**A black and grey logo

Description automatically generated**

**CÔNG TY CỔ PHẦN KATALYST**

147 NGUYỄN CỬU VÂN, P.17, BÌNH THẠNH | WWW.KATALYST.VN

**KATALYST DBT PROJECT TESTING DOCUMENT**

dÀNH CHO ….

[1. Dbt generic test 1](#_Toc172645263)

[1.1 Not null 1](#_Toc172645264)

[1.2 Unique 1](#_Toc172645265)

[1.3 Relationship 2](#_Toc172645266)

[2. Customer generic test 3](#_Toc172645267)

[2.1 Ghost\_record 3](#_Toc172645268)

[2.1.1 Source code 3](#_Toc172645269)

[2.1.2 Detail và áp dụng 3](#_Toc172645270)

[2.2 Hub\_w\_source 3](#_Toc172645271)

[2.2.1 Source code 3](#_Toc172645272)

[2.2.2 Detail và áp dụng 4](#_Toc172645273)

[2.3 Sat\_der\_w\_source 4](#_Toc172645274)

[2.3.1 Source code 4](#_Toc172645275)

[2.3.2 Detail và áp dụng 4](#_Toc172645276)

[2.4 Sat\_w\_source 5](#_Toc172645277)

[2.4.1 Source code 5](#_Toc172645278)

[2.4.2 Detail và áp dụng 5](#_Toc172645279)

[2.5 Sat\_snp\_w\_source 5](#_Toc172645280)

[2.5.1 Source code 5](#_Toc172645281)

[2.5.2 Detail và áp dụng 5](#_Toc172645282)

[2.6 Lsat\_der\_w\_source 6](#_Toc172645283)

[2.6.1 Source code 6](#_Toc172645284)

[2.6.2 Detail và áp dụng 6](#_Toc172645285)

[2.7 Lsat\_w\_source 7](#_Toc172645286)

[2.7.1 Source code 7](#_Toc172645287)

[2.7.2 Detail và áp dụng 7](#_Toc172645288)

[2.8 Lsat\_snp\_w\_source 7](#_Toc172645289)

[2.8.1 Source code 7](#_Toc172645290)

[2.8.2 Detail và áp dụng 7](#_Toc172645291)

***DANH SÁCH SAMPLE CODE***

[Exp\_Code 1: test not null 1](#_Toc172645292)

[Exp\_Code 2: test unique 2](#_Toc172645293)

[Exp\_Code 3: test relationship 2](#_Toc172645294)

[Exp\_Code 4: test ghost record 3](#_Toc172645295)

[Exp\_Code 5: test hub\_w\_source 4](#_Toc172645296)

[Exp\_Code 6: Test sat\_der\_w\_source 4](#_Toc172645297)

[Exp\_Code 7: test sat\_w\_source 5](#_Toc172645298)

[Exp\_Code 8: Test sat\_snp\_w\_source 6](#_Toc172645299)

[Exp\_Code 9: test lsat\_der\_w\_source 6](#_Toc172645300)

[Exp\_Code 10: Test last\_w\_source 7](#_Toc172645301)

[Exp\_Code 11: Test lsat\_snp\_w\_source 8](#_Toc172645302)

Katalyst dbt project hiện tại sử dụng dbt generic test và dbt customer generic test phục vụ cho việc testing và kiểm soát chất lượng data.

1. Dbt generic test

Generic test là các macros test có sẵn của dbt. Các project datavault hiện tại sử dụng các loại generic test: not null, unique, và relationship.

