

انبار داده آموزشگاه مجازی

استاد درس:

دکتر علیرضا بصیری

اعضای تیم:

محمد هادی حاجی حسینی، غزاله زهتاب، سجاد

عباسی، آرش ماری اوریاد

پاییز 99

مقدمه : توضیح سورس

- موضوع: آموزشگاه مجازی (آنلاین)
- افراد میتوانند در دوره های مختلف شرکت کنند و به محتوای آن کورس دسترسی پیدا کنند.
- هر کورس میتواند چندین مدرس داشته باشد.
- شرکت کنندگان میتوانند در صورت بروز مشکل به راحتی با پشتیبانی ارتباط بگیرند.
- و ...
- سورس دارای ۲۲ جدول است.

فهرست مطالب

1. توضیح انبار داده
2. بُعدها
3. مارت Course_Education
4. مارت User Behavior
5. مارت Human Resources

1. توضیح انبار داده

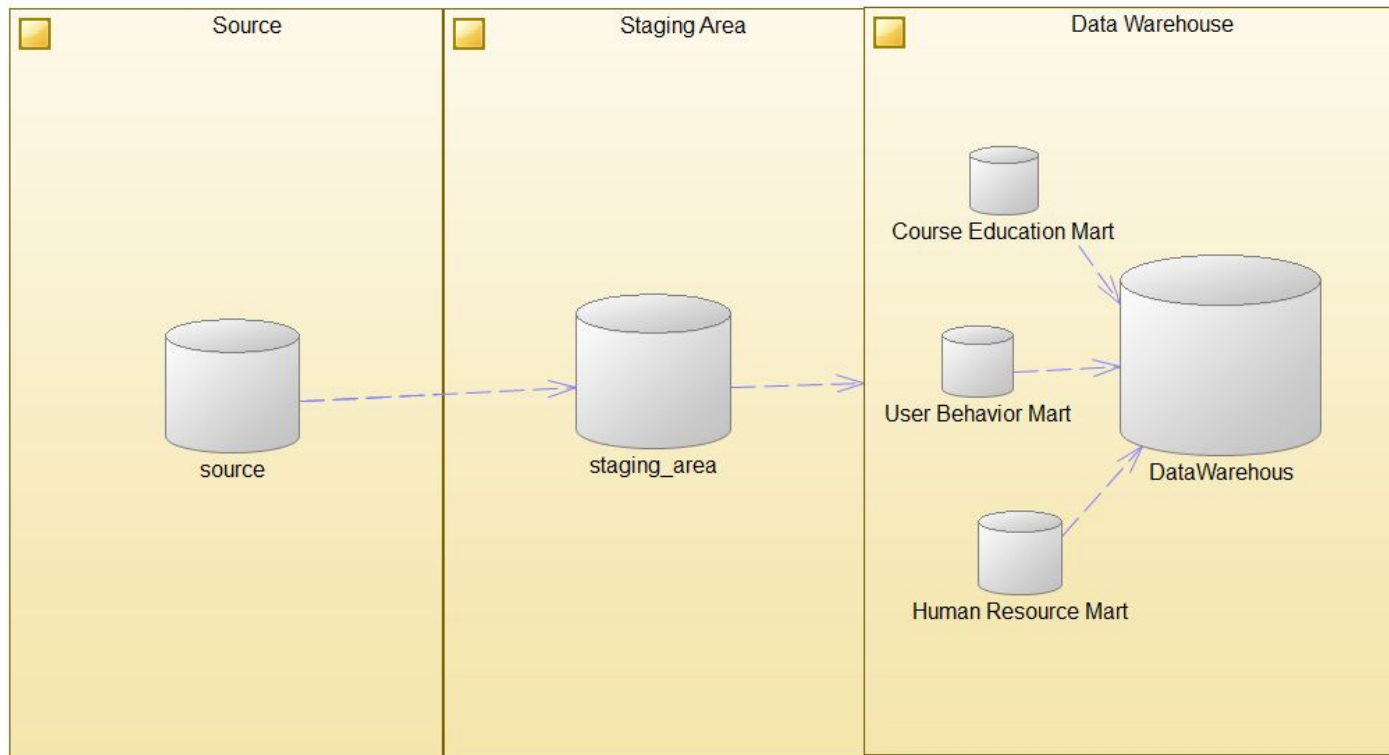
2. بعدها

3. مارت Course Education

4. مارت User Behavior

5. مارت Human Resources

توضیح انبار داده



فهرست مطالب

1. توضیح انبار داده

2. بعدها

a. S_Dim_User

b. S_Dim_Course

c. S_Dim_Date

d. C_Dim_CourseTopic

e. HR_Dim_Instructor

f. HR_Dim_Staff

g. HR_Dim_TicketCategory

3. مارٹ Course

4. مارٹ User

5. مارٹ Human Resources

بعد S_Dim_User

- این بعد اطلاعات کاربران سایت را در انبار داده نگهداری می‌کند.
- یک SCD نوع یک است.
- به دلیل کاربرد در چندین مارت یک بعد اشتراکی (S_) است.

بعد S_Dim_User

[S_Dim_User]

#	<u>user_id</u>	<u>Integer</u>
o	first_name	Variable characters (70)
o	last_name	Variable characters (70)
o	username	Variable characters (150)
o	credit	Number
o	phone_number	Variable characters (25)
o	email_address	Variable characters (355)
o	is_email_verified	Boolean
o	description_verified	Variable characters (70)
o	datetime_signed_up	Date & Time
o	gender	Variable characters (10)
o	date_of_birth	Date

بعد S_Dim_Course

- مسئول نگهداری اطلاعات مربوط به دوره‌ها در انبار داده
- این بُعد SCD نوع دوم می‌باشد. (بر اساس فیلد price)
- به دلیل استفاده‌ی چندین مارت از این بُعد، یک بُعد اشتراکی به حساب می‌آید.
- فیلدهای این بعد منطبق بر فیلدهای جدول OnlineCourse از سورس می‌باشد.

