**Spread**

يتناول هذا الدرس الموضع التانى من الاحصاء الوصفيه وهوالانتشار وكيفية حسابه

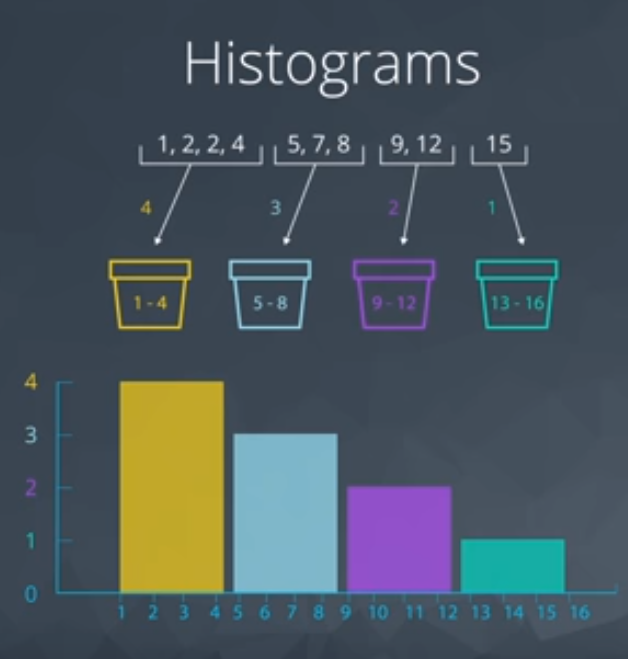
**Measures of Spread** are used to provide us an idea of how spread out our data are from one another. Common measures of spread include:

و يمكن قياس الانتشار عن طريق اربع مقاييس

1. **Range النطاق**
2. **Interquartile Range (IQR) النطاق الربعى**
3. **Standard Deviation الانحارف المعيارى**
4. **Variance التباين**

**Histograms**:

هو التعبير عن الداتا عن طريق استخدام الرسومات البيانيه



**Five Number Summary** : طريقه تستخدم لحساب النطاق و النطاق الربعى

فى هذا الطريقه نتبع عدة خطوات للحصول على الخمس ارقام المطلوبه

(minimum – first quartile – second quartile[median] – third quartile – maximum)

مثال :

Set1 = [ 5 , 8 , 3 , 2 , 1 , 3 , 10 ]

اولا نقم بترتيب العناصر تصاعديا

Set1 = [ 1 ,2 , 3 , 3 , 5 , 8 , 10 ]

Max = 10

Min = 1

Median = 3

يتبقى رقمين فقط هما

First and third quartile

First quartile هو الوسط للمجموعه التى تقع على يسار الوسط للمجموعه الام

Set1 = [1 ,2 , 3 , 3 , 5 , 8 , 10 ]

Subset1 = [1 , 2 , 3]

First quartile = 2

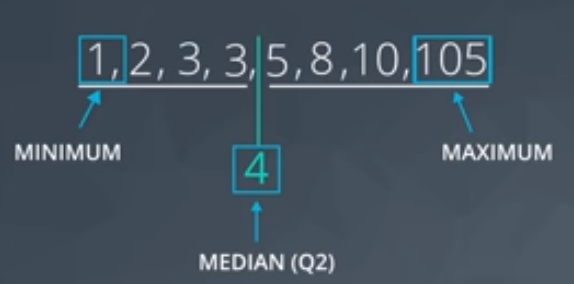
Third quartile هو الوسط للمجموعه التى تقع على بمين الوسط الخاص بالمجموعه الام

subSet2 = [ 5 , 8 , 10 ]

