

به نام خدا



دانشگاه صنعتی اصفهان

دانشکده مهندسی برق و کامپیوتر

## ETL Document

موضوع انبار داده:

فروشگاه آنلاین بازی (مانند steam)

استاد درس:

دکتر علیرضا بصیری

توسط:

رسول کامکار

## فهرست مطالب

۳.....	توضیحات سورس و پروژه.....
۴.....	دیتامارت سفارش کاربران.....
۴.....	بعد زمان.....
۵.....	جدول بعد سفارشات.....
۶.....	جدول بعد بازی ها.....
۸.....	جدول بعد کاربران.....
۱۰.....	جدول فکت تراکنشی.....
۱۲.....	جدول فکت دوره ای ماهانه.....
۱۴.....	جدول فکت تجمیعی.....

## توضیحات سورس و پروژه

سورس این پروژه دیتابیس یک فروشگاه بازی آنلاین مانند steam و epicgames است. در این سورس اطلاعات کاربران، بازی‌ها، دسته‌بندی بازی‌ها، نظر کاربران درباره بازی‌ها، کدهای تخفیف، سفارش کاربران و بازی‌های هر سفارش موجود می‌باشد.

از لحاظ مدیریتی، سفارشات کاربران برای تصمیم‌گیری حائز اهمیت است. دیتامارت موجود در این پروژه اطلاعات سفارش کاربران را برای تصمیمات مدیریتی ارائه می‌دهد.

در این دیتامارت، اطلاعات کاربران، بازی‌ها و سفارشات موجود است و برای کاربران اطلاعات سفارشات به صورت تراکنشی، ماهانه و تجمیعی ارائه می‌شود.

در این دیتامارت اطلاعات سفارشات، کاربران و بازی‌ها به عنوان ابعاد طراحی شده‌اند.

در بعد کاربران اطلاعات تاریخچه‌ای کشور کاربر به صورت scd نوع سوم فراهم شده است، علت این نوع طراحی، اهمیت وضعیت اقتصادی کاربران براساس کشور آن‌ها است. به طور مثال، ممکن است کاربران کشورهایی با وضعیت اقتصادی ضعیف‌تر، سفارشات کمتری داشته باشند که می‌توان برای این مشکل چاره‌اندیشی کرد، مانند کم کردن قیمت بازی‌ها به صورت منطقه‌ای. همینطور ممکن است کاربران به دلیل تفاوت قیمت منطقه‌ای یا مسائل دیگر، کشور خود را مرتباً تغییر دهند که این مسئله نیز می‌تواند در تصمیمات مهم باشد. نمونه این تصمیمات مدیریتی بر اساس کشور در فروشگاه steam هم انجام شد که می‌توانید [در این مقاله](#) راجع به آن مطالعه کنید.

## دیتامارت سفارش کاربران

بعد زمان

date_dim		
time_key 🔑		int4
full_date_alterate_key		date
calendar_year		int2
day_number_of_week		int2
english_day_name_of_week		varchar(30)
day_number_of_month		int2
day_number_of_year		int2
week_number_of_year		int2
english_month_name		varchar(10)
month_number_of_year		int2


## جدول بعد سفارشات

orders_dim_scd1	
order_id 🔗	int4
coupon_code	varchar(70)
coupon_effective_until	date
coupon_discount	decimal(6,3)
order_status	varchar(70)

**orders\_dim\_scd1**

این بعد حاوی اطلاعات مربوط به سفارشات اعم از کد تخفیف (کوپن)، تاریخ اتمام کد تخفیف، درصد تخفیف و وضعیت سفارش است. scd نوع ۱ بر روی فیلد order\_status اعمال شده که پس از ثبت سفارش میتواند تغییر وضعیت دهد (از unpaid به paid تغییر کند)

نام فیلد	نوع فیلد	شرح
order_id	int4	از ستون order_id جدول orders بدست می‌آید
coupon_code	varchar(70)	از ستون coupon_code جدول orders بدست می‌آید
coupon_effective_until	date	از ستون coupon_effective_until جدول حاصل از جویین orders و coupons روی coupon_code بدست می‌آید
coupon_discount	decimal(6,3)	از ستون coupon_discount جدول حاصل از جویین orders و coupons روی coupon_code بدست می‌آید
order_status	varchar(70)	از ستون order_status جدول orders بدست می‌آید

game_dim_scd2	
game_key 	serial
game_id	int4
game_name	varchar(350)
min_memory_size	int2
is_online	bool
max_no_players	int2
details	varchar(350)
price	decimal(8,2)
price_start_date	date
price_end_date	date
price_current_flag	bool

game\_dim\_scd2

این بعد حاوی اطلاعات مربوط به بازی‌ها اعم از کلید، شناسه، نام بازی، حداقل حافظه مورد نیاز، آنلاین بودن بازی، حداکثر بازیکن‌ها، توضیحات و قیمت است. SCD نوع ۲ بر روی فیلد قیمت اعمال شده است.

