AMRITA VISHWA VIDYAPEETHAM

AHEAD Online

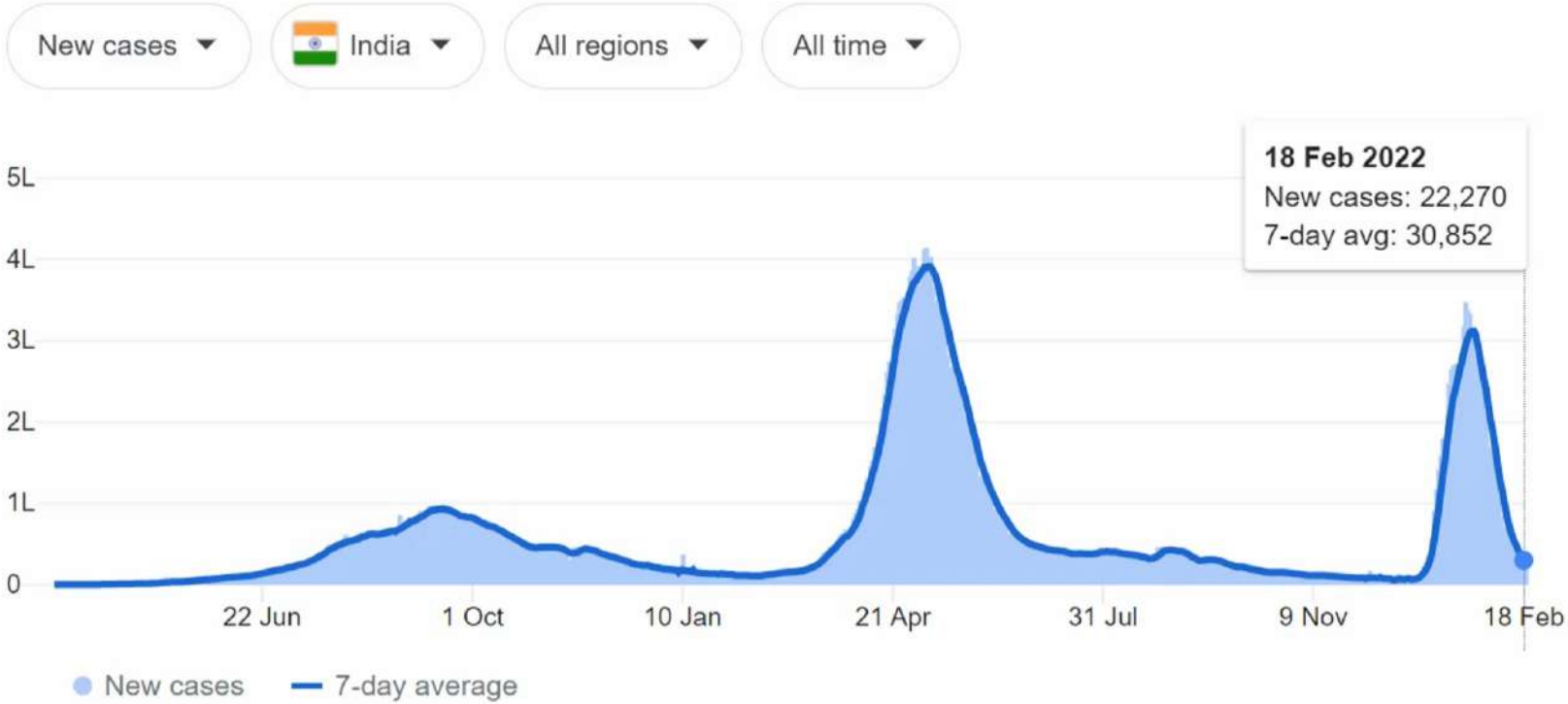
Probability and Statistics

Dr. Ajay Vadakkepatt, Jayakumar T V, Vidya M
Department of Computer Science
Amrita Vishwa Vidyapeetham