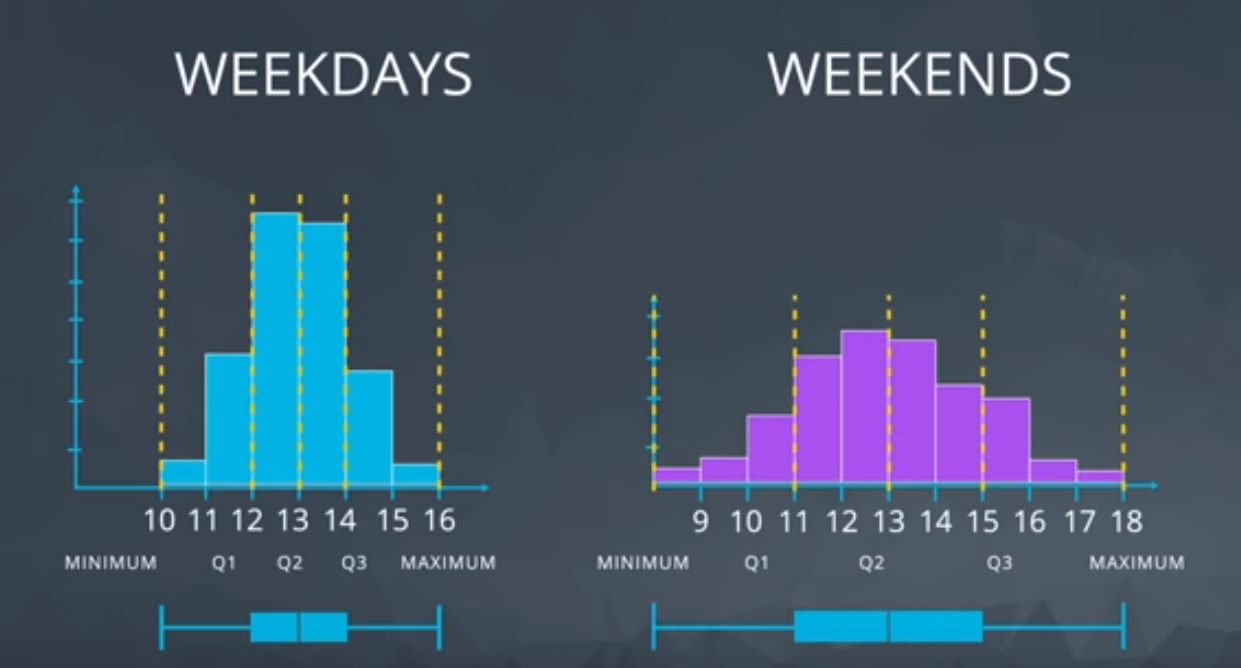
Third quartile = 8

و بالتالى الخمس ارقام المطلوبه قد حصلنا عليهم

اذا كان عدد الارقام زوجى فاكما تعلمنا فالدرس السابق كيفية حساب الوسط لمجموعه الارقام الزوجيه   
نقوم بحساب الوسط ونقسم المجموعات الفرعيبه حوله كالصوره التاليه



عن طريق الخمس ارقام نحصل على

Range and quartile range

وبتمثيل الخمس نقط على الرسم البيانى فالمدى هو فرق اكبررقم مع اصغر رقم

Q1 ,Q2 يقع النطاق الربعى بين

Standard Deviation الانحراف المعيارى

هى طريقه افضل للمعرفة انتشاء الداتا عن طريق رقم واحد فقط

النحراف المعيارى هو مدى تبعد كل نقطه عن المتوسط

مثال

كما فالصوره نفرض ان هناك مجموعه من الافراض تعمل فى مكان واحد وتبعد المسافة التاليه كما بالصوره

فاذا كمنا بحاب المتوسط للبعد الجميع عن العمل يكون الناتج 18



المتوسط يعرض كم يبعض مكان العمل عن كل فرد ولكن اذا ارضنا عمل مقارنة الفرقات بين مسافات كل موظف عن العمل

او بمعنى اخر كم يبعد كل فرد عن متوسط المسافات وبذالك نعرف مدى انتشار الداتا 

**Variance التباين**

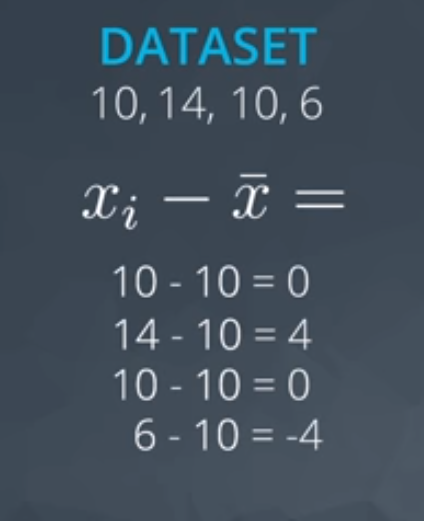
التباين هو متوسط مربعات الانحرافات المعياريه للنقاط

"اكيد انت مش فاهم حاجه من الكلام الكبير ده ههههههههه "

