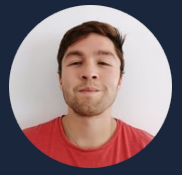
Walmart Chile

Merchandising Analytics

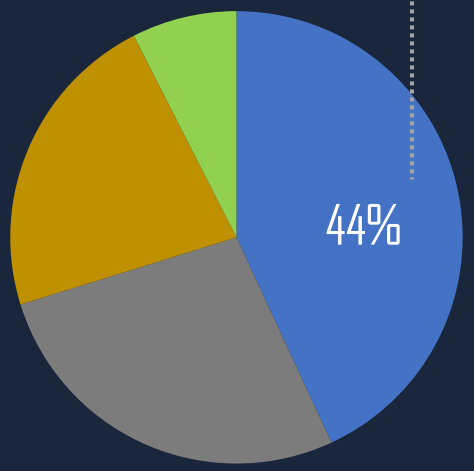
Sebastián Torres
Subgerente de Analítica Comercial



Francisco Alfaro
Senior Data Scientist

Nicolás Fredes
Senior Data Scientist

Walmart Chile



lider

express de LIDER

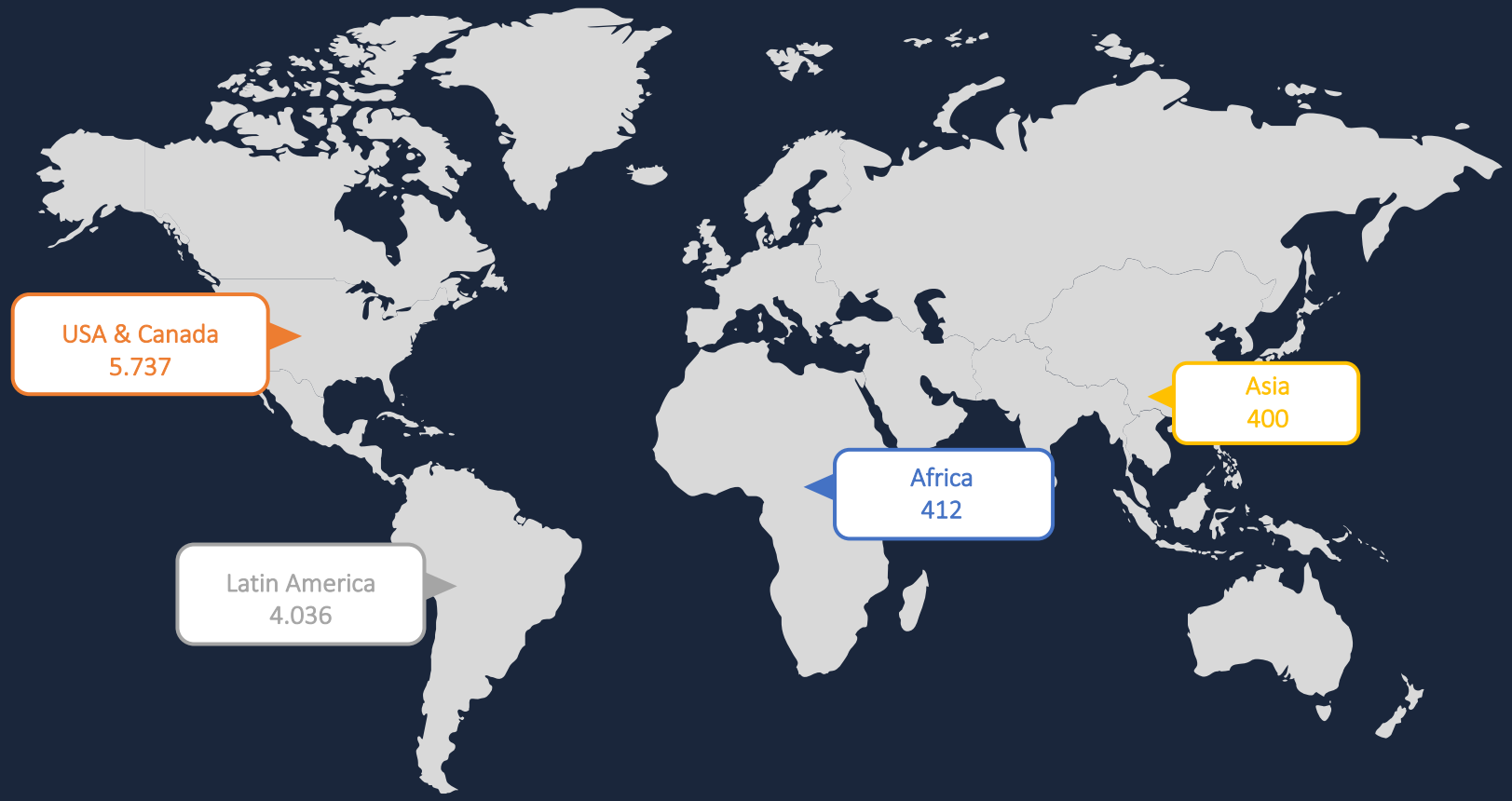
SuperBodega aCuenta

central mayorista

lider.cl

Walmart

Save money. Live better.



10.585 stores
2.3 million associates
Revenue of US$572 billions

¿Qué es el Retail?

