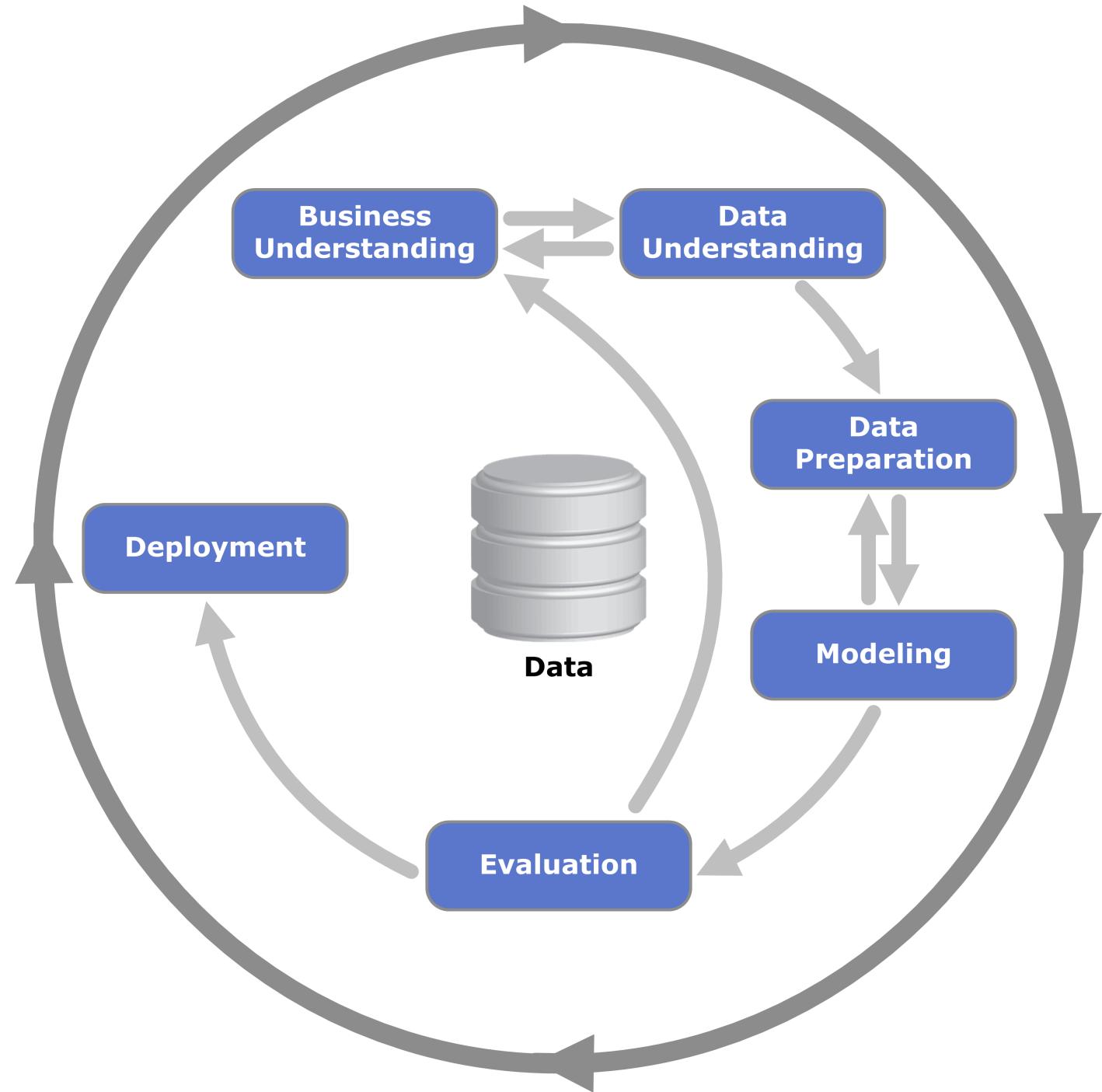


Clasificación de tipos de viajes Walmart



Metodología CRISP-DM

1. Comprensión del Negocio
 2. Comprensión de los Datos
 3. Preparación de los datos
 4. Modelado
 5. Evaluación
 6. Implantación



