

EDA - Ironhack Payments

1. Introduccion

IronHack Payments, un pionero en el ámbito de los servicios financieros, ha revolucionado el concepto de adelantos de efectivo desde su lanzamiento en 2020. En un esfuerzo por optimizar continuamente sus servicios y comprender mejor el comportamiento de sus usuarios, IronHack Payments ha iniciado un proyecto de análisis de cohortes.

Este proyecto implica la realización de un análisis exhaustivo de las cohortes de usuarios, segmentadas según el mes en que realizaron su primer adelanto de efectivo. La investigación seguirá la evolución de cuatro métricas clave mes a mes, permitiendo a IronHack Payments obtener insights profundos sobre la dinámica de sus clientes y la eficacia de sus servicios.

2. Descripción del conjunto de datos

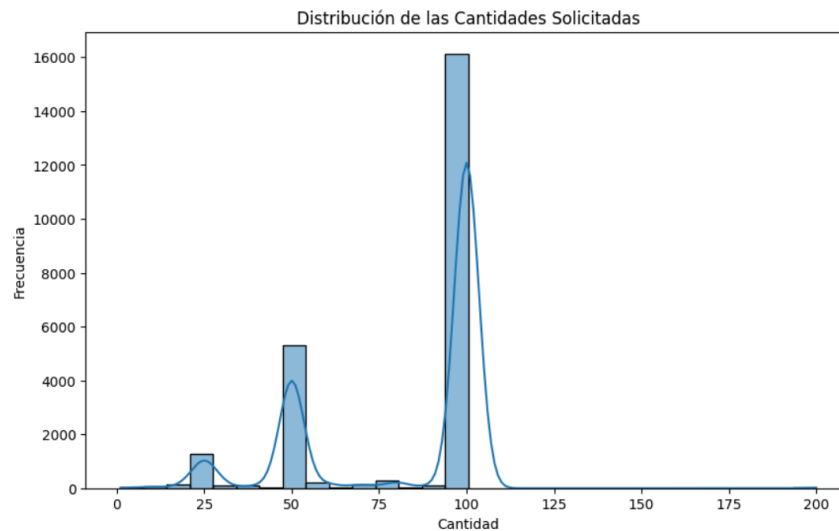
Los datos analizados provienen de dos archivos principales:

- Solicitudes de Efectivo (extract - cash request - data analyst.csv);
- Tarifas Asociadas (extract - fees - data analyst - .csv);
- El archivo Lexique - Data Analyst.xlsx proporciona el significado de cada columna en los conjuntos de datos.

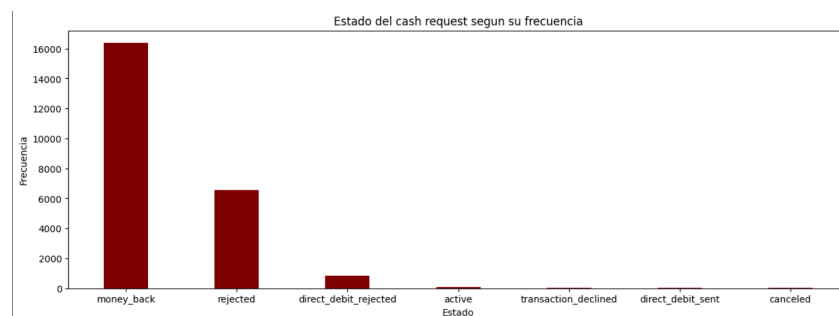
3. Limpieza y Transformación de Datos

Se realizaron las siguientes transformaciones y limpiezas de datos:

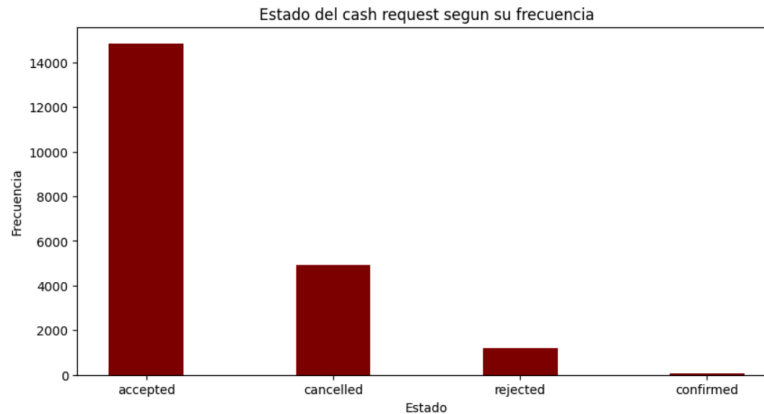
- Conversión de columnas de fecha a formato datetime.
- Imputación de valores faltantes en user_id utilizando deleted_account_id cuando user_id estaba vacío.
- Eliminación de duplicados y columnas innecesarias en el conjunto de datos de tarifas.
- Unión de los conjuntos de datos de solicitudes de efectivo y tarifas utilizando el identificador id.
- Las cantidades solicitadas mayoritariamente se agrupan en 3 tipos: alrededor de los 25 euros, de los 50 euros y de los 100 euros. Por esta razón, se han agrupado los datos en estas categorías que facilitan el análisis.



