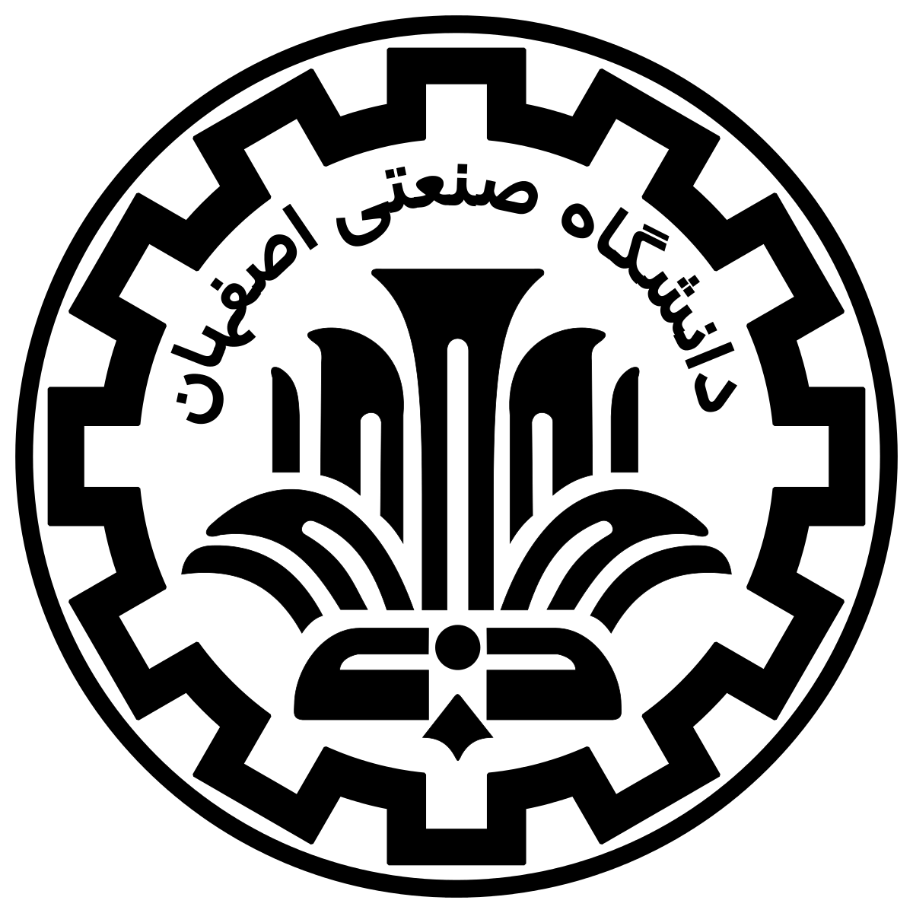
باسمه تعالی



دانشگاه صنعتي اصفهان

دانشکده برق و کامپيوتر

**طراحی انبار داده اجاره فیلم Sakila**

درس:

پایگاه داده 2

استاد درس:

دکتر علیرضا بصیری

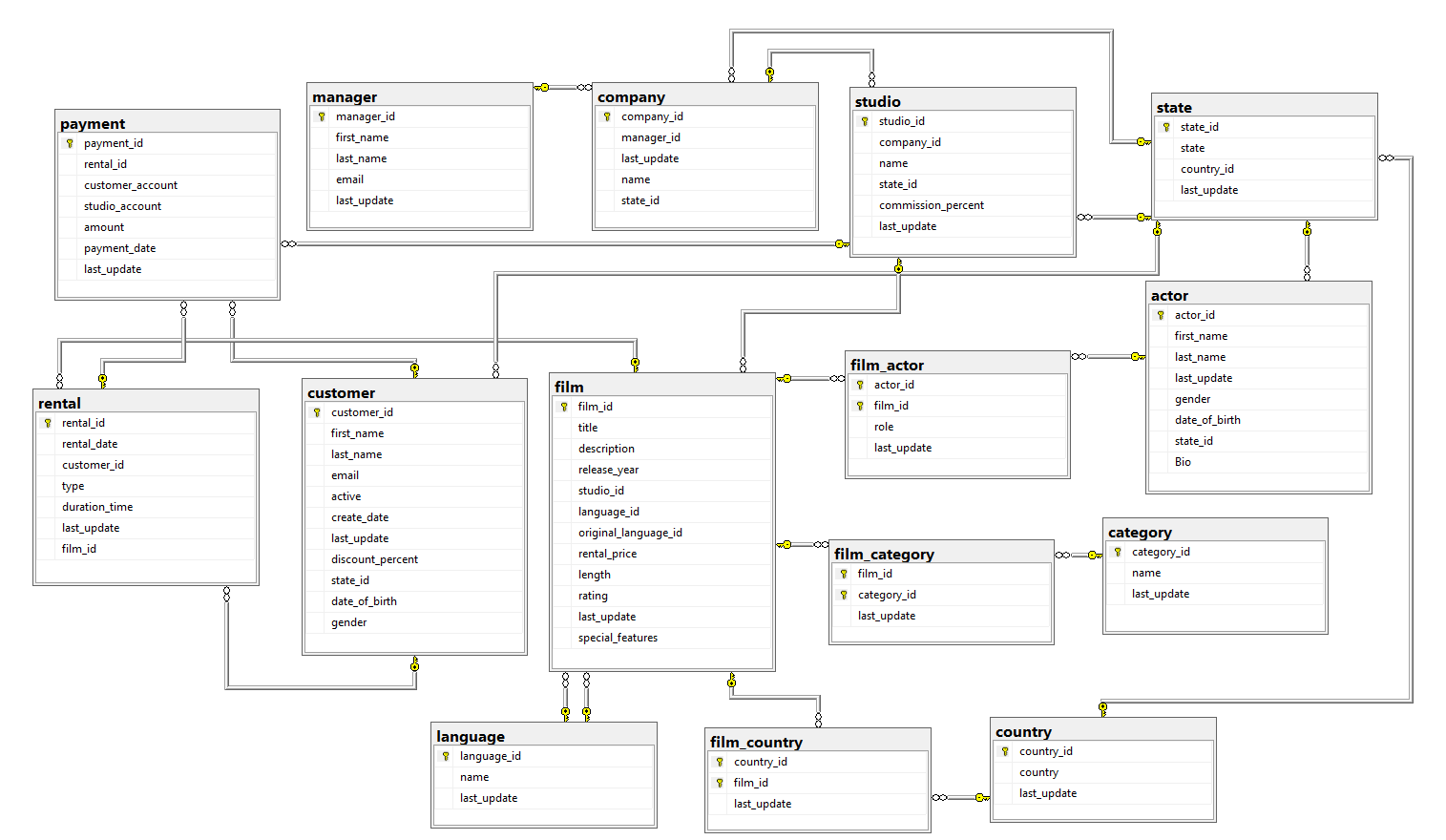
اعضای تیم:

