

به نام خدا

**ETL Document**

پروژه پایگاه داده 2

انبار داده دانشگاه

با تمرکز روی مارت آموزشی دانشجو ها

مرضیه علیدادی

**9631983**

استاد بصیری

## بعد ها

• dimDate :

universitydw dimdate	
TimeKey	: int(11)
FullDateAlternateKey	: varchar(60)
PersianFullDateAlternateKey	: varchar(60)
DayNumberOfWeek	: int(11)
PersianDayNumberOfWeek	: int(11)
EnglishDayNameOfWeek	: varchar(20)
PersianDayNameOfWeek	: varchar(20)
DayNumberOfMonth	: int(11)
PersianDayNumberOfMonth	: int(11)
DayNumberOfYear	: int(11)
PersianDayNumberOfYear	: int(11)
WeekNumberOfYear	: int(11)
PersianWeekNumberOfYear	: int(11)
EnglishMonthName	: varchar(60)
PersianMonthName	: varchar(60)
MonthNumberOfYear	: int(11)
PersianMonthNumberOfYear	: int(11)
CalendarQuarter	: int(11)
PersianCalendarQuarter	: int(11)
CalendarYear	: int(11)
PersianCalendarYear	: int(11)
CalendarSemester	: int(11)
PersianCalendarSemester	: int(11)

✓ سطر های این جدول، از فایل Date.txt بدست آمده است.

• dimStudent :

universitydw dimstudent	
student_code	: int(11)
ID	: varchar(15)
first_name	: varchar(30)
last_name	: varchar(30)
dept_name	: varchar(30)
national_code	: varchar(20)
previous_phone_number	: varchar(25)
phone_number_change_date	: date
current_phone_number	: varchar(25)
tot_cred	: decimal(13,10)
student_type	: bit(1)
student_type_description	: varchar(250)
student_degree	: smallint(6)
student_degree_description	: varchar(250)

- ✓ student\_code : یک surrogate key است. به این منظور از این نوع کلید استفاده شده است، که یک پارامتر از نوع int کلید اصلی جدول باشد.
- ✓ ID : از روی ستون ID جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ first\_name : از روی ستون first\_name جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ last\_name : از روی ستون last\_name جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ dept\_name : از روی ستون dept\_name جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ national\_code : از روی ستون national\_code جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ previous\_phone\_number : این ستون، شماره تلفن قبلی دانشجو را نشان می دهد. برای لحاظ کردن SCD3 روی ستون شماره تلفن، در نظر گرفته شده است. با کمک ستون phone\_number جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ phone\_number\_change\_date : این ستون، تاریخ تغییر شماره ی دانشجو را نشان می دهد. که با استفاده از پراسیجر های dimStudent\_firstLoad و dimStudent\_NextLoad درج می شود. برای لحاظ کردن SCD3 روی ستون شماره تلفن، در نظر گرفته شده است.
- ✓ current\_phone\_number : این ستون، شماره تلفن فعلی دانشجو را نشان می دهد. برای لحاظ کردن SCD3 روی ستون شماره تلفن، در نظر گرفته شده است. از روی ستون phone\_number جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ tot\_cred : از روی ستون tot\_cred جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ student\_type : از روی ستون student\_type جدول student موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می آید.
- ✓ student\_type\_description : از روی ستون student\_type جدول student موجود

در StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می‌آید. و توضیح بیشتری درباره ی student\_type می‌دهد.

✓ student\_degree : از روی ستون student\_degree جدول student موجود در

StageArea که حاوی اطلاعات جدول student موجود در Source است، بدست می‌آید.

✓ student\_degree\_description : از روی ستون student\_degree جدول student

موجود در StageArea که حاوی اطلاعات جدول student موجود در Source است،

بدست می‌آید. و توضیح بیشتری درباره ی student\_degree می‌دهد.

