1)

The correlation coefficient of 0.7 indicates a strong positive relationship between the SAT score and the college GPA. when the correlation is positive, it means if SAT scores increase, the college GPA tends to increase.

2)

$\bar{x}$ = 170 cm, σ = 10 cm,
Sample size (n) = 1000,
normally distributed = True

a)

We can say, By using Emperical formula (*because the dataset is normally distributed and* $\bar{x}$ & σ *is given*) the percentage of individuals in the dataset have between 160 cm and 180 cm is **68%** approx.

b)

Sample size (n) = 100
$\bar{x}$ = 170
σ for the sample = $10 \div \sqrt{100}$ = 1 cm
Z-score = (X - μ) /(σ / $\sqrt{n}$) = 175 -170 / 1 = 5
Using z-table we get 0.99 ≃ 100 %

c)

x =185
Z-score = (X - μ) /(σ / $\sqrt{n}$) = (185-175) / 10 = 1.5

d)

1 - 0.025 = 9750
On checking z-table we get = -1.645
Therefore,

Z-score = (X - μ) /(/ $\sqrt{n}$)
X = z * σ + μ
X = -1.645 * 10 +170
X = 153.55 cm

e)

cv = $\frac{\sigma}{\mu}$ * 100
cv = 10/170 *100
cv = 5.88 %

f)

Skewness = 3*(mean - median) / σ
Since data is normally distributed, mean = median
Skewness = 3 * (mean - mean) / σ = 0

3)

a)

Dispersion in "Blood Pressure Before":
Range: The range is the difference between the maximum and minimum values.
Maximum value: 148
Minimum value: 120
Range: 148 - 120 = 28

Variance = $\Sigma(Xi-\bar{X})^2 / N$

Variance for Blood Pressure Before (mmHg) = 43.1019
Variance for Blood Pressure After (mmHg) = 46.9704

Standard deviation = variance ^ 0.5

Standard deviations for Blood Pressure Before (mmHg) = 6.565203729
Standard deviation for Blood Pressure After(mmHg) = 6.853495459

Range:
The range provides a measure of how much the values spread out from the minimum to the maximum. A larger range indicates a wider dispersion of values.
Variance:
The variance measures the average squared deviation from the mean. A higher variance indicates a greater dispersion or spread of values around the mean.
Standard Deviation:
The standard deviation is the square root of the variance. It measures the dispersion or spread of values from the mean. A higher standard deviation indicates a wider spread of values.

b)

For Blood Pressure Before (mmHg)

Mean = (130 + 142 + 120 + ... + 135) / 100 = 132.58
Standard Deviation= 6.565203729
Standard Error = Standard Deviation / Square root of the number of values
= 0.6598278012
Margin of Error = 1.96 * Standard Error
1.96 * 6.598278012 = 12.93262490352 $\simeq 13$

To calculate the confidence interval:
Lower Bound = Mean - Margin of Error = 131.28
Upper Bound = Mean + Margin of Error = 133.88

For Blood Pressure After(mmHg)

Mean = (120 + 135 + 118 ... + 130) / 100 = 128.36
Standard Deviation= 6.565203729
Standard Error = Standard Deviation / Square root of the number of values
= 6.888022103
Margin of Error = 1.96 * Standard Error
1.96 * 6.888022103 = 13.50052332188 $\simeq 13.5$

To calculate the confidence interval:
Lower Bound = Mean - Margin of Error = 114.86
Upper Bound = Mean + Margin of Error = 141.86

C)

MAD = (Σ |xi – x|) / n)

Where xi is the ith value of column
X is mean
N is the count

For Blood Pressure Before (mmHg)

MAD = `3.4106051316484808e-15`

For Blood Pressure After(mmHg)

MAD = `1.3642420526593923e-14`

4)

Perfect squares = 1,4,9 &16

No of favorable outcome = 4
Total no of outcome = 20

Probability = $\frac{No\ of\ favorable\ outcome}{Total\ no\ of\ outcome}$ = $\frac{4}{20}$ = 0.2 or 20 %

5)

A → Taxi belongs to company A
B → Taxi belongs to company B
L → Taxi is Late

P(A/L) = ?

Using Baye's theorem,
P(A/L) = P(L/A) * P(A) /P(B)

P(L) = P(L/A) * P(A) + P(L/B) * P(B)
= 1-0.95 *0.80 + 1-0.90 * 0.2
= 0.05 *0.80 + 0.10 * 0.2
= 0.04 + 0.02
= 0.06
Therefore,
P(A/L) = 0.06 * 0.80/0.06 = 80 %

Probability of Late taxi belongs to company A is 80 %

7)

2x + 3 - 8 = 0 ⇒ x = 5/2
2y + x - 5 = 0 ⇒ y = 5/4

Variance of x = 4
Var(y) = Var(x) * (slope of y)^2
= 4 * (5/4) ^2
= 4 * 25/16 = 25/4

8)

Here are the steps to perform the Wilcoxon signed-rank test:

Step 1: Calculate the absolute values of the differences and rank them from smallest to largest, preserving the sign:

Difference (absolute)	Rank
1	4
1	5
1	6
2	7
2	8
2	9
2	10
2	11
3	12
3	13

Step 3: Calculate the sum of positive ranks (W+) and the sum of negative ranks (W-):

W+ = 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 = 85

W- = 0 (there are no negative ranks)

Step 4: Calculate the test statistic (T):

T = min (W+, W-) = min (85, 0) = 0

Step 5: Determine the critical value or p-value:

The critical value or p-value is obtained from the Wilcoxon signed-rank table based on the sample size and desired significance level (e.g., $\alpha = 0.05$). Since the sample size is 10, we look for the critical value in the row corresponding to $n = 10$.

