Film Data Analysis

Case – How to improve the quality of stretch films?

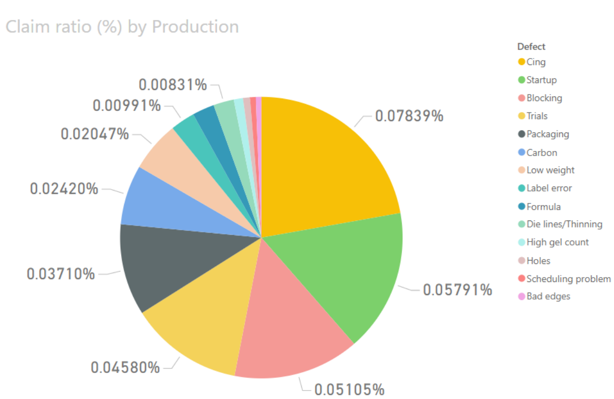
As a quality assurance engineer, understanding the quality status of the products our plant produces is important. Our company wants to know what factors will affect quality during the process. When the quality was improved according to the analysis below, production shipping quality can potentially increase by.

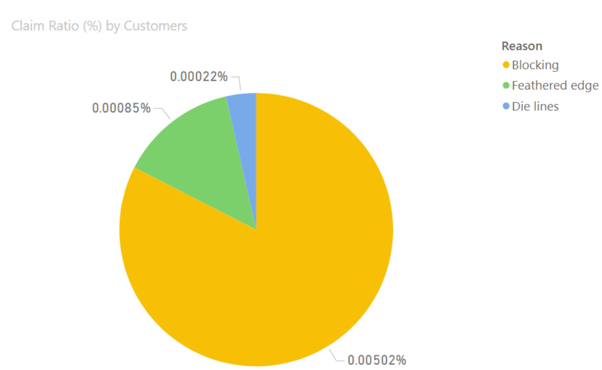
What kinds of products do I analyze for stretch film?

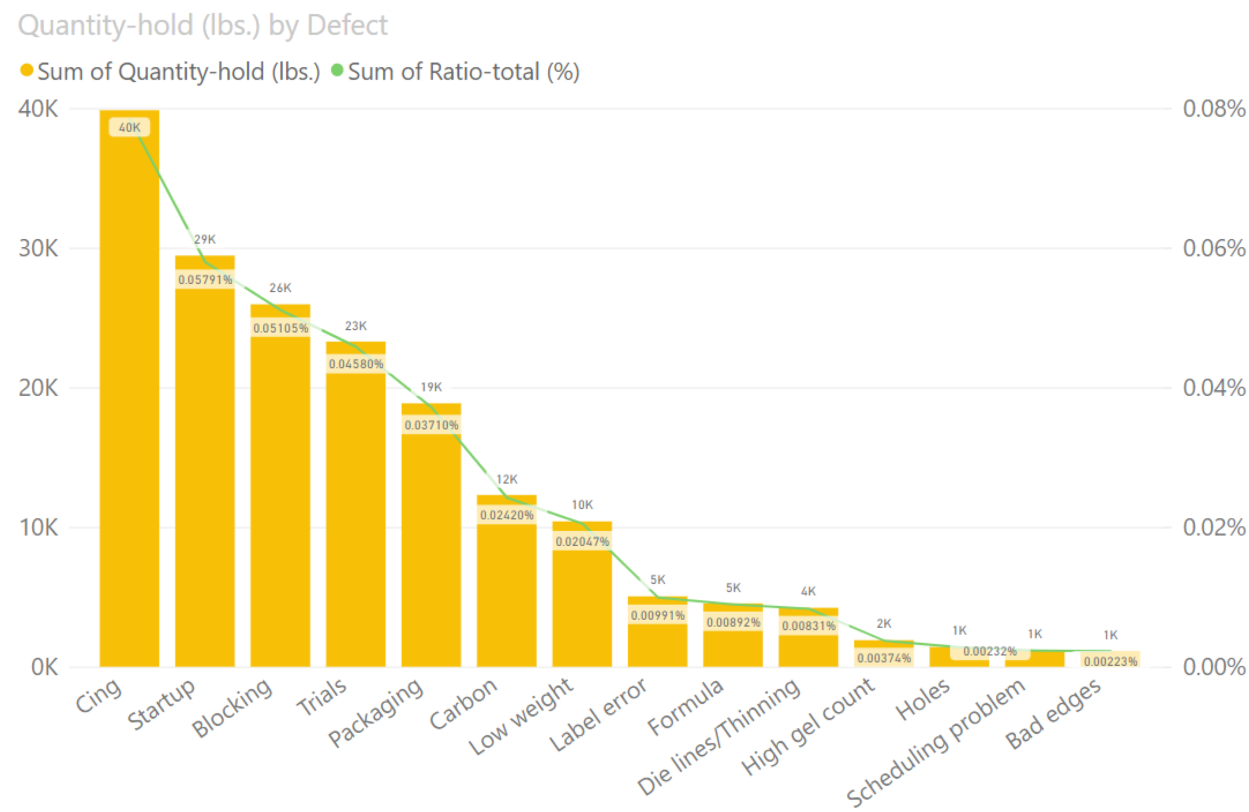
Machine film is the main product in our plant. It is one kind of stretch film which is used for high-speed equipment and applications that require films to pre-stretch over 300%. The specs of machine films are very different from hand films and they will be more serious.