• dimInstructor :

universitydw diminstructor	
🔑	instructor_code : int(11)
🔑	ID : varchar(15)
🔑	first_name : varchar(30)
🔑	last_name : varchar(30)
🔑	dept_name : varchar(30)
📅	salary_start_date : date
📅	salary_end_date : date
#	salary_change_flag : int(11)
#	salary : decimal(18,12)
🔑	national_code : varchar(20)
🔑	previous_phone_number : varchar(25)
📅	phone_number_change_date : date
🔑	current_phone_number : varchar(25)
#	instructor_type : smallint(6)
🔑	instructor_type_description : varchar(250)

✓ instructor\_code : یک surrogate key است. با توجه به اینکه ستون salary از نوع SCD2 است، این نوع کلید در نظر گرفته شده است.

✓ ID : از روی ستون ID جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می‌آید.

✓ first\_name : از روی ستون first\_name جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می‌آید.

✓ last\_name : از روی ستون last\_name جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می‌آید.

✓ dept\_name : از روی ستون dept\_name جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می‌آید.

✓ salary\_start\_date : این ستون، نشان می دهد که از چه تاریخی، مقدار موجود در ستون Salary همین سطر، درج شده است و از چه تاریخی حقوق استاد به این مقدار تغییر پیدا کرده است. که با استفاده از پراسیجر های dimInstructor\_firstLoad و dimInstructor\_NextLoad درج می شود. برای لحاظ کردن SCD2 روی ستون salary، در نظر گرفته شده است.

✓ salary\_end\_date : این ستون، نشان می دهد که تا چه تاریخی، مقدار موجود در ستون Salary همین سطر، درج شده بوده است و تا چه تاریخی حقوق استاد این مقدار بوده است. که با استفاده از پراسیجر های dimInstructor\_firstLoad و dimInstructor\_NextLoad درج می شود. برای لحاظ کردن SCD2 روی ستون salary، در نظر گرفته شده است.

✓ salary\_change\_flag : این ستون، دارای محتوای 0 یا 1 است. نشان می دهد که آیا مقدار موجود در ستون Salary همین سطر، در حال حاضر معتبر است یا خیر. که با استفاده از پراسیجر های dimInstructor\_firstLoad و dimInstructor\_NextLoad درج می شود. برای لحاظ کردن SCD2 روی ستون salary، در نظر گرفته شده است.

✓ salary : از روی ستون salary جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می آید. که دارای SCD2 است.

✓ national\_code : از روی ستون national\_code جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می آید.

✓ previous\_phone\_number : این ستون، شماره تلفن قبلی استاد را نشان می دهد. برای لحاظ کردن SCD3 روی ستون شماره تلفن، در نظر گرفته شده است. با کمک ستون phone\_number جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می آید.

✓ phone\_number\_change\_date : این ستون، تاریخ تغییر شماره ی استاد را نشان می دهد. که با استفاده از پراسیجر های dimInstructor\_firstLoad و dimInstructor\_NextLoad درج می شود. برای لحاظ کردن SCD3 روی ستون شماره تلفن، در نظر گرفته شده است.

✓ current\_phone\_number : این ستون، شماره تلفن فعلی استاد را نشان می دهد.

برای لحاظ کردن SCD3 روی ستون شماره تلفن، در نظر گرفته شده است. از روی ستون phone\_number جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می‌آید.

✓ instructor\_type : از روی ستون instructor\_type جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می‌آید.

✓ instructor\_type\_description : از روی ستون instructor\_type\_description جدول instructor موجود در StageArea که حاوی اطلاعات جدول instructor موجود در Source است، بدست می‌آید. و توضیح بیشتری درباره ی instructor\_type می‌دهد.

• dimCourse :

universitydw dimcourse	
🔑	course_code : int(11)
📄	course_id : varchar(18)
📄	title : varchar(60)
📄	dept_name : varchar(30)
📅	credits_start_date : date
📅	credits_end_date : date
#	credits_change_flag : int(11)
#	credits : decimal(12,10)

✓ course\_code : یک surrogate key است. با توجه به اینکه ستون credits از نوع SCD2 است، این نوع کلید در نظر گرفته شده است.