Objective

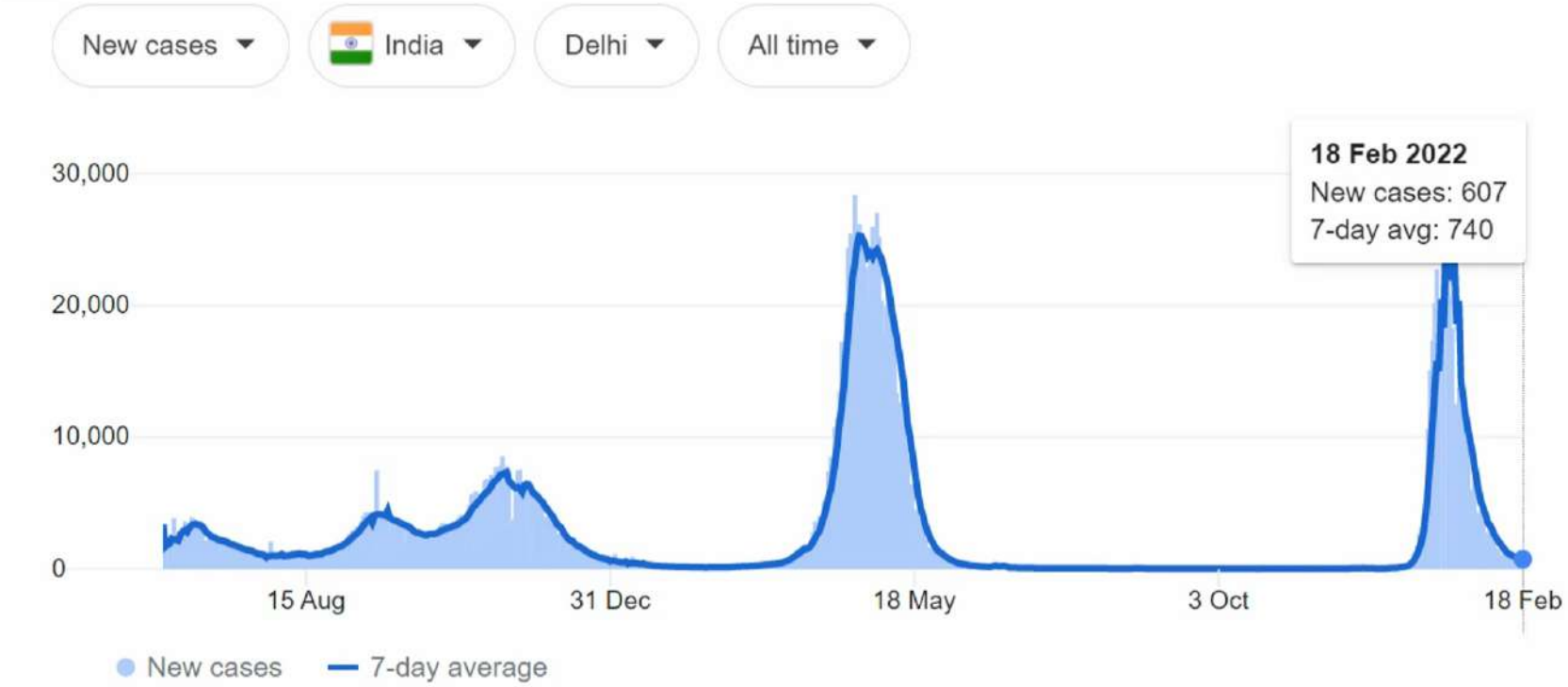
- COVID-19 Statistics
- Result Analysis
- Google Scholar