* 1. Not null

Cần check not null với các columns thuộc các tables sau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **not null** | biz\_key | hkey\_hub | hkey\_lnk | hkey\_sat/lsat |
| hub | *not null* | *not null* |  |  |
| lnk |  |  | *not null* |  |
| sat(der/main/snp) |  |  |  | *not null* |
| lsat(der/main/snp) |  |  |  | *not null* |

Test not null được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test not null table hub\_account:

* + - Biz\_key: ln\_ac\_nbr
    - Hkey\_hub: dv\_hkey\_hub\_account

Ta gán test:

Exp\_Code 1: test not null

|  |  |
| --- | --- |
| 1  2  3  4  5 | - name: hub\_account  columns:  - name: dv\_hkey\_hub\_account  tests:  - not\_null |

* 1. Unique

Cần check unique với các columns thuộc các tables sau:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **unique** | biz\_key | hkey\_hub | hkey\_lnk | hkey\_sat/lsat |
| hub | *unique* | *unique* |  |  |
| lnk |  |  | *unique* |  |
| sat(der/main/snp) |  |  |  | *unique* |
| lsat(der/main/snp) |  |  |  | *unique* |

Test unique được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test unique table hub\_account:

* + - Biz\_key: ln\_ac\_nbr
    - Hkey\_hub: dv\_hkey\_hub\_account

Ta gáng test:

Exp\_Code 2: test unique

|  |  |
| --- | --- |
| 1  2  3  4  5 | - name: hub\_account  columns:  - name: dv\_hkey\_hub\_account  tests:  - unique |

* 1. Relationship

Cần check relationship các column của các bảng như sau:

|  |  |  |
| --- | --- | --- |
| **con** | **cha** | **column** |
| lnk | *hub1* | *dv\_hkey\_hub\_1* |
| *hub2* | *dv\_hkey\_hub\_2* |
| sat | *hub* | *dv\_hkey\_hub* |
| lsat | *lnk1* | *dv\_hkey\_lnk\_1* |
| *lnk2* | *dv\_hkey\_lnk\_2* |

* Con: bảng con, là table được gáng test.
* Cha: table cha
* Test relationship sẽ check tất cả các giá trị của column được gáng test tại bảng con có thuộc bảng cha không, và **không check** quan hệ ngược lại.
* Test relationship được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.
* Vd:

Test relationship table lnk\_customer\_account và hub\_account, hub\_customer với

lnk\_customer có:

* + - Dv\_hkey\_lnk\_customer\_account
    - Dv\_hhey\_hub\_customer (hkey\_hub\_1)
    - Dv\_hkey\_hub\_account (hkey\_hub\_2)

Hub\_customer (hub1) có:

* + - Dv\_hkey\_hub\_customer

Hub\_account (hub2) có:

* + - Dv\_hkey\_hub\_account

Ta gáng test:

Exp\_Code 3: test relationship

|  |  |
| --- | --- |
| 1  2  3  4  5  6  7  8  9  10  11  12 | - name: lnk\_customer\_account  columns:  - name: dv\_hkey\_hub\_customer  tests:  - relationships:  to: ref('hub\_customer')  field: dv\_hkey\_hub\_customer  - name: dv\_hkey\_hub\_account  tests:  - relationships:  to: ref('hub\_account')  field: dv\_hkey\_hub\_account |

Với:

* + - To: ref(parents table name)
    - Field: **column name tại parents table** cần check relationship

1. Customer generic test

Customer generic test trong dbt là các tests mà customer tự tạo ra để sử dụng lặp đi lặp lại tương tự như các generic test có sẵn của dbt. Customer generic test (CGT) có dạng sql model, nằm trong directory /test/generic của dbt project.

Các CGT hiện có:

* Ghost\_record
* Hub\_w\_source
* Lnk\_w\_source
* Sat\_der\_w\_source
* Sat\_w\_source
* Sat\_snp\_w\_source
* Lsat\_der\_w\_source
* Lsat\_w\_source
* Lsat\_snp\_w\_source
  1. Ghost\_record

Ghost\_record test check sự tồn tại của ghost\_record trong table được gáng test. Nếu test fail, table này không tồn tại ghost record hoặc có hơn 1 dòng có giá trị là ghost record.

* + 1. Source code

Xem detail code của ghost record test tại /test/generic/ghost\_record.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test ghost\_record được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test ghost\_record table hub\_account:

* + - Hkey\_hub: dv\_hkey\_hub\_account

Ta gán test:

Exp\_Code 4: test ghost record

|  |  |
| --- | --- |
| 1  2  3  4  5 | - name: hub\_account  columns:  - name: dv\_hkey\_hub\_account  tests:  - ghost\_record |

* 1. Hub\_w\_source

Hub\_w\_source test dùng để test table hub có đúng và đủ với table source tương ứng không. Nếu hub\_w\_source test fail, hub có nhiều hơn hoặc ít hơn giá trị cần có.