سید وحید فاطمی (9927543)

حوری دهش (9821413)

تیر 1403

دیتابیس سورس



دیتابیس سورس ما راجب یک سایتی همچون youtube (یکی از امکانات youtube) و ... است که فیلم اجاره می‌دهد. این فیلم ها برای سایر کمپانی ها است که در استودیوهای خودشون ساخته شده‌اند.

هر فیلم متعلق به یک استودیو است و هر استودیو متعلق به یک کمپانی است.

برای فیلم، زبان اصلی، زبان دوبله شده (در صورت وجود)، بازیگران کشور ساخت و کتگوری فیلم مشخص شده است. (همچنین یک سری اطلاعات ریز دیگر نیز دارد)

هر مشتری میتواند فیلم مورد علاقه خودش را اجاره کند برای خودش یا دیگری. (اگر شخص A فیلم را برای شخص B اجاره کرده باشد، آیدی شخص A در جدول Payment میخورد و آیدی شخص B در جدول Rental میخورد.)

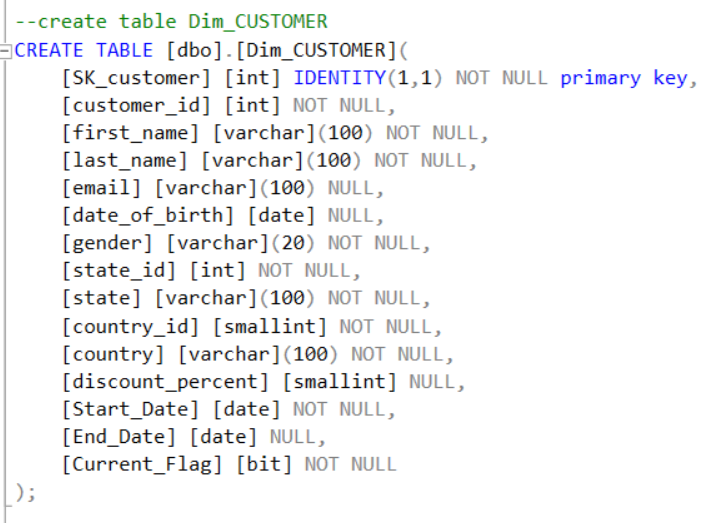
به هر مشتری ممکنه درصد مشخصی تخفیف داده شده باشد.

در هر استودیو یک درصد کمیسیون وجود دارد که نشان دهنده این است که آن درصد از فروش آن استودیو باید به ما پرداخته شود.

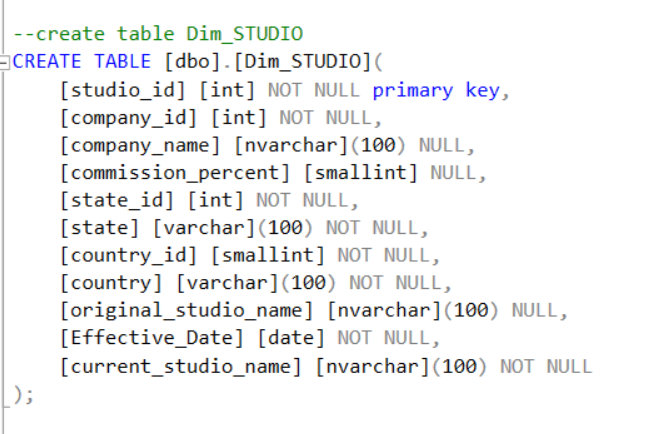
دیتابیس بر پایه دیتابیسی به نام Sakila که در اینترنت وجود دارد، است اما تغییراتی در آن لحاظ شده است.

