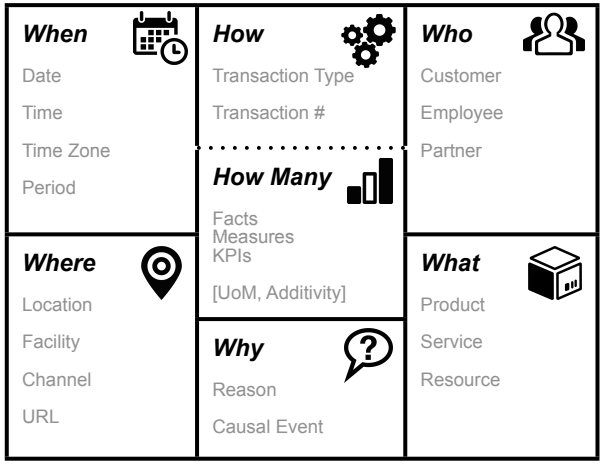
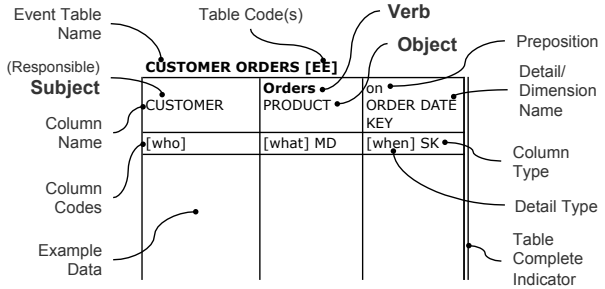
BEAM✲

Business Event Analysis & Modeling

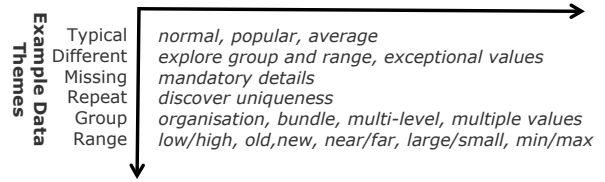
Agile Dimensional Modeling

Who does what?
When and where?
How much / how many?
Why and how?





7W Details: Who, What, When, Where, How Many, Why, How



© DecisionOne Consulting +44 (0)7971 964794
www.decisionone.co.uk information@decisionone.co.uk

Event and Fact Table Types

[DE] **Discrete Event.** Point in time or short duration (completed) transaction.

[EE] **Evolving Event.** (multi-verb) process that takes time to complete.

[RE] **Recurring Event.** Measurements taken at predictable regular intervals.

[TF] **Transaction Fact table.** Physical equivalent of **DE**. Typically maintained by insert only.

[AS] **Accumulating Snapshot.** Physical equivalent of **EE**. Maintained by insert and update. Typically contains multiple milestone date/time dimensions and duration facts.

[PS] **Periodic Snapshot.** Physical equivalent of **RE**. Typically contains semi-additive facts.

[AG] **Aggregate.** Fact table that pre-summarizes an existing detailed fact table.

[DF] **Derived Fact table.** Fact table constructed by merging, slicing, or pivoting existing fact tables.

Dimension Types

[CV] **Current Value.** Contains current value only dimensional attributes. *Type 1 SCD*.

[HV] **Historic Value.** Contains at least one historical value dimensional attribute. *Type 2 slowly changing dimension (SCD)*.

[RP] **Role-Playing.** Used to play multiple roles.

[RU] **Roll-Up.** Derived from a more granular dimension.

[SD] **Swappable Dimension.** Part of a set of dimensions with a common surrogate key that can be used in place of each other.

[ML] **Multi-Level.** Dimension containing additional members representing higher levels in the dimension's hierarchy.

[HM] **Hierarchy Map.** Table used to resolve a recursive relationship. Represents a variable-depth hierarchy.

[MV] **Multi-Valued.** Bridge table used to resolve a many-to-many relationship between a fact table and a multi-valued dimension.

[PD] **Pivoted Dimension.** Contains column flags built from the row values of another dimension.

General Column Types

MD **Mandatory.** Value is present under normal conditions. Can be nullable to handle errors.

NN **Not Null.** Column does not allow nulls. All **SK** and **FK** columns are **NN** by default.