* + 1. Source code

Xem detail code của hub\_w\_source test tại /test/generic/hub\_w\_source.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test hub\_w\_source được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test table hub\_account:

* + - Hkey\_hub: dv\_hkey\_hub\_account

Ta gán test:

Exp\_Code 5: test hub\_w\_source

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | - name: hub\_account  columns:  - name: dv\_hkey\_hub\_account  tests:  - hub\_w\_source:  model\_name: hub\_account |

Với:

* + - model\_name: tên model (hub table) cần test
    1. Lnk\_w\_source

Link\_w\_source test dùng để test table lnk có đúng và đủ với table source tương ứng không. Nếu lnk\_w\_source test fail, lnk có nhiều hơn hoặc ít hơn giá trị cần có.

* + 1. Source code

Xem detail code của lnk\_w\_source test tại /test/generic/lnk\_w\_source.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test lnk\_w\_source được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test table lnk\_customer\_account:

* + - Hkey\_lnk: dv\_hkey\_lnk\_customer\_account

Ta gán test:

Exp\_Code 5: test hub\_w\_source

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | - name: lnk\_customer\_account  columns:  - name: dv\_hkey\_lnk\_customer\_account  tests:  - lnk\_w\_source:  model\_name: lnk\_customer\_account |

Với:

* + - model\_name: tên model (lnk table) cần test
  1. Sat\_der\_w\_source

Sat\_der\_w\_source test dùng để test table sat der có đúng và đủ với table source tương ứng không. Nếu sat\_der\_w\_source test fail, sat der đang thiếu so với source.

* + 1. Source code

Xem detail code của sat\_derr\_w\_source test tại /test/generic/sat\_der\_w\_source.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test sat\_der\_w\_source được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test table sat\_der\_account:

* + - Hkey\_hub: dv\_hkey\_sat\_account

Ta gán test:

Exp\_Code 6: Test sat\_der\_w\_source

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | - name: sat\_der\_account  columns:  - name: dv\_hkey\_sat\_account  tests:  - sat\_der\_w\_source:  model\_name: sat\_account |

Với:

* + - model\_name: tên model (sat\_der table) cần test
    - Lưu ý là (L)sat\_derivativity (sat/lsat-der) và (L)sat\_snapshot sử dụng cùng config với tên (L)sat (không có “der” hoặc “snp” prefix), nên ở đây các customer generic test cần đến model\_name, **model\_name chỉ sử dụng tên sat/lsat mà không phải là sat/lsat\_der/snp.**
  1. Sat\_w\_source

Sat\_w\_source test dùng để test table sat main có đúng và đủ với table source tương ứng không. Nếu sat\_w\_source test fail, sat main đang thiếu so với source.

* + 1. Source code

Xem detail code của sat\_w\_source test tại /test/generic/sat\_w\_source.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test sat\_w\_source được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test table sat\_account:

* + - Hkey\_hub: dv\_hkey\_sat\_account

Ta gán test:

Exp\_Code 7: test sat\_w\_source

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | - name: sat\_account  columns:  - name: dv\_hkey\_sat\_account  tests:  - sat\_w\_source:  model\_name: sat\_account |

Với:

* + - model\_name: tên model (sat table) cần test.
  1. Sat\_snp\_w\_source

Sat\_snp\_w\_source test dùng để test table sat snp có đúng và đủ với table source tương ứng không. Nếu sat\_snp\_w\_source test fail, sat snp đang thiếu so với source.

