tugger

Power BI:

The Tugger Toolkit

tugger

I. Foundation & ETL

1.2 Fact Granularity & Real-World Choices

Timothy Amison
Power BI Developer

The Problem

1. Data at **different levels**, for example:
   - Order **Header** (Total)
   - Order **Line** (Product/Item)
2. User needs to **filter** everything by the line

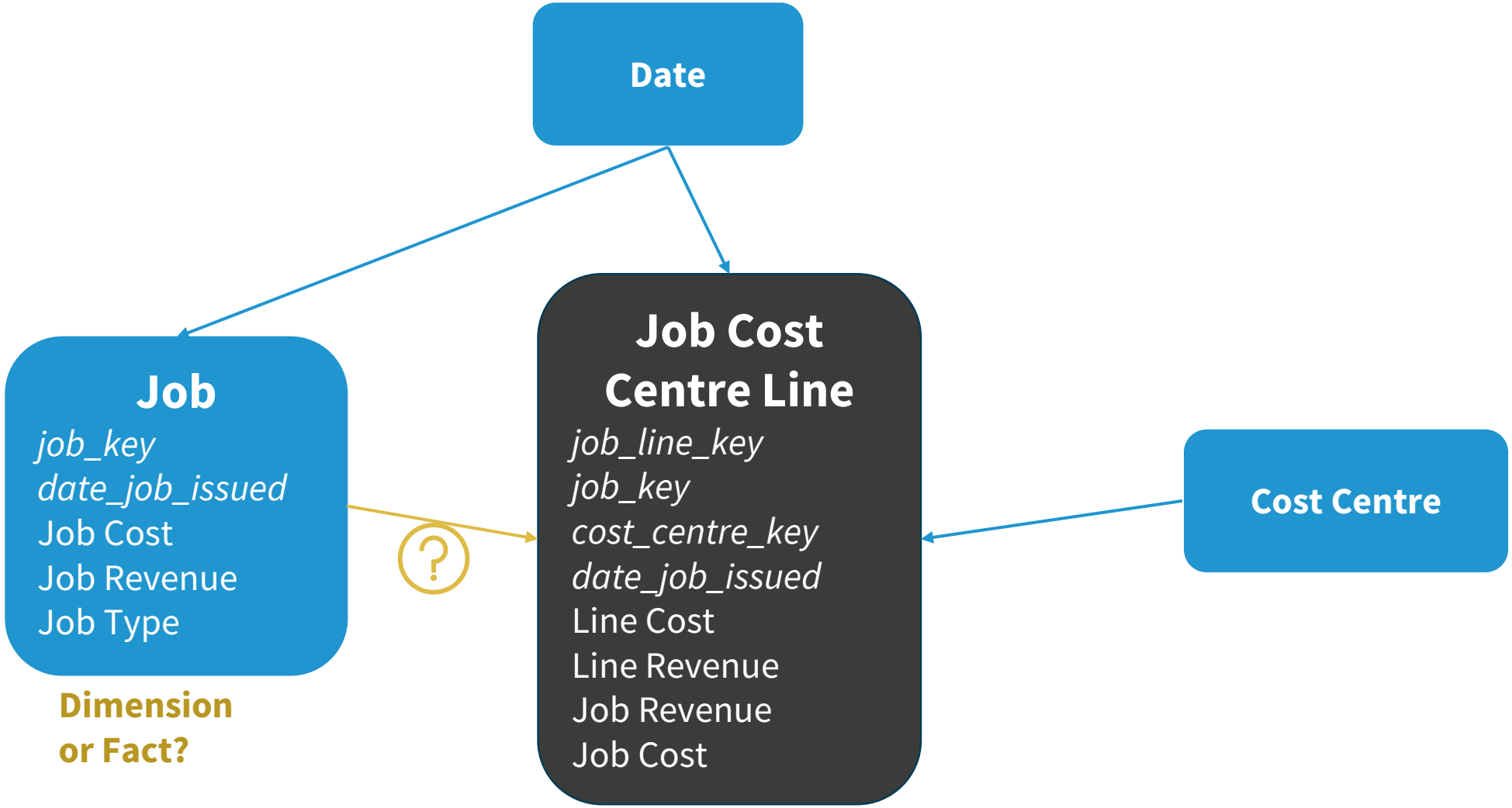
Job (Header)