ND **NDn** **No Duplicates.** Numbered to define combinations of column values that must be unique. **PK** columns are **ND** by default.

Xn **Exclusive.** Column is not valid in combination with other **X** columns. Numbered to identify mutually exclusive groups and identify the specific **DC** which controls validity.

DC **DC**n,n **Defining Characteristic.** Column value dictates which **X** columns are valid. E.g., `Product Type DC` defines which exclusive product dimension attributes are valid. Number list relates multiple defining characteristics in the same table to specific **X**n exclusive columns or groups.

[Wtype**]** **[**dimension**]** **Dimension type or name**. The *W type (who, what, when, where, why, how)* of an event detail or the dimension name when a detail is a role; e.g., `Salesperson [Employee]` where `Salesperson` is a role of the `Employee` dimension. Also used to describe recursive relationships.

Event and Fact Table Column Types

DD **Degenerate Dimension**. Dimensional attribute stored in a fact table. Typically used for transaction IDs (*how* details).

GD **GD**n **Granularity Dimension.** Dimension combination that defines the granularity of a fact table. Numbered when alternative combinations exist.

MV **Multi-Valued.** Event detail contains multiple values that must be resolved using a bridge table. Fact table **FK** that references a multi-value bridge table.

ML **Multi-Level.** Event detail can represent various levels in a hierarchy; e.g., individual employee or teams/branches. Fact table **FK** that points to a multi-level dimension *and* makes use of the additional levels.

Fact Types

FA — **Fully Additive.** Fact that produces a correct total when summed across any combination of its dimensions. For a fact to be (fully) additive it must be expressed in a single unit of measure. Percentages and unit prices are not additive.

SA / **SA**n — **Semi-Additive**. Fact that can be correctly totaled by some dimensions but not by at least one non-additive (**NA**) dimension: e.g., an account balance cannot be summed over time: its **NA** dimension. **SA** facts are often averaged over their **NA** dimension.

SA is always used in conjunction with at least one **NA** dimension to relate the semi-additive fact to its non-additive dimension(s).

Numbering relates multiple **SA**n facts in the same table to their specific **NA**n dimension(s).

NA / **NA**n — **Non-Additive.** Fact that cannot be aggregated using sum; e.g., `Temperature NA`. Non-additive facts can be aggregated using functions such as min, max, average.

Non-additive dimension of a semi-additive fact. Numbering relates multiple non-additive dimensions in the same table to specific semi-additive (**SA**n) facts.

DF / **DF=** formulae — **Derived Fact**. Value can be derived from other columns within the same table. May be followed by a simple formula referencing other facts or date/time details by number; e.g., `Unit Price DF=Revenue/Quantity`.

[UoM**]** / **[**U1, U2... **]** — **Unit of Measure**. Unit of measure symbol or description; e.g., `Order Revenue [$]` or `Delivery Delay [days]`.

List denotes that multiple units can be recorded for a quantity. They must be converted into a standard unit (**U1**) to produce an additive fact. Can also be use to document the list of conversion factors required at reporting time.

Dimensional Attribute Types

CV / **CV**n — **Current Value.** Attribute records current values only. Changes overwrite previous values. Supports *"as is"* reporting. Also known as a *type 1 slowly changing dimension (SCD)*.

Combined with **HV** to define hybrid **CV/HV** attributes with default **CV** behavior listed first. Implemented as separate **CV** & **HV** attributes.

Combined with **PV** to define hybrid **CV/PV** attributes or numbered to relate separate **CV**n attributes to matching **PV**n attributes.

HV / **HV**n — **Historic Value.** Attribute records historical values. Changes cause new versions of dimension members to be created: preserving their historically correct values. Supports *"as was"* reporting. Also known as a *type 2 SCD*.

Combined with **CV** to define hybrid **HV/CV** attributes with default **HV** behavior listed first. Implemented as separate **HV** & **CV** attributes.

Numbering defines *conditional* **HV**n attributes groups: combinations of attributes that only act as **HV** when every member of their *n* group changes at the same time. Used in combination with **CV** to treat small changes or corrections as **CV**; e.g., `Street CV, HV1` and `Zip Code CV, HV1` will be treated as **CV** individually but as **HV** if both change at once.