1 BUY

- Sourcing & Negotiation
- Planning & Revenue
- Space & Assortment
- Pricing & Promotions
- Item & Catalog

2 MOVE

- Imports
- Warehouses
- Logistics & Routing
- Supply chain
- Last mile delivery

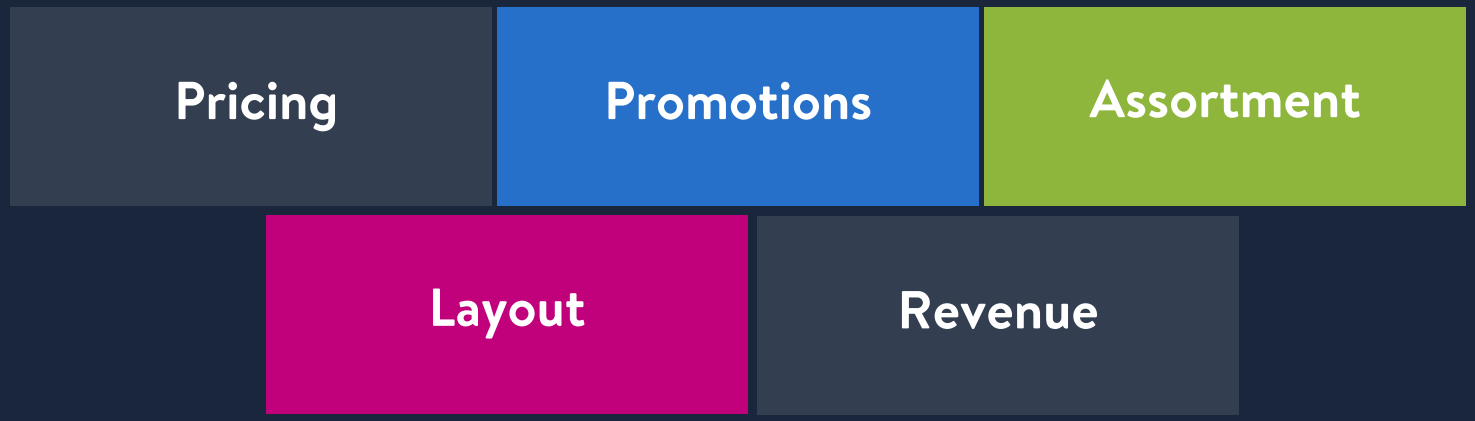
3 SELL

- Store location
- Workforce scheduling
- Digital experience
- Customer Relationship Management (CRM)

385 stores, 60000 products, 50000 associates
4.5 millions customers
8000 millions scans

Merchandising Analytics

To provide digital solutions for Walmart Chile business issues,
by using **new technologies**, **mathematical modeling**,
and the ability to process **large amounts of data**.



Association rules

Association of products or sets of products

7804642160138	PAN ARABE..	$ 990
7809611707444	PECHUGA DESH	$ 3.690
7802575201218	NIDO FETTUC	$ 1.030
7808743602146	AVENA INSTAN	$ 770
7805045000144	JABON LAVAR	$ 3.490
7803468003025	MOLDE LIGHT	$ 1.390
7802920000916	YOGHURT LIGH	$ 220
7802920000916	YOGHURT LIGH	$ 220
7802920000916	YOGHURT LIGH	$ 220
7802920000916	YOGHURT LIGH	$ 220
7613036756440	NUBE PINA	$ 380
7613036756440	NUBE PINA	$ 380
7613036756440	NUBE PINA	$ 380
7613036756440	NUBE PINA	$ 380
7802200400047	MERM DURAZNO	$ 970
7613036443166	NUBE COCO	$ 380
7613036443166	NUBE COCO	$ 380
7613036443166	NUBE COCO	$ 380
7613036443166	NUBE COCO	$ 380
7802710350504	CASSATTA BRI	$ 2.190
7802200400061	MERM DAMASCO	$ 970
7613035180703	PACK CONDENS	$ 1.490
7802615006551	ARROZ GRADO	$ 1.150
7802926000682	HELADO AGUA	$ 2.690