Job Key	Cost	Revenue
1	80	100
2	150	200

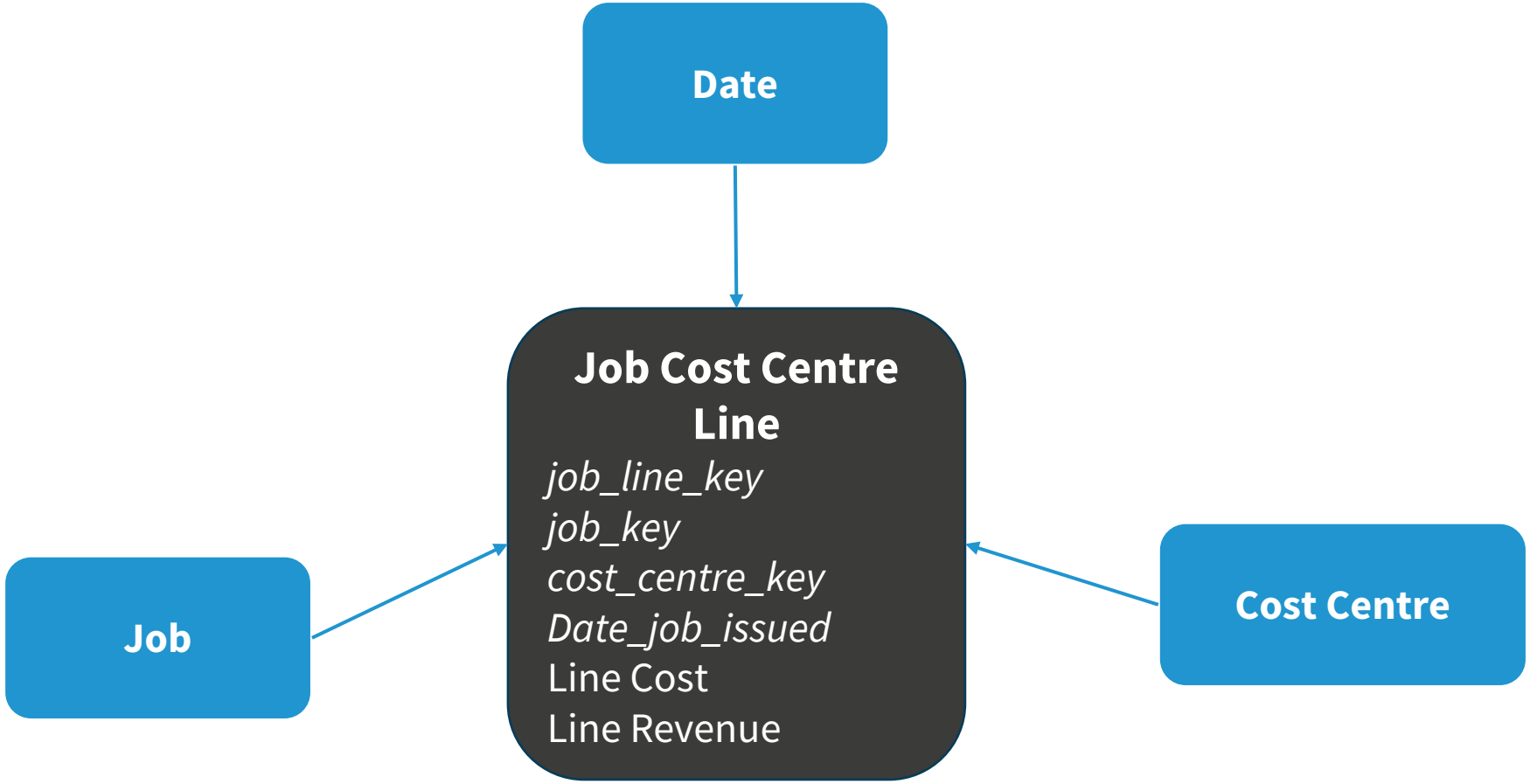
Job Cost Centre Line (Line)

Job Line Key	Job Key	Cost Centre Key	Cost	Revenue
1	1	1	30	50
2	1	2	50	50
3	2	1	70	90
4	2	2	80	110