فى الانحراف المعيارى عرفنا كم يبعد كل فرد او نقطه عن المتوسط للمجموعة الارقام كلها

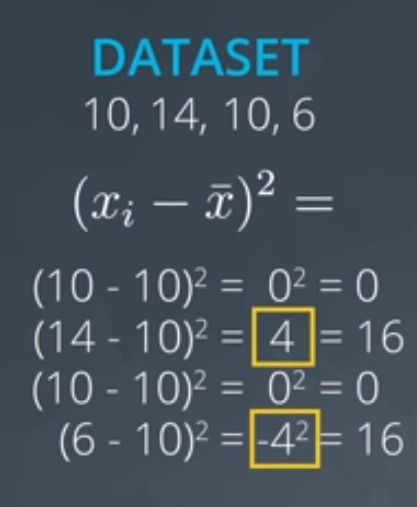
اى ان اصبح لكل نقطه انحراف معيارى خاص بها ماذا لو ارضنا معرفه متوسط الانحرافات المعياريه او رقم واحد يدل على الانحراف المعيارى لكل النقاط

نقوم بجمع كل الانحرافات المعياريه ونقسمها على عددها



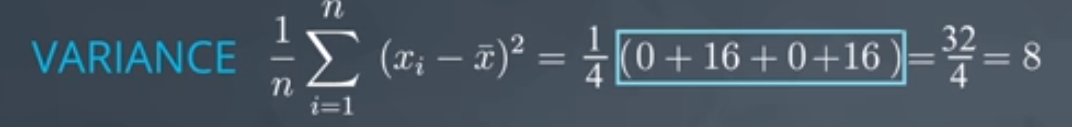
فى الصوره السابقه مجموعة ارقام 6 10 14 10 متوسط هذه الارقام هو 10

و فالصوره الانحراف المعيارى لكل رقم نلاحط ان هناك انحرافات معياريه نتيجته فيمه سالبه للتخلص من القيم السالبه نقم بتربيع كل النواتج