دیتامارت اجاره

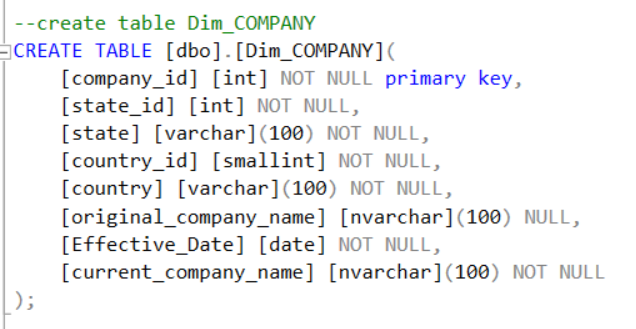
Dimention Tables



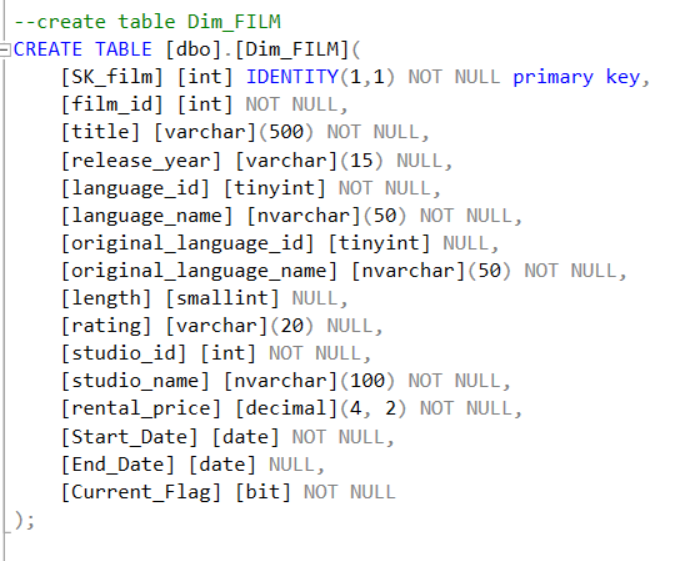
|  |  |
| --- | --- |
| Dim\_CUSTOMER | |
| ستون | شرح |
| SK\_customer | این ستون برای لحاظ کردن scd نوع 2 ایجاد شده است و توسط خود دیتابیس مقداردهی میشود (identity است) |
| customer\_id | از ستون customer\_id جدول customer موجود در Staging Area که حاوی داده‌های جدول customer منبع داده است به دست می‌آید. |
| first\_name | از ستون first\_name جدول customer موجود در Staging Area که حاوی داده‌های جدول customer منبع داده است به دست می‌آید. |
| last\_name | از ستون last\_name جدول customer موجود در Staging Area که حاوی داده‌های جدول customer منبع داده است به دست می‌آید. |
| email | از ستون email جدول customer موجود در Staging Area که حاوی داده‌های جدول customer منبع داده است به دست می‌آید. |
| date\_of\_birth | از ستون date\_of\_birth جدول customer موجود در Staging Area که حاوی داده‌های جدول customer منبع داده است به دست می‌آید. |
| gender | از ستون gender جدول customer موجود در Staging Area که حاوی داده‌های جدول customer منبع داده است به دست می‌آید.  این ستون در پایگاه داده سورس مقدار 0 و1 داشته و وقتی که داخل Staging Area لودش میکنیم مقدار male و female میگیرد. |
| state\_id | از ستون state\_id جدول customer موجود در Staging Area که حاوی داده‌های جدول customer منبع داده است به دست می‌آید. |
| state | از ستون state جدول state موجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country\_id | از ستون country\_id جدول state موجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country | از ستون country جدول country موجود در Staging Area که حاوی داده‌های جدول country منبع داده است به دست می‌آید. |
| discount\_percent | از ستون discount\_percent جدول customer موجود در Staging Area که حاوی داده‌های جدول customer منبع داده است به دست می‌آید.  Scd نوع 2 بر روی این ستون لحاظ شده است |
| Start\_Date | تاریخی است که این رکورد ایجاد شده است. |
| End\_Date | تاریخی است که این رکورد منقضی شده است و یک رکورد دیگر با همین customer\_id ولی با SK\_ customer متفاوت ایجاد شده است |
| Current\_Flag | نشان می‌دهد که این رکورد معتبر است (1) یا خیر (0) |



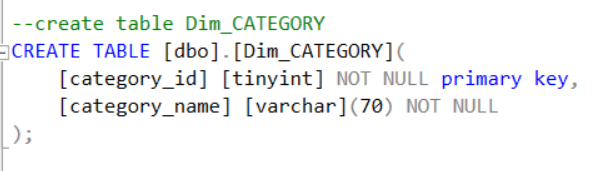
|  |  |
| --- | --- |
| Dim\_STUDIO | |
| ستون | شرح |
| studio\_id | از ستون studio\_id جدول studio موجود در Staging Area که حاوی داده‌های جدول studio منبع داده است به دست می‌آید. |
| company\_id | از ستون company\_id جدول studio موجود در Staging Area که حاوی داده‌های جدول studio منبع داده است به دست می‌آید. |
| company\_name | از ستون name جدول company موجود در Staging Area که حاوی داده‌های جدول company منبع داده است به دست می‌آید. |
| commission\_percent | از ستون commission\_percent جدول studio موجود در Staging Area که حاوی داده‌های جدول studio منبع داده است به دست می‌آید.  بر روی این ستون scd نوع 1 لحاظ شده است |
| state\_id | از ستون state\_id جدول studio موجود در Staging Area که حاوی داده‌های جدول studio منبع داده است به دست می‌آید. |
| state | از ستون state جدول state موجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country\_id | از ستون country\_id جدول state موجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country | از ستون country جدول country موجود در Staging Area که حاوی داده‌های جدول country منبع داده است به دست می‌آید. |
| original\_studio\_name | از ستون name جدول studio موجود در Staging Area که حاوی داده‌های جدول studio منبع داده است به دست می‌آید.  این رکورد نام قبلی این استودیو را نشان می‌دهد (scd نوع 3 بر روی نام استودیو) |
| Effective\_Date | تاریخی است که نام استودیو تغییر کرده است |
| current\_studio\_name | این ستون نام فعلی استودیو را نشان می‌دهد. |