Metodología CRISP-DM



1 Comprendión del Negocio	2 Comprendión de los datos	3 Preparación de los Datos	4 Modelado	5 Evaluación
Objetivo	Calidad de los datos	Limpieza de los Datos	Modelos	Función de Perdida
El objeto del presente trabajo es recrear la clasificación realizada por Walmart, mediante el empleo de técnicas de ciencia de datos, con el fin de ayudar a mejorar la toma de decisiones en cuanto al plan de mercadotecnia de la empresa y mejorar la experiencia de compra del cliente.	Los datos aportados por Kaggle presentaron algunas inconsistencias como registros mal escritos, datos faltantes y no tenían formato Tidy. EDA Mediante el Análisis Exploratorio de Datos se pudo identificar a algunas de las variables relevantes que explican la clasificación de los viajes de los clientes en función del consumo.	Mediante herramientas estadísticas se procedió a limpiar los datos capturados con el fin de prepararlos para la etapa de transformación. Transformación de los Datos Con el fin de obtener un mayor poder de predicción se le realizaron diversas transformaciones a los datos .	Con base en el desarrollo de distintos modelos estadísticos y de aprendizaje automático se buscó clasificar el tipo de viaje con base en las probabilidades obtenidas aplicando los modelos.	Evaluamos nuestros modelos con la función de perdida logarítmica multiclas.
				Comparamos los resultados de los modelos propuestos y elegimos el que tenía mejor puntuación. $-\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} \log(p_{ij}),$

6. Implementación

Intentamos correr Flask pero por limitaciones de tiempo no se logró ese parte del proyecto.