بعد S_Dim_Course

S_Dim_Course			
<u>course_key</u>	<pi>	Integer	<M>
course_id		Integer	
title		Variable characters (400)	
description		Variable characters (1024)	
price		Money	
course_type		Variable characters (255)	
course_level		Variable characters (255)	
pre_req		Variable characters (1024)	
language		Variable characters (1024)	
file_size		Integer	
how_to_download		Variable characters (1024)	
contact_way		Variable characters (1024)	
completion_percentage		Decimal	
length_in_calendar_time		Variable characters (1024)	
start_date		Date & Time	
datetime_content_update		Date & Time	
datetime_created		Date & Time	
price_starting_date		Date & Time	
price_end_date		Date & Time	
current_flag		Boolean	

بعد S_Dim_Date

- این بعد برای نگهداری تاریخ‌ها، اسامی ماه‌ها (به فارسی و انگلیسی)، شماره‌ی سال‌ها (فارسی و انگلیسی) و ... استفاده می‌شود.
- بُعد تاریخ به علت کاربرد در اکثر مارت‌ها، یک بُعد اشتراکی به حساب می‌آید.
- تاریخ‌های این بعد را طوری آپدیت کردیم که تاریخ‌های فعلی نیز در این بعد وجود داشته باشند (تا سال ۲۰۲۰).

بعد C_Dim_CourseTopic

- از این بُعد برای نگهداری ارتباط میان course و topic و subtopic استفاده می‌شود.
- هر subtopic به وسیله‌ی کلید خارجی به یک topic متصل است و هر topic خود به وسیله‌ی کلید خارجی به یک course منتسب می‌شود.
- میان course و topic و subtopic یک رابطه‌ی سلسله‌مراتبی وجود دارد.

بعد C_Dim_CourseTopic

C_Dim_CourseTopic			
<u>subtopic_id</u>	<u><pi></u>	<u>Integer</u>	<u><M></u>
topic_id		Integer	
course_id		Integer	
topic_title		Variable characters (1024)	
topic_priority		Integer	
subtopic_title		Variable characters (1024)	
subtopic_priority		Integer	
course_title		Variable characters (1024)	

بعد HR_Dim_Instructor

- این بعد اطلاعات مدرس را در انبار داده نگهداری می‌کند.
- یک SCD نوع سوم روی Phone_nubmer
- در تمامی گزارش گیری هایی که به منظور بررسی وضعیت مدرس ها طراحی می شود کاربرد دارد.

بعد HR_Dim_Instructor

HR_Dim_Instructor

#	<u>inst_id</u>	<u>Integer</u>
o	first_name	Variable characters (70)
o	last_name	Variable characters (70)
o	email_address	Variable characters (355)
o	is_email_verified	Boolean
o	short_description	Variable characters (70)
o	description	Variable characters (500)
o	datetime_signed_up	Date & Time
o	original_phone_num	Variable characters (25)
o	effective_date	Date & Time
o	current_phone_num	Variable characters (25)

- اطلاعات کارکنان در این دایمنشن ذخیره شده است.
- مثلاً برنامه نویسان یا افرادی که در بخش پشتیبانی هستند. اگر شخص دیگری به عنوان یکی از کارکنان به شرکت اضافه شود اینجا اضافه میشود.

بعد HR_Dim_Staff

HR_Dim_Staff			
<u>staff_id</u>	<u><pi></u>	<u>Integer</u>	<u><M></u>
first_name		Variable characters (70)	
last_name		Variable characters (70)	
date_of_birth		Date	
email_address		Variable characters (355)	
is_email_verified		Boolean	
description_is_email_verified		Variable characters (200)	
phone_number		Variable characters (25)	
datetime_joined		Date & Time	
gender		Variable characters (10)	
short_description		Variable characters (500)	
username		Variable characters (150)	
job_title		Variable characters (400)	
job_description		Variable characters (1000)	

بعد HR_Dim_TicketCategory

- این بعد دارای لیستی از انواع بلیط هایی که ممکن است برای کاربر صادر شود می باشد.
- در گزارش گیری هایی از جنس بررسی پشتیبانی سیستم از این بعد استفاده می شود.

بعد HR_Dim_TicketCategory

HR_Dim_TicketCategory

#	<u>ticket_category_id</u>	<u>Integer</u>
o	ticket_cat_title	Variable characters (250)
o	[description]	Variable characters (500)