$$supp(X) = \frac{|\{t \in D; X \subseteq t\}|}{|D|} = \frac{C_x}{|D|} = P(X)$$

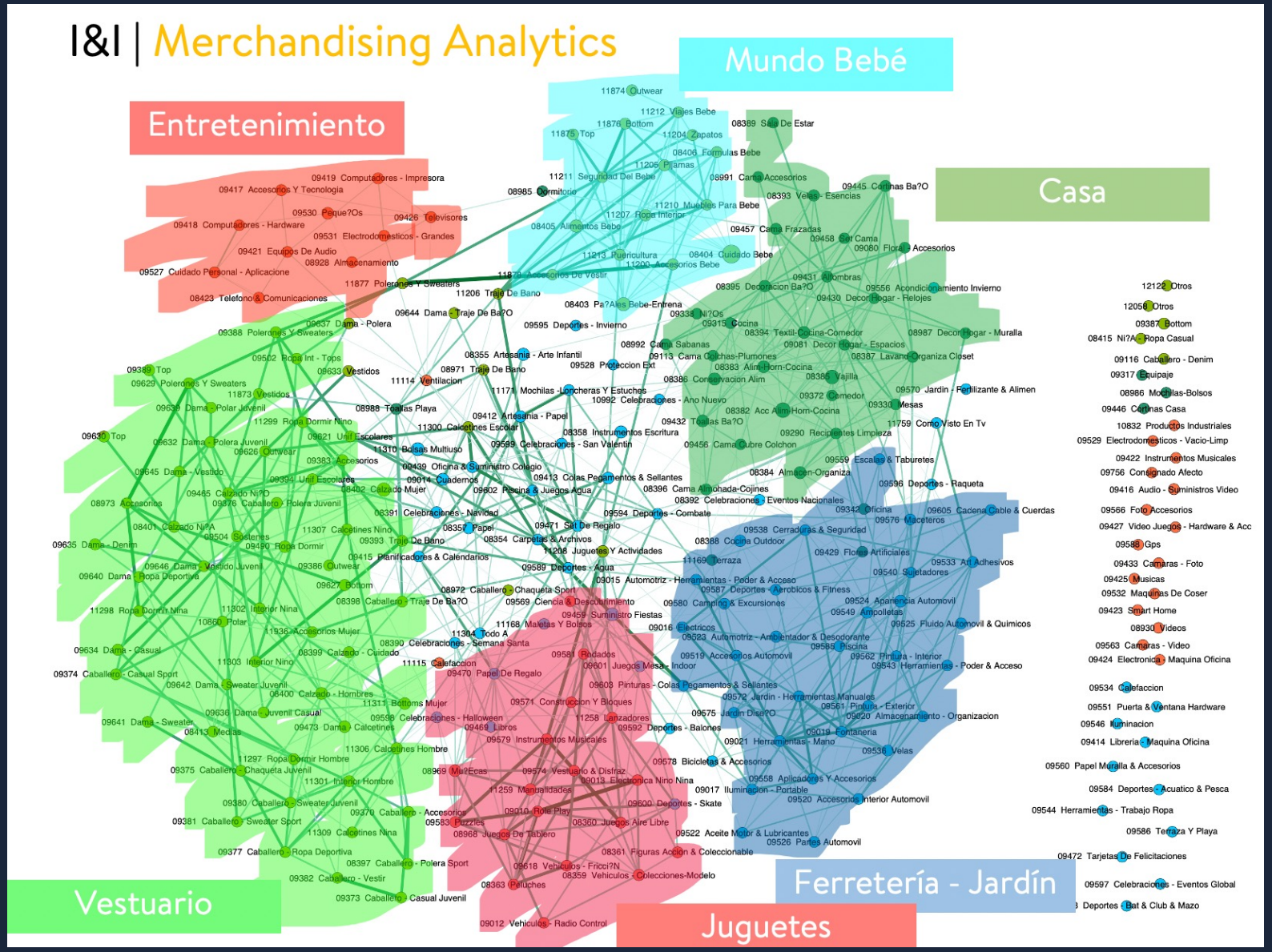
$$conf(X \Rightarrow Y) = \frac{supp(X \cup Y)}{supp(X)} = \frac{P(X \cap Y)}{P(X)} = \frac{C_{xy}}{C_x} = P(Y|X)$$

$$lift(X \Rightarrow Y) = \frac{conf(X \Rightarrow Y)}{supp(X)} = \frac{P(X \cap Y)}{P(X)P(Y)} = \frac{C_{xy}}{C_x C_y}|D|$$

Association rules

STORE_NBR	DPFL_X	Fineline_Desc_X	DPFL_Y	Fineline_Desc_Y	C_xy	C_x	C_y	Lift
3	21100	ACOND FAMILIAR	23600	SHAMPOO FAMILIAR	14991	30140	43910	52.3
3	21075	ACOND TOD/TIPO CAB	23525	SHAMPOO CABELLO SECO	10035	62870	24799	29.7
3	924500	CREMAS	924525	SOPAS	24650	58176	78410	25
3	21075	ACOND TOD/TIPO CAB	23475	SHAMPOO CABELLO NORM	23979	62870	72489	24.3
8	922750	TODA OCASION	928350	TENTACION	17077	38506	54085	23.6
616	924500	CREMAS	924525	SOPAS	21864	42975	58263	23.5
616	21100	ACOND FAMILIAR	23600	SHAMPOO FAMILIAR	29500	52975	64427	23.3
3	921875	DOYPACK KETCHUP	921950	MOSTAZA REGULAR	13060	73408	39251	20.9
616	21075	ACOND TOD/TIPO CAB	23525	SHAMPOO CABELLO SECO	10182	57335	23325	20.5

Association rules



Pricing optimization

Optimal price
for each product
at each store.

$$\max \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} (p_{i,t} - c_{i,t}) q_{i,t}$$

s.t.

