

معایب شکل دوم نرمال

ابتدا اجازه دهید که مثالی از یک جدول بیاورم که به شکل دوم نرمال بوده ولی به شکل سوم نرمال نباشد. برای این منظور دو جدول زیر که هر دو در شکل سوم نرمال به سر می‌برند را با هم ترکیب می‌کنیم. ستون هایی از جدول که با نقاط قرمز رنگ علامت گذاری شده اند کلیدهای اولیه جدول می‌باشند.

| کد دانشجو | نام دانشجو |
|-----------|------------|
| | |

| نام رشته | نوع رشته | تعداد کل واحدها |
|----------|----------|-----------------|
| | | |

اگر این دو جدول را با هم ترکیب کنیم، جدولی حاصل می‌شود که به فرم دوم نرمال است یعنی تمام خصیصه‌های غیر کلیدی وابسته به کلید اولیه (کد دانشجو) می‌باشند. اما همانطور که در بخش بعدی گفته خواهد شد، به شکل سوم نرمال نمی‌باشد.

| کد دانشجو | نام دانشجو | نام رشته | نوع رشته |
|-----------|------------|----------|----------|
| دانشجو ۱ | نام ۱ | رشته ۱ | نوع ۱ |
| دانشجو ۲ | نام ۲ | رشته ۱ | نوع ۱ |
| دانشجو ۳ | نام ۳ | رشته ۱ | نوع ۱ |
| دانشجو ۴ | نام ۴ | رشته ۲ | نوع ۳ |
| دانشجو ۵ | نام ۵ | رشته ۲ | نوع ۳ |
| دانشجو ۶ | نام ۶ | رشته ۳ | نوع ۲ |

خصیصه "نوع رشته" به کلید اولیه جدول وابستگی تابعی دارد ولی از نوع متعددی (یعنی وابستگی از طریق خصیصه نام دانشجو می‌تواند بدست باید، چرا که نوع رشته به نام رشته و نام رشته به نام دانشجو وابستگی تابعی دارد)، این موضوع علاوه بر افزونگی اطلاعات باعث بی نظمی در به هنگام سازی خواهد شد. بطور نمونه

ایراد در عمل insert: این واقعیت که یک رشته خاص دارای یک نوع رشته خاص است را نمی‌توان اضافه کنیم، مثلاً نمی‌توانیم بیان کنیم که رشته ریاضی از نوع علوم پایه است مگر آن که دانشجویی باشد در رشته ریاضی مشغول به تحصیل است.