FV — **Fixed Value.** Attribute values do not change over time; e.g., `Date of Birth FV`. Corrections overwrite previous incorrect values: behaves like a **CV** attribute. Also know as a *type 0 SCD*.

PV / **PV**n — **Previous Value.** Attribute records previous values. Supports *"as previously"* or *"as at"* reporting. Also known as a *type 3 SCD*.

Combined with **CV** to define hybrid **CV/PV** attributes or numbered to relate separate **PV**n attributes to their matching **CV**n attributes; e.g., `Previous Territory PV1` and `Territory CV1`.

PV attributes can also hold initial or "as at date" values; e.g., `Initial Territory PV1` or `YE2010 Territory PV1`.

Key Types

PK — **Primary Key**. A column or group of columns that uniquely identifies each row in a table.

FK — **Foreign Key**. A column that references the primary key of another table.

SK — **Surrogate Key**. Anonymous integer assigned by the data warehouse as the primary key for a dimension table. Dimensional foreign key in fact tables. Denotes that example data will be replaced by integer keys.

BK — **Business Key**. Source system key.

NK — **Natural Key.** Key used in the real world.

RK — **Recursive Key.** Foreign key that references the primary key of its own table. Often used to represent variable-depth hierarchies. Used to build [HM] hierarchy maps.

Data Types

Cn — **Character**. Number defines the maximum length, overriding any default length.

DTn — **Date/Time**. Number is used in duration formulas for derived facts; e.g., `Delivery Delay DF=DT2-DT1`. Number can denote default order of milestones within an [EE].

Dn — **Date**. Number is used in duration formulas for derived facts. Number can denote default order of milestones within an [EE].

Nn.n — **Numeric**. Number defines precision, overriding the default precision.

Tn — **Text**. Long character data used to hold free format text. Number defines the maximum length, overriding any default length.

B — **Blob.** Binary long object used to hold documents, images, sound, objects, etc.

Data Profile Annotation

{Source**}** — **Data source**. system, table, column or file, field source name. / delimited choices.

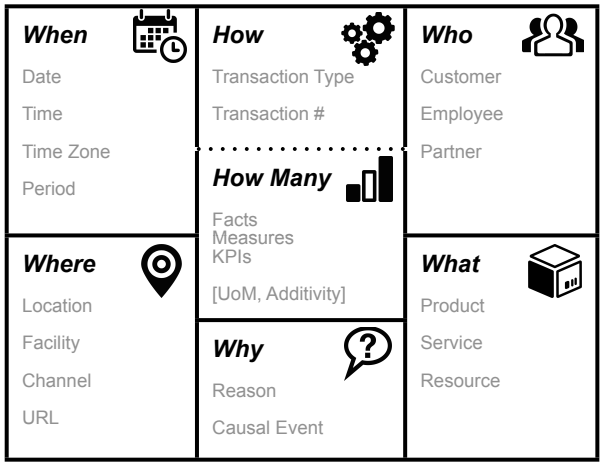
~~Unavailable~~ / ~~**MD**~~ — **Unavailable or incorrect.** Data source for table or column is unavailable or does not comply with the column type code.

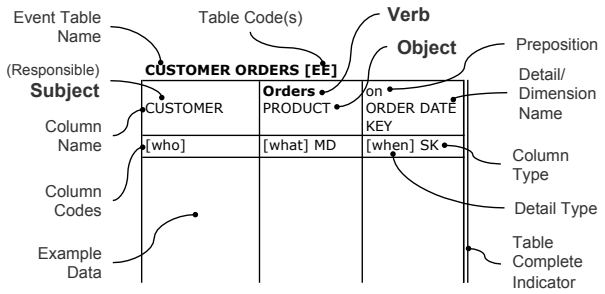
BEAM✲

Business Event Analysis & Modeling

Agile Dimensional Modeling

Who does what?
When and where?
How much / how many?
Why and how?





