



با سلام و احترام،

## فصل چهاردهم: خلاصه آماره ها با استفاده از توصیفی ها (Descriptives)

روش توصیفی برای به دست آوردن خلاصه سنجش ها که به صورت تخمینی، متغیرهای Scale را به صورت نرمال توزیع کرده است و به آسانی موارد غیر عادی را در میان آن متغیرها با محاسبه Z Score (نمره Z) شناسایی می کند (Z Score به مقادیر استاندارد شده نیز معروفند. برای به دست آوردن Z-Score یک متغیر، برای هر مورد مقدار میانگین متغیر را کم کرده و بر انحراف معیار تقسیم نمایید).

$$Z = \frac{X - \bar{X}}{\sigma}$$

به عبارت دیگر، مقادیر استاندارد، وضعیت افراد یا موارد را نسبت به میانگین تعیین می کند.

### استفاده از آمار توصیفی برای مطالعه داده های کمی

یک شرکت مخابراتی اطلاعات مربوط به یک مشتری را که شامل مکالمات راه دور مشتری، مکالمه تلفن رایگان، اجاره تجهیزات، کارت تلفن، و سرویس بی سیم (Wireless) در ماه گذشته می باشد نگهداری می کند. با استفاده از آمار توصیفی به بررسی مقرون به صرفه بودن خدمات این شرکت برای مشتری پرداخته می شود.

### اجرای تحلیل

۱. برای تحلیل توصیفی، مسیر زیر را از منوی اصلی برگزینید:

Analyze > Descriptive Statistics > Descriptive

۲. گزینه های **Equipment Last month** , **Toll Free Last month** , **Long distance Last** , **Calling Card Last month** و **Wireless Last month** را به عنوان متغیرهای تحلیل انتخاب نمایید.

۳. روی **Ok** کلیک کنید.

این مراحل یک جدول آماره توصیفی را تشکیل می دهد که برای مقایسه مقادیری که برای هر سرویس پرداخته پرداخته شده است به کار می رود.

## جدول آماره توصیفی

از جدول، تشخیص اینکه کدام سرویس به صرفه تر است مشکل می باشد. به طور میانگین، مشتریها بیشترین مقدار را روی اجاره تجهیزات پرداخت کرده اند، اما کمی پراکندگی در مقادیر پرداخت شده وجود دارد. سرویس کارت تلفن مشتریها به طور میانگین فقط یک مقدار جزئی نسبت به تجهیزات اجاره ای کمتر است، و پراکندگی خیلی کمی در مقادیر وجود دارد.

|                          | N    | Minimum | Maximum | Mean    | Std. Deviation |
|--------------------------|------|---------|---------|---------|----------------|
| Long distance last month | 1000 | 90      | 99.95   | 11.7231 | 10.36349       |
| Toll free last month     | 1000 | 00      | 173.00  | 13.2740 | 16.90212       |
| Equipment last month     | 1000 | 00      | 77.70   | 14.2198 | 19.06854       |
| Calling card last month  | 1000 | 00      | 109.25  | 13.7810 | 14.08450       |
| Wireless last month      | 1000 | 00      | 111.95  | 11.5839 | 19.71943       |
| Valid N (listwise)       | 1000 |         |         |         |                |

مشکل واقعی در اینجا این است که بیشتر مشتریان برخی سرویسها را ندارند، بنابراین تعداد زیادی 0 شماره انداخته شده است. یک راه حل برای این مشکل آن است که عدد 0 به عنوان مقادیر غیر معتبر ( **missing** **Values** ) محسوب شود به طوری که به حساب آوردن هر سرویس در تحلیل، مشروط به داشتن آن سرویس باشد.

## تبدیل متغیرها

۱. برای ضبط 0 به صورت مقدار غیر معتبر، مسیر زیر را از منوی اصلی برگزینید:

**Transform > Recode Into Same Variables**

۲. گزینه های **Equipment Last month** , **Toll Free Last month** , **Long distance Last** , **Calling Card Last month** و **Wireless Last month** را به عنوان متغیرهای عددی برگزینید.

۳. دکمه **Old and New Values** را کلیک کنید.

۴. مقدار **0** را برای **Old Value** مقدار قدیمی تایپ نمایید.

۵. گزینه **System- missing** را برای قسمت **New Value** (مقدار جدید) فعال نمایید.

۶. روی **Add** کلیک کنید.

۷. دکمه **Continue** را کلیک کنید. عبارتی در کادر مربوط به آن ظاهر می شود تا مقدار جدید را به متغیر نسبت دهد.

۸. در کادر محاره **Recode into same Variables** روی **Ok** کلیک کنید.

