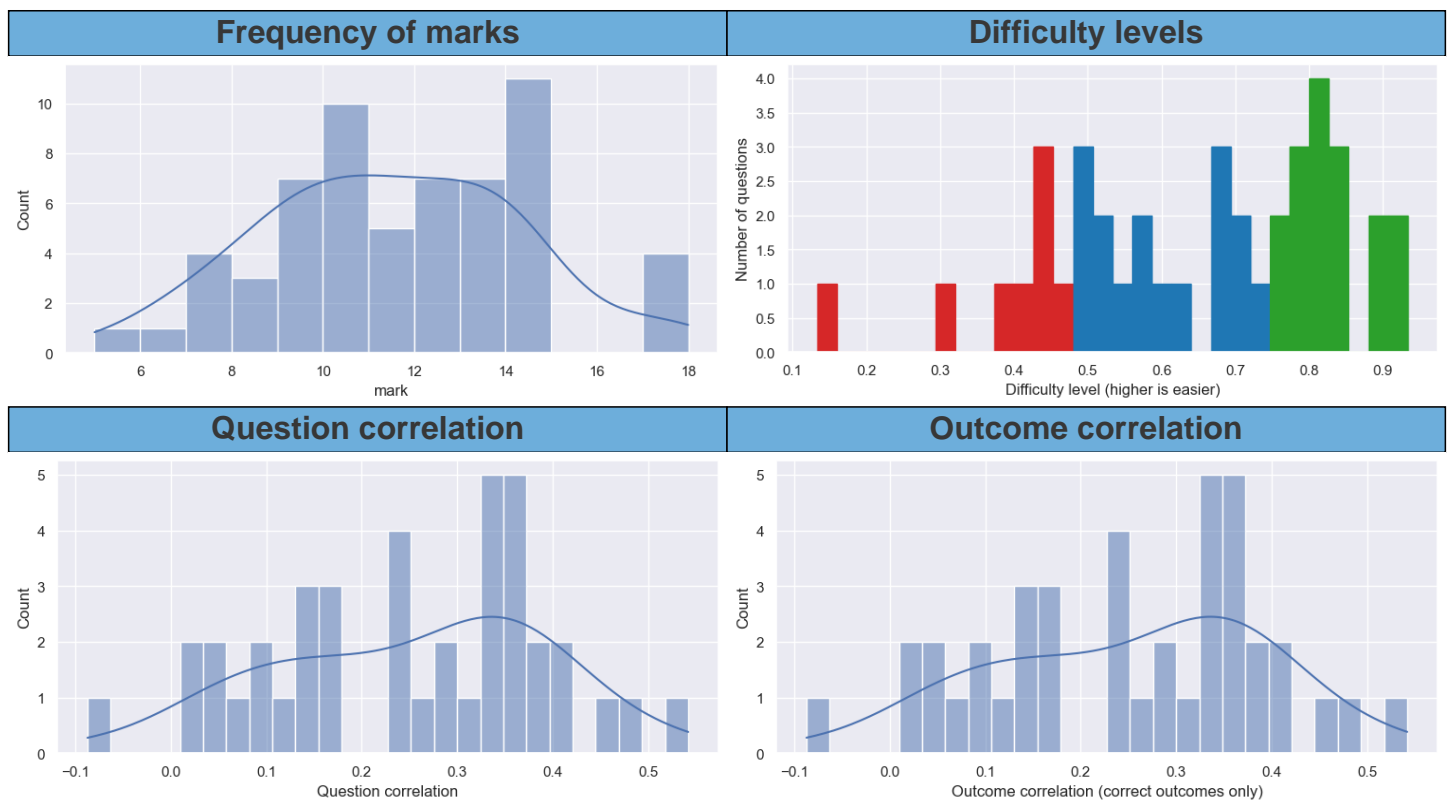




## 202304-G1P02E\_18 - Examination report

Number of examinees	Number of questions	Maximum possible mark	Minimum possible mark	Maximum achieved mark
60	40	20.0	5.0	18.0
Mean	Median	Mode	Std Dev	Variance
11.52	11.5	9.0	2.85	8.14
Std error of mean	Std error of measurement	Skew	Kurtosis	Test reliability
0.37	0.37	0.09	-0.28	0.003208



### TL;DR

On average, the exam results were 11.5 out of 20. The most common score was 9 out of 20. The scores varied quite a bit, with a standard deviation of 2.85. The distribution of scores was slightly skewed to the right, meaning that there were more low scores than high scores. The kurtosis was negative, indicating that the distribution was flatter than a normal distribution.