✓ course\_id : از روی ستون course\_id جدول course موجود در StageArea که حاوی اطلاعات جدول course موجود در Source است، بدست می‌آید.

✓ title : از روی ستون title جدول course موجود در StageArea که حاوی اطلاعات جدول course موجود در Source است، بدست می‌آید.

✓ dept\_name : از روی ستون dept\_name جدول course موجود در StageArea که حاوی اطلاعات جدول course موجود در Source است، بدست می‌آید.

✓ credits\_start\_date : این ستون، نشان می‌دهد که از چه تاریخی، مقدار موجود در ستون credits همین سطر، درج شده است و از چه تاریخی تعداد واحد درس به این مقدار تغییر پیدا کرده است. که با استفاده از پراسیجر های dimCourse\_firstLoad و

dimCourse\_NextLoad درج می شود. برای لحاظ کردن SCD2 روی ستون credits، در نظر گرفته شده است.

✓ credits\_end\_date : این ستون، نشان می دهد که تا چه تاریخی، مقدار موجود در ستون credits همین سطر، درج شده بوده است و تا چه تاریخی تعداد واحد درس این مقدار بوده است. که با استفاده از پراسیجر های dimCourse\_firstLoad و dimCourse\_NextLoad درج می شود. برای لحاظ کردن SCD2 روی ستون dimCourse، در نظر گرفته شده است.

✓ credits\_change\_flag : این ستون، دارای محتوای 0 یا 1 است. نشان می دهد که آیا مقدار موجود در ستون credits همین سطر، در حال حاضر معتبر است یا خیر. که با استفاده از پراسیجر های dimCourse\_firstLoad و dimCourse\_NextLoad درج می شود. برای لحاظ کردن SCD2 روی ستون credits، در نظر گرفته شده است.

✓ credits : از روی ستون credits جدول course موجود در StageArea که حاوی اطلاعات جدول course موجود در Source است، بدست می آید. که دارای SCD2 است.

• dimClass :

universitydw dimclass	
class_code	: int(11)
building	: varchar(25)
room_number	: varchar(17)
capacity	: decimal(4,0)

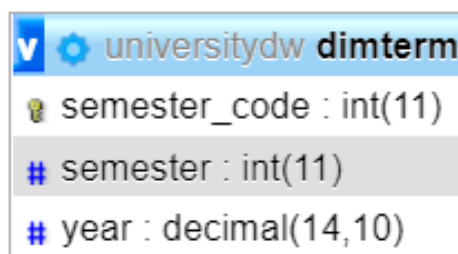
✓ class\_code : یک surrogate key است. به این منظور از این نوع کلید استفاده شده است، که یک پارامتر از نوع int کلید اصلی جدول باشد.

✓ building : از روی ستون building جدول classroom موجود در StageArea که حاوی اطلاعات جدول classroom موجود در Source است، بدست می آید.

✓ room\_number : از روی ستون room\_number جدول classroom موجود در StageArea که حاوی اطلاعات جدول classroom موجود در Source است، بدست می آید.

✓ capacity : از روی ستون capacity جدول classroom موجود در StageArea که حاوی اطلاعات جدول classroom موجود در Source است، بدست می آید.

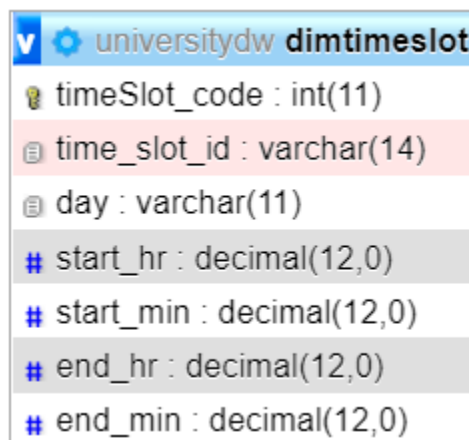
• dimTerm :



universitydw dimterm	
🔑	semester_code : int(11)
#	semester : int(11)
#	year : decimal(14,10)

- ✓ semester\_code : یک surrogate key است. به این منظور از این نوع کلید استفاده شده است، که یک پارامتر از نوع int کلید اصلی جدول باشد.
- ✓ Semester : از روی ستون semester جدول section موجود در StageArea که حاوی اطلاعات جدول section موجود در Source است، بدست می آید.
- ✓ year : از روی ستون year جدول section موجود در StageArea که حاوی اطلاعات جدول section موجود در Source است، بدست می آید.

