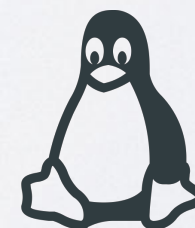
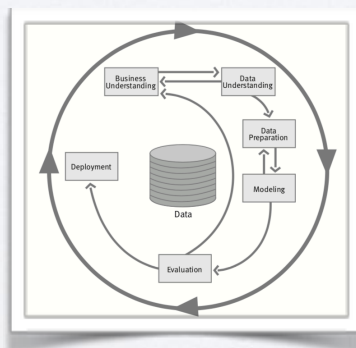


WALMART RECRUITING:TRIP TYPE CLASSIFICATION

Itzel Muñoz | 22803
Héctor Corro | 14350
Uriel Miranda | 77508



BUSINESS UNDERSTANDING



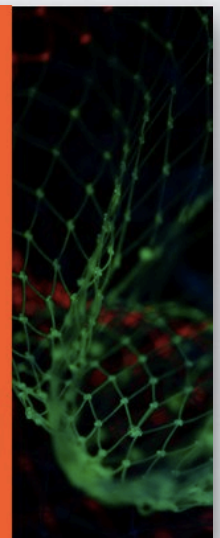
- 20,000 Tiendas
- 28 Países
- THE DATA CAFE
- 2.5 PB / Hr

“

A **Petabyte** is the equivalent of **4,000** digital photos per day, over your entire life.

Lifewire.com

”



