Cold Storage - Case study

Submitted by,
Ajuna John

The Problem

Problem 1

Cold Storage started its operations in Jan 2016. They are in the business of storing Pasteurized Fresh Whole or Skimmed Milk, Sweet Cream, Flavoured Milk Drinks. To ensure that there is no change of texture, body appearance, separation of fats the optimal temperature to be maintained is between 2º - 4º C.
In the first year of business they outsourced the plant maintenance work to a professional company with stiff penalty clauses. It was agreed that if it were statistically proven that 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 case it exceeded 5% then the penalty would be 25% of the AMC fee. The average temperature data at date level is given in the file "Cold_Storage_Temp_Data_.csv"

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

2. Find overall mean for the full year.

3. Find Standard Deviation for the full year.

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

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

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

Dataset to be used for Problem 1: Cold_Storage_Temp_Data_.csv

Problem 2

In Mar 2018, Cold Storage started getting complaints from their Clients that they have been getting complaints from end consumers of the dairy products going sour and often smelling. On getting these complaints, the supervisor pulls out data of last 35 days' temperatures. As a safety measure, the Supervisor has been vigilant to maintain the temperature equal to 3.9º C.
Assume 3.9º C as the acceptable temperature and at alpha = 0.1 do you feel that there is need for some corrective action in the Cold Storage Plant or is it that the problem is from procurement side from where Cold Storage is getting the Dairy Products. The data of the last 35 days is in "Cold_Storage_Mar2018_.csv"

1. Which Hypothesis test shall be performed to check the if corrective action is needed at the cold storage plant? Justify your answer.

2. State the Hypothesis and do the necessary calculations to accept or reject the corresponding null hypothesis.

3. Give your inference.

Dataset to be used for Problem 2: Cold_Storage_Mar2018_.csv

Solution – Problem 1

Analysis of DataSet

The dataset has 365 observations with 4 variables i.e "Season", "Month", "Date" and "Temperature".

Structure

Season: Factor w/ 3 levels "Rainy","Summer",..: 3 3 3 3 3 3 3 3 3 3 ...

Month: Factor w/ 12 levels "Apr","Aug","Dec",..: 5 5 5 5 5 5 5 5 5 5 ...

Date: int 1 2 3 4 5 6 7 8 9 10 ...

**Temperature**: num 2.3 2.2 2.4 2.8 2.5 2.4 2.8 3 2.4 2.9 ...

Dataset has records for 12 months and is grouped into 3 seasons as Rainy, Summer and Winter.

Summary

| Season	Month	Date	Temperature
Rainy :122	Aug : 31	Min. : 1.00	Min. :1.700
Summer:120	Dec : 31	1st Qu.: 8.00	1st Qu.:2.700
Winter:123	Jan : 31	Median :16.00	Median :3.000
	Jul : 31	Mean :15.72	Mean :3.002
	Mar : 31	3rd Qu.:23.00	3rd Qu.:3.300
	May : 31	Max. :31.00	Max. :4.500
	(Other):179		

Summary has the **minimum** temperature as **1.7** Degree Celsius, **maximum** temperature as **4.5** Degree Celsius and **mean** temperature as **3.002**

Boxplot and histogram

Cold Storage Temperature Histogram
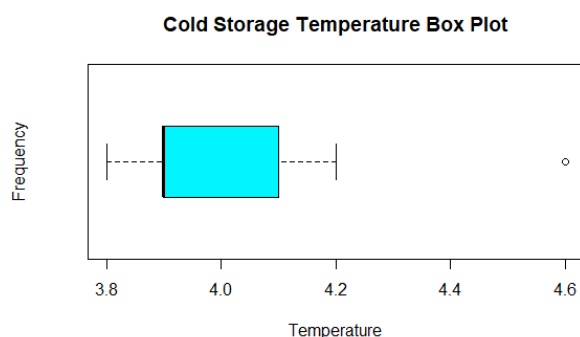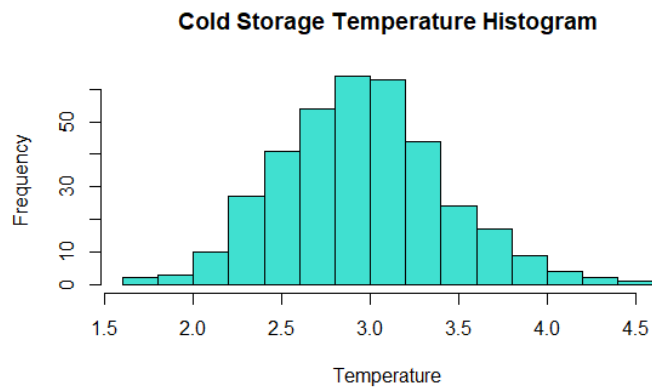
There are possible outliers/extreme values in the temperature
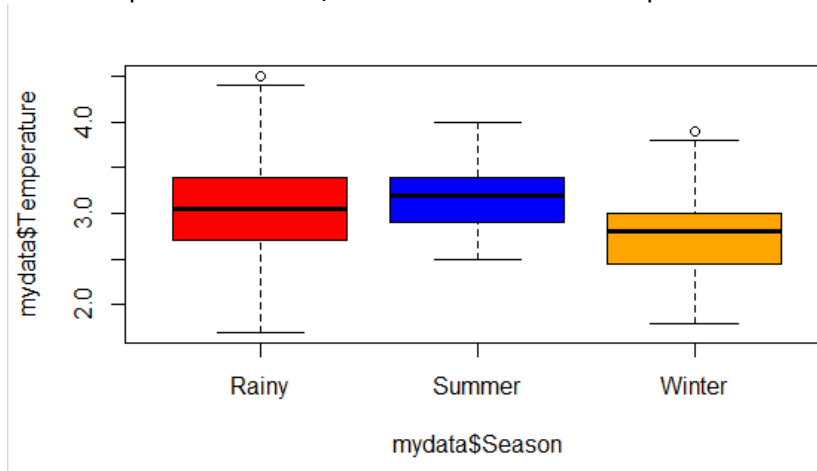
1. Find mean cold storage temperature for Summer, Winter and Rainy Season

