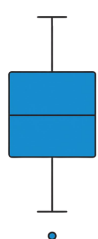


NEWSLETTER

SEMAKAN KUALITI DATA MENGGUNAKAN

BOX PLOT

Box plot merupakan kaedah **grafik/visualisasi data** yang digunakan dalam statistik bagi **memaparkan taburan data dan mengenalpasti outliers**



APA ITU BOX PLOT?

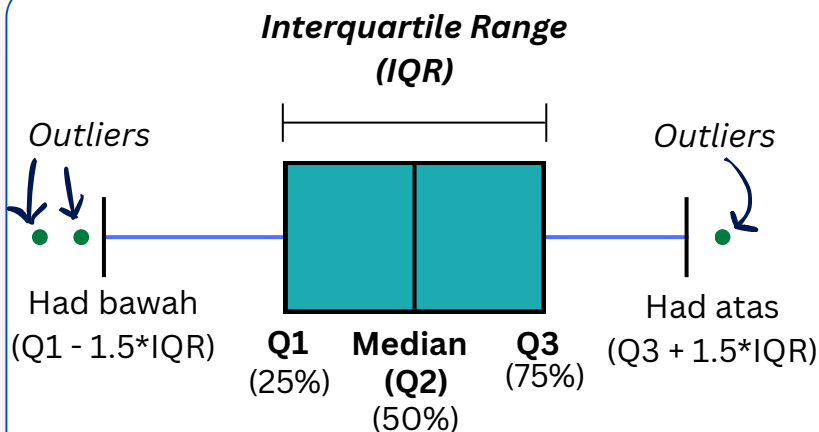
Box plot (juga dikenali sebagai whisker plot) menggambarkan lima (5) ringkasan statistik utama sesuatu taburan data

RINGKAS TETAPI EFEKTIF



Ia membantu untuk melihat kedudukan median sesuatu data, taburan data dan *outliers* dengan cepat

MEMAHAMI BOX PLOT



INTERQUARTILE RANGE (IQR)

$$IQR = Q3 - Q1$$

- Mengukur julat 50% data pertengahan, iaitu antara persentil ke-25 hingga ke-75
- Semakin besar IQR, semakin besar serakan data

KOMPONEN BOX PLOT

- | | |
|---|---|
| 1 Median (Q2)
Nilai penengah data | 4 Had bawah
Nilai minimum dalam julat normal |
| 2 Kuartil 1 (Q1)
Nilai persentil ke-25 | 5 Had atas
Nilai maksimum dalam julat normal |
| 3 Kuartil 3 (Q3)
Nilai persentil ke-75 | 6 Outliers
Nilai yang terletak diluar julat normal |

Box plot memberikan gambaran menyeluruh tentang taburan data dan merupakan visualisasi penting dalam semakan serta penambahbaikan kualiti data





MENGAPA BOX PLOT?



Mengesan nilai
outliers



Mengenalpasti
kesilapan input data



Mengenalpasti
sama ada data
simetri atau
condong/pencong



Membandingkan
kualiti antara
kumpulan data



Meningkatkan
ketepatan statistik
rasmi

MENTAFSIR CORAK TABURAN DATA MELALUI BOX PLOT

Condong ke Kanan (Right Skew)



Median berada hampir
dengan kuartil pertama.
Kebanyakan data tertumpu
pada nilai yang lebih rendah

Variasi Tinggi



Kotak besar, julat IQR
besar; menunjukkan
taburan data yang luas

Data Tidak Bersih



Banyak outliers; petunjuk
utama kemungkinan ralat
input yang tinggi

“ Data yang berkualiti ialah asas kepada statistik yang boleh dipercayai dan keputusan yang tepat ”