COVID-19 Statistics



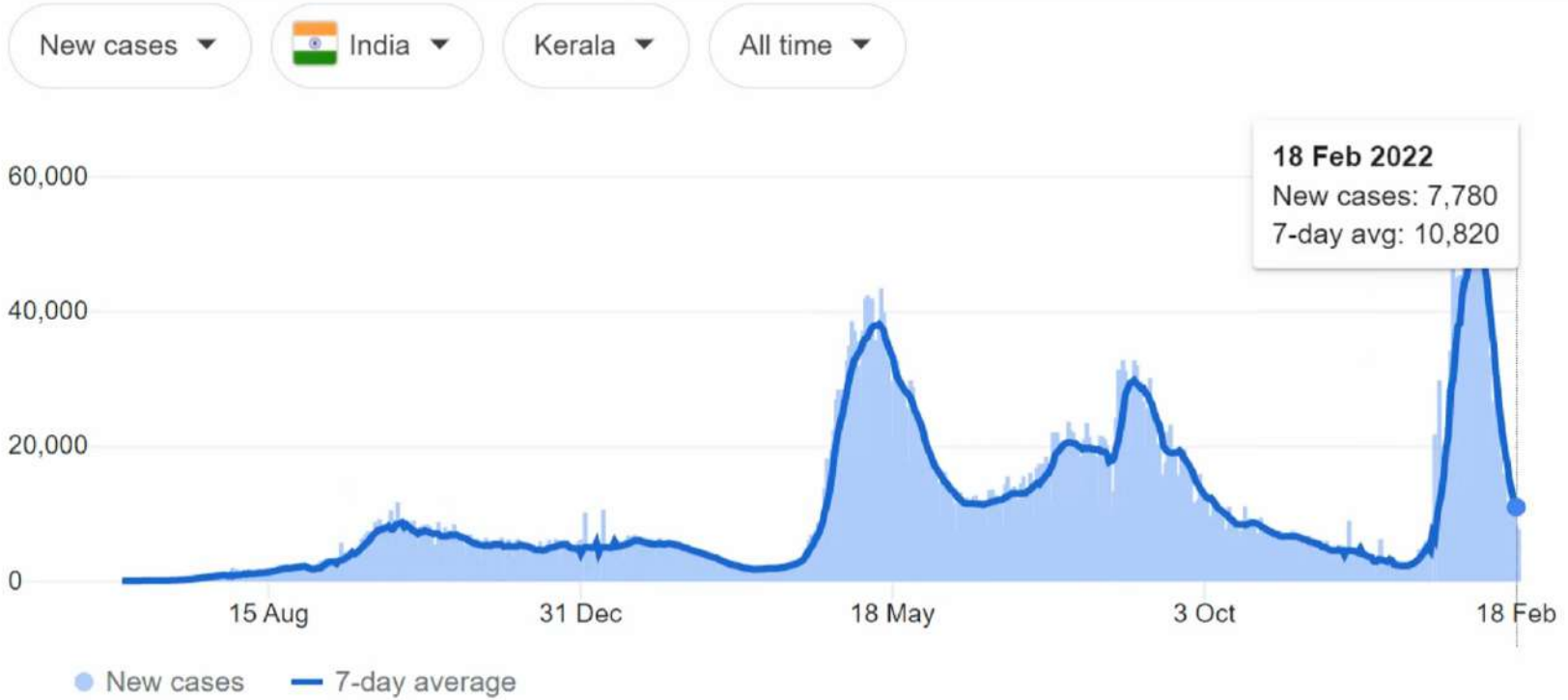
https://www.google.com/search?q=covid-19+statistics&oq=covid-19+statistics&aqs=chrome..69i57j0i512l9.5582j0j15&sourceid=chrome&ie=UTF-8

