PRIME INTUIT
Finishing School

Stem and Leaf Plots
Scatter Plots

PRIME INTUIT
Finishing School

Stem and Leaf Plots

Sachin's score in his last 30 ODI's

[40, 175, 10, 69, 43, 96, 8, 4, 200, 7, 24, 28, 120, 38, 27, 111, 2, 53, 85, 18, 2, 48, 15, 3, 22, 14, 39, 6, 114, 52]

Stem	Leaf
6	9
8	5
9	6
11	1

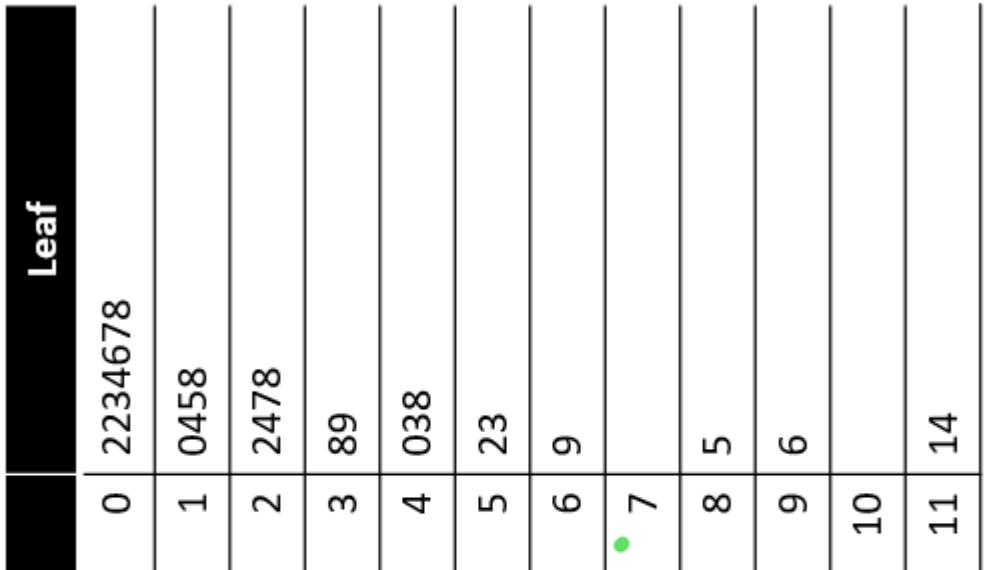
PRIME INTUIT
Finishing School

Stem and Leaf Plots

Sachin's score in his last 30 ODI's

[2,2,3,4,6,7,8,10,14,15,18,22,24
,27,28,38,39,40,43,48,52,53,69,
85,96,11,114,120,175,200]

Stem	Leaf
0	2234678
1	0458
2	2478
3	89
4	038
5	23
6	9
7	
8	5
9	6
10	
11	14



PRIME INTUIT
Finishing School

Stem and Leaf Plots (continuous data)

Sachin's strike rate in his las 30 ODIs

[58.82, 124.11, 58.82,109.52, 82.69, 92.3, 100, 80, 136.05, 63.63, 54.54, 96.55, 104.34, 67.85, 122.72, 109.9, 50.11, 77.94, 73.91,128.57, 33.33, 76.19, 62.5, 25, 95.65, 93.33, 130, 31.57, 77.55, 108.33]

"Efficient way of describing small and medium data"

Stem	Leaf
2	5
3	23
4	
5	0599
6	248
7	4688
8	03
9	2367
10	048
11	00
12	349
13	06

PRIME INTUIT
Finishing School

Stem and Leaf Plots (for Large Values)

State / Union Territory	NSDP(INR)
Goa	467998
Delhi	365529
Sikkim	357643
Chandigarh	297313
Haryana	226644
Puducherry	220461
Karnataka	210887
Telangana	205696
Uttatakhand	198738
Tamil Nadu	193750
Maharashtra	180596
Kerala	179523

Stem	Leaf
17	9523
18	0596
19	3750 8738
20	5696
21	0887
22	0461 6644
29	7313
35	7643
36	5529
46	7798

What if the data contains bigger values ?

difficult

PRIME INTUIT
Finishing School

Stem and Leaf Plots (Splitting Rows)

What if a row has many values ?