What quality problems do CFP have and where it got reports?

In 2020, our company got defect reports both from production and customers. One is production that finds some defects or is concerned with products before sending to customers, the other one is complaints by customers after using products. Hand Cling (40K, 0.0783%) was the main defect from production, and blocking (0.00502%) was the main complaint from customers.

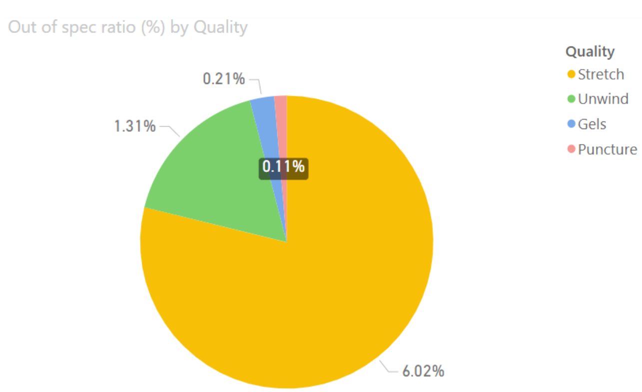




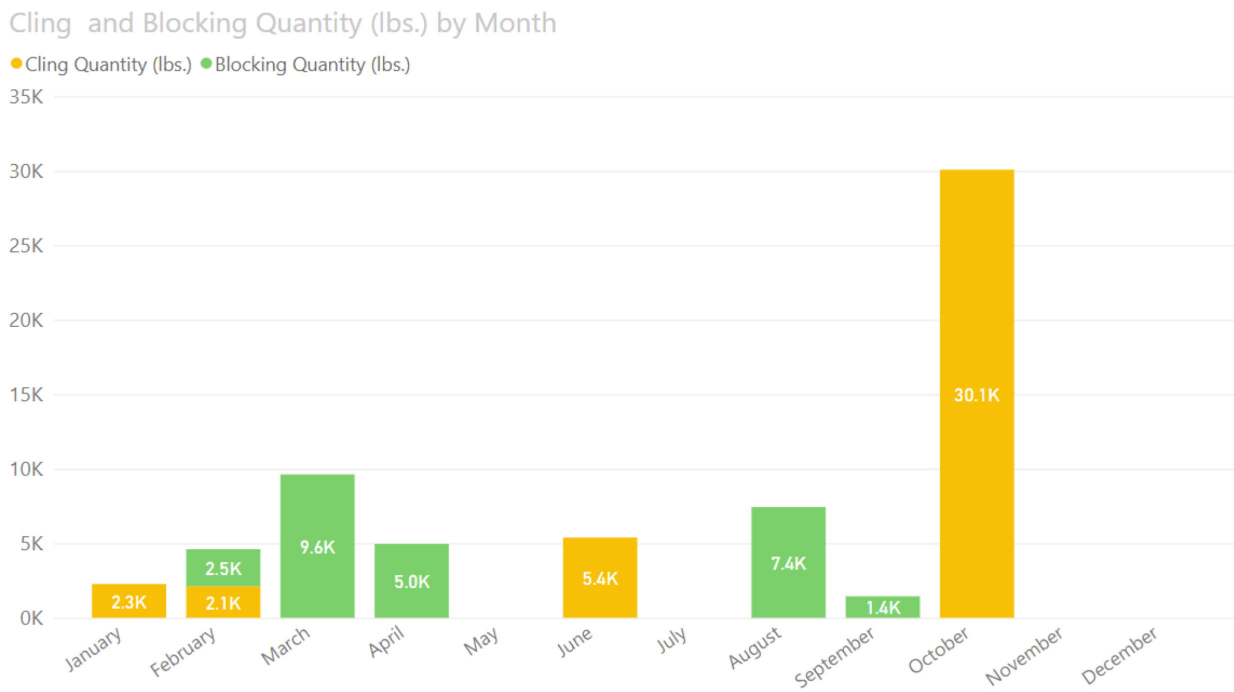


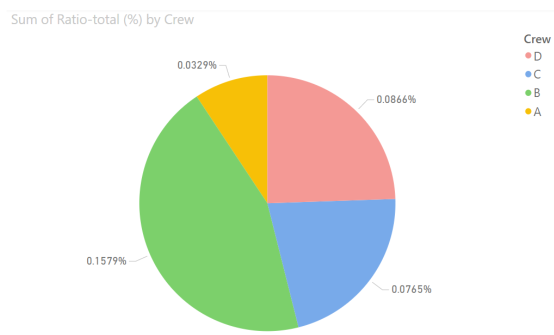
What are the potential reasons for quality defects?