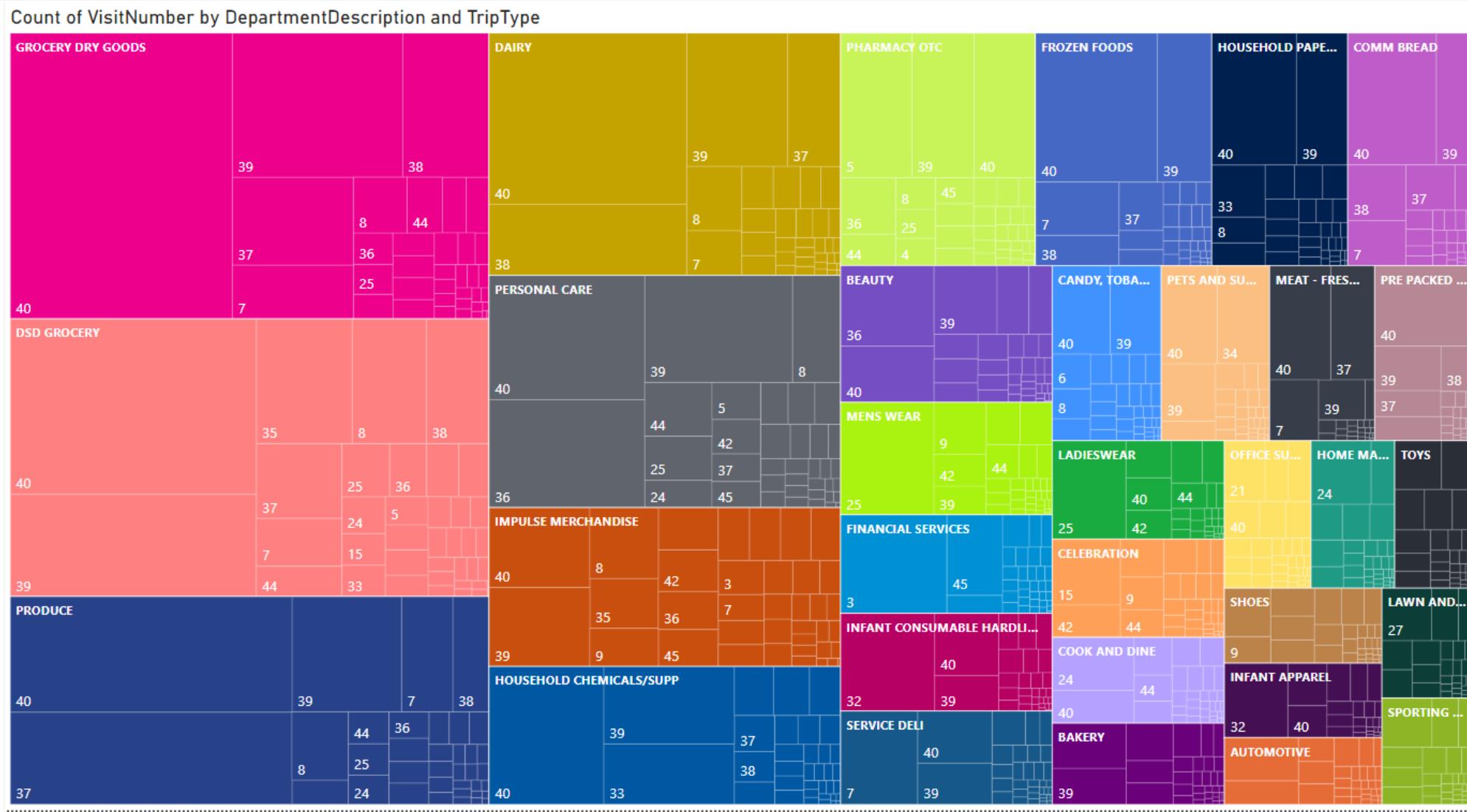
1.- COMPRENSIÓN DEL NEGOCIO

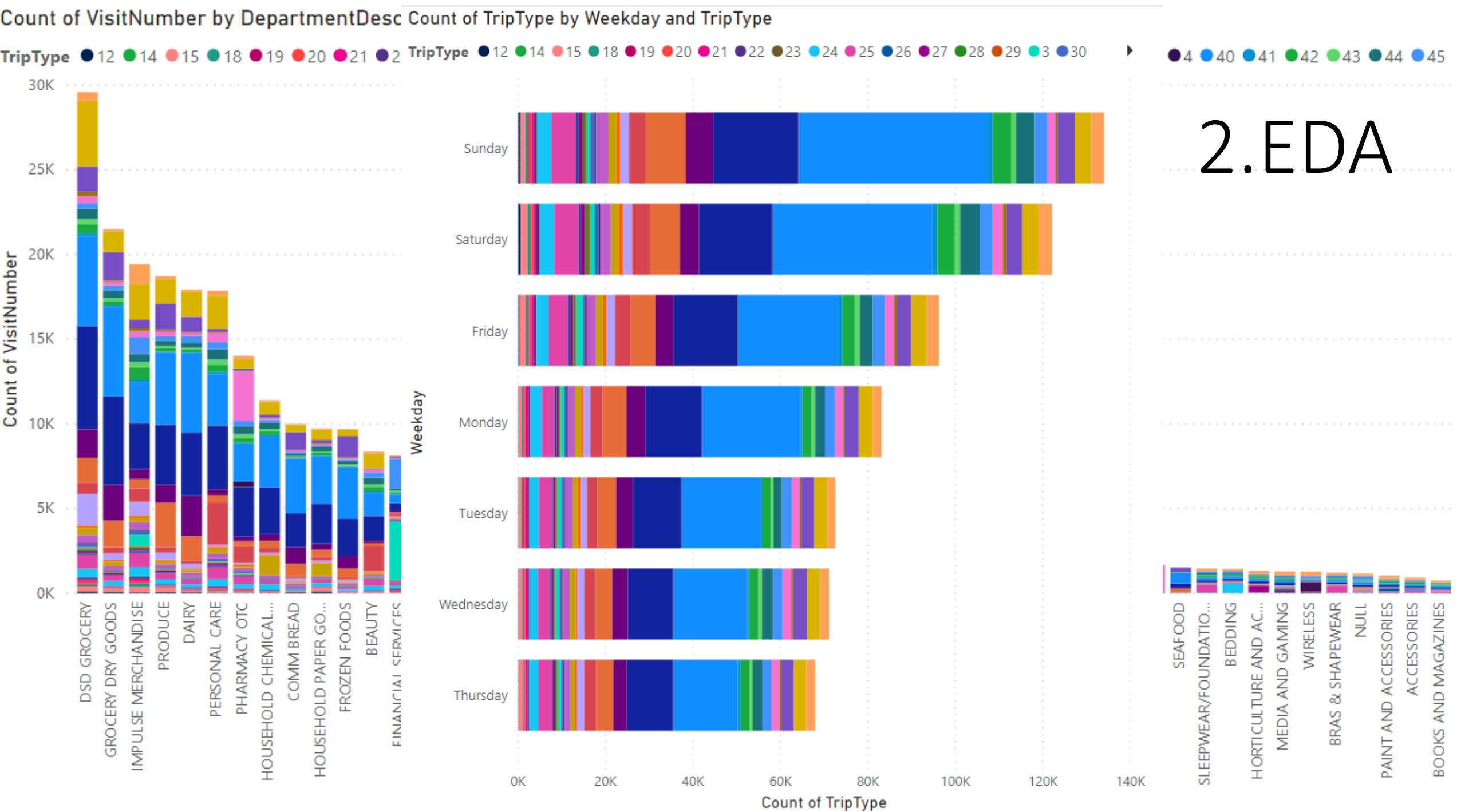
Objetivo

El objeto del presente trabajo es recrear la clasificación realizada por Walmart, mediante el empleo de técnicas de machine learning, con el fin de ayudar a mejorar la toma de decisiones en cuanto al plan de mercadotecnia de la empresa y mejorar la experiencia de compra del cliente.