• dimTimeSlot :



universitydw dimtimeslot	
🔑	timeSlot_code : int(11)
#	time_slot_id : varchar(14)
#	day : varchar(11)
#	start_hr : decimal(12,0)
#	start_min : decimal(12,0)
#	end_hr : decimal(12,0)
#	end_min : decimal(12,0)

- ✓ timeSlot\_code : یک surrogate key است. به این منظور از این نوع کلید استفاده شده است، که یک پارامتر از نوع int کلید اصلی جدول باشد.
- ✓ time\_slot\_id : از روی ستون time\_slot\_id جدول time\_slot موجود در StageArea که حاوی اطلاعات جدول time\_slot موجود در Source است، بدست می آید.
- ✓ day : از روی ستون day جدول time\_slot موجود در StageArea که حاوی اطلاعات جدول time\_slot موجود در Source است، بدست می آید.
- ✓ start\_hr : از روی ستون start\_hr جدول time\_slot موجود در StageArea که حاوی اطلاعات جدول time\_slot موجود در Source است، بدست می آید.



- ✓ start\_min : از روی ستون start\_min جدول time\_slot موجود در StageArea که حاوی اطلاعات جدول time\_slot موجود در Source است، بدست می‌آید.
- ✓ end\_hr : از روی ستون end\_hr جدول time\_slot موجود در StageArea که حاوی اطلاعات جدول time\_slot موجود در Source است، بدست می‌آید.
- ✓ end\_min : از روی ستون end\_min جدول time\_slot موجود در StageArea که حاوی اطلاعات جدول time\_slot موجود در Source است، بدست می‌آید.



• factlessPrereq :

universitydw factlessprereq
course_id : varchar(18)
prereq_id : varchar(18)

- این جدول ارتباط بین دایمنشن dimCouse با خودش را نشان می‌دهد. به این منظور در نظر گرفته شده است، که پیش نیاز هر درس را مشخص کند.
- ✓ course\_id : کلید اصلی دایمنشن dimCouse است.
  - ✓ prereq\_id : کلید اصلی دایمنشن dimCouse است.

• factlessAvailableCorses :

universitydw factlessavailablecorses
# TimeKey : int(11)
# instructor_code : int(11)
# instructor_ID : varchar(15)
# course_code : int(11)
# course_id : varchar(18)
# class_code : int(11)
# semester_code : int(11)

- این جدول، دروس ارائه شده را نمایش می‌دهد.
- ✓ TimeKey : کلید اصلی دایمنشن dimDate است.
  - ✓ instructor\_code : کلید اصلی (Surrogate key) دایمنشن dimInstructor است.

- ✓ instructor\_ID : natural key دایمنشن dimInstructor است.
- ✓ course\_code : کلید اصلی (Surrogate key) دایمنشن dimCourse است.
- ✓ course\_id : natural key دایمنشن dimCourse است.
- ✓ class\_code : کلید اصلی (Surrogate key) دایمنشن dimClass است.
- ✓ semester\_code : کلید اصلی دایمنشن dimTerm است.

• factTransactionTakeCorse :

universitydw facttransactiontakecorse	
#	TimeKey : int(11)
#	student_code : int(11)
@	student_ID : varchar(15)
#	instructor_code : int(11)
@	instructor_ID : varchar(15)
#	course_code : int(11)
@	course_id : varchar(18)
#	class_code : int(11)
#	semester_code : int(11)
#	timeSlot_code : int(11)
@	time_slot_id : varchar(14)
@	grade : varchar(12)
@	grade_description : varchar(30)

این جدول، دروسی که توسط هر دانشجو اخذ شده را نمایش می‌دهد. همچنین نمره ی دانشجو در هر درس اخذ شده و وضعیت نمره ی او را نمایش می‌دهد.

