
Entrepôt de Données Pinterest •

Massila LAKAF, Imane LAZIZI, Kanzy YOUSSEF



Analyse des Besoins Métier

Contexte Pinterest

- 📌 +480M utilisateurs actifs mensuels.
- 📌 Plateforme de découverte visuelle entre réseaux sociaux et e-commerce.
- 📌 Modèle économique basé sur la publicité (promoted pins).

Actions à Analyser :

Conversions &
Achats Externes



Clics & Visites Site
Marchands



Modèle Détaillé : Conversions •



Deux mesures principales

📌 conversions_values : additive

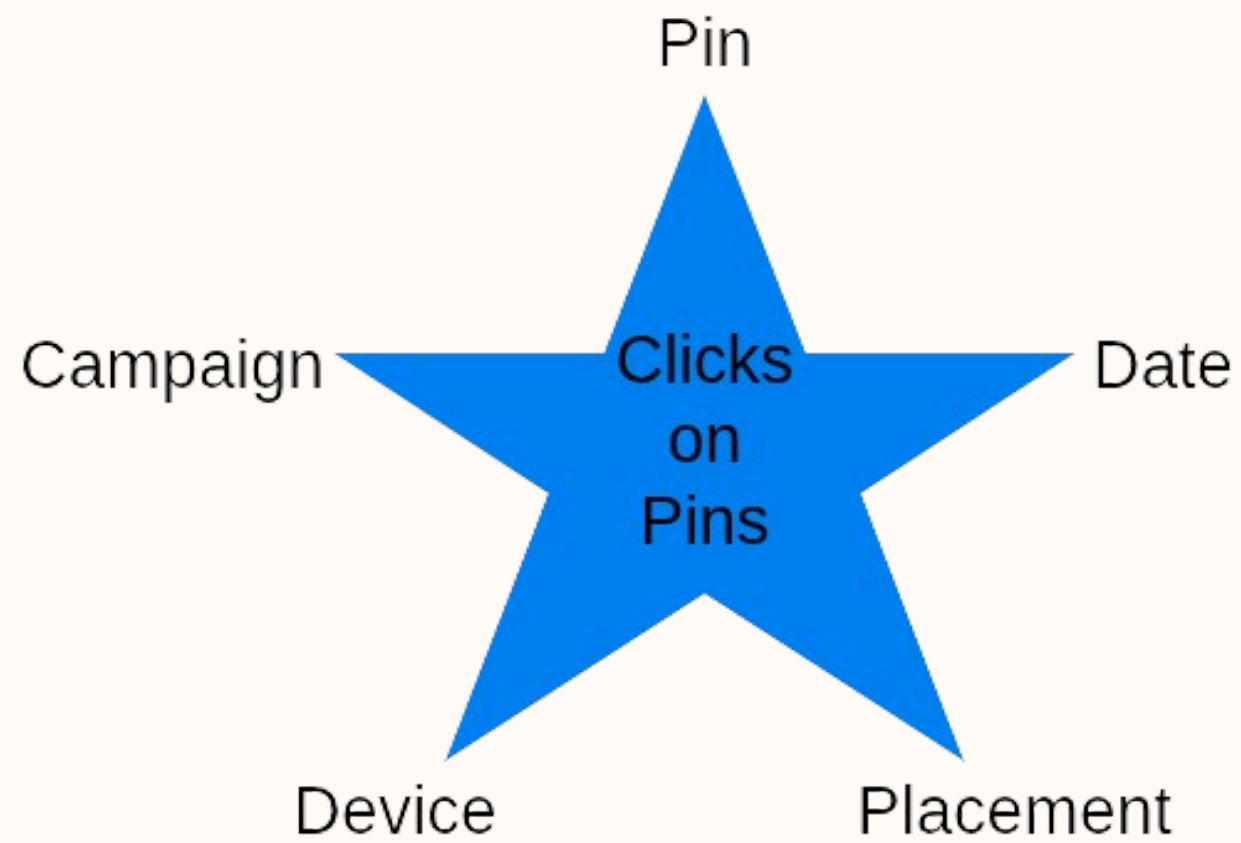
📌 conversion_count : additive

Taille de l'Entrepôt

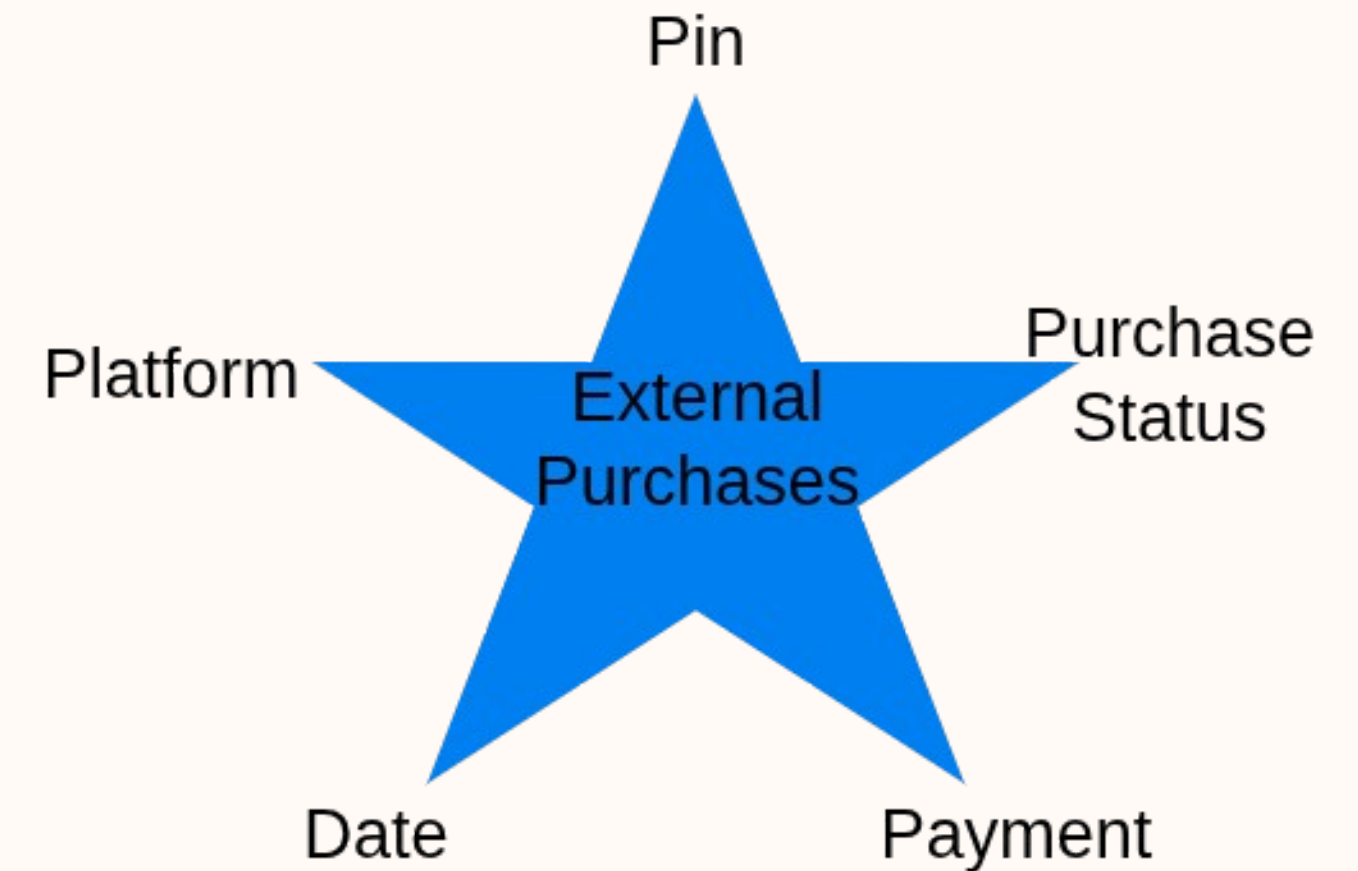
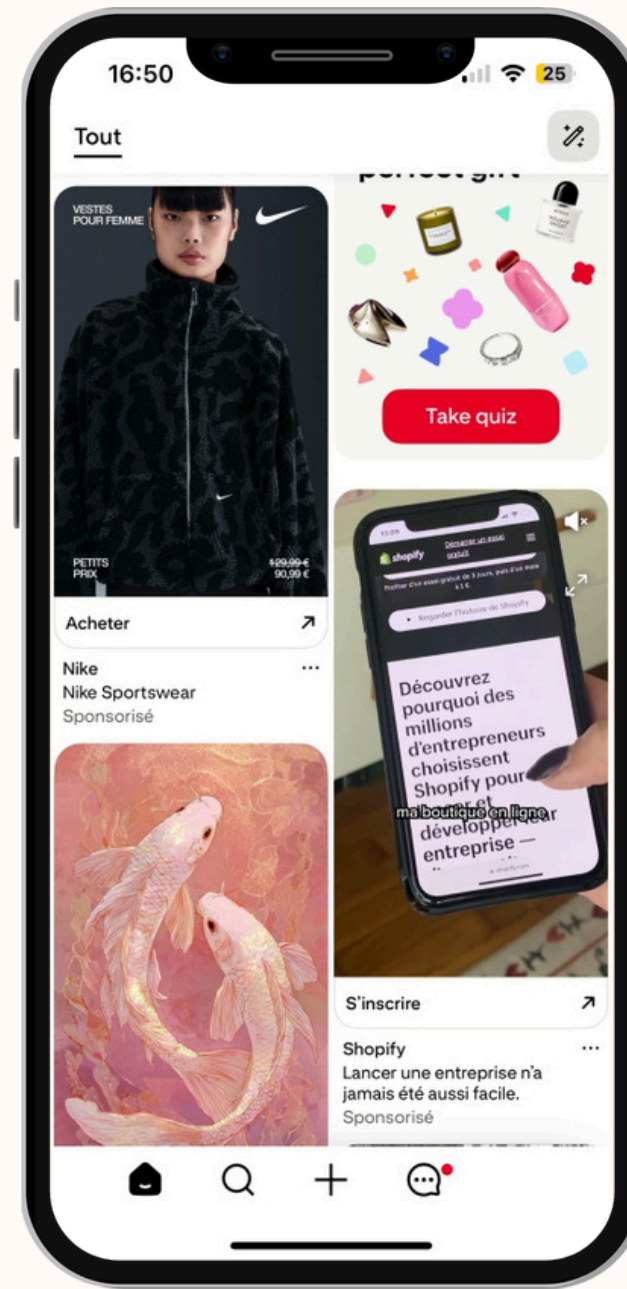
📌 115 Millions conversions / an