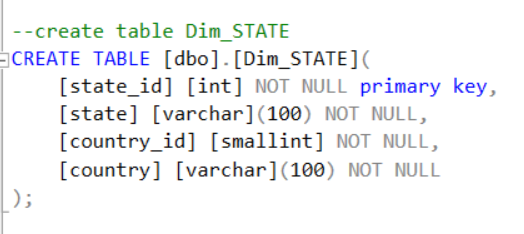
|  |  |
| --- | --- |
| Dim\_COMPANY | |
| ستون | شرح |
| company\_id | از ستون company\_id جدول company موجود در Staging Area که حاوی داده‌های جدول company منبع داده است به دست می‌آید. |
| state\_id | از ستون state\_id جدول company موجود در Staging Area که حاوی داده‌های جدول company منبع داده است به دست می‌آید. |
| state | از ستون state جدول state موجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country\_id | از ستون country\_id جدول state موجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country | از ستون country جدول country موجود در Staging Area که حاوی داده‌های جدول country منبع داده است به دست می‌آید. |
| original\_company\_name | از ستون original\_company\_name جدول company موجود در Staging Area که حاوی داده‌های جدول company منبع داده است به دست می‌آید.  این رکورد نام قبلی این کمپانی را نشان می‌دهد (scd نوع 3 بر روی نام کمپانی) |
| Effective\_Date | تاریخی است که نام کمپانی تغییر کرده است |
| current\_company\_name | این ستون نام فعلی کمپانی را نشان می‌دهد. |



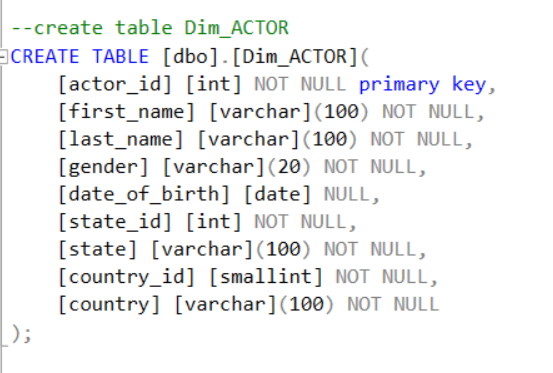
|  |  |
| --- | --- |
| Dim\_FILM | |
| ستون | شرح |
| SK\_film | این ستون برای لحاظ کردن scd نوع 2 ایجاد شده است و توسط خود دیتابیس مقداردهی میشود (identity است) |
| film\_id | از ستون film\_id جدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید. |
| title | از ستون title جدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید. |
| release\_year | از ستون release\_year جدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید. |
| language\_id | از ستون language\_id جدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید. |
| language\_name | از ستون name جدول language موجود در Staging Area که حاوی داده‌های جدول language منبع داده است به دست می‌آید. |
| original\_language\_id | از ستون original\_language\_id جدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید. |
| original\_language\_name | از ستون name جدول language موجود در Staging Area که حاوی داده‌های جدول language منبع داده است به دست می‌آید. |
| length | از ستون length جدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید. |
| rating | از ستون rating جدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید. |
| studio\_id | از ستون studio\_idجدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید. |
| studio\_name | از ستون name جدول studioموجود در Staging Area که حاوی داده‌های جدول studio منبع داده است به دست می‌آید. |
| rental\_price | از ستون rental\_price جدول film موجود در Staging Area که حاوی داده‌های جدول film منبع داده است به دست می‌آید.  Scd نوع 2 بر روی این ستون لحاظ شده است |
| Start\_Date | تاریخی است که این رکورد ایجاد شده است. |
| End\_Date | تاریخی است که این رکورد منقضی شده است و یک رکورد دیگر با همین film\_id ولی با SK\_film متفاوت ایجاد شده است |
| Current\_Flag | نشان می‌دهد که این رکورد معتبر است (1) یا خیر (0) |



|  |  |
| --- | --- |
| Dim\_CATEGORY | |
| ستون | شرح |
| category\_id | از ستون category\_idجدول categoryموجود در Staging Area که حاوی داده‌های جدول category منبع داده است به دست می‌آید. |
| category\_name | از ستون nameجدول categoryموجود در Staging Area که حاوی داده‌های جدول category منبع داده است به دست می‌آید. |

