



# المحاضرة السادسة: إدارة البيانات العلمية (استمرار ضمان البيانات)

د ماجد أحمد عسکر  
قسم علوم الحاسوب، كلية الحاسوبات والمعلومات

جامعة أسيوط  
[majid.askar@aun.edu.eg](mailto:majid.askar@aun.edu.eg)

# المزيد عن الرسومات R

• بوكسبلوتس

## ج) بوكسبلوتس

- يظهر الكثير من المعلومات حول متغير في قطعة

-الوسيط

-معدل الذكاء

-القيم المتطرفة

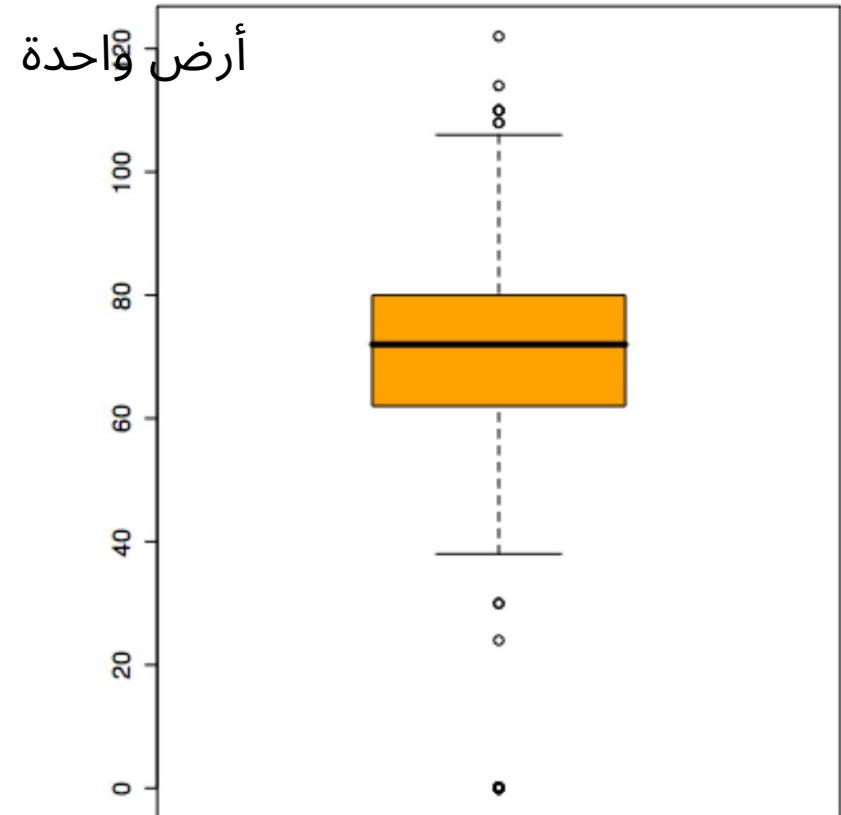
-يتراوح

-انحراف

• السلبيات

-الإفراط في التخطيط

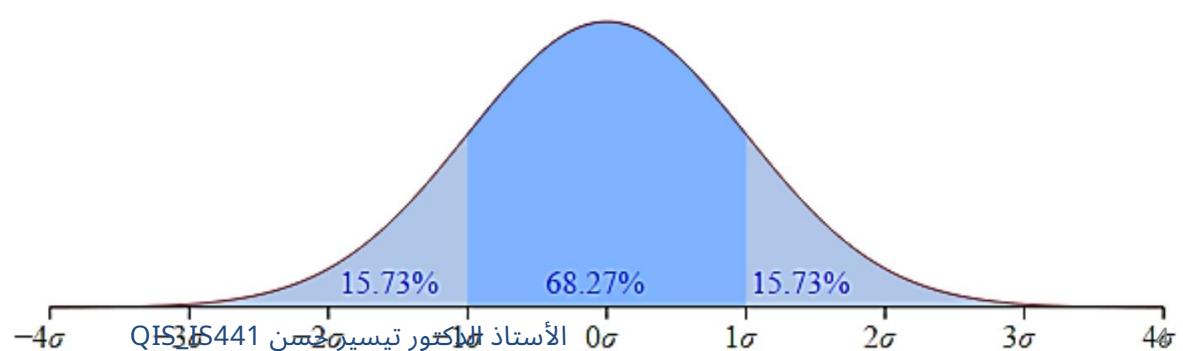
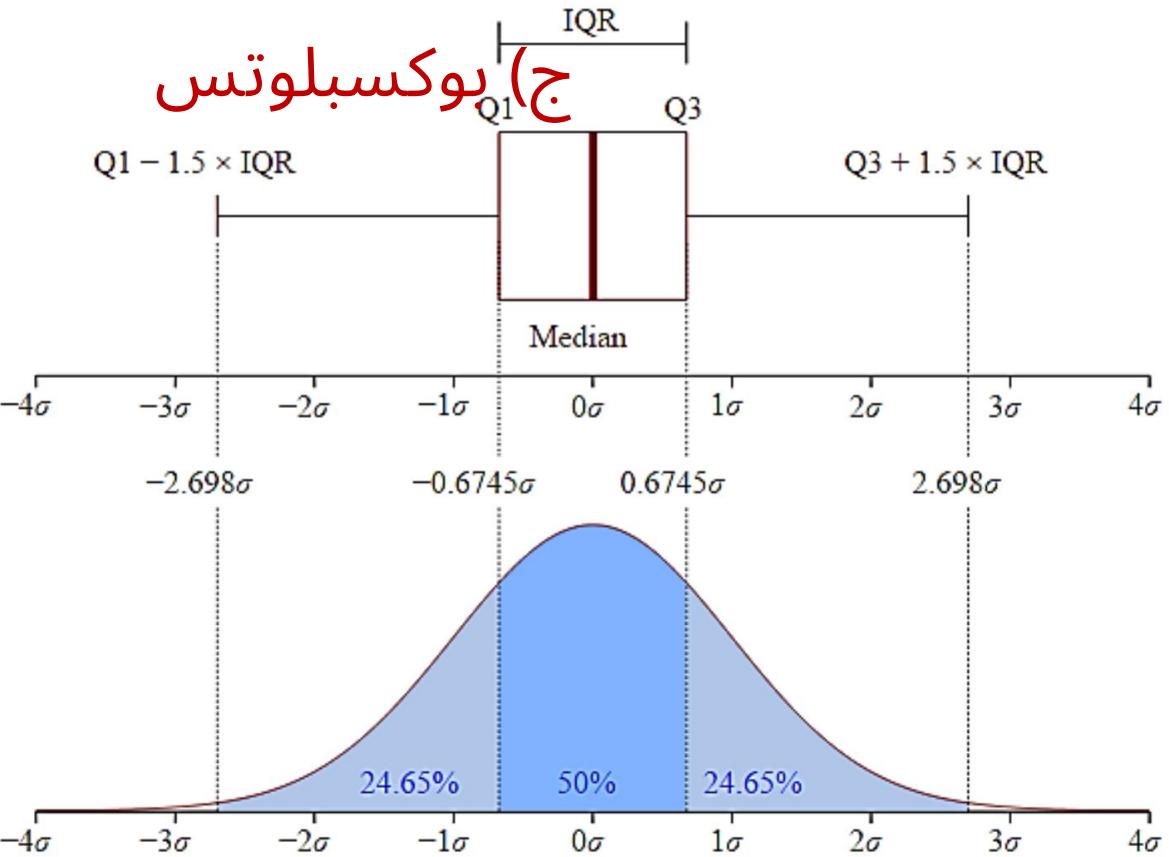
-من الصعب معرفة الشكل التوزيعي



الاحتمال والإحصاء Boxplot

دالة الكثافة (pdf) لـ  $N(0,1\sigma^2)$

ج) بوكسبلوتس



## ج) بوكسبلوتس

• يظهر الكثير من المعلومات حول متغير في قطعة

أرض واحدة

-الوسيط-

-معدل الذكاء (الرباعيات)

-القيم المتطرفة

-يتراوح

-انحراف-

# الربعيات



• هي القيم التي تقسم قائمة الأرقام

إلى أرباع:

- وضع قائمة الأرقام بالترتيب

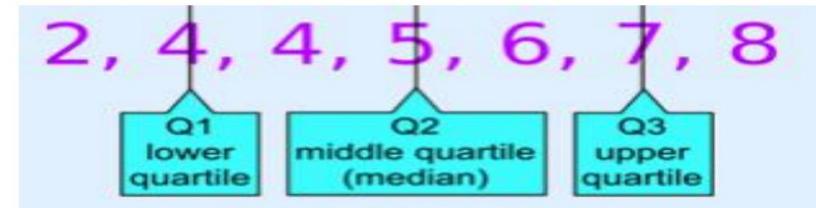
- ثم قم بتقسيم القائمة إلى أربعة أجزاء متساوية

- الربعيات في "التحفيضات"

• مثال: 5، 7، 4، 4، 6، 2، 8

• رتبهم بالترتيب: 2، 4، 4، 5، 6، 7، 8

• قسّم القائمة إلى أرباع:



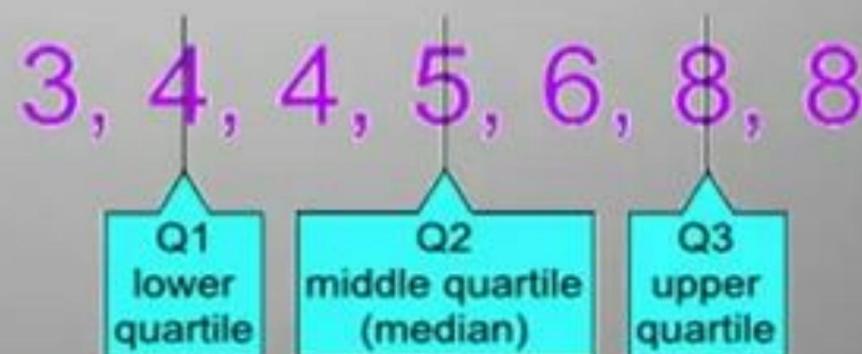
## والنتيجة هي

الربع  $\bullet = 4$  • (Q1) 1الربع ، (Q2) 2 وهو أيضاً الوسيط، • 5 =الربع 7

3 (Q3)

# How to Find Quartiles

- Put the data set in order from least to greatest
- Find the median of the data set, this is  $Q_2$
- Find the median of the values positioned before  $Q_2$ , this is  $Q_1$
- Find the median of the values positioned after  $Q_2$ , this is  $Q_3$



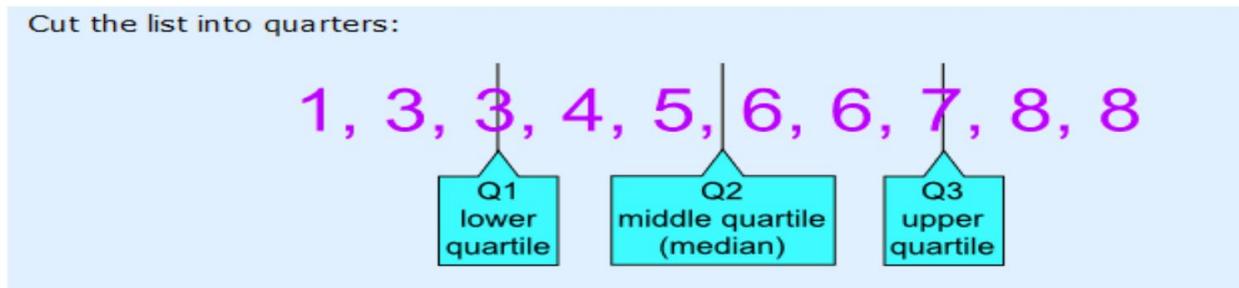
# ماذا عن "القطع" بين رقمين

• الربع هو متوسط الرقمين .