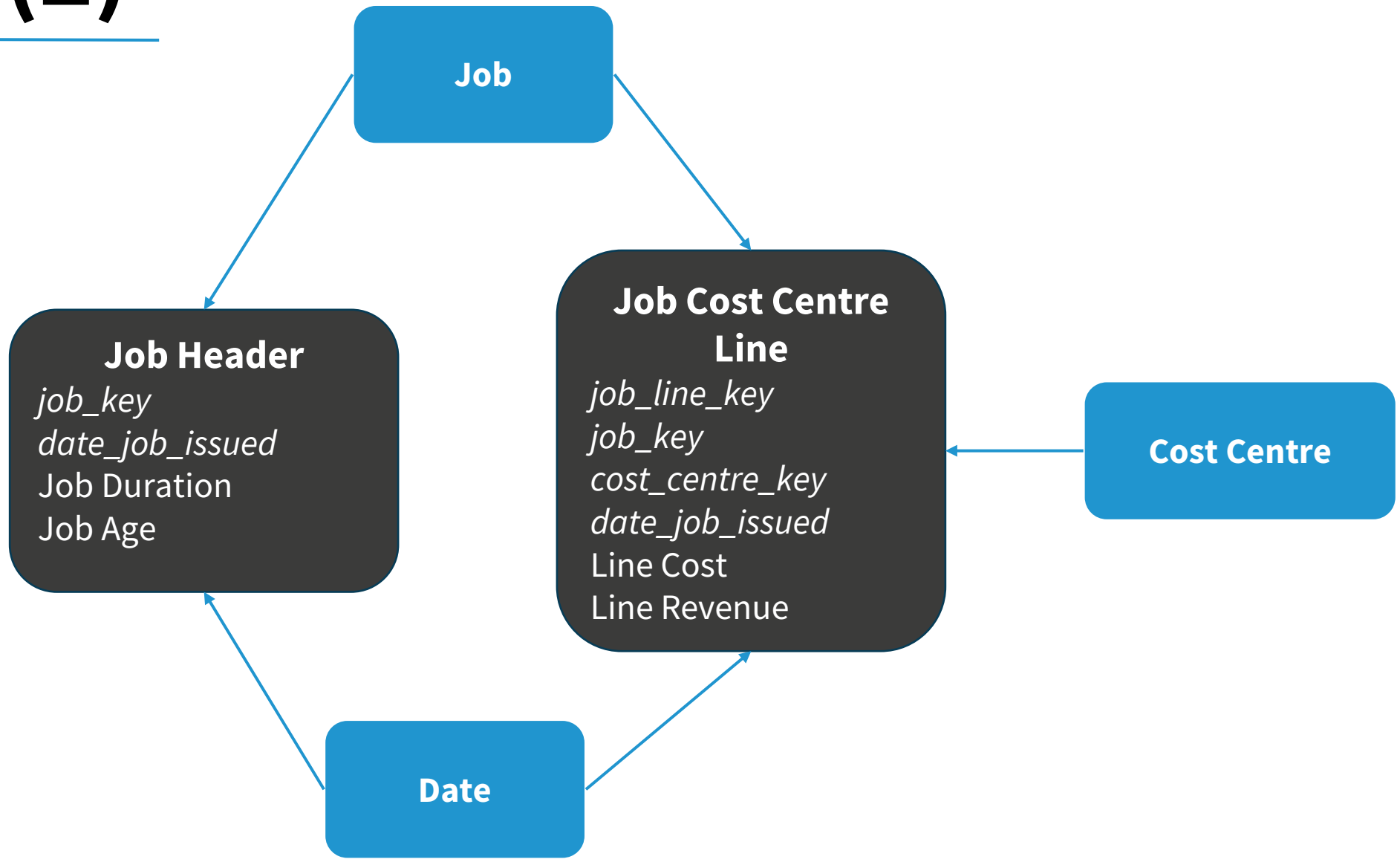
Source Model (Simplified)



Our Goal (1)


Date
Job Cost Centre Line
job_line_key
job_key
cost_centre_key
Date_job_issued
Line Cost
Line Revenue
Job
Cost Centre

Our Goal (2)


Job
Job Header
job_key
date_job_issued
Job Duration
Job Age
Job Cost Centre Line
job_line_key
job_key
cost_centre_key
date_job_issued
Line Cost
Line Revenue
Cost Centre
Date