ANALISIS BOX PLOT MENGGUNAKAN DATA BANKI EKONOMI

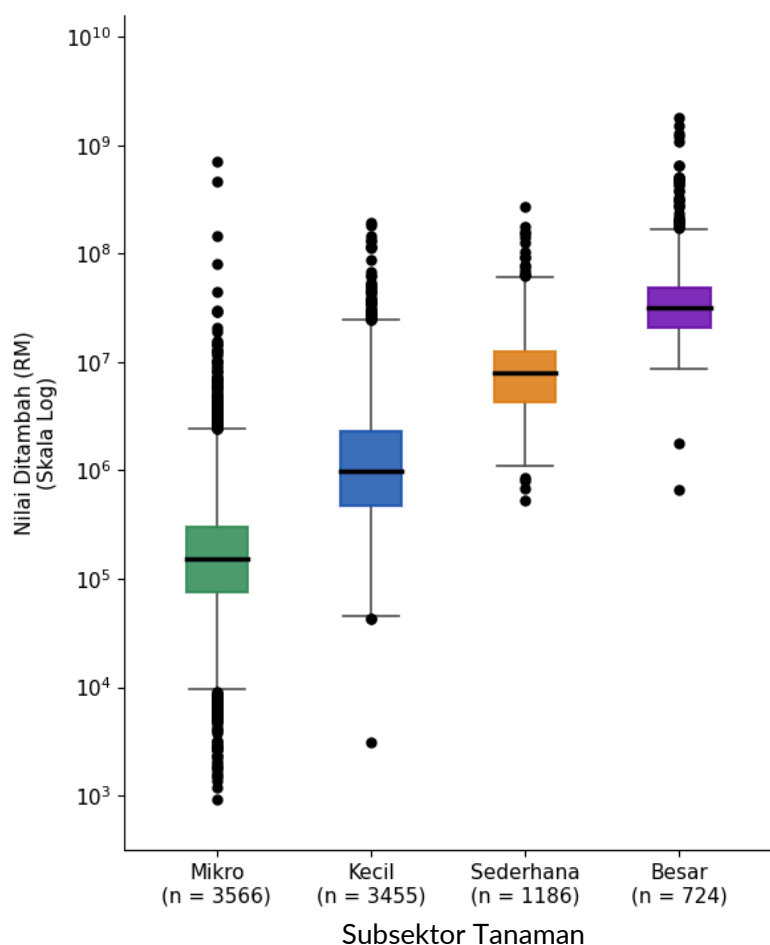
Perbandingan Nilai Ditambah (RM) bagi Kategori PMKS subsektor Tanaman:

MIKRO

KECIL

SEDERHANA

BESAR



DAPATAN UTAMA

1

PERTINDIHAN PRESTASI ANTARA SAIZ FIRMA

- Taburan nilai ditambah bertindih antara semua kategori PMKS
- Terdapat firma kecil yang mencatat prestasi setara firma lebih besar

2

VARIASI BERBEZA MENGIKUT SAIZ FIRMA

- Variasi dan julat firma kecil dan mikro lebih tinggi
- Median meningkat mengesahkan semakin besar saiz firma, semakin tinggi Nilai Ditambah

3

OUTLIERS MEWAKILI FIRMA SEBENAR

- Nilai ekstrem wujud dalam semua kategori, termasuk firma Mikro
- Dalam data banci, outliers mencerminkan struktur ekonomi sebenar



Skala logaritma digunakan bagi memaparkan taburan data yang mempunyai julat nilai yang sangat besar dengan lebih jelas



INTERPRETASI

Box plot menunjukkan bahawa hubungan antara saiz firma dan nilai ditambah tidak hanya bergantung kepada median. Taburan yang bertindih dan variasi dalaman yang besar mencerminkan ekonomi yang heterogen dan tidak seimbang.



KESIMPULAN

- Saiz firma mempengaruhi nilai ditambah, namun bukan satu-satunya penentu prestasi.
- box plot penting untuk memahami struktur ekonomi sebenar dalam data banci
- Outliers perlu disemak bagi memastikan ketepatan data

