

به نام خدا

گزارش تمرین اول

در سوال ۱,۱ گفته شده است که مواردی را محاسبه کنیم:

در اکسل ارسال شده فایل main شامل این مورد است که فایل پایتونی نیز برای آن ارسال میکنم که همه ی این موارد را در آن با توجه به ایت دیتاست محاسبه کرده ام.

Number of element: 920

The median:1.8

Minimum:0.7

Maximum:6.3

Range:5.6

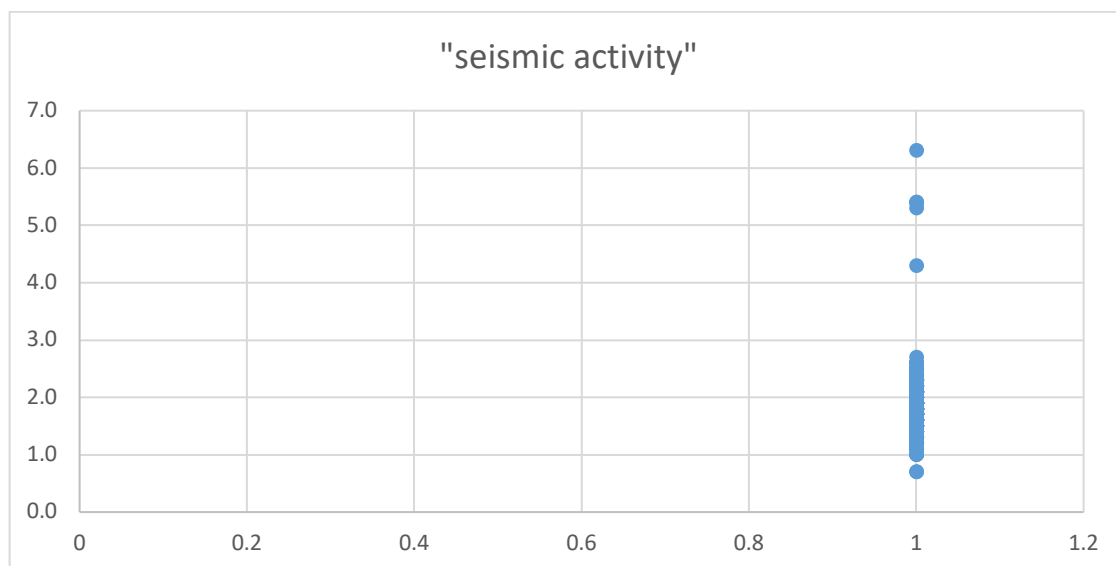
First quartile:1.6

Third quartile:2

IQR:0.4

95% confidence interval: $1.7 < \mu < 1.9$

در صورت سوال ۱,۲ گفته شده است که نمودار scatter plot را رسم کنیم در این مکان قرار دهیم.



در صورت سوال ۱,۳ گفته شده است :

توزیع مجموعه داده ها : که پراکندگی داده ها در نمودار نشان داده میشود و در نمودار هم قابل فهمیدن است که داده ی پرت وجود دارد.

بله به نظر میرسد مجموعه داده ها به دنبال توزیع نرمال سازمان یافته باشند.

در قسمت ۱,۴ با استفاده از فرمولی می‌خواهیم outlierها را تشخیص بدهیم:

که طبق فرمول گفته شده ۹ عدد outlier تشخیص داده شد که تعداد داده ها با این وجود از ۹۲۰ به ۹۱۱ تغییر پیدا کرد .

فرمول ان در قسمتی که نام ان outlier هست در اکسل وجود دارد.

=IF(OR(A1>\$E\$4;A1<\$E\$5);"Outlier";"Not Outlier")

در قسمت ۱,۵ گفته شده است outlierها را حذف کنید و مواردی را محاسبه کنید که با توجه به اطلاعات موجود در قسمت outlierمن این موارد محاسبه شده است .