“

The **User’s Filter** defines the
Model’s Grain

”

Our Progress

1. **“Build for Power BI**, not the Source Database”

2. **4 Rules** to Guide Modelling Choices:
    1. Same Fact Grain
    2. Thin Fact
    3. One-Table Dimension
    4. No Messy Joins

1.1 – Star Schema Architecture & Integrity

1.2 – Fact Granularity & Real-World Choices

1.3 – Modelling Degenerate Dimensions

1.4 – Event-Based Modelling for HR/CRM

1.5 – Data Quality Strategies

1.6 – ETL Architecture & Staging Layer

Today

1. **“Build for Power BI**, not the Source Database”

2. **4 Rules** to Guide Modelling Choices:
   1. Same Fact Grain
      - “**User’s Filter** defines the Model’s Grain”
   2. Thin Fact
   3. One-Table Dimension
   4. No Messy Joins

1.1 – Star Schema Architecture & Integrity

1.2 – Fact Granularity & Real-World Choices

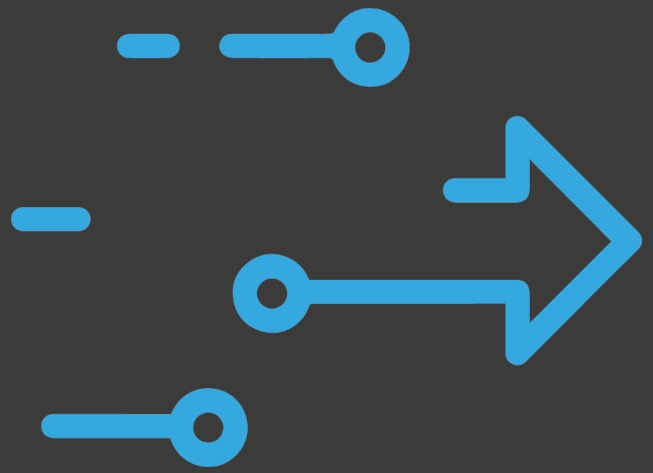
1.3 – Modelling Degenerate Dimensions

1.4 – Event-Based Modelling for HR/CRM

1.5 – Data Quality Strategies

1.6 – ETL Architecture & Staging Layer

tugger



What is Granularity?

Granularity

The level of detail of each row

Example:
Orders

- Header Fact: One row per **order**
- Line Fact: One row per **product/item** line

Caution:
Mixing Grains can cause issues, especially with:

- Filters
- Aggregation

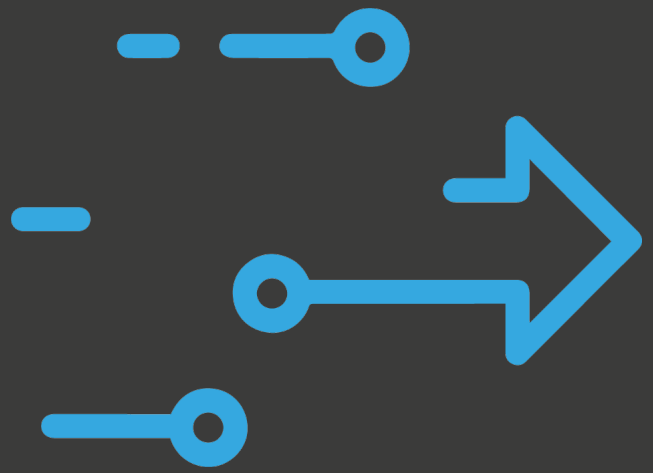
Job (Header)

Job Key	Cost	Revenue
1	80	100
2	150	200

Job Cost Centre Line (Line)

Job Line Key	Job Key	Cost Centre Key	Cost	Revenue
1	1	1	30	100
2	1	2	50	100
3	2	1	70	200
4	2	2	80	200

tugger



Why Didn’t I Use a Header Fact?

Job Cost Centre Example

Problem

- Users need to **filter all** measures by **Cost Centre**.
- **Can't filter Header Facts** by Cost Centre accurately.

Solution

- **One Fact** at **line level**
- Calculate header-level measures (totals) from the line-level fact.

Outcome

- All filters, including **cost centre**, **work consistently**.