Stem	Leaf
43	012,019,213,314,419
43	514, 625,765,822,991

Row 1: leaf starting from 0-5
Row 2: leaf starting from 5-9

inference

PRIME INTUIT
Finishing School

Stem and Leaf Plots Vs Histogram

Stem	Leaf
0	2234678
1	0458
2	2478
3	89
4	038
5	23
6	9
7	
8	5
9	6
10	
11	14
12	0

Stem & Leaf Plot looks like a histogram rotated on its side

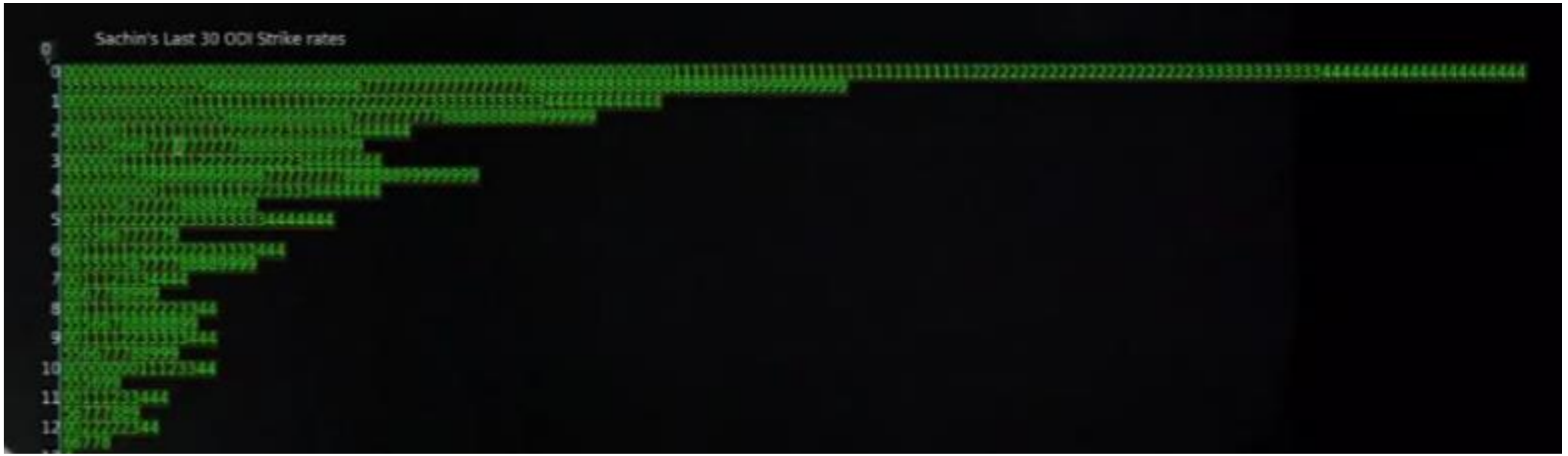
More informative: displays values with group

SQL

Histogram

PRIME INTUIT
Finishing School

Stem and Leaf Plots Vs Histogram



Stem & Leaf Plot is not preferred for large data set histogram is better

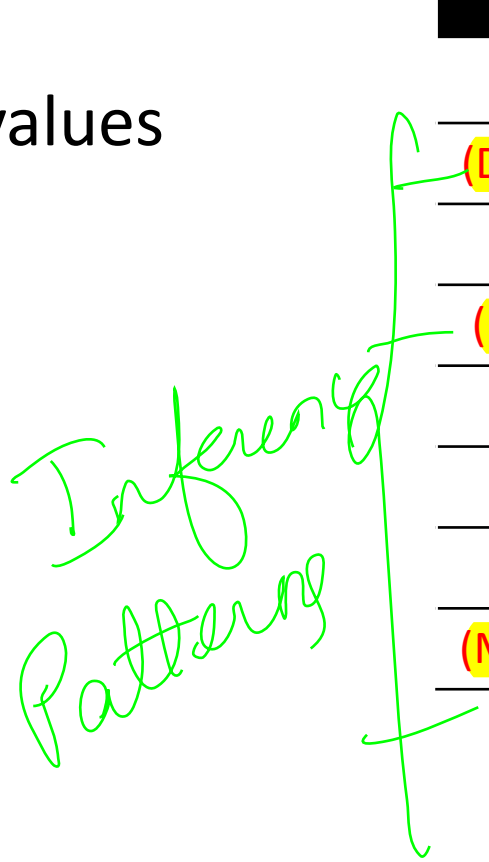
where

PRIME INTUIT
Finishing School

Stem and Leaf Plots Vs Histogram

Displaying individual values makes it easy to spot patterns

Stem		Leaf
	1	
(Diff of 3)	2	222 5555 888
	3	
(Diff of 3)	4	111 4444 777
	5	
	6	11 44444 579
	7	
(Multi of 2)	8	000 22 444 66 888