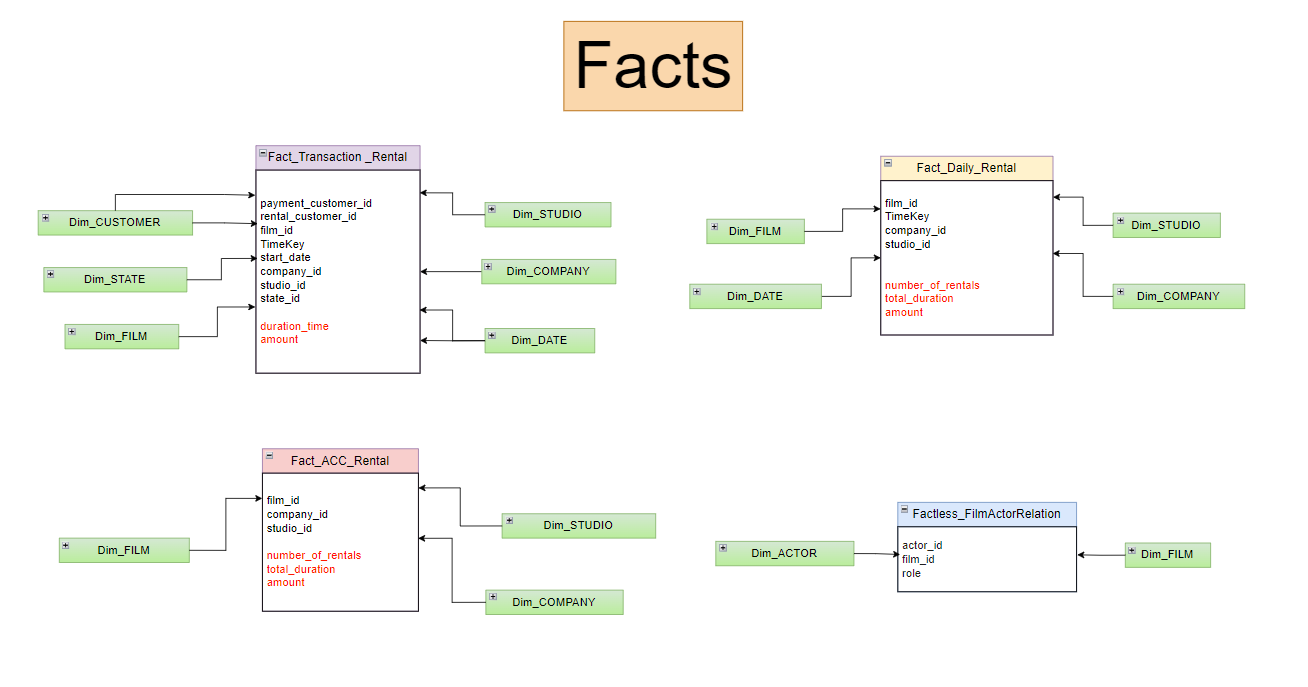


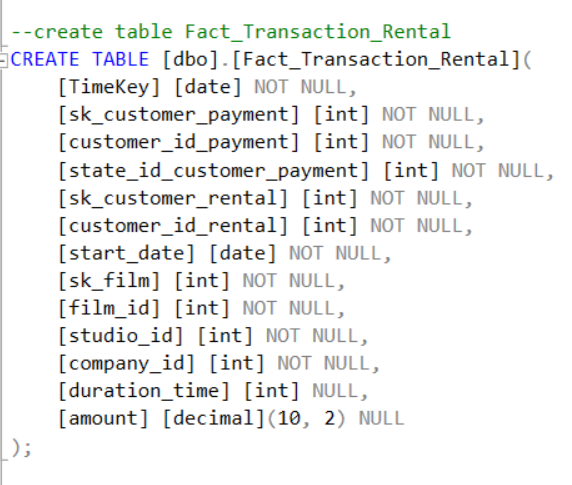
|  |  |
| --- | --- |
| Dim\_STATE | |
| ستون | شرح |
| state\_id | از ستون state\_idجدول stateموجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| state | از ستون stateجدول stateموجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country\_id | از ستون country\_idجدول stateموجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country | از ستون countryجدول countryموجود در Staging Area که حاوی داده‌های جدول country منبع داده است به دست می‌آید. |



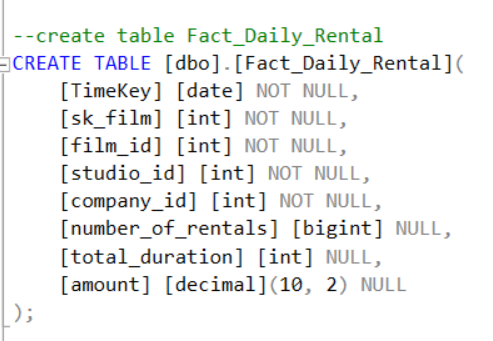
|  |  |
| --- | --- |
| Dim\_ACTOR | |
| ستون | شرح |
| actor\_id | از ستون actor\_idجدول actorموجود در Staging Area که حاوی داده‌های جدول actor منبع داده است به دست می‌آید. |
| first\_name | از ستون first\_nameجدول actorموجود در Staging Area که حاوی داده‌های جدول actor منبع داده است به دست می‌آید. |
| last\_name | از ستون last\_nameجدول actorموجود در Staging Area که حاوی داده‌های جدول actor منبع داده است به دست می‌آید. |
| gender | از ستون genderجدول actorموجود در Staging Area که حاوی داده‌های جدول actor منبع داده است به دست می‌آید. |
| date\_of\_birth | از ستون date\_of\_birthجدول actorموجود در Staging Area که حاوی داده‌های جدول actor منبع داده است به دست می‌آید. |
| state\_id | از ستون state\_idجدول actorموجود در Staging Area که حاوی داده‌های جدول actor منبع داده است به دست می‌آید. |
| state | از ستون stateجدول stateموجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country\_id | از ستون country\_idجدول stateموجود در Staging Area که حاوی داده‌های جدول state منبع داده است به دست می‌آید. |
| country | از ستون countryجدول countryموجود در Staging Area که حاوی داده‌های جدول country منبع داده است به دست می‌آید. |

