

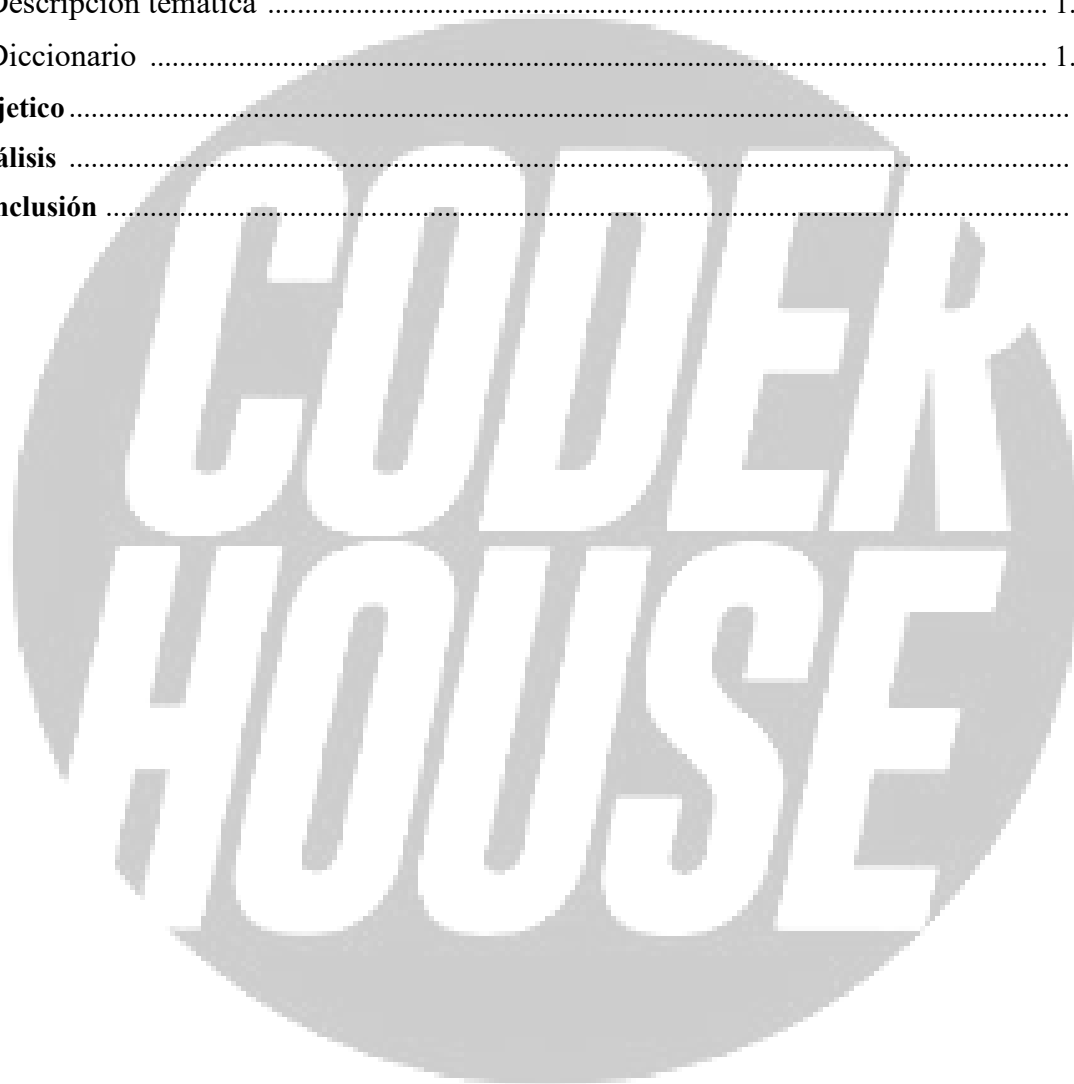
Data Science I

Comisión 67465

Blanco Sabrina

TABLA DE CONTENIDOS

Introducción.....	1
Descripción temática	1.1
Diccionario	1.2
Objetivo	2
Análisis	3
Conclusión	4



1. INTRODUCCIÓN

1.1 Descripción temática

El presente análisis se basa en un conjunto de datos que representa las ventas del Centro de Producción (CDP) del grupo gastronómico Urban Group, empresa en la que trabajo. Urban Group gestiona diversas marcas de restaurantes, tanto propios como franquiciados. El CDP centraliza la compra de insumos y la producción de artículos que luego se distribuyen a los distintos locales. Estos artículos incluyen desde platos elaborados y envasados (listos para regenerar), descartables y porcionado de carnes.

