%
Región metropolitana	2651	66,71%
Centro sur	268	6,74%
Sur	240	6,04%

**Tabla X:** Composición de la muestra por zona geográfica del usuario.

Se utilizó como base aquellas encuestas de usuarios localizados en la región metropolitana.

### Nivel de satisfacción

Para el modelo de retención se utiliza el nivel de estrellas de la satisfacción del consumidor como una variable explicativa.

<b>Nivel de satisfacción</b>	<b>Número de observaciones</b>	<b>Porcentaje del total</b>
------------------------------	--------------------------------	-----------------------------

1	1236	31,10%
2	270	6,79%
3	399	10,04%
4	466	11,73%
5	1603	40,34%

**Tabla X:** Composición de la muestra por nivel de satisfacción.

Esta variable es una variable continua.

## Herramientas

Para la obtención y descarga de los datos, se utilizó el sistema interno de monitoreo de DiDi, el cual está dividido en dos:

1. CS Global: Sistema de gestión y registro de tickets de servicio al cliente.
2. Magellan: Sistema de registro de viajes.

Para la preparación, cruce y limpieza de la base de datos, se utilizó R. Esto consistió en:

- Cruce de datos:
  - Entre las bases de datos de los viajes y la base de datos del sistema de servicio al cliente, es decir los datos correspondientes a los tickets.
- Limpieza de la base de datos:
  - Eliminación de observaciones deficientes, es decir, observaciones las cuales no tengan datos para todas las columnas necesarias.
  - Eliminación de observaciones repetidas.
- Preparación de la base de datos:
  - Creación de variables categóricas a partir de variables continuas.
  - Creación de variables binarias a partir de variables continuas.

Para la estimación de los modelos se utilizó R, con el paquete *Ordinal*.

## Resultados

### Modelo de satisfacción del consumidor

En la tabla X, X y X a continuación, se observan los resultados del modelo *Ordered Probit* para explicar la satisfacción del consumidor, el modelo considera todos aquellos factores que fueron significativos al 90% o más.

Variable	Estimación	Error Estándar	Estadístico-t	Valor p
Canal de contacto				
Phone	-0,22999	0,10331	-2,226	0,0260
FAQ	-0,54526	0,08294	-6,575	0,0000
Email	-0,27872	0,11238	-2,480	0,0131
Canal de respuesta				
Dentro de la aplicación	-0,26592	0,04175	-6,470	0,0000
Zona geográfica				
Centro Sur	0,018439	0,07574	2,434	0,0149
Centro Norte	0,013146	0,05369	2,448	0,0143
Número de contactos para la resolución				
Una vez	2,00513	0,05720	35,054	0,0000
Varias veces	1,04182	0,06056	17,203	0,0000
Presencia de comentario	-0,39217	0,03961	-9,901	0,0000
Tiempo de resolución del ticket				
Corto	0,43995	0,08115	5,421	0,0000
Número de parámetros		10		
Número de observaciones		3974		



**Tabla X:** Resultados del modelo *Ordered Probit* para la satisfacción del consumidor.

Umbral	Estimación	Error estándar	Estadístico-t
1   2	0,41830	0,06203	6,744
2   3	0,68479	0,06241	10,973
3   4	1,06182	0,06343	16,741
4   5	1,47502	0,06476	22,775

**Tabla X:** Resultados de los umbrales del modelo *Ordered Probit* para la *satisfacción* del consumidor.

Métrica de bondad de ajuste	Valor
<i>Log Likelihood</i> (Modelo)	-4585.93
<i>Log Likelihood</i> (0)	-5540,86
Pseudo - Rho cuadrado	17,234%
Prueba Chi Cuadrado	0,000

**Tabla X:** Resultados de las métricas de bondad de ajuste del modelo *Ordered Probit* para la satisfacción del consumidor.