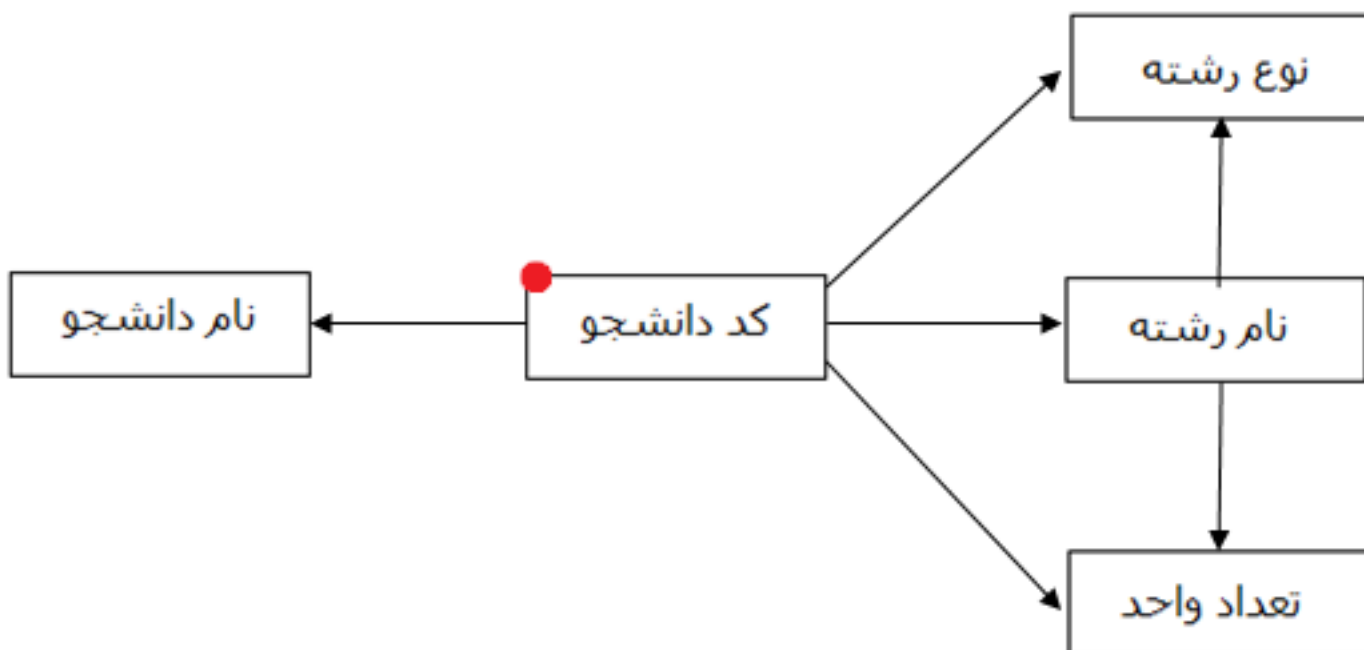
ایراد در عمل delete: با حذف یک دانشجو نه تنها اطلاعات مربوط به دانشجو بلکه اطلاعات مربوط به رشته تحصیلی نیز ممکن است حذف شود. مثلاً با حذف سطر مربوط به دانشجوی شماره 6 تمام اطلاعات مربوط به رشته شماره 3 نیز حذف خواهد شد.

ایراد در عمل update: اگر فرضاً بخواهیم نوع رشته ای به نام رشته 1 را تغییر دهیم به جای یک سطر باید چندین سطر (سه سطر در داده‌های نمونه) را بروز رسانی کنیم.

تعریف شکل نرمال سوم

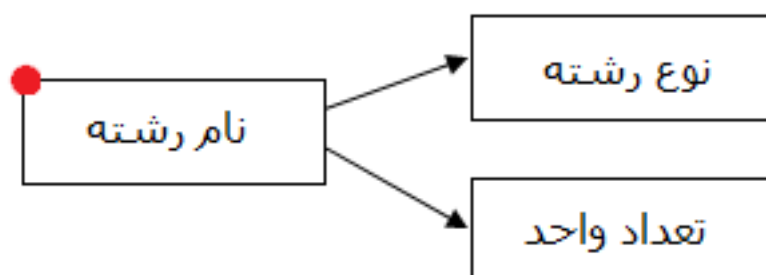
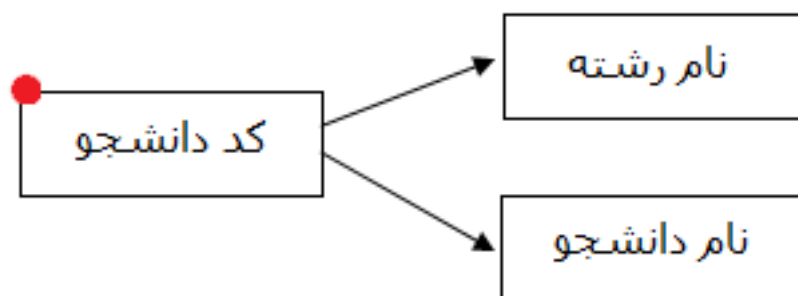
یک متغیر رابطه ای به شکل سوم نرمال است اگر به شکل دوم نرمال بوده و وابستگی‌های با واسطه (متعدی) نداشته باشد.

بر می‌گردیم به جدول ترکیبی قبل، نمودار FD جدول مورد نظر به صورت زیر است:



در این نمودار واضح است که وابستگی خصیصه نوع رشته به کد دانشجو از طریق خصیصه نام رشته بدست می‌آید. همینطور برای خصیصه "تعداد واحد". پس دو خصیصه‌ی نوع رشته و تعداد واحد با واسطه به کد دانشجو مرتبط هستند.