Cada fila del dataset representa una transacción de venta realizada por el CDP e incluye la siguiente información:

- Fecha, codigo, nombre, apellido, razón social del cliente
- Comprobante, prefijo y numero
- Tipo de local: Propio, franquicia o venta al personal interno
- Tipo de factura, numero de mesa, mozo
- Dia de entrega
- Precio, cantidad, codigo, nombre, rubro del articulo vendido
- Si fue facturado y el importe total de dicha factura

1.2 Diccionario

Columna | Tipo de dato | Descripción |

- Fecha | Fecha (datetime) | Fecha de la venta realizada por el CDP
- Cod_cliente | Int | Código interno del cliente
- Nombre | Texto | Nombre del cliente
- Apellido | Texto | Apellido del cliente
- Razón social | Texto | Razon social, si es personal nombre y apellido
- Comprobante | Texto/Numero | Numero de factura
- Prefijo | Numero | Identificadores únicos del comprobante
- Número | Numero | Identificadores únicos del comprobante
- Tipo de local | Texto | Puede ser "Propio", "Franquicia" o "Personal"
- Cod_cpb | Texto | Tipo de la factura emitida
- Concepto | Texto | Por donde ingreso la venta

- Mesa | Numérico (int) | Identificador del control de mesa (esto porque el * sistema está adaptado de un restaurante)
- Cod_mozo | Numero | Nombre del mozo que cargó el pedido (esto porque el sistema está adaptado de un restaurante)
- Envio | Texto | Dia de entrega
- Precio | Numérico (float) | Precio por unidad del artículo
- Cantidad | Numérico (float/int) | Unidades vendidas
- Nombre | Texto | Nombre del artículo
- Código artículo | Texto/Numérico | Código interno del artículo vendido
- Rubro | Texto | Rubro al que pertenece el articulo
- Facturado | Texto | Indica si fue o no facturado
- Importe | Numero (float) | Total de la factura que contiene el articulo

2. Objetivo final

El objetivo es poder predecir si un cliente va a pedir de un rubro específico.

Utilizare método de clasificación usando RandomForestClassifier

Para cada combinación de cliente + rubro + fecha: si el cliente hizo un pedido ($CANTIDAD > 0$) = 1 si no pidió nada ese día = 0

Features:

- Fecha
- Cliente
- Rubro

3. Análisis

En la primera ejecución el modelo predice 100% que se pide, esto es por los datos que se tienen ya que todos los locales suelen pedir todos los productos.

Agrego casos en los que no se pida para ver como funciona y se obtiene el caso contrario al inicial.

Pruebo que pasa con una muestra mas chica, ahora se observa de ambos

4. Conclusión

Podría servir para predecir si el cliente pedirá cierto rubro en una próxima fecha pero solo para hacer recomendaciones de pedidos, no para tener en cuenta para el stock de cada local (debido a los casos que se agregaron de no pedido)

Estoy haciendo toda la carrera de data science para poder aplicarlo a mi trabajo, en cada curso veo modificaciones desde la base de información que tenemos y son los próximos cambios para hacer, quise ver y usar el dataset que tenía porque es la información que veo día a día.

Luego de haber realizado esta entrega se me ocurre tomar las ventas de los locales para predecir a partir de las ventas que rubros elaborar en el centro de producción, en lugar de hacerlo al revés ya que contarían con stock de mercadería en cada local, sacando la información de las ventas se puede reforzar en cdp el personal en ese rubro.

