**Research Project**

Data set based on the daily number of visitors to a popular amusement park over a span of 30 days.

**Visitor Data:**

Day 1: 1000 visitors

Day 2: 1200 visitors

Day 3: 800 visitors

Day 4: 900 visitors

Day 5: 1100 visitors

Day 6: 800 visitors

Day 7: 900 visitors

Day 8: 600 visitors

Day 9: 1200 visitors

Day 10: 1000 visitors

Day 11: 700 visitors

Day 12: 1200 visitors

Day 13: 800 visitors

Day 14: 900 visitors

Day 15: 400 visitors

Day 16: 550 visitors

Day 17: 640 visitors

Day 18: 1250 visitors

Day 19: 900 visitors

Day 20: 1100 visitors

Day 21: 800 visitors

Day 22: 600 visitors

Day 23: 450 visitors

Day 24: 760 visitors

Day 25: 1250 visitors

Day 26: 1000 visitors

Day 27: 900 visitors

Day 28: 800 visitors

Day 29: 1200 visitors

Day 30: 950 visitors

**Word Problems:**

**Question 1:** What is the average number of daily visitors to the amusement park over the 30-day period?

**Answer:** To find the sum, you can add up all the visitor counts:

Total visitors = 1000 + 1200 + 800 + 900 + 1100 + 800 + 900 + 600 + 1200 + 1000 + 700 + 1200 + 800 + 900 + 400 + 550 + 640 + 1250 + 900 + 1100 + 800 + 600 + 450 + 760 + 1250 + 1000 + 900 + 800 + 1200 + 950

Total visitors = 27,050

Divide the total number of visitors by the total number of days:

Average number of daily visitors = Total visitors / Number of days

Average number of daily visitors = 27,050 / 30

Average number of daily visitors = 901.67

Therefore, the average number of daily visitors to the amusement park over the 30-day period is approximately 901.67.

**Question 2:** How much does the number of daily visitors vary from the average?

**Answer:** First, calculate the average number of daily visitors using the previous calculation:

Average number of daily visitors = 901.67

Next, calculate the deviation of each daily visitor count from the average. To do this, subtract the average from each visitor count:

Day 1: 1000 - 901.67 = 98.33

Day 2: 1200 - 901.67 = 298.33

Day 3: 800 - 901.67 = -101.67

Day 4: 900 - 901.67 = -1.67

Day 5: 1100 - 901.67 = 198.33

Day 6: 800 - 901.67 = -101.67

Day 7: 900 - 901.67 = -1.67

Day 8: 600 - 901.67 = -301.67

Day 9: 1200 - 901.67 = 298.33

Day 10: 1000 - 901.67 = 98.33

Day 11: 700 - 901.67 = -201.67

Day 12: 1200 - 901.67 = 298.33

Day 13: 800 - 901.67 = -101.67

Day 14: 900 - 901.67 = -1.67

Day 15: 400 - 901.67 = -501.67

Day 16: 550 - 901.67 = -351.67

Day 17: 640 - 901.67 = -261.67

Day 18: 1250 - 901.67 = 348.33

Day 19: 900 - 901.67 = -1.67

Day 20: 1100 - 901.67 = 198.33

Day 21: 800 - 901.67 = -101.67

Day 22: 600 - 901.67 = -301.67

Day 23: 450 - 901.67 = -451.67

Day 24: 760 - 901.67 = -141.67

Day 25: 1250 - 901.67 = 348.33

Day 26: 1000 - 901.67 = 98.33

Day 27: 900 - 901.67 = -1.67

Day 28: 800 - 901.67 = -101.67

Day 29: 1200 - 901.67 = 298.33

Day 30: 950 - 901.67 = 48.33

Now, calculate the average deviation by summing up all the deviations and dividing by the number of days:

Average deviation = (98.33 + 298.33 - 101.67 - 1.67 + 198.33 - 101.67 - 1.67 - 301.67 + 298.33 + 98.33 - 201.67 + 298.33 - 101.67 - 1.67 - 501.67 - 351.67 - 261.67 + 348.33 - 1.67 + 198.33 - 101.67 - 301.67 - 451.67 - 141.67 + 348.33 + 98.33 - 1.67 - 101.67 + 298.33 + 48.33) / 30

Average deviation ≈ 105.11

The average deviation from the average number of daily visitors is approximately 105.11. This value indicates the average amount by which the daily visitor counts deviate from the mean. It provides a measure of the variability or spread of the data points around the average.