Based on the table, if the test statistic (T) is less than or equal to the critical value, we reject the null hypothesis (H0) and conclude that the therapy had a significant effect on anxiety levels. Otherwise, if the test statistic is greater than the critical value, we fail to reject H0.

Since $T = 0$ and there is no critical value less than or equal to 0, we can conclude that the therapy had a significant effect on anxiety levels.

9)

Step 1:
Set up the null and alternative hypotheses:
Null hypothesis (H0): The mean scores of all the students are the same.
Alternative hypothesis (H1): The mean scores of the students are not all the same.

Step 2:
Calculate the mean score for each exam:
Mean Exam 1 = (85 + 70 + 90 + 75 + 95) / 5 = 83
Mean Exam 2 = (90 + 80 + 85 + 70 + 92) / 5 = 83.4
Mean Final Exam = (92 + 85 + 88 + 75 + 96) / 5 = 87.2

Step 3:
Calculate the sum of squares between groups (SSG):
SSG = (n1 * (mean1 - overall_mean)^2) + (n2 * (mean2 - overall_mean)^2) + (n3 * (mean3 - overall_mean)^2)
= (5 * (83 - 84.12)^2) + (5 * (83.4 - 84.12)^2) + (5 * (87.2 - 84.12)^2)
= 5.56 + 0.1344 + 8.7616
= 14.455

Step 4:
Calculate the sum of squares within groups (SSW):
SSW = (n1 - 1) * variance1 + (n2 - 1) * variance2 + (n3 - 1) * variance3
= (4 * 45.2) + (4 * 45.6) + (4 * 53.2)

= 180.8 + 182.4 + 212.8
= 575.6

Step 5:
Calculate the degrees of freedom between groups (dfG) and within groups (dfW):
dfG = number of groups - 1 = 3 - 1 = 2
dfW = total number of observations - number of groups = 15 - 3 = 12

Step 6:
Calculate the mean squares between groups (MSG) and within groups (MSW):
MSG = SSG / dfG = 14.455 / 2 = 7.2275
MSW = SSW / dfW = 575.6 / 12 = 47.9667

Step 7:
Calculate the F-statistic:
F = MSG / MSW = 7.2275 / 47.9667 = 0.1504

Step 8:
Determine the critical value or p-value:
To determine the critical value or p-value for the F-statistic, we need to consult an F-distribution table or use statistical software. We can compare the calculated F-value with the critical value or obtain the p-value.
If the calculated F-value is greater than the critical value or the p-value is less than the chosen significance level (e.g., $\alpha = 0.05$), we reject the null hypothesis and conclude that the mean scores of the students are not all the same. Regarding the student with the highest score, based on the given data, Jeevan has the highest scores in all three exams (Exam 1: 95, Exam 2: 92, Final Exam: 96)

10)

Using Binomial probability,

P(x=k) = (nCk * p^k * (1-p)^(n-k)

Given,
n = 500
k = 20
p(x=20) = (500C20) * 0.05^20 * (1-0.05)^(500-20)

12)

Step 1: Define the null and alternative hypotheses:

Null hypothesis (H_0): There is no significant difference in the mean improvement scores between Group A and Group B.
Alternative hypothesis (H_1): There is a significant difference in the mean improvement scores between Group A and Group B.

Step 2: Set the significance level (α) to 0.05.

Step 3: Perform the t-test:
We'll use a two-sample independent t-test because we have two independent groups (Group A and Group B) and want to compare their means.

Using the given information:

For Group A:

Sample size (n_1) = 30
Mean improvement score ($\bar{x}_1$) = 2.5
Standard deviation (s_1) = 0.8
For Group B:

Sample size (n_2) = 30
Mean improvement score ($\bar{x}_2$) = 2.2
Standard deviation (s_2) = 0.6
Using the formula for the two-sample independent t-test:

$t = (\bar{x}_1 - \bar{x}_2) / \sqrt{[(s_1^2 / n_1) + (s_2^2 / n_2)]}$

Substituting the values:

$t = (2.5 - 2.2) / \sqrt{[(0.8^2 / 30) + (0.6^2 / 30)]}$

Calculating the result:

$t \approx 1.86$

Step 4: Determine the critical value and compare it with the calculated t-value.
Since we have a significance level (α) of 0.05 and the test is two-tailed, we divide α by 2 to get 0.025. With degrees of freedom (df) = $n_1 + n_2 - 2 = 30 + 30 - 2 = 58$, we find the critical value using a t-distribution table or statistical software.

For a two-tailed test and $\alpha/2 = 0.025$, the critical t-value is approximately ±2.000.

Since |t| (absolute value of t) = 1.86 < 2.000, the calculated t-value does not exceed the critical t-value.

Step 5: Make a decision and interpret the results:
The calculated t-value does not fall in the critical region, so we fail to reject the null hypothesis (H_0). We do not have enough evidence to suggest a significant difference in the mean improvement scores between Group A and Group B.