۹. برای تحلیل یک آماره توصیفی بر مبنای متغیرهای تبدیل شده، کادر محاوره **Descriptive** را بازخوانی نمایید.

۱۰. دکمه **Options** را در این کادر محاوره کلیک کنید.

۱۱. گزینه های **Minimum** و **Maximum** را غیر فعال نمایید.

۱۲. گزینه های **Skewness** و **Kurtosis** را فعال کنید.

۱۳. دکمه **Continue** را کلیک کنید.

۱۴. **Ok** را در کادر محاوره **Descriptive** کلیک کنید.

## آماره توصیفی

وقتی تحلیل، مشروط به سرویس هایی است که مشتری از آنها استفاده می کند، نتایج به طور چشمگیری متفاوت می شود. سرویس **Wireless** (بی سیم) و اجاره تجهیزات، درآمد بیشتری نسبت به سرویس های دیگر برای شرکت دارد (a). از این گذشته، مادامی که سرویس بی سیم متغیری با بیشترین مشتری باشد، تجهیزات اجاره ای یکی از پایین ترین انحراف معیارها را خواهد داشت (b).

|                          | N         | Mean      | Std.      | Skewness  |            | Kurtosis  |            |
|--------------------------|-----------|-----------|-----------|-----------|------------|-----------|------------|
|                          | Statistic | Statistic | Statistic | Statistic | Std. Error | Statistic | Std. Error |
| Long distance last month | 1000      | 11.7231   | 10.38349  | 2.966     | .077       | 14.052    | .155       |
| Toll free last month     | 475       | 27.9453   | 13.82910  | 3.465     | .112       | 26.735    | .224       |
| Equipment last month     | 386       | 36.8389   | 10.39563  | 7.56      | .124       | .641      | .248       |
| Calling card last month  | 678       | 20.3260   | 12.62916  | 2.150     | .094       | 7.572     | .187       |
| Wireless last month      | 296       | 39.1348   | 15.32916  | 1.359     | .142       | 3.079     | .282       |
| Valid N (listwise)       | 131       |           |           |           |            |           |            |

## یافتن موارد غیر عادی

با مطالعه مقادیر استاندارد (یا نمرات  $Z$ ) متغیرها می‌توانید مشتریهایی که مبالغ بیشتر یا کمتری را نسبت به مشتریهای دیگر در هر سرویس پرداخت کرده‌اند پیدا کنید.

اگر چه، نیاز به استفاده از نمره  $Z$  به این دلیل است که توزیع هر متغیر به صورت غیرنرمال ( $\text{non-normal}$ ) تعیین نشده است. مقادیر  $\text{Skewness}$  و  $\text{Kurtosis}$  (چولگی و کشیدگی) در جدول آماره خیلی بزرگ گزارش شده‌اند، و این نشان می‌دهد که توزیع این متغیرها بدون شک غیرنرمال می‌باشد.

با توجه به مقادیر مثبت متغیرها، یک راه حل ممکن این است که نمره‌های  $Z$  ( $Z \text{ Score}$ ) متغیرهای تبدیل لگاریتم مطالعه شود. (تبدیل لگاریتم یا  $\text{log-transformed}$  همان لگاریتم طبیعی است که با اعمال روی متغیرهای مقدار مثبت، برای تصحیح غیرنرمال بودن استفاده می‌شود. این روش در مواقعی که غیرنرمال بودن ناشی از چولگی ( $\text{Skewness}$ ) مثبت باشد مؤثرتر است).

## اجرای تحلیل

۱. برای یافتن نمره‌های  $Z$  برای متغیرهای تبدیل لگاریتم، کادر محاوره  $\text{Descriptive}$  را بازخوانی کنید.
۲. از گزینه  $\text{log distance Last month}$  تا  $\text{Wireless Last month}$  را از فهرست متغیرهای تحلیل خارج کنید.
۳. از گزینه  $\text{log-long distance}$  تا  $\text{log-Wireless}$  را به عنوان متغیر تحلیل انتخاب کنید.
۴. گزینه  $\text{Save Standardized Values as Variables}$  را فعال نمایید. با فعال کردن این گزینه، مقادیر استاندارد به عنوان متغیر در نظر گرفته می‌شوند.
۵.  $\text{Ok}$  را کلیک کنید.