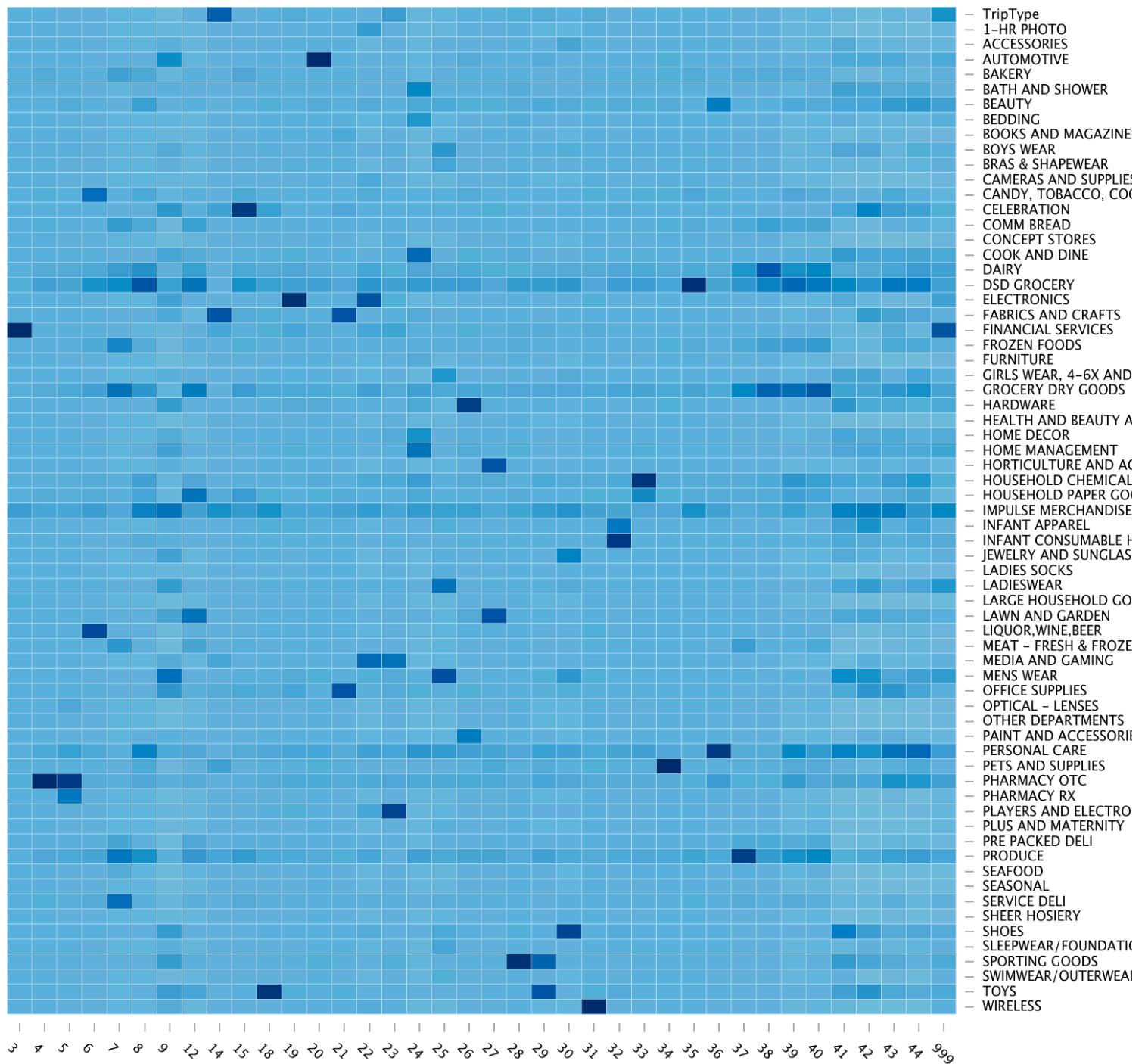
Mediante estrategias diversas como el reacomodo de productos, promociones por sector, ventas atadas, entre muchas otras, se busca mejorar las ventas a partir de un conocimiento de los hábitos de consumo de los clientes.