فهرست مطالب

1. توضیح انبار داده

2. بعدها

3. **مارت Course_Education**

a C_Fact_CourseBuying

b C_Fact_CourseBuying_Periodic

c C_Fact_CategoryBuying_Acc

d C_Fact_CourseDownloading

e C_Fact_Course_Downloading_Periodic

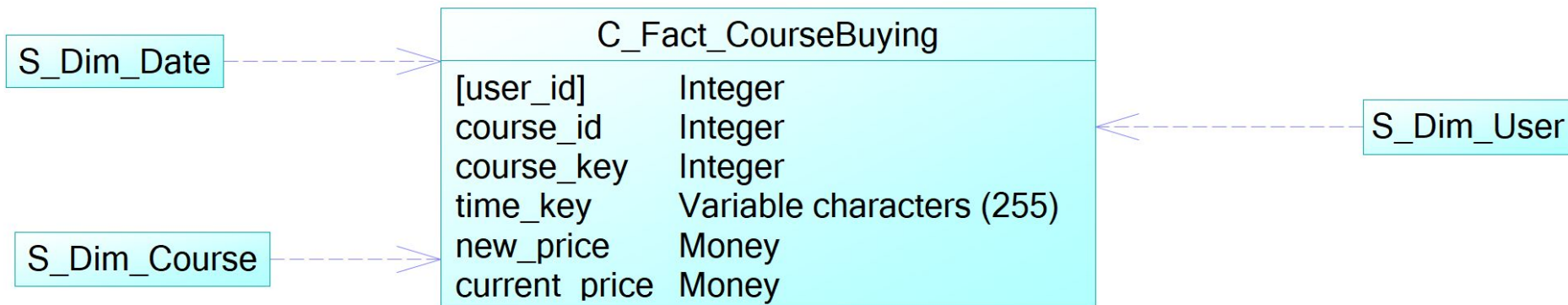
4. **مارت User Behavior**

5. **مارت Human Resources**

فکت تراکنشی C_Fact_CourseBuying

- در این جدول به ازای هر خرید کورس توسط هر کاربر، یک رکورد وجود دارد.
- به ازای هر خرید رکوردی ثبت میشود، پس این فکت به صورت تراکنشی است.
- برای هر خرید، هم قیمت اصلی کورس و هم قیمت آن پس از تخفیف ثبت میشود. یعنی اینکه کورس چه قیمتی داشته ولی کاربر با چه قیمتی خریده (تخفیف)
- گزارشی که چه کسی چه کورسی را در چه تاریخی با چه مبلغ اصلی و چه مبلغ نهایی به خاطر تخفیف خریده است.
- بعد ها: S_Dim_Date , S_Dim_User , S_Dim_Course
- Measures: قیمت خرید کورس بدون اعمال تخفیف، قیمت پرداختی با تخفیف

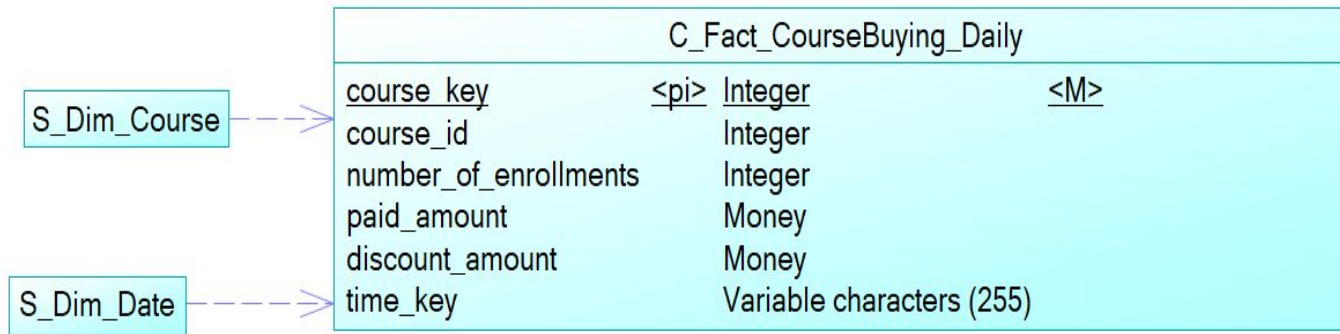
فکت تراکنشی C_Fact_CourseBuying



فکت C_Fact_CourseBuying_Daily

- در این جدول فکت که از نوع Daily Snapshot می‌باشد، برای هر کورس به صورت روزانه موارد زیر محاسبه و ذخیره می‌شوند:
 - تعداد enrollmentها (ثبت‌نام‌ها) برای آن کورس
 - هزینه‌ی پرداخت شده برای آن کورس
 - میزان تخفیفی که کاربران در فرایند ثبت‌نام کورس از آن بهره‌مند شده‌اند
- این فکت از روی فکت تراکنشی C_Fact_CourseBuying به دست می‌آید.

فکت C_Fact_CourseBuying_Daily



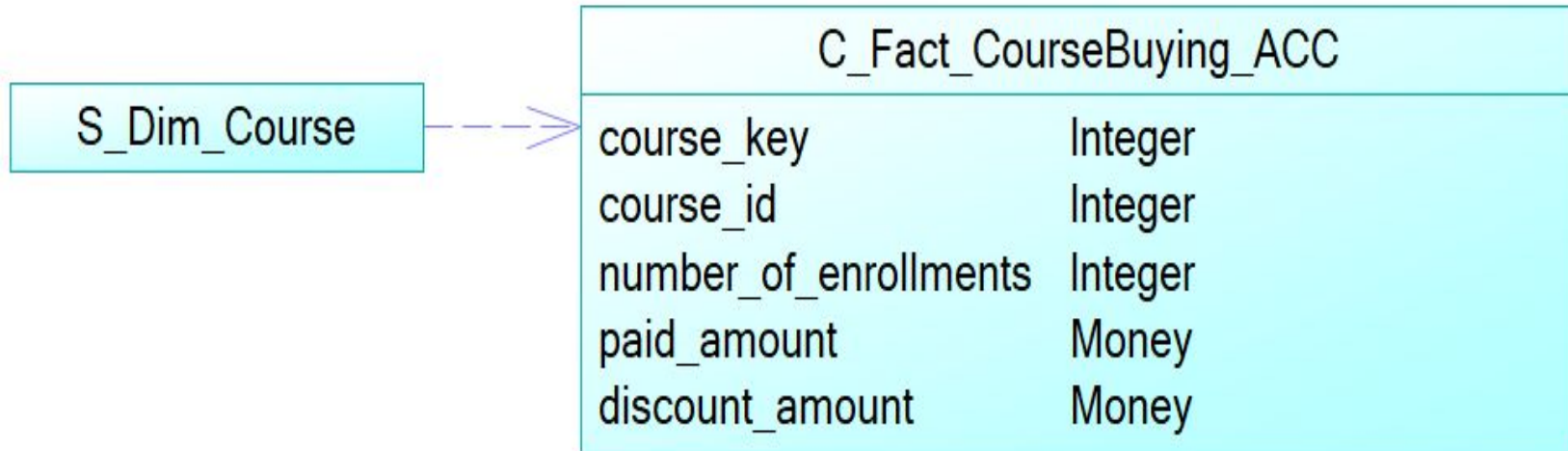
فکت C_Fact_CourseBuying_Acc

- هر رکورد اين جدول گزارشی است در اين باب که برای يك کورس خاص:
 - تا کنون چند نفر در اين کورس ثبت نام کرده اند؟
 - از نظر مالی اين کورس چه ميزان فروش داشته است؟
 - ثبت نام کنندگان در اين کورس تا کنون از چه ميزان تخفيف بهره مند شده اند.