## جدول آماره توصیفی

به استثنای متغیر  $\text{log-toll free}$ ، چولگی و کشیدگی به طور قابل توجهی برای متغیرهای تبدیل لگاریتم کوچک‌تر است. خدمات تبدیل لگاریتمی تلفن رایگان ( $\text{toll-free}$ )، ممکن است در مورد مقادیر زیاد چولگی و کشیدگی ادامه یابد زیرا یک مشتری مبلغ زیادی که غیرعادی است برای ماه گذشته پرداخت کرده است. بررسی نمودار کادری ( $\text{box plot}$ ) این را تأیید می‌نماید.

|                    | N         | Mean      | Std.      | Skewness  |            | Kurtosis  |            |
|--------------------|-----------|-----------|-----------|-----------|------------|-----------|------------|
|                    | Statistic | Statistic | Statistic | Statistic | Std. Error | Statistic | Std. Error |
| Log-long distance  | 1000      | 2.1821    | .73455    | .166      | .077       | -.001     | .155       |
| Log-toll free      | 475       | 3.2397    | .41381    | .304      | .112       | 1.107     | .224       |
| Log-equipment      | 386       | 3.5681    | .27756    | .037      | .124       | -.344     | .248       |
| Log-calling card   | 678       | 2.8542    | .55729    | .081      | .094       | .109      | .187       |
| Log-wireless       | 296       | 3.5983    | .36729    | .200      | .142       | -.168     | .282       |
| Valid N (listwise) | 131       |           |           |           |            |           |            |

## نمودارهای کادری نمره‌های Z

زمانی که بخواهیم میانگین یک متغیر فاصله‌ای را بین دو یا چند گروه از پاسخ‌گویان مقایسه کنیم، از نمودار کادری استفاده می‌کنیم. و دو نوع نمودار کادری وجود دارد: نمودار کادری ساده (Simple) و نمودار خوشه‌ای (Clustered).

۱. برای مشاهده نمره‌های Z و یافتن مقادیر غیر عادی، مسیر **Graphs > Legacy Dialog > Boxplot** را از منوی اصلی برگزینید.

۲. در کادر محاوره **Boxplot**، گزینه **Summaries of separate Variables** را فعال نمایید.

۳. دکمه **Define** را کلیک کنید. توجه داشته باشید که نوع نمودار (در کادری ساده (Simple) در حالت انتخاب باشد.

۴. گزینه‌های **Z Score: log- long distance** تا **Z Score: log- Wireless** را به عنوان متغیرهایی که نمودار نشان می‌دهد انتخاب نمایید.

۵. دکمه **Options** را کلیک کنید.

۶. گزینه **Exclude Cases Variable by Variable** را در کادر محاوره **Options** فعال نمایید.

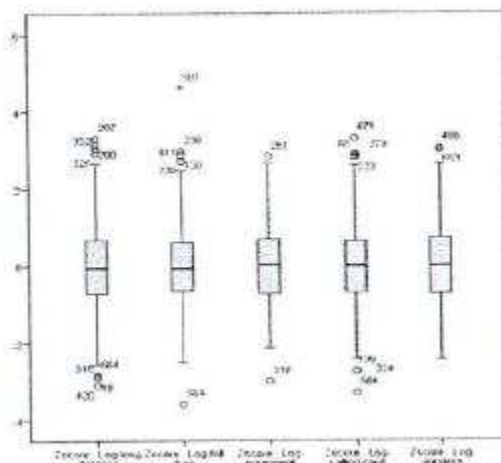
۷. دکمه **Continue** را کلیک کنید.

۸. **Ok** را در کادر محاوره **Define Simple Boxplot** کلیک کنید.

نمودار مستطیلی نمره‌های Z (**Z Score**) نشان می‌دهد که مشتری 567، به طور میانگین بیشترین استفاده را از سرویس تلفن رایگان (toll-free) در ماه گذشته برده است. این نکته باید برای چولگی و کشیدگی زیاد مشاهده شده در متغیر **Toll free Last month** در نظر گرفته شود.







## خروجی های آماره های توصیفی

جدول **Descriptive Statistics** نیز مشابه جداول فراوانی دارای همان شاخص ها یعنی میانگین، انحراف معیار، چولگی و کشیدگی است.

|                | 1995 Sales   | 1996 Sales    |
|----------------|--------------|---------------|
| Mean           | \$183,773    | \$371,893     |
| Minimum        | \$97,500     | \$189,000     |
| Maximum        | \$799,800    | \$1,350,000   |
| Sum            | \$71,671,440 | \$145,038,250 |
| Std. Deviation | \$78,142     | \$171,311     |
| Skewness       | 2.995        | 2.112         |
| Kurtosis       | 13.528       | 5.247         |

مبحث را در مرجع کاربردی **SPSS 20** (۲۱) دنبال نمایید.



**Telegram.me/iepnu**  
 کانال تخصصی مهندسی صنایع دانشگاه پیام نور