PRIME INTUIT
Finishing School

Back to back Stem and Leaf Plots

2 sides

Leaf	Stem	Leaf
8875210	0	2234678
30	1	0458
1	2	2478
72	3	89
	4	038
	5	23
	6	9
76	7	
1	8	5
	9	6
	10	

Test Scores

ODI Scores

Can be used to compare 2 different sets of data

PRIME INTUIT
Finishing School

How to describe relationship between variables? Scatter Plots

PRIME INTUIT
Finishing School

Multiple attributes in datasets

Cricket	Runs, Balls, Minutes, Strike Rate, dismissal
Agriculture	State, district, crop, area, yield
E Commerce	Price, Color, Pattern, Size, Discount

We often expect certain relationship between attributes

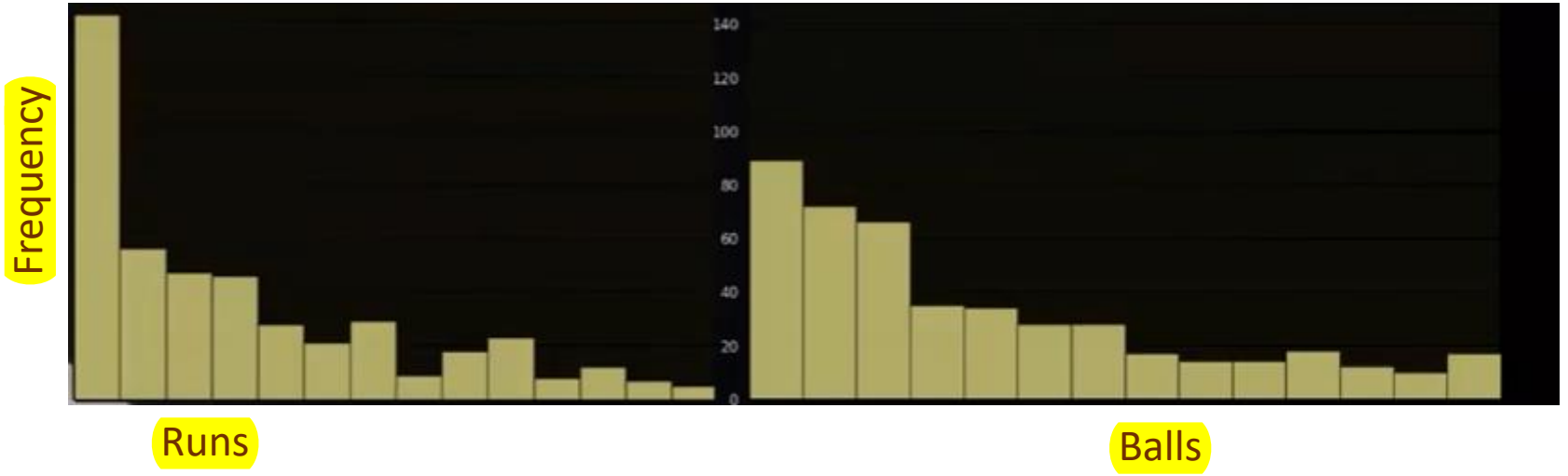
Runs Scored = f(Balls faced)

Total yield = g(area)

Price = h(Size)