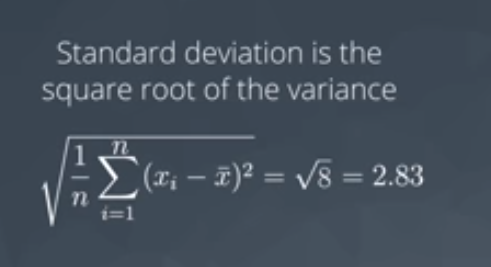


نقوم بجمعهم ونقسمهم على عددهم فيكون الناتج هو 8

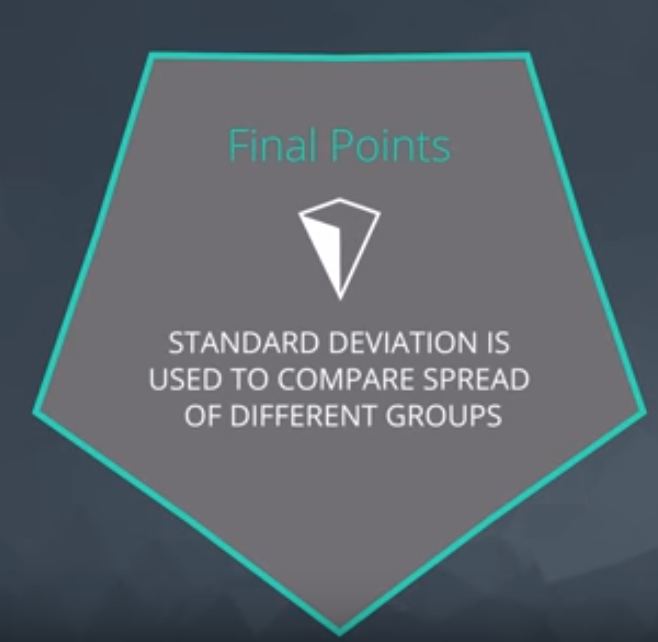
Variance = 8



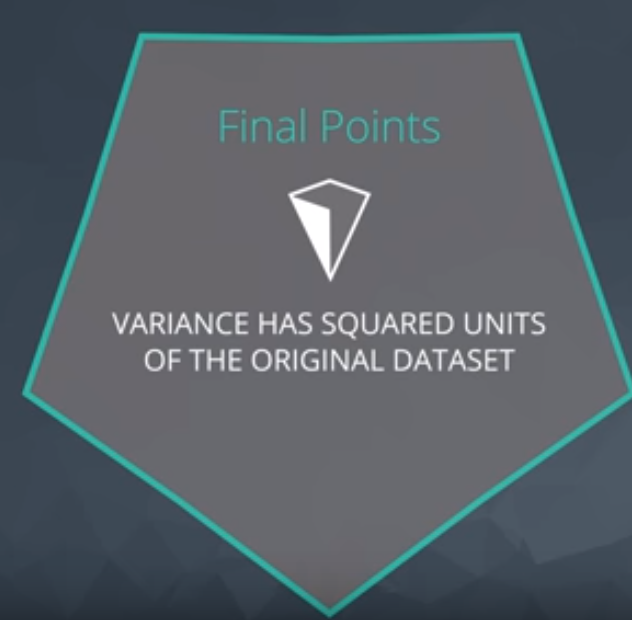
الانحارف المعيارى هو الجزرالتربيعى للتباين

= 

الان و عن طريق الانحراف المعيارى نستطيع ان نقارن مجموعتين مختلفتين من البيانات عن طريق رقم واحد فقط ونعرف اى من الاثنين اكثر انتشارا على الرسم البيانى



يجب استخدام الانحراف المعيارى وعمل المقارنات بين مجموعتين من نفس النوع اى لا يمكن ان ان نقارن مجموعات بيانات البرصه التى تتعامل بالدولار مع بيانات برصه اخرى تتعامل باليورو



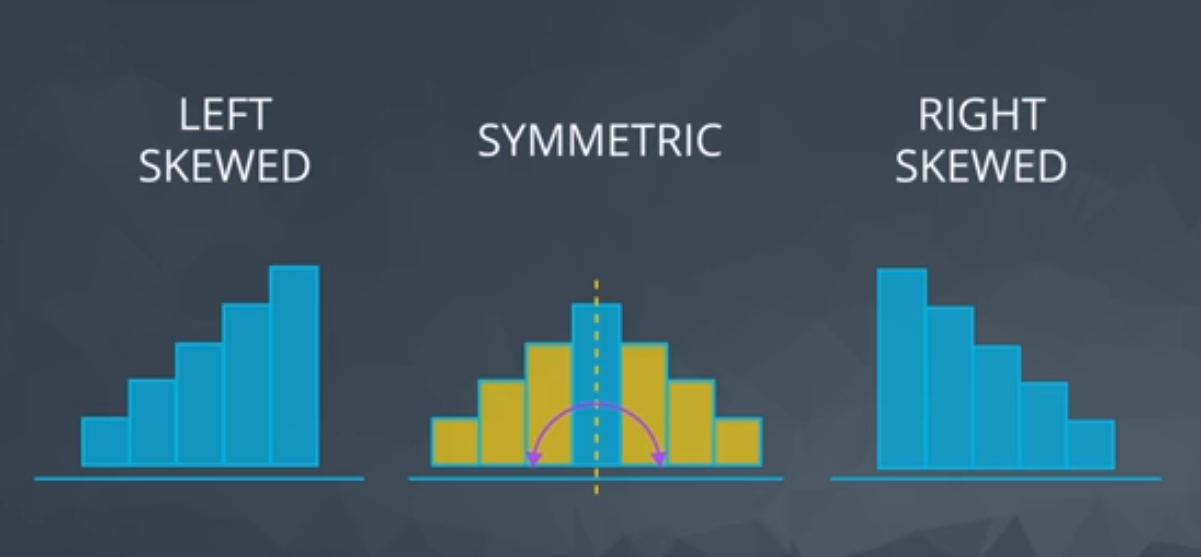
لو لدى شركتين يمكننى ان استثمر بهم ومتوسط ربح الاولى يساوى متوسط ربح الثانيه لا يجب ان استثمر بها لاننى لا اعرف هل توزيع الربح متساوى بين الكل ام هناك فرقات كبيره ولذلك يجب ان اعرف الانحراف المعيارى ايضا