DATA UNDERSTANDING

```

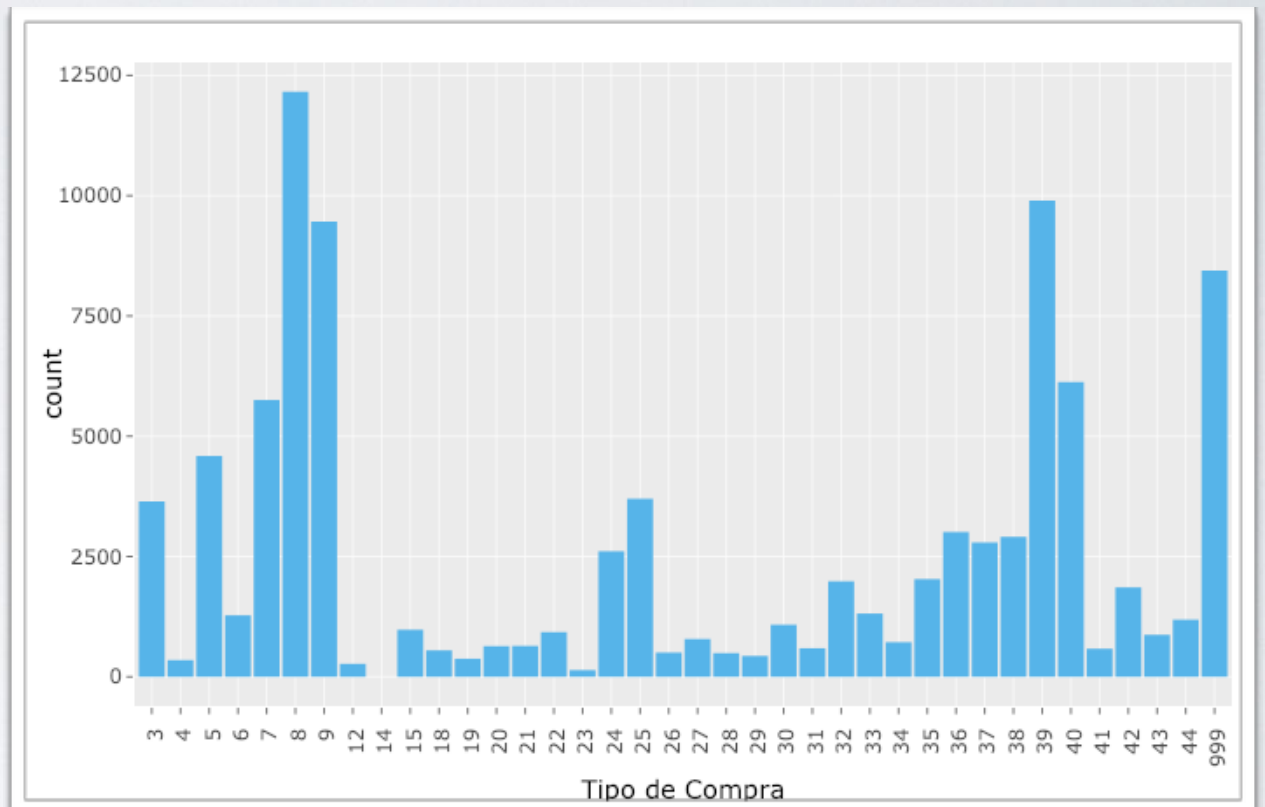
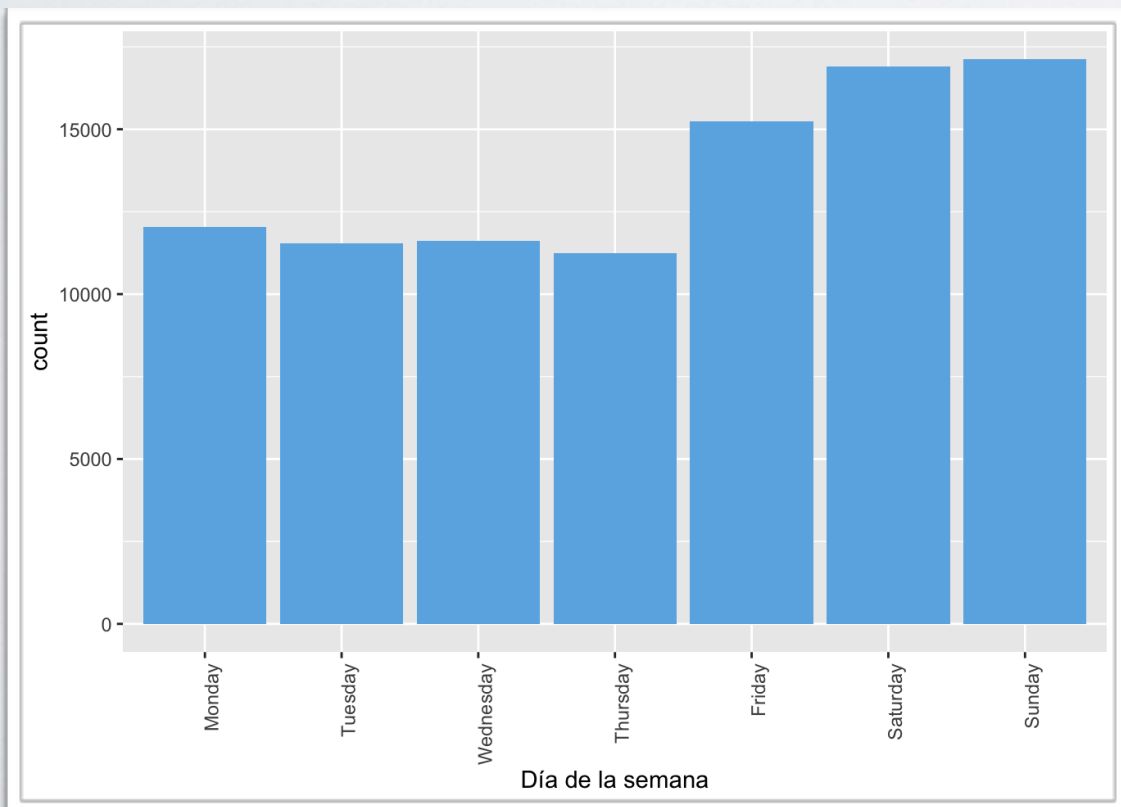
graph TD
    TripType --> VisitNumber
    TripType --> Weekday
    VisitNumber --> Upc
    Weekday --> Upc
  