- ✓ TimeKey : کلید اصلی دایمنشن dimDate است.
- ✓ student\_code : کلید اصلی (Surrogate key) دایمنشن dimStudent است.
- ✓ student\_ID : natural key دایمنشن dimStudent است.
- ✓ instructor\_code : کلید اصلی دایمنشن dimInstructor است.
- ✓ instructor\_ID : natural key دایمنشن dimInstructor است.
- ✓ course\_code : کلید اصلی (Surrogate key) دایمنشن dimCourse است.
- ✓ course\_id : natural key دایمنشن dimCourse است.
- ✓ class\_code : کلید اصلی (Surrogate key) دایمنشن dimClass است.
- ✓ semester\_code : کلید اصلی دایمنشن dimTerm است.
- ✓ time\_slot\_code : کلید اصلی دایمنشن dimTimeSlot است.
- ✓ time\_slot\_ID : natural key دایمنشن dimTimeSlot است.

✓ grade : از روی ستون grade جدول takes موجود در StageArea که حاوی اطلاعات

جدول takes موجود در Source است، بدست می‌آید.

✓ grade\_description : از روی ستون grade جدول takes موجود در StageArea که حاوی اطلاعات جدول takes موجود در Source است، بدست می‌آید. و توضیح بیشتری درباره ی grade و وضعیت قبولی دانشجو می‌دهد.

• factTermicReportCard :

universitydw facttermicreportcard	
#	TimeKey : int(11)
#	student_code : int(11)
#	student_ID : varchar(15)
#	semester_code : int(11)
#	average : varchar(12)
#	max_grade : varchar(12)
#	min_grade : varchar(12)
#	probation : bit(1)
#	total_credits : decimal(13,10)
#	total_passed_credits : decimal(13,10)
#	total_failed_credits : decimal(13,10)

این جدول، کارنامه ی هر ترم دانشجو را نشان می‌دهد. به صورت ترمی و دوره ای اطلاعات دروس اخذ شده ی دانشجو در آن ترم خاص را ثبت می‌کند.

✓ TimeKey : کلید اصلی دایمنشن dimDate است.

✓ student\_code : کلید اصلی (Surrogate key) دایمنشن dimStudent است.

✓ student\_ID : natural key دایمنشن dimStudent است.

✓ semester\_code : کلید اصلی دایمنشن dimTerm است.

✓ average : میانگین نمرات دروس دانشجو در آن ترم را نشان می‌دهد. با AVG گرفتن

روی ستون grade مربوط به جدول takes (در جدول حاصل از جویین جداول takes,

student, course, section، گروپ بای student\_ID, student\_code, TimeKey,

semester\_code)، بدست می‌آید.

✓ max\_grade : بیشترین نمره ی دانشجو در دروس دانشجو در آن ترم را نشان می‌دهد. با

MAX گرفتن روی ستون grade مربوط به جدول takes (در جدول حاصل از جویین

جداول student, course, section, takes، گروپ بای student\_ID, TimeKey,

student\_code)، بدست می‌آید.

✓ min\_grade : کمترین نمره ی دانشجو در دروس دانشجو در آن ترم را نشان می‌دهد. با

MIN گرفتن روی ستون grade مربوط به جدول takes (در جدول حاصل از جوین جداول takes, student, course, section، گروپ بای student\_code بای TimeKey, semester\_code)، بدست می‌آید.

✓ probation : این ستون، دارای محتوای 0 یا 1 است. نشان می‌دهد که آیا دانشجو با توجه به معدل ترمش، مشروط شده است یا خیر. با بررسی ستون average، و مقایسه‌ی مقدار آن با 12 تعیین می‌شود.

✓ total\_credits : مجموع واحدهای اخذ شده ی دانشجو در آن ترم را نشان می‌دهد. با SUM گرفتن روی ستون credits مربوط به جدول course (در جدول حاصل از جوین جداول takes, student, course, section، گروپ بای student\_code بای TimeKey, semester\_code)، بدست می‌آید.

