

پروژه آمار

این فایل فقط برای مشاهده است برای خرید به سایت کافی نت مراجعه کنید فقط سه هزار

تومان

موضوع مورد بررسی:

ساعاتی که هر فرد در طول هفته برای گوش دادن به موسیقی صرف می کند

جامعه آماری: مدرسه

فهرست مطالب

۱	پروژه آمار
۱	جامعه آماری: مدرسه
۳	نمونه: کلاس
۴	دسته بندی داده ها
۴	جدول فراوانی ها
۴	فراوانی مطلق
۵	شاخص های هندسی:
۶	نمودار مستطیلی :
۷	نمودار میله ای
۸	نمودار چند بر فراوانی
۹	نمودار ساقه و برگ
۱۰	نمودار دایره ای
۱۱	میانگین :
۱۲	نمودار جعبه
۱۳	شاخص های پراکندگی
۱۳	دامنه تغییرات :
۱۴	واریانس (پراش)
۱۴	انحراف معیار :
۱۵	ضریب تغییرات :
۱۵	پراکندگی :

نمونه : کلاس

۲۵.۳۰.۲۵.۲۰.۱۰.۳۰.۱۵.۳۵.۱۴

۵.۵.۱۰.۱۵.۱۵.۲۰.۲۵.۲۵.۲۰.۵.۱۵.۲۵

۳۰.۵.۱۰.۱۰.۲۰.۵.۱۵.۲۵.۲۵.۳۰.۳۵.۲۰

۱۵.۲۰.۳۰.۲۵.۵.۱۰.۱۵.۲۰.۱۵.۵.۳۰.۲۵

۲۰.۱۵.۵.۱۰

دسته بندی داده ها کمترین داده - بیشترین داده = دامنه تغییرات

$$R=35-5=30 \quad \text{طول دسته} = \frac{30}{5}=6$$

دامنه تغییرات = طول دسته

تعداد دسته

جدول فراوانی ها :

دسته	فراوانی	مرکز دسته	فراوانی نسبی	درصد فراوانی نسبی	فراوانی مطلق
(۱۰،۵)	۹	$\frac{5+10}{2} = 7.5$	$\frac{9}{50} = 0.18$	$0.18 \times 100 = 18\%$	۹
۱۵،۱۰)	۶	۱۲/۵	$\frac{6}{50} = 0.12$	$0.12 \times 100 = 12\%$	۱۵
۱۵و۲۰)	۱۰	۱۷/۵	$\frac{10}{50} = 0.2$	$0.2 \times 100 = 20\%$	۲۵
(
(۲۵،۲۰)	۸	۲۲/۵	$\frac{8}{50} = 0.16$	$0.16 \times 100 = 16\%$	۳۳
(۳۰،۲۵)	۹	۲۷/۵	$\frac{9}{50} = 0.18$	$0.18 \times 100 = 18\%$	۴۲
(۳۵،۳۰)	۶	۳۲/۵	$\frac{6}{50} = 0.12$	$0.12 \times 100 = 12\%$	۴۸
(۴۰،۳۵)	۲	۳۷/۵	$\frac{2}{50} = 0.04$	$0.04 \times 100 = 4\%$	۵۰

شاخص های هندسی:

نمودارهای شاخص های هندسی وسیله ای سودمند برای به تصویر در آوردن و تجسم جامعه

می باشد .

نمودارها :