PRIME INTUIT
Finishing School

Can individual plots reveal relations

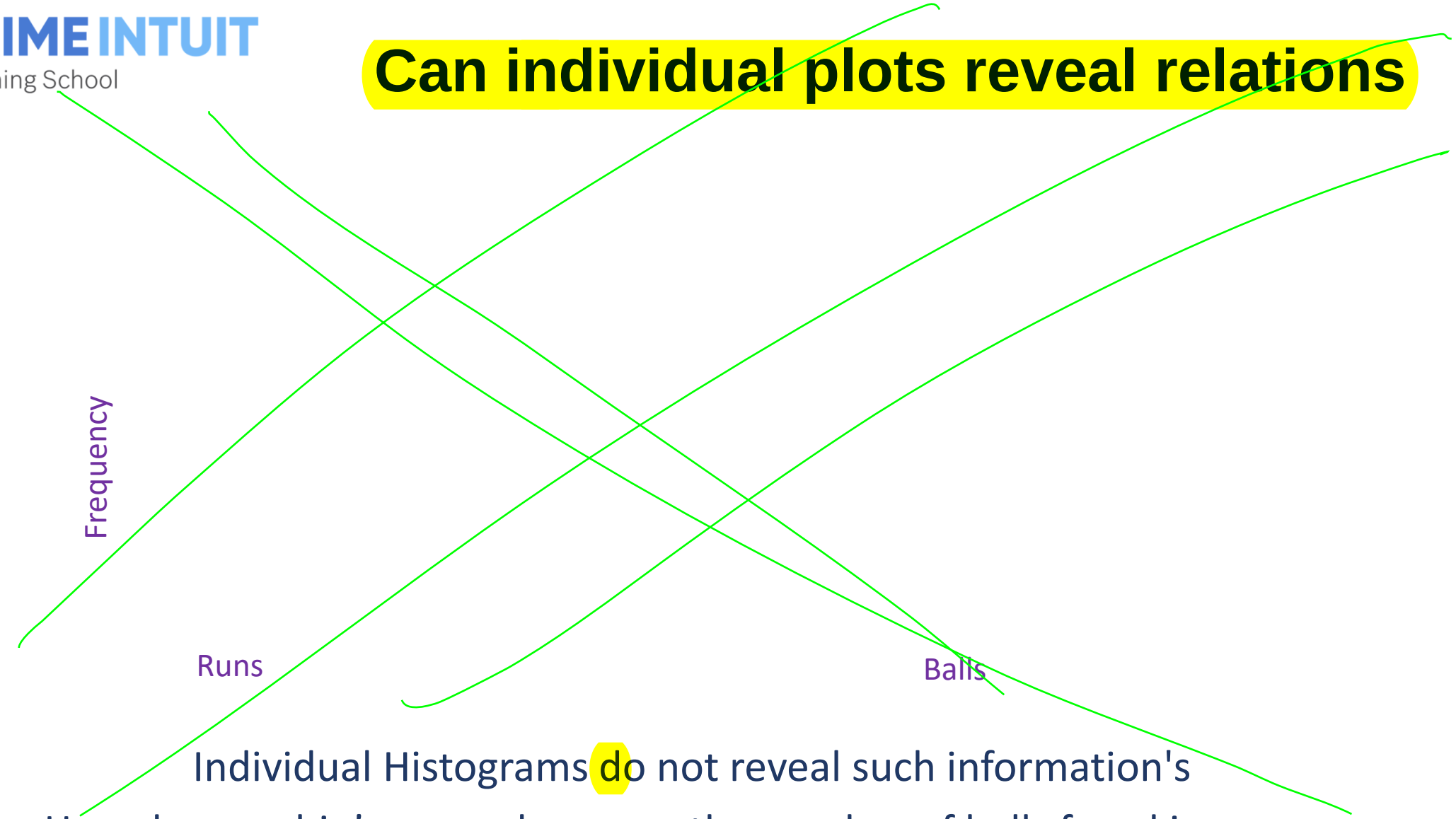


Individual Histograms do not reveal such information's

How does sachin's score change as the number of balls faced increases

PRIME INTUIT
Finishing School

Can individual plots reveal relations

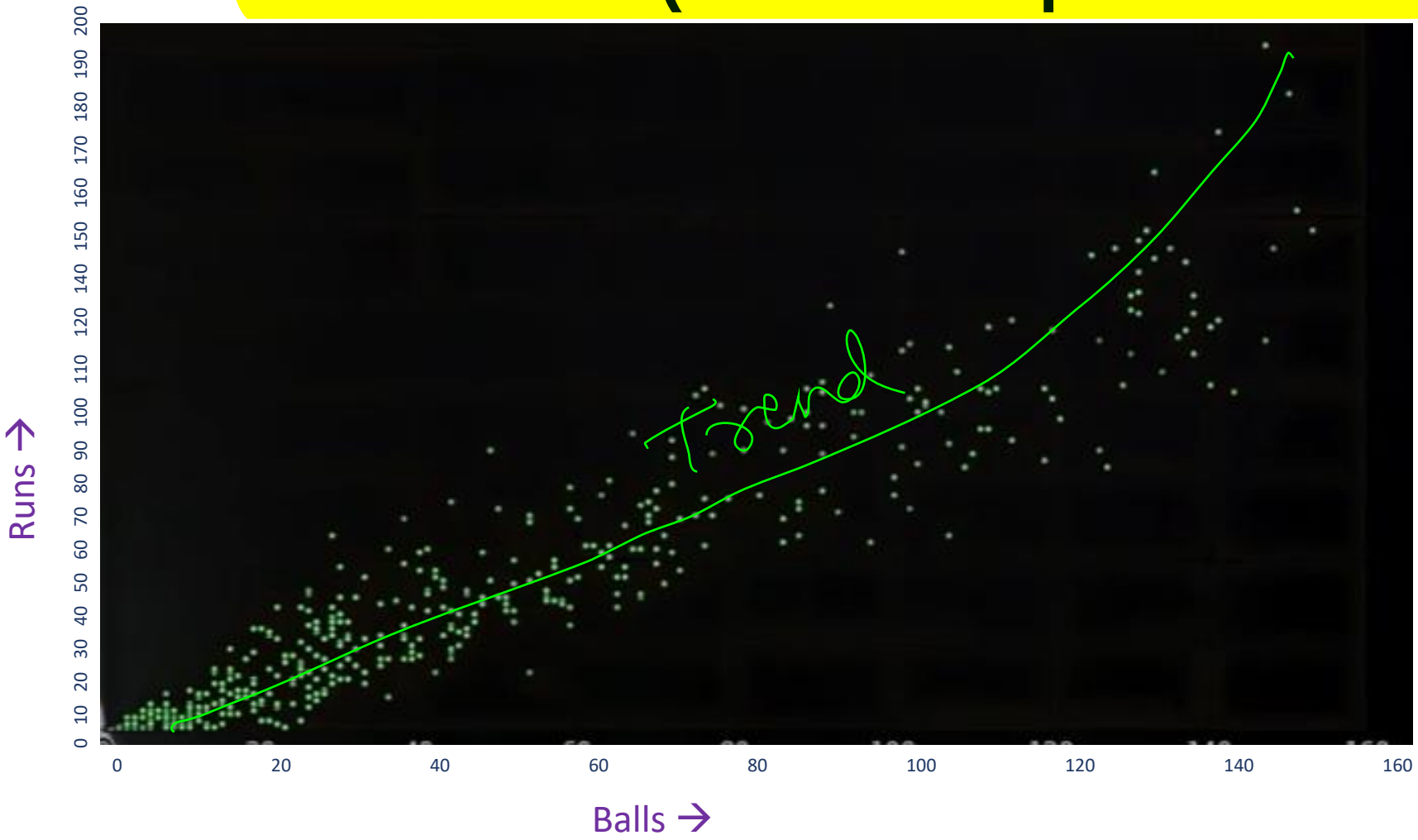