$$(1) \quad \frac{\sum_{i \in \mathcal{I}_k} \frac{p_{i,t}}{cp_{i,t}} v_{i,t}}{\sum_{i \in \mathcal{I}_k} v_{i,t}} \leq pok_k \qquad \forall k \in \mathcal{K}, t \in \mathcal{T},$$

$$(2) \quad \frac{p_{i,t}}{cp_{i,t}} \leq poi_{i,t} \qquad \forall i \in \mathcal{I}, t \in \mathcal{T},$$

$$(3) \quad p_{i,t} = p_{j,t} + \delta^1_{i,j} \qquad \forall (i,j) \in \mathcal{L}_p, t \in \mathcal{T},$$

$$(4) \quad p_{i,t} = p_{j,t} \delta^2_{i,j} \qquad \forall (i,j) \in \mathcal{L}_f, t \in \mathcal{T},$$

$$(5) \quad p_{i,t} = p_{j,t} \delta^3_{i,j} \qquad \forall (i,j) \in \mathcal{L}_{mp}, t \in \mathcal{T},$$

$$(6) \quad p_{i,t} = fp_{i,t} \qquad \forall i \in \mathcal{F}, t \in \mathcal{T},$$

$$(7) \quad p_{i,t} \geq 1.19 c_{i,t} \qquad \forall i \in \mathcal{I}, t \in \mathcal{T},$$

$$(8) \quad \frac{op_{i,t} - p_{i,t}}{pp_{i,t}} \geq 0 \qquad \forall i \in \mathcal{I}, t \in \mathcal{T},$$

$$(9) \quad p_{i,t} \leq \max \left\{ \frac{1.19 c_{i,t}}{1 - (om_{i,t} + mg)}, pp_{i,t}(1 - mv^{\flat}_k) \right\} \qquad \forall (i,t) \in \mathcal{NP}, k \in \mathcal{K}, v_{i,t} > 0,$$

$$(9^*) \quad p_{i,t} \geq \min \left\{ \frac{1.19 c_{i,t}}{1 - (om_{i,t} + mg)}, pp_{i,t}(1 + mv^{\sharp}_k) \right\} \qquad \forall (i,t) \in \mathcal{NP}, k \in \mathcal{K}, v_{i,t} = 0,$$

Forecast modeling

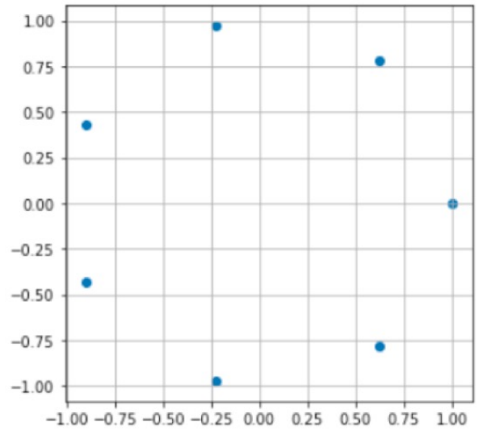
1. ⛏️ Data pre-processing
2. 👷 Feature Engineering (*)
3. 📈 Trend Estimation (*)
4. 💥 Model Training (*)
5. 🔧 Post-Processing

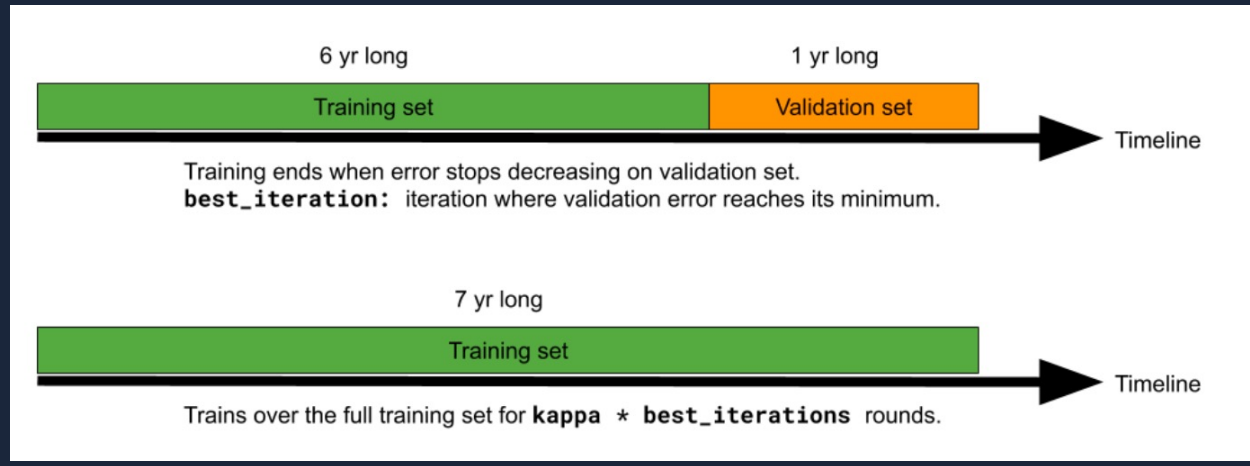
$$y(t) = \underbrace{T(t)}_{\text{Trend}} + \underbrace{S(t)}_{\text{Seasonality}} + \underbrace{H(t)}_{\text{Holidays}} + \underbrace{E(t)}_{\text{Exogenous*}} + \epsilon_t$$