• مثال: 1، 3، 3، 4، 5، 6، 6، 7، 8، 8

• الأرقام بالترتيب بالفعل

• قسّم القائمة إلى أرباع:



في هذه الحالة، يقع الربع 2 في المنتصف بين 6 و 6.5:  $Q2 = (5+6)/2 = 5.5$

والنتيجة هي

• الربعى 1 (Q1) = 3 •

= 7 الربعى 2 (Q2) = 5.5 •

3 (Q3)

# مثال

5    6    } 12    13    } 15    18    } 22    50  
 ↓              ↓              ↓  
 Q<sub>1</sub>            Q<sub>2</sub>            Q<sub>3</sub>

$$Q_2 = \frac{13+15}{2} = \frac{28}{2} = 14$$

$$Q_1 = \frac{6+12}{2} = \frac{18}{2} = 9$$

$$Q_3 = \frac{18+22}{2} = \frac{40}{2} = 20$$

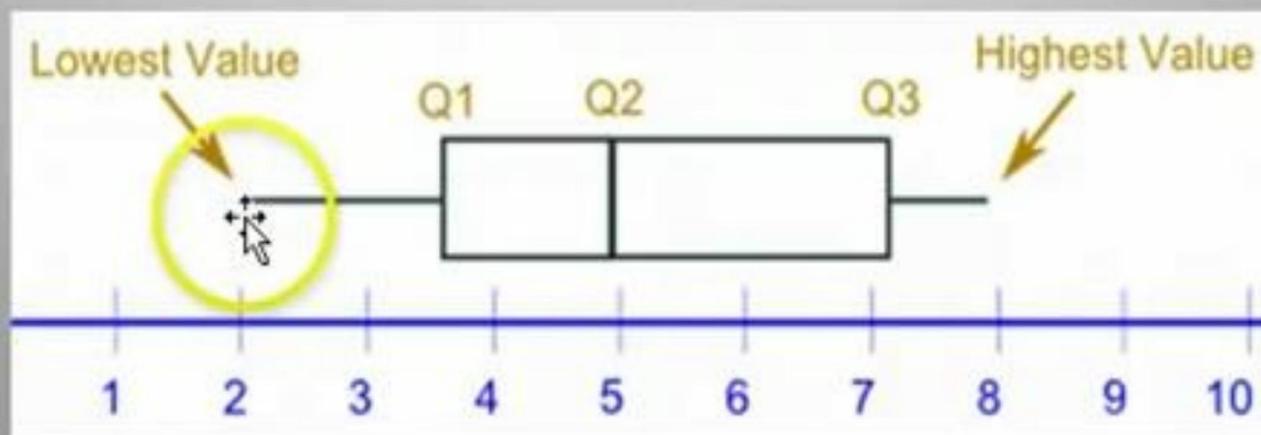
# 5 Number Summary

1. Minimum Value
2.  $Q_1$
3.  $Q_2$
4.  $Q_3$
5. Maximum Value

These can be displayed graphically as a **boxplot** (also called a box and whisker plot).

# How to Construct a Boxplot

- Determine the 5 Number Summary
- Arrange the 5 Number Summary on a number line in the following fashion



5    6    |    12    13    |    15    18    |    22    50  
 ↓              ↓              ↓  
 Q<sub>1</sub>            Q<sub>2</sub>            Q<sub>3</sub>

$$Q_2 = \frac{13+15}{2} = \frac{28}{2} = 14$$

$$Q_1 = \frac{6+12}{2} = \frac{18}{2} = 9$$

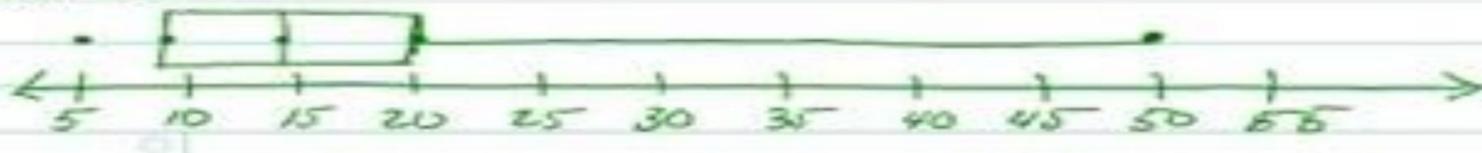
$$Q_3 = \frac{18+22}{2} = \frac{40}{2} = 20$$

$Q_1 = 9$
$Q_2 = 14$
$Q_3 = 20$

### Box Plot

### 5 number summary

1. Min = 5
2. Q<sub>1</sub> = 9
3. Q<sub>2</sub> = 14
4. Q<sub>3</sub> = 20
5. Max = 50



# Interquartile Range (IQR) and Outliers

- An **Outlier** is a data value that is *much smaller* or *much larger* than the other values in the data set.
- $IQR = Q_3 - Q_1$
- Test for Outliers:
  1. Find IQR
  2. Multiply  $1.5(IQR)$
  3. Subtract  $Q_1 - 1.5(IQR)$
  4. Add  $Q_3 + 1.5(IQR)$
  5. Any value less than the value in step 3 or more than the value in step 4 is an **outlier**.

5 6 12 13 15 18 22 50

$$IQR = Q_3 - Q_1 = 20 - 9 = 11$$

$$1.5(11) = 16.5$$

$$Q_1 - 16.5 = 9 - 16.5 = -7.5$$

$$Q_3 + 16.5 = 20 + 16.5 = 36.5$$

anything outside of the interval

-7.5 to 36.5

is an outlier

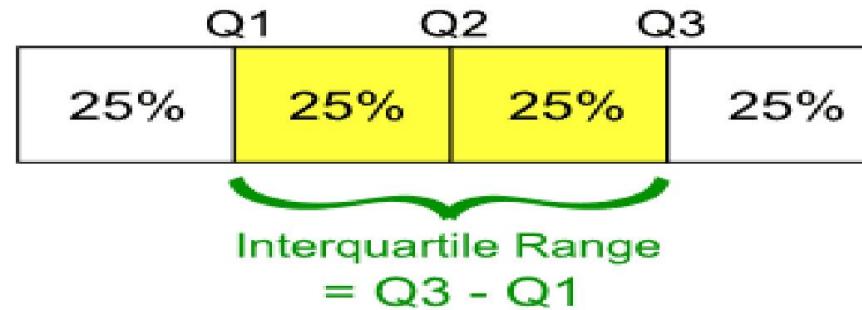
50 is larger than 36.5

50 is an outlier!

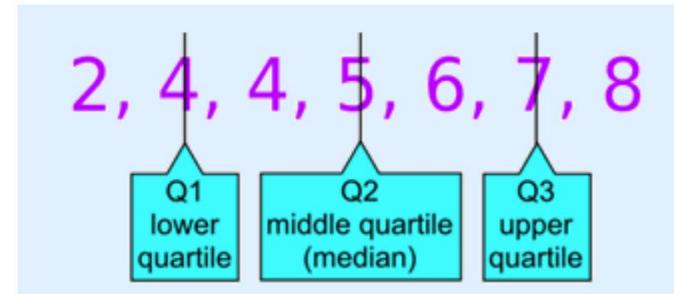
↑  
outlier?

# النطاق الرباعي

• "المدى الرباعي" هو من Q1 إلى Q3:



لحساب ذلك، فقط قم بطرح الرابع من الرابع ، مثل هذا:



$$\text{المدى الرباعي} = Q_3 - Q_1 = 7 - 4 = 3$$

# ج) بوكس بلوت

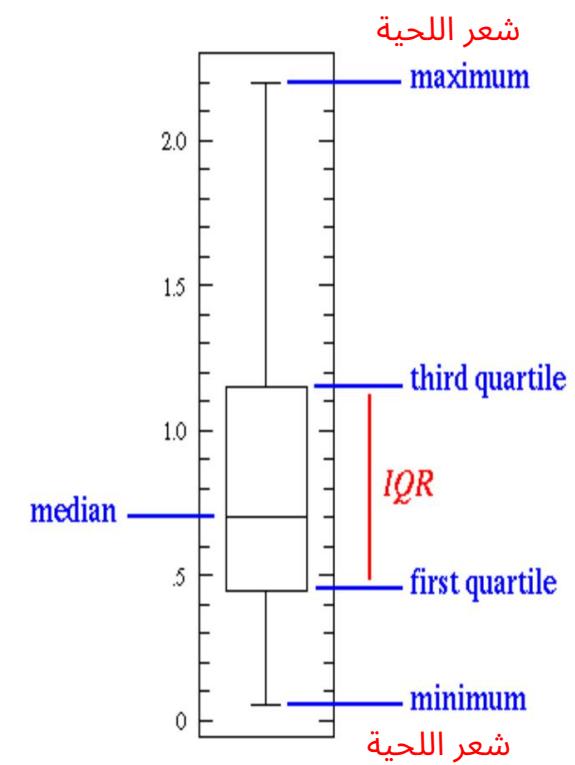
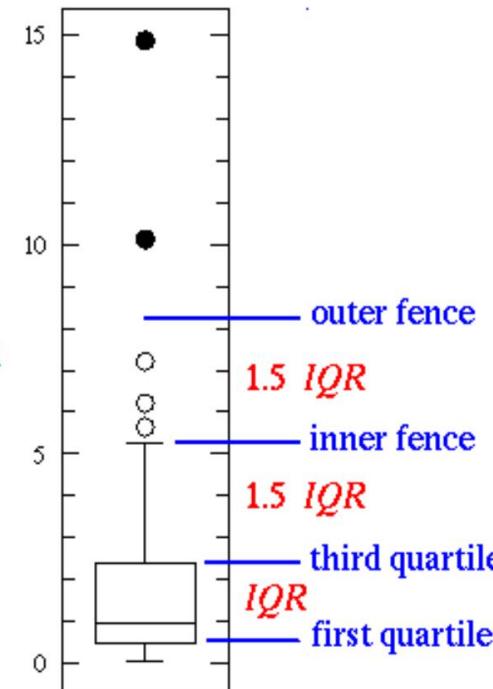
- يظهر الكثير من المعلومات حول متغير في قطعة أرض واحدة
- الوسيط

-معدل الذكاء (الربعيات)  
القيم المتطرفة  
*outliers*

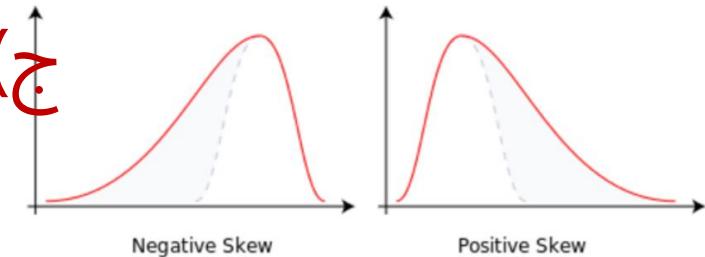
- يتراوح  
انحراف

*suspected outliers*

 القيم المتطرفة بشكل افتراضي هي الملاحظات  $(Q1 - 1.5 \times IQR)$  \* أكثر من  $(Q3 + 1.5 \times IQR)$  بعيدة عن الربع الأقرب