Job (Header)

Job Key	Cost	Revenue
1	80	100
2	150	200

Job Cost Centre Line (Line)

Job Line Key	Job Key	Cost Centre Key	Cost	Revenue
1	1	1	30	50
2	1	2	50	50
3	2	1	70	90
4	2	2	80	110

Employee Absence Example

Problem

- Users need to **filter all** measures by **Weekday**.
- **Can't filter Absence Facts** by Weekday accurately.

Solution

- **One Fact** at **day level**
- Calculate header-level measures (totals) from the line-level fact.

Outcome

- All filters, including **weekday**, **work consistently**.

Leave (Sickness & Paid Time Off)

Leave Key	Start Date	Days Off
1	2024-01-01	1
2	2025-05-01	2.5

Daily Leave

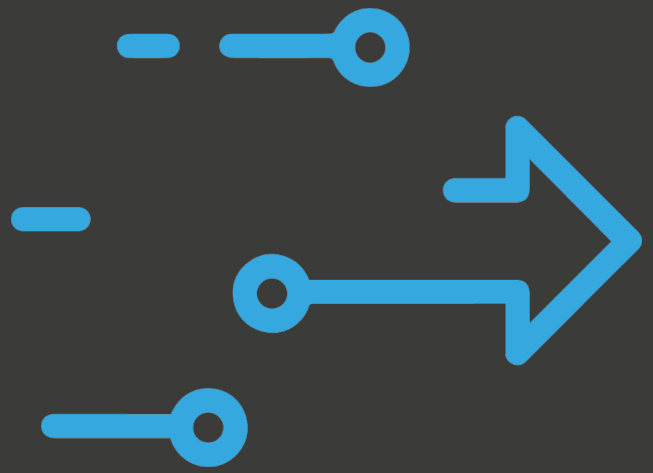
Leave Key	Date	Days Off
1	2024-01-01	1
2	2025-05-**01**	1
2	2025-05-**02**	1
2	2025-05-**03**	0.5

Implications & Trade-offs

- **Simpler** Model.
- Fewer Tables.
- **Filters work** as expected
- **If you need to filter by the lowest grain**, this is ideal.

- **Header-level facts** might be **different** to the sum of lines.
- **Can't show** or reconcile **differences** between header and line totals.
- **If header-level facts are independently important**, consider having 2 fact tables (header and line).

tugger



When *Would* I Use a Header Fact?

The Allocation Trade-Off and a Dual-Fact Model

The Allocation Problem

*Using **Header Facts** at the Line Grain*

- Need to report on **header-level** metrics.
- **Missing data** at line-level.
- Repeating the total for all lines can make **measures wrong or difficult.**

Job (Header)

Job Key	Cost	Revenue
1	80	100
2	150	200

Job Cost Centre Line (Line)

Job Line Key	Job Key	Cost Centre Key	Cost	Revenue
1	1	1	30	100
2	1	2	50	100
3	2	1	70	200
4	2	2	80	200

The Dual-Fact Solution

Problem

- Need to report on **header-level** metrics
- **Missing data** at line-level
- Repeating the total for all lines can make **measures wrong or difficult**

Solution

- Reintroduce the **header-level** table
- 2 fact tables: Header & Line

Outcome

- **Accurate** Header-level metrics
- **Can't filter** Revenue or Revenue derived measures e.g. Profit, Margin.

Job (Header)

Job Key	Cost	Revenue
1	80	100
2	150	200

Job Cost Centre Line (Line)

Job Line Key	Job Key	Cost Centre Key	Cost	Revenue
1	1	1	30	?
2	1	2	50	?
3	2	1	70	?
4	2	2	80	?

The Custom Allocation Solution

Problem

- Need to report on **header-level** metrics
- **Missing data** at line-level
- **Dual-fact solution: Can't filter** Revenue, Profit or Margin.