Individual Histograms do not reveal such information's

How does sachin’s score change as the number of balls faced increases

PRIME INTUIT
Finishing School

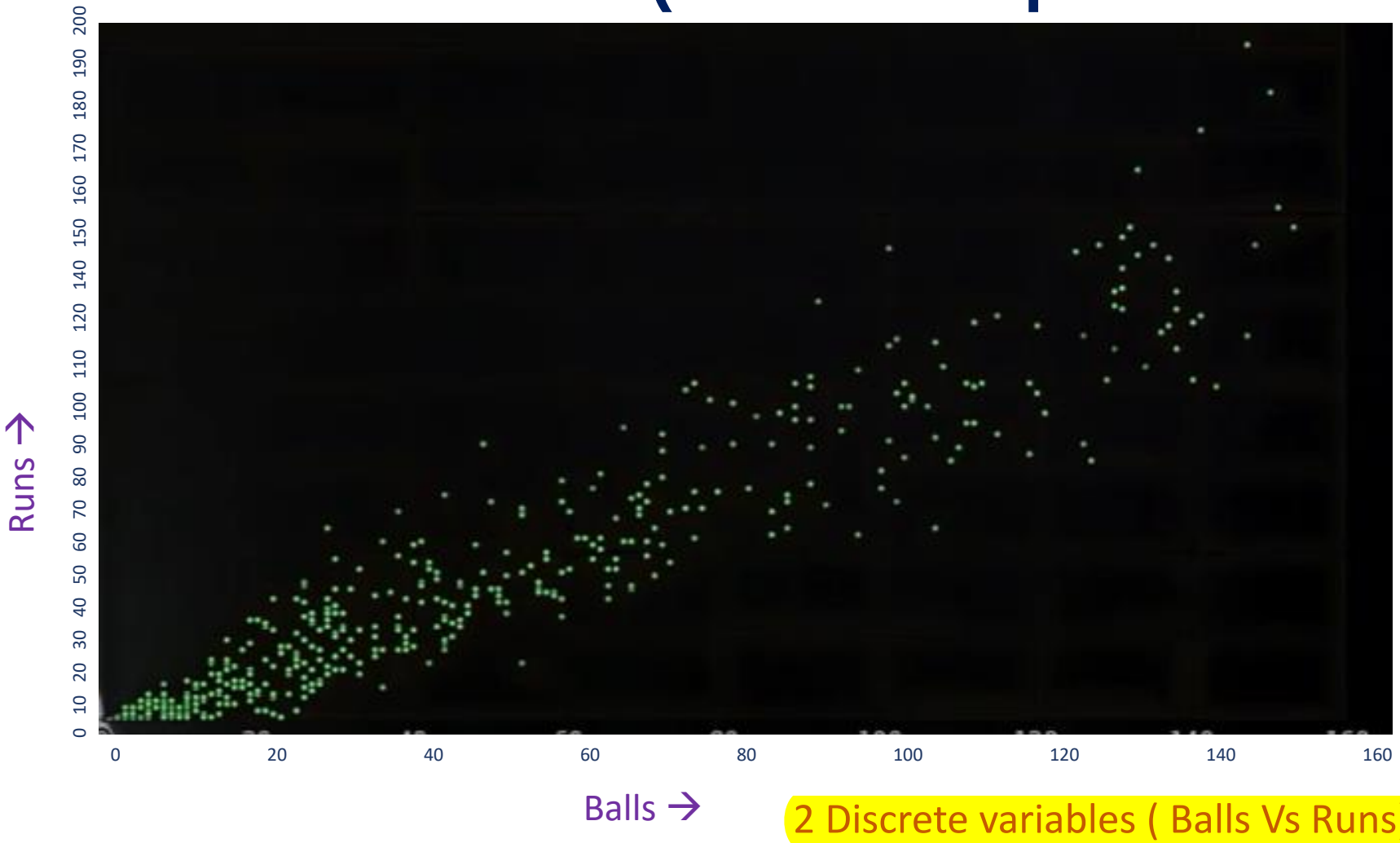
Scatter Plot (relationship b/n Variables)



X – co ordinates balls faced, Y – co ordinates runs scored

PRIME INTUIT
Finishing School

Scatter Plot (relationship b/n Variables)