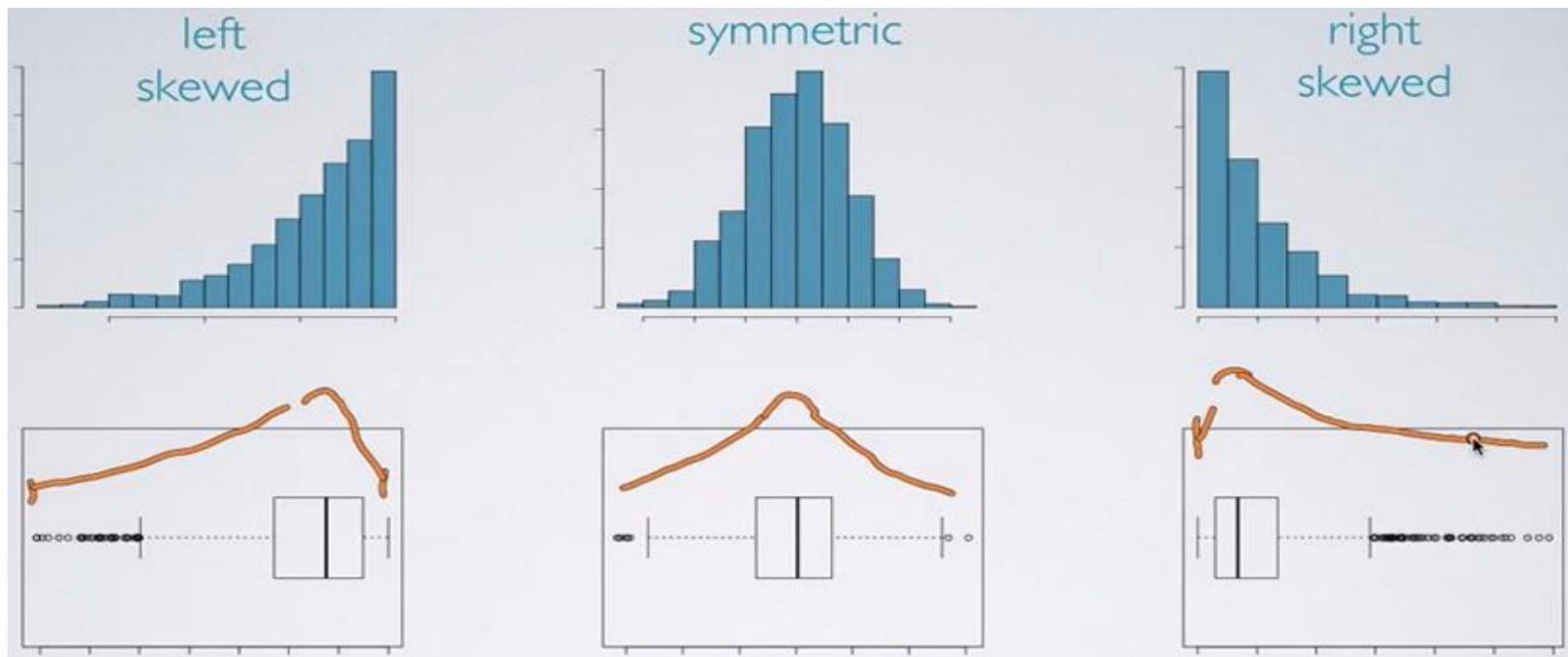
## ج) بوكسبلوتس



•توزيع الانحراف

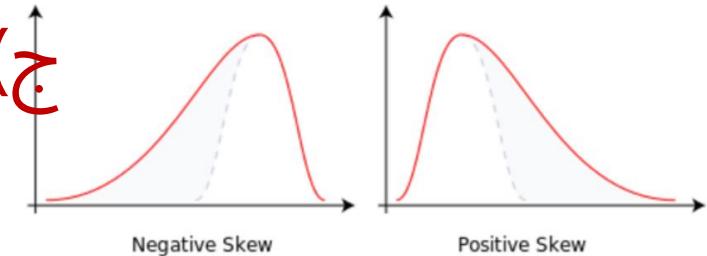
-انحراف سلبي -انحراف يسار -انحراف إيجابي -انحراف

يمين (ذيل مكتوب)



## ج) بوكسبلوتس

• توزيع الانحراف



-انحراف سلبي -انحراف يسار -انحراف إيجابي -انحراف

يمين (ذيل مكتوب)

### Normal Distribution

$(\text{Quartile 3} - \text{Quartile 2}) = (\text{Quartile 2} - \text{Quartile 1})$



يمكن حساب الانحراف على القيم المتوسطة

### Positive Skew

$(\text{Quartile 3} - \text{Quartile 2}) > (\text{Quartile 2} - \text{Quartile 1})$



### Negative Skew

$(\text{Quartile 3} - \text{Quartile 2}) < (\text{Quartile 2} - \text{Quartile 1})$



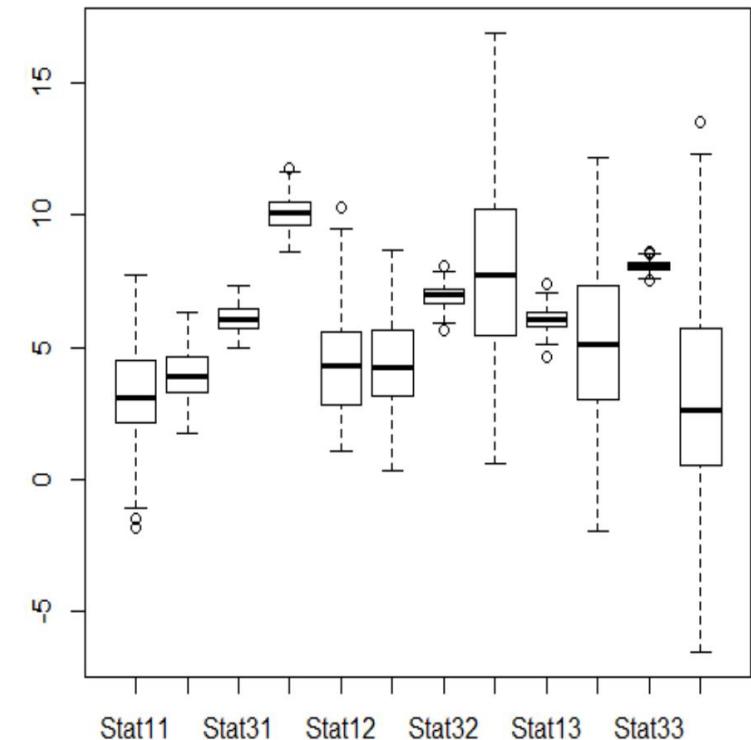
# R Box مخططات

## • مخططات الصندوق • مثال:

- `Stat21=rnorm(100)` .
- `data<-data.frame(Stat11=rnorm(100,mean=3,sd=2),`  
`sd = 1),` • Stat31 = rnorm (100, يعني `rnorm(n=100,mean=3,sd=1)`  
`= rnorm (100, sd = 0.5).` • Stat41 = rnorm (100, يعني `= 4.`  
`Stat43=rnorm(100,mean=4,sd=4))` يعني `(10, sd = 0.5).` • Stat12

- `Stat33=rnorm(100,mean=8,sd =0.2),`  
`Stat13=rnorm(100,mean=6,sd=0.5),` • Stat23=rnorm(100,mean=5,sd=3),  
`Stat42=rnorm(100,mean=8 ,sd=3),` •  
`Stat32=rnorm(100,mean=7,sd=0.5),`  
`=4,sd=2),` • Stat22=rnorm(100,mean=4.5,sd=2),

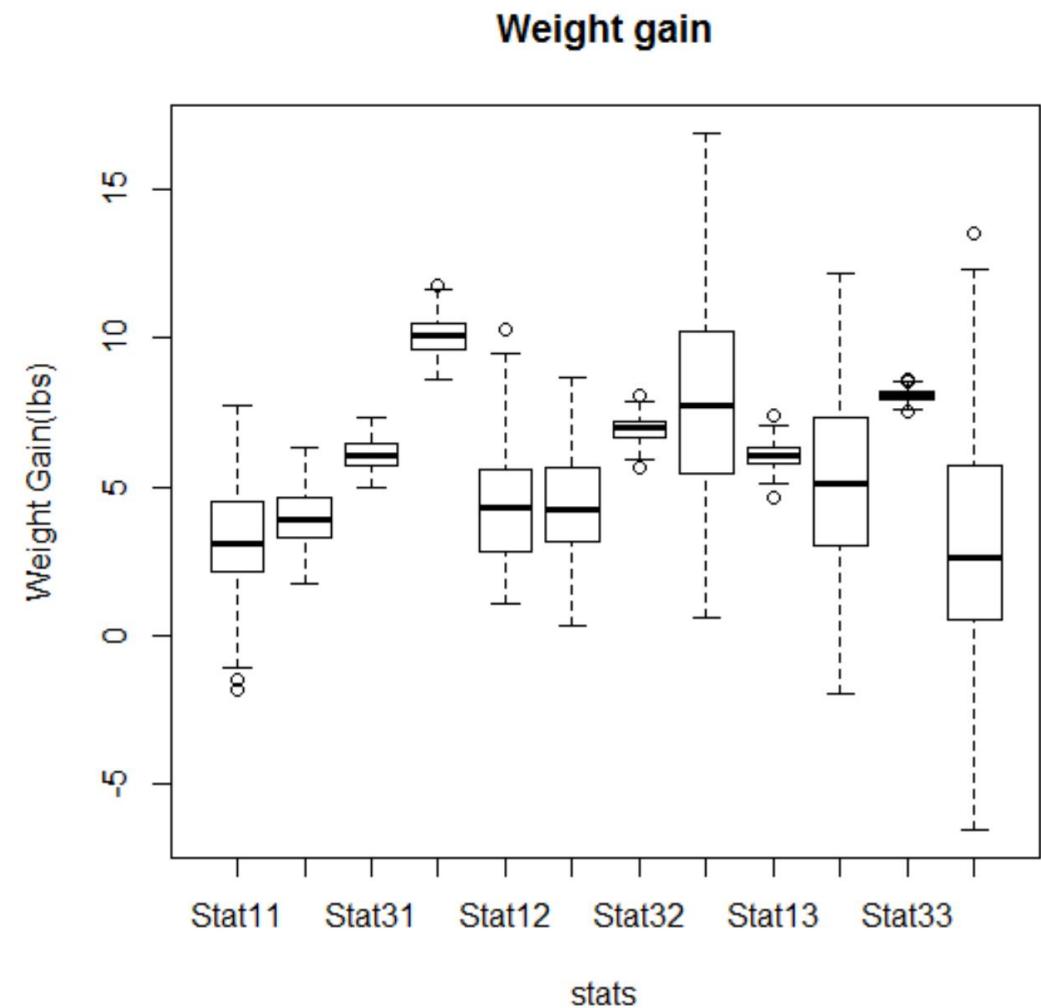
• بوكسيلوت (البيانات)



# بوكسلوتس

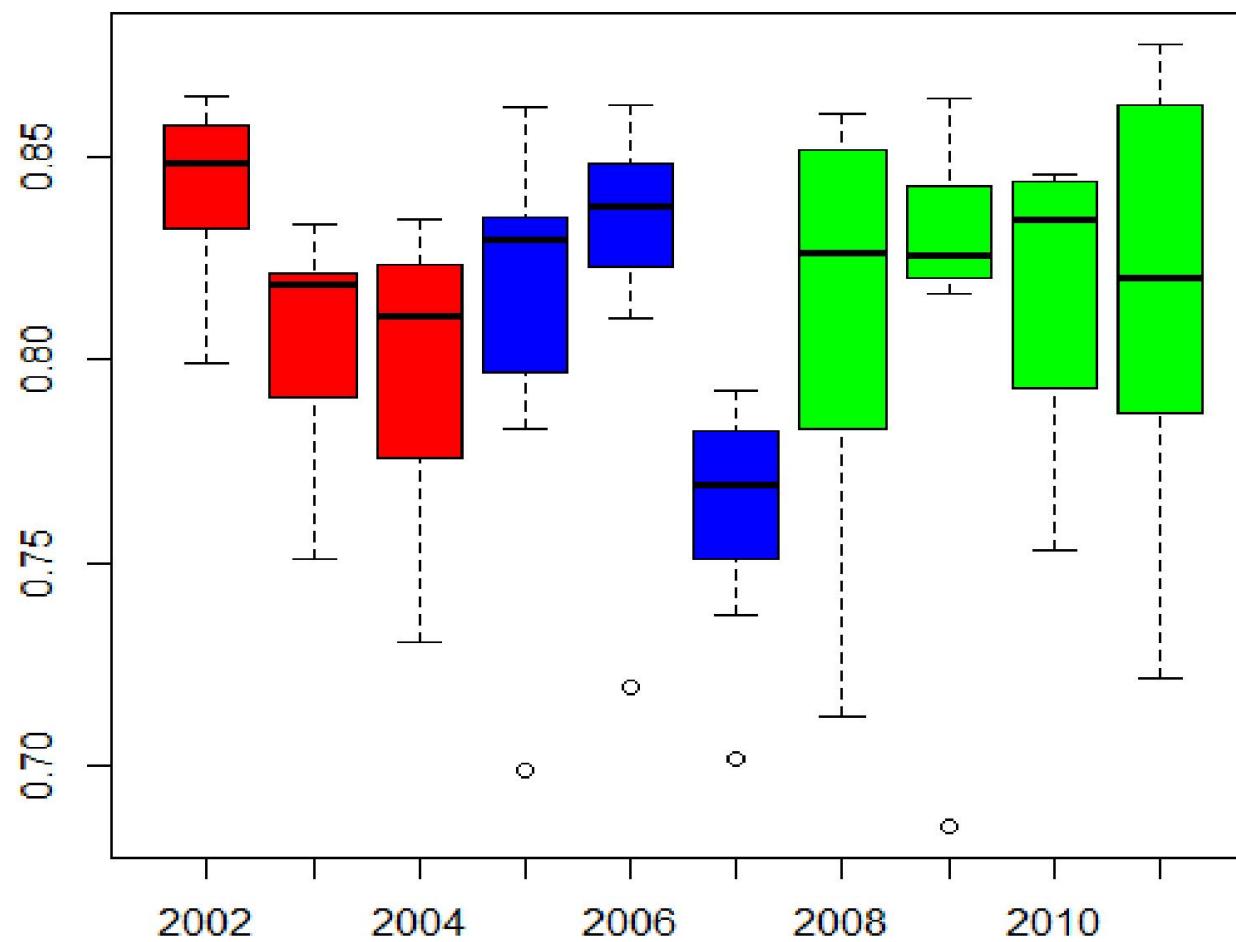
تغیره!