Cyclical representation for day of week

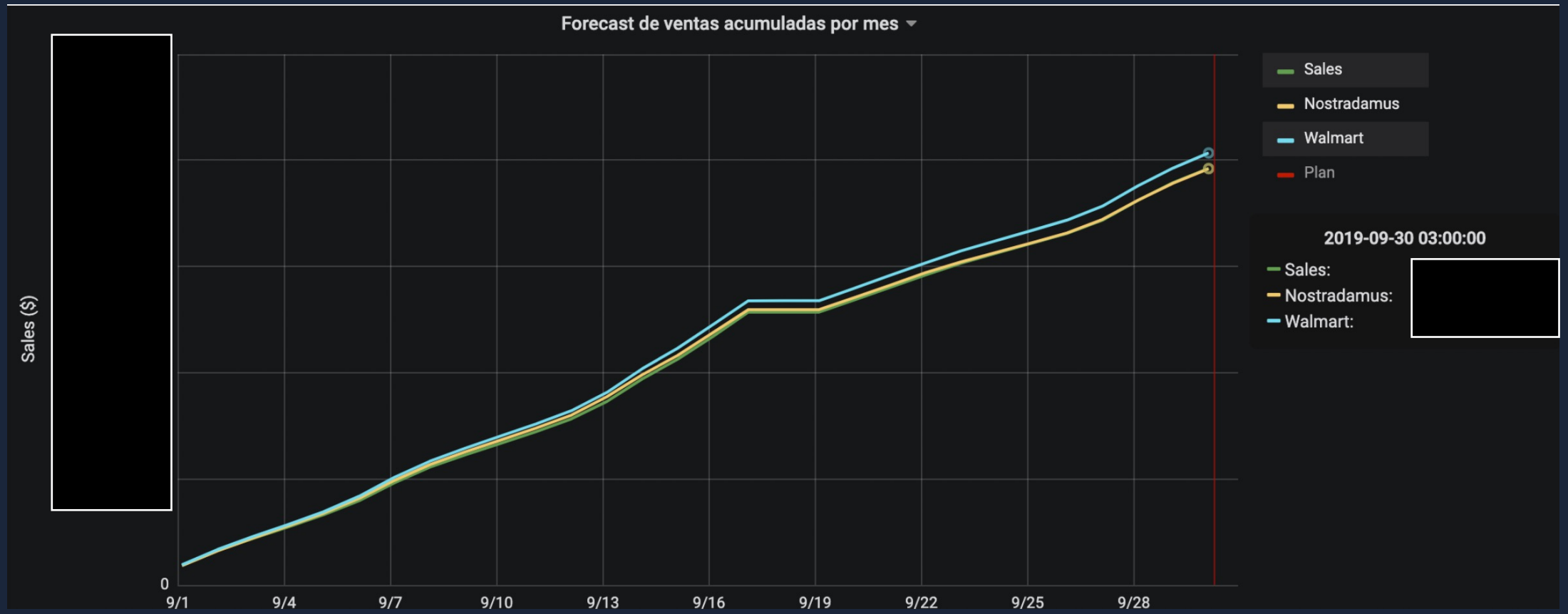
```
day_of_week = np.array([0,1,2,3,4,5,6])
dw_sin = np.sin(day_of_week*(2.*np.pi/7))
dw_cos = np.cos(day_of_week*(2.*np.pi/7))

plt.figure(figsize=(5,5))
plt.grid()
plt.scatter(dw_cos, dw_sin)
plt.show()
```





Forecast modeling



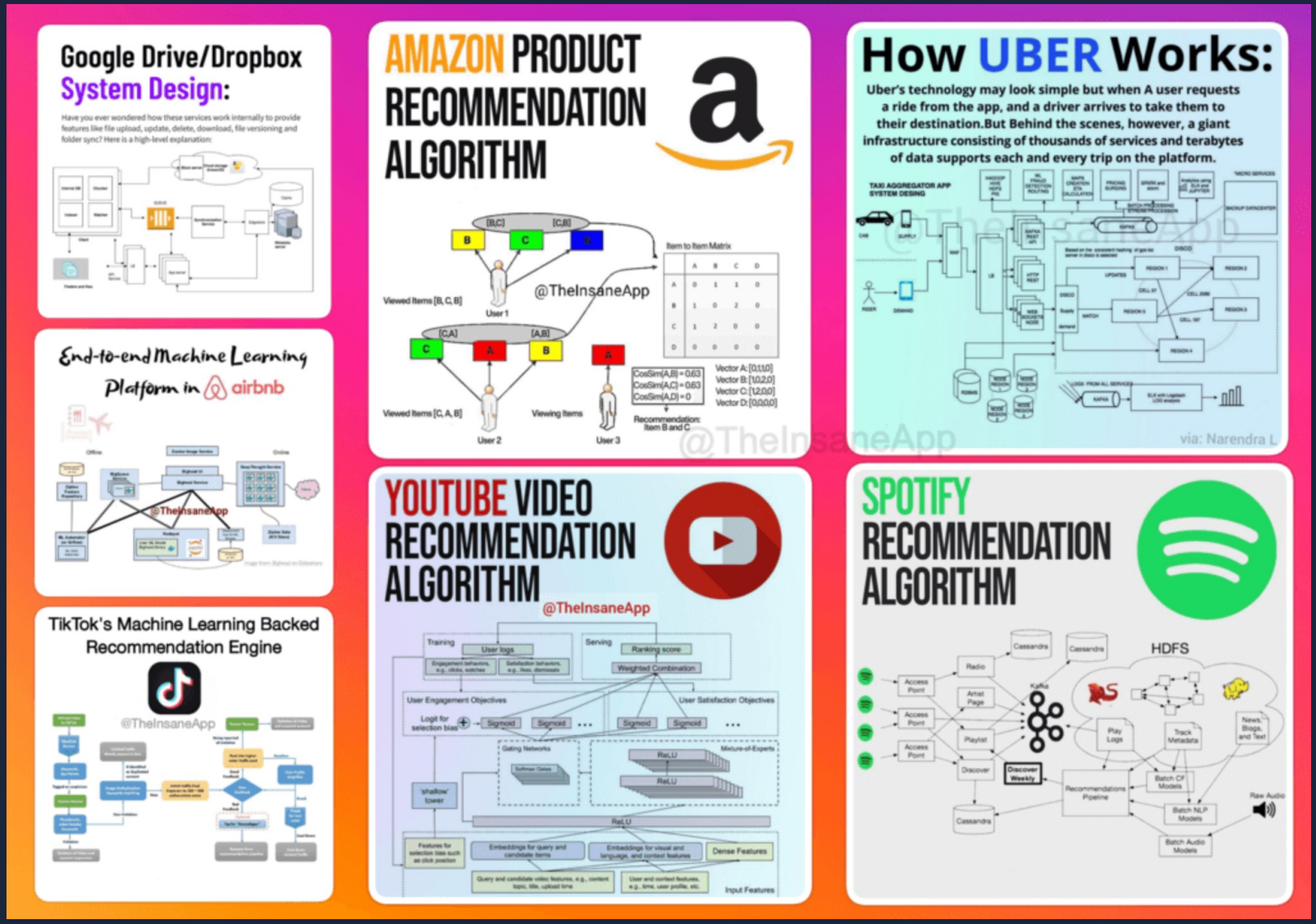
September 2019:

