

راهنمای آماده‌سازی پروژه‌ی پایان‌ترم

مبانی احتمال – ترم اول سال تحصیلی ۹۶-۱۳۹۵

مثال: تعداد زیادی از بازی‌های رایانه‌ایِ فروشگاه‌های هنوز به فروش نرفته است. فروشنده علاقه‌مند به بررسی توزیع میزان درآمد حاصل از فروش این بازی‌ها بر حسب عواملی همچون زمان فروش، قیمت بازی‌ها و رده‌ی سنی شده بر روی بازی‌ها است تا بتواند در سفارش‌های بعدی خود رویه‌ای اتخاذ کند که با بیشترین سود حداقل مرجوعی را داشته باشد.

الف. طرح سؤال پژوهشی

آیا میزان درآمد حاصل از فروش بازی‌های کامپیوتری به فروش رسیده در یک فروشگاه اینترنتی در ماه‌های مختلف سال متفاوت است؟

آیا تعداد بازی‌های کامپیوتری به فروش رسیده در این فروشگاه به قیمت بازی‌ها بستگی دارد؟

آیا تعداد بازی‌های کامپیوتری به فروش رسیده در این فروشگاه به رده‌ی سنی آن بازی بستگی دارد؟ به عبارتی، آیا بازی‌های مخصوص به یک رده‌ی سنی خاص فروش بیشتری داشته‌اند؟

ب. متغیرها

به کمک سؤالات پژوهشی درج شده در قسمت الف می‌توان دریافت که برای پاسخ به این سؤالات به متغیرهایی مانند نمونه‌های زیر نیاز داریم:

۱. یک متغیر پیوسته که قیمت‌های آیت‌ها را در بر داشته باشد.

۲. یک متغیر گسسته‌ی چند رده‌ای که هر رده‌اش معرف یکی از ماه‌های سال باشد؛ مثلاً متغیری که یکی از مقادیر یک تا دوازده را می‌پذیرد. اگر داده‌ها را به فرم جدول داده‌های اکسل تصور کنید، ستون مربوط به این متغیر به قسمی است که برای هر آیت‌می که به فروش رفته است یکی از عددهای یک تا دوازده درج شده که نشان می‌دهد این آیت‌م در کدامیک از ماه‌های سال به فروش رفته است.

۳. یک متغیر گسسته‌ی چند رده‌ای که رده‌ی سنی مخصوص خریداران بازی‌ها را مشخص کرده است؛ مثلاً متغیری سه‌رده‌ای که فقط سه عدد یک (زیر هفت سال)، دو (بین هفت تا هجده سال)، و سه (بالای هجده سال) را می‌پذیرد.

ج. تعداد داده‌ها (حجم نمونه) چقدر باشد؟

اهمیت چندانی ندارد؛ هر چه بیشتر بهتر؛ آن قدر هم کم نباشد که وقتی نمودارهایی همچون هیستوگرام (بافت‌نگار)شان را رسم می‌کنید، در هر رده از متغیرهای گسسته‌ی معرفی شده در قسمت ب، سه چهار عدد بیشتر نداشته باشیم.

ه. چه مباحثی از آمار توصیفی باید مورد بحث قرار بگیرند؟

این دیگر هنر شماسست که روش‌هایی را که آموخته‌اید کدام‌ها را و به چه نحوی به کار ببرید که شما را به جواب سؤال پژوهشی برساند. ضمناً، اگر مثلاً فرض می‌کنید باکس پلات (نمودار جعبه‌ای) به تنهایی معرف داده‌های شما و دربردارنده‌ی اطلاعات لازم است، بد نیست در کنارش

نمودارهای میله‌ای چندرده‌ای را هم امتحان کنید و نشان دهید که این نمودار هم مؤید یافته‌های حاصل از باکس پلات است. وقتی قصد دارید به این سؤال پژوهشی پاسخ دهید که با افزایش رده‌ی سنی چه تغییری در میزان فروش حاصل می‌شود، بد نیست از یک نمودار پراکنش هم استفاده کنید. در کنار نمودارها، جدول‌های شامل مقادیر آماره‌های توصیفی همچون میانگین، میانه و انحراف معیار را فراموش نکنید. هیچ میانگینی را بدون درج انحراف معیار داده‌ها در کنار آن رها نکنید. اگر توزیع شما چوله است، در تفسیر نتایج حتماً بنویسید که کدام یک از آماره‌های مرکزی بهتر از همه داده‌ها را توصیف می‌کند؛ یعنی، با استناد به مقادیر درج شده در جدول‌های عددی، بنویسید که اگر به جای میانه از میانگین استفاده می‌کردید ممکن بود چه اشتباهی پیش بیاید.

ی. از چه نرم‌افزاری استفاده کنیم؟

استفاده از هر نرم‌افزاری که نمودارها و خروجی‌های عددی مورد بحث را به صورت مطلوبی به دست دهد مجاز و بلا مانع است!

موفق باشید

صفوی‌منش