* + 1. Source code

Xem detail code của sat\_snp\_w\_source test tại /test/generic/sat\_snp\_w\_source.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test sat\_snp \_w\_source được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test table sat\_snp\_account:

* + - Hkey\_hub: dv\_hkey\_sat\_account

Ta gán test:

Exp\_Code 8: Test sat\_snp\_w\_source

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | - name: sat\_snp\_account  columns:  - name: dv\_hkey\_sat\_account  tests:  - sat\_snp\_w\_source:  model\_name: sat\_account |

Với:

* + - model\_name: tên model (sat\_snp table) cần test
    - Lưu ý là (L)sat\_derivativity (sat/lsat-der) và (L)sat\_snapshot sử dụng cùng config với tên (L)sat (không có “der” hoặc “snp” prefix), nên ở đây các customer generic test cần đến model\_name, **model\_name chỉ sử dụng tên sat/lsat mà không phải là sat/lsat\_der/snp.**
  1. Lsat\_der\_w\_source

Lsat\_der\_w\_source test dùng để test table lsat der có đúng và đủ với table source tương ứng không. Nếu lsat\_der\_w\_source test fail, lsat der đang thiếu so với source.

* + 1. Source code

Xem detail code của lsat\_der\_w\_source test tại /test/generic/lsat\_der\_w\_source.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test lsat\_der\_w\_source được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test table lsat\_der\_customer\_account:

* + - Hkey\_sat: dv\_hkey\_lsat\_customer\_account

Ta gán test:

Exp\_Code 9: test lsat\_der\_w\_source

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | - name: lsat\_der\_customer\_account  columns:  - name: dv\_hkey\_lsat\_customer\_account  tests:  - lsat\_der\_w\_source:  model\_name: lsat\_customer\_account |

Với:

* + - model\_name: tên model (lsat\_der table) cần test
    - Lưu ý là (L)sat\_derivativity (sat/lsat-der) và (L)sat\_snapshot sử dụng cùng config với tên (L)sat (không có “der” hoặc “snp” prefix), nên ở đây các customer generic test cần đến model\_name, **model\_name chỉ sử dụng tên sat/lsat mà không phải là sat/lsat\_der/snp.**
  1. Lsat\_w\_source

Lsat\_w\_source test dùng để test table lsat main có đúng và đủ với table source tương ứng không. Nếu lsat\_w\_source test fail, lsat main đang thiếu so với source.

* + 1. Source code

Xem detail code của lsat\_w\_source test tại /test/generic/lsat\_w\_source.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test lsat\_w\_source được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test table lsat\_customer\_account:

* + - Hkey\_sat: dv\_hkey\_lsat\_customer\_account

Ta gán test:

Exp\_Code 10: Test last\_w\_source

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | - name: lsat \_customer\_account  columns:  - name: dv\_hkey\_lsat\_customer\_account  tests:  - lsat\_w\_source:  model\_name: lsat\_customer\_account |

Với:

* + - model\_name: tên model (lsat table) cần test
  1. Lsat\_snp\_w\_source

Lsat\_snp\_w\_source test dùng để test table lsat snp có đúng và đủ với table source tương ứng không. Nếu lsat\_snp\_w\_source test fail, lsat snp đang thiếu so với source.

* + 1. Source code

Xem detail code của lsat\_snp\_source test tại /test/generic/lsat\_snp\_w\_source.sql của dbt project tương ứng.

* + 1. Detail và áp dụng

Test lsat\_snp\_w\_source được khai báo vào file schema.yaml, dưới tên table và column tương ứng cần test.

* Vd:

Test table lsat\_snp\_customer\_account:

* + - Hkey\_sat: dv\_hkey\_lsat\_customer\_account

Ta gán test:

Exp\_Code 11: Test lsat\_snp\_w\_source

|  |  |
| --- | --- |
| 1  2  3  4  5  6 | - name: lsat\_snp\_customer\_account  columns:  - name: dv\_hkey\_lsat\_customer\_account  tests:  - lsat\_snp\_w\_source:  model\_name: lsat\_customer\_account |

Với:

* + - model\_name: tên model (lsat\_snp table) cần test
    - Lưu ý là (L)sat\_derivativity (sat/lsat-der) và (L)sat\_snapshot sử dụng cùng config với tên (L)sat (không có “der” hoặc “snp” prefix), nên ở đây các customer generic test cần đến model\_name, **model\_name chỉ sử dụng tên sat/lsat mà không phải là sat/lsat\_der/snp.**