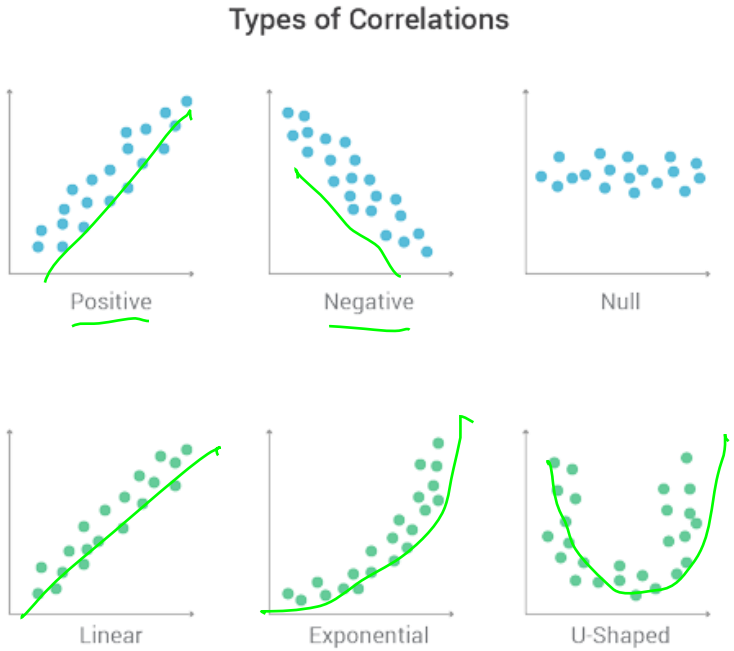
Not for qualitative variables

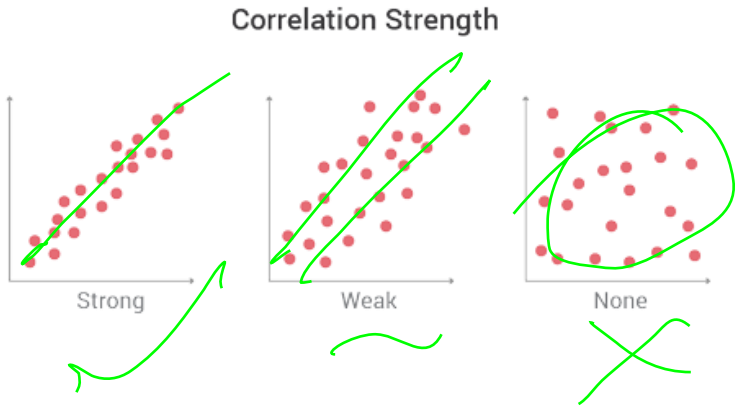
2 Discrete variables (Balls Vs Runs)
2 Continuous variables (farm data)
1 Continuous and 1 discrete (strike rate Vs Runs)