✓ total\_passed\_credits : مجموع واحدهای قبول شده ی دانشجو در آن ترم را نشان می‌دهد. با SUM گرفتن روی ستون credits مربوط به جدول course، به شرط اینکه ستون grade موجود در جدول takes بیشتر از 10 باشد (در جدول حاصل از جوین جداول takes, student, course, section، گروپ بای student\_code بای TimeKey, semester\_code)، بدست می‌آید.

✓ total\_failed\_credits : مجموع واحدهای مردود شده ی دانشجو در آن ترم را نشان می‌دهد. با SUM گرفتن روی ستون credits مربوط به جدول course، به شرط اینکه ستون grade موجود در جدول takes کمتر از 10 باشد (در جدول حاصل از جوین جداول takes, student, course, section، گروپ بای student\_ID, student\_code بای TimeKey, semester\_code)، بدست می‌آید.

#### • factAccumulativeReportCard :

universitydw factaccumulativereportcard
# student_code : int(11)
# student_ID : varchar(15)
# average : varchar(12)
# max_grade : varchar(12)
# min_grade : varchar(12)
# probation_num : decimal(13,10)
# total_credits : decimal(13,10)
# total_passed_credits : decimal(13,10)
# total_failed_credits : decimal(13,10)

این جدول، کارنامه ی دانشجو، به صورت تجمیعی از ابتدا تا حال را نشان می‌دهد.

- ✓ student\_code : کلید اصلی (Surrogate key) دایمنشن dimStudent است.
- ✓ student\_ID : natural key دایمنشن dimStudent است.
- ✓ average : میانگین نمرات دروس دانشجو از ابتدا تا حال را نشان می‌دهد. با AVG گرفتن روی ستون grade مربوط به جدول takes (در جدول حاصل از جویین جداول student, course, section, گروه بای student\_code, student\_ID)، بدست می‌آید.
- ✓ max\_grade : بیشترین نمره ی دانشجو در دروس دانشجو از ابتدا تا حال را نشان می‌دهد. با MAX گرفتن روی ستون grade مربوط به جدول takes (در جدول حاصل از جویین جداول student, course, section, گروه بای student\_code, student\_ID)، بدست می‌آید.
- ✓ min\_grade : کمترین نمره ی دانشجو در دروس دانشجو از ابتدا تا حال را نشان می‌دهد. با MIN گرفتن روی ستون grade مربوط به جدول takes (در جدول حاصل از جویین جداول student, course, section, گروه بای student\_code, student\_ID)، بدست می‌آید.
- ✓ probation\_num : این ستون، مجموع مشروطی های دانشجو در کل تحصیلاتش را نشان می‌دهد. با SUM گرفتن روی probation مربوط به فکت قبلی (فکت اسنپشات ترمی)، هر جا مقدارش 1 باشد، یعنی مشروط بوره و یکی به این شمارش اضافه می‌شود.
- ✓ total\_credits : مجموع واحدهای اخذ شده ی دانشجو از ابتدا تا حال را نشان می‌دهد. با SUM گرفتن روی ستون credits مربوط به جدول course (در جدول حاصل از جویین جداول student, course, section, گروه بای student\_code, student\_ID)، بدست می‌آید.
- ✓ total\_passed\_credits : مجموع واحدهای قبول شده ی دانشجو از ابتدا تا حال را نشان می‌دهد. با SUM گرفتن روی ستون credits مربوط به جدول course، به شرط اینکه ستون grade موجود در جدول takes بیشتر از 10 باشد (در جدول حاصل از جویین جداول student, course, section, گروه بای student\_code, student\_ID)، بدست می‌آید.
- ✓ total\_failed\_credits : مجموع واحدهای مردود شده ی دانشجو از ابتدا تا حال را نشان می‌دهد. با SUM گرفتن روی ستون credits مربوط به جدول course، به شرط اینکه ستون grade موجود در جدول takes کمتر از 10 باشد (در جدول حاصل از جویین جداول

student\_code, student\_ID بای takes, student, course, section بدست می آید.