۱- نمودار مستطیلی

۲- نمودار میله ای

۳- نمودار دایره ای

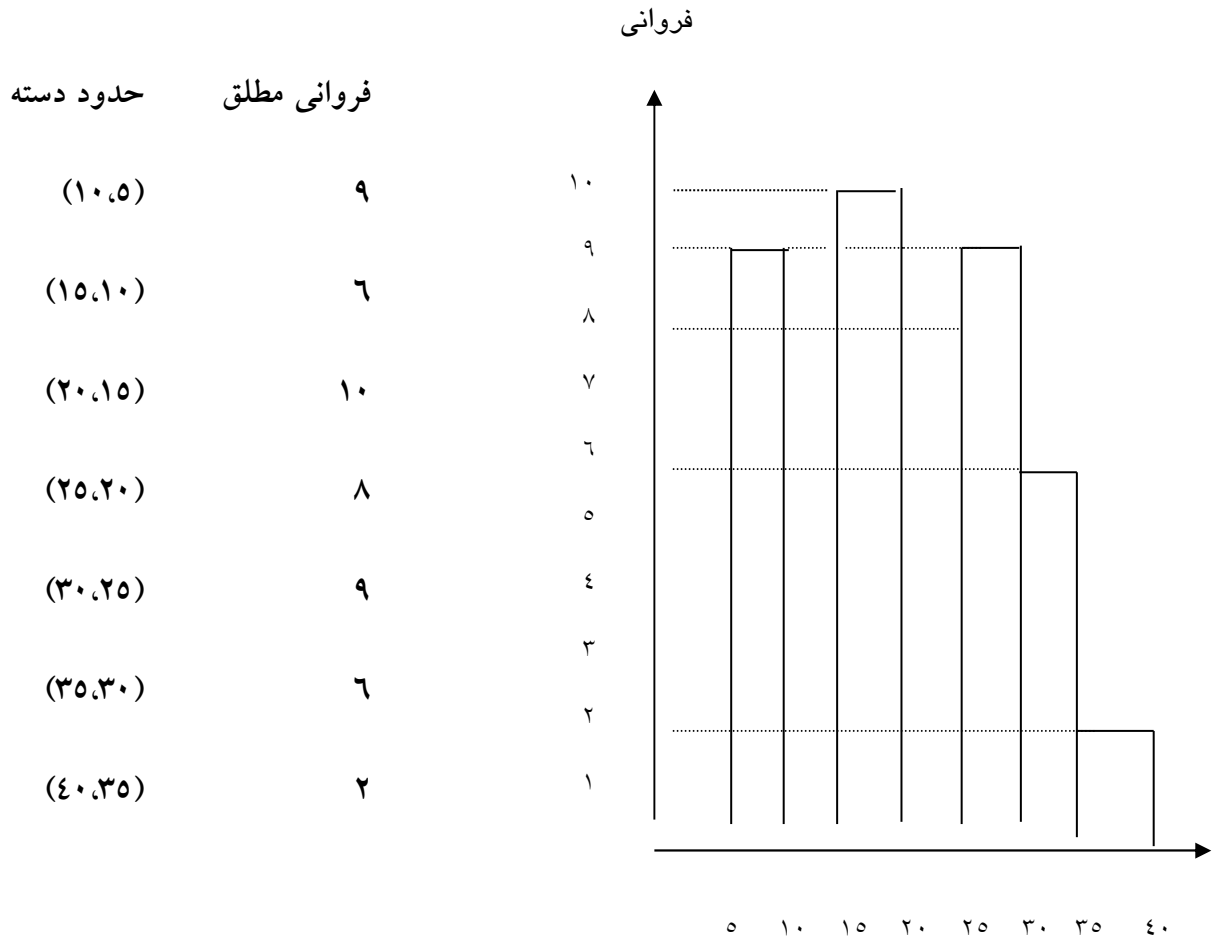
۴- نمودار چند بر فراوانی

۵- نمودار ساقه ای و برگ

نمودار مستطیلی :

نمایشی از داده های دسته بندی شده که در آن سطح مستطیل ها متناسب با فراوانی دسته

هست به نمودار مستطیلی نمودار ستونی یا نمودار هیستوگرام نیز می گویند.