فکت C_Fact_CourseBuying_Acc

- این فکت از نوع تجمیعی می باشد و اطلاعات آن از جدول فکت تراکنشی C_Fact_CourseBuying به دست می آید.
- بُعدها: C_Dim_Course
- Measure ها: تعداد ثبت نامی ها، مبلغ پرداخت شده برای کورس، تخفیفی که مشمول ثبت نام کنندگان شده است

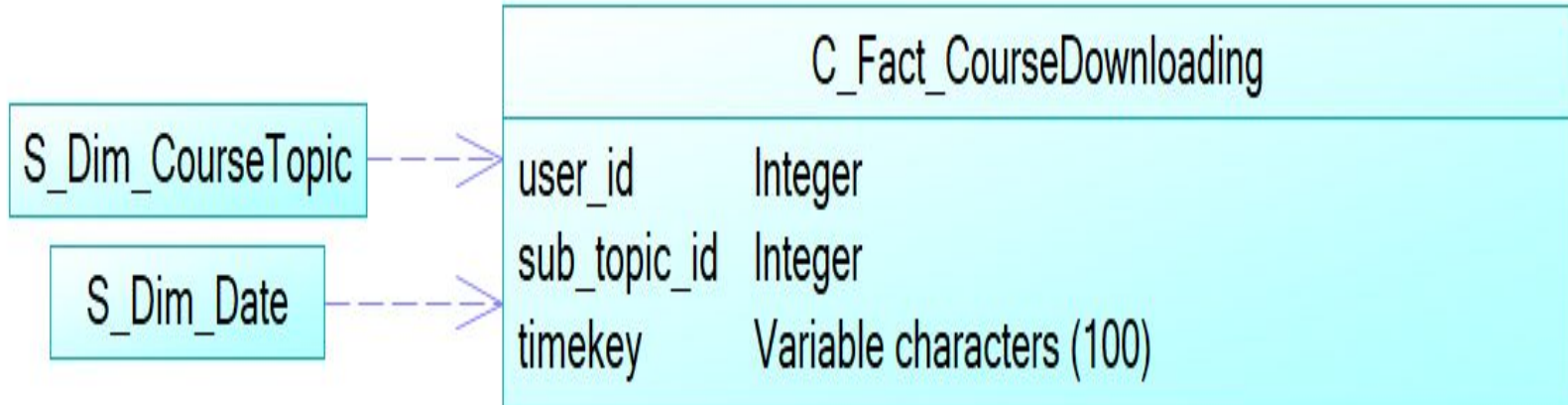
فکت C_Fact_CourseBuying_ACC



فکت تراکشی C_Fact_CourseDownloading

- این فکت که تراکشی به حساب می آید جهت ثبت و بررسی درخواست های دانلود کاربران طراحی شده است.
- در واقع هر رکورد این فکت نشان دهنده ی آن است که کاربری در تاریخی مشخص اقدام به دانلود محتوای یک subtopic مشخص از یک دوره کرده است.

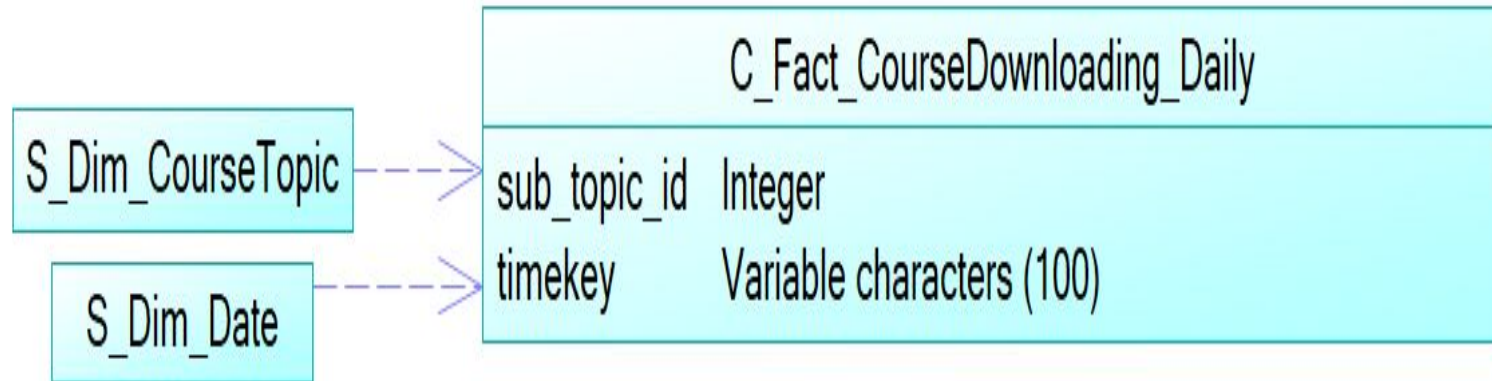
فکت تراکنشی C_Fact_CourseDownloading



فکت C_Fact_CourseDownloading_Daily

- این فکت از نوع Daily Snapshot می باشد.
- از روی فکت تراکنشی C_Fact_CourseDownloading به دست می آید.
- در این فکت می خواهیم گزارش کنیم که برای هر subtopic (و به تبع آن برای هر course) در هر روز، چه تعداد تلاش برای دانلود توسط کاربران صورت گرفته است.
- این گزارش از جهت کنترل و بررسی منابع سرور حائز اهمیت می باشد.