- Al analizar el estado de los cash request, la mayoría fueron reembolsados o directamente rechazados. La alta proporción de solicitudes reembolsadas sugiere que la mayoría de los usuarios cumplen con sus compromisos de reembolso, lo que es positivo para la salud financiera del servicio. A su vez, una alta tasa de rechazos también podría reflejar problemas o confusiones en el proceso de solicitud que llevan a los usuarios a no cumplir con los requisitos necesarios. Esto podría ser un área de mejora para facilitar el proceso y reducir la tasa de rechazos injustificados.



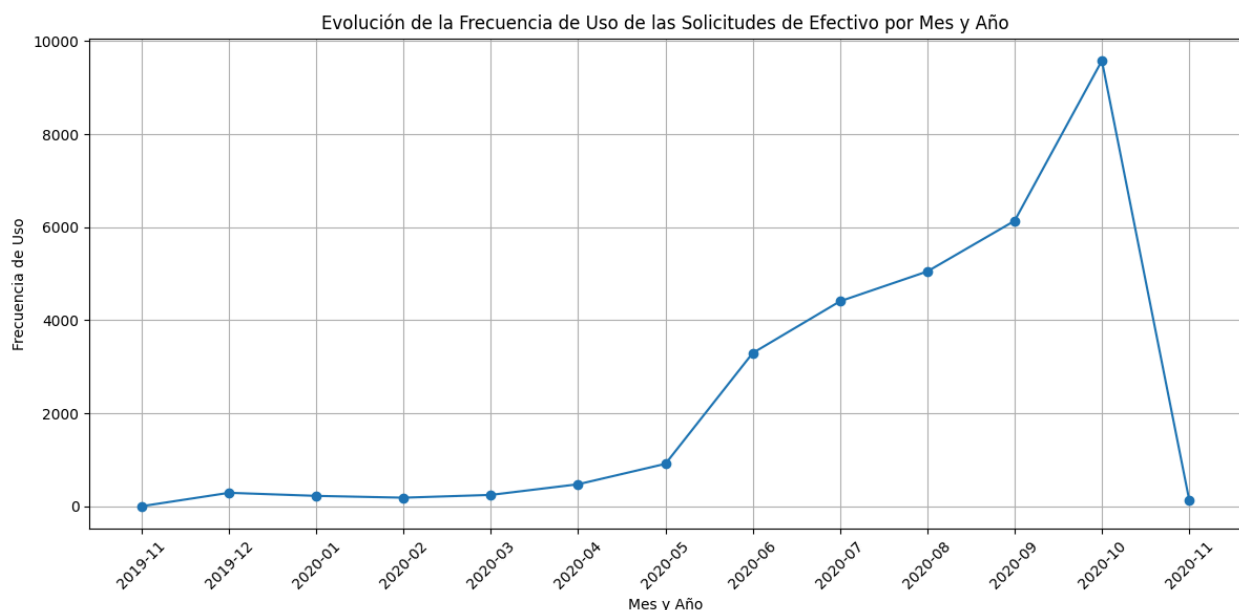
- Una alta proporción de comisiones aceptadas indica que el sistema de cobro es eficiente y efectivo. Los usuarios están cumpliendo con sus obligaciones de pago, lo que es un buen indicativo de la salud financiera del servicio.