2 COMPRENSIÓN DE LOS DATOS





3. PREPARACIÓN DE LOS DATOS



4. MODELADO

Microsoft Azure Machine Learning Studio (classic)

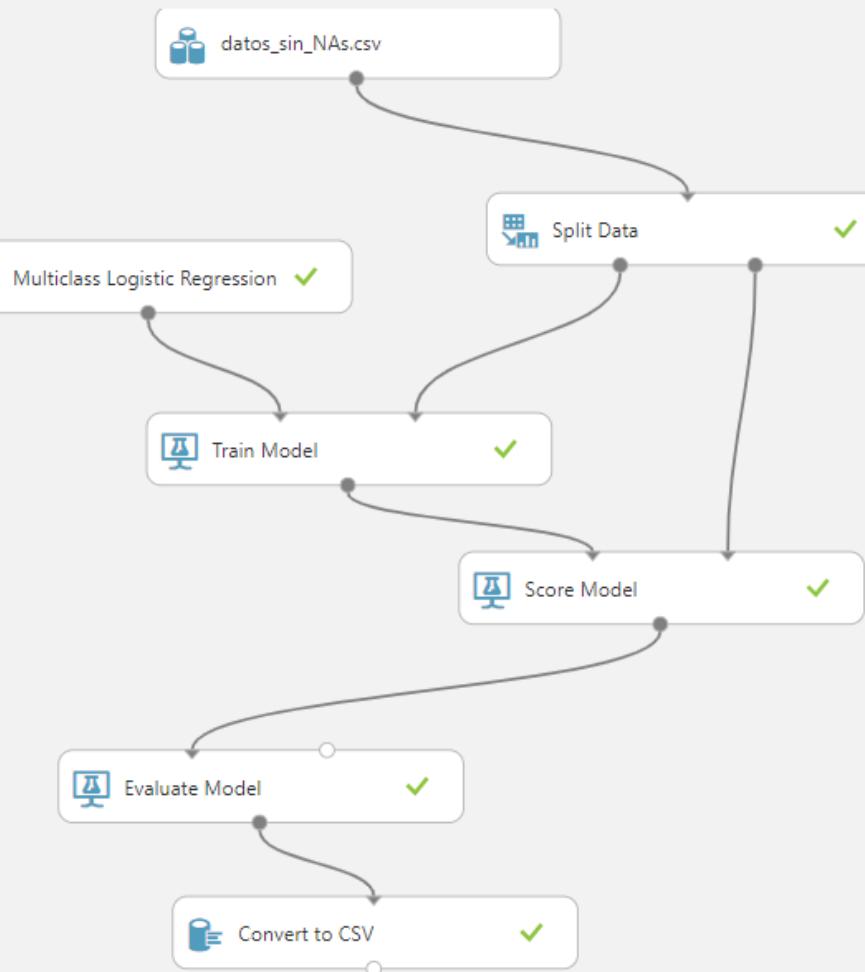
Javier Vale

- Search experiment items
- Saved Datasets
- Trained Models
- Data Format Conversions
- Data Input and Output
- Data Transformation
- Feature Selection
- Machine Learning
- OpenCV Library Modules
- Python Language Modules
- R Language Modules
- Statistical Functions
- Text Analytics
- Time Series
- Web Service

Training experiment

Predictive experiment

Experiment created on 12/18/2019



5. EVALUACIÓN Y RESULTADOS

9 submissions for Daniela Pinto Veizaga		Sort by	Private Score	▼	
All	Successful	Selected			
Submission and Description		Private Score	Public Score	Use for Final Score	
pre_GB_std.csv	6 hours ago by Daniela Pinto Veizaga	1.32845	1.33002	<input type="checkbox"/>	
add submission details					
pre_GB.csv	6 hours ago by Daniela Pinto Veizaga	1.40185	1.40690	<input type="checkbox"/>	
XGboost					
pre_logit.csv	8 hours ago by Daniela Pinto Veizaga	1.42031	1.42349	<input type="checkbox"/>	
third try					

6.
▶ CONCLUSIONES