**Question 3:** How consistent or scattered are the daily visitor counts?

**Answer:** First, let's calculate the squared deviation for each day:

Day 1: (1000 - 901.67)^2 = 79025.56

Day 2: (1200 - 901.67)^2 = 98460.11

Day 3: (800 - 901.67)^2 = 10345.56

Day 4: (900 - 901.67)^2 = 0.11

Day 5: (1100 - 901.67)^2 = 38055.56

Day 6: (800 - 901.67)^2 = 10345.56

Day 7: (900 - 901.67)^2 = 0.11

Day 8: (600 - 901.67)^2 = 91222.22

Day 9: (1200 - 901.67)^2 = 98460.11

Day 10: (1000 - 901.67)^2 = 9596.11

Day 11: (700 - 901.67)^2 = 40806.11

Day 12: (1200 - 901.67)^2 = 98460.11

Day 13: (800 - 901.67)^2 = 10345.56

Day 14: (900 - 901.67)^2 = 0.11

Day 15: (400 - 901.67)^2 = 169840.11

Day 16: (550 - 901.67)^2 = 126850.11

Day 17: (640 - 901.67)^2 = 68115.56

Day 18: (1250 - 901.67)^2 = 197575.56

Day 19: (900 - 901.67)^2 = 0.11

Day 20: (1100 - 901.67)^2 = 38055.56

Day 21: (800 - 901.67)^2 = 10345.56

Day 22: (600 - 901.67)^2 = 91222.22

Day 23: (450 - 901.67)^2 = 180740.11

Day 24: (760 - 901.67)^2 = 20095.56

Day 25: (1250 - 901.67)^2 = 197575.56

Day 26: (1000 - 901.67)^2 = 9596.11

Day 27: (900 - 901.67)^2 = 0.11

Day 28: (800 - 901.67)^2 = 10345.56

Day 29: (1200 - 901.67)^2 = 98460.11

Day 30: (950 - 901.67)^2 = 2206.11

Next, calculate the sum of squared deviations:

Sum of squared deviations = 79025.56 + 98460.11 + 10345.56 + 0.11 + 38055.56 + 10345.56 + 0.11 + 91222.22 + 98460.11 + 9596.11 + 40806.11 + 98460.11 + 10345.56 + 0.11 + 169840.11 + 126850.11 + 68115.56 + 197575.56 + 0.11 + 38055.56 + 10345.56 + 91222.22 + 180740.11 + 20095.56 + 197575.56 + 9596.11 + 0.11 + 10345.56 + 98460.11 + 2206.11

Sum of squared deviations = 1789710.89

Calculate the variance by dividing the sum of squared deviations by the number of observations (30 in this case):

Variance = 1789710.89 / 30

Variance ≈ 59657.03

The variance measures the average squared deviation from the mean, and in this case, it indicates the level of scatter or consistency in the daily visitor counts. A higher variance value suggests greater variability or scatter in the data, while a lower value indicates more consistency.

**Question 4:** In how many different orders can the top three busiest days be arranged?

**Answer:** In this case, the top three busiest days are the days with the highest number of visitors:

Day 18: 1250 visitors

Day 25: 1250 visitors

Day 2: 1200 visitors

To calculate the number of permutations, you can use the formula for permutations of n objects taken r at a time:

P(n, r) = n! / (n - r)!

Where n is the total number of objects (in this case, the total number of days) and r is the number of objects taken at a time (in this case, 3).

Substituting the values, we get:

P(30, 3) = 30! / (30 - 3)!

= 30! / 27!

= (30 \* 29 \* 28 \* 27!) / 27!

= 30 \* 29 \* 28

= 24,360

Therefore, there are 24,360 different orders in which the top three busiest days can be arranged.

**Question 5:** How many different groups of five days can be selected from the 30 days?

**Answer:** The combination formula, also known as "n choose k" or denoted as C(n, k), calculates the number of ways to choose k elements from a set of n elements.

In this case, we want to choose 5 days from a total of 30 days, so we can calculate it as:

C(30, 5) = 30! / (5! \* (30 - 5)!)

where "!" represents the factorial function.

Plugging in the values:

C(30, 5) = 30! / (5! \* 25!)

Calculating the factorials:

C(30, 5) = (30 \* 29 \* 28 \* 27 \* 26) / (5 \* 4 \* 3 \* 2 \* 1)

Simplifying:

C(30, 5) = 142,506

Therefore, there are 142,506 different groups of five days that can be selected from the 30-day period.