پس با تجزیه این نمودار به صورت زیر شرط شکل سوم نرمال هم محقق خواهد شد:



کافیه خصیصه کلید اولیه جدول "رشته ها" را به جدول "دانشجو" اضافه کنیم تا هر دو جدول به شکل نرمال سوم در بیایند. نقطه قرمز به معنای کلید اولیه و نقطه آبی به معنای کلید خارجی می باشد:

| کد دانشجو | نام دانشجو | نام رشته |
|-----------|------------|----------|
| | | |

| نام رشته | نوع رشته | تعداد واحد |
|----------|----------|------------|
| | | |

موفق باشید

نظرات خوانندگان

نویسنده: سعید
تاریخ: ۱۳۹۱/۱۱/۱۴ ۱:۷

خوب، اگر این سه قسمت رو بخوایم با EF Code first مدل کنیم:

```
using System.Collections.Generic;
using System.ComponentModel.DataAnnotations;

public class Student
{
    public int Id { set; get; }
    public string Name { set; get; }

    //هر دانشجو چند ترم در دانشگاه خواهد بود
    public virtual ICollection<Semester> Semesters { set; get; }
    //هر دانشجو چندین واحد دارد
    public virtual ICollection<Unit> Units { set; get; }
}

public class Semester
{
    public int Id { set; get; }
    public string Name { set; get; }
    public int Average { set; get; }

    [ForeignKey("StudentId")]
    public virtual Student Student { set; get; }
    public int StudentId { set; get; }
}

public class Unit
{
    public int Id { set; get; }
    public string Name { set; get; }
    public string UnitType { set; get; }
    public int NumberOfUnits { set; get; }

    [ForeignKey("StudentId")]
    public virtual Student Student { set; get; }
    public int StudentId { set; get; }
}
```

به نظر می‌رسد که خاصیت Average جاش در کلاس Semester نیست. حتی به Unit هم نباید به صورت مستقیم ارتباط پیدا کنه. نیاز به یک کلاس دیگر هست که بتونه به ازای هر دانشجو، ترم و واحد، نمره ثبت کرد. میانگین، یک خاصیت آماری است که می‌تونه اصلاً لحاظ نشه و در گزارشات محاسبه بشه. و یا هر ترم یک سری واحد داره. اینطوری چطور؟ چون الان مشخص نیست در هر ترم چه واحدهایی برداشته.

نویسنده: محمد سلم ابادی
تاریخ: ۱۳۹۱/۱۱/۱۴ ۸:۵۴

موضوعی که شما مطرح می‌کنید خارج از بحث مطرح شده است. من تصمیم نداشتم که یک محیط عملیاتی را پیاده سازی کنم. تنها مثال هایی برای درک بهتر موضوع آوردم. بله میانگین یک خاصیت آماری است، اما ما می‌توانیم برای سرعت بخشیدن به query هایمان برای بدست آوردن معدل، آن را بصورت فیزیکی ذخیره داشته باشیم. چون معدل بعد از ثابت و تعیین شدن دیگر تغییر نخواهد کرد.

نویسنده: سعید
تاریخ: ۱۳۹۱/۱۱/۱۴ ۹:۱۰

مشکلی که بودن میانگین در کلاس ترم ایجاد می‌کنه وابسته کردن آن به دانشجو است درحالیکه ترم باید یک موجودیت واحد و مستقل باشد. با این طراحی فعلی باید کل اطلاعات ثابت یک ترم به ازای هر دانشجو یکبار دیگر هم ثبت شود.

یک نکته‌ای رو چند وقت قبل حین کار با ef بهش برخوردم که جالب بود. از دید ef ، کلید خارجی یا کلید اصلی «فقط خواندنی» هستند. یعنی اگر در اینجا کسی بخواد نام رشته رو تغییر بده مشکل ساز خواهد شد.

به دو دلیل:

- استفاده از رشته‌ها نسبت به یک عدد چون طولانی‌تر هستند کندتر است برای حالت تعریف کلید

- اگر تغییری قرار است رخ دهد، باید به تمام جداول اعمال شود. (تعریف اطلاعات تکراری)