'زيادة الوزن'=niam,atad)tolpxob  
'زيادة الوزن (رطل)'=baly  
'إحصائيات'=balx



# تلوين

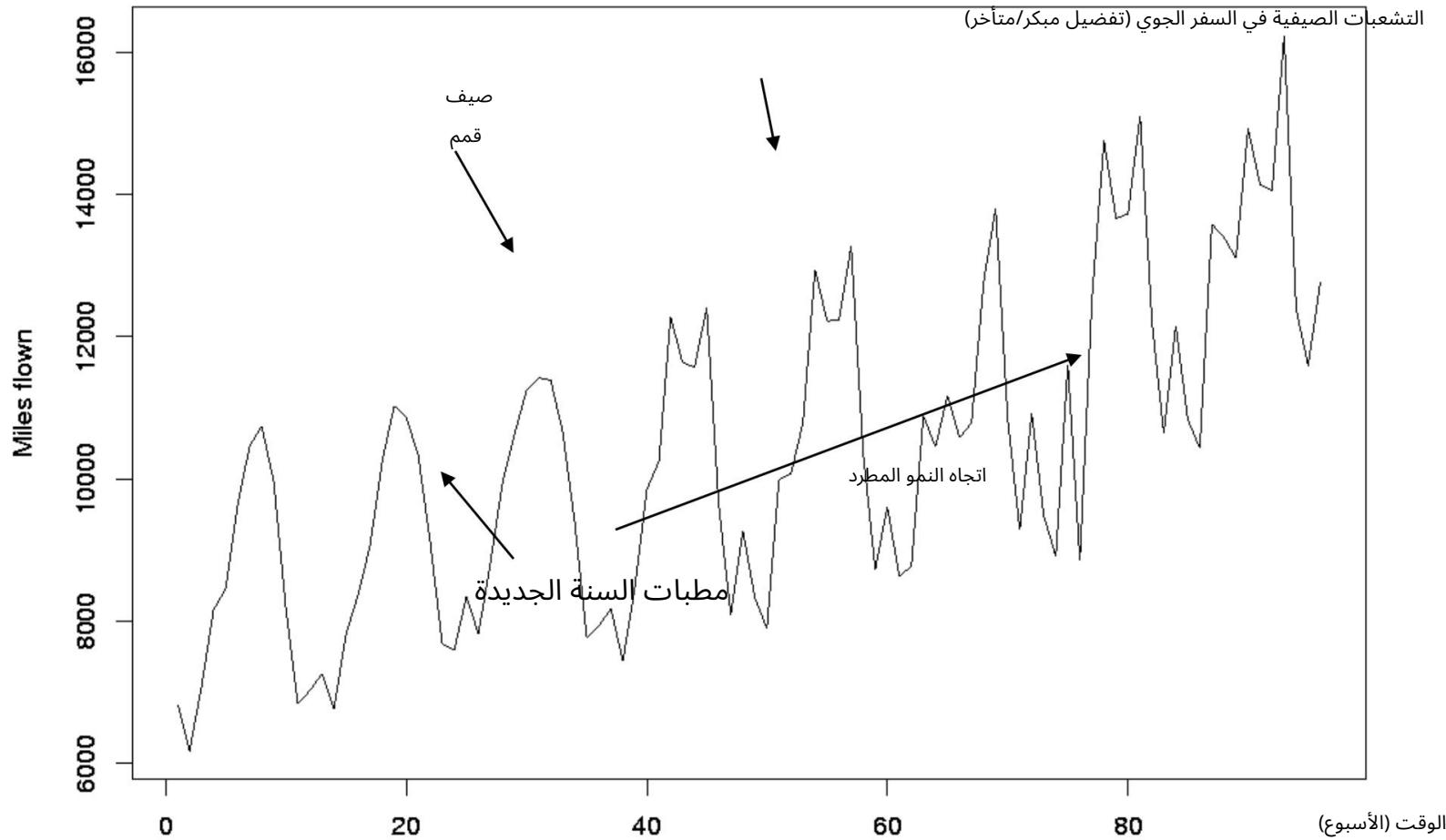
```
0.8401 0.8426 0.7777 0.8608 0.8243 0.8426 0.8602'), header=TRUE) • colors =  
0.8009 0.6988 0.719 0.7645 0.8142 0.8164 0.7529 0.7214 0.8483 0.8159 0.833  
0.7653 0.7512 0.8106 0.8354 0.7013 0.752 0.8256 0.8235 0.7579 0.8572 0.8186  
0.8651 0.8211 0.8346 0.8624 0.8629 0.7926 0.8543 0.8648 0.8456 0.8656 0.7991  
0.8159 0.8584 0.8217 0.8105 0.8295 0.8382 0.769 0.8265 0.8287 0.8344 0.8204  
0.8459 0.8778 0.8251 0.7509 0.7301 0.7829 0.8103 0.7368 0.7122 0.685 0.7623  
2008 2009 2010 2011 0.8392 0.8337 0.8145 0.8306 0.8545 0.7876 0.8491 0.8568  
• df <- read.table(textConnection('2002 2003 2004 2005 2006 2007  
ج(4)"أحمر",3,"أزرق",4)"أخضر",3))per,df))per  
• boxplot(df,names=2002:2011,col=colors)
```



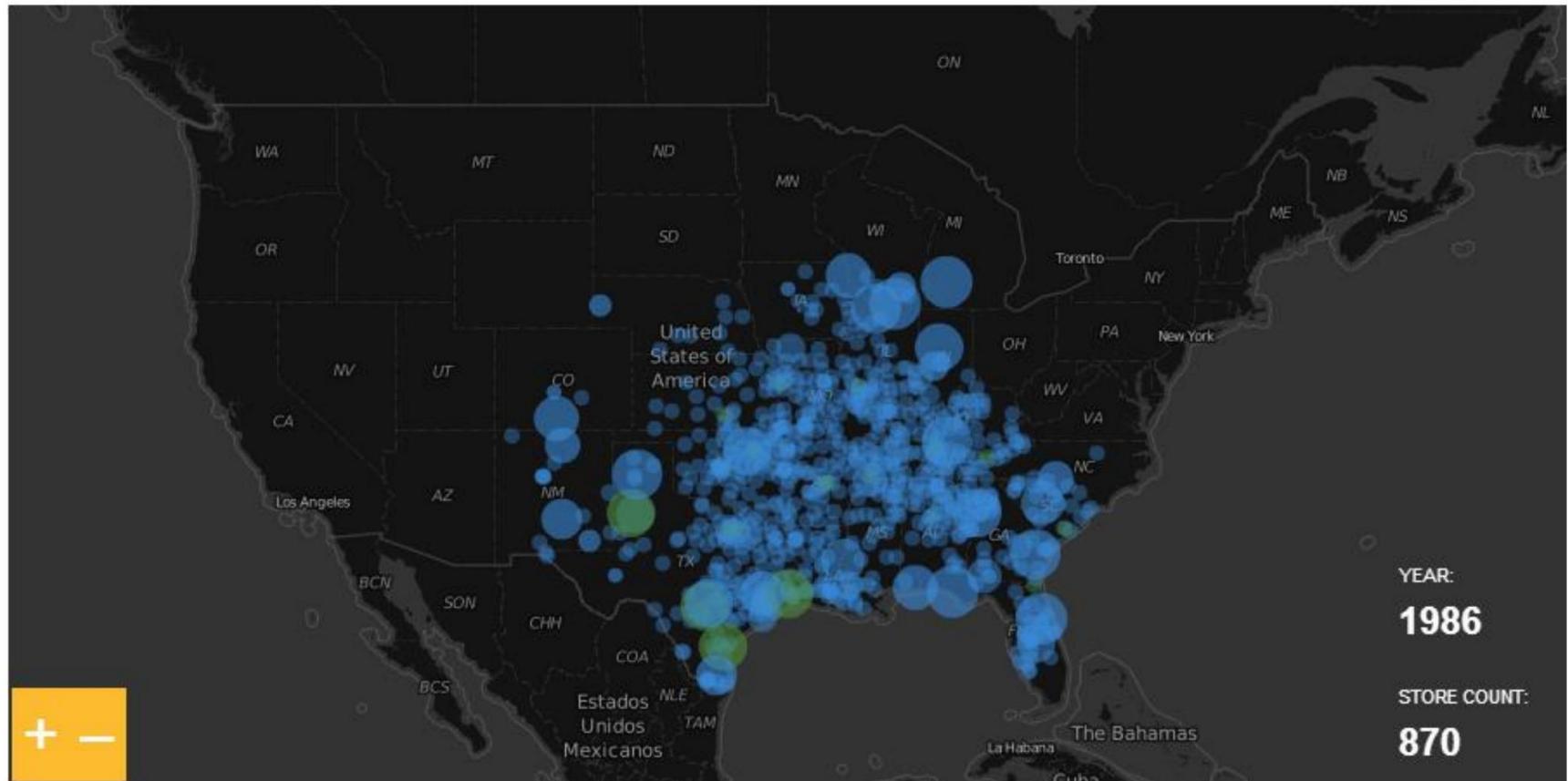
## د) السلسل الزمنية



إذا كانت بياناتك تحتوي على مكون مؤقت، فتأكد من استغلاله



## (\*)البيانات المكانية والزمانية



على سبيل المثال، نمو السوق بمرور الوقت [projects.flowingdata.com/walmart/](http://projects.flowingdata.com/walmart/)

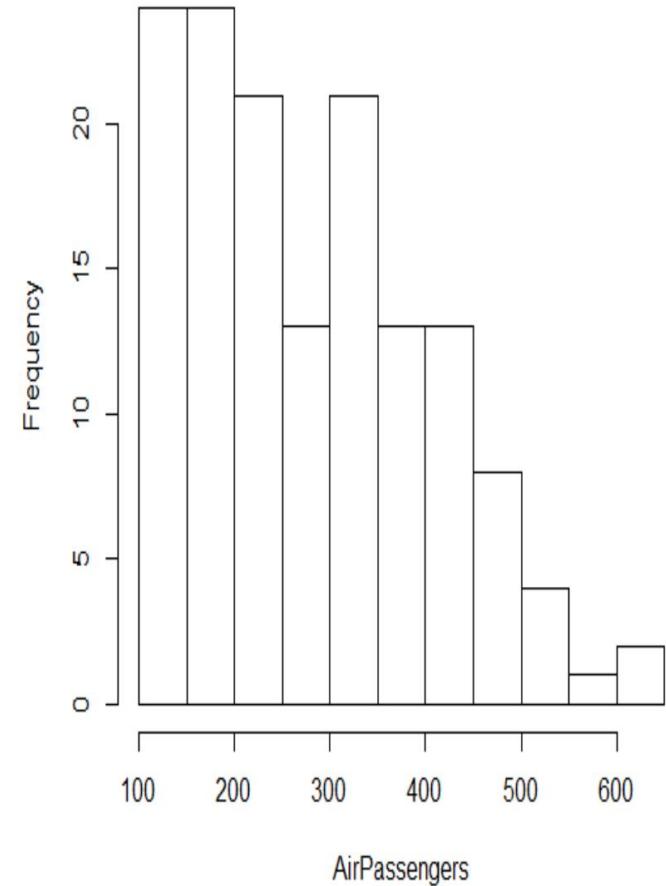
يمكنك فقط القيام بمخططات مبعثرة متتالية للحصول على نفس التأثير <http://>