COVID-19 Statistics



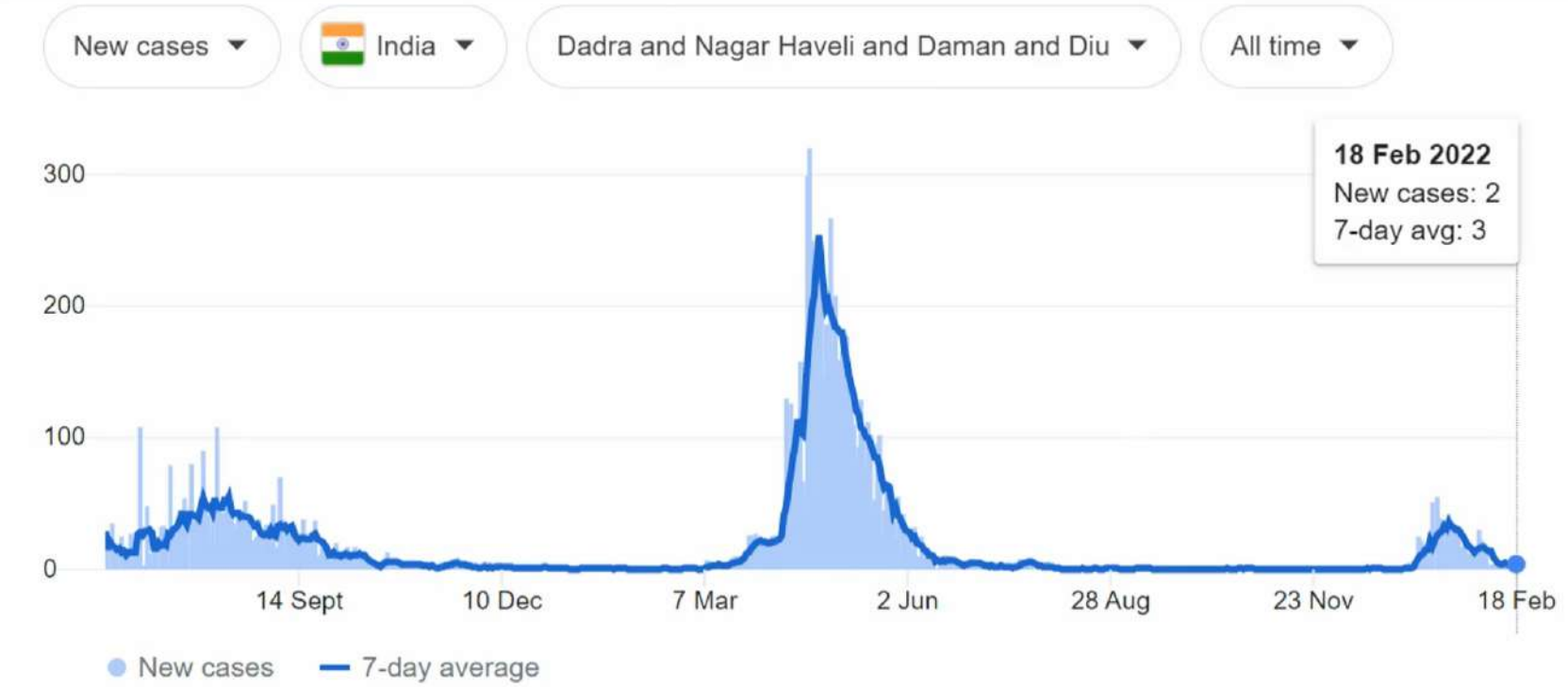
https://www.google.com/search?q=covid-19+statistics&oq=covid-19+statistics&aqs=chrome..69i57j0i512l9.5582j0j15&sourceid=chrome&ie=UTF-8

COVID-19 Statistics



https://www.google.com/search?q=covid-19+statistics&oq=covid-19+statistics&aqs=chrome..69i57j0i512l9.5582j0j15&sourceid=chrome&ie=UTF-8

COVID-19 Statistics



https://www.google.com/search?q=covid-19+statistics&oq=covid-19+statistics&aqs=chrome..69i57j0i512l9.5582j0j15&sourceid=chrome&ie=UTF-8