Except for material problems, we found out line speed, output, vacuum, temperature, and seasons can be the possible process reasons to influence quality. Also, we noticed that packaging and label error are not low (24K lbs), we can assume that human error is one of the impacts.



After analyzing the numbers of out of spec on products from quality testing data, it did show stretch values have the highest ratio of out of spec. Even though not every out of spec will cause quality, it did match with cling, blocking problems from production and potential breaking problems from customers.





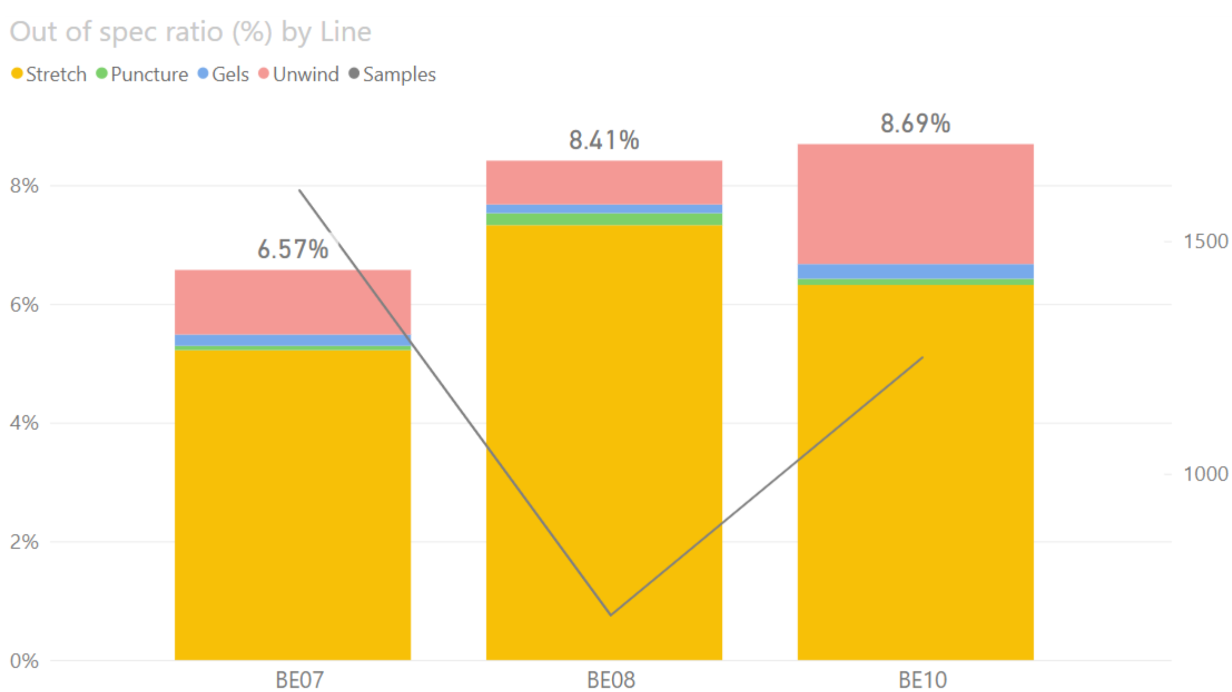
1. Cling and blocking are related to seasons and unwind values that films are stickier during summer (Aug, 7.4K lbs.) and the time temperature increases (March, 9.4K lbs.), so we can see cling problems increase when the temperature is going to get cooler in October.
   Solution: Adjusted the ratio of cling resin when changing seasons.
2. Stretch has the highest out of spec ratio (6.02%) in quality values which has probability to break when customers use it.
   Solution: one of the main factors is vacuum, we can use that to improve stretch values.
3. Human error is also from both production and customer complaints.
   Solution: Shift leads should make sure operators and packers are followed by sops , especially B crew (0.1579%).

By applying the above solutions, production shipment can increase 0.18% (90K lbs.) and save potential breaking problems by 6.02%.

Which production line is more effective?

There are three different machine lines that produce this type of film. All three lines have small differences with each other which may affect the film quality. Line 7 and line 8 are very similar, both of them have 4 extruders and installed at the same time. Line 10 is the newest one that only has 5 extruders and unmoving chill roll.

Line 7 here shows that it tested the largest numbers of samples, but has the lowest out of spec ratio (6.57%).



Film Data Analytics

Case – Why stretch values?

The ability to predict stretch values of stretch film is valuable for production with processing procedures. As our products are stretch film, stretch data are very important for improving the efficiency of production. In order to build a model to predict stretch values, processing factors that might affect stretch should be known.

What factors affect values of stretch?

At first, the processing factors which have possibilities to influence stretch values based on production experience are chosen to do the predictions.