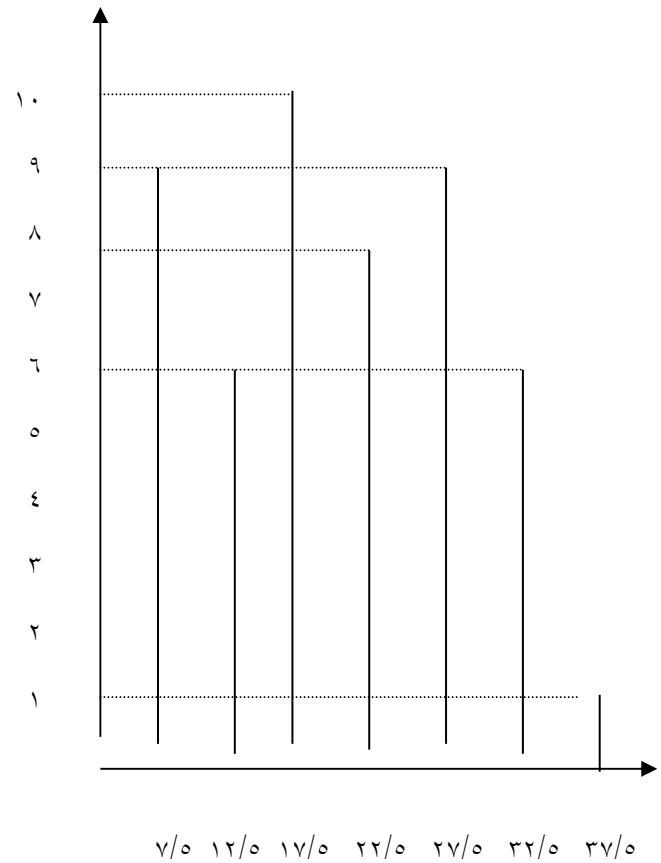


نمودار میله ای

این نمودار بیشتر به این متغیرهای گسسته و کیفی مناسب است آنچه در این نمودار مهم است

مقایسه فراوانی داده هاست .

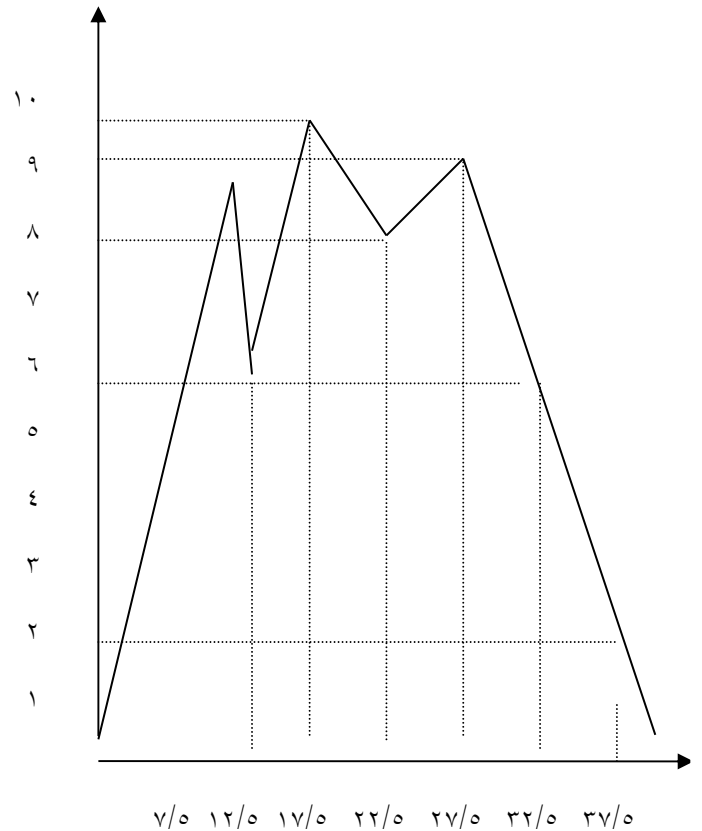
حدود دسته	فراوانی مطلق
۷/۵	۹
۱۲/۵	۶
۱۷/۵	۱۰
۲۲/۵	۸
۲۷/۵	۹
۳۲/۵	۶
۳۷/۵	۲



نمودار چند بر فراوانی

در نمودار مستطیلی فراوانی ها روی دسته ها تغییر نمی کنند و نمودار مستطیل بیشتر شبیه به پلکانی به صورت رو به رو است اگر بخواهیم تغییرات متغیر را فاصله باشد بهتر نشان دهیم از نمودار چند بر فراوانی استفاده می کنیم چند بر فراوانی برای این داده ها پیوسته متناسب است به نمودار چند بر فراوانی نمودار خط شکسته نیز می گویند .