mydata$Season: Rainy
[1] **3.087705**
----
mydata$Season: Summer
[1] **3.1475**
----
mydata$Season: Winter
[1] **2.776423**

2. Find overall mean for the full year
3.002466

3. Find Standard Deviation for the full year
0.4658319

4. Assume Normal distribution, what is the probability of temperature having fallen below 2º C?
0.01569906

5. Assume Normal distribution, what is the probability of temperature having gone above 4º C?
0.01612075

6. What will be the penalty for the AMC Company?
The penalty condition in terms of probability of temperature going outside the 2 - 4 ºC:
(i) Above 2.5 % and less than 5 % then the penalty would be 10% of AMC (annual maintenance case).
(ii) In case it exceeded 5 % then the penalty would be 25% of the AMC fee.

the total probability for condition less than 2 Degree Celsius or greater than4 Degree Celsius is calculated as **0.03181981** ºC

Now use this total probability is used to check against the two conditions mentioned above to get the penalty as **10%**

Solution – Problem 2

Analysis of DataSet

The dataset has 35 observations with 4 variables i.e "Season", "Month", "Date" and "Temperature".

Structure

```
'data.frame':    35 obs. of  4 variables:
 $ Season    : Factor w/ 1 level "Summer": 1 1 1 1 1 1 1 1 1 1 ...
 $ Month     : Factor w/ 2 levels "Feb","Mar": 1 1 1 1 1 1 1 1 1 1 ...
 $ Date      : int  11 12 13 14 15 16 17 18 19 20 ...
 $ Temperature: num  4 3.9 3.9 4 3.8 4 4.1 4 3.8 3.9 ...
```

Summary

Season	Month	Date	Temperature
Summer:35	Feb:18	Min.: 1.0	Min. :3.800
	Mar:17	1st Qu.: 9.5	1st Qu.:3.900
		Median:14.0	Median :3.900
		Mean :14.4	Mean :3.974
		3rd Qu.:19.5	3rd Qu.:4.100
		Max. :28.0	Max. :4.600

Sample has 35 records, with minimum temperature as 3.8 ºC, mean temperature as 3.974ºC and Maximum temperature as 4.600 ºC.

Histogram and Boxplot

Cold Storage Temperature Histogram

Cold Storage Temperature Box Plot
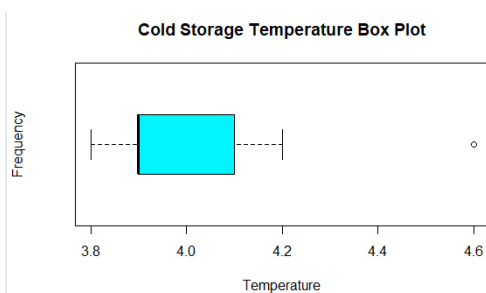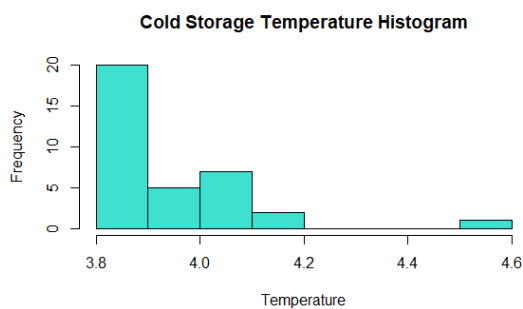
There are possible outliers/extreme values in the sample

1. Which Hypothesis test shall be performed to check the if corrective action is needed at the cold storage plant? Justify your answer.

One sample t test can be performed. As here, one sample is being compared with a standard. Also, n>30 and population std deviation is not known.

2. State the Hypothesis and do the necessary calculations to accept or reject the corresponding null hypothesis

Null hypothesis H0: Temperature <= 3.9 degree Celsius
Alternate Hypothesis Ha: Temperature > 3.9 degree Celsius
As alternate hypothesis is with a greater than sign, it is a Right tailed test.

One - sample t -test

data: coldData$Temperature
t = 2.7524, df = 34, p-value = 0.004711
alternative hypothesis: true mean is greater than 3.9
90 percent confidence interval:
 3.939011   Inf
sample estimates:
mean of x
 3.974286

p value obtained is **0.004711** which is much less than alpha=0.1
#Hence **null hypothesis is rejected**. i.e The temperature is above 3.9 Degree Celsius
#So, it is recommended that corrective action is required.

3. Give your inference.

One sample t test is performed. This is because, here one sample is being compared with a standard. Also, n>30 and population std deviation is not known. From the hypothesis test it is recommended that a corrective action must be done.