Introduction

Cold Storage Company deals in business of storing Pasteurized Fresh Whole or Skimmed Milk, Sweet Cream, Flavoured Milk Drinks since 2016. To maintain perfect texture, body appearance, separation of fats of these products the company maintains optimal temperature between 2º - 4º C. In this problem statement, we will explore how much a penalty will be imposed on company if the Annual maintenance case i.e the required temperature is not maintained.

This problem set will help us to analyse the dataset. Examine the change in temperature with change in season. Explore the dataset using central tendency and other parameters.

This data is maintained for recording all the changes taking place in the whole year (365days) for all the 3 seasons – summer, Rainy, winter.

Data set:

There are total 365 rows and 4 columns in the dataset.

Out of 4 columns 2 columns are of object type (Season & Month) 1 is integer (Date) and 1 is float (Temperature) data type.

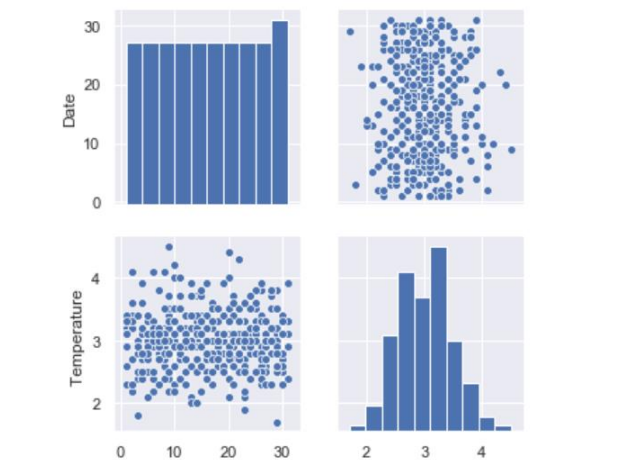|	Season	Month	Date	Temperature
0	Winter	Jan	1	2.3
1	Winter	Jan	2	2.2
2	Winter	Jan	3	2.4
3	Winter	Jan	4	2.8
4	Winter	Jan	5	2.5

There is no null data in data set

```
df.isnull().sum()
```

```
Season         0
Month          0
Date           0
Temperature    0
dtype: int64
```

Pair Plot –

Pair Plot shows whether relationship between the variables can be continuous or categorical.

Data description –

Short summary of the full data

Highest temperature recorded was 4.5 degrees C.

	Season	Month	Date	Temperature
count	365	365	365.000000	365.000000
unique	3	12	NaN	NaN
top	Winter	May	NaN	NaN
freq	123	31	NaN	NaN
mean	NaN	NaN	15.720548	3.002466
std	NaN	NaN	8.808321	0.465832
min	NaN	NaN	1.000000	1.700000
25%	NaN	NaN	8.000000	2.700000
50%	NaN	NaN	16.000000	3.000000
75%	NaN	NaN	23.000000	3.300000
max	NaN	NaN	31.000000	4.500000

Q1 . Find mean cold storage temperature for Summer, Winter, and Rainy Season.

➔ The **mean** (average) of a **data** set **is** found by adding all numbers in the **data** set and then dividing by the number of values in the set. The **mean** is essentially a model of your **data** set.
  Therefore, in our data set the average temperature maintained by company in each season is as follows.

- Mean cold storage temperature for Winter season is 2.78
- Mean cold storage temperature for Rainy season is 3.09
- Mean cold storage temperature for Summer season is 3.15

Q2. Find the overall mean for the full year.

  The average temperature maintained by company throughout the year is 3.00
  (overall Mean cold storage temperature for full year is 3.00)

Q3. Find Standard Deviation for the full year.

➔ **Standard deviation tells you** how spread out the data is. It is a measure of how far each observed value is from the mean.

- overall Standard Deviation for full year is 0.47

➔ A **probability distribution** is a statistical function that describes all the possible values and likelihoods that a random variable can take within a given range.

Q4. Assume Normal distribution, what is the probability of temperature having fallen below 2º C?

- Under normal distribution the probability that the temperature has fallen below 2º C is 1.5%.
  This mainly happens in Rainy and Winter season as the overall climate of environment is already cold as compared to summer season.

Q5. Assume Normal distribution, what is the probability of temperature having gone above 4º C?

- Under normal distribution the probability that the temperature has fallen below 2º C is 1.6%.
  This mainly happens in Rainy season as the climate keeps changing depending on rain.

Q6. What will be the penalty for the AMC Company?

- As per the company policy if the The probability of temperature going outside the 2º - 4º C during the one-year contract was above 2.5% and less than 5% then the penalty would be 10% of AMC (annual maintenance case) .

- In this year the probability that the temperature went outside 2º - 4º C is 3% , hence the penalty for the AMC of Company will be 10%.