4. Conclusiones

Del análisis realizado pueden deducirse las siguientes conclusiones:

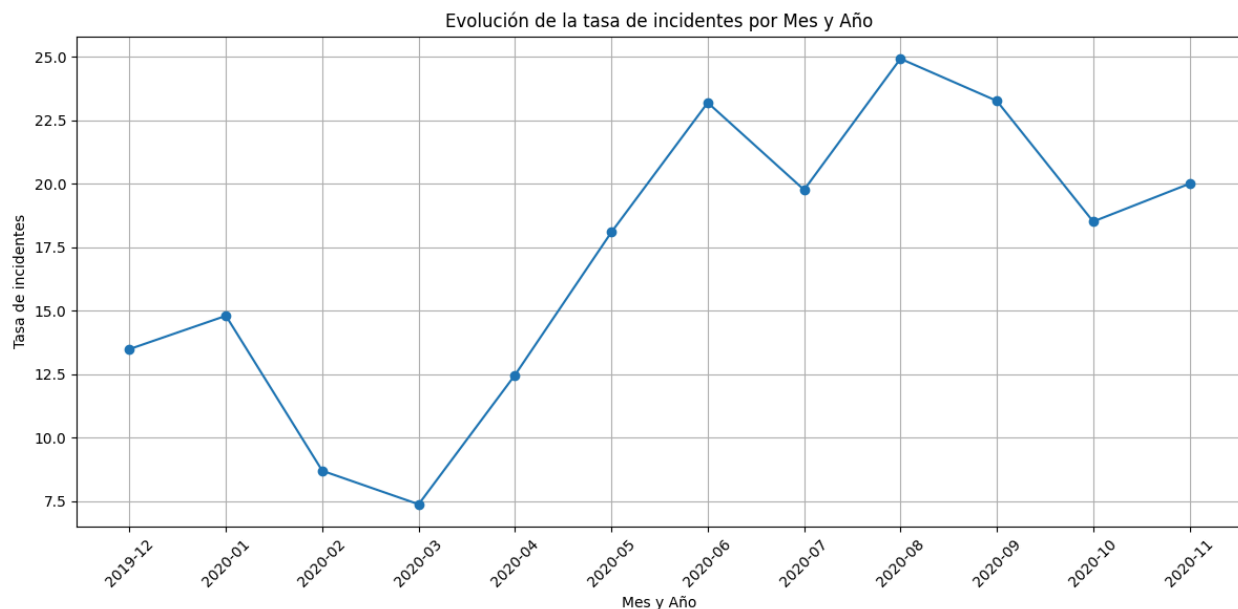
- **Frecuencia de uso del servicio:** los datos indican variaciones significativas en la frecuencia de uso del servicio entre los diferentes meses y años. En particular, se observa una tendencia creciente a partir de mayo de 2020 y un pico en el mes de octubre del 2020, cuando se alcanzaron las 9581 utilizations.



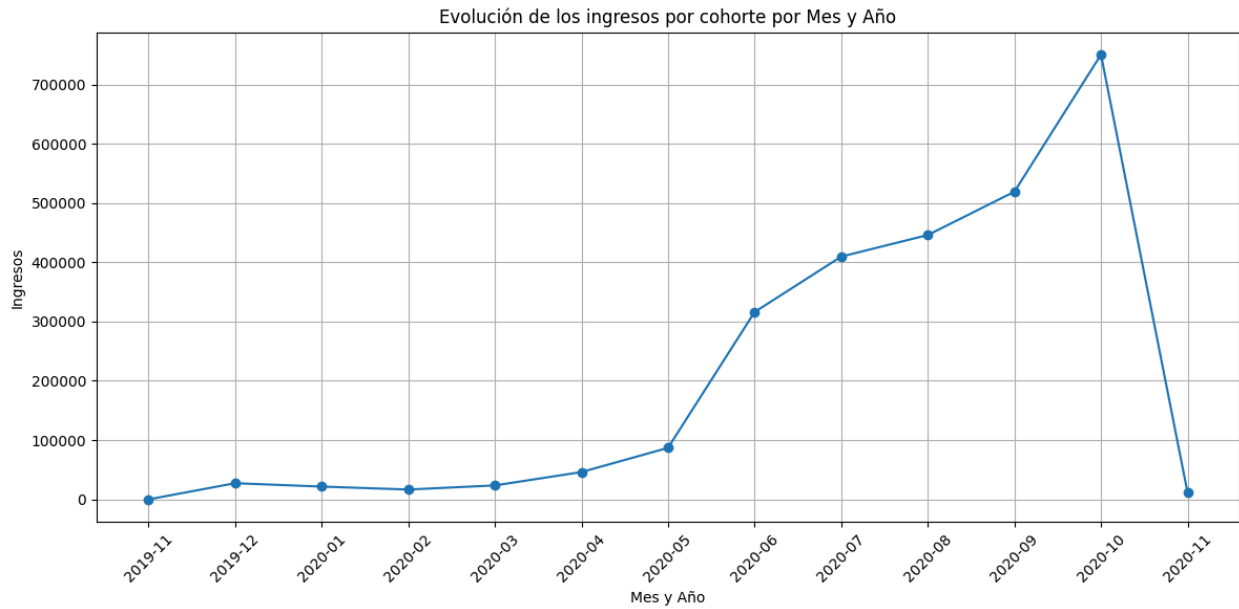
- **Ratio de incidentes:** se observa que el ratio de incidentes varía entre diferentes cohortes, lo cual puede indicar áreas de mejora en el servicio. A partir de marzo de 2020, observamos un notable aumento en la tasa de incidentes. Este incremento es coherente con el aumento en la frecuencia de uso del servicio durante el mismo período. Sin embargo, es importante destacar que el incremento en