7W Details: Who, What, When, Where, How Many, Why, How

Example Data Themes

Typical	*normal, popular, average*
Different	*explore group and range, exceptional values*
Missing	*mandatory details*
Repeat	*discover uniqueness*
Group	*organisation, bundle, multi-level, multiple values*
Range	*low/high, old,new, near/far, large/small, min/max*

© DecisionOne Consulting +44 (0)7971 964794
www.decisionone.co.uk information@decisionone.co.uk

Event and Fact Table Types

[DE] **Discrete Event.** Point in time or short duration (completed) transaction.

[EE] **Evolving Event.** (multi-verb) process that takes time to complete.

[RE] **Recurring Event.** Measurements taken at predictable regular intervals.

[TF] **Transaction Fact table.** Physical equivalent of **DE**. Typically maintained by insert only.

[AS] **Accumulating Snapshot.** Physical equivalent of **EE**. Maintained by insert and update. Typically contains multiple milestone date/time dimensions and duration facts.

[PS] **Periodic Snapshot.** Physical equivalent of **RE**. Typically contains semi-additive facts.

[AG] **Aggregate.** Fact table that pre-summarizes an existing detailed fact table.

[DF] **Derived Fact table.** Fact table constructed by merging, slicing, or pivoting existing fact tables.

Dimension Types

[CV] **Current Value.** Contains current value only dimensional attributes. *Type 1 SCD*.

[HV] **Historic Value.** Contains at least one historical value dimensional attribute. *Type 2 slowly changing dimension (SCD).*

[RP] **Role-Playing.** Used to play multiple roles.

[RU] **Roll-Up.** Derived from a more granular dimension.

[SD] **Swappable Dimension.** Part of a set of dimensions with a common surrogate key that can be used in place of each other.

[ML] **Multi-Level.** Dimension containing additional members representing higher levels in the dimension's hierarchy.

[HM] **Hierarchy Map.** Table used to resolve a recursive relationship. Represents a variable-depth hierarchy.

[MV] **Multi-Valued.** Bridge table used to resolve a many-to-many relationship between a fact table and a multi-valued dimension.

[PD] **Pivoted Dimension.** Contains column flags built from the row values of another dimension.

General Column Types

MD **Mandatory.** Value is present under normal conditions. Can be nullable to handle errors.

NN **Not Null.** Column does not allow nulls. All **SK** and **FK** columns are **NN** by default.

ND **NDn** **No Duplicates.** Numbered to define combinations of column values that must be unique. **PK** columns are **ND** by default.

Xn **Exclusive.** Column is not valid in combination with other **X** columns. Numbered to identify mutually exclusive groups and identify the specific **DC** which controls validity.

DC **DCn,n** **Defining Characteristic.** Column value dictates which **X** columns are valid. E.g., `Product Type DC` defines which exclusive product dimension attributes are valid. Number list relates multiple defining characteristics in the same table to specific **X**n exclusive columns or groups.

[W$_{type}$] [$_{dimension}$] **Dimension type or name**. The *W type (who, what, when, where, why, how)* of an event detail or the dimension name when a detail is a role; e.g., `Salesperson [Employee]` where `Salesperson` is a role of the `Employee` dimension. Also used to describe recursive relationships.

Event and Fact Table Column Types

DD **Degenerate Dimension**. Dimensional attribute stored in a fact table. Typically used for transaction IDs (*how* details).

GD **GDn** **Granularity Dimension.** Dimension combination that defines the granularity of a fact table. Numbered when alternative combinations exist.

MV **Multi-Valued.** Event detail contains multiple values that must be resolved using a bridge table. Fact table **FK** that references a multi-value bridge table.

ML **Multi-Level.** Event detail can represent various levels in a hierarchy; e.g., individual employee or teams/branches. Fact table **FK** that points to a multi-level dimension *and* makes use of the additional levels.

Fact Types

FA **Fully Additive.** Fact that produces a correct total when summed across any combination of its dimensions. For a fact to be (fully) additive it must be expressed in a single unit of measure. Percentages and unit prices are not additive.

SA
SAn **Semi-Additive**. Fact that can be correctly totaled by some dimensions but not by at least one non-additive (**NA**) dimension: e.g., an account balance cannot be summed over time: its **NA** dimension. **SA** facts are often averaged over their **NA** dimension.

SA is always used in conjunction with at least one **NA** dimension to relate the semi-additive fact to its non-additive dimension(s).

Numbering relates multiple **SA**n facts in the same table to their specific **NA**n dimension(s).