# R رسومات

رسم بياني - احصمت (AirPassengers)

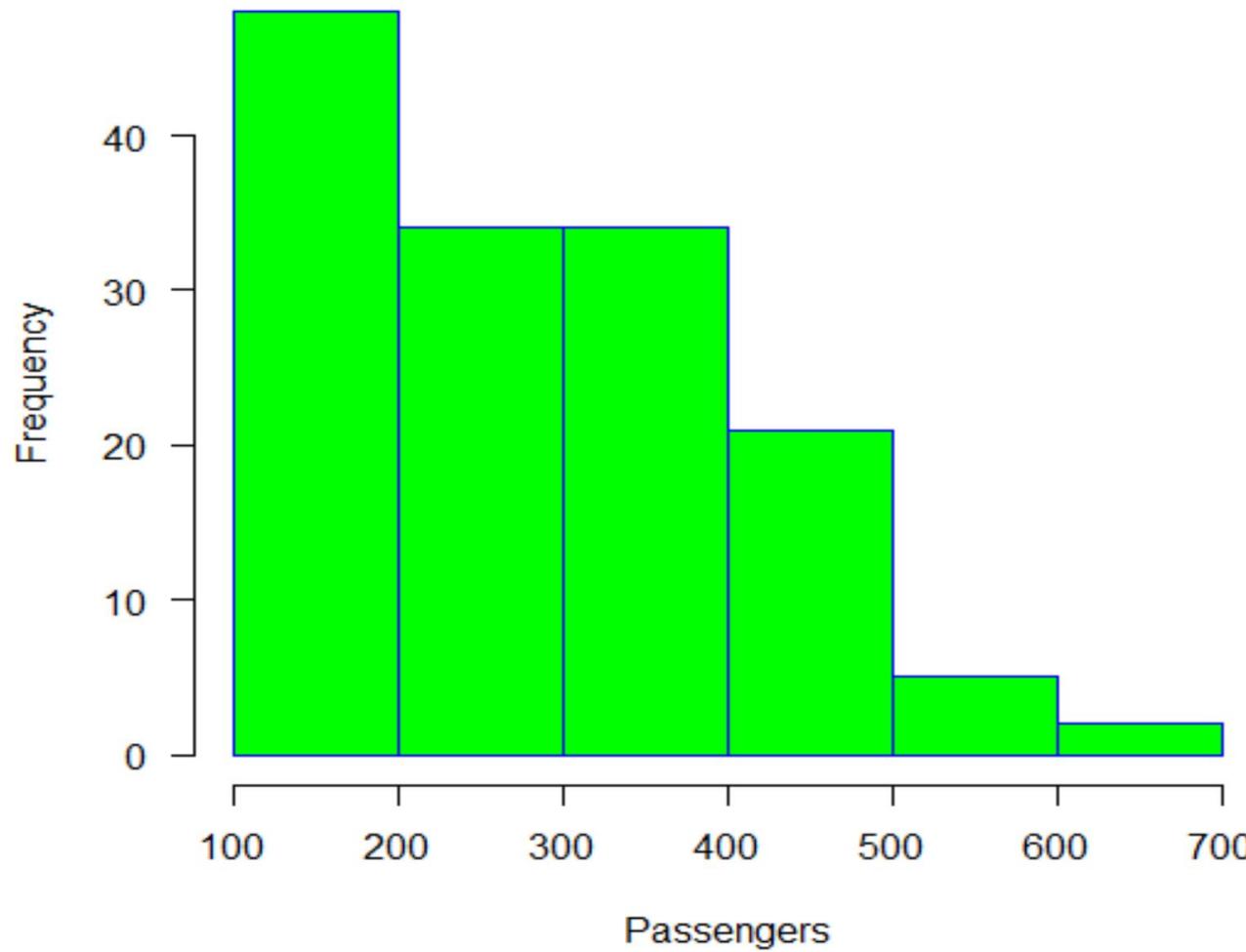
Histogram of AirPassengers



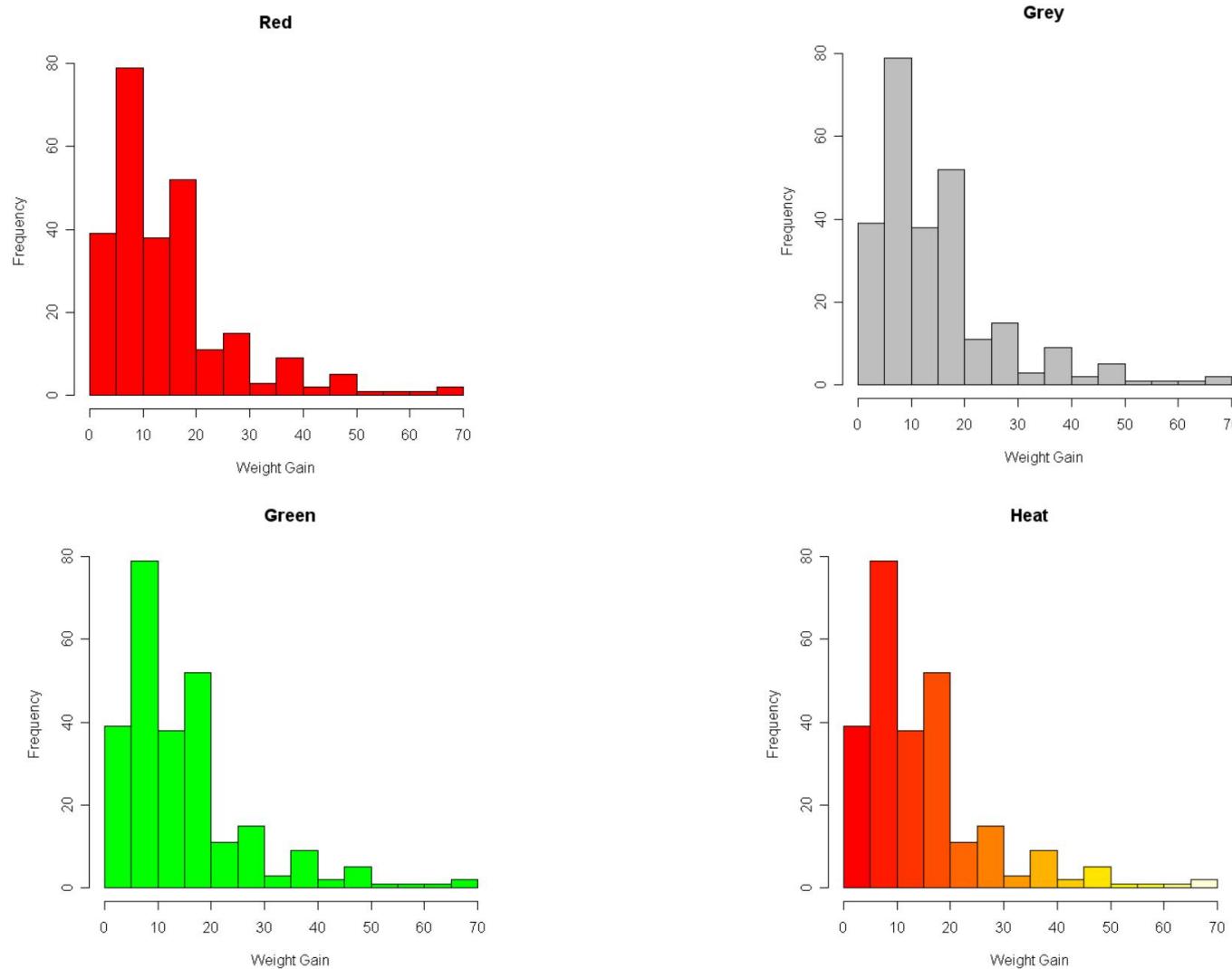
## إضافة العناوين والألوان

```
• الرسم البياني للركاب = niam • hist(AirPassengers, • +  
• + border="blue", • + col="green", " • الجويين • +  
• + xlab="Passengers",  
• + xlim=c(100,700),  
• + las=1, • فواصل=5)
```

## Histogram for Air Passengers



# رسومات - IIR الألوان



## الوظائف المشتركة لإنشاء المؤامرات



• مؤامرة، boxplot، barplot، النجوم، أزواج، matplot.

اصمت، الصورة، كفاف.

تعد كل وظيفة من وظائف الرسم هذه جزءاً من الحزمة الأساسية، ولمعرفة التفاصيل الخاصة بكل وظيفة، يرجى الرجوع إلى صفحات المساعدة (باستخدام وظيفة المساعدة، على سبيل المثال >help(function\_name) ).

## وظيفة المؤامرة

نبدأ الحبكة بإخبار الرسومات أولاً

النظام أننا على وشك بدء مؤامرة جديدة. • `plot.new()` . • هذا يدل على **أننا على وشك البدء** بمشروع جديد

مؤامرة ويجب أن يحدث قبل حدوث أي رسومات.

• يختار استدعاء `Plot.new()` منطقة رسم مستطيلة افتراضية لظهور قطعة الأرض فيها. • منطقة الرسم محاطة بأربعة هوامش.

MACHINE TRANSLATED BY GOOGLE

File Edit View Misc Packages Windows Help

R Console R Graphics: Device 2 (ACTIVE)

```
tried_all_packages = tried_all_packages, type = help_type)
class(paths) <- "help_files_with_topic"
paths
}
<bytecode: 0x00000000f7ced68>
<environment: namespace:utils>
> summary(trees)
   Girth      Height      Volume
Min. : 8.30  Min. :63  Min. :10.20
1st Qu.:11.05 1st Qu.:72  1st Qu.:19.40
Median :12.90 Median :76  Median :24.20
Mean   :13.25 Mean   :76  Mean   :30.17
3rd Qu.:15.25 3rd Qu.:80  3rd Qu.:37.30
Max.   :20.60 Max.   :87  Max.   :77.00
> BA(X)
Error: could not find function "BA"
> abs(x0
+ abs(x)
Error: unexpected symbol in:
"abs(x0
abs"
> plot.new()
>
> |
```