**Question 6:** What is the total number of unique visitors who visited the amusement park during the 30-day period?

**Answer:** To find the total number of unique visitors, we can create a set of all the visitors and then calculate the size of that set.

Visitor Set = {1000, 1200, 800, 900, 1100, 600, 700, 400, 550, 640, 1250, 450, 760, 950}

Total number of unique visitors = Size of Visitor Set

Calculating the size of the Visitor Set:

Total number of unique visitors = 14

Therefore, the total number of unique visitors who visited the amusement park during the 30-day period is 14..

**Question 7:** How many visitors attended the amusement park on both Day 15 and Day 22?

**Answer:** To determine how many visitors attended the amusement park on both Day 15 and Day 22, we need to find the intersection of the visitors on those two specific days.

Visitor Data:

Day 15: 400 visitors

Day 22: 600 visitors

Let's find the intersection of the visitors on Day 15 and Day 22.

Visitor Set on Day 15 = {400}

Visitor Set on Day 22 = {600}

Intersection of the Visitor Sets = Visitor Set on Day 15 ∩ Visitor Set on Day 22

Calculating the intersection:

Visitor Set on Day 15 ∩ Visitor Set on Day 22 = {}

The intersection of the Visitor Sets is an empty set, indicating that there were no visitors who attended the amusement park on both Day 15 and Day 22.

Therefore, the number of visitors who attended the amusement park on both Day 15 and Day 22 is 0.

**Question 8:** What is the probability that a randomly chosen day had fewer than 1000 visitors?

**Answer:** Number of days with fewer than 1000 visitors: 14 (Day 3, 4, 6, 8, 11, 13, 15, 16, 17, 21, 22, 23, 24, 28)

Total number of days: 30

Probability = Number of days with fewer than 1000 visitors / Total number of days

Probability = 14 / 30

Probability = 0.4667 or 46.67%

Therefore, the probability that a randomly chosen day had fewer than 1000 visitors is approximately 0.4667 or 46.67%.

**Question 9:** What is the probability of exactly 25 days having more than 900 visitors?

**Answer:** We need to determine the probability of a day having more than 900 visitors and then apply the binomial distribution formula.

Given that there are 30 days in total and 15 days have more than 900 visitors, the probability of success (a day having more than 900 visitors) is:

Probability of success (p) = Number of successful outcomes / Total number of outcomes

= 15 / 30

= 0.5

Now, we can use the binomial distribution formula:

P(X = k) = (n choose k) \* p^k \* (1-p)^(n-k)

Where:

n is the total number of trials (30 days)

k is the number of successful trials (25 days)

p is the probability of success on each trial (0.5)

Let's calculate the probability:

P(X = 25) = (30 choose 25) \* (0.5)^25 \* (1-0.5)^(30-25)

Using the binomial coefficient (n choose k):

(30 choose 25) = 30! / (25! \* (30-25)!)

= (30 \* 29 \* 28 \* 27 \* 26) / (5 \* 4 \* 3 \* 2 \* 1)

= 142,506

Now, we can substitute the values into the formula:

P(X = 25) = 142,506 \* (0.5)^25 \* (0.5)^5

Calculating this, we find that the probability of exactly 25 days having more than 900 visitors is approximately 0.0916 or 9.16%.

**Question 10:** Given that it rained on Day 10 and it rained 8 times throughout the month, what is the probability that the number of visitors was less than 1100?

**Answer:** Total count of rainy days: 8 out of 30 days and It rained on Day 10

We want to find the probability that the number of visitors was less than 1100 given the above conditions.

Let's define the events:

A: The event that the number of visitors is less than 1100.

B: The event that it rained on Day 10.

C: The event that there were 8 rainy days in total.

We want to calculate P(A | B ∩ C), the probability of A given that both B and C occurred.

To calculate this, we need to determine P(A ∩ B ∩ C) and P(B ∩ C).

P(A ∩ B ∩ C) represents the probability that the number of visitors is less than 1100, it rained on Day 10, and there were 8 rainy days in total. From the given data, the number of visitors on Day 10 is 1000, which satisfies the condition of being less than 1100. We also know that there were 8 rainy days in total.