حدود دسته	فراوانی مطلق
۷/۵	۹
۱۲/۵	۶
۱۷/۵	۱۰
۲۲/۵	۸
۲۷/۵	۹
۳۲/۵	۶
۳۷/۵	۲



نمودار ساقه و برگ

خوبی این نمودار در این است که تمام داده ها را در بر دارد . عدد ۲ که در واقع همان عدد

۲۰ است و قسمت اصلی داده ها ست ساقه می گوئیم. سایه ارقام را که متصل به این ساقه

هستند برگ می گوئیم .

ساقه	برگ
۱	۰۰۰۰۴۵ ۵۵۵ ۵۵۵۵۵ ۵
۲	.
۳	۰۰۰۰۰۰۰۰ ۰۵۵۵۵۵۵۵۵۵
	۰۰۰۰۰۰۰۵۵

نمودار دایره ای

نموداری که می تواند اطلاعات موجود در داده ها رابه سرعت در معرض دید قرار دارد .

زاویه مرکزی بر حسب درجه = $۳۶۰ \times$ فراوانی آن دسته

کل فراوانی

فراوانی	
۹	$\frac{9}{50} \times ۳۶۰ = ۸/۶۴$
۶	$\frac{6}{50} \times ۳۶۰ = ۴۳۲$
۱۰	$\frac{10}{50} \times ۳۶۰ = ۷۲$
۸	$\frac{8}{50} \times ۳۶۰ = ۵۷/۶$
۹	$\frac{9}{50} \times ۳۶۰ = ۶۴/۸$
۶	$\frac{6}{50} \times ۳۶۰ = ۴۳۲$
۲	$\frac{2}{50} \times ۳۶۰ = ۱۴/۴$

میانگین :

میانگین معدل داده ها می باشد میانگین در مرکز داده ها قرار دارد یعنی همان مقداری که داده

ها از میانگین اضافه دارند، همان مقدار هم از میانگین کم دارند .

میانگین شاخص خوبی برای نشان دادن مرکزیت داده هاست در بعضی از مسائل این شاخص

نمی تواند مؤثر باشد .

داده ها = x_i

فروانی = F_i

$$X = \frac{5^9 + 10^0 + 14 + 15^1 + 20^8 + 25^9 + 30^6 + 35^2}{50} = 17/38$$

$$x = \frac{4 x_i f_i}{4 f_i} = \frac{1015}{50} = 20/3$$

دسته	F_i	x_i	$x_i f_i$
(۵-۱۰)	۹	۷/۵	۶۷/۵
(۱۰-۱۵)	۶	۱۲/۵	۷۵
(۱۵-۲۰)	۱۰	۱۷/۵	۱۷۵
(۲۰-۲۵)	۸	۲۲/۵	۱۸۰
(۲۵-۳۰)	۹	۲۷/۵	۲۴۷/۵
(۳۰-۳۵)	۶	۳۲/۵	۱۹۵
(۳۵-۴۰)	۲	۳۷/۵	۷۵