Legacy foreacst accuracy: 3.46%

Nostradamus accuracy: -0.10%

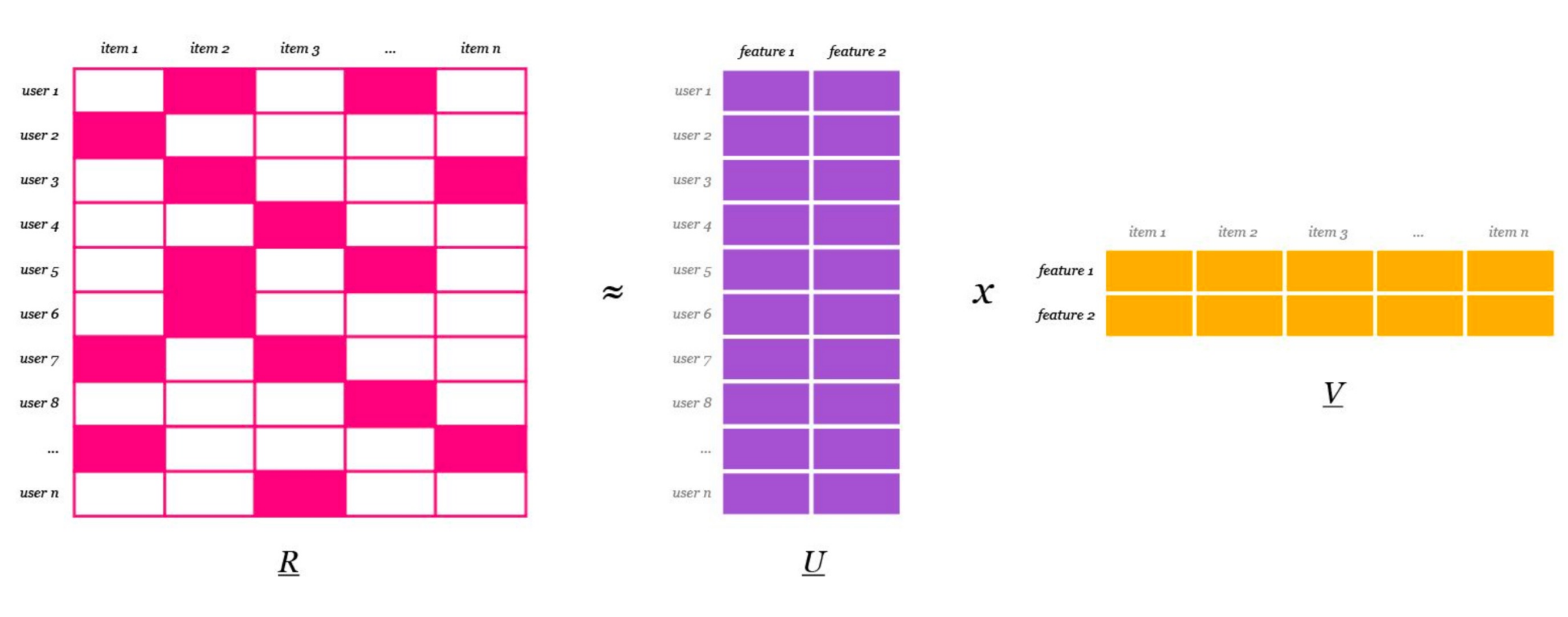
30x improvement
In forecast accuracy

Recommendation systems



Source: TheInsaneApp 2021

Recommendation systems

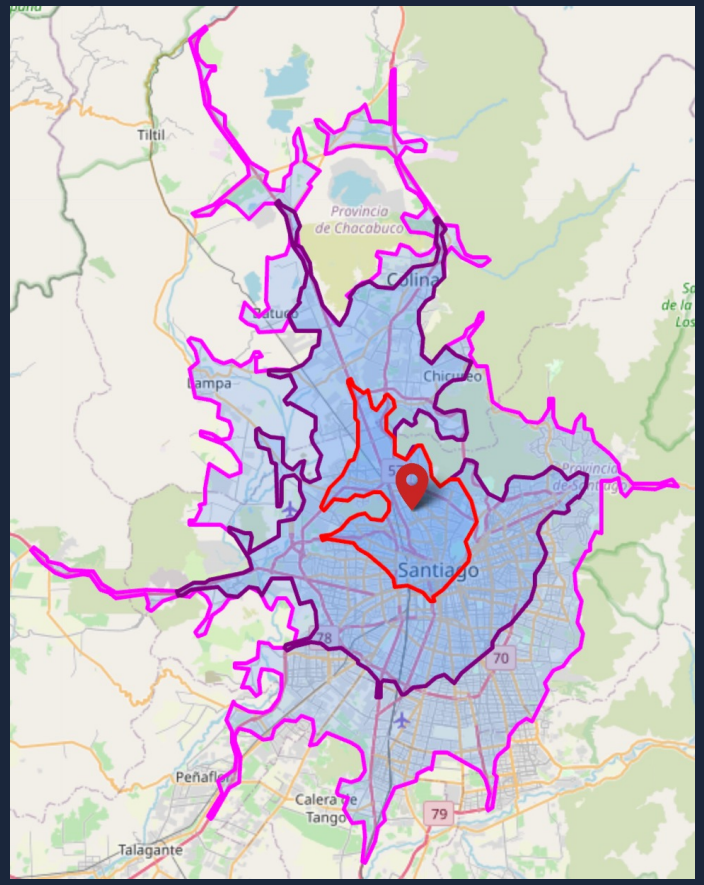


Georeferenced data

Use of publicly available data to leverage our solutions to georeferenced problems.