## Interpretación

### Bondad de ajuste

La Prueba Chi Cuadrado muestra que el Log Likelihood del modelo propuesto es significativamente diferente al Log Likelihood del modelo sin variables explicativas con una significancia por sobre el 99% dado el valor p de 0,000 indicado en la tabla X

Este modelo arrojó un Pseudo - Rho Cuadrado de un 17,24%.

### Umbrales

Todos los umbrales del modelo son significativos al 95%, como se puede apreciar en la Tabla X, donde todos los estadísticos t son superiores a 1,96, lo cual indica que estos son significativos.

### Factores

En principio, se observa en la Tabla X la ausencia de factores anteriormente mencionados, esto es debido a que ninguno de los factores ausentes fue significativo al 90% que se utilizó como umbral para su permanencia en el modelo.

Se ve que por ejemplo factores tales como la razón de contacto, la línea de negocios, la etapa a la cual llegó el viaje, el tiempo de respuesta de la encuesta, no fueron significativos al 90% para explicar la satisfacción del consumidor.

1. Canal de contacto: Este factor fue significativo, indicando que aquellos contactos provenientes de los canales *Phone*, *FAQ* y *Email*, eran significativamente diferentes del canal base, es decir *Self Help*.

Se observa que todos los factores poseen coeficientes negativos en este modelo, esto quiere decir que, a diferencia del canal base, un contacto entrante por estos canales es más probable que tenga una satisfacción menor. Así también se observa que dentro de los canales significativos aquel con menor coeficiente es el canal FAQ.

Esto puede explicarse porque el canal de FAQ es aquel con menos interacción directa entre usuario y agente, debido a que los agentes tienen scripts definidos de respuesta que deben seguirse muy de cerca, esto da una sensación de atención poco personalizada.

Se observa también que el canal con el coeficiente negativo menor es el canal Phone, naturalmente, los usuarios sienten que interactuando directamente con

un agente, la atención es mucho más personalizada y sus problemas son escuchados directamente.

Los coeficientes de estos factores son negativos, debido a la diferencia en *Resolution Rate* de los diferentes canales, dado que los canales en los que se interactúa con un agente los lineamientos de compensaciones son mucho más estrictos que aquellos programados en el canal *Self Help*:

Canal	Problema/duda resuelto(a)	Problema/duda no resuelto(a)
Email	65,99%	34,01%
FAQ	78,88%	21,12%
Phone	71,94%	28,06%
Self Help	82,60%	17,40%

**Tabla X:** *Resolution rate* por canal de contacto

- Canal de respuesta: Para que un usuario sea capaz de responder la encuesta de satisfacción mediante la aplicación, este tiene que tener activadas las notificaciones provenientes de la aplicación de DiDi, estas pueden llegar en un momento inoportuno o molesto para el usuario, mientras que aquellos usuarios que responden por Email, en muchos casos no reciben notificaciones invasivas y solamente ven que hay una encuesta disponible al momento de revisar su correo electrónico, es decir, en sus propios términos. En resumen, en lo que a canal de respuesta respecta, los usuarios se pueden agrupar en 2 segmentos (No mutuamente excluyentes, ya que todos los usuarios tienen la posibilidad de responder por ambos medios): Usuarios activamente notificados para responder la encuesta (Notificación de la aplicación de DiDi) y usuarios pasivamente notificados para responder la encuesta (Email).

DiDi

¿Nos podrías contar sobre tu experiencia con el área de soporte de DiDi?

\* 1. Califica tu experiencia con nuestro servicio al cliente

★ ★ ★ ★ ★

\* 2. ¿Cuántas veces contactaste a servicio al cliente para resolver tu problema?

☐ Una vez

☐ Varias veces

☐ Mi problema no se resolvió.