According to the data collected, the following questions should probably be reviewed:

- Questions Q001, Q002, Q005, Q006, Q007, Q008, Q009, Q010, Q011, Q012, Q013, Q014, Q015, Q016, Q017, Q018, Q019, Q023, Q024, Q025, Q026, Q028, Q029, Q030, Q031, Q035, Q037 and Q040 have distractors that have never been chosen.

## Question - Q001

replied	correct	empty
54 (90.0%)	46 (76.67%)	6 (10.0%)
difficulty		correlation
0.7667 (easy)		0.1495 (Low)
answer	correct	ticked
A		0 (0.0%)
B		0 (0.0%)
C		8 (13.33%)
D	*	46 (76.67%)

## Question - Q002

replied	correct	empty
58 (96.67%)	56 (93.33%)	2 (3.33%)
difficulty		correlation
0.9333 (easy)		0.0724 (None)
answer	correct	ticked
A	*	56 (93.33%)
B		0 (0.0%)
C		0 (0.0%)
D		2 (3.33%)

## Question - Q003

replied	correct	empty
40 (66.67%)	25 (41.67%)	20 (33.33%)
difficulty		correlation
0.4167 (intermediate)		0.3954 (Strong)
answer	correct	ticked
A	*	25 (41.67%)
B		8 (13.33%)
C		2 (3.33%)
D		5 (8.33%)

## Question - Q004

replied	correct	empty
50 (83.33%)	30 (50.0%)	10 (16.67%)
difficulty		correlation
0.5 (intermediate)		0.1296 (Low)
answer	correct	ticked
A	*	30 (50.0%)
B		2 (3.33%)
C		4 (6.67%)
D		14 (23.33%)

## Question - Q005

replied	correct	empty
39 (65.0%)	26 (43.33%)	21 (35.0%)
difficulty		correlation
0.4333 (intermediate)		0.3694 (Strong)
answer	correct	ticked
A		0 (0.0%)
B		0 (0.0%)
C	*	26 (43.33%)
D		13 (21.67%)

## Question - Q006

replied	correct	empty
56 (93.33%)	51 (85.0%)	4 (6.67%)
difficulty		correlation
0.85 (easy)		0.1427 (Low)
answer	correct	ticked
A		0 (0.0%)
B		1 (1.67%)
C	*	51 (85.0%)
D		4 (6.67%)

## Question - Q007

replied	correct	empty
50 (83.33%)	46 (76.67%)	10 (16.67%)
difficulty		correlation
0.7667 (easy)		-0.0873 (To be reviewed!)
answer	correct	ticked
A	*	46 (76.67%)
B		2 (3.33%)
C		0 (0.0%)
D		2 (3.33%)

## Question - Q008

replied	correct	empty
51 (85.0%)	31 (51.67%)	9 (15.0%)
difficulty		correlation
0.5167 (intermediate)		0.1767 (Low)
answer	correct	ticked
A		8 (13.33%)
B		0 (0.0%)
C		12 (20.0%)
D	*	31 (51.67%)

## Question - Q009

replied	correct	empty
46 (76.67%)	37 (61.67%)	14 (23.33%)
difficulty		correlation
0.6167 (easy)		0.2652 (Moderate)
answer	correct	ticked
A		5 (8.33%)
B		4 (6.67%)
C	*	37 (61.67%)
D		0 (0.0%)

## Question - Q010

replied	correct	empty
45 (75.0%)	36 (60.0%)	15 (25.0%)
difficulty		correlation
0.6 (intermediate)		0.4138 (Strong)
answer	correct	ticked
A		4 (6.67%)
B		5 (8.33%)
C	*	36 (60.0%)
D		0 (0.0%)

## Question - Q011

replied	correct	empty
37 (61.67%)	27 (45.0%)	23 (38.33%)
difficulty		correlation
0.45 (intermediate)		0.3619 (Strong)
answer	correct	ticked
A		1 (1.67%)
B	*	27 (45.0%)
C		9 (15.0%)
D		0 (0.0%)

## Question - Q012

replied	correct	empty
46 (76.67%)	41 (68.33%)	14 (23.33%)
difficulty		correlation
0.6833 (easy)		0.4474 (Strong)
answer	correct	ticked
A		5 (8.33%)
B	*	41 (68.33%)
C		0 (0.0%)
D		0 (0.0%)

## Question - Q013

replied	correct	empty
49 (81.67%)	49 (81.67%)	11 (18.33%)
difficulty		correlation
0.8167 (easy)		0.1475 (Low)
answer	correct	ticked
A	*	49 (81.67%)
B		0 (0.0%)
C		0 (0.0%)
D		0 (0.0%)