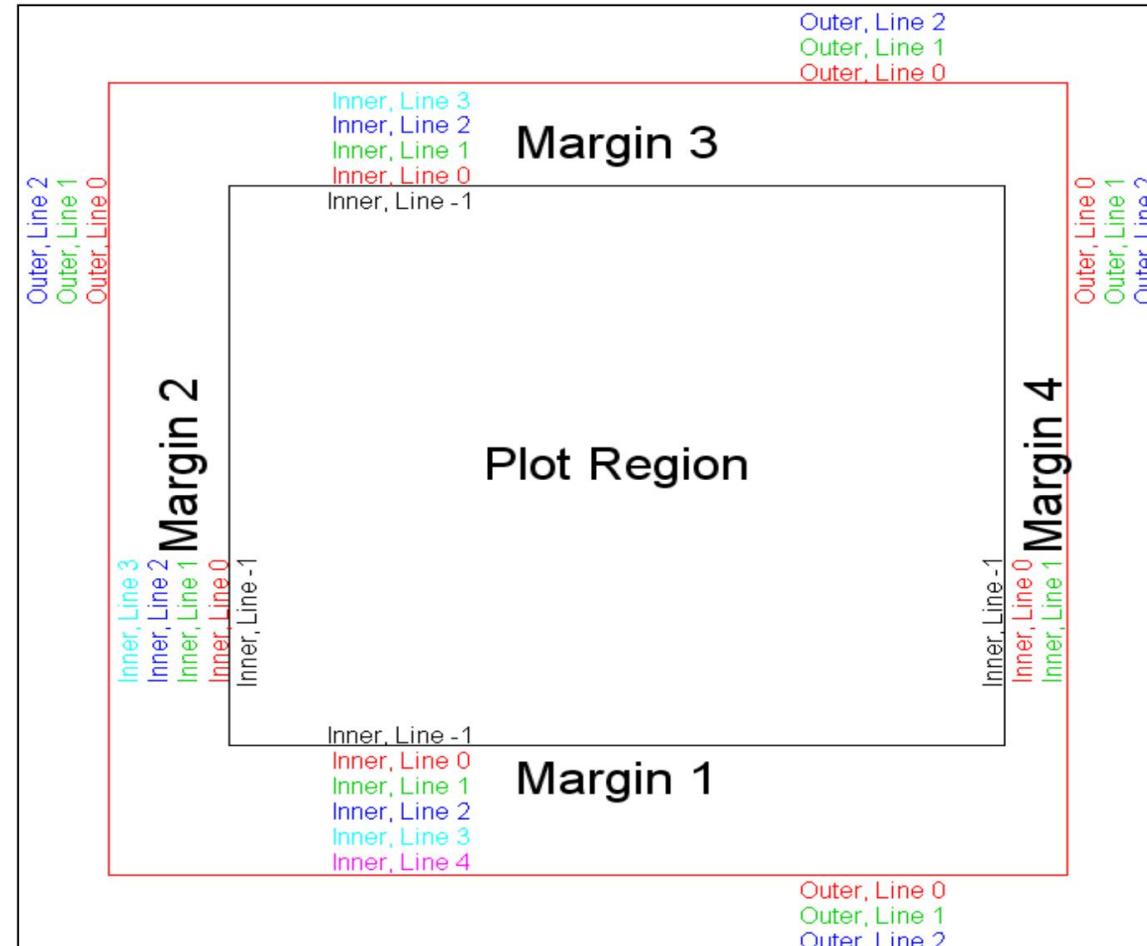
• يقوم `mtext()` بإضافة نص في هوامش

## الأشكال

• يتم ترقيم الهوامش 1:4 باتجاه عقارب الساعة من الأسفل

• تتم طباعة النص بعدد محدد من الأسطر بعيداً عن حافة منطقة الرسم

• يمكن الطباعة في الخارج  
الهوامش إذا كانت متوفرة

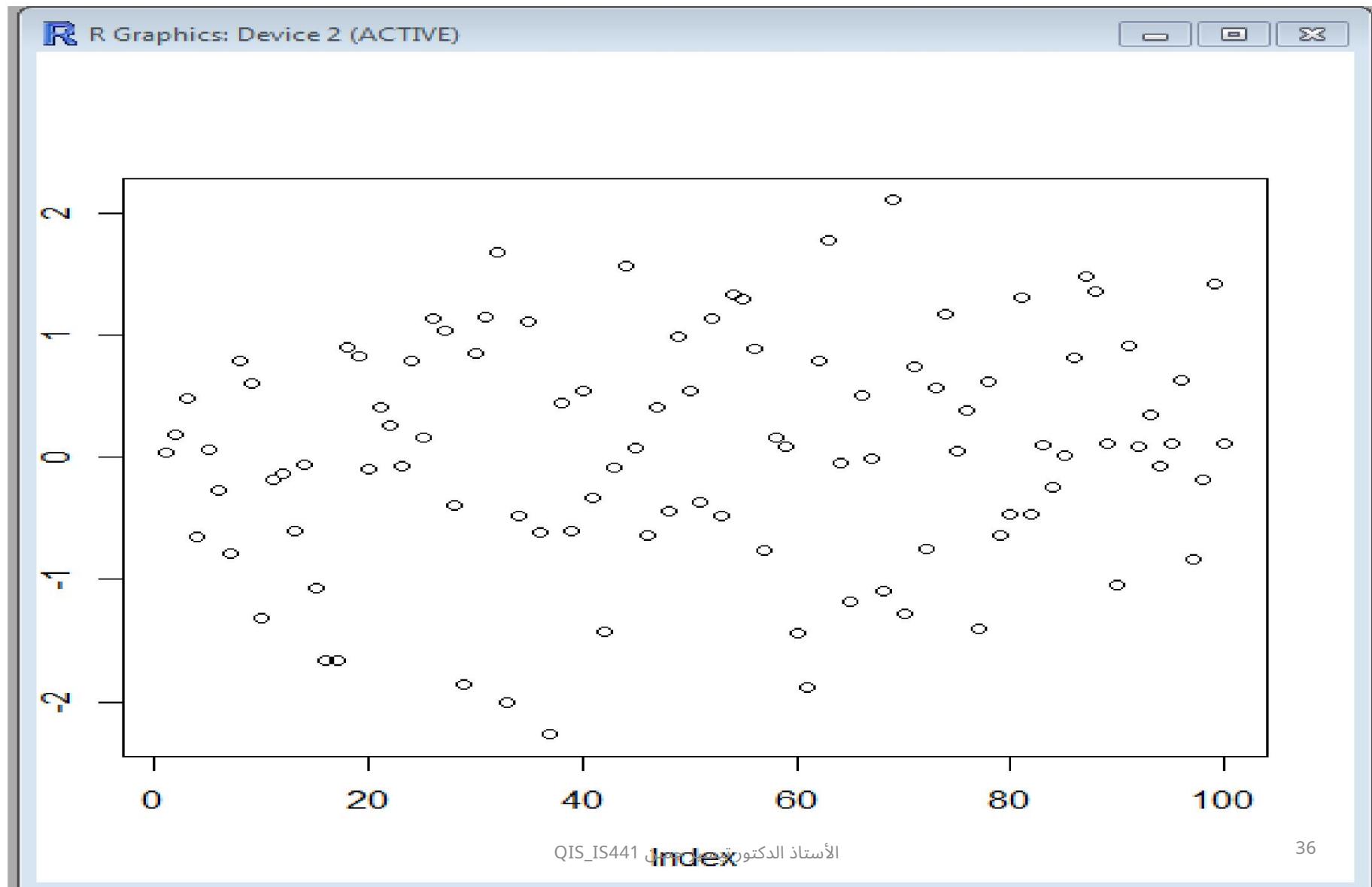




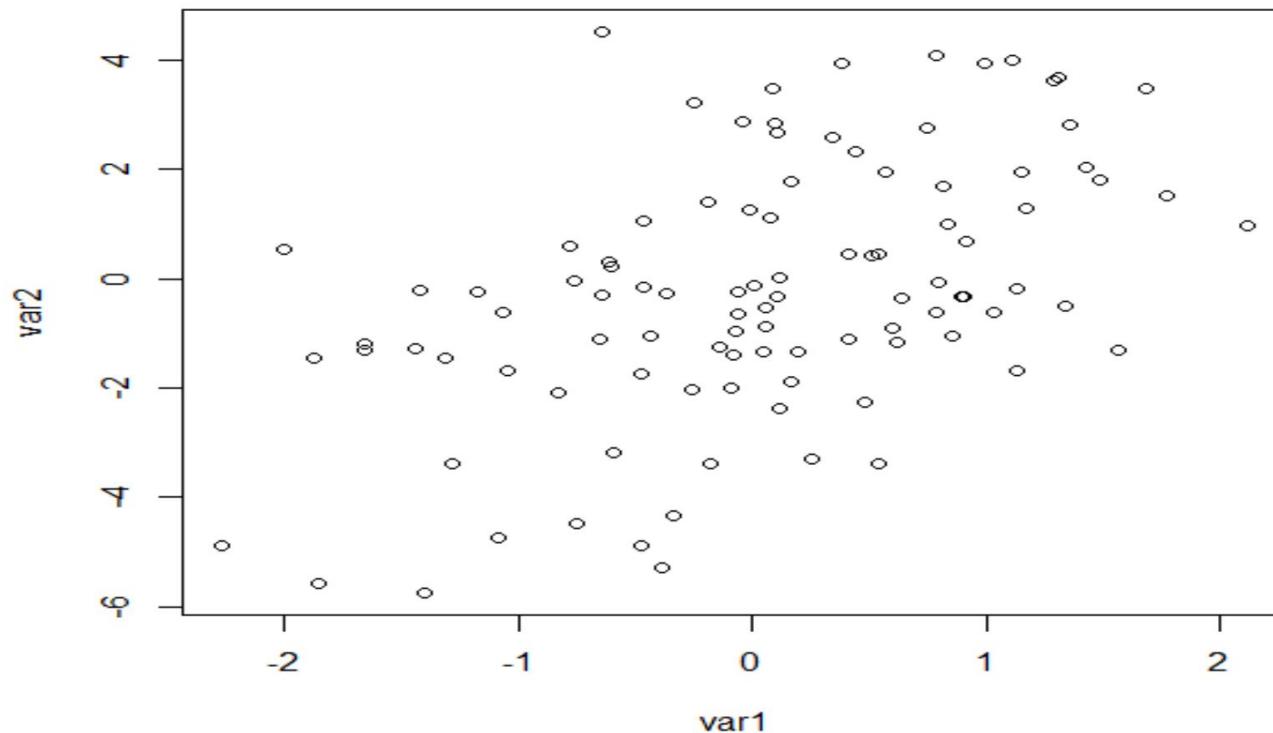
• وظيفة الرسم هي وظيفة الرسم الأساسية.  
سيؤدي استخدام دالة مؤامرة مع متغير واحد فقط إلى رسم فهرس. • سيؤدي استخدام متغيرين في دالة الرسم إلى رسم بياني حيث يتم رسم الوسيطة الأولى كمتغير على المحور  $x$  والوسيلة الثانية باعتبارها المتغير المرسوم على المحور . وفي المثال التالي، قمنا بإنشاء متغيرين، var1 و var2، وقمنا برسمهما بيانيًا باستخدام المخطط والدوال .

•

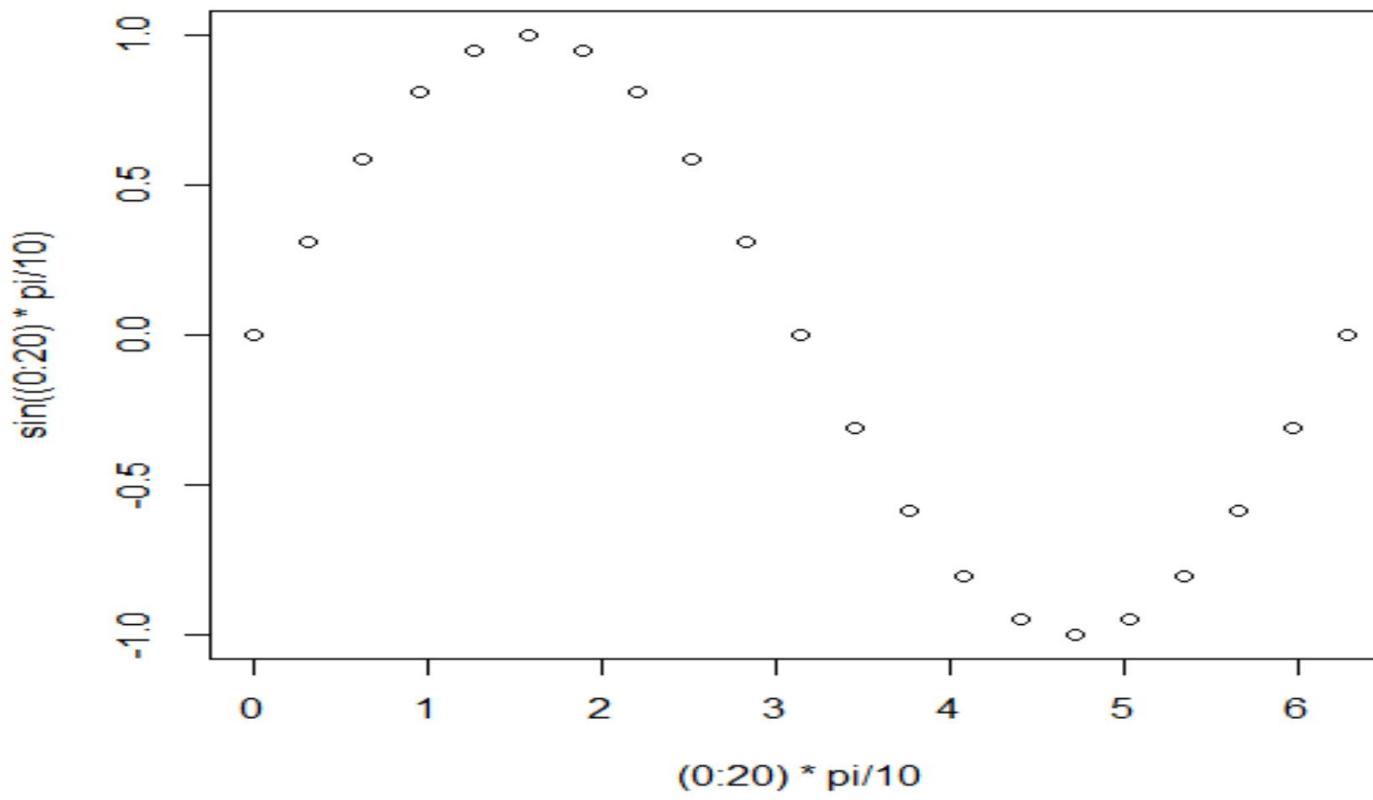
(var1) **مخطّط var1 <-  
rnnorm (100)**



var2 مارمورة (var3<100) var1\*.