نام فیلد	نوع فیلد	شرح
game_key	sequence(int4)	این ستون surr_key است و به صورت خودکار توسط DBMS مقدار گرفته و افزایش میابد.
game_id	int4	از ستون game_id جدول games بدست می‌آید.
game_name	varchar(350)	از ستون game_name جدول games بدست می‌آید.
min_memory_size	int2	از ستون min_memory_size جدول games بدست می‌آید.
is_online	bool	از ستون online جدول games بدست می‌آید.
max_no_players	int2	از ستون max_no_players جدول games بدست می‌آید.
details	varchar(350)	از ستون details جدول games بدست می‌آید.
price	decimal(8,2)	از ستون price جدول games بدست می‌آید.

این فیلد کمکی برای scd2 تعریف شده و نشان دهنده تاریخ شروع قیمت این رکورد است	date	price_start_date
این فیلد کمکی برای scd2 تعریف شده و نشان دهنده تاریخ پایان قیمت این رکورد است	date	price_end_date
این فیلد کمکی برای scd2 تعریف شده و نشان دهنده قیمت فعال بازی است	bool	price_current_flag

## جدول بعد کاربران

users_dim_scd3	
user_id 🔗	int4
email_address	varchar(200)
first_name	varchar(200)
last_name	varchar(200)
birth_date	date
original_country_code	varchar(2)
original_country_name	varchar(70)
original_effective_until	date
previous_country_code	varchar(2)
previous_country_name	varchar(70)
previous_effective_until	date
current_country_code	varchar(2)
current_country_name	varchar(70)

users\_dim\_scd3

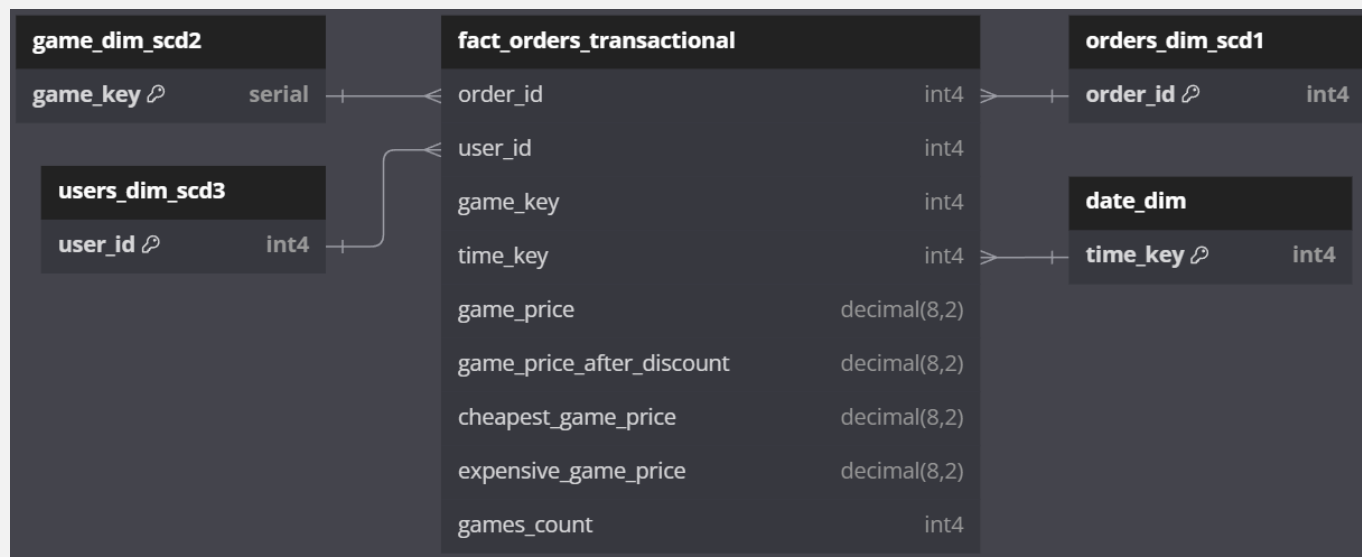
این بعد حاوی اطلاعات مربوط به کاربران اعم از شناسه، ایمیل، نام، نام خانوادگی، تاریخ تولد، کد کشور و نام کشور است. SCD نوع ۳ برای کد و نام کشور در نظر گرفته شده که فیلدهای کمکی کد و نام کشور اولیه و تاریخ تغییر آن، کد و نام کشور قبلی و تاریخ تغییر آن و کد و نام کشور فعلی برای این SCD استفاده می‌شوند.