Potential factors:

- Line speed
- Output
- Vacuum 1
- Vacuum 2
- Temperature 1
- Temperature 2

After analyzing correlations of these factors, we found out line speed and output have strong relationship and values in V2, T1, and T2 didn't change much. Therefore, we use output and V1 to be our predict factors first.

```
Call:
lm(formula = stretch ~ output + V1, data = summer)

Residuals:
    Min      1Q  Median      3Q     Max 
-29.764 -14.034  -7.381   6.768 168.499 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)   
(Intercept) 640.89750  225.10206   2.847  0.00724 **
output       -0.07643    0.06597  -1.159  0.25428   
V1           -0.36157    0.96279  -0.376  0.70946   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 31.56 on 36 degrees of freedom
Multiple R-squared:  0.0459,	Adjusted R-squared:  -0.007107 
F-statistic: 0.8659 on 2 and 36 DF,  p-value: 0.4292
```

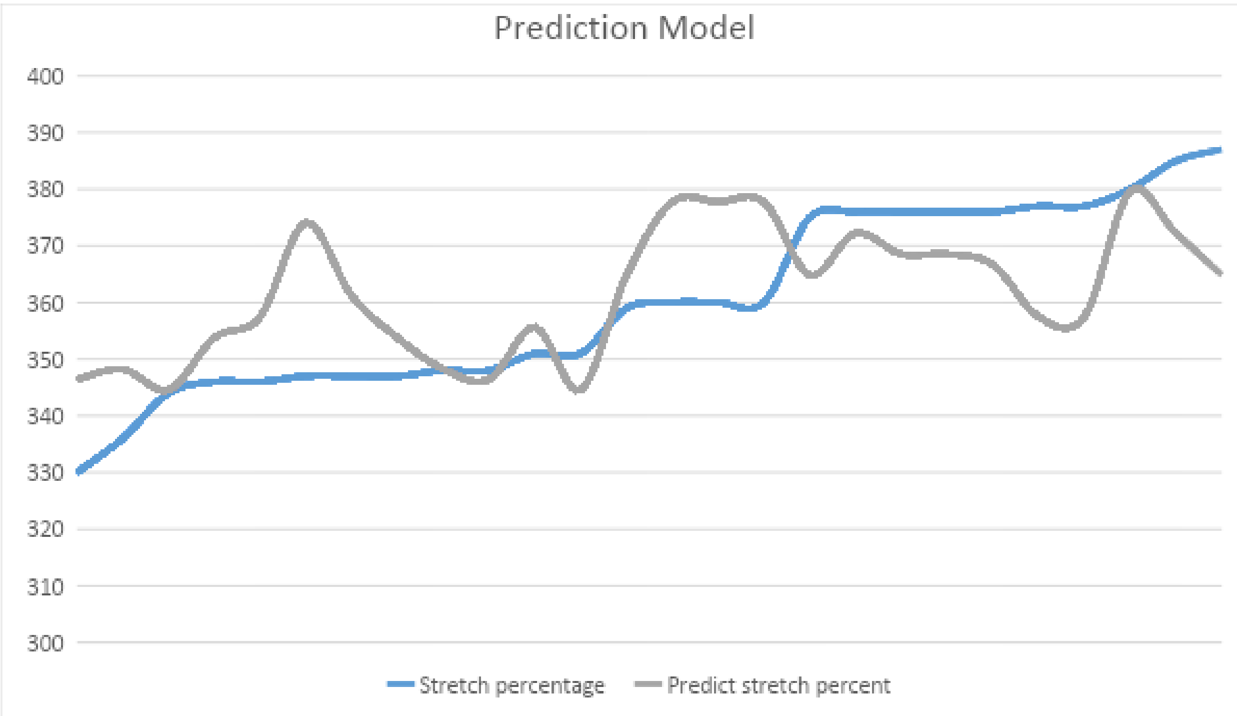
Since p-value in this table is high that we cannot use this model to do prediction, we decided to remove output values and got the table below.

```
Call:
lm(formula = stretch ~ V1, data = summer)

Residuals:
    Min      1Q  Median      3Q     Max 
-19.645  -8.501   1.199   7.543  25.265 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 446.3289    24.4506  18.254  < 2e-16 ***
V1           -1.8443     0.4769  -3.867  0.00049 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.2 on 33 degrees of freedom
Multiple R-squared:  0.3119,	Adjusted R-squared:  0.291 
F-statistic: 14.96 on 1 and 33 DF,  p-value: 0.0004897
```

This summary shows low enough p value and the estimate of intercept is more reasonable. Than we use this model to predict the value of stretch percent base on the primary vacuum data.

How can we predict stretch values?

The model here uses vacuum 1 as selected factor. After developing the model, only two values show +/- 20 in residuals which means this model is good to use for prediction.