Result Analysis

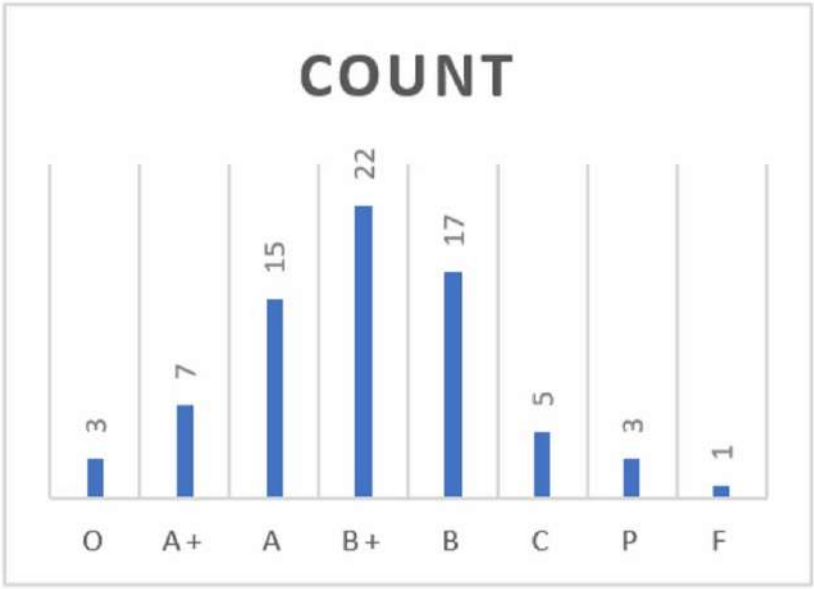
Continuous Evaluation(70.0)											Project (30.0)	Total (100.0)	Total (100.0)	Grade
C1 (6)	C2 (6)	C3 (6)	C4 (6)	C5 (6)	A1 (8)	A2 (8)	A3 (8)	A4 (8)	A5 (8)	Total(70.0)				
6	5.4	5.06	4.8	4	7.6	7.6	7.6	7.4	7.4	62.86	28	90.86	91	O
5.4	5.4	4.8	4.66	4.1	7.6	7.6	7.4	7.4	7.2	61.56	29	90.56	91	O
6	5.7	4.5	4.26	4	7.6	7.6	7.6	7.6	7.5	62.36	28	90.36	90	O
6	6	5.46	5.2	4.8	7.6	7.6	7.6	7.4	7.4	65.06	22	87.06	87	A+
6	5.4	5.2	4.3	4.26	7.6	7.6	7.6	7.5	7.4	62.86	24	86.86	87	A+
6	5.6	5.4	4.5	4.26	7.6	7.6	7.4	7.2	7.2	62.76	24	86.76	87	A+
6	6	4.8	3.9	3.2	7	7	6.9	6.8	6.8	58.4	28	86.4	86	A+
6	5.4	4.5	4.4	4.26	7.6	7.6	7.6	7.4	7.3	62.06	24	86.06	86	A+
5.4	5.4	4.8	4.8	4.4	7.6	7.6	7.5	7.2	6.8	61.5	23	84.5	85	A+
6	5.4	5.2	4.5	4.26	7	7	6.8	6.8	6.6	59.56	25	84.56	85	A+
6	5.4	5.2	4.5	4.26	7	6.8	6.8	6.8	6.6	59.36	24	83.36	83	A
6	5.4	4.66	4.5	4.4	7	7	7	6.8	6.6	59.36	24	83.36	83	A
6	6	5.4	5.06	3.9	5.6	5.6	5.6	5.4	5.4	53.96	29	82.96	83	A
6	6	5.4	5.2	5.06	6.2	6.2	6	6	5.8	57.86	24	81.86	82	A

Detailed Mark Split-up of 73 students in a class
(Note : Only 14 rows displayed)