Tables of rental fact

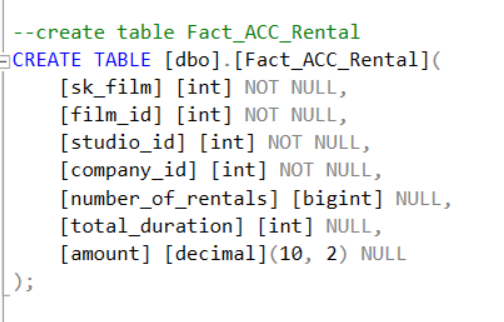




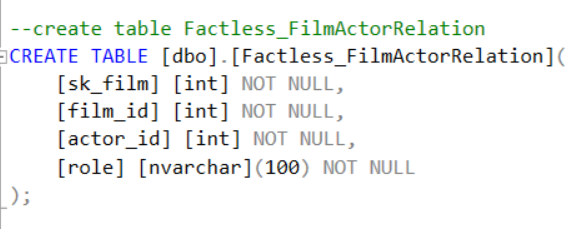
|  |  |
| --- | --- |
| Fact\_Transaction\_Rental | |
| ستون | شرح |
| TimeKey | از ستون TimeKey جدول Dim\_Date در Data Warehouse به دست می‌آید. |
| sk\_customer\_payment | از ستون SK\_customer جدول Dim\_CUSTOMER در Data Warehouse به دست می‌آید. |
| customer\_id\_payment | از ستون customer\_account جدول payment در Staging Area به دست می‌آید. (برای بحث maintenance آورده شده) |
| state\_id\_customer\_payment | از ستون state\_id جدول Dim\_CUSTOMER در Data Warehouse به دست می‌آید. |
| sk\_customer\_rental | از ستون SK\_customer جدول Dim\_CUSTOMER در Data Warehouse به دست می‌آید. |
| customer\_id\_rental | از ستون customer\_id جدول rental در Staging Area به دست می‌آید. (برای بحث maintenance آورده شده) |
| start\_date | از ستون rental\_date جدول rental در Staging Area به دست می‌آید. |
| sk\_film | از ستون SK\_film جدول Dim\_FILM در Data Warehouse به دست می‌آید. |
| film\_id | از ستون film\_id جدول rental در Staging Area به دست می‌آید.  (برای بحث maintenance آورده شده) |
| studio\_id | از ستون studio\_account جدول payment در Staging Area به دست می‌آید. |
| company\_id | از ستون company\_id جدول Dim\_STUDIO در Data Warehouse به دست می‌آید. |
| duration\_time | از ستون duration\_time جدول rental در Staging Area به دست می‌آید. |
| amount | از ستون amount جدول payment در Staging Area به دست می‌آید. |



|  |  |
| --- | --- |
| Fact\_Daily\_Rental | |
| ستون | شرح |
| TimeKey | از ستون TimeKey جدول Dim\_Date در Data Warehouse به دست می‌آید. |
| sk\_film | از ستون SK\_film جدول Dim\_FILM در Data Warehouse به دست می‌آید. |
| film\_id | از ستون film\_id جدول Dim\_FILM در Data Warehouse به دست می‌آید.  (برای بحث maintenance آورده شده) |
| studio\_id | از ستون studio\_id جدول Dim\_FILM در Data Warehouse به دست می‌آید. |
| company\_id | از ستون company\_id جدول Dim\_STUDIO در Data Warehouse به دست می‌آید. |
| number\_of\_rentals | در جدول Fact\_Transaction\_Rental، Group by میزنیم بر روی چهار فیلد  TimeKey، film\_id، studio\_id و company\_id و count(\*) را محاسبه میکنیم. (برای هر روز) |
| total\_duration | در جدول Fact\_Transaction\_Rental، Group by میزنیم بر روی چهار فیلد  TimeKey، film\_id، studio\_id و company\_id و sum(duration\_time) را محاسبه میکنیم. (برای هر روز) |
| amount | در جدول Fact\_Transaction\_Rental، Group by میزنیم بر روی چهار فیلد  TimeKey، film\_id، studio\_id و company\_id و sum(amount) را محاسبه میکنیم. (برای هر روز) |