Mostly good for

PRIME INTUIT
Finishing School

Recap on fuctions





Linear Y = mx+c

Exponential Y = e^x

Quadratic Y = ax^2

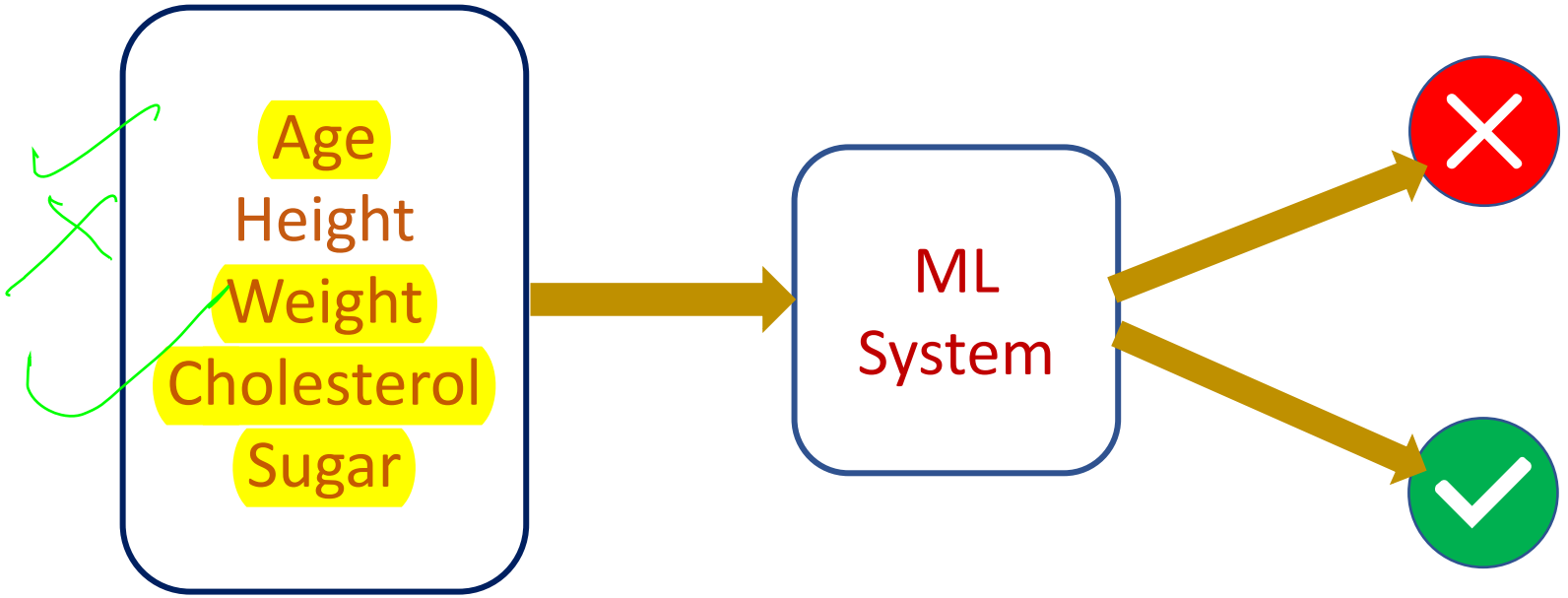
PRIME INTUIT
Finishing School

Use of Scatter Plots in ML

PRIME INTUIT
Finishing School

Use of Scatter plot in ML

1. Identifying discriminatory features

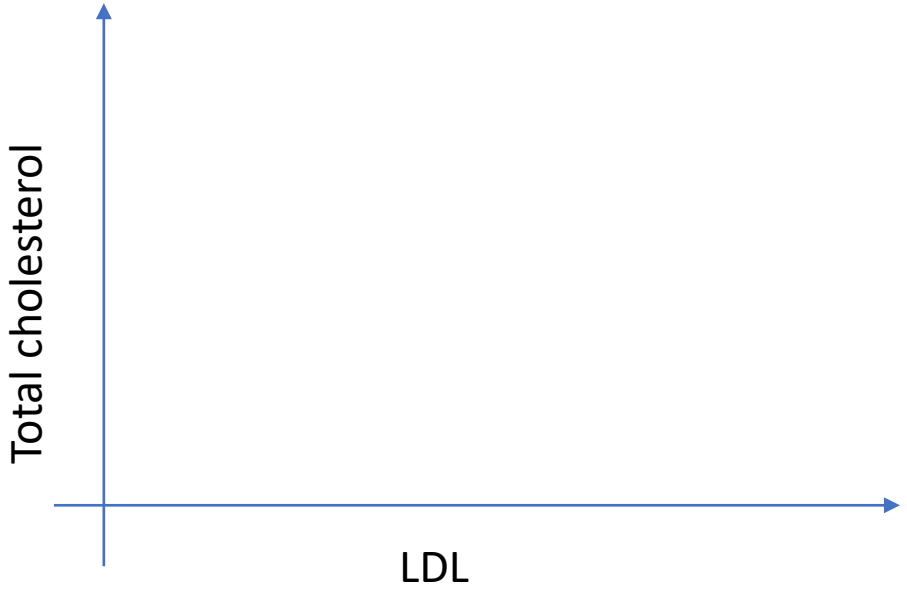


Age	Height	…….	Risk
38	6		✗
47	5.8		✓
53	5.10		✓

PRIME INTUIT
Finishing School