Result Analysis

Start	End	Grade	Count	%
90	100	O	3	4.11
85	89	A+	7	9.59
78	84	A	15	20.6
70	77	B+	22	30.1
60	69	B	17	23.3
50	59	C	5	6.85
40	49	P	3	4.11
0	39	F	1	1.37

Grade distribution table



Clustered Column Chart of Grade counts

Google Scholar

Jayakumar T V

FOLLOWING

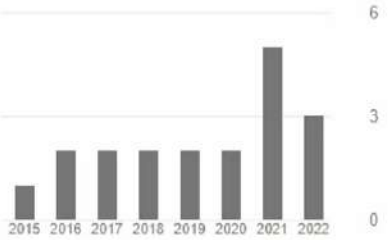
Unknown affiliation
No verified email
Medical image processing Image Compression Machine Learning

TITLE	CITED BY	YEAR
Analysis of different automatic cataract detection and classification methods CP Niya, TV Jayakumar 2015 IEEE International Advance Computing Conference (IACC), 696-700	12	2015
A study on melanoma skin cancer detection techniques PM Amulya, TV Jayakumar 2017 International Conference on Intelligent Sustainable Systems (ICISS ...	5	2017
A review on how human aging influences facial expression recognition (FER) R Mary, TV Jayakumar Innovations in bio-inspired computing and applications, 313-322	3	2016
Preserving the Privacy of Audio File Using Ideal Secret Sharing Scheme with Cloud Storage MA Mohan, TV Jayakumar, K Praveen International Conference on Soft Computing and Signal Processing, 209-216		2019

Articles 1–4 SHOW MORE

Cited by

	All	Since 2017
Citations	20	16
h-index	3	3
i10-index	1	0



Co-authors EDIT

No co-authors

https://scholar.google.com/citations?user=vYSoNJEAAAAJ&hl=en

Conclusion

- COVID-19 Statistics
- Result Analysis
- Google Scholar

AMRITA VISHWA VIDYAPEETHAM

AHEAD Online

Probability and Statistics

Dr. Ajay Vadakkepatt, Jayakumar T V, Vidya M
Department of Computer Science
Amrita Vishwa Vidyapeetham

Objective

- To introduce descriptive and inferential statistics
- Identify level of measurements
  - Nominal
  - Ordinal
  - Interval
  - Ratio

Descriptive Statistics

- Used to organize and describe a collection of data

Name	Mode	Age
Karthik	Online	35
Ram	Offline	26
Parvathi	Offline	21
Christina	Online	28
Ishaan	Offline	21
Tahir	Offline	22
John	Online	26
Tom	Online	32
Sunainah	Online	33
Shafira	Online	34

What is the most frequent choice for mode of study?

What is the average age?

Inferential Statistics

- Inferential statistics are used to make inferences about large group of data based on smaller group of data
- Smaller group is called **sample**, which is a subset of **population**
- Inferring 60% of students prefer online mode
- Which of the several names will be most appealing for our course?
  - AHEAD Online, VidyAmrita, OnLyceum
- How to find best treatment for a particular disease?

Nominal

- Defined by characteristics of an outcome that match into one and only one category
- No order between categories
- Least Precise
- Qualitative variables
- Religion?
- Political Affiliation ?
- Area of living?
- Brand of clothes?

Ordinal

- In this type of measurement things are ordered
- Cannot say anything about intervals between rankings
- More precise than nominal
- Qualitative variable
- Gold, Silver, Bronze of a 100 m sprint
- Knowledge of teacher (Poor, Satisfactory, Good, Excellent)

Interval

- Equals intervals between neighboring data points
- More precise than nominal and ordinal
- Quantitative variable
- There is no true zero point
- Zero means absence of any of the trait that is being measured
- Temperature
- Test Scores

Ratio

- Characterized by the presence of an absolute zero on the scale
- Most precise level of measurement
- Quantitative variables
- Age
- Weight
- Height

Level of Measurements

- Nominal
  - Gender, Brand of clothes
- Ordinal
  - Top 3 in a race, Feedback questions
- Interval
  - Temperature, Test Scores
- Ratio
  - Age, Height

