

تکلیف اضافه درس آمار مهندسی (گروه ۱): موعده تحویل ۱۴۰۲/۳/۳۰ ساعت ۲۳:۰۰

ارائه نتایج آماری با استفاده از نرم افزار Minitab

داده‌های ارائه شده در فایل اکسل به نام Customer_data.xlsx مربوط به مشتریان یک شرکت بیمه در ایالات متحده آمریکا می‌باشد. هر کدام از دانشجویان می‌بایست بر اساس اعدادی که به ایشان اختصاص داده شده و در برگه یا شیت Students Number موجود است؛ به روش زیر ردیف‌هایی را که باید تحلیل‌های آماری را بر روی آنها انجام دهند، جداسازی نمایند.

به هر دانشجو سه عدد تصادفی اختصاص داده شده است. دانشجویان می‌بایست با استفاده از هر یک از اعداد تصادفی که به آنها اختصاص داده شده ۴۰ سطر را انتخاب نمایند.

$$i = ((n + 8) \times 9) + 1$$

در رابطه بالا n عدد تصادفی اختصاص داده شده به دانشجو و i شماره ردیفی است که دانشجو اطلاعات خود را از آن آغاز می‌کند. برای مثال اگر به دانشجویی عدد ۲۷ اختصاص داده شده است، این دانشجو باید از سطر ۳۱۶ الی ۳۵۶ را جدا نموده و مورد بررسی قرار دهد. همین عمل عیناً باید برای عدد تصادفی دوم و سوم نیز انجام شود و نهایتاً ۱۲۰ سطر انتخاب شده و تحلیل‌های آماری بر روی آنها انجام گیرد.

$$i = ((27 + 8) \times 9) + 1 = 316 \quad , \quad 316 + 40 = 356$$

در صورتی که ردیف‌های انتخابی از سطر ۵۰۰ عبور کرد، تعداد سطر باقی‌مانده را از سطر اول شروع به انتخاب نمایید.

در این نمونه، ستون‌های زیر وجود دارد که توصیف هر کدام در جدول زیر ارائه شده است.

عنوان متغیر	توصیف متغیر
id	شماره شناسایی هر فرد
Gender	جنسیت
is.employed	آیا فرد شاغل می‌باشد؟ (۱ به معنای شاغل بودن می‌باشد).
income	درآمد سالانه
marital.stat	وضعیت تاهل
health.ins	آیا فرد دارای بیمه درمانی می‌باشد؟ (۱ به معنای داشتن بیمه درمانی می‌باشد).
housing.type	نوع منزل محل سکونت افراد
num.vehicles	تعداد وسایل نقلیه
age	سن
state.of.res	ایالت محل زندگی

(الف)

- درصد افراد شاغل و افرادی که بیمه درمانی دارند را بیابید.
- چند درصد از افراد هم آقا هستند و هم شاغل؟
- چند درصد از آقایان شاغل هستند؟
- مشخص نمایید چند درصد از آقایان هیچ گاه ازدواج نکرده‌اند (Never Married).
- نمودار جعبه‌ای و هیستوگرام درآمد و تعداد وسایل نقلیه را رسم نموده و چارک‌های اول و سوم را مشخص نمایید.
- داده‌های پرت مربوط به درآمد افراد را حذف نموده و مجدداً نمودار جعبه‌ای و هیستوگرام مربوط به درآمد را رسم نمایید. در صورت وجود تغییر بین نمودارها، علل تغییرات را بیان نمایید.
- کشیدگی و چولگی درآمد را قبل و بعد از حذف داده‌های پرت مقاسه کنید.

(ب) آزمون فرض‌های زیر را در سطح معناداری ۰/۰۵ انجام دهید.

فرضهای صفر و مقابل را مشخص کنید و مقدار آماره آزمون و p -value را نیز به صورت جداگانه گزارش دهید.

- مشخص نمایید آیا شواهد کافی برای رسیدن به این نتیجه که میانگین درآمد در ایالت متحده بیشتر از ۲۰۰۰۰ دلار است، وجود دارد؟
- مشخص نمایید آیا شواهد کافی برای رسیدن به این نتیجه که افراد به طور میانگین بیش از یک وسیله نقلیه دارند، وجود دارد؟
- مشخص کنید آیا شواهد کافی برای رسیدن به این نتیجه که میانگین درآمد افرادی که بیمه درمانی دارند، بیش از کسانی است که بیمه درمانی ندارند؟
- مشخص نمایید آیا شواهد کافی برای رسیدن به این نتیجه که میانگین درآمد آقایان بیشتری از خانم‌ها است، وجود دارد؟
- مشخص نمایید آیا شواهد کافی برای رسیدن به این نتیجه که میانگین درآمد افراد در ایالت کالیفرنیا بیش از میانگین درآمد افراد در سایر ایالت‌ها می‌باشد، وجود دارد؟
- مشخص نمایید آیا شواهد کافی برای رسیدن به این نتیجه که میانگین درآمد افراد در ایالت پنسیلوانیا کمتر از میانگین درآمد افراد در سایر ایالت‌ها می‌باشد، وجود دارد؟

توضیحات:

- برای پاسخ به سوالات قسمت ب داده‌های پرت^۱ را حذف نکنید.
- اگر در سطری از داده‌ها، اطلاعاتی از یک متغیر کمی که مورد نیاز شما است وجود نداشت، می‌توانید میانگین کل داده‌ها را برای آن متغیر در آن سطر در نظر بگیرید. اگر آن متغیر کیفی بود، مقداری از آن متغیر که بیشترین تکرار را داشته است را در نظر بگیرید.
- طراحی تکلیف با هدف یادگیری نرم‌افزار Minitab و فصول آزمون فرض و آمار توصیفی صورت گرفته است.

¹ Outlier

- برای پاسخ به برخی از سوالات ممکن است دانشجویان احتیاج به جستجو در اینترنت داشته باشند؛ این جستجو بخشی از هدف تکلیف می باشد.

روش تحویل تکلیف: لطفا فایل پی دی اف مربوط به گزارش حل تکلیف را در سامانه یکتا بارگذاری کرده و فایل مربوط به نرم افزار Minitab را نیز در سامانه IUTBox بارگذاری کرده و لینک آن را در انتهای گزارش خود یادداشت نمایید.