**Figura X:** Encuesta de satisfacción al cliente dentro de la app.

3. Zona geográfica: Se observa que aquellos usuarios en el centro norte y centro sur tienen una mayor probabilidad de tener una satisfacción más alta que la zona base (Región Metropolitana), ambas con una significancia mayor al 98%, sin embargo no fue posible hallar una causa raíz identificable con los datos disponibles actualmente del porque estas regiones puedan tener una mayor satisfacción con el servicio en comparación a la zona base. Esto amerita un estudio más a fondo para encontrar la causa raíz de el porque estas zonas poseen usuarios con mayor probabilidad de tener una satisfacción más alta en comparación a la Región Metropolitana.
4. Número de contactos para la resolución: Se observa que este factor está dentro de los más significativos dentro de los factores considerados en el modelo, es decir, aquellos usuarios que necesitaron contactarse una o varias veces para resolver su problema, tienen una diferencia significativa en términos de satisfacción con aquellos usuarios a los que su problema no se resolvió, situación que fue usada como base.

En el modelo, el coeficiente  $\beta$  tanto de usuarios que se contactaron una vez, como aquellos que se contactaron varias veces para resolver su problema, es

positivo, esto indica, que en términos generales, los usuarios que pudieron resolver su problema al contactarse una o más veces con el servicio al cliente, tienen una satisfacción mayor que aquellos que no lograron resolver su problema. Además, el coeficiente de los usuarios que lograron resolver problemas en el primer contacto es casi el doble que el de los usuarios que tuvieron que contactarse más veces. Por lo tanto, la satisfacción de los usuarios que lograron resolver su problema en el primer contacto es muy superior a los demás, como se puede ver en la tabla X.

Número de contactos para la resolución	Porcentaje de usuarios que marcaron 4 o 5 estrellas en la encuesta de satisfacción
Mi problema no se resolvió	4,51%
Varias veces	35,42%
Una vez	76,38%

**Tabla X:** Usuarios satisfechos (Aquellos que marcaron 4 o 5 estrellas) por número de contactos con el servicio al cliente para resolver su problema.

5. Presencia de comentario: El que un usuario escribiera un comentario en la encuesta de satisfacción, probó tener una diferencia muy significativa con aquellos usuarios que decidieron no dejar un comentario.

Se observa que el coeficiente es negativo, esto proporciona una información muy valiosa: los usuarios insatisfechos son mucho más propensos a dejar un comentario que aquellos usuarios satisfechos.

Tipo de usuario	Porcentaje de encuestas con comentarios
Usuarios satisfechos (4 o 5 estrellas)	42,85%
Usuarios insatisfechos (1, 2 o 3 estrellas)	57,15%

**Tabla X:** Porcentaje de encuestas con comentarios por tipo de usuario.

6. Tiempo de resolución de ticket: Como se explicó anteriormente el tiempo de resolución de resolución de ticket se dividió entre corto tiempo, mediano tiempo y largo tiempo. Para propósitos del modelo se utilizaron como base aquellos usuarios que tuvieron una resolución de ticket en un mediano tiempo . El modelo arrojó que no existe una diferencia significativa al 90% entre los usuarios que tuvieron su ticket resuelto en un mediano tiempo y aquellos que tuvieron la resolución de su ticket en un largo tiempo, sin embargo, aquellos usuarios a los cuales se les resolvió un ticket en un corto plazo, mostraron una diferencia altamente significativa (Por sobre el 99%) con aquellos usuarios que se les resolvió su ticket en un mediano plazo. El coeficiente de este factor es positivo, indicando que por regla general, aquellos usuarios a los cuales se les resuelven sus tickets más rápido tienden a tener una satisfacción mucho mayor que aquellos usuarios cuyos tickets son resueltos en un mediano o largo plazo.

La evidencia actual respalda lo observado, donde 2 de las 5 razones con mayor correlación con la satisfacción tienen que ver la rapidez con la respuesta a las dudas o quejas que un usuario pueda tener (Press, Irwin; Ganey, Rodney F; Hall, Melvin F.; American Bankers Association. ABA Banking Journal; New York Tomo 89, N.º 9, (Sep 1997): 73-74.)

### **Modelo de retención: Cambio en el nivel de uso**

En la tabla X, X y X a continuación, se observan los resultados del modelo *Ordered Probit* para explicar el cambio en el nivel de uso de la aplicación, el modelo considera todos aquellos factores que fueron significativos al 90% o más.