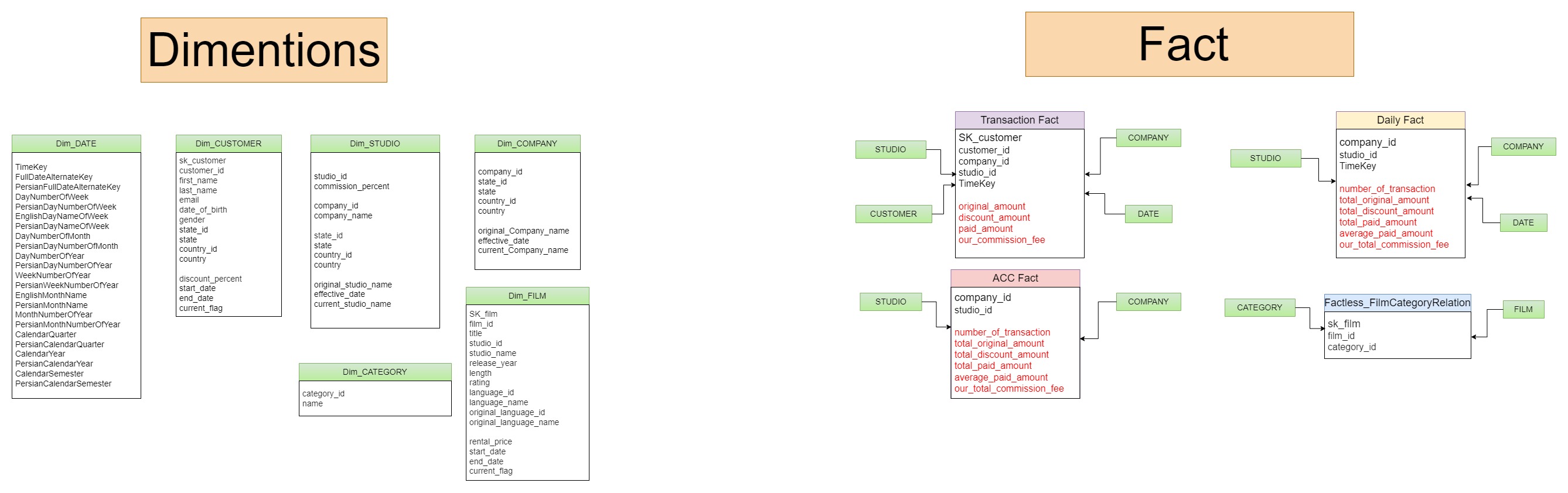


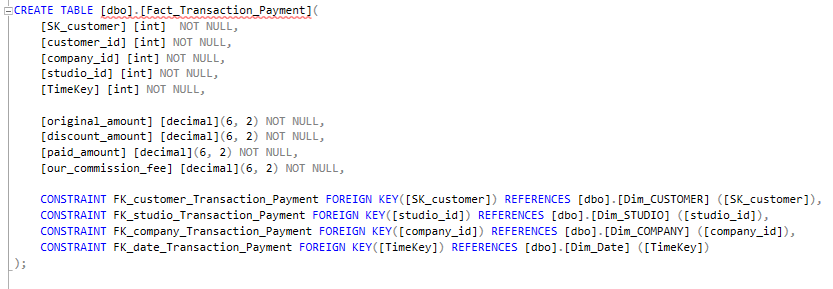
|  |  |
| --- | --- |
| Fact\_ACC\_Rental | |
| ستون | شرح |
| sk\_film | از ستون SK\_film جدول Dim\_FILM در Data Warehouse به دست می‌آید. |
| film\_id | از ستون film\_id جدول Dim\_FILM در Data Warehouse به دست می‌آید.  (برای بحث maintenance آورده شده) |
| studio\_id | از ستون studio\_id جدول Dim\_FILM در Data Warehouse به دست می‌آید. |
| company\_id | از ستون company\_id جدول Dim\_STUDIO در Data Warehouse به دست می‌آید. |
| number\_of\_rentals | در جدول Fact\_Daily\_Rental، Group by میزنیم بر روی سه فیلد film\_id، studio\_id و company\_id و count(\*) را محاسبه میکنیم. (برای هر روز) |
| total\_duration | در جدول Fact\_Daily\_Rental، Group by میزنیم بر روی سه فیلد film\_id، studio\_id و company\_id و sum(duration\_time) را محاسبه میکنیم. (برای هر روز) |
| amount | در جدول Fact\_Daily\_Rental، Group by میزنیم بر روی سه فیلد film\_id، studio\_id و company\_id و sum(amount) را محاسبه میکنیم. (برای هر روز) |



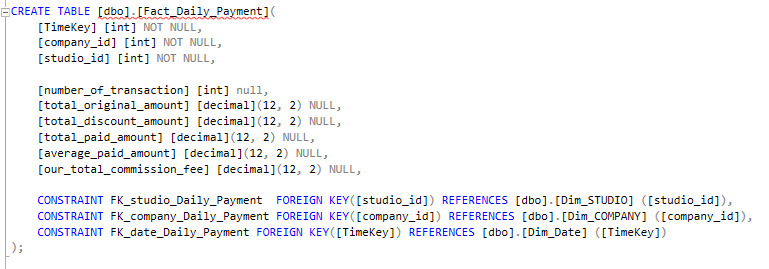
|  |  |
| --- | --- |
| Factless\_FilmActorRelation | |
| ستون | شرح |
| sk\_film | از ستون SK\_film جدول Dim\_FILM در Data Warehouse به دست می‌آید. |
| film\_id | از ستون film\_id جدول Dim\_FILM در Data Warehouse به دست می‌آید. |
| actor\_id | از ستون actor\_id جدول film\_actor در Staging Area به دست می‌آید. |
| role | از ستون role جدول film\_actor در Staging Area به دست می‌آید. |

