


(Big) Data Engineering In Depth

From Beginner to Professional

Mostafa Alaa Mohamed

Senior Big Data Engineer

 MoustafaAlaa  Moustafa Alaa  @Moustafa_alaa22

 mustafa.alaa.mohamed@gmail.com

¹Big Data & Analytics Department, Epam Systems

The Definitive Guide to Big Data Engineering Tasks

Videos classification

Watching Method / Audience	Computer	Mobile/Tablet	Just listening
Developer	●		
DevOps	●		
Business	●		

Table: Video classification

The green circle ● means short video.

The blue circle ● means medium video.

The red circle ● means long video

Dimensions Types: Shrunken Rollup Dimensions

Shrunken Rollup Dimensions

- Shrunken Rollup dimension is used for developing aggregate (higher level of summary) fact tables.
- It required that the data model has a lower level of granularity.

Example

- We have a daily usage fact table, and we need to have a higher level of monthly usage. So, we use the monthly dimension to get a summary of the daily.
- We have a daily usage fact table aggregated on area-id, and we need to create another summary table aggregated based on city id. So, the new grain level here is the new dimension for the city.

Shrunken Rollup Dimensions Cont.

OrderDate	AreaID	TotalOrders
123456789	123	20
123456789	123	30
123456789	678	10
123456789	678	12

AreaID	AreaName	CityID
123	Al-Matareya	1
678	Ain shams	1

CityID	CityName
1	Cairo

OrderDate	CityID	TotalOrders
123456789	1	72