NA
NAn **Non-Additive.** Fact that cannot be aggregated using sum; e.g., `Temperature NA`. Non-additive facts can be aggregated using functions such as min, max, average.

Non-additive dimension of a semi-additive fact. Numbering relates multiple non-additive dimensions in the same table to specific semi-additive (**SA**n) facts.

DF
DF= formulae **Derived Fact**. Value can be derived from other columns within the same table. May be followed by a simple formula referencing other facts or date/time details by number; e.g., `Unit Price DF=Revenue/Quantity`.

[UoM**]**
[U1, U2...**]** **Unit of Measure**. Unit of measure symbol or description; e.g., `Order Revenue [$]` or `Delivery Delay [days]`.

List denotes that multiple units can be recorded for a quantity. They must be converted into a standard unit (**U1**) to produce an additive fact. Can also be use to document the list of conversion factors required at reporting time.

Dimensional Attribute Types

CV
CVn **Current Value.** Attribute records current values only. Changes overwrite previous values. Supports *"as is"* reporting. Also known as a *type 1 slowly changing dimension (SCD)*.

Combined with **HV** to define hybrid **CV/HV** attributes with default **CV** behavior listed first. Implemented as separate **CV** & **HV** attributes.

Combined with **PV** to define hybrid **CV/PV** attributes or numbered to relate separate **CV**n attributes to matching **PV**n attributes.

HV
HVn **Historic Value.** Attribute records historical values. Changes cause new versions of dimension members to be created: preserving their historically correct values. Supports *"as was"* reporting. Also known as a *type 2 SCD*.

Combined with **CV** to define hybrid **HV/CV** attributes with default **HV** behavior listed first. Implemented as separate **HV** & **CV** attributes.

Numbering defines *conditional* **HV**n attributes groups: combinations of attributes that only act as **HV** when every member of their *n* group changes at the same time. Used in combination with **CV** to treat small changes or corrections as **CV**; e.g., `Street CV, HV1` and `Zip Code CV, HV1` will be treated as **CV** individually but as **HV** if both change at once.

FV **Fixed Value.** Attribute values do not change over time; e.g., `Date of Birth FV`. Corrections overwrite previous incorrect values: behaves like a **CV** attribute. Also know as a *type 0 SCD*.

PV
PVn **Previous Value.** Attribute records previous values. Supports *"as previously"* or *"as at"* reporting. Also known as a *type 3 SCD*.

Combined with **CV** to define hybrid **CV/PV** attributes or numbered to relate separate **PV**n attributes to their matching **CV**n attributes; e.g., `Previous Territory PV1` and `Territory CV1`.

PV attributes can also hold initial or "as at date" values; e.g., `Initial Territory PV1` or `YE2010 Territory PV1`.

Key Types

PK **Primary Key**. A column or group of columns that uniquely identifies each row in a table.

FK **Foreign Key**. A column that references the primary key of another table.

SK **Surrogate Key**. Anonymous integer assigned by the data warehouse as the primary key for a dimension table. Dimensional foreign key in fact tables. Denotes that example data will be replaced by integer keys.

BK **Business Key**. Source system key.

NK **Natural Key.** Key used in the real world.

RK **Recursive Key.** Foreign key that references the primary key of its own table. Often used to represent variable-depth hierarchies. Used to build [HM] hierarchy maps.

Data Types

Cn **Character**. Number defines the maximum length, overriding any default length.

DTn **Date/Time**. Number is used in duration formulas for derived facts; e.g., `Delivery Delay DF=DT2-DT1`. Number can denote default order of milestones within an [EE].

Dn **Date**. Number is used in duration formulas for derived facts. Number can denote default order of milestones within an [EE].

Nn.n **Numeric**. Number defines precision, overriding the default precision.

Tn **Text**. Long character data used to hold free format text. Number defines the maximum length, overriding any default length.

B **Blob.** Binary long object used to hold documents, images, sound, objects, etc.

Data Profile Annotation

{Source**}** **Data source**. system, table, column or file, field source name. / delimited choices.

~~Unavailable~~
~~**MD**~~ **Unavailable or incorrect.** Data source for table or column is unavailable or does not comply with the column type code.