

Exploración de los Datos

Nuria Arroyo Bustamante y Heriberto Espino Montelongo

Universidad de las Américas Puebla

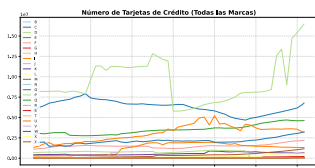
P25 LDS1021 1: Minería de Datos

Dr. Hector Saib Maravillo Gomez

3 de marzo de 2025

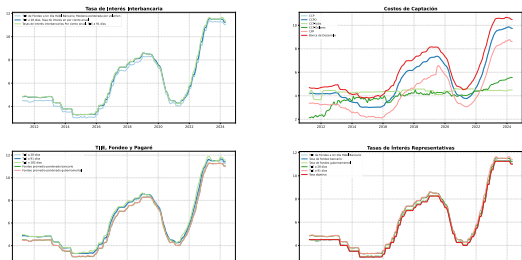
Exploración de los Datos

En esta etapa se lleva a cabo un análisis de sensibilidad del número de tarjetas de crédito, desglosadas por tipo, en relación con diversas variables macroeconómicas. El objetivo es identificar cuáles de estas variables ejercen un mayor impacto en cada segmento de tarjetas de crédito.



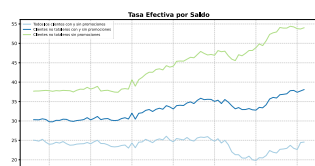
Se inició evaluando el volumen de tarjetas disponibles en el mercado, observándose que cada tipo sigue distribuciones particulares. Por ejemplo, las tarjetas tipo D experimentaron una notable caída a finales de 2017, mientras que las de tipo F mostraron una tendencia ascendente durante el mismo periodo.

Identificación de Variables Redundantes



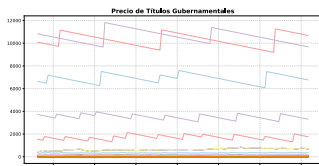
Se detectó que varias variables presentaban patrones similares, lo que se confirma en el correlograma al observar correlaciones iguales a 1. Dado que la inclusión de variables redundantes no aporta información adicional, se optó por conservar únicamente la TIIE a 28 días.

Tasa Efectiva por Saldo



Aún no se ha determinado si la tasa efectiva por saldo causa el volumen de tarjetas de crédito activas o si se produce al revés. Este aspecto será objeto de análisis, por lo que se mantendrá en el estudio.

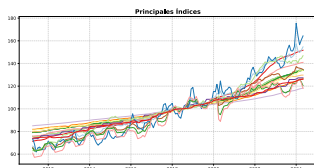
Precios de Títulos Gubernamentales



Estos precios presentan un comportamiento cíclico, pero se mantienen estables en términos de estacionariedad y estacionalidad. En el correlograma final se muestra que se formaron dos grupos, uno formado por las tarjetas G, U, M, E, ..., D, tienen una correlación negativa contra los CETES, que es opuesto a las tarjetas V, W, P, H, C, que cuentan con una correlación positiva, mientras que con las otras variables las correlaciones son al revés, es primer grupo cuenta con la mayoría de las correlaciones positivas mientras el segundo cuenta con correlaciones negativas.

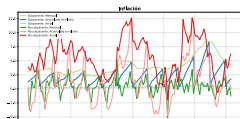
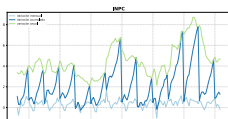
D, tienen una correlación negativa contra los CETES, que es opuesto a las tarjetas V, W, P, H, C, que cuentan con una correlación positiva, mientras que con las otras variables las correlaciones son al revés, es primer grupo cuenta con la mayoría de las correlaciones positivas mientras el segundo cuenta con correlaciones negativas.

Índices Económicos



Los principales índices han mostrado una tendencia al alza, acompañada de fluctuaciones. Se contempla la posibilidad de aplicar diferencias a una de estas series, para hacerlas estacionarias y realizar una regresión entre series de tiempo, con el fin de identificar patrones o relaciones causales con el volumen de tarjetas de crédito.

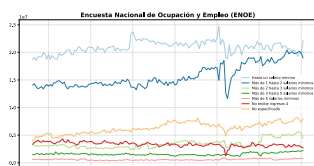
INPC e Inflación



Tanto el INPC como la inflación exhiben una alta varianza, caracterizada por marcadas subidas, bajadas y numerosos picos, que se mantiene

cercana al 2.5, lo que puede indicar estacionariedad.

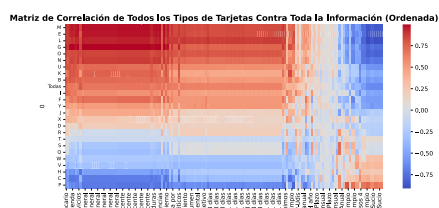
Aspectos Relacionados con el Mercado Laboral



El número de empleados formales ha registrado un crecimiento moderado durante el periodo, aunque presenta pequeñas fluctuaciones que podrían ser significativas. Es importante destacar que el rango de crédito asignado a una persona está

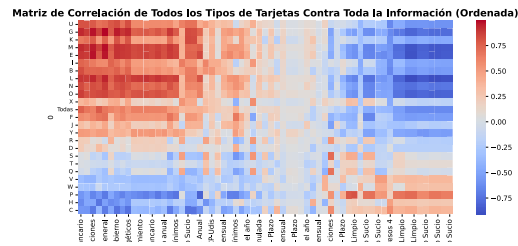
directamente ligado a sus ingresos, y la ENOE proporciona datos desglosados en términos de salarios mínimos percibidos.

Selección Final de Variables



Se hizo una matriz de correlación lineal del volumen de tarjetas de crédito vs las otras variables y se identificaron grupos de variables correlación superior a 0.95.

Por lo que se compararon las variables que venían de la misma fuente de información y se optó por conservar aquella variable que ofrecía una interpretación más sencilla. Por ejemplo, entre la “TIIE a 28 días” y la “Tasa de fondeo ponderada por BPAS y BREMS”, se eligió mantener la TIIE a 28 días, otro ejemplo es la matriz de correlación de los principales índices, donde los valores con menor correlación tienen 0.95. De todas formas se optó por mantener INPC Por Objeto al gasto (primera fila), SSC e INPC, no subyacente, energéticos y tarifas autorizadas por el gobierno, energético y no energético (penúltima y antepenúltima fila).



Tras eliminar las variables redundantes, se obtuvo el correlograma final con las variables que se utilizarán en el análisis.