Variable	Estimación	Error Estándar	Estadístico-t	Valor p
Tiempo para responder				
Largo	0,10473	0,05460	1,918	0,055
Canal de respuesta				
Dentro de la aplicación	0,12181	0,03720	3,274	0,0010
Zona geográfica				

Sur	-0,42117	0,07207	-5,844	0,0000
Centro Norte	-0,10740	0,04832	-2,223	0,0262
Norte	0,28981	0,07600	3,814	0,0001
Razón de contacto				
Información	-0,22561	0,10465	-2,156	0,0311
Número de contactos para la resolución				
Una vez	0,15075	0,05557	2,713	0,0067
Varias veces	0,11673	0,05387	2,167	0,0303
Tiempo de resolución				
Corto	0,05944	0,03550	1,675	0,0940
Etapas a la cual llegó el viaje				
Finalizado	-0,11282	0,03411	-3,307	0,0009
Nivel de satisfacción	0,03768	0,01261	2,987	0,0028
<b>Número de parámetros</b>		10		
<b>Número de observaciones</b>		3974		

**Tabla X:** Resultados del modelo *Ordered Probit* para el cambio en el nivel de uso.

Umbral	Estimación	Error estándar	Estadístico-t
-2   -1	-1,69457	0,06227	-27,21
-1   0	-0,64224	0,05328	-12,05
0   1	0,76466	0,05360	14,27
1   2	1,89902	0,06055	31,36

**Tabla X:** Resultados de los umbrales del modelo *Ordered Probit* para el cambio en el nivel de uso.

Métrica de bondad de ajuste	Valor
<i>Log Likelihood</i> (Modelo)	-4883,68
<i>Log Likelihood</i> (0)	-4942,63
Pseudo - Rho cuadrado	1,192%
Prueba Chi Cuadrado	0,000

**Tabla X:** Resultados de las métricas de bondad de ajuste del modelo *Ordered Probit* para el cambio en el nivel de uso.

## Interpretación

### Bondad de ajuste

El Log Likelihood del modelo propuesto es significativamente diferente al Log Likelihood del modelo sin variables explicativas, como es indicado por la Prueba Chi Cuadrado, la cual muestra una significancia por sobre el 99% dado el valor p de 0,000 indicado en la tabla X.

Este modelo arrojó un Pseudo - Rho Cuadrado de un 1,192%.

### Umbrales

Todos los umbrales del modelo son significativos al 95%, como se puede apreciar en la Tabla X, donde todos los estadísticos t son superiores a 1,96, lo cual indica que estos son significativos.

### Factores

1. Tiempo para responder: Aquellos usuarios que respondieron su encuesta en más de un día tuvieron una diferencia positiva y significativa en el uso posterior de la aplicación con respecto a aquellos usuarios que respondieron la encuesta en menos de una hora. Este factor, fue significativo al 90%, lo cual si bien denota cierta correlación entre la variable dependiente y este factor, no es



significativa al 95% como se espera de una variable altamente correlacionada. De todos modos una explicación a la significancia por sobre un 90%, es que los usuarios que responden la encuesta de satisfacción más de un día después de que esta es enviada demuestran un uso de la aplicación más frecuente. Por ejemplo, un usuario que responde la encuesta 2 días después de recibirla es un usuario que tiene en mente a DiDi constantemente o que ha vuelto a usar el servicio.

2. Canal de respuesta: Aquellos usuarios que respondieron la encuesta por medio de la aplicación (y no por el correo), tienen una mayor probabilidad de aumentar el nivel de uso del servicio. Esto podría explicarse porque estos usuarios, a diferencia de aquellos que responden por correo, son usuarios que ya están lo suficientemente familiarizados con la aplicación.

Es importante mencionar también, que para que un usuario se entere de que ha recibido la encuesta de satisfacción por la aplicación, este tiene que tener las notificaciones de la aplicación de DiDi activadas y permitidas en su dispositivo, por tanto, a estos usuarios no solo tendrían una visibilidad mucho mayor de las encuestas de satisfacción, sino también de las promociones y descuentos que DiDi hace de cuando en cuando, haciendo muy probable así que estos usuarios vuelvan a usar el servicio.

