





InstaCart Fact Dim - Snowflake

app.snowflake.com/tsdbsqw/zh79962/w2lFe9Eg59eT/query

WorksheetsInstaCart Fact Dimsnowpipestorage_integrations2023-09-21 2:10am2023-09-21 2:15am2023-09-21 2:38am2023-09-21 2:38am+

DatabasesWorksheets

Pinned (0)
No pinned objects

All Objects
SCD_DEMO
SNOWFLAKE
SNOWFLAKE_SAMPLE_DATA
YASH_DW1

YASH_DW1.INSTACART_PROJECT

Settings

Latest Version

```
1  -- This is the process which is necessary to have access to our aws account(i.e connecting snowflake to my aws account)
2  CREATE STAGE my_stage1
3  URL = 's3://snowflake-yash-dimmodelling-project/instacart-dm-project/'
4  CREDENTIALS = (AWS_KEY_ID = 'AKIAYQRCIOB442IHUUUS' AWS_SECRET_KEY = '8Zmhc1BZF10AeuUwxtahDnr6Nu6YvmQk90qBdauI');
5
6  CREATE OR REPLACE FILE FORMAT csv_file_format
7  TYPE = 'CSV'
8  FIELD_DELIMITER = ','
9  SKIP_HEADER = 1
10 FIELD_OPTIONALLY_ENCLOSED_BY='"; -- If the CSV file has a header row, skip it
11
12
13 CREATE TABLE aisles (
14     aisle_id INTEGER PRIMARY KEY,
15     aisle VARCHAR
16 );
17
18 COPY INTO aisles (aisle_id, aisle) -- Copy Cmd to load our data from the s3 bucket
19 FROM @my_stage1/aisles.csv
20 FILE_FORMAT = (FORMAT_NAME = 'csv_file_format');
21 --SELECT * FROM aisles LIMIT 10
22
23 CREATE TABLE departments (
24     department_id INTEGER PRIMARY KEY,
25     department VARCHAR
26 );
27
28 COPY INTO departments (department_id, department)
29 FROM @my_stage1/departments.csv
30 FILE_FORMAT = (FORMAT_NAME = 'csv_file_format');
31
32 CREATE OR REPLACE TABLE products (
33     product_id INTEGER PRIMARY KEY,
34     product_name VARCHAR,
```

InstaCart Fact Dim - Snowflake

+

← → ↺

app.snowflake.com/tsdbsqw/zh79962/w2lFe9Eg59eT/query

🔖 ☆ 📁 ⬇️ 🗑️

Y

⋮

Worksheets

InstaCart Fact Dim

snowpipe

storage_integration

2023-09-21 2:10am

2023-09-21 2:15am

2023-09-21 2:38am

2023-09-21 2:40am

+

Databases

Worksheets

🔍

ACCOUNTADMIN • COMPUTE_WH

Share

▶️

⌵

Pinned (0)

No pinned objects

🔍 All Objects

⋮

> 📁 SCD_DEMO

> 📁 SNOWFLAKE

> 📁 SNOWFLAKE_SAMPLE_DATA

> 📁 YASH_DW1

YASH_DW1.INSTACART_PROJECT

Settings

Latest Version

🔍

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

COPY INTO products (product_id, product_name, aisle_id, department_id)

FROM @my_stage1/products.csv

FILE_FORMAT = (FORMAT_NAME = 'csv_file_format');

CREATE OR REPLACE TABLE orders (

order_id INTEGER PRIMARY KEY,

user_id INTEGER,

eval_set STRING,

order_number INTEGER,

order_dow INTEGER,

order_hour_of_day INTEGER,

days_since_prior_order INTEGER

);

COPY INTO orders (order_id, user_id, eval_set, order_number, order_dow, order_hour_of_day, days_since_prior_order)

FROM @my_stage1/orders.csv

FILE_FORMAT = (FORMAT_NAME = 'csv_file_format');

CREATE OR REPLACE TABLE order_products (

order_id INTEGER,

product_id INTEGER,

add_to_cart_order INTEGER,

reordered INTEGER,

PRIMARY KEY (order_id, product_id)

);

COPY INTO order_products (order_id, product_id, add_to_cart_order, reordered)

FROM @my_stage1/order_products.csv

FILE_FORMAT = (FORMAT_NAME = 'csv_file_format');

📄 📄 📄

InstaCart Fact Dim - Snowflake

app.snowflake.com/tsdbsqw/zh79962/w2lFe9Eg59eT/query

WorksheetsInstaCart Fact Dimsnowpipestorage_integration2023-09-21 2:10am2023-09-21 2:15am2023-09-21 2:38am2023-09-21 2:38am+

DatabasesWorksheets

Pinned (0)

No pinned objects

All Objects

SCD_DEMO

SNOWFLAKE

SNOWFLAKE_SAMPLE_DATA

YASH_DW1

YASH_DW1.INSTACART_PROJECT

Settings

Latest Version

72

73

--Now the table that we have in raw format, we need to convert that into FACT&DIM tables(The middle process of converting fact to dim table is called as transformation)

