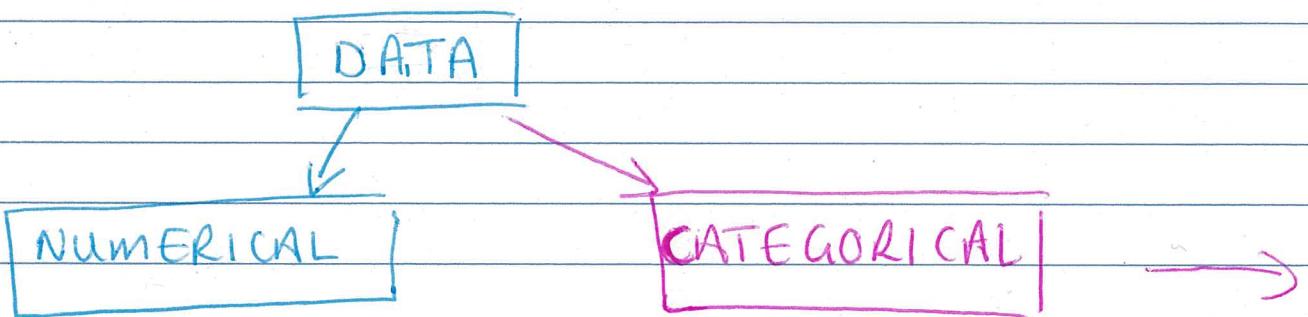


Summary Notes for Chapter 1

Types of Data

The BASICS



Discrete - can only have fixed values

- use frequency tables (and percent frequency)
- then plot using histograms
- also can plot using box, stem & leaf, dot
- ungrouped data

e.g. Continuous - no gaps in the possible values

• weight

Be careful as sometimes it is

• height

hard to tell the difference

• salary

- again use frequency tables and histograms

$$\frac{f}{\sum f} \times 100 = \%$$

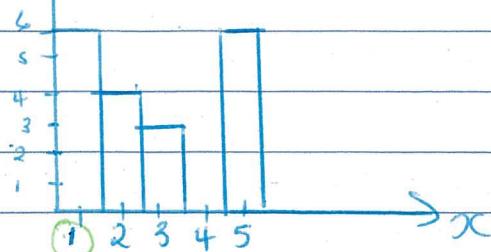
- grouped data now - otherwise representing simply would be IMPOSSIBLE!!

x	f	%
1	1	10
2	2	20
3	3	30
4	2	20
5	1	10
6	1	10
7	1	10
8	1	10
9	1	10
10	1	10
11	1	10
12	1	10
13	1	10
14	1	10
15	1	10
16	1	10
17	1	10
18	1	10
19	1	10
20	1	10
21	1	10
22	1	10
23	1	10
24	1	10
25	1	10
26	1	10
27	1	10
28	1	10
29	1	10
30	1	10
31	1	10
32	1	10
33	1	10
34	1	10
35	1	10
36	1	10
37	1	10
38	1	10
39	1	10
40	1	10
41	1	10
42	1	10
43	1	10
44	1	10
45	1	10
46	1	10
47	1	10
48	1	10
49	1	10
50	1	10
51	1	10
52	1	10
53	1	10
54	1	10
55	1	10
56	1	10
57	1	10
58	1	10
59	1	10
60	1	10
61	1	10
62	1	10
63	1	10
64	1	10
65	1	10
66	1	10
67	1	10
68	1	10
69	1	10
70	1	10
71	1	10
72	1	10
73	1	10
74	1	10
75	1	10
76	1	10
77	1	10
78	1	10
79	1	10
80	1	10
81	1	10
82	1	10
83	1	10
84	1	10
85	1	10
86	1	10
87	1	10
88	1	10
89	1	10
90	1	10
91	1	10
92	1	10
93	1	10
94	1	10
95	1	10
96	1	10
97	1	10
98	1	10
99	1	10
100	1	10

Σf means add up all numbers
 ↪ with equal total number of results
 must add up to 100%

(2)

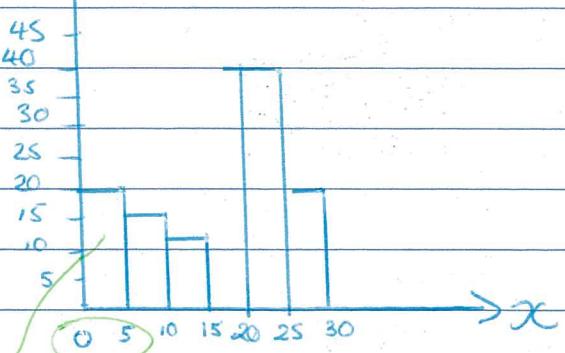
freq. ↑ Example - ungrouped



- possible value goes in centre of column

- no gaps
- always label axes
- give a title
- use appropriate scale

freq. ↑ Example - grouped



- lower limit of group goes on left hand side of column

- so there are 20 results between 0 and 5
- and no values between 15 and 20
- same

Key points for Categorical

- how many levels and what are they?
(these are the possible choices/results)

e.g

• gender

• eye colour

• postcode

• rego number

- use frequency / percent freq table then bar chart (most used way to present data)

- bar chart has gaps between the columns

- mode is the only "average" that is meaningful.