3. Zona geográfica: Se observan en los resultados que la variable binaria correspondiente a la zona sur del país tiene un parámetro negativo y con una significancia superior al 99%, esto significa que aquellos pasajeros que se contactaron con el servicio al cliente tienen una menor probabilidad de aumentar el nivel de uso de la aplicación. Esto se puede explicar ya que las comunas que agrupan la mayor cantidad de población de la zona sur (y por consiguiente demanda) tales como Puerto Montt, Osorno, Temuco, Concepción, Talcahuano, Los Ángeles y Chillán, sufrieron de cuarentenas que restringieron la movilidad durante los meses estudiados. La línea de negocios Express fue desactivada en las comunas que sufrían de cuarentenas, siguiendo así los lineamientos de las autoridades.

Así, los pasajeros en estas ciudades tenían menores oportunidades en términos de tiempo y oferta para poder mantener o aumentar el nivel de uso del servicio.

4. Razón de contacto: En la tabla X se indica que la variable binaria que toma valor 1 para aquellos usuarios que se contactaron para preguntar información básica acerca de la aplicación está acompañada por un parámetro negativo, es

decir, aquellos usuarios que se contactan por esta razón tienen mayores posibilidades de disminuir el uso de la aplicación en el futuro.

Esto puede deberse a que dentro de esta categoría se clasifican las preguntas más básicas con respecto al funcionamiento y características de la aplicación, naturalmente, aquellos usuarios que consultan por esto tienen menos conocimiento de la aplicación, de la marca y de la industria del *ridesharing* en general, esto debido a la naturaleza de las preguntas. Es por esto que estos usuarios podrían tener una menor probabilidad de aumentar el uso de un servicio del cual no tienen suficiente conocimiento, y por consiguiente una menor o nula lealtad a la marca.

5. Número de contactos para la resolución: Como se puede apreciar en la Tabla X, las variables que describen a aquellos usuarios tuvieron una resolución de su problema o consulta, tiene una constante positiva, esto indica que estos usuarios tienen una mayor probabilidad de aumentar el uso del servicio luego de un contacto con el servicio al cliente en comparación a aquellos que no recibieron una solución a su problema o duda, los cuales fueron utilizados como base para este factor. Además, la constante de la variable que describe aquellos usuarios que recibieron una solución en el primer contacto, es mayor a aquella que describe los usuarios que tuvieron que contactarse más de una vez para resolver su consulta o problema, indicando que hay dos aspectos que si se cumplen aumentan la probabilidad de que un usuario aumente el nivel de uso en el futuro:
  - a. Que su problema o duda sea resuelta.
  - b. Que su problema o duda sea resuelta en el primer contacto con el departamento de servicio al cliente.
6. Tiempo de resolución: Como se mencionó anteriormente, Jahanshahi, Gashti, Mirdamadi, Nawaser, y Khaksar. (2011) identificaron la calidad de la atención recibida por parte del servicio al cliente como uno de los factores que eventualmente afecta en la retención, por otra parte, la evidencia demuestra que el tiempo es uno de los factores determinantes tras la satisfacción de un usuario (Press, Irwin; Ganey, Rodney F; Hall, Melvin F.; American Bankers Association. ABA Banking Journal; New York Tomo 89, N.º 9, (Sep 1997): 73-74.). El resultado aquí observado, es decir, un coeficiente positivo para aquellos usuarios que tuvieron un tiempo de resolución corto, se condice con la evidencia antes presentada, ya que este coeficiente implica que estos usuarios

tienen una mayor probabilidad de aumentar el uso futuro de la aplicación en comparación al grupo de usuarios usado como base, aquellos que tuvieron un tiempo de resolución medio.