In [1]:
```
import pandas as pd
import numpy as np
import statsmodels.api as sm
from sklearn import datasets
from sklearn import linear_model
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn import preprocessing
```

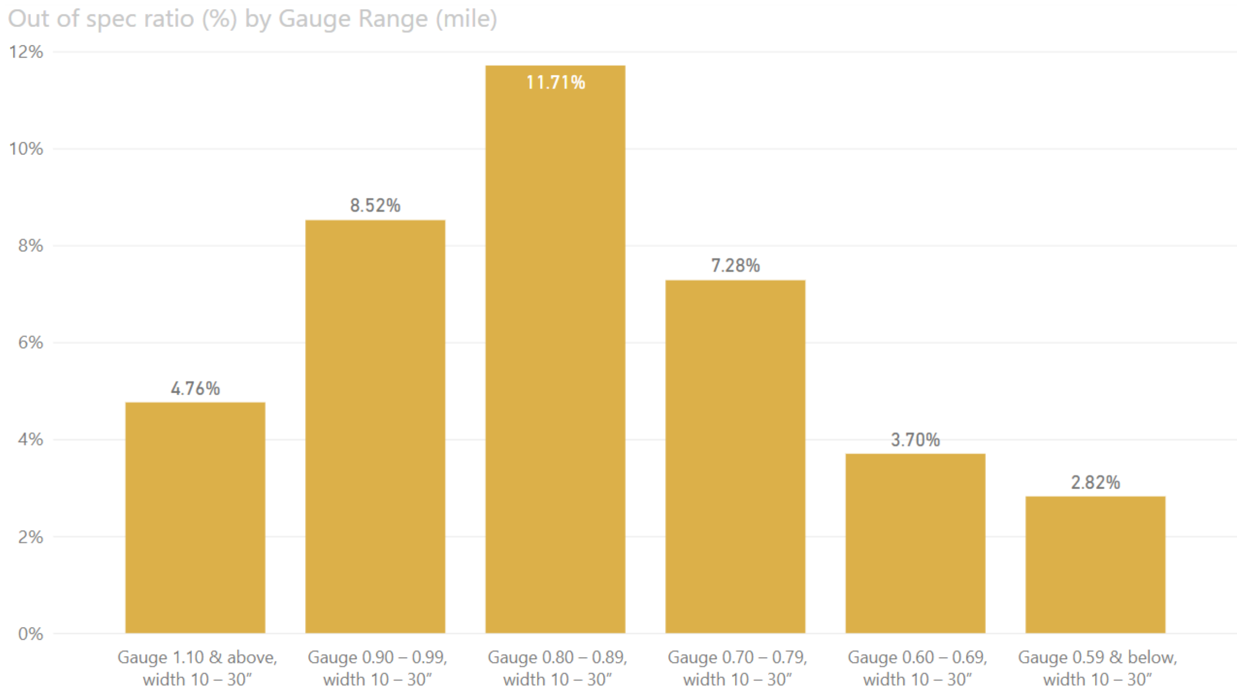
In [2]:
```
df = pd.read_csv("HLB(855).csv")
```

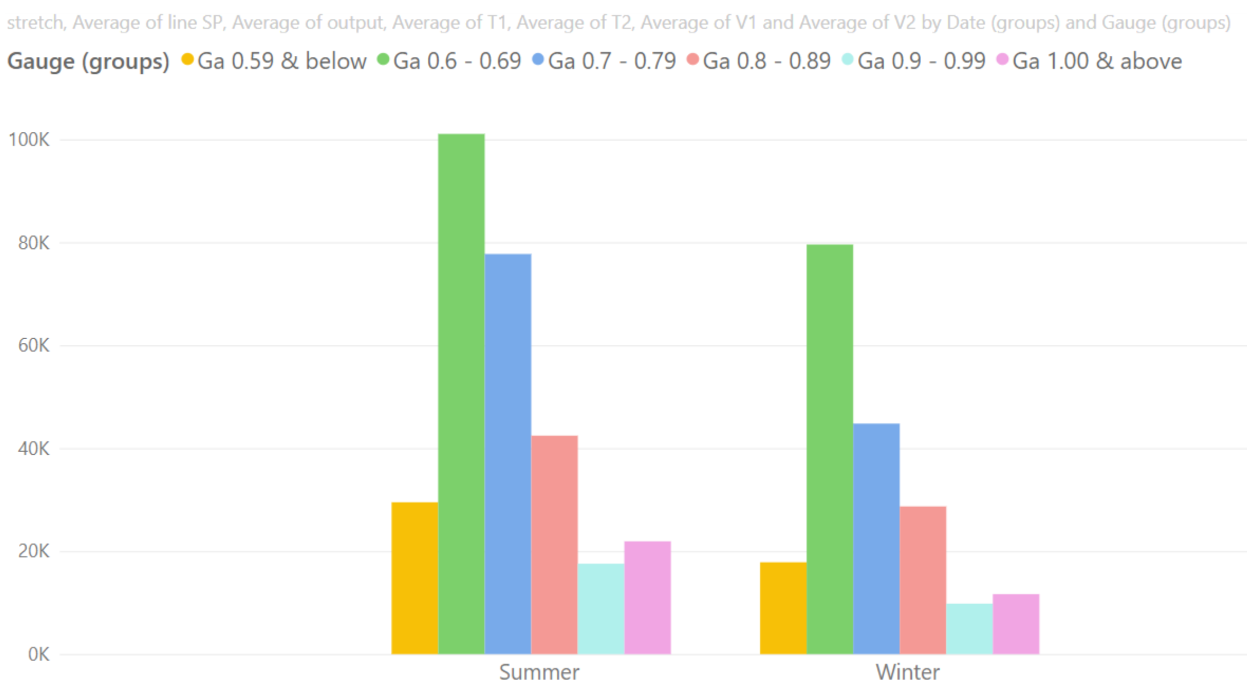
In [3]:
```
for i in range(0, len(df)):
    if df['Film'][i][0:3] == "HLP":
        df['Film'][i] = round(float(df['Film'][i][6:9]) / 25.4 , 1)
    elif df['Film'][i][0:3] == "HLB":
        df['Film'][i] = float(df['Film'][i][6:9]) / 10
df
```