$$\sum f_i = 50$$

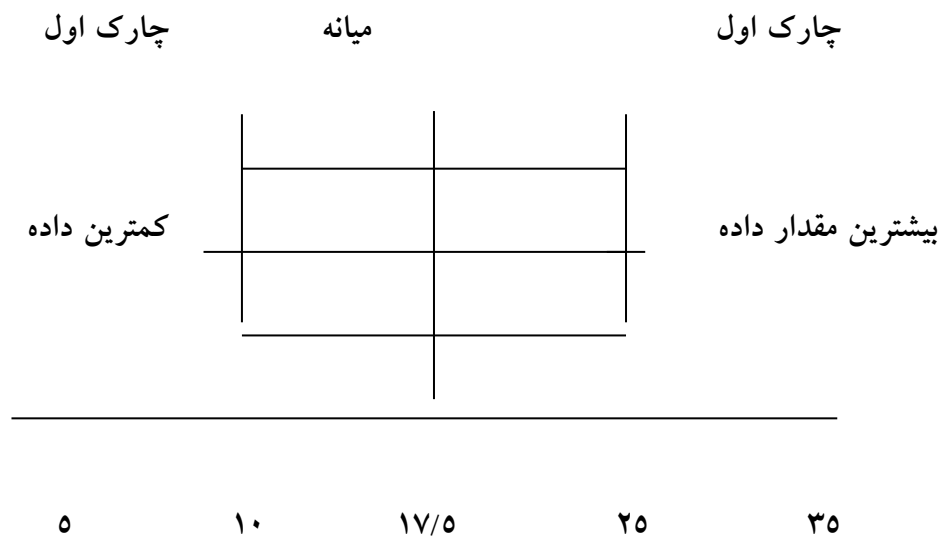
$$\sum x_i f_i = 1015$$

می بینیم که اختلاف بسیار زیاد است

نمودار جعبه

این نمودار مشخص کننده وضعیت داده ها است یعنی اینکه آیا داده ها بیشتر اطراف میانگین متمرکزند یا بیشتر اطراف کمترین داده یا بیشترین داده متمرکزند و یا اینکه آیا داده ها با هم نزدیک هستند یا خیر؟

میانۀ اول داده ها را چارک اول Q_1 میانۀ دوم داده ها را چارک دوم Q_2 می نامند .



شاخص های پراکندگی

- (۱) دامنه ی تغییرات
- (۲) چارک های بالا و پایین
- (۳) واریانس
- (۴) انحراف معیار
- (۵) ضریب تغییرات

دامنه تغییرات :

بزرگترین داده - کمترین داده $R=35-5=30$

چارک های بالا و پایین .

دامنه ی تغییرات ممکن است در بعضی مواقع تعبیرهای نامناسب از جامعه ارائه کند مثلاً در جامعه ای که عملاً داده ها به هم نزدیک اند به علت وجود مقدار خیلی کوچک و خیلی بزرگ در جامعه دامنه ی تغییرات عدد بزرگی به دست آید و حال آن که جامعه از دید آماری جامعه متمرکزی باشد در این حالت ها از این روش استفاده می کنیم .

واریانس (پراش)

برابر میانگین مجذور از میانگین است و آن را با σ^2 نشان می دهیم بنا براین : اگر داده ها

برابر باشند واریانس آن ها صفر است و بالعکس

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + 000(x_n - \bar{x})^2}{2}$$

$$\sigma^2 = \frac{(14 - 20/3)^2 + (35 - 20/3)^2 + (15 - 20/3)^2 + 000 + 920 - 20/3)^2}{50}$$

$$\sigma^2 = 28/5$$

انحراف معیار :

بانماد σ نشان داده می شود برابر جذر واریانس است .

واحد آن همان واحد متغیر است

تفاوت عمده در واحد واریانس و واحد میانگین را با جذر گرفتن از واریانس از بین می بریم

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + 000(x_n - \bar{x})^2}{n}}$$

باتوجه به واریانس که در قسمت بالا به دست آوردیم انحراف معیار برابر است با

$$\sigma = \sqrt{28/5}$$

ضریب تغییرات :

نماد CV است یعنی خارج قسمت انحراف معیار بر میانگین

برای از بین بردن واحد اندازه گیری از معیار ضریب تغییرات استفاده می کنیم .

ضریب تغییرات یعنی میزان پراکندگی به ازای یک واحد از میانگین

انحراف معیار

$$CV = \frac{v}{x} = \sqrt{\frac{28/5}{20/3}} \quad CV = \sqrt{\frac{28/5}{20/3}}$$

میانگین

پراکندگی :

یعنی این که داده ها از مرکز خود چقدر دور هستند برای این کار تک تک داده ها را از

میانگین کم می کنیم این تفاضل را انحراف از میانگین می نامیم

$$X_1 - \bar{X}, X_2 - \bar{X}, \dots, X_n - \bar{X}$$

$$۱۴ - ۲۰/۳ = ۶/۳ و ۰۰۰۰۰$$