C_Fact_CourseDownloading_Daily



فهرست مطالب

1. توضیح انبار داده

2. بعدها

3. مارت course

4. مارت User Behavior

a. U_Fact_UserRating

b. U_Fact_PassedCourses

c. U_Fact_UserRate_Acc

d. U_Fact_Comments

e. U_Fact_CommentRating

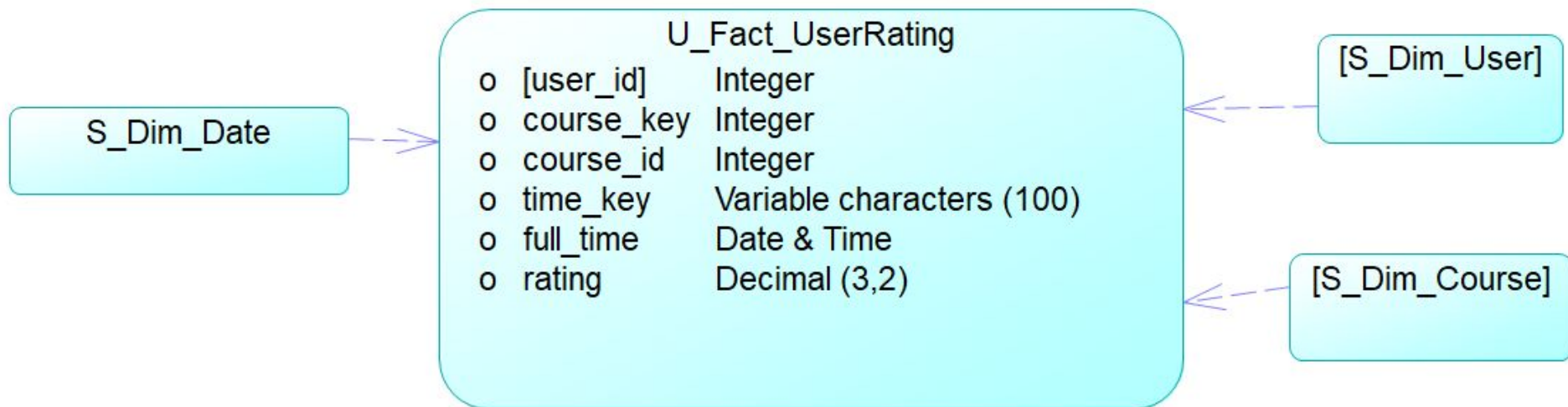
f. U_Fact_InfluentialUsers_Acc

5. مارت Human Resources

فکت تراکنشی U_Fact_UserRating

- فکت تراکنشی در مارت User Behavior است .
- هدف آن این است که rating یوزر ها به هر course را در طول زمان نگهداری کند.
- برای تکمیل آن از جدول UserOnlineCourse در staging area استفاده می شود.

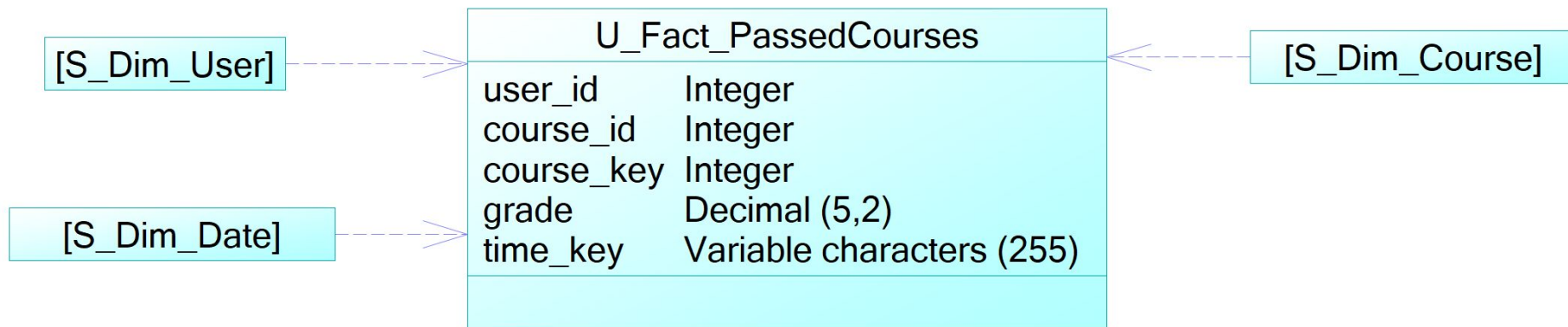
فکت تراکنشی U_Fact_UserRating



فکت تراکنشی U_Fact_PassedCourses

- یک فکت تراکنشی در مارت User Behavior است.
- هدف این فکت نگهداری دوره های پاس شده از هر یوزر است.
- از این فکت در فکت تجمیعی U_Fact_UserRate_Acc نیز استفاده می شود.

فکت تراکنشی U_Fact_PassedCourses



فکت تجمیعی U_Fact_UserRate_Acc

- یک فکت تجمیعی در مارت User Behavior است.
- هدف آن این است که مشخص کند هر کاربر کل خریدی که تاکنون کرده است چقدر بوده و چه تعداد دوره‌ای شرکت کرده است و میانگین ریتینگ که تاکنون داده چند بوده است.
- برای تکمیل از فکت تراکنشی U_Fact_UserRating و یک ویو گرفته شده از C_Fact_CourseBuying استفاده می‌کند تا به کورس‌های که کاربران اخذ کرده‌اند دسترسی داشته باشد.
- از این فکت می‌توان به عنوان داده‌های رکامندر سیستم استفاده کرد.

فکت تجمیعی U_Fact_UserRate_Acc

U_Fact_UserRate_Acc	
[user_id]	Integer
sum_of_courses	Integer
sum_of_passed_courses	Integer
average_rating	Decimal (3,2)
number_of_rating	Integer
max_rating	Decimal (3,2)