📌 10 Milliards de pins / an

Snapshot & Updated-Records ●



Snapshot



Updated-Records

Techniques Avancées de Modélisation

Table Pont (Relation N:M)

PinCategory
(pin_id,category_id)
pin_id
category_id

```
SELECT c.category_name, SUM(f.conversion_value) AS revenue
FROM fact_conversions f
JOIN pin_category_bridge b ON f.pin_id = b.pin_id
JOIN dim_category c ON b.category_id = c.category_id
JOIN dim_date d ON f.date_id = d.date_id
WHERE d.full_date >= DATE '2025-01-01'
GROUP BY c.category_name;
```

Partitionnement Hybride dim_pin

- 📌 **Colonnes** : dim_pin_static + dim_pin_dynamic
- 📌 **Lignes** : RANGE par created_at + HASH par
pin_id + index sur is_promoted

Implémentation & Requêtes Analytiques

SELECT

c.category_id, t.period_of_the_day,

SUM(f.conversion_value) **as** revenue,

COUNT(*) **as** conversion_events

FROM fact_conversion f

JOIN dim_category c **ON** f.category_id
= c.category_id

JOIN dim_time t **ON** f.time_id =
t.time_id

GROUP BY c.category_id,
t.period_of_the_day;

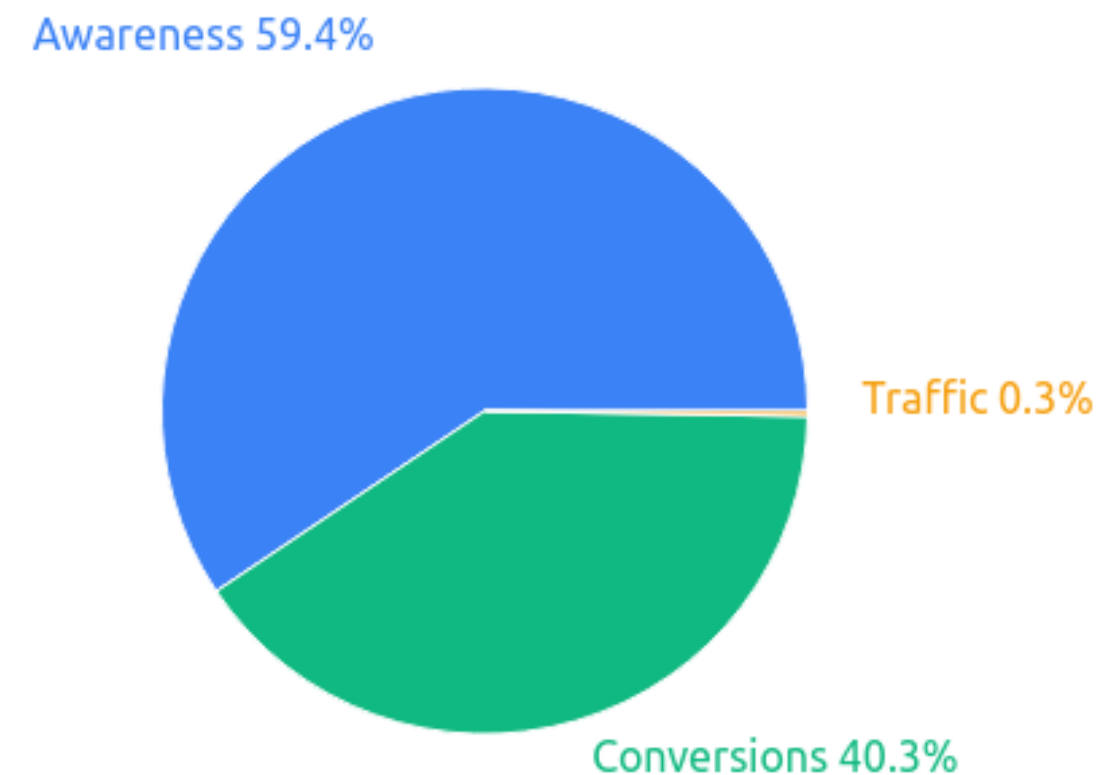
CATEGORY_ID	PERIOD_OF_	REVENUE	CONVERSION_EVENTS
8	Morning	1110	2
5	Afternoon	1299,99	3
10	Evening	349,98	5
3	Evening	140	1
4	Morning	0	1
4	Afternoon	1198,99	4
10	Night	175	1
3	Night	220	1
9	Afternoon	125,5	1
6	Night	599,99	1
5	Evening	450	1

CATEGORY_ID	PERIOD_OF_	REVENUE	CONVERSION_EVENTS
9	Evening	230	4
9	Morning	110	3
2	Morning	120	2
2	Afternoon	350,49	6
10	Afternoon	409,99	2
8	Evening	750	1
7	Morning	113	4
10	Morning	0	2
8	Afternoon	1499,98	5