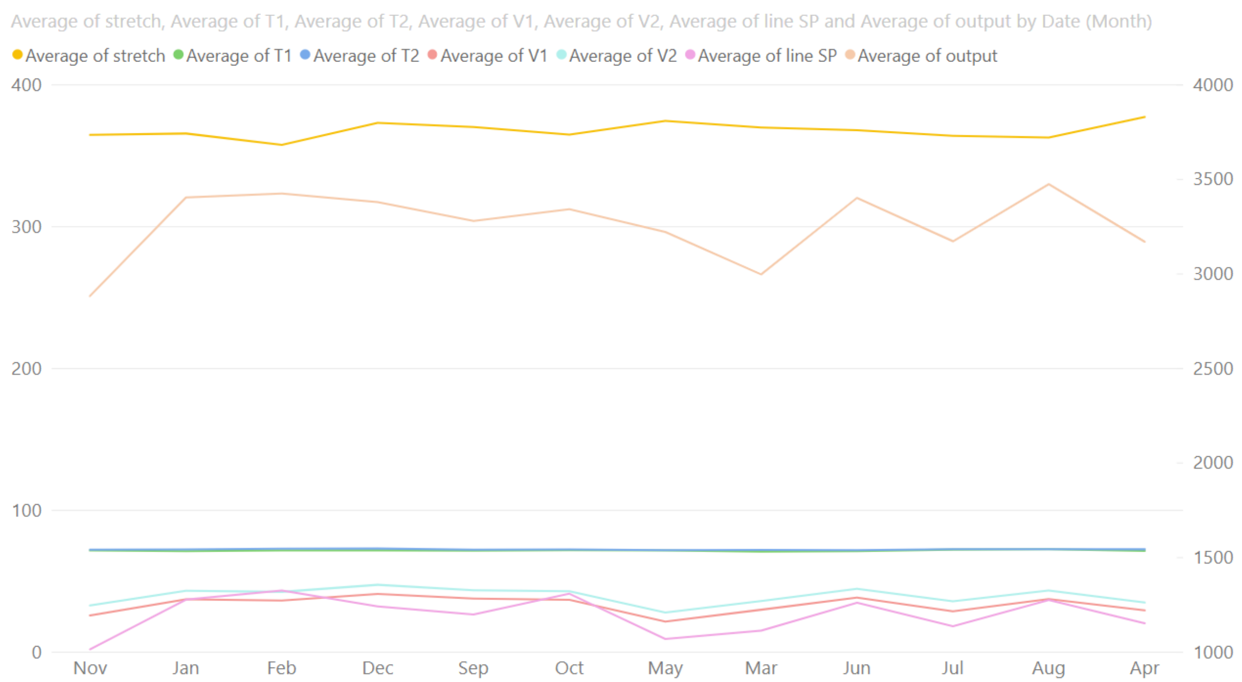
In [19]:
```
df2 = df[['Date', 'Film', 'stretch', "UW", "SF", "Puncture", "Gel count", "MAC
for col in ['stretch', 'UW', 'SF', 'Puncture', 'Gel count']:
    df2 = df2[df2[col] != 0]
df2 = df2.reset_index(drop=True)
df2
```