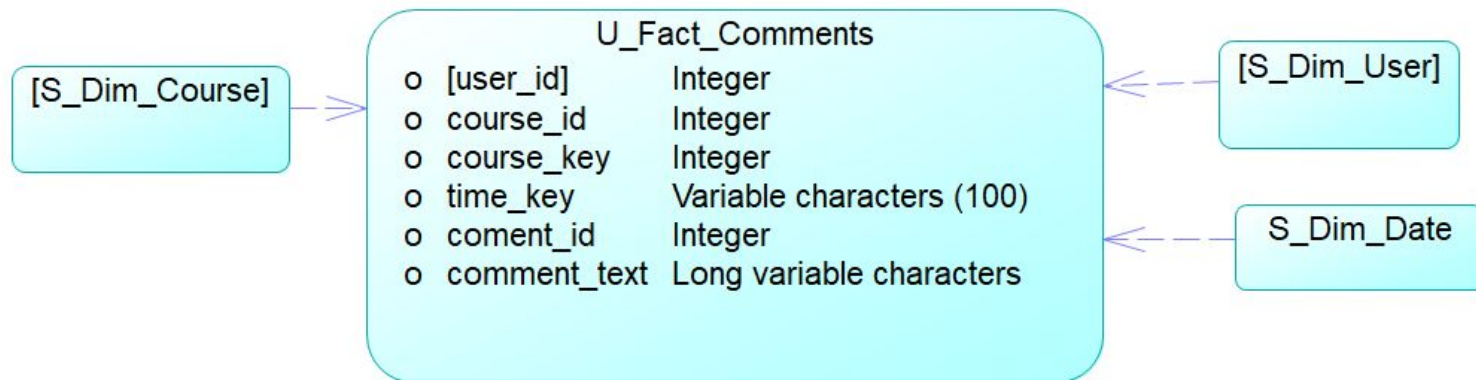
S_Dim_User



فکت تراکنشی U_Fact_Comments:

- یک فکت تراکنشی در مارت User Behavior است.
- هدف آن نگهداری این است که یوزر به کدام course در چه تاریخی comment قرار داده است.
- داده آن از Comment در staging area آورده می شود.
- در فکت های تجمعی از آن استفاده شده است.

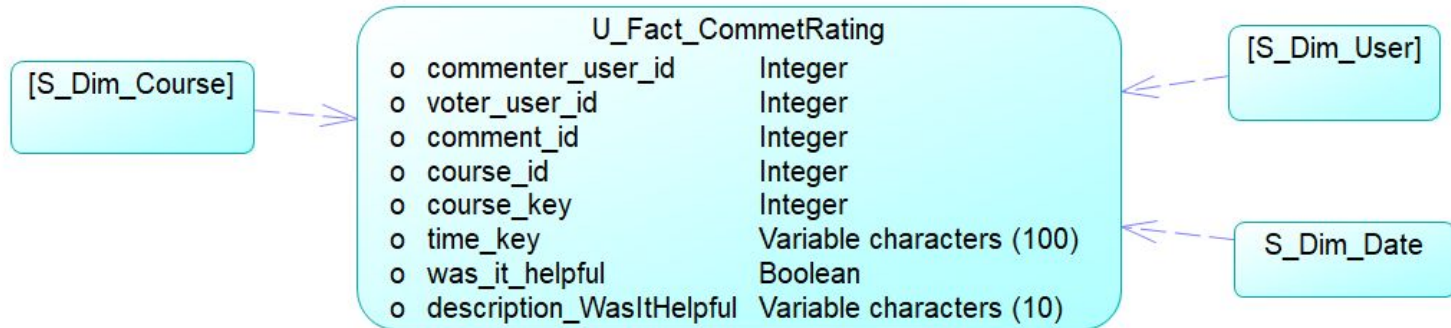
فکت تراکنشی :U_Fact_Comments



فکت تراکنشی U_Fact_CommetRating

- یک فکت تراکنشی در مارت User Behavior است.
- این فکت بر محور was_it_helpful در دیتابیس source است و هدف آن این است که به صورت تراکنشی was_it_helpful ها را نگهداری کند تا در نهایت بتواند کاربرانی که بیشترین کامنت مفید را دارند شناسایی کند.
- این فکت برای محاسبه‌ی فکت تجمیعی U_Fact_InfluentialUser استفاده می‌شود.

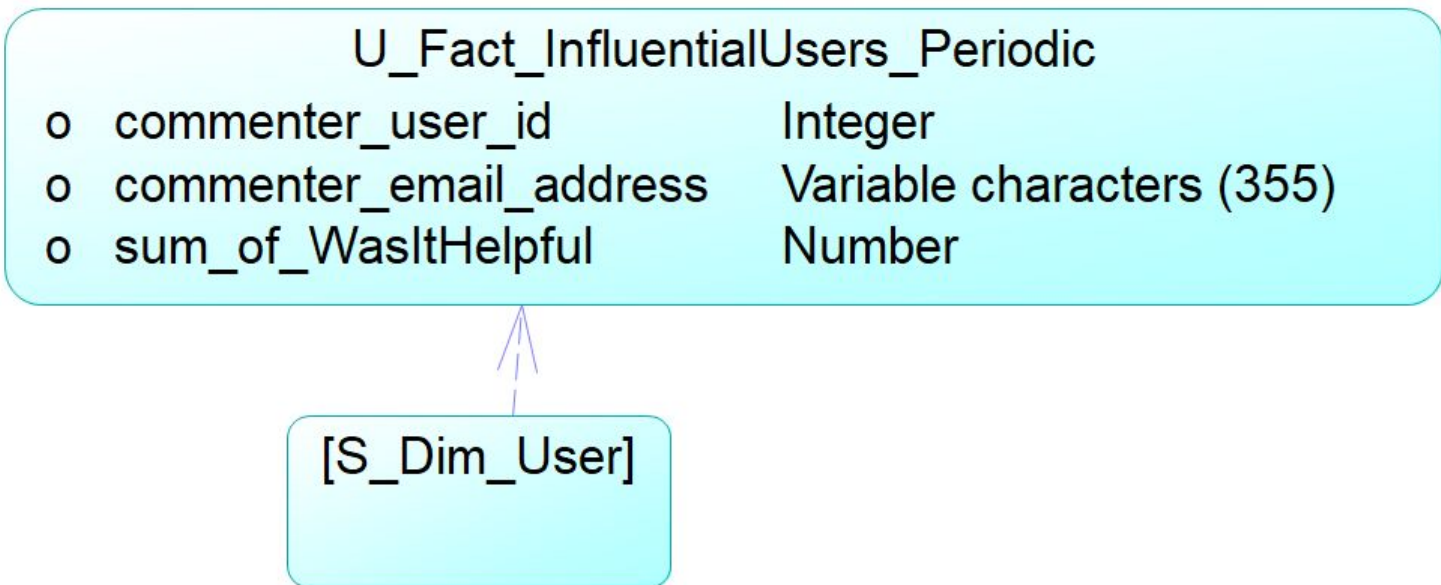
فکت تراکنشی U_Fact_CommetRating



فکت دوره‌ای **U_Fact_InfluentialUsers_Periodic**

- یک فکت تجمیعی در مارت User Behavior است.
- این فکت با محوریت فکت U_Fact_CommentRating پر می‌شود.
- هدف آن پیدا کردن کاربرانی است که بیشترین کامنت تاثیرگذار را داشته اند.