# مؤامرة ((0:20)\*بأي/، 10، خطيبة((02:0)\*بأي/(10))



# السيطرة على الهوامش



- يمكن تغيير أحجام الهوامش حسب إجراء مكالمه إلى الدالة `par(wen.tolp)` قبل استدعاء `wen`. • ضبط أحجام الهوامش بالبوصة.

<الاسمية(ماي=ج(2, 1, 1))>

ضبط أحجام الهوامش في أسطر النص.

<اضبط عرض القطعة وارتفاعها بالبوصة> `par(mar=c(4, 4, 2, 2))`

<الاسمية (دبوس = ج ((5, 4)

# High-Level Plot Functions

---

<code>plot()</code>	Scatterplot
<code>hist()</code>	Histogram
<code>boxplot()</code>	Boxplot
<code>qqplot()</code> , <code>qnorm()</code> , <code>qline()</code>	Quantile plots
<code>interaction.plot()</code>	Interaction plot
<code>sunflowerplot()</code>	Sunflower scatterplot
<code>pairs()</code>	Scatter plot matrix
<code>symbols()</code>	Draw symbols on a plot
<code>dotchart()</code> , <code>barplot()</code> , <code>pie()</code>	Dot chart, bar chart, pie chart
<code>curve()</code>	Draw a curve from a given function
<code>image()</code>	Create a grid of colored rectangles with colors based on the values of a third variable
<code>contour()</code> , <code>filled.contour()</code>	Contour plot
<code>persp()</code>	Plot 3-D surface

---

# Low-Level Plot Functions

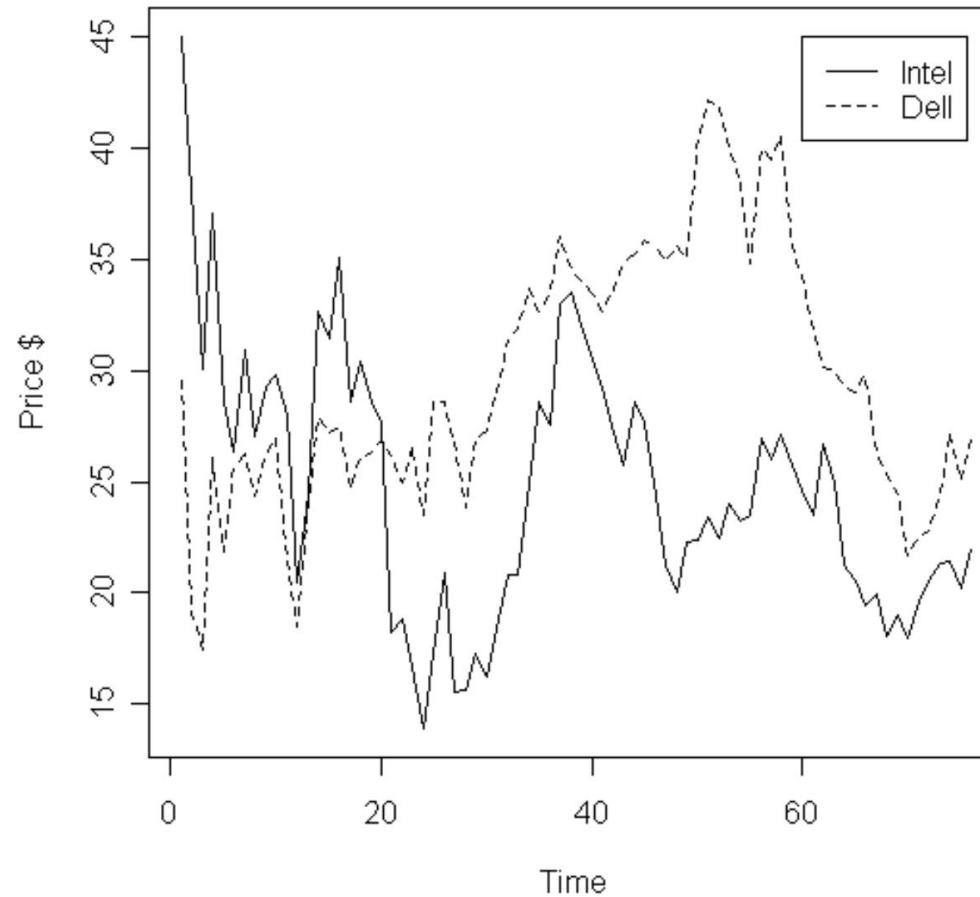
---

points()	Add points to a figure
lines()	Add lines to a figure
text()	Insert text in the plot region
mtext()	Insert text in the figure and outer margins
title()	Add figure title or outer title
legend()	Insert legend
axis(), axis.Date()	Customize axes
abline()	Add horizontal and vertical lines or a single line
box()	Draw a box around the current plot
rug()	Add a 1-D plot of the data to the figure
polygon()	Draw a polygon
rect()	Draw a rectangle
arrows()	Draw arrows
segments()	Draw line segments
trans3d()	Add 2-D components to a 3-D plot

---

# إضافة أسطورة

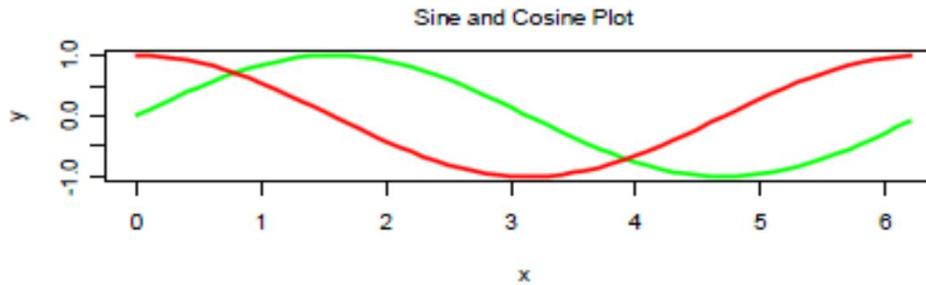
**Closing Stock Prices**



```
legend(60,45,c('Intel','Dell'),lty=c(1,2))
```

## Plotting Two Data Series

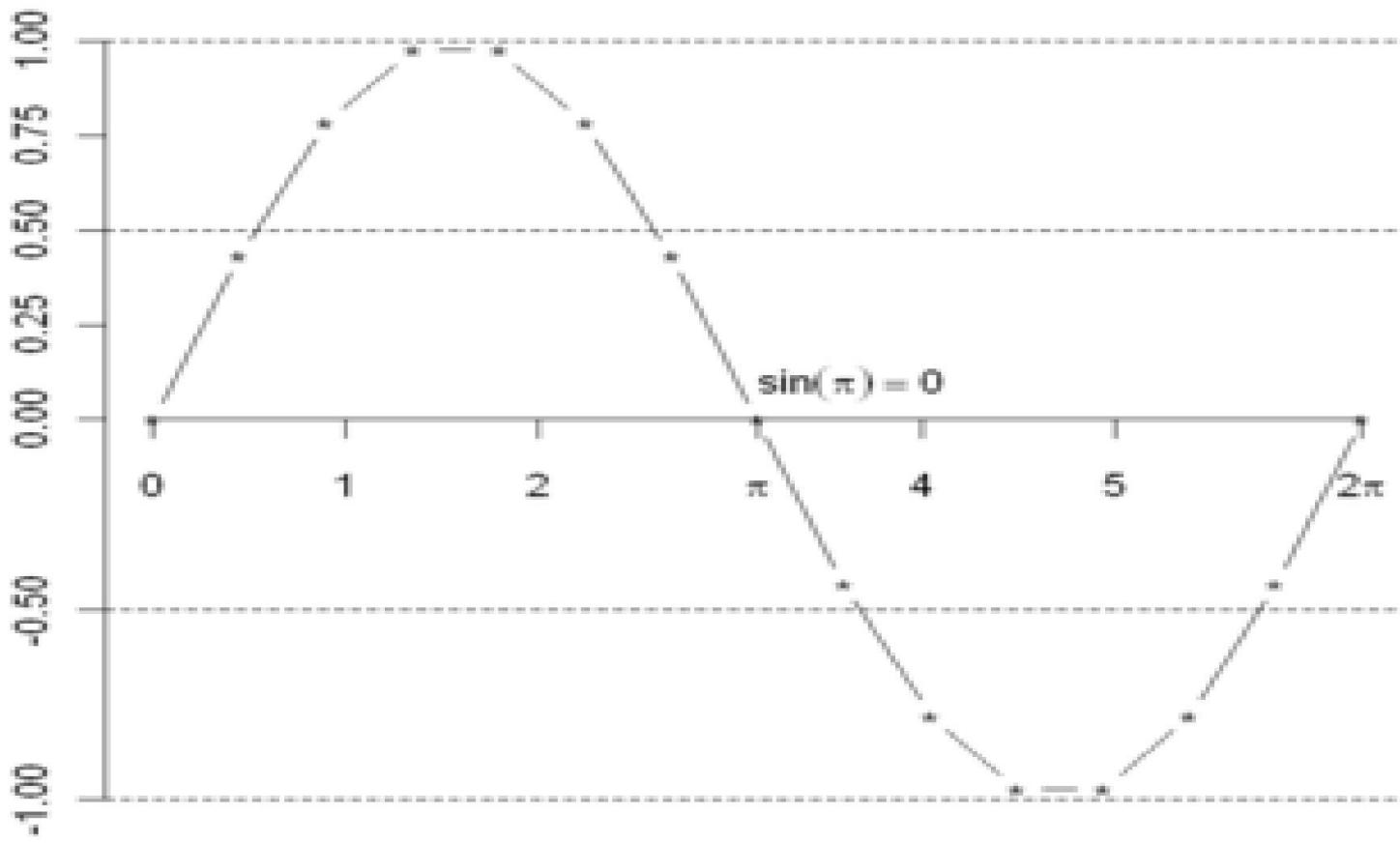
```
> x <- seq(0,2*pi, by = 0.1)
> y <- sin(x)
> y1 <- cos(x)
> plot(x,y, col = "green", type = "l", lwd = 3)
> lines(x,y1, col = "red", lwd = 3)
> mtext("Sine and Cosine Plot", side = 3, line = 1)
```





- `text(pi, 0.1, Expression(sin(pi) == 0), adj = 0)`
  - `abline(h = c(-1, -0.5, 0.5, 1), lty = 2)`