	30min	
Fase	**Acumulado**	**Porcentaje**
Total compañía	15 878 851	90.8%

Total personas	17 482 610



THE NEW YORKER

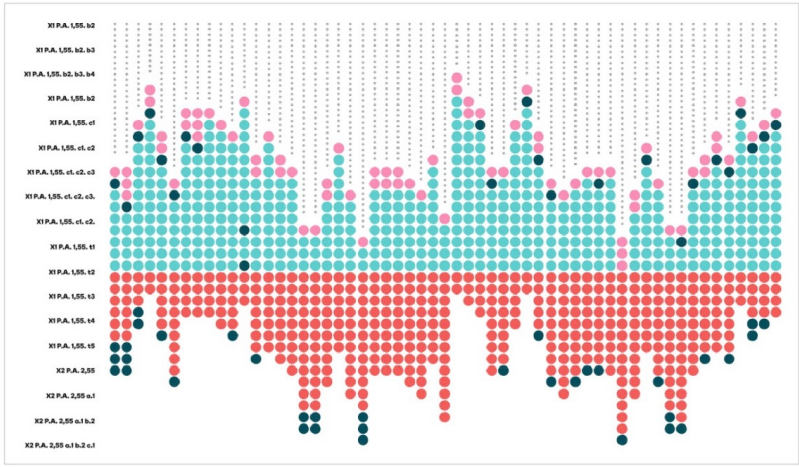
ANNALS OF COMMUNICATIONS

HOW THE MATH MEN OVERTHREW THE MAD MEN

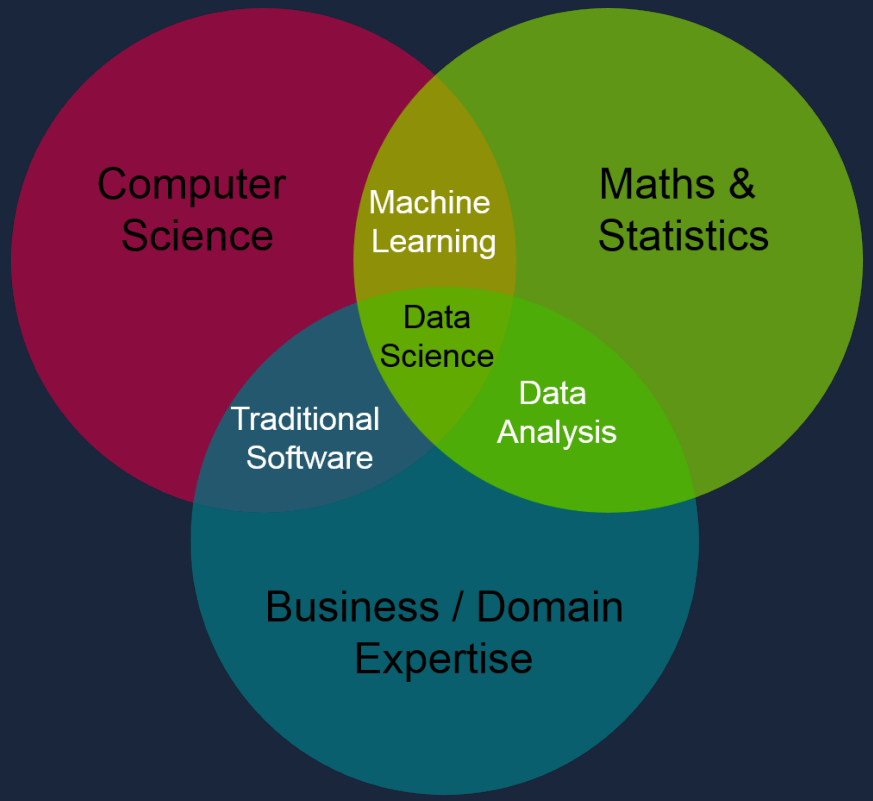
By Ken Auletta May 21, 2018

Ediciones anteriores viernes La revista de laSegunda laSegunda

viernes_reportaje



LA NUEVA INFLUENCIA DE LOS MATEMÁTICOS



Hay empleo!

Data Scientist Área Analítica Avanzada (Las Condes)
Coca-Cola Embonor
Gran Santiago, Región Metropolitana de Santiago, Chile

Data Scientist
EY
Gran Santiago, Región Metropolitana de Santiago, Chile

Data Scientist - Santiago
Teck Resources Limited
Las Condes, Región Metropolitana de Santiago, Chile

Digital Data Scientist
Nestlé
Gran Santiago, Región Metropolitana de Santiago, Chile

Data Scientist
Banco de Chile
Gran Santiago, Región Metropolitana de Santiago, Chile

CL_CIENTIFICO/A DE DATOS
Telefónica