فکت دوره‌ای U_Fact_InfluentialUsers_Periodic



1. توضیح انبار داده

2. بعدها

3. مارت Course

4. مارت User Behavior

5. مارت Human Resources

a. HR_Fact_InstructorCourse_T

b. HR_Fact_InstructorCourse_ACC

c. HR_Fact_InstructorRate_Daily

d. HR_Fact_StaffPayment_T

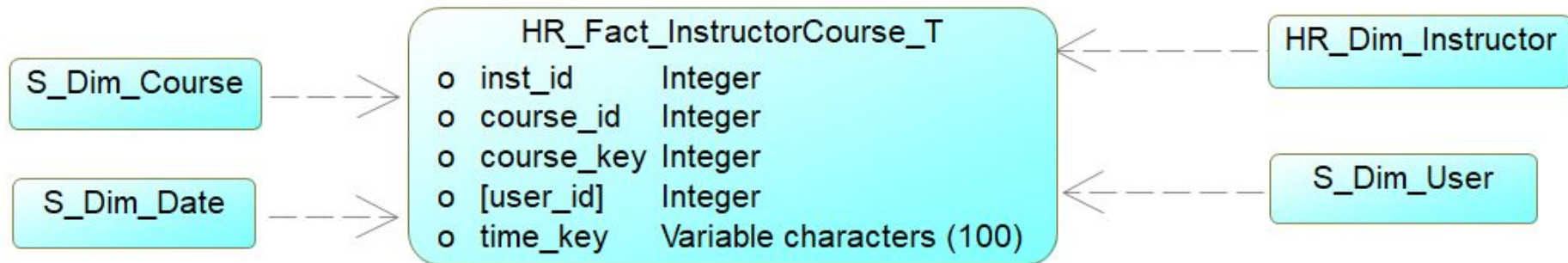
e. HR_Fact_StaffPayment_Yearly

f. HR_Fact_Tickets_Daily

فکت تراکنشی HR_Fact_InstructorCourse_T

- در این جدول فکت، به ازای ثبت نام هر کاربر در یک کورس، برای تمامی مدرس ها مشارکت کننده در آن کورس یک رکورد ثبت می شود.
- با کمک رکورد های این فکت امکان گزارش گیری های کاربردی به صورت تجمعی فراهم می شود.

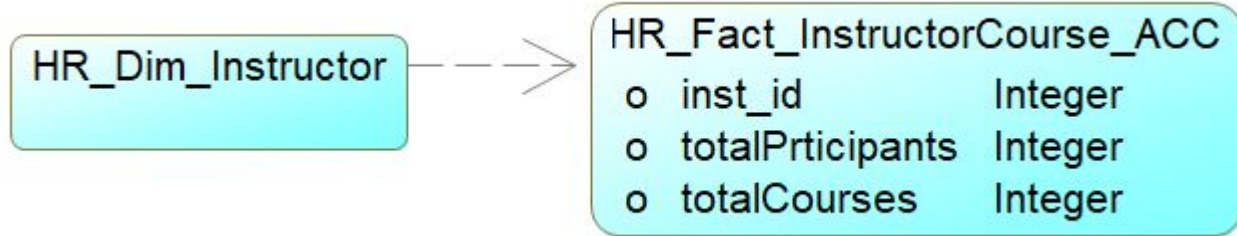
فکت تراکنشی HR_Fact_InstructorCourse_T



فکت تجمعی HR_Fact_InstructorCourse_ACC

- در این جدول فکت تجمعی، برای هر مدرس تعداد دانشجویان او و تعداد کورس هایی که او در آن ها مشارکت دارد مشخص شده است.
- این جدول از جدول HR_Fact_InstructorCourse_T تغذیه می شود.

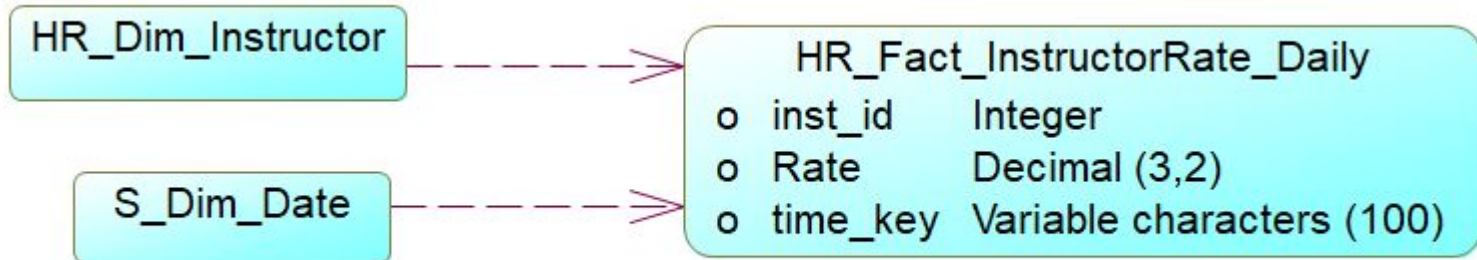
فکت تجمعی HR_Fact_InstructorCourse_ACC



فکت روزانه HR_Fact_InstructorRate_Daily

- در این جدول فکت به صورت روزانه Rate مدرس ها ثبت می شود.
- در سمت منبع، امتیاز مدرس ها با کمک یک تریگر آپدیت می شود.
- با نگهداری مسیر تغییرات Rate در انبار داده می توانیم به گزارشاتی در خصوص عملکرد مدرس ها برسیم.