74

75

-- Building Dim Fact Tables

76

77

CREATE OR REPLACE TABLE dim_users AS (

78

SELECT

79

user_id

80

FROM

81

orders

82

);

83

84

CREATE OR REPLACE TABLE dim_products AS (

85

SELECT

86

product_id,

87

product_name

88

FROM

89

products

90

);

91

92

93

CREATE OR REPLACE TABLE dim_aisles AS (

94

SELECT

95

aisle_id,

96

aisle

97

FROM

98

aisles

99

);

100

101

CREATE OR REPLACE TABLE dim_departments AS (

102

SELECT

103

department_id,

104

department

105

FROM

InstaCart Fact Dim - Snowflake

app.snowflake.com/tsdbsqw/zh79962/w2lFe9Eg59eT/query

Worksheets

InstaCart Fact Dim

snowpipe

storage_integration

2023-09-21 2:10am

2023-09-21 2:15am

2023-09-21 2:38am

2023-09-21 2:38am

+

Databases

Worksheets

ACCOUNTADMIN

COMPUTE_WH

Share

▶

▼

Pinned (0)

No pinned objects

Q All Objects

SCD_DEMO

SNOWFLAKE

SNOWFLAKE_SAMPLE_DATA

YASH_DW1

YASH_DW1.INSTACART_PROJECT

Settings

Latest Version

Q

101 CREATE OR REPLACE TABLE dim_departments AS (
102 SELECT
103 department_id,
104 department
105 FROM
106 departments
107);
108
109 CREATE OR REPLACE TABLE dim_orders AS (
110 SELECT
111 order_id,
112 order_number,
113 order_dow,
114 order_hour_of_day,
115 days_since_prior_order
116 FROM
117 orders
118);
119
120 CREATE TABLE fact_order_products AS (
121 SELECT
122 op.order_id,
123 op.product_id,
124 o.user_id,
125 p.department_id,
126 p.aisle_id,
127 op.add_to_cart_order,
128 op.reordered
129 FROM
130 order_products op
131 JOIN
132 orders o ON op.order_id = o.order_id
133 JOIN
134 products p ON op.product_id = p.product_id

InstaCart Fact Dim - Snowflake

app.snowflake.com/tsdbsqw/zh79962/w2lFe9Eg59eT/query

WorksheetsInstaCart Fact Dimsnowpipestorage_integration2023-09-21 2:10am2023-09-21 2:15am2023-09-21 2:38am2023-09-21 2:38am+

DatabasesWorksheets

Pinned (0)
No pinned objects

All Objects
SCD_DEMO
SNOWFLAKE
SNOWFLAKE_SAMPLE_DATA
YASH_DW1

YASH_DW1.INSTACART_PROJECTSettings

131 JOIN
132 orders o ON op.order_id = o.order_id
133 JOIN
134 products p ON op.product_id = p.product_id
135);
136
137
138
139
140 --Analytics
141
142 -- Query to calculate the total number of products ordered per department:
143 SELECT
144 d.department,
145 COUNT(*) AS total_products_ordered
146 FROM
147 fact_order_products fop
148 JOIN
149 dim_departments d ON fop.department_id = d.department_id
150 GROUP BY
151 d.department;
152
153
154 -- Query to find the top 5 aisles with the highest number of reordered products:
155 SELECT
156 a.aisle,
157 COUNT(*) AS total_reordered
158 FROM
159 fact_order_products fop
160 JOIN
161 dim_aisles a ON fop.aisle_id = a.aisle_id
162 WHERE
163 fop.reordered = TRUE
164 GROUP BY

ACCOUNTADMIN

COMPUTE_WH

Share

▶

▼

Latest Version

Q

InstaCart Fact Dim - Snowflake

app.snowflake.com/tsdbsqw/zh79962/w2lFe9Eg59eT/query

WorksheetsInstaCart Fact Dimsnowpipestorage_integration2023-09-21 2:10am2023-09-21 2:15am2023-09-21 2:38am2023-09-21 2:38am+

DatabasesWorksheets

Pinned (0)

No pinned objects

All Objects

SCD_DEMO

SNOWFLAKE

SNOWFLAKE_SAMPLE_DATA

YASH_DW1

YASH_DW1.INSTACART_PROJECT

Settings

Latest Version

162

WHERE

163

fop.reordered = TRUE

164

GROUP BY

165

a.aisle

166

ORDER BY

167

total_reordered DESC

168

LIMIT 5;

169

170

171

-- Query to calculate the average number of products added to the cart per order by day of the week:

172

SELECT

173

o.order_dow,

174

AVG(fop.add_to_cart_order) AS avg_products_per_order

175

FROM

176

fact_order_products fop

177

JOIN

178

dim_orders o ON fop.order_id = o.order_id

179

GROUP BY

180

o.order_dow;

181

182

-- Query to identify the top 10 users with the highest number of unique products ordered:

183

SELECT

184

u.user_id,

185

COUNT(DISTINCT fop.product_id) AS unique_products_ordered

186

FROM

187

fact_order_products fop

188

JOIN

189

dim_users u ON fop.user_id = u.user_id

190

GROUP BY

191

u.user_id

192

ORDER BY

193

unique_products_ordered DESC

194

LIMIT 10;

195