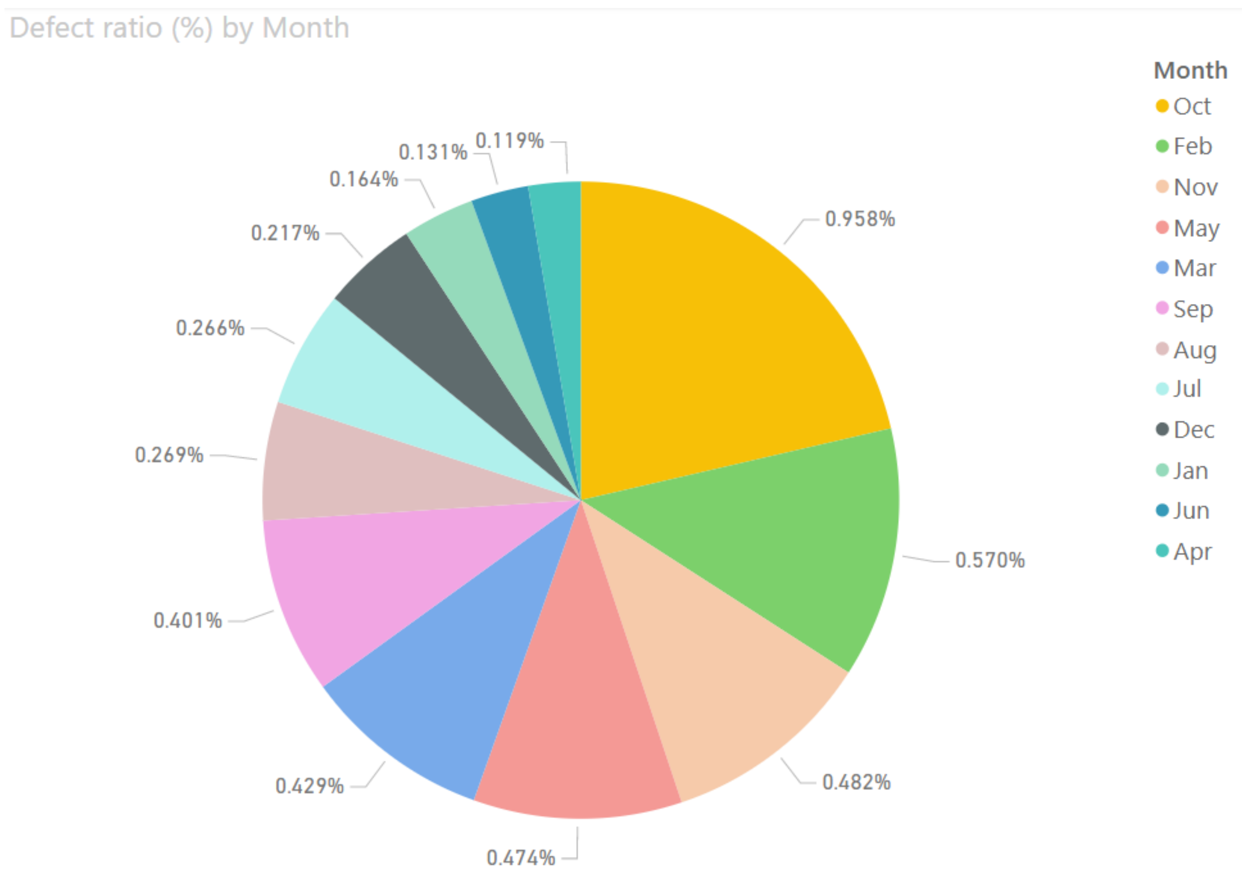
In [27]:
```
df2_fourthW.describe(percentiles = [.25, .5, .75, .95])
```