Let's calculate P(B ∩ C), the probability that it rained on Day 10 and there were 8 rainy days in total. Assuming each day's weather is independent, the probability of rain on Day 10 is 1/30, and the probability of having 8 rainy days out of 30 days is given by the binomial distribution:

P(B ∩ C) = (30 choose 8) \* (1/30)^8 \* (29/30)^22

Now, let's calculate P(A | B ∩ C) using the formula:

P(A | B ∩ C) = P(A ∩ B ∩ C) / P(B ∩ C)

Since P(A ∩ B ∩ C) = P(B ∩ C), the conditional probability simplifies to:

P(A | B ∩ C) = P(B ∩ C) / P(B ∩ C) = 1

Therefore, the probability that the number of visitors was less than 1100 given that it rained on Day 10 and there were 8 rainy days in total is 1 or 100%

**Question 11:** What is the probability of having exactly 1100 visitors on a randomly selected day?

**Answer:** Create the frequency distribution:

Count the number of occurrences of each visitor count. For example, there is 1 occurrence of 1100 visitors, 1 occurrence of 1200 visitors, 3 occurrences of 800 visitors, and so on.

Visitor count | Frequency

400 | 1

450 | 1

550 | 1

600 | 2

640 | 1

700 | 1

760 | 1

800 | 5

900 | 5

950 | 1

1000 | 3

1100 | 2

1200 | 3

1250 | 2

Calculate the probability mass function (PMF):

The PMF assigns a probability to each possible visitor count. To calculate the PMF, we divide the frequency of each visitor count by the total number of days.

Visitor count | Frequency | PMF

400 | 1 | 1/30

450 | 1 | 1/30

550 | 1 | 1/30

600 | 2 | 2/30

640 | 1 | 1/30

700 | 1 | 1/30

760 | 1 | 1/30

800 | 5 | 5/30

900 | 5 | 5/30

950 | 1 | 1/30

1000 | 3 | 3/30

1100 | 2 | 2/30

1200 | 3 | 3/30

1250 | 2 | 2/30

Calculate the probability of exactly 1100 visitors:

From the PMF, we can see that the probability of having exactly 1100 visitors on a randomly selected day is 2/30 or approximately 0.0667 (6.67%).

Using the probability mass function, we can explicitly assign probabilities to each visitor count based on the given data. This allows us to calculate the probability of specific outcomes, such as exactly 1100 visitors, by dividing the frequency of that outcome by the total number of days.

In this case, the probability of having exactly 1100 visitors is 2/30 or approximately 0.0667 (or 6.67%). This means that if we randomly select a day from the 30-day period, there is a 6.67% chance that the number of visitors on that day will be exactly 1100.

**Question 12:** What is the probability of having an average of 1000 visitors per day over the 30-day period?

**Answer:** The Poisson probability formula is:

P(x, λ) = (e^(-λ) \* λ^x) / x!

In this case, we want to calculate the probability of having x = 1000 visitors per day for 30 days, assuming an average rate of λ = 1000.

Calculate the value of e^(-λ):

e^(-λ) = e^(-1000)

Note: The value of e is approximately 2.71828.

Calculate the value of λ^x:

λ^x = 1000^30

Calculate the value of x! (factorial of x):

x! = 30!

Note: The factorial of a number is the product of all positive integers less than or equal to that number.

Substitute the calculated values into the Poisson probability formula:

P(1000, 1000) = (e^(-1000) \* 1000^30) / 30!

Calculate e^(-1000):

e^(-1000) ≈ 4.5399929762484854e-435

Calculate 1000^30:

1000^30 = 1e90 (1 followed by 90 zeros)

Calculate 30! (factorial of 30):

30! = 265252859812191058636308480000000

Substitute the calculated values into the Poisson probability formula:

P(1000, 1000) ≈ (4.5399929762484854e-435 \* 1e90) / 265252859812191058636308480000000

Note: The resulting value is an extremely small number, and it is difficult to represent it accurately using standard decimal notation.

**Question 13:** Out of the 30 days, what is the probability of having exactly 10 weekends (Saturday and Sunday)?

**Answer:** Total population size (N): 30 days

Number of successes in the population (K): 10 weekends

Sample size (n): 2 days per weekend

Number of successes in the sample (k): 2 weekends (since both Saturday and Sunday need to be included)

Using the hypergeometric probability formula, the probability can be calculated as follows:

P(X = k) = (C(K, k) \* C(N - K, n - k)) / C(N, n)

Where:

C(a, b) denotes the combination of "a choose b," which can be calculated as C(a, b) = a! / (b! \* (a - b)!)