نام فیلد	نوع فیلد	شرح
user_id	int4	از ستون user_id جدول users بدست می‌آید
email_address	varchar(200)	از ستون email_address جدول users بدست می‌آید
first_name	varchar(200)	از ستون first_name جدول users بدست می‌آید
last_name	varchar(200)	از ستون last_name جدول users بدست می‌آید
birth_date	date	از ستون birth_date جدول users بدست می‌آید
original_country_code	varchar(2)	این فیلد کمکی SCD3 برای ذخیره کد کشور اولیه کاربر است و مقدار اولیه آن تغییری نمی‌کند، از ستون country_code جدول users بدست می‌آید.



این فیلد کمکی scd3 برای ذخیره نام کشور اولیه کاربر است و مقدار اولیه آن تغییری نمی‌کند، از ستون country_name جدول حاصل از جویین users و countries بر روی فیلد country_code بدست می‌آید.	varchar(70)	original_country_name
این فیلد کمکی scd3 برای ذخیره تاریخ کشور اولیه کاربر است و در صورت تغییر اولیه کشور کاربر، مقدار تاریخ تغییر را می‌گیرد.	date	original_effective_until
این فیلد کمکی scd3 برای ذخیره کد کشور قبلی کاربر است، در صورت تغییر کشور کاربر، مقدار current_country_code را می‌گیرد.	varchar(2)	previous_country_code
این فیلد کمکی scd3 برای ذخیره نام کشور قبلی کاربر است، در صورت تغییر کشور کاربر، مقدار current_country_name را می‌گیرد.	varchar(70)	previous_country_name
این فیلد کمکی scd3 برای ذخیره تاریخ کشور کاربر است و در صورت تغییر کشور کاربر، مقدار تاریخ تغییر را می‌گیرد.	date	previous_effective_until
از ستون country_code جدول users بدست می‌آید.	varchar(2)	current_country_code
از ستون country_name جدول حاصل از جویین users و countries بر روی فیلد country_code بدست می‌آید.	varchar(70)	current_country_name

## جدول فکت تراکنشی



### fact\_orders\_transactional

این فکت به صورت تراکنشی مژرها را برای هر بازی در هر سفارش برای هر کاربر نشان می‌دهد. این مژرها عبارت اند از:

**قیمت بازی:** قیمت بازی ثبت شده در سفارش

**قیمت بعد از تخفیف:** قیمت بازی ثبت شده در سفارش پس از اعمال تخفیف سفارش

**ارزان ترین بازی:** قیمت ارزان ترین بازی در این سفارش

**گران ترین بازی:** قیمت گران ترین بازی در این سفارش

**تعداد بازی‌ها:** تعداد کل بازی‌های این سفارش

نام فیلد	نوع فیلد	شرح
order_id	int4	از ستون order_id جدول orders بدست می‌آید
user_id	int4	از ستون user_id جدول orders بدست می‌آید
game_key	int4	از ستون game_key جدول حاصل از جویین order_games با game_dim_scd2 بر روی game_id و شرط روی price_current_flag بدست می‌آید
time_key	int4	از اعمال تابع to_char(order_date, 'YYYYMMDD')::int بر روی order_date جدول orders بدست می‌آید
game_price	decimal(8,2)	از ستون game_price جدول حاصل از جویین orders با order_games بر روی order_id بدست می‌آید
game_price_after_discount	decimal(8,2)	از ستون جدول حاصل از جویین orders با order_games بر روی order_id بدست می‌آید
cheapest_game_price	decimal(8,2)	از اعمال تابع min بر روی ستون game_price جدول order_games با گروه بندی order_id بدست می‌آید
expensive_game_price	decimal(8,2)	از اعمال تابع max بر روی ستون game_price جدول order_games با گروه بندی order_id بدست می‌آید