la tasa de incidentes es desproporcionadamente mayor en comparación con el aumento en la frecuencia de uso.

En otras palabras, la proporción de incidentes está aumentando, lo que sugiere que el servicio está experimentando más problemas por solicitud de efectivo en comparación con períodos anteriores.

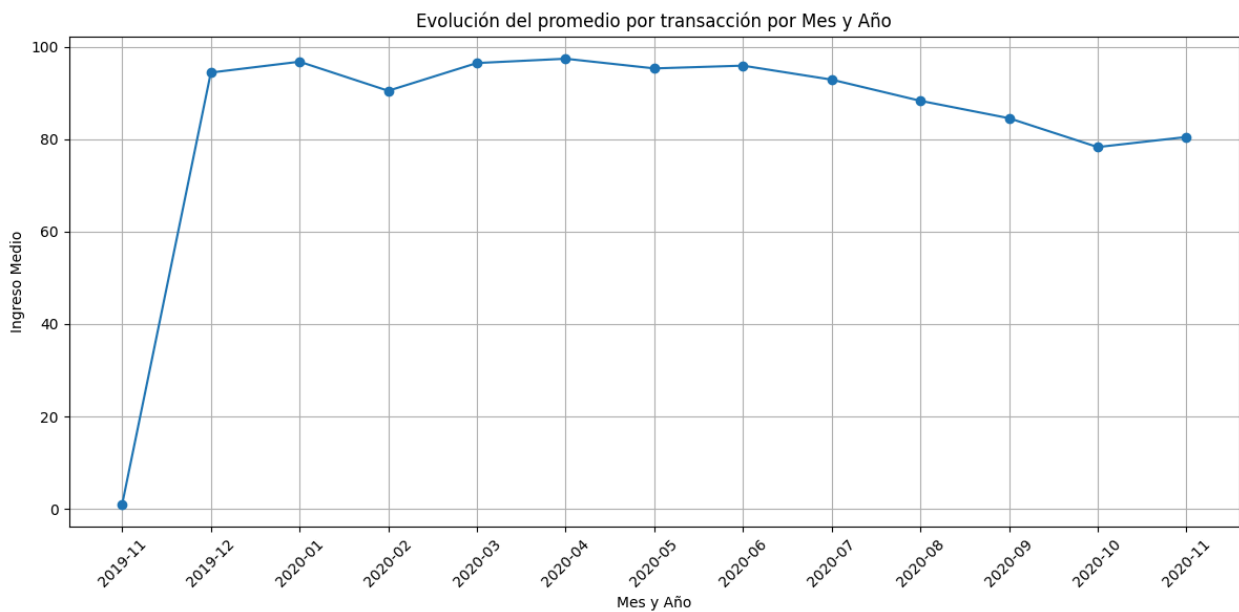


- **Ingresos generados:** se observa que los ingresos generados están totalmente relacionados con la frecuencia de uso del servicio. Por tanto, en futuros estudios podría estudiarse la relación entre ellos y generar un modelo de aprendizaje automático que nos ayudara a predecir estas variables.



- **Promedio por transacción:** esta métrica proporciona una visión del comportamiento de gasto promedio de los usuarios, lo cual es valioso para la planificación de precios y ofertas.

El gasto promedio se ha quedado constante a lo largo de los meses, lo que nos indica que el aumento total de los ingresos generados no está debido a un aumento del gasto promedio por usuario, sino a un incremento de la cartera de clientes (aumento de frecuencia de uso del servicio).



- **Tiempo medio de cobro:** Observamos un periodo inicial hasta mayo de 2020 donde no se producía un cobro de comisión por solicitar un cash request. Esto podría ser debido a un periodo promocional para atraer clientes. Posteriormente, ambas métricas muestran variabilidad significativa a lo largo del tiempo, lo cual sugiere que factores externos e internos pueden haber influido en la capacidad de cobro tanto de las solicitudes de efectivo como de las comisiones.