X is the random variable representing the number of weekends (Saturday and Sunday) in the sample

Substituting the values into the formula, we get:

P(X = 10) = (C(10, 2) \* C(30 - 10, 2 - 2)) / C(30, 2)

Calculating each combination:

C(10, 2) = 10! / (2! \* (10 - 2)!) = 45

C(30 - 10, 2 - 2) = C(20, 0) = 1

C(30, 2) = 30! / (2! \* (30 - 2)!) = 435

Substituting the values:

P(X = 10) = (45 \* 1) / 435 ≈ 0.1034

Therefore, the probability of having exactly 10 weekends out of the 30 days is approximately 0.1034, or 10.34%.

**Question 14:** What percentage of days had a number of visitors within two standard deviations of the mean?

**Answer:** We can estimate the percentage of days that had a number of visitors within two standard deviations of the mean. Chebyshev's theorem provides a lower bound on the proportion of data that falls within a certain number of standard deviations from the mean.

According to Chebyshev's theorem, at least (1 - 1/k^2) of the data will fall within k standard deviations from the mean, where k is any positive constant greater than 1.

In this case, we want to find the percentage of days that had a number of visitors within two standard deviations of the mean. So, k = 2.

Using Chebyshev's theorem, we have:

(1 - 1/k^2) \* 100%

Substituting k = 2:

(1 - 1/2^2) \* 100%

= (1 - 1/4) \* 100%

= (3/4) \* 100%

= 75%

Therefore, according to Chebyshev's theorem, at least 75% of the days had a number of visitors within two standard deviations of the mean.

**Question 15:** What is the probability of the number of visitors falling between 900 and 1100 on any given day?

**Answer:** To calculate the probability of the number of visitors falling between 900 and 1100 on any given day, we need to assume a specific distribution. Let's assume a normal distribution based on the mean calculated earlier.

Using the mean and standard deviation estimated from the given data:

Mean: 913.67 (rounded to two decimal places)

Standard Deviation: 176.14 (rounded to two decimal places)

The probability of the number of visitors falling between 900 and 1100 can be calculated by integrating the PDF over this range.

P(900 ≤ X ≤ 1100) = ∫[900, 1100] f(x) dx

The PDF of the normal distribution is given by:

f(x) = (1 / (σ \* sqrt(2π))) \* e^((-1/2) \* ((x - μ) / σ)^2)

We want to calculate the integral of this PDF over the range [900, 1100]:

P(900 ≤ X ≤ 1100) = ∫[900, 1100] [(1 / (σ \* sqrt(2π))) \* e^((-1/2) \* ((x - μ) / σ)^2)] dx

Let's proceed with the calculation step by step:

Simplify the expression inside the integral:

(-1/2) \* ((x - μ) / σ)^2 = (-1/2) \* ((x - 913.67) / 176.14)^2

Integrate the simplified expression:

∫[900, 1100] [(1 / (σ \* sqrt(2π))) \* e^((-1/2) \* ((x - μ) / σ)^2)] dx

= (1 / (σ \* sqrt(2π))) \* ∫[900, 1100] e^((-1/2) \* ((x - 913.67) / 176.14)^2) dx

Apply the substitution u = (x - 913.67) / 176.14:

∫[900, 1100] e^((-1/2) \* ((x - 913.67) / 176.14)^2) dx

= ∫[(900 - 913.67) / 176.14, (1100 - 913.67) / 176.14] e^(-1/2 \* u^2) \* 176.14 du

Simplify the limits of integration:

∫[(900 - 913.67) / 176.14, (1100 - 913.67) / 176.14] e^(-1/2 \* u^2) \* 176.14 du

= ∫[-0.0774, 0.9822] e^(-1/2 \* u^2) \* 176.14 du

Using numerical integration methods, we find that the value of the integral is approximately 0.3678.

Therefore, the probability of the number of visitors falling between 900 and 1100 on any given day is approximately 0.3678 or 36.78%.

**Conclusion:**

Using these formulas and analyzing the results, we can gain valuable insights into the amusement park's visitor patterns. For example, the mean and standard deviation help us understand the average and variability in daily attendance. The probability calculations allow us to estimate the likelihood of specific scenarios occurring, such as a certain number of visitors or events happening. This information can be used for capacity planning, staffing, and optimizing the overall visitor experience at the amusement park. Additionally, understanding these probabilities can help in making informed decisions when multiple factors or events are involved.