Categorize

Rank

Equally spaced intervals

True zero point

Conclusion

- To introduce descriptive and inferential statistics
- Identify level of measurements
  - Nominal
  - Ordinal
  - Interval
  - Ratio

Probability and Statistics

Dr. Ajay Vadakkepatt, Jayakumar T V, Vidya M
Department of Computer Science
Amrita Vishwa Vidyapeetham

Objective

- To understand measures of central tendency
  - Mean
  - Median
  - Mode

Measures of central tendency

- Measure to describe whole set of data with a single data
- Average is the one value that best represents an entire data
- Mean, Median and Mode are three classes of average
- Each provides a different types of information
- Class average for particular course

Mean

- Sum of values in a group, divided by number of values
- Most common type of average computed (arithmetic mean)

$$\bar{x} = \frac{\sum x}{n}$$

- It is very sensitive to extreme scores
- Arithmetic mean is the point about which sum of deviations is equal to zero
- Other types - harmonic, geometric

Median

- It is defined as midpoint in a set of scores
- To compute the median
  - List the values in order (highest to lowest or lowest to highest)
  - Find the score at middle position (Find average of two values at middle if total number of observation is even)
- It is also known as 50'th percentile, point below which 50% of the cases in scores fall
- Extreme scores or outliers don't count

Mean vs Median

1	55	55-37.5	17.5
2	37	37-37.5	-0.5
3	33	33-37.5	-4.5
4	25	25-37.5	-12.5

140	140-58	82
55	55-58	-3
37	37-58	-21
33	33-58	-25
25	25-58	-33

- Mean $= \frac{55+37+33+25}{4} = 37.5$
- Median $= \frac{33+37}{2} = 35$

- Mean $= \frac{140+55+37+33+25}{5} = 58$
- Median $= 37$

Mode

- It is the value that occurs more frequently
- Least precise measure

Disease	Count
Normal Fever	34
Covid-19	124
Throat Pain	27

- Mode is Covid-19

Conclusion

- To understand measures of central tendency
  - Mean
  - Median
  - Mode

Probability and Statistics

Dr. Ajay Vadakkepatt, Jayakumar T V, Vidya M
Department of Computer Science
Amrita Vishwa Vidyapeetham

Objective

- To understand variability measures like
  - Range
  - Standard Deviation
  - Variance

Understanding Variability

- Average is a representative score of a set of scores
- Variability means how different scores are from one score

	Set-1	Set-2	Set-3
	4	4	5
	4	5	5
	7	5	5
	2	6	5
	8	5	5
Mean	5	5	5

Range

- It is the difference between highest and lowest values
- It should not be used regarding how different individual scores are from one another

37
3
22
78
19

Range=78-3=75

Standard Deviation

- Frequently used measure of variability
- Average distance from mean

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$$

	Set-1	$(x - \bar{x})^2$	**Set-2**	$(x - \bar{x})^2$	**Set-3**	$(x - \bar{x})^2$
	4	1	4	1	5	0
	4	1	5	0	5	0
	7	4	5	0	5	0
	2	9	6	1	5	0
	8	9	5	0	5	0
Mean	5		5		5	
s	$\sqrt{6}$		$\sqrt{0.5}$		0	

Variance

- It is the square of standard deviation

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

	Set-1	$(x - \bar{x})^2$	**Set-2**	$(x - \bar{x})^2$	**Set-3**	$(x - \bar{x})^2$
	4	1	4	1	5	0
	4	1	5	0	5	0
	7	4	5	0	5	0
	2	9	6	1	5	0
	8	9	5	0	5	0
Mean	5		5		5	
s	$\sqrt{6}$		$\sqrt{0.5}$		0	
s^2	6		0.5		0	

Conclusion

- To understand variability measures like
  - Range
  - Standard Deviation
  - Variance