Solution

1. Custom **Allocation at line level**

$$Total\ Revenue * \frac{Line\ Cost}{Total\ Cost}$$

Outcome

- All filters, including cost centre, **work consistently**.
- Relies on **assumed allocation** metric

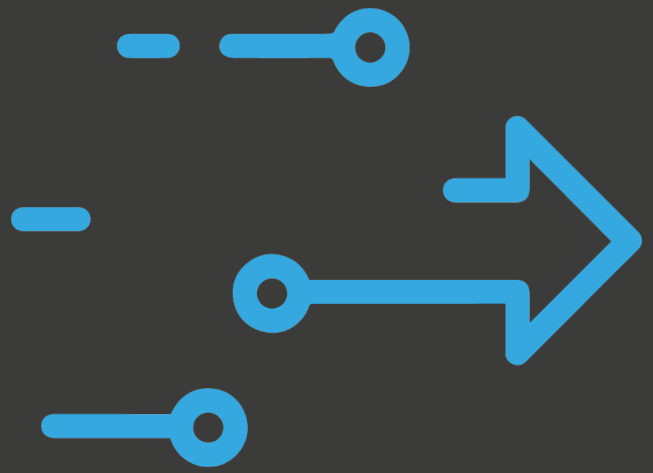
Job (Header)

Job Key	Cost	Revenue
1	80	100
2	150	200

Job Cost Centre Line (Line)

Job Line Key	Job Key	Cost Centre Key	Cost	Revenue
1	1	1	30	100 x (30/80)
2	1	2	50	100 x (50/80)
3	2	1	70	200 x (70/150)
4	2	2	80	200 x (80/150)

tugger



When I *Did* Use a Header Fact

The Allocation Trade-Off and a Dual-Fact Model

The Dual-Fact Necessity

Problem

- Need to report on **header-level** metrics
- **Missing data** at line-level
- Repeating the total for all lines can make **measures wrong or difficult**

Solution

- Reintroduce the **header-level** table
- 2 fact tables: Header & Line

Outcome

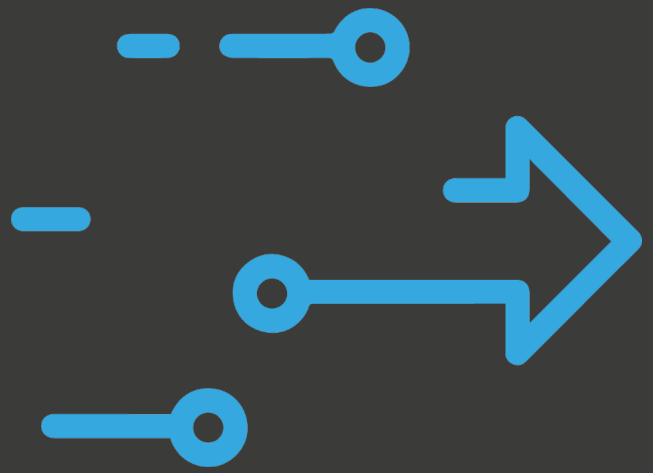
- **Accurate** Header-level metrics
- **Can't filter** Days to Win or other Header values by Cost Centre.

Quotes (Header)

Quote Key	Days to Win
1	80
2	150

Quote Cost Centre Line (Line)

Quote Line Key	Quote Key	Cost Centre Key	Cost	Revenue
1	1	1	30	50
2	1	2	50	50
3	2	1	70	90
4	2	2	80	110



Key Takeaways

And Actionable Steps Going Forward

Jargon Decoder

Term	Plain English
Grain / Granularity	The **level of detail** chosen for the Fact Table (e.g., per line item vs. per entire invoice)
Header Fact	A measure that applies to the **entire transaction** (e.g., total shipping charge for the whole job).
Line Fact	A measure that applies to an **individual item or line** within a transaction (e.g., job cost per specific cost centre).

Key Takeaways

1. **The User's Filter** is the Model's Grain.

2. **Choose the lowest grain** which all dimensions can connect and filter consistently.

3. **Data isn't always perfect**: your job is to be aware of the options and decide what is your priority for your goals.

What’s Next

1.1 – Star Schema Architecture & Integrity

1.2 – Fact Granularity & Real-World Choices

1.3 – Modelling Degenerate Dimensions

1.4 – Event-Based Modelling for HR/CRM

1.5 – Data Quality Strategies

1.6 – ETL Architecture & Staging Layer

Q&A

Let’s Chat!

- Any Questions?
- Topics you would like to be covered in future sessions?

tugger

See the full webinar schedule

https://tuggerapp.com/Webinars

Website: **https://tuggerapp.com**
LinkedIn: **www.linkedin.com/in/timothy-amison**
Email: **timothy.amison@tuggerapp.com**

tuggerapp.com