

# Three concepts about the Star Schema

every beginner should know



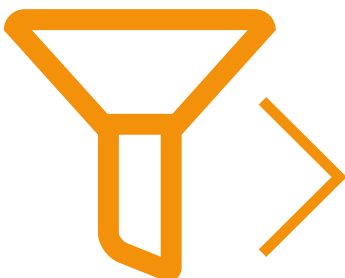
## Fact Table

- There is (*usually*) only one per model
- Records what you are reporting on
- **Should** have the fewest number of columns
- Will (*usually*) have the most number of rows
- Numeric data types should be preferred
- Every row will have a date, maybe a time



## Dimension tables

- Describe the facts in detail
- Usually will have lots of columns
- Text fields are expected and normal
- Can be many dimension tables per model
- Will have many fewer rows than the fact table
- Only the "key" column should have high cardinality



## Relationships

- Can be auto-detected when data is imported
- Gold standard is one-to-many/many-to-one
- Single direction filters are preferable
- Dim tables sit on the one side of the relationship
- The fact table sits on the many side
- Relationships transmit filters between tables