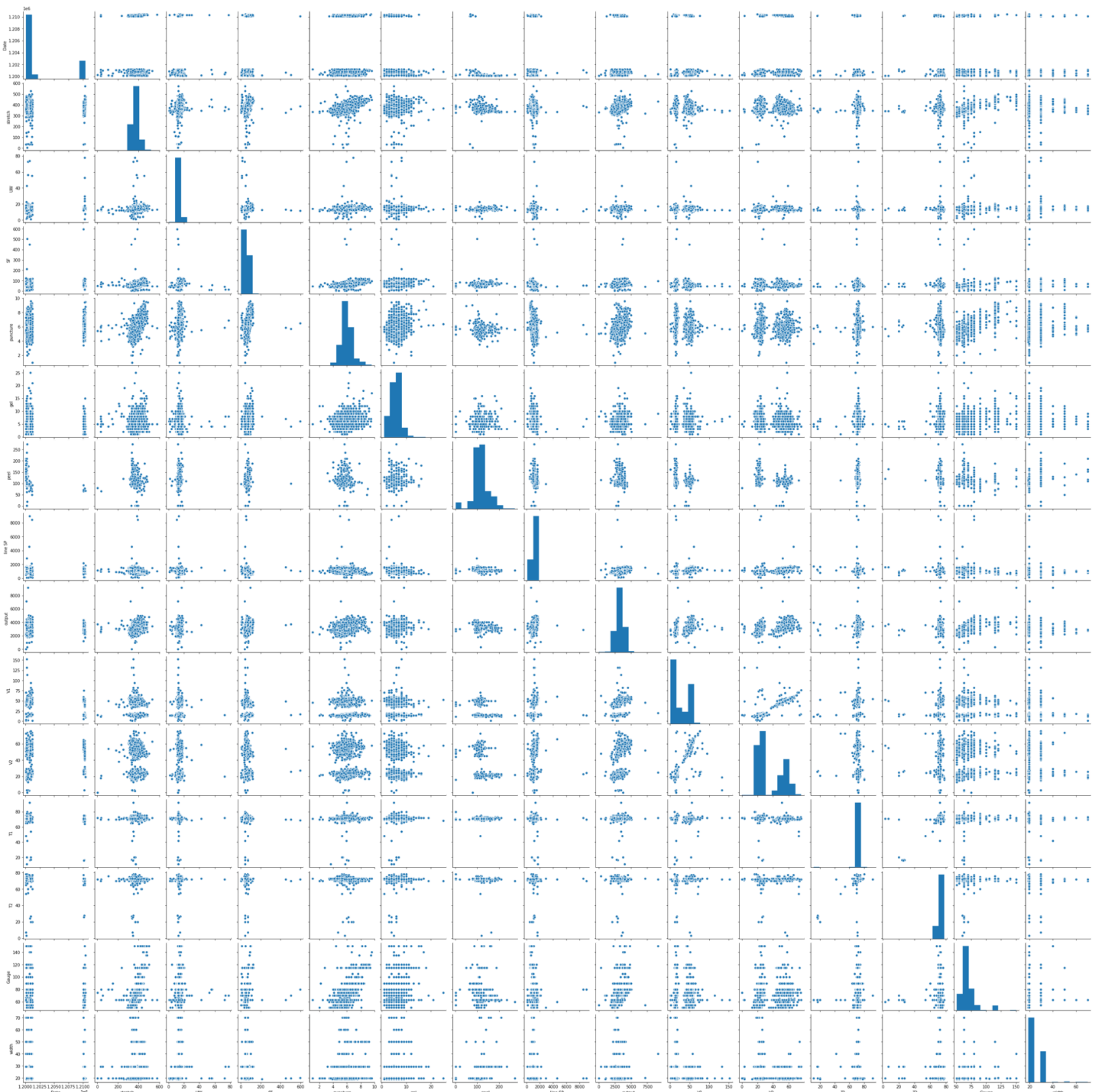
In [26]:
```
first = (df2["gauge"] <= 44)
second = (df2["gauge"].between(45, 49))
third = (df2["gauge"].between(50, 59))
fourth = (df2["gauge"] >= 60)
# Assign gauge classified dataframe in to 4 variables
df2_first = df2[first]
df2_second = df2[second]
df2_third = df2[third]
df2_fourth = df2[fourth]
firstN = (df2_first['Film'] <= 13)
firstM = (df2_first['Film'].between(14,16))
firstW = (df2_first['Film'] > 16)
secondN = (df2_second['Film'] <= 13)
secondM = (df2_second['Film'].between(14,16))
secondW = (df2_second['Film'] > 16)
thirdN = (df2_third['Film'] <= 13)
thirdM = (df2_third['Film'].between(14,16))
thirdW = (df2_third['Film'] > 16)
fourthN = (df2_fourth['Film'] <= 13)
fourthM = (df2_fourth['Film'].between(14,16))
fourthW = (df2_fourth['Film'] > 16)

df2_firstN = df2_first[firstN]
df2_firstM = df2_first[firstM]
df2_firstW = df2_first[firstW]
df2_secondN = df2_second[secondN]
df2_secondM = df2_second[secondM]
df2_secondW = df2_second[secondW]
df2_thirdN = df2_third[thirdN]
df2_thirdM = df2_third[thirdM]
df2_thirdW = df2_third[thirdW]
df2_fourthN = df2_fourth[fourthN]
df2_fourthM = df2_fourth[fourthM]
df2_fourthW = df2_fourth[fourthW]
```


Out of spec ratio (%) by Gauge Range (mile)
12%
10%
8%
6%
4%
2%
0%
4.76%
8.52%
11.71%
7.28%
3.70%
2.82%
Gauge 1.10 & above, width 10 – 30"
Gauge 0.90 – 0.99, width 10 – 30"
Gauge 0.80 – 0.89, width 10 – 30"
Gauge 0.70 – 0.79, width 10 – 30"
Gauge 0.60 – 0.69, width 10 – 30"
Gauge 0.59 & below, width 10 – 30"


stretch, Average of line SP, Average of output, Average of T1, Average of T2, Average of V1 and Average of V2 by Date (groups) and Gauge (groups)
Gauge (groups)
Ga 0.59 & below
Ga 0.6 - 0.69
Ga 0.7 - 0.79
Ga 0.8 - 0.89
Ga 0.9 - 0.99
Ga 1.00 & above
100K
80K
60K
40K
20K
0K
Summer
Winter


Average of stretch, Average of T1, Average of T2, Average of V1, Average of V2, Average of line SP and Average of output by Date (Month)
Average of stretch
Average of T1
Average of T2
Average of V1
Average of V2
Average of line SP
Average of output
400
300
200
100
0
4000
3500
3000
2500
2000
1500
1000
Nov
Jan
Feb
Dec
Sep
Oct
May
Mar
Jun
Jul
Aug
Apr


Defect ratio (%) by Month
Month
Oct
Feb
Nov
May
Mar
Sep
Aug
Jul
Dec
Jan
Jun
Apr
0.958%
0.570%
0.482%
0.474%
0.429%
0.401%
0.269%
0.266%
0.217%
0.164%
0.131%
0.119%



Date
stretch
UW
SF
puncture
gel
peel
line SP
output
V1
V2
T1
T2
Gauge
width