از اعمال تابع count بر روی ستون game_id جدول order_games با گروه بندی order_id بدست می آید	int4	games_count
-----------------------------------------------------------------------------------------------	------	-------------

## جدول فکت دوره ای ماهانه

users_dim_scd3		fact_orders_periodic_monthly		date_dim	
user_id ↻	int4	user_id	int4	time_key ↻	int4
		time_key	int4		
		cheapest_game_price	decimal(8,2)		
		expensive_game_price	decimal(8,2)		
		average_game_price	decimal(10,2)		
		total_game_price	decimal(12,2)		
		number_of_orders	int4		
		total_games_bought	int4		

### fact\_orders\_periodic\_monthly

این فکت به صورت ماهانه مژرها را برای هر کاربر نشان می‌دهد. این مژرها عبارت اند از:

**ارزانترین بازی:** ارزانترین بازی خریداری شده توسط کاربر در این ماه

**گرانترین بازی:** گرانترین بازی خریداری شده توسط کاربر در این ماه

**میانگین قیمت بازیها:** میانگین قیمت بازیهایی که این کاربر در این ماه خریده است

**مجموع قیمت بازیها:** مجموع قیمت بازیهایی که این کاربر در این ماه خریده است

**تعداد سفارشات:** تعداد کل سفارشات کاربر در این ماه

**تعداد بازیها:** تعداد کل بازیهای خریده شده توسط این کاربر در این ماه

این جدول با استفاده از یک loop بر روی زمان به صورت ماهانه پر می‌شود. کلیدهای زمانی این فکت اولین روز هر ماه را نشان می‌دهد، به طور مثال 20230101 نشان دهنده ماه اول سال ۲۰۲۳ است

نام فیلد	نوع فیلد	شرح
user_id	int4	از ستون user_id جدول orders بدست می‌آید
time_key	int4	از اعمال تابع to_char(var, 'YYYYMMDD')::int روی متغیر loop بدست می‌آید
cheapest_game_price	decimal(8,2)	از اعمال تابع min بر روی ستون game_price جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید
expensive_game_price	decimal(8,2)	از اعمال تابع max بر روی ستون game_price جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید
average_game_price	decimal(10,2)	از اعمال تابع avg بر روی ستون game_price جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید
total_game_price	decimal(12,2)	از اعمال تابع sum بر روی ستون game_price جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید

از اعمال تابع count بر روی ستون order_id جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید	int4	number_of_orders
از اعمال تابع count بر روی ستون game_id جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید	int4	games_count

## جدول فکت تجمیعی

fact_orders_acc			users_dim_scd3	
user_id	int4	➤	user_id	int4
cheapest_game_price	decimal(8,2)			
expensive_game_price	decimal(8,2)			
average_game_price	decimal(10,2)			
total_game_price	decimal(12,2)			
number_of_orders	int4			
total_games_bought	int4			

### fact\_orders\_acc

این فکت به صورت تجمیعی مژرها را برای هر کاربر نشان می‌دهد. این مژرها عبارت اند از:

**ارزانترین بازی:** ارزانترین بازی خریداری شده توسط کاربر

**گرانترین بازی:** گرانترین بازی خریداری شده توسط کاربر

**میانگین قیمت بازیها:** میانگین قیمت بازیهایی که این کاربر خریده است

**مجموع قیمت بازیها:** مجموع قیمت بازیهایی که این کاربر خریده است

**تعداد سفارشات:** تعداد کل سفارشات کاربر

**تعداد بازیها:** تعداد کل بازیهای خریده شده توسط این کاربر

نام فیلد	نوع فیلد	شرح
user_id	int4	از ستون user_id جدول orders بدست می‌آید
cheapest_game_price	decimal(8,2)	از اعمال تابع min بر روی ستون game_price جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید
expensive_game_price	decimal(8,2)	از اعمال تابع max بر روی ستون game_price جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید
average_game_price	decimal(10,2)	از اعمال تابع avg بر روی ستون game_price جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید
total_game_price	decimal(12,2)	از اعمال تابع sum بر روی ستون game_price جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید

از اعمال تابع count بر روی ستون order_id جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید	int4	number_of_orders
از اعمال تابع count بر روی ستون game_id جدول حاصل از جویین order_games با orders بر روی order_id با گروه بندی user_id بدست می‌آید	int4	games_count