## Question - Q014

replied	correct	empty
55 (91.67%)	51 (85.0%)	5 (8.33%)
difficulty		correlation
0.85 (easy)		0.052 (None)
answer	correct	ticked
A		2 (3.33%)
B		2 (3.33%)
C		0 (0.0%)
D	*	51 (85.0%)

## Question - Q015

replied	correct	empty
37 (61.67%)	19 (31.67%)	23 (38.33%)
difficulty		correlation
0.3167 (difficult)		0.0847 (None)
answer	correct	ticked
A		16 (26.67%)
B		2 (3.33%)
C	*	19 (31.67%)
D		0 (0.0%)

## Question - Q016

replied	correct	empty
38 (63.33%)	31 (51.67%)	22 (36.67%)
difficulty		correlation
0.5167 (intermediate)		0.5422 (Very strong)
answer	correct	ticked
A		0 (0.0%)
B	*	31 (51.67%)
C		6 (10.0%)
D		1 (1.67%)

## Question - Q017

replied	correct	empty
50 (83.33%)	47 (78.33%)	10 (16.67%)
difficulty		correlation
0.7833 (easy)		0.3535 (Strong)
answer	correct	ticked
A		2 (3.33%)
B	*	47 (78.33%)
C		0 (0.0%)
D		1 (1.67%)

## Question - Q018

replied	correct	empty
60	55 (91.67%)	0
difficulty		correlation
0.9167 (easy)		0.3322 (Strong)
answer	correct	ticked
A		1 (1.67%)
B		4 (6.67%)
C	*	55 (91.67%)
D		0 (0.0%)

## Question - Q019

replied	correct	empty
56 (93.33%)	35 (58.33%)	4 (6.67%)
difficulty		correlation
0.5833 (intermediate)		0.3576 (Strong)
answer	correct	ticked
A		0 (0.0%)
B	*	35 (58.33%)
C		21 (35.0%)
D		0 (0.0%)

## Question - Q020

replied	correct	empty
33 (55.0%)	26 (43.33%)	27 (45.0%)
difficulty		correlation
0.4333 (intermediate)		0.3397 (Strong)
answer	correct	ticked
A		1 (1.67%)
B		3 (5.0%)
C		3 (5.0%)
D	*	26 (43.33%)

## Question - Q021

replied	correct	empty
47 (78.33%)	23 (38.33%)	13 (21.67%)
difficulty		correlation
0.3833 (difficult)		0.0378 (None)
answer	correct	ticked
A		1 (1.67%)
B		15 (25.0%)
C		8 (13.33%)
D	*	23 (38.33%)

## Question - Q022

replied	correct	empty
55 (91.67%)	47 (78.33%)	5 (8.33%)
difficulty		correlation
0.7833 (easy)		0.0961 (None)
answer	correct	ticked
A		1 (1.67%)
B	*	47 (78.33%)
C		2 (3.33%)
D		5 (8.33%)

## Question - Q023

replied	correct	empty
56 (93.33%)	54 (90.0%)	4 (6.67%)
difficulty		correlation
0.9 (easy)		0.2868 (Moderate)
answer	correct	ticked
A		0 (0.0%)
B		2 (3.33%)
C		0 (0.0%)
D	*	54 (90.0%)
E		0 (0.0%)

## Question - Q024

replied	correct	empty
51 (85.0%)	43 (71.67%)	9 (15.0%)
difficulty		correlation
0.7167 (easy)		0.2391 (Moderate)
answer	correct	ticked
A		0 (0.0%)
B	*	43 (71.67%)
C		1 (1.67%)
D		7 (11.67%)

## Question - Q025

replied	correct	empty
51 (85.0%)	41 (68.33%)	9 (15.0%)
difficulty		correlation
0.6833 (easy)		0.3777 (Strong)
answer	correct	ticked
A		3 (5.0%)
B		7 (11.67%)
C		0 (0.0%)
D	*	41 (68.33%)

## Question - Q026

replied	correct	empty
55 (91.67%)	49 (81.67%)	5 (8.33%)
difficulty		correlation
0.8167 (easy)		0.414 (Strong)
answer	correct	ticked
A		5 (8.33%)
B		1 (1.67%)
C	*	49 (81.67%)
D		0 (0.0%)

## Question - Q027

replied	correct	empty
59 (98.33%)	49 (81.67%)	1 (1.67%)
difficulty		correlation
0.8167 (easy)		0.0332 (None)
answer	correct	ticked
A		7 (11.67%)
B	*	49 (81.67%)
C		1 (1.67%)
D		2 (3.33%)