Gran Santiago, Región Metropolitana de Santiago, Chile

Data Scientist
IFG Capital
Viña del Mar, Región de Valparaíso, Chile (Presencial)

Data Scientist
Gimnasios Energy Club
Estación Central, Región Metropolitana de Santiago, Chile

Walmart Chile | Este es el lugar

www.somoswalmartchile.cl

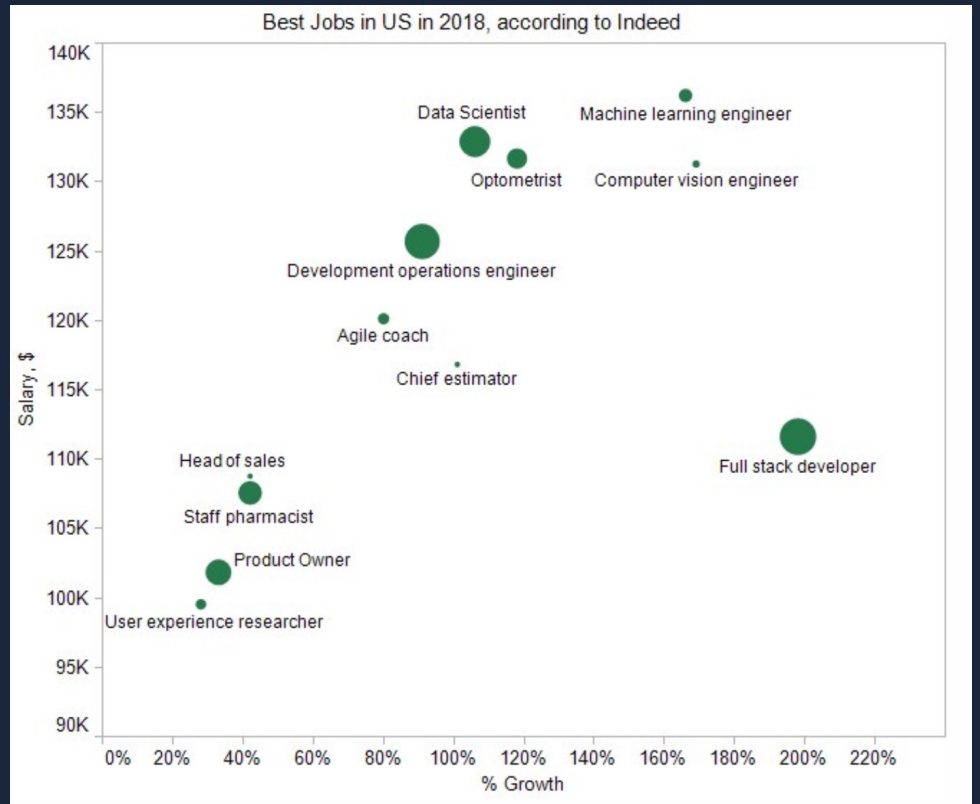


Fig. 1: Top US Jobs in 2018, based on Salary (Y-Axis) and Growth (X-Axis).
Circle size is proportional to the number of job titles.

Recursos:

https://www.kdnuggets.com/

https://datasciencecl.slack.com/

https://www.datasciencecentral.com/

http://blog.kaggle.com/

Teamwork


$1.990
$1.790
Walmart Chile


La primera pla ta de h drógeno verde
a niv l industrial de Latinoamérica

Optimization:

Mathematical modeling and optimization of product assortment.

Price-demand elasticity and price optimization.

Time series:

Hierarchical or grouped forecast of time series.

Data mining:

Asymmetrical association rules.

sebastian.torres@walmart.com