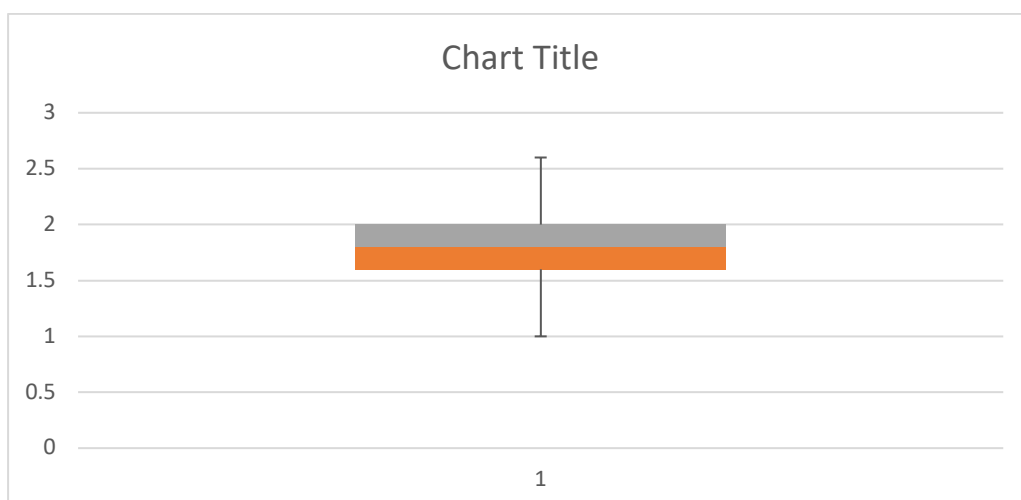
که انحراف معیار از ۰,۴ به 2.9 تغییر پیدا کرد.

میانۀ همان مقدار ۱,۸ است و تغییری نکرد.

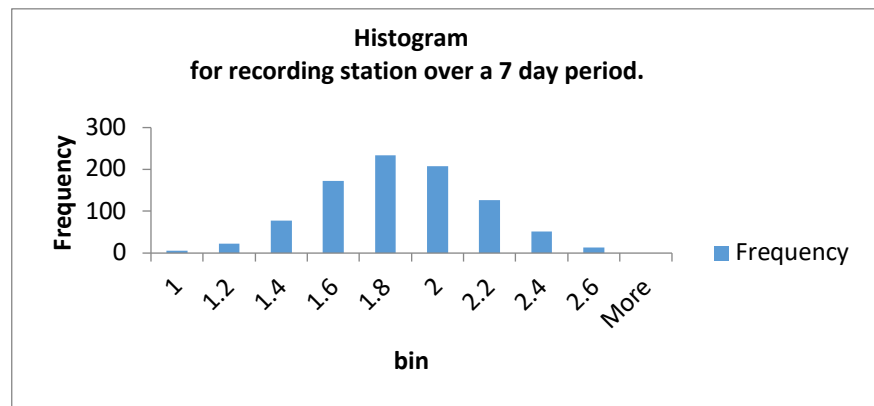
میانگین هم تغییری نکرده است و برابر ۱,۸ است.

بنظرم زیاد این مجموعه داده ها تحت تاثیر outlier ها قرار نگرفته اند و با توجه به تعداد زیاد داده ها نه عدد داده ی پرت قابل جبران است.

۱,۶: رسم whiskers plot با داده های جدید:



در قسمت ۱,۷ رسم هیستوگرام داده های جدید:



در قسمت ۲ :

N=22

Mean=39

Standard deviation:3.4

Sqrt(n):4.6

Error:3.4/4.6=0.73

$39-0.73 < \mu < 39+0.73$

$39-1.96*(0.73) < \mu < 39+1.96*(0.73)$

$39-1.43 < \mu < 39+1.43$

$۳۷,۶ < \mu < 40.4$

شکل نمودار کد زده شده را در زیر آورده شد:

