

1) نرم افزار اکسل کدامیک از شاخص های آماری زیر را انجام میدهد (10نمره)

مود - میانه - میانگین - واریانس - انحراف معیار - رگرسیون

حل : همگی

2) نرم افزار SPSS کدامیک از شاخص های آماری زیر را انجام میدهد (10نمره)

مود - میانه - میانگین - واریانس - انحراف معیار - رگرسیون

حل : همگی

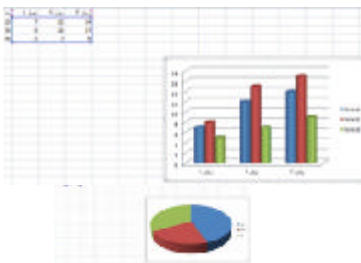
3) انواع روشهای مقیاس گذاری را نام ببرید هر کدام نیم خط توضیح دهید (10نمره)

مقیاس اسمی (NOMINAL) - کد گذاری عددی - اعداد نسبت داده شده ارجحیت و برتری ندارند (مثل جنسیت - زن 2 مرد 1)
مقیاس رتبه ای (ORDINAL) (ترتیبی) - با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند ولی تناسب ندارند (مثل پرسشنامه
لینکرت شامل موافق 3 - بی نظر 2 - مخالف 1)

مقیاس نسبتی (SCALE) (وزنی) با کد گذاری عددی - اعداد نسبت داده شده ارجحیت دارند و تناسب هم دارند (مثل سن و مثل وزن - مثلا
وزن احمد 35 و وزن عباس 70 کیلوگرم یعنی عباس بیشتر از احمد است و به نسبت دو برابر هم میباشد)

مقیاس فاصله ای - مثلا به یک بیمار بگوییم اگر حداکثر درد مثل یک خط کش عدد 100 باشد درد شما چقدر است و بیمار بگوید 65

4) چهار مورد از انواع نمودارها نام ببرید و هر یک نیم خط توضیح دهید و شکل نمودار را ترسیم کنید (10نمره)



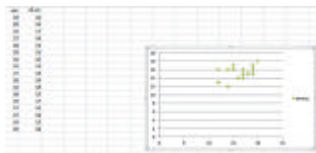
نمودار میله ای Bar Chart: برای مشاهده فراوانی داده ها و مقایسه داده ها نسبت به هم

نمودار دایره ای Pie Chart: برای مشاهده سهم هر مورد از داده ها

نمودار هیستوگرام Histogram: برای مشاهده نحوه توزیع داده ها بصورت پیوسته

نمودار خطی Line Chart: برای نشان دادن رابطه بین دوسری اعداد مثلا خط رگرسیون

نمودار پراکنندگی Scatter Plot: برای نشان دادن پراکنندگی بین دوسری اعداد مثلا در رگرسیون



5) جدول سوال زیر قبلا حل شده است و میانگین 13.2 و واریانس 3.36 میباشد. مجدد حل نکنید

X_i	11	13	12	15	17
f_i	2	3	2	2	1

الف) انحراف معیار را حساب کنید

ج) میانگین با انحراف معیار جمع و تفریق کنید و یکبار دیگر میانگین با دو برابر انحراف معیار جمع و تفریق

کنید و برای داده فوق که نرمال است هر دو جواب را کامل توضیح دهید. د) جدول زیر شباهت زیادی به

جدول بالا در همین سوال دارد با استفاده از جواب قسمت بالا میانگین و انحراف معیار جدول زیر را بدون حل و

بشکل سریع بدست آورید و (ضریب تغییرات این داده جدید را بدست آورید (20)

X_i	121	123	122	125	127
f_i	2	3	2	2	1

در داده های فوق هر یک از داده جدید باندازه 60 واحد بیش از داده قدیم است

در خصوص x ها خود مسئله عنوان کرده که

$$\bar{x} = 13.2 \quad \sigma_x = \sqrt{3.36} = 1.83$$

تحلیل $\mu \pm 2\sigma$ $\mu \pm \sigma$

$$\mu \pm \sigma = 13.2 \pm 1.83 = 11.38 \quad \text{تا} \quad 15.02$$

یعنی بیش از 68٪ داده ها بین 11.38 تا 15.02 میباشند

$$\mu \pm 2\sigma = 13.2 \pm (2 * 1.83) = 13.2 \pm 3.66 = 9.56 \quad \text{تا} \quad 16.84$$

یعنی بیش از 96٪ داده ها بین 9.56 تا 16.84 میباشند

طبق جدول فوق در خصوص y ها میتوان نوشت که

$$y_i = x_i + 110$$

و طبق قانون میتوان نوشت

$$\bar{y} = \bar{x} + 110 = 13.2 + 110 \rightarrow \bar{y} = 123.2$$

$$\sigma_y^2 = \sigma_x^2 \rightarrow$$

$$\sigma_y = \sigma_x = \sqrt{3.36} = 1.83 \rightarrow \sigma_y = 1.83$$

چون داده های جدید همگی 110 واحد اضافه شده اند پس میانگین داده های جدید 110 واحد بیشتر از میانگین داده های قبلی میشود و انحراف معیار داده های جدید با انحراف معیار داده های قدیم هیچ فرقی نمیکند