## Question - Q028

replied	correct	empty
55 (91.67%)	54 (90.0%)	5 (8.33%)
difficulty		correlation
0.9 (easy)		0.1689 (Low)
answer	correct	ticked
A		0 (0.0%)
B		1 (1.67%)
C		0 (0.0%)
D	*	54 (90.0%)



## Question - Q029

replied	correct	empty
52 (86.67%)	50 (83.33%)	8 (13.33%)
difficulty		correlation
0.8333 (easy)		0.3189 (Strong)
answer	correct	ticked
A		0 (0.0%)
B		1 (1.67%)
C		1 (1.67%)
D	*	50 (83.33%)

## Question - Q030

replied	correct	empty
48 (80.0%)	47 (78.33%)	12 (20.0%)
difficulty		correlation
0.7833 (easy)		0.0246 (None)
answer	correct	ticked
A		1 (1.67%)
B		0 (0.0%)
C		0 (0.0%)
D	*	47 (78.33%)

## Question - Q031

replied	correct	empty
55 (91.67%)	48 (80.0%)	5 (8.33%)
difficulty		correlation
0.8 (easy)		0.2828 (Moderate)
answer	correct	ticked
A		1 (1.67%)
B		0 (0.0%)
C	*	48 (80.0%)
D		6 (10.0%)

## Question - Q032

replied	correct	empty
45 (75.0%)	30 (50.0%)	15 (25.0%)
difficulty		correlation
0.5 (intermediate)		0.33 (Strong)
answer	correct	ticked
A		6 (10.0%)
B		5 (8.33%)
C		4 (6.67%)
D	*	30 (50.0%)

## Question - Q033

replied	correct	empty
47 (78.33%)	28 (46.67%)	13 (21.67%)
difficulty		correlation
0.4667 (intermediate)		0.2307 (Moderate)
answer	correct	ticked
A		19 (31.67%)
B	*	28 (46.67%)

## Question - Q034

replied	correct	empty
42 (70.0%)	40 (66.67%)	18 (30.0%)
difficulty		correlation
0.6667 (easy)		0.1604 (Low)
answer	correct	ticked
A		2 (3.33%)
B	*	40 (66.67%)

## Question - Q035

replied	correct	empty
50 (83.33%)	42 (70.0%)	10 (16.67%)
difficulty		correlation
0.7 (easy)		0.3382 (Strong)
answer	correct	ticked
A		1 (1.67%)
B		0 (0.0%)
C		7 (11.67%)
D	*	42 (70.0%)

## Question - Q036

replied	correct	empty
46 (76.67%)	8 (13.33%)	14 (23.33%)
difficulty		correlation
0.1333 (difficult)		0.249 (Moderate)
answer	correct	ticked
A	*	8 (13.33%)
B		11 (18.33%)
C		2 (3.33%)
D		25 (41.67%)

## Question - Q037

replied	correct	empty
52 (86.67%)	44 (73.33%)	8 (13.33%)
difficulty		correlation
0.7333 (easy)		0.2367 (Moderate)

answer	correct	ticked
A		5 (8.33%)
B		0 (0.0%)
C		3 (5.0%)
D	*	44 (73.33%)

Question - Q038		
replied	correct	empty
50 (83.33%)	32 (53.33%)	10 (16.67%)
difficulty		correlation
0.5333 (intermediate)		0.3599 (Strong)
answer	correct	ticked
A	*	32 (53.33%)
B		3 (5.0%)
C		1 (1.67%)
D		14 (23.33%)

Question - Q039		
replied	correct	empty
44 (73.33%)	30 (50.0%)	16 (26.67%)
difficulty		correlation
0.5 (intermediate)		0.4714 (Very strong)
answer	correct	ticked
A		9 (15.0%)
B	*	30 (50.0%)
C		5 (8.33%)

Question - Q040		
replied	correct	empty
40 (66.67%)	35 (58.33%)	20 (33.33%)
difficulty		correlation
0.5833 (intermediate)		0.3397 (Strong)
answer	correct	ticked
A		4 (6.67%)
B		0 (0.0%)
C		1 (1.67%)
D	*	35 (58.33%)

## Definitions

### Standard Deviation

Standard deviation is a statistical measure that describes the amount of variability or dispersion in a set of data. It is a measure of how spread out the data is from the mean or average value. A higher standard deviation means that the data points are more spread out and a lower standard deviation indicates that the data points are closer to the mean.

In the context of Classical Test Theory, standard deviation is used to measure the variability of scores on a test. It helps to evaluate the consistency and quality of a test by showing the spread of scores around the mean. If there is a high standard deviation, it indicates that the test has a wide range of scores and may have items that are too difficult or too easy for the test takers. On the other hand, a low standard deviation suggests that the test has a narrow range of scores and may not be sensitive enough to distinguish between high and low performing test takers.