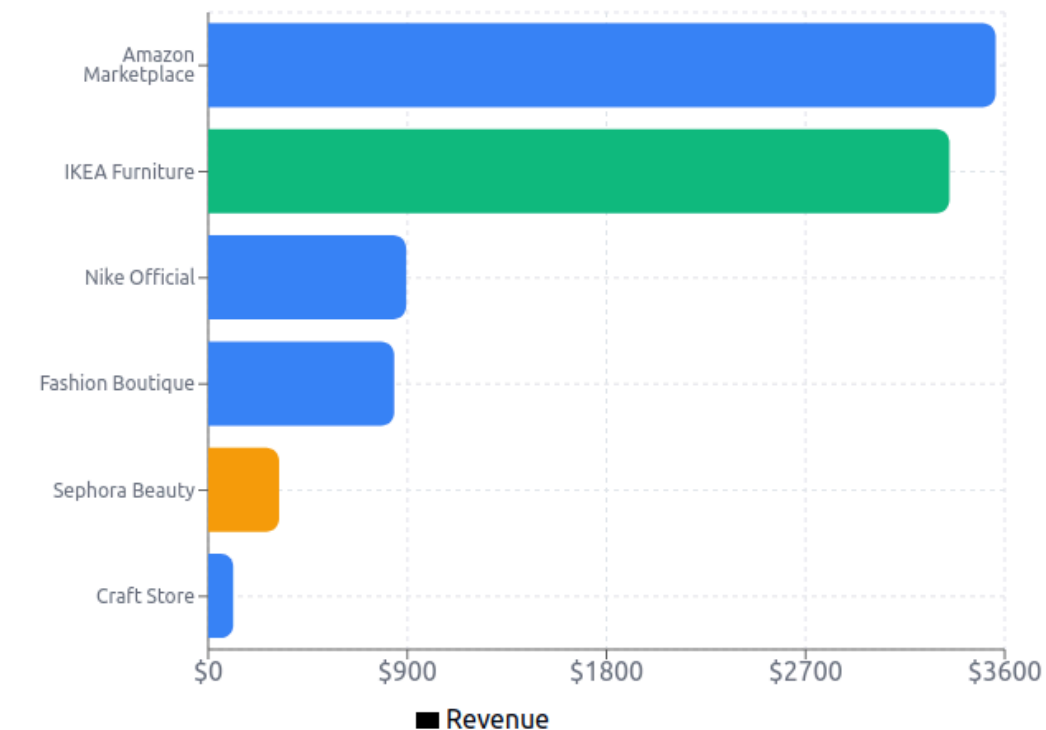
20 lignes selectionnees.

**Performance des catégories par
moment de la journée**

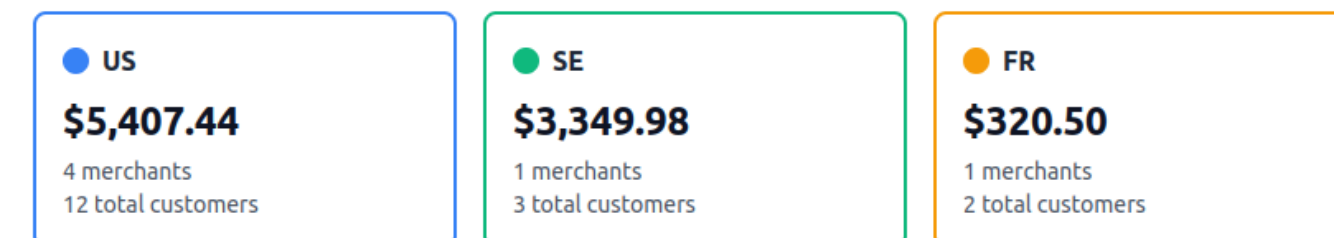
Autres Exemples de Requêtes



**Tendance des revenus
par objectif de campagne en
Janvier 2025.**



Performance by Country



Top Marchant par pays.

Optimisation

Vues Matérialisées (Treillis d'Agrégation)

MV1 : mv_daily_campaign_merchant

MV2 : mv_user_cohort

MV3 : mv_category_time

Index Bitmap Join

📌 bji_campaign_objective : Optimise filtres sur objectif campagne (5-10 valeurs)

📌 bji_is_weekend : Optimise filtres booléens weekend/semaine



Conclusion

Merci pour votre écoute !



Bibliographie

Pinterest Help Center : <https://help.pinterest.com/en/business/article/install-the-pinterest-tag>

Track Conversions with Pinterest Tag. :<https://help.pinterest.com/en/business/article/track-conversions-with-pinterest-tag>

Add Event Codes: <https://help.pinterest.com/en/business/article/add-event-codes>

Pinterest for Business. Available on : <https://business.pinterest.com/en-gb/>

Pinterest statistics: <https://truelist.co/blog/pinterest-statistics>

Kimball, R. & Ross, M. The Data Warehouse Toolkit. 3rd Edition, Wiley, 2013.

Inmon, W. H. Building the Data Warehouse. 4th Edition, Wiley, 2005.

Oracle Documentation : <https://docs.oracle.com/en/database/oracle>

