بسمه تعالی

موضوع : مبانی آمار و احتمال برای ورود به علم داده و تحلیل داده

بخش پنجم:

- آمار توصیفی
- جدول توزیع فراوانی: فراوانی نسبی/فراوانی تجمعی
- نمونه جدول فراوانی

در این پارت از اموزش به بررسی **حوزه آمار توصیفی و جدول فراوانی به عنوان یک ابزار پرکاربرد در این حوزه** می پردازیم .

در جلسات گذشته با **مباحث اولیه آمار و احتمالات آشنا شدیم** .

اگر داده‌ها را فقط به صورت خام نگهداری کنیم، استخراج اطلاعات از آن‌ها بسیار دشوار خواهد بود. بنابراین اولین قدم بعد از جمع‌آوری داده‌ها، خلاصه‌سازی و توصیف آن‌ها است. تا بتوانیم رفتار داده ها و جامعه آماری را بدست بیاریم به این حوزه آمار توصیفی می گویند.

کاربرد اصلی آمار توصیفی :

* **خلاصه‌سازی داده‌ها** : مانند محاسبه میانگین، میانه، نما و انحراف معیار.
* **نمایش داده‌ها** : از طریق جدول‌ها و نمودارها.
* **توصیف ویژگی‌های داده‌ها:** بدون اینکه بخواهیم درباره کل جامعه نتیجه‌گیری کنیم.

نکته مهم این است که آمار توصیفی فقط داده‌های موجود را توصیف می‌کند و وارد پیش‌بینی یا استنباط نمی‌شود.

برای کارکردن در این حوزه اولین و پرکاربردترین ابزار جدول فراوانی است .

جدول توزیع فراوانی چیست؟

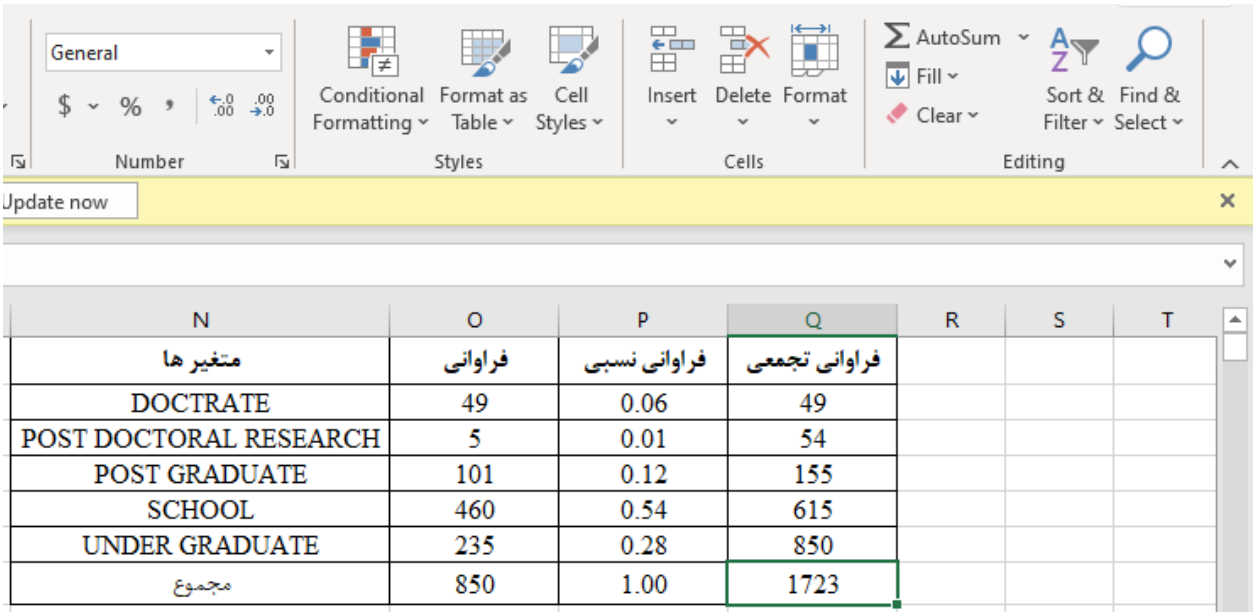
جدول توزیع فراوانی یا Frequency Distribution Table جدولی است که نشان می‌دهد هر مقدار یا هر گروه از داده‌ها، چند بار تکرار شده است.

به زبان ساده: فراوانی یعنی تعداد دفعات مشاهده یک مقدار.

انواع فراوانی : نسبی/ فراوانی تجمعی و ترکیبی از اینها

انواع جدول فراوانی : جدول فراوانی تک متغیر / جدول فراوانی چند متغیر بهش می گن cross table

نمونه هایی از جدول فراوانی:



متغیر ها	فراوانی	فراوانی نسبی	فراوانی تجمعی
DOCTRATE	49	0.06	49
POST DOCTORAL RESEARCH	5	0.01	54
POST GRADUATE	101	0.12	155
SCHOOL	460	0.54	615
UNDER GRADUATE	235	0.28	850
مجموع	850	1.00	1723