### Variance

Variance is another statistical measure of the spread or variability in a set of data. It is the average of the squared differences from the mean or average value. In other words, it quantifies how much the data points differ from their mean value.

Variance is used to evaluate the reliability of a test score. It measures the extent to which the scores obtained by different test takers vary from each other, or from their mean score. A higher variance indicates that the test scores are more spread out and less consistent, while a lower variance suggests that the test takers' scores are more consistent and closer to each other.

Variance is an important statistical concept in understanding the quality of a test score and diagnosing measurement errors. A reliable test should have a low variance, indicating that the test scores are less likely to be affected by random errors or chance factors.

### Skewness and kurtosis

Skewness and kurtosis are two statistical measures that describe the shape of a probability distribution or frequency distribution of a set of data.

Skewness is a measure of the asymmetry of the distribution, or how much it deviates from the normal or symmetrical distribution. If the tail of the distribution is longer on the left, it is negatively skewed, while if it is longer on the right, it is positively skewed. A perfectly symmetrical distribution has zero skewness.

Kurtosis, on the other hand, measures the peakedness of the distribution, or how much it deviates from the normal distribution in terms of the distribution's tails. A distribution with high kurtosis has more extreme values (in the tails) than a normal distribution, while a distribution with low kurtosis is less peaked and has lighter tails. A normal distribution has a kurtosis of 3, which is sometimes called "mesokurtic."

In the context of Classical Test Theory, skewness and kurtosis are used to evaluate the normality or symmetry of score distributions. For example, if a set of scores on a test has a high positive skewness, this may indicate that the test items were too easy and many test takers scored high, or that there might be some measurement error in the test. Conversely, if the scores have a high

negative skewness, it suggests that the test items were too difficult or that there might be systematic negative biases in the test results. A normal score distribution is ideal for most applications, as it indicates that the test measures the intended construct effectively and there are no biases or measurement errors.

Kurtosis is also relevant for test score distributions, as high kurtosis or high peakedness can indicate that the test may lack discriminating power and have ceiling or floor effects. Ceiling effects occur when many test takers obtain the highest possible score, making it difficult to differentiate between high-achieving test takers. Floor effects occur when many test takers obtain the lowest possible score, making it difficult to differentiate between low-achieving test takers. Thus, evaluating the skewness and kurtosis of test score distributions can help diagnose measurement issues and improve the quality of the test.

## **Discrimination**

'Discrimination' refers to the ability of a test item to differentiate between high performing and low performing test takers. Discrimination is a measure of the extent to which an item measures the intended construct effectively.

Typically, items with a high level of discrimination are those that are answered correctly by high-performing test takers but answered incorrectly by low-performing test takers. If an item is answered correctly by all test takers or answered incorrectly by all, it indicates a lack of discrimination. A well-designed test should have a balance of easy, moderate, and difficult items that can differentiate between test takers with different levels of ability and accurately measure the intended construct.

Discrimination is evaluated using the point-biserial correlation coefficient, which measures the correlation between the scores of an item and the total score on the test. An item with a high point-biserial correlation is considered to be a good discriminator.

## **Correlation**

Correlation is a statistical measure that describes the degree of association or relationship between two variables. It measures the direction and strength of the linear relationship between the variables. A correlation coefficient is a numerical value between -1 and +1 that indicates the degree of correlation between the variables.

In positive correlation, both variables move in the same direction (either both increase or both decrease). In negative correlation, the variables move in opposite directions (one increases while the other decreases). A correlation coefficient of 0 indicates that there is no linear relationship between the two variables.

## **Test Reliability**

Test reliability is a measure of the consistency or stability of a test over time. If a test is reliable, it should give consistent results when administered to the same group of individuals at different times or when different raters score the test independently. There are several methods for evaluating test reliability, and one of the most widely used is Cronbach's alpha coefficient. Cronbach's alpha coefficient is a statistical measure of the internal consistency of a test. It measures the extent to which items in a test that are intended to measure the same construct are similar in nature or

correlated with each other. Alpha coefficient ranges from 0 to 1, with higher values indicating better internal consistency. A Cronbach's alpha coefficient of 0.70 or higher is generally considered to be acceptable for most purposes. The alpha coefficient is calculated by analyzing the correlations between all possible pairs of items in a test. If the items are highly correlated, it suggests that they are measuring the same construct, and the test has strong internal consistency. If the items are not highly correlated, it indicates that the test has low internal consistency and may need to be revised or improved.