

پاسخنامه

پاسخنامه کونیز جلسه 07 دوره علم داده، یادگیری ماشین و هوش مصنوعی - فرزاد مینویی

✓ ۱* - طبق مدل CRSP-DM اولین گام در پروژه‌های داده چیست؟

✓ ۱ فهم مسئله

۲ فهم داده‌ها

۳ آماده‌سازی داده

۴ مدل‌سازی

✓ ۲* - برای آنکه نتایج تحلیل داده روی نمونه آماری قابل تعمیم به جامعه باشد، نمونه باید کدام ویژگی را داشته باشد؟

۱ تصادفی باشد.

۲ از منظر متغیرهای کلیدی و ویژگی‌های مورد مطالعه، به جامعه شباهت داشته باشد.

۳ اندازه نمونه برای استنباط از داده‌ها به حد کافی باشد.

✓ ۴ همه موارد

✓ ۳* - نخستین گام در فاز فهم داده‌ها کدام است؟



۱ درک فرآیند جمع‌آوری داده

۲ مستندسازی فراداده (Meta Data)

۳ تحلیل تک متغیره (Univariate Analysis)

۴ بررسی داده‌های گمشده

✓ ۴* - کدام گزینه درباره مستندسازی فراداده (Meta Data) نادرست است؟

۱ تعریف متغیرها شفاف می‌شود.

۲ نوع متغیرها (رسته‌ای یا عددی) مشخص می‌گردد.



۳ تحلیل‌های ساده آماری صورت می‌گیرد.

۴ تعریف رسته‌های مختلف متغیر رسته‌ای روشن می‌شود.

✓ ۵* - کدام گزینه درباره تحلیل تک‌متغیره (Uni-variate Analysis) نادرست است؟

۱ برای متغیرهای رسته‌ای جدول فراوانی وقوع محاسبه می‌شود.

۲ برای متغیرهای عددی، شاخص‌های مرکزی و پراکندگی مورد توجه قرار می‌گیرد.

۳ توزیع متغیرهای عددی در قالب نمودار هیستوگرام بررسی می‌شود.



۴ رابطه دو به دو متغیرها بررسی می‌شود.

❌ *۶- کدام گزینه درباره تابع توزیع چگالی احتمال (Probability Density Function) نادرست است؟



۱ مقدار چگالی احتمال می‌تواند بزرگتر از 1 باشد.



۲ با تغییر bin در نمودار هیستوگرام، شکل آن تغییر می‌کند.

۳ سطح زیر نمودار تابع توزیع چگالی احتمال 1 است.

۴ براساس نمونه، برآوردی درباره توزیع متغیر در جامعه می‌دهد.

✅ *۷- کدام گزینه درباره توزیع نرمال (Normal Distribution) درست است؟

۱ برای متغیرهای رسته‌ای قابل استفاده است.

۲ میانه توزیع همواره از میانگین آن بزرگتر است.



۳ یک توزیع متقارن است.

۴ همه پدیده‌ها از توزیع نرمال پیروی می‌کنند.

۸* - کدام گزینه درباره ضریب همبستگی پیرسون نادرست است؟

۱ شدت رابطه خطی بین دو متغیر عددی را اندازه‌گیری می‌کند.

۲ همواره عددی بین 1 و -1 است.

۳ اگر با افزایش یک متغیر عددی، متغیر عددی دیگر کاهش یابد، ضریب همبستگی بین این دو منفی است.

۴ می‌توان از آن بر اندازه‌گیری شدت رابطه بین دو متغیر رسته‌ای نیز استفاده کرد.

۹* - داده‌های زیر را در نظر بگیرید. کدام گزینه درباره آن نادرست است؟

9.42 , 8.64 , 9.09 , 7.33 , 8.76 , 11.42 , 11.39 , 9.35 , 10.22
13.35 , 8.31 , 7.9 , 11.09 , 7.3 , 11.27 , 11.2 , 9.97 , 11.63 , 9.82 , 5.23

۱ تابع توزیع این متغیر در جامعه چولگی زیادی به سمت راست دارد.

۲ براساس نمودار $q-q$ به نظر می‌رسد این متغیر از جامعه‌ای با توزیع نرمال آمده است.

۳ براساس نمودار توزیع چگالی احتمال به نظر می‌رسد این متغیر از جامعه‌ای با توزیع نرمال آمده است.

۴ براساس داده‌ها، برآورد از میانگین این متغیر در جامعه 9.63 است.

✓ ۱۰- ضریب همبستگی پیرسون بین x و y را محاسبه کنید.

x	y
9.45	23.71
8.85	17.19
8.01	15.20
11.79	22.48
7.35	14.06
12.91	28.89
11.34	22.60
10.12	20.19
11.58	23.86
9.02	18.39
14.22	25.69
9.53	26.22
10.44	19.99
6.99	13.64
11.76	21.53

۱ 0.923



۲ 0.822

۳ 0.752

۴ 0.615

Porsline

با پرسلاین فرم خود را طراحی و ارسال کنید و گزارش‌های آماری دریافت نمایید.