=====

ضریب تغییرات

$$\rho = \frac{\sigma}{\mu} = \frac{1.83}{123.2} = 0.014 = 1.4\%$$

یعنی پراکندگی داده ها کم است

همیشه $0 \leq \rho \leq 1$ است

اگر ρ صفر شود یعنی داده ها اصلا پراکنده نیستند داده ها بسیار بهم نزدیک هستند

اگر ρ یک شود یعنی داده ها خیلی پراکنده هستند

6) در یک منطقه میزان مبتلایان به ازدیاد قندخون در چند سال گذشته بشرح ذیل میباشد معادله خط

رگرسیون را نوشته پیش بینی سال 1400 چقدر میباشد (20نمره)

x=سال	1395	1396	1397	1398	1399	1400
میزان قند	2	2	3	3	5	??

x=سال	1395	1396	1397	1398	1399	1400
میزان قند	2	2	3	3	5	??
	-2	-1	0	1	2	3

$$y = a + bx$$

$$b = \frac{\sum x_i y_i - \frac{\sum x_i * \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$= \frac{(-2 * 2) + (-1 * 2) + (0 * 3) + (1 * 3) + (2 * 5) - \frac{(-2 - 1 + 0 + 1 + 2)(2 + 2 + 3 + 3 + 5)}{5}}{((-2)^2 + (-1)^2 + (0)^2 + (1)^2 + (2)^2) - \frac{(-2 - 1 + 0 + 1 + 2)^2}{5}}$$

$$= \frac{7}{10} = 0.7$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{-2 - 1 + 0 + 1 + 2}{3} = 0 \qquad \bar{y} = \frac{\sum y_i}{n} = \frac{2 + 2 + 3 + 3 + 5}{5} = \frac{15}{5} = 3$$

$$\bar{y} = a + b\bar{x} \qquad 3 = a + (0.7 * 0) \qquad a = 3$$

$$y = 3 + 0.7x$$

$$Xm = 1400 \rightarrow x = 1400 - 1397 = 3 \rightarrow y = 3 + (0.7 * 3) = 5.1$$

7) نمرات دانشجویان بشرح زیر است داده‌ها را در جدول زیر مرتب کنید. الف) مد(نما) داده‌ها چه مقدار است (چرا) ب). میانه داده‌ها چه مقدار است؟ ج) میانگین و واریانس و انحراف معیار نمرات دانشجویان را بدست آورید. و) با تکمیل جدول، فراوانی نسبی و فراوانی نسبی تجمعی را محاسبه کنید و با یک مثال نسبی‌ها را توضیح دهید.. نمرات: 13 و 12 و 12 و 13 و 17 و 13 و 16 و 16 و 11 و 13 (20نمره)

نمره = x					
تعداد = f					

حل:

x داده	11	12	13	16	17
f فراوانی	1	2	4	2	1
F فراوانی تجمعی	1	3	7	9	10
r = $\frac{f_i}{\sum f_i}$ فراوانی نسبی	1/10=10%	2/10=20%	4/10=40%	2/10=20%	1/10=10%
R = $\frac{\sum f_{i_i}}{\sum f_i}$ فراوانی نسبی تجمعی	1/10=10%	3/10=30%	7/10=70%	9/10=90%	10/10=100%

مد داده ای بمقدار 13 میباشد زیرا بیشترین تعداد را دارد

$$M=13$$

میانه داده‌ها

$$Q \sum f_i = \frac{1}{2}(1 + 2 + 4 + 2 + 1) = \frac{10}{2} = 5 \rightarrow \text{در سطر } F_i \rightarrow 5+ \rightarrow F_i=7 \rightarrow m=13$$

یعنی نصف داده‌ها 13 یا کمتر از 13 هستند

$$\bar{x} = \mu = \frac{\sum x_i * f_i}{\sum f_i} = \frac{(11 * 1) + (12 * 2) + (13 * 4) + (16 * 2) + (17 * 1)}{1 + 2 + 4 + 2 + 1} = \frac{136}{10} = 13.6$$

$$\sigma^2 = \frac{\sum (x - \bar{x})^2 f_i}{\sum f_i}$$

$$\sigma^2 = \frac{(11 - 13.6)^2 * 1 + (12 - 13.6)^2 * 2 + (13 - 13.6)^2 * 4 + (16 - 13.6)^2 * 2 + (17 - 13.6)^2 * 1}{1 + 2 + 4 + 2 + 1} = 3.64 \text{ واریانس}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{3.64} = 1.9 \text{ انحراف معیار}$$

جدول در بالا تکمیل شد

نسبی r یعنی اینکه مثلاً از داده به مقدار 13 بتعداد 4 تا داریم یعنی 40٪ داده‌ها 13 هستند

نسبی تجمعی R یعنی اینکه مثلاً از داده به مقدار 13 و کمتر از 13 بتعداد 7 تا داریم یعنی 70٪ داده‌ها 13 یا کمتر از 13

هستند