

۱- داده های فایل blood را درنظر بگیرید. تابعی بنویسید که اطلاعات افراد را بگیرید و مشخص کند که در رابطه با هر یک از فاکتورهای خونی، فرد دچار مشکل هست یا خیر و این مطلب را در قالب یک پیام در رابطه با آن مشکل، اعلام کند. از طرفی این تابع قابلیت این را داشته باشد که اگر فردی در وضعیت بحرانی قرار دارد یعنی حداقل بر اساس سه فاکتور خونی اش دچار مشکل باشد، وی را در لیست افراد در ریسک بیماری های قلبی و عروقی قرار دهد.
محدوده خطر:

BMI>2.5

Chol>200 کلسترول

HDL<40 در مردان و HDL<50 در زنان

LDL>160

TG>150 تری گلیسیرید

تابع نوشته شده را بر مجموعه داده های blood اعمال کنید و نتیجه را برای ده نفر اول گزارش دهید. همچنین لیست کاملی از تمام افرادی که در فایل قرار دارند و در ریسک بیماری های قلبی و عروقی هستند ارائه دهید.

۲- داده های bank را درنظر بگیرید. اطلاعات مربوط به حساب های بانکی افراد هست. متغیرهای سن و موجودی حساب (balance) را درنظر بگیرید. ابتدا نمودار پراکنش داده ها را رسم کنید. سپس در این نمودار با دو رنگ مختلف مشخص کنید که چه افرادی در ریسک پرداخت نکردن بدی های کارت های اعتباری خود در آینده قرار دارند. همچنین ده نفر از پر ریسک ترین و ده نفر از کم ریسک ترین افراد را برای رییس بانک با ذکر تمام مشخصات آنها تعیین نمایید. برای این افراد اطلاعات شغل، سن، میزان موجودی حساب و وضعیت تاهلشان را مشخص نمایید. اطلاعات بر اساس میزان موجودی حسابشان گزارش شوند.

از بررسی های قبلی مشخص شده که افرادی که بدی های کارت اعتباریشان را پرداخت نمی کنند به طور متوسط ۲۵ ساله با انحراف معیار ۸ هستند و متوسط موجودی حساب این افراد 200- با انحراف معیار ۲۰۰ است. راهنمایی: فرض می کنیم توزیع سن و موجودی حساب، در هر یک از دو دسته ریسک زیاد و ریسک کم، نرمال دو متغیره باشد با میانگین و انحراف معیارهای بالا و کوواریانس صفر. سپس اطلاعات هر صاحب حساب را در درستی های مربوط به هر دسته قرار میدهم. در هر دسته که درستی بیشتری داشت، صاحب حساب متعلق به آن دسته می شود.

۳- تابعی بنویسید که به وسیله آن با وارد کردن بردار مشاهدات و مقدار μ_0 و α و همچنین واریانس جامعه و مشخص کردن اینکه آزمون دو طرفه، کوچکتری در مقابل بزرگتری و یا برعکس است، آزمون فرض برای میانگین جمعیت انجام دهد و نتیجه را به صورت، مقدار آماره آزمون و پی مقدار گزارش دهد. تابع شما این قابلیت را داشته باشد که اگر داده گمشده داشتید، آن را حذف کند. فرض کنید توزیع جمعیت نرمال است.

به وسیله این تابع آزمون فرض زیر را انجام دهید.

$$H_0: \mu < 20 \quad H_1: \mu > 20$$

مشاهدات:

۱۶ ۹ ۸ ۱۰ ۲۶ ۲۴ ۲۳ ۵۳ ۲۰ ۱۰ ۱۲

موفق باشید