Tables of payment fact

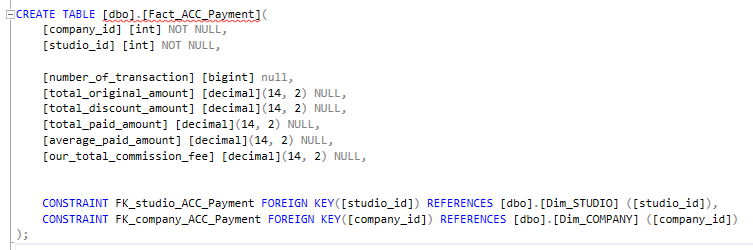




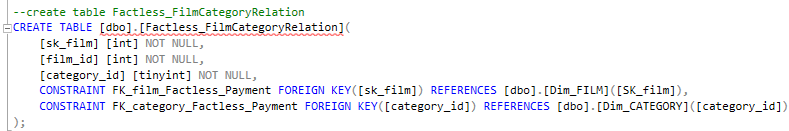
|  |  |
| --- | --- |
| Fact\_Transaction\_Payment | |
| ستون | شرح |
| SK\_customer | از ستون SK\_customer جدول Dim\_CUSTOMER در Data Warehouse به دست می‌آید. |
| customer\_id | از ستون customer\_id جدول Dim\_CUSTOMER در Data Warehouse به دست می‌آید. (برای بحث maintenance آورده شده) |
| company\_id | از ستون company\_id جدول Dim\_COMPANY در Data Warehouse به دست می‌آید. |
| studio\_id | از ستون studio\_id جدول Dim\_STUDIO در Data Warehouse به دست می‌آید. |
| TimeKey | از ستون TimeKey جدول Dim\_Date در دیتابیس DW\_sakila گرفته شده است |
| original\_amount | ازحاصل ضرب مقدار ستون amount جدول payment به دست می اید:  مقدار discount\_percent از جدول کاستومر برداشته میشود) |
| discount\_amount |  |
| paid\_amount | از ستون amount جدول payment در Staging Area به دست می‌آید. نشان دهنده مقداری است که مشتری پرداخت کرده است (با لحاظ کردن تخفیف) |
| our\_commission\_fee |  |



|  |  |
| --- | --- |
| Fact\_Daily\_Payment | |
| ستون | شرح |
| company\_id | از ستون company\_id جدول Dim\_COMPANY در Data Warehouse به دست می‌آید. |
| studio\_id | از ستون studio\_id جدول Dim\_STUDIO در Data Warehouse به دست می‌آید. |
| TimeKey | از ستون TimeKey جدول Dim\_Date در دیتابیس DW\_sakila گرفته شده است |
| number\_of\_transaction | در جدول Fact\_Transaction\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و count(\*) را محاسبه میکنیم. (برای هر روز) |
| Total\_original\_amount | در جدول Fact\_Transaction\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(original\_amount) را محاسبه میکنیم. (برای هر روز) |
| Total\_discount\_amount | در جدول Fact\_Transaction\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(discount\_amount)) را محاسبه میکنیم. (برای هر روز) |
| Total\_paid\_amount | در جدول Fact\_Transaction\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(paid\_amount) را محاسبه میکنیم. (برای هر روز) |
| average\_paid\_amount | در جدول Fact\_Transaction\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و avg(paid\_amount) را محاسبه میکنیم. (برای هر روز) |
| our\_ total\_commission\_fee | در جدول Fact\_Transaction\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(our\_commission\_fee) را محاسبه میکنیم. (برای هر روز) |



|  |  |
| --- | --- |
| Fact\_ACC\_Payment | |
| ستون | شرح |
| company\_id | از ستون company\_id جدول Dim\_COMPANY در Data Warehouse به دست می‌آید. |
| studio\_id | از ستون studio\_id جدول Dim\_STUDIO در Data Warehouse به دست می‌آید. |
| TimeKey | از ستون TimeKey جدول Dim\_Date در دیتابیس DW\_sakila گرفته شده است |
| number\_of\_transaction | در جدول Fact\_Daily\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(number\_of\_transaction) را محاسبه میکنیم. |
| Total\_original\_amount | در جدول Fact\_Transaction\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(total\_original\_amount) را محاسبه میکنیم. |
| Total\_discount\_amount | در جدول Fact\_Daily\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(total\_discount\_amount) را محاسبه میکنیم. |
| Total\_paid\_amount | در جدول Fact\_Daily\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(total\_paid\_amount) را محاسبه میکنیم. |
| average\_paid\_amount | در جدول Fact\_Daily\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و مقدار را محاسبه میکنیم. |
| our\_ total\_commission\_fee | در جدول Fact\_Daily\_Payment، Group by میزنیم بر روی دو فیلد company\_id و studio\_id و sum(our\_total\_commission\_fee) را محاسبه میکنیم. |



|  |  |
| --- | --- |
| Factless\_FilmCategoryRelation | |
| ستون | شرح |
| sk\_film | از ستون SK\_film جدول Dim\_Film در Data Warehouse به دست می‌آید. |
| film\_id | از ستون film\_id جدول Dim\_Film در Data Warehouse به دست می‌آید. |
| category\_id | از ستون category \_id جدول Dim\_ category در Data Warehouse به دست می‌آید. |