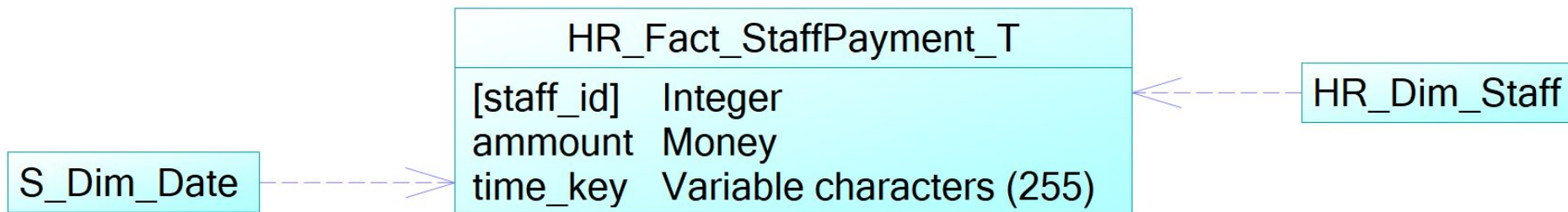
فکت روزانه HR_Fact_InstructorRate_Daily



فکت تراکنشی HR_Fact_StaffPayment_T

- این جدول فکت اطلاعات پرداختی به کارکنان را ذخیره می کند.
- هدف این فکت این است که بتوانیم گزارش های مربوط به پرداختی های کارکنان را تهیه کنیم.
- حقوق دریافتی هر شخص در هر تاریخ در این جدول ذخیره می شود.

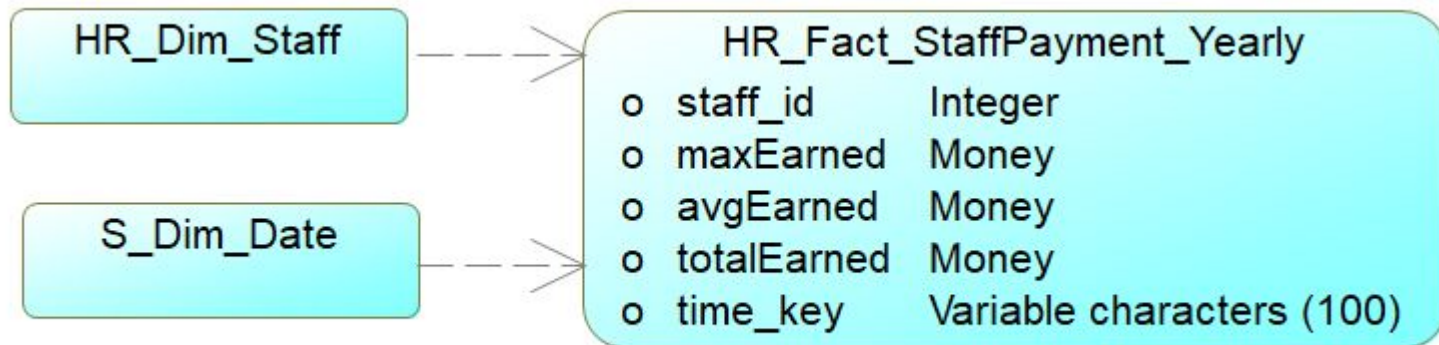
فکت تراکنشی HR_Fact_StaffPayment_T



فکت سالانه HR_Fact_StaffPayment_Yearly

- این فکت شامل اطلاعات پرداختی به هر یک از کارمندان به صورت سالیانه است.
- هدف این فکت دریافت گزارش های مربوط به پرداخت ها به کارکنان است.
- از این گزارشات برای تغییر و بهبود سیستم دستمزد و پاداش ها می توان استفاده کرد.

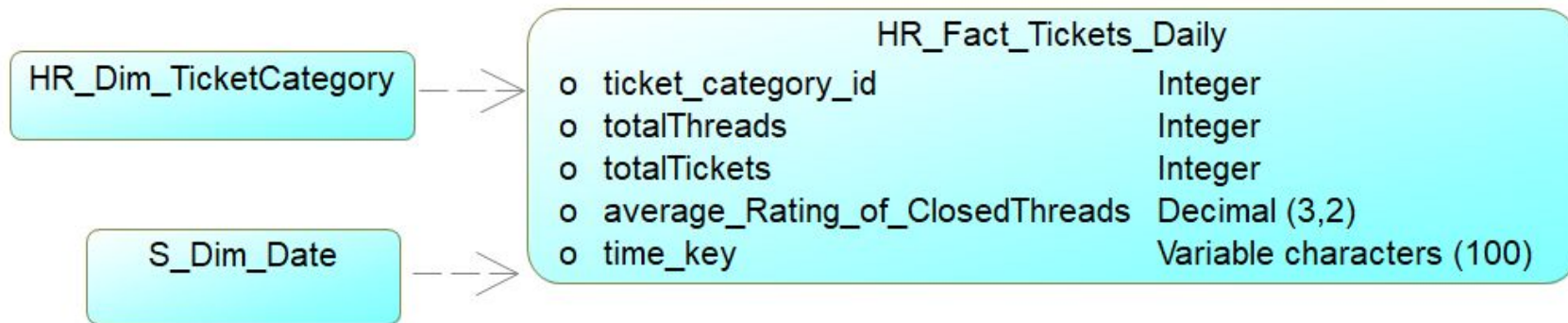
فکت سالانه HR_Fact_StaffPayment_Yearly



فکت روزانه HR_Fact_Tickets_Daily

- این جدول به صورت روزانه مجموع تعداد تردهای ایجاد شده، مجموع تعداد بلیط های ایجاد شده، و میانگین رای داده شده به تردهای پایان یافته را بر اساس دسته بندی موجود برای بلیط ها ثبت می کند.
- هدف آن دریافت گزارش به منظور تحلیل بازخورد کاربران سیستم و بررسی روند پاسخگویی به دغدغه های کاربران از طرف تیم پشتیبانی است.

فکت روزانه HR_Fact_Tickets_Daily



سیاس از توجهتون