7. Estado del viaje: Es de destacar que aquellos usuarios que reportan un viaje finalizado tienen una mayor probabilidad de reducir el uso de la aplicación el futuro que aquellos usuarios que no reportaron un viaje o reportaron un viaje no finalizado, basándose en los hallazgos de Jahanshahi, Gashti, Mirdamadi, Nawaser, y Khaksar. (2011), aquí se evidencia que la calidad del servicio otorgado no sea la óptima y que por lo tanto esto esté afectando el nivel de uso posterior al contacto con el servicio al cliente de los pasajeros. Si bien esto va más allá del foco de un departamento de servicio al cliente, es necesario tener presente la gran influencia de la experiencia de los usuarios a bordo del vehículo como uno de los mayores factores que afectan la retención.
8. Nivel de satisfacción: Finalmente, ya que la significancia de este factor está por sobre el 99%, fue posible confirmar una de las principales hipótesis de este estudio: Que un nivel de satisfacción alto de un usuario con el departamento de servicio al cliente de DiDi tiene una influencia positiva sobre el posterior nivel de uso de la aplicación.

## Conclusiones y recomendaciones

A lo largo del estudio se ha demostrado la importancia que tiene el proveer una atención eficaz y eficiente a los usuarios pasajeros de una empresa de *ridesharing*, y para poder lograr este objetivo, es necesario enfocarse en los factores que explican significativamente las elecciones de los usuarios pasajeros de DiDi.

Un factor común tanto en el modelo de satisfacción como en el modelo de retención, es el número de contactos necesarios para la resolución del problema o duda del usuario: Los usuarios quieren sus problemas resueltos y al primer contacto, esto sirve de guía para enfocar los esfuerzos a:

1. Investigar cuales son las soluciones que los usuarios realmente buscaban encontrar, evaluar el impacto y la factibilidad que tendría implementar estas soluciones e implementar aquellas que no tengan un impacto insignificante o una factibilidad muy baja.
2. Redefinir los procesos que requieran dos o más contactos para resolver la problemática del usuario, y de ser posible resolver estas consultas en el primer contacto.

Además es de destacar, que los usuarios que responden las encuestas de satisfacción mediante la aplicación tienen una satisfacción considerablemente menor a aquellos que eligen responderlas por Email, se debería considerar mapear un “viaje del cliente” a través de estos dos canales de respuesta de la encuesta, para así confirmar o rechazar la teoría aquí presentada, de ser confirmada, es necesario enfocar los esfuerzos en desarrollar notificaciones menos invasivas e idear nuevos modos de incentivar las respuestas a estas encuestas.

Por otra parte, se requiere un mayor estudio de la experiencia a bordo, para explicar la causa raíz de él porque los usuarios pasajeros de la aplicación son más propensos reducir su uso luego de reportar un viaje finalizado en comparación a aquellos que no han finalizado el viaje, es muy posible que hayan falencias críticas en cuanto al servicio otorgado que necesitan ser analizadas para su posible solución, lo cual, va más allá del rol del departamento de servicio al cliente.

Por otro lado, es necesario un estudio más a fondo de el porqué los usuarios del centro norte y centro sur de Chile tienen una mayor probabilidad de tener mayor satisfacción que los usuarios de la región metropolitana. Encontrar las razones de que es lo que está provocando esto, y de ser así, qué es lo que el departamento de servicio al cliente está haciendo bien puede ser crucial para expandir estas iniciativas al resto del país y por consiguiente una mayor satisfacción.

Finalmente, tener en consideración que el éxito que empresas como Amazon han tenido gracias a su enfoque *customer centric* (Baboolal-Frank, 2021), puede ser replicado en la industria del *ridesharing*, ya que ha quedado demostrado que la influencia de la satisfacción del usuario con el servicio al cliente tiene un impacto directo y positivo sobre el posterior nivel de uso de la aplicación, por lo tanto centrarse en otorgar el mejor servicio posible en cuanto a resolución de problemas y dudas debería ser uno de los ejes principales para poder aumentar el nivel de uso de los usuarios y por consiguiente su lealtad.