```

Weekday

Upc

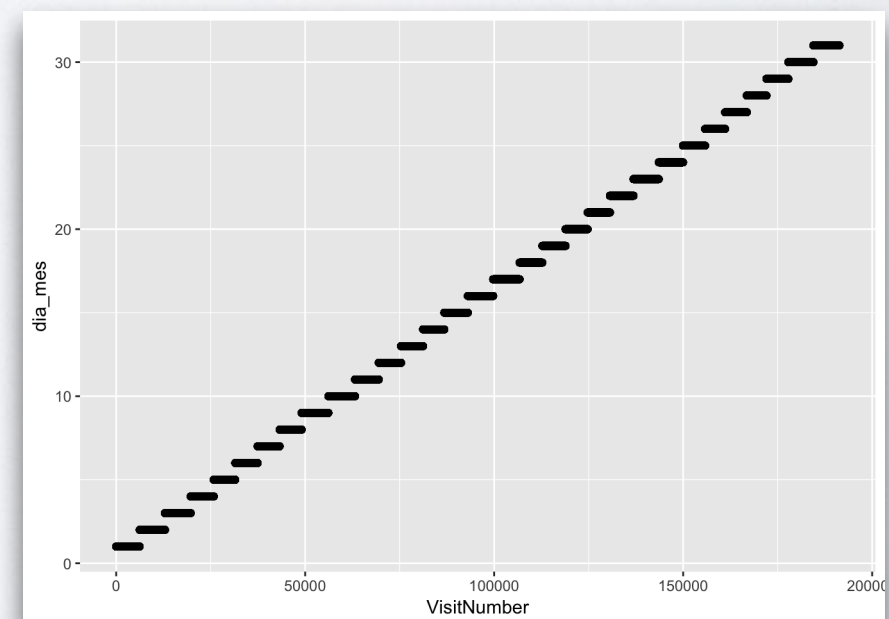
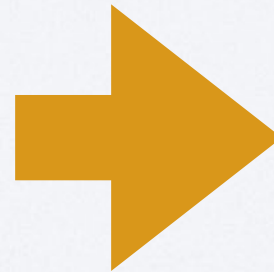
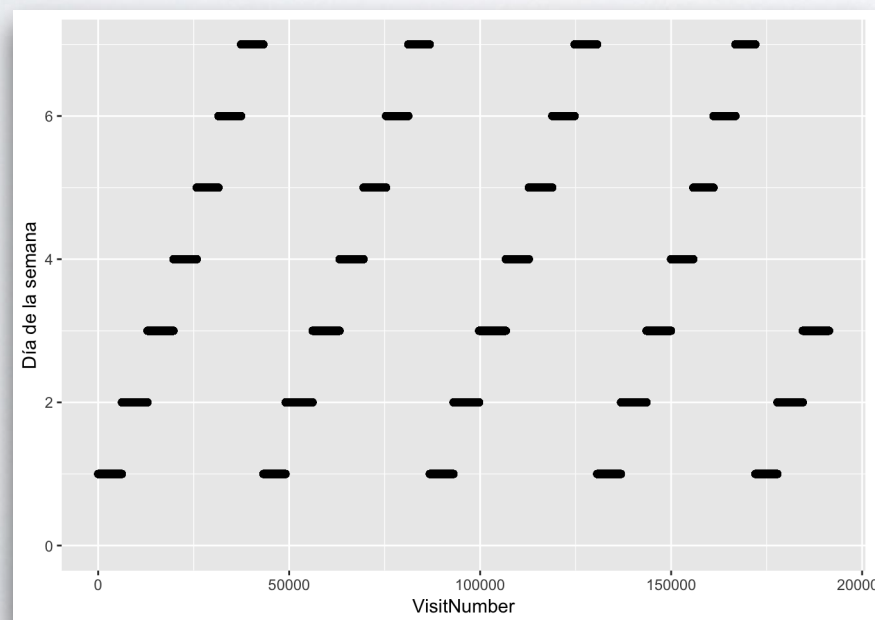
ScanCount	DepartmentDescription	FinelineNumber
-----------	-----------------------	----------------

FinelineNumber

[illegible]

DATA PREPARATION

- Total de productos comprados en la visita.
- Indicadora de devolución de productos.
- Total de productos devueltos en la visita.
- Total de departamentos visitados, si se compró al menos un producto en el departamento éste cuenta cómo 1a en la suma de departamentos visitados.
- Productos diferentes comprados respecto a UPC, es decir cuántos productos con diferente código UPC se compraron en la visita
- Día del mes en que se realizó la visita, se explica más adelante esta variable



MODELS

■ ■ ■



Modelo	Parámetros	Tiempo Ajuste	Public Score
Regresión Multinomial	penalización L1, C = 1	3 hrs	1.3185
Bosque Aleatorio	depth = 100, feat = sqrt, split = 10, n_est = 1500	4 hrs	2.1320
Máquina de Soporte Vectorial	kernel lineal, C = 10	5 hrs	1.0820
Red Neuronal	alpha 100, 30 capas, 1000 iter, semilla 0	4 hrs	1.0370

EVALUATION