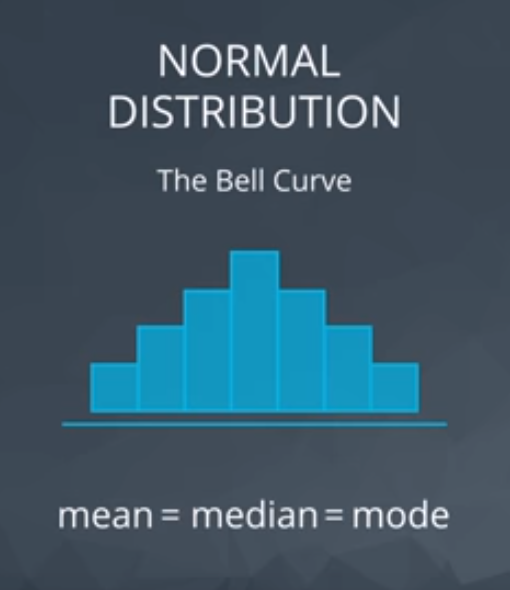
وعلى نفس المثال اذا كنت اعرف الانحراف المعيارى فقط لا يجب ان استثمر حتى اعرف متوسط الربح فقد يكون الشركتان متسوايتان فى الانحراف المعيارى اى ان الفرقات بين جميع المستثمرين واحده لكن لا الربح نفسه مختلف مثال

1, 2, 3, 4 and 5, 6, 7, 8.

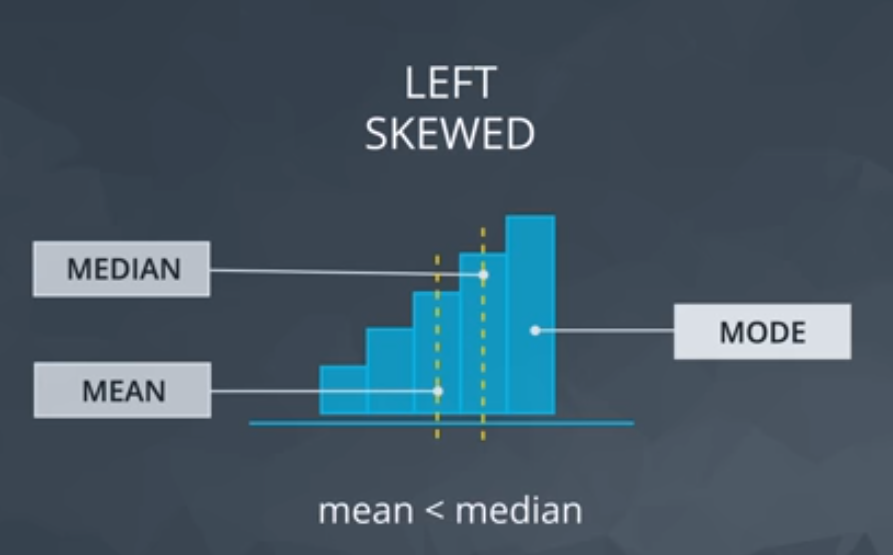
فى هذا المثال الانحراف المعيارى واحد لكن متوسط الربح اعلى فالاولى اى ان العائد اليك من الربح فى الاولى اعلى من الثانيه

**Shap**

****

****

****

****

**Outliers**

هى القيم التى تقع فى مجال بعييد عن باقى القيم

مقال اذا جمعنا بيانات رواتب رؤساء شركات عاديه ستكون متقاربه ثم وضعنا معهم راتب رائيس الفيس بوك ستكون معظم الرواتب متقاربه ماعدا راتب "مارك زجربرك " سيكون بعيد عن مقال باقى الداتا

القيم الخارجه outlier وتسمى ال

وتسأثر على المتوسط و الانحارف المعيارى بشكل كبير واذا كانت هذه القيم الشاذه اخطاء مطبيعه او اخطاء فى الادخال فيجب اذالتها او التعامل معها و وضع القيم الصحيحه فرقم واحد فقط قد يغير نتائجنا الى ارقام خاطئيه

الخطوات العمليه للتعامل مع الداتا

1. قم برسم الداتا على رسم بييانى
2. اذا كنت تملك قيم خارجه يجب ان تتعامل معاها سواء باذالتها او تعديلها
3. اذا كنت تعمل مع داتا بالتوزيع الاعتيادى يمكنك استخدام المتوسط والانحارف المعيارى لمعرفة كل شئ عن الداتا
4. اذا كنت تعمل مع داتا منحرفه الشكل الاعداد الخمس كما شرحنا فى الاعلى ستعطيك كل ما تحتاج من معلومات