Use of Scatter plot in ML

1. Identifying discriminatory features

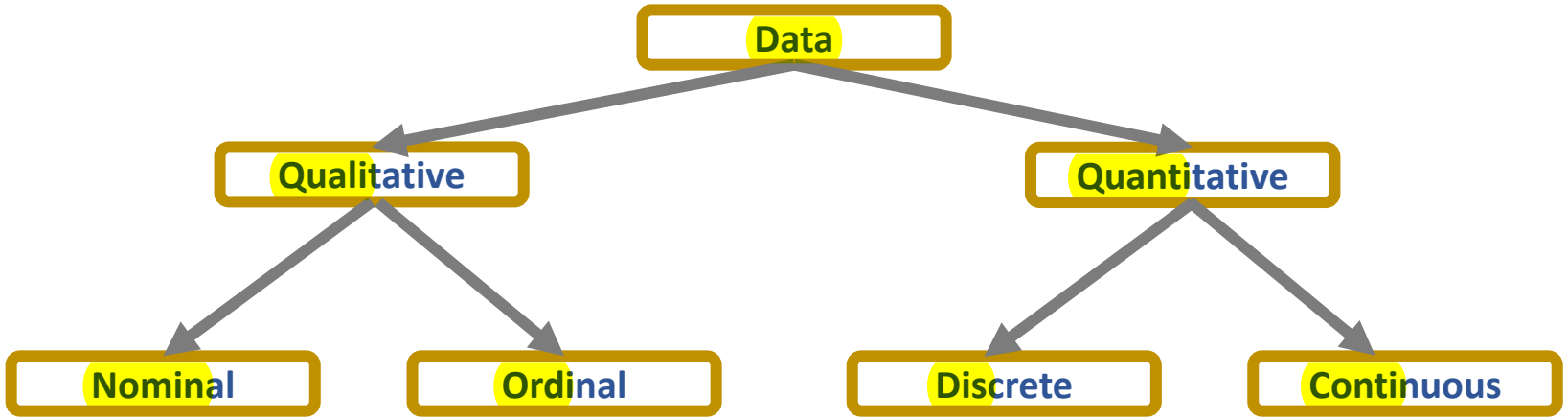


Use uncorrelated or non redundant features for classification

PRIME INTUIT
Finishing School

Summary

Summary



(Relative) Frequency Tables
(Relative) Frequency Bar Charts
(Relative) Grouped Bar Charts

(Relative) Histograms
(Relative) Frequency Polygons
(Relative) Cumulative Frequency Polygons
Steam and Leaf Plots
Scatter Plots

Thank You!