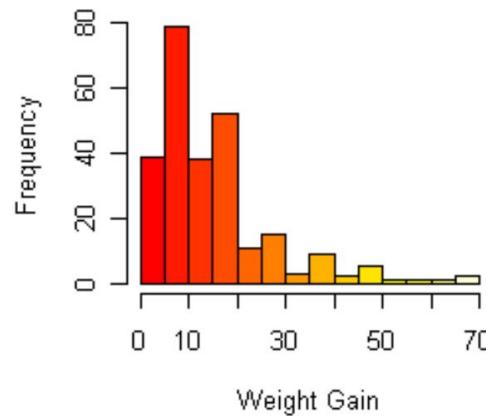
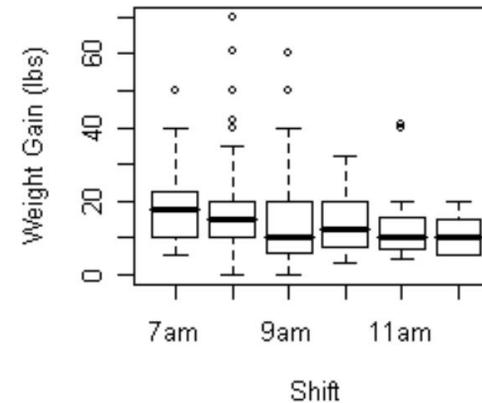
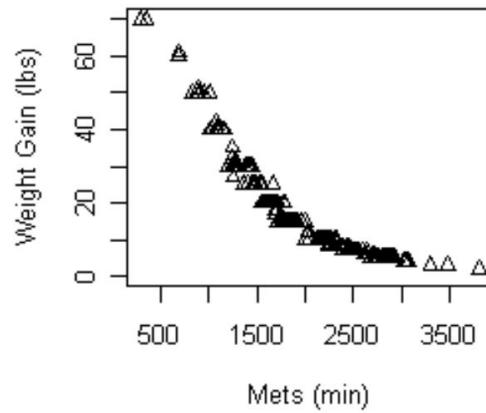
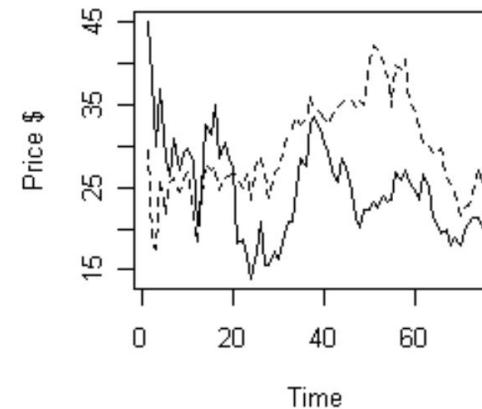
• تضييف الدالة `abline` خطوطاً إلى الرسم البياني وعند استخدام خيارات `h` نشير إلى أننا نريد أن تكون الخطوط خطوطاً مرجعية أفقية وبالتالي نحتاج فقط إلى تحديد تقاطع على كل سطر. • يحدد خيار `adj` في وظيفة النص كيفية ضبط النص = 0 (مضبوط على اليسار، = 0.5 في المنتصف، و 1 = مضبوط على اليمين).





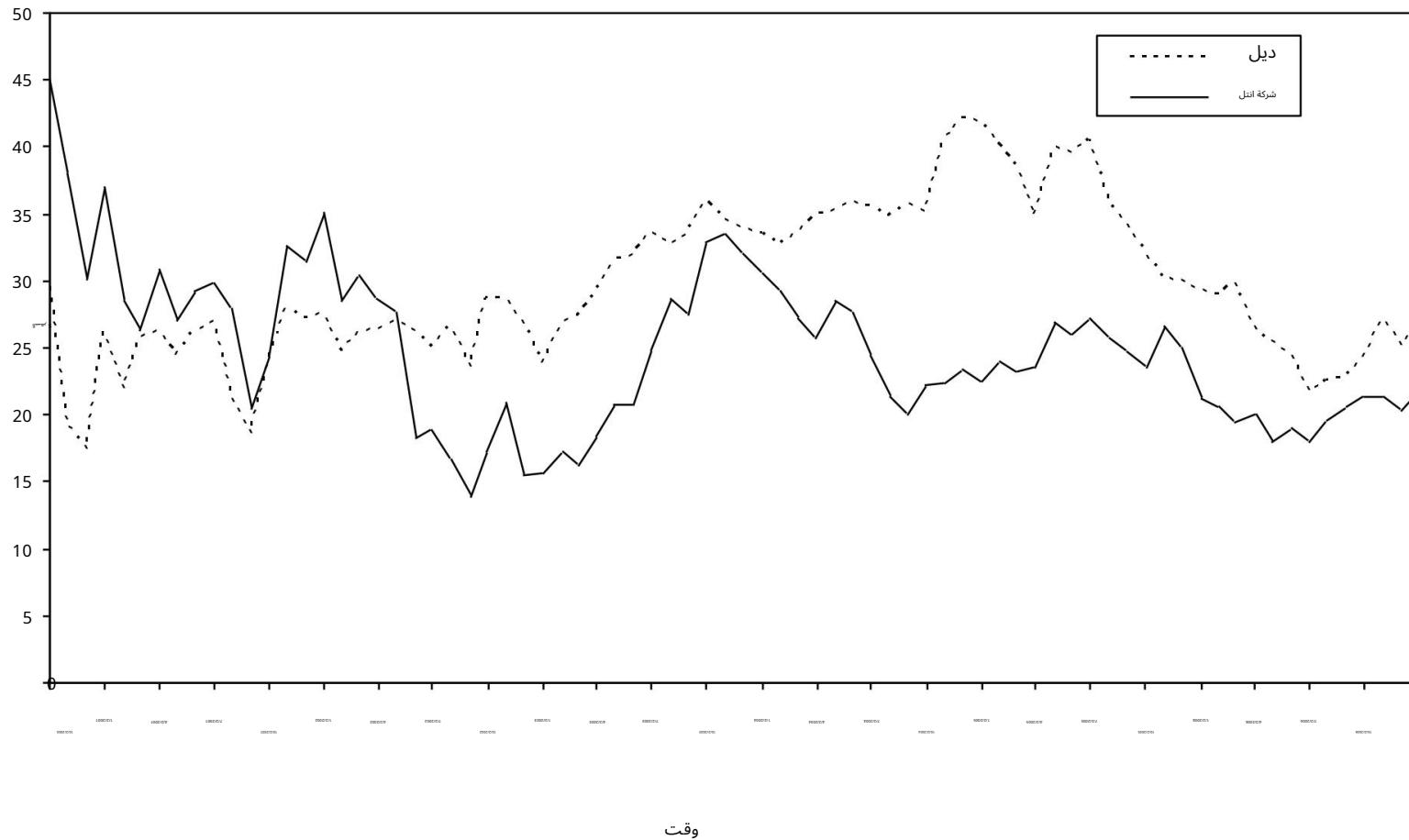
• يضيف خط المنحدر  $b$  ويعرض  $a$  إلى المخطط الحالي.  
• يضيف خطًا أفقيًا ( $v=x$ )  $abline(v=x)$  عموديًا  
• يضيف خطًا مائلًا محدودًا بالقمم المرتبة في  $(y, x)$  و (اختياريًا) قم  
مضلغاً ( $x, y$ , ...), يرسم مضلغاً محدوداً بالقمم المرتبة في  $(y, x)$  و (اختياريًا) قم  
بتظليلها بخطوط التظليل، أو املأها إذا كان جهاز الرسومات يسمح بذلك  
الأشكال.

# رسومات الألواح

**Histogram****Weight Gain****Met Minutes vs. Weight Gain****Closing Stock Prices**

# الجودة - اكسل

إغلاق أسعار الأسهم



## ملخص

• يمكن العثور على جميع رموز وملفات R على:

- <http://www.cran.r-project.org/>
-

# المعلمات الرسمية

• حجم النص والرمز • يمكن استخدام الخيارات التالية للتحكم في حجم النص والرمز

الرسوم البيانية.

## خيارات

## وصف

`cex.com` رقم يشير إلى المقدار الذي يجب من خلاله تغيير حجم النص والرموز بالنسبة إلى الإعداد الافتراضي. 1.5 = افتراضي، 0.5 = أصغر بنسبة 50%， أكبر بنسبة 50%، الخ.

`cex.axis`

تكبير الشرح المحوري بالنسبة لـ `cex`

`cex.lab`

تكبير تسميات `x` و `y` بالنسبة لـ `cex`

`cex.main`

تكبير العناوين نسبة إلى `cex`

`cex.sub`

تكبير الترجمات نسبة إلى CEX

# المعلمات الرسمية

## الخطوط

- يمكنك بسهولة ضبط حجم الخط ونمطه، ولكن عائلة الخطوط أكثر تعقيداً بعض الشيء.

الخيار	وصف
الخط	عدد صحيح يحدد الخط الذي سيتم استخدامه للنص. 1=عادي، 2=غامق، 3=مائل، 4=غامق ومائل، 5=رمز
الخط. المحور	الخط للتعليق التوضيحي للمحور
Font.lab	الخط للتسميات x و y
الخط.الرئيسي	الخط للعناوين
Font.sub	الخط للترجمات
ملاحظة	<p>حجم نقطة الخط (حوالي 1/72 بوصة)  <math>xec*sp</math>=حجم النص</p>
عائلة	<p>عائلة الخطوط لرسم النص. القيم القياسية هي "sans", "mono", "symbol".</p> <p>الأستاذ الدكتور تيسير جعشن <sup>QIS IS441</sup> رسم الخطوط يعتمد على الجهاز.</p>

# المعلمات الرسمية

## محاور

يمكنك إنشاء محاور مخصصة باستخدام الدالة `axis()`. المحور (الجانب، عند=)، التسميات=، `pos=`، `lty=`، `col=`، `las=`، `tck=`، (...، حيث

خيارات

وصف

جانب

عدد صحيح يشير إلى جانب الرسم البياني لرسم المحور (1=أسفل، 2=يسار، 3=أعلى، 4=يمين)

في

متجه رقمي يشير إلى مكان رسم علامات التشنج الإلإرادي

تسميات

ناقل أحرف للتسميات التي سيتم وضعها على علامات التجزئة

(إذا كان ،`NULL`سيتم استخدام القيم (at

نقط البيع

الإحداثيات التي سيتم رسم خط المحور عليها. (أي القيمة على المحور الآخر الذي يتقطع فيه)

`lty`

نوع الخط

العقيد

لون الخط وعلامة التجزئة

`las`

التسميات متوازية (0)=) أو متعامدة (2)=) على المحور

`tck`

طول علامة التجزئة كجزء من منطقة التخطيط (الرقم السادس خارج الرسم البياني، والرقم الموجب بداخله، 0 يمنع علامات التجزئة، 1ينشئ خطوط الشبكة) الافتراضي هو 0.01-الأستاذ الدكتور تيسير حسن QIS\_IS441

(...)



## حفظ الرسومات الخاصة بك

• يقوم R عادةً بإنشاء مخرجات على الشاشة • في نظام التشغيل Windows، يمكنك الإشارة والنقر فوق رسم بياني لنسخه إلى الحافظة • ومع ذلك، يمكن لـ R أيضًا حفظ مخرجات الرسومات الخاصة به في ملف يمكنك توزيعه أو تضمينه في مستند تم إعداده باستخدام Word أو LATEX

# إعادة توجيه إخراج الرسومات الخاصة بك



• حدد الجهاز أولاً:

• جودة عالية، تنسيق محمول - pdf() -

الي الجودة

• جودة منخفضة، ولكنها مناسبة للويب - png()

• بعد إنشاء الرسومات الخاصة بك، ما عليك سوى إغلاق الجهاز

• dev.off()

```
# This graph is plotted to the PDF file
> pdf("my_graph.pdf")
> plot(x, y, ylab = "This is a simple graph")
> dev.close()
```

## مراجع

**Thank You**