Dimensionality Reduction and Visualization
Exercise Set 5 Solutions

Md. Abdullah-Al Mamun

Part C: Graphical Excellence

Problem C1: Appealing Visualization

We have chosen an appealing scientific visualization of pandemic protections from nature article "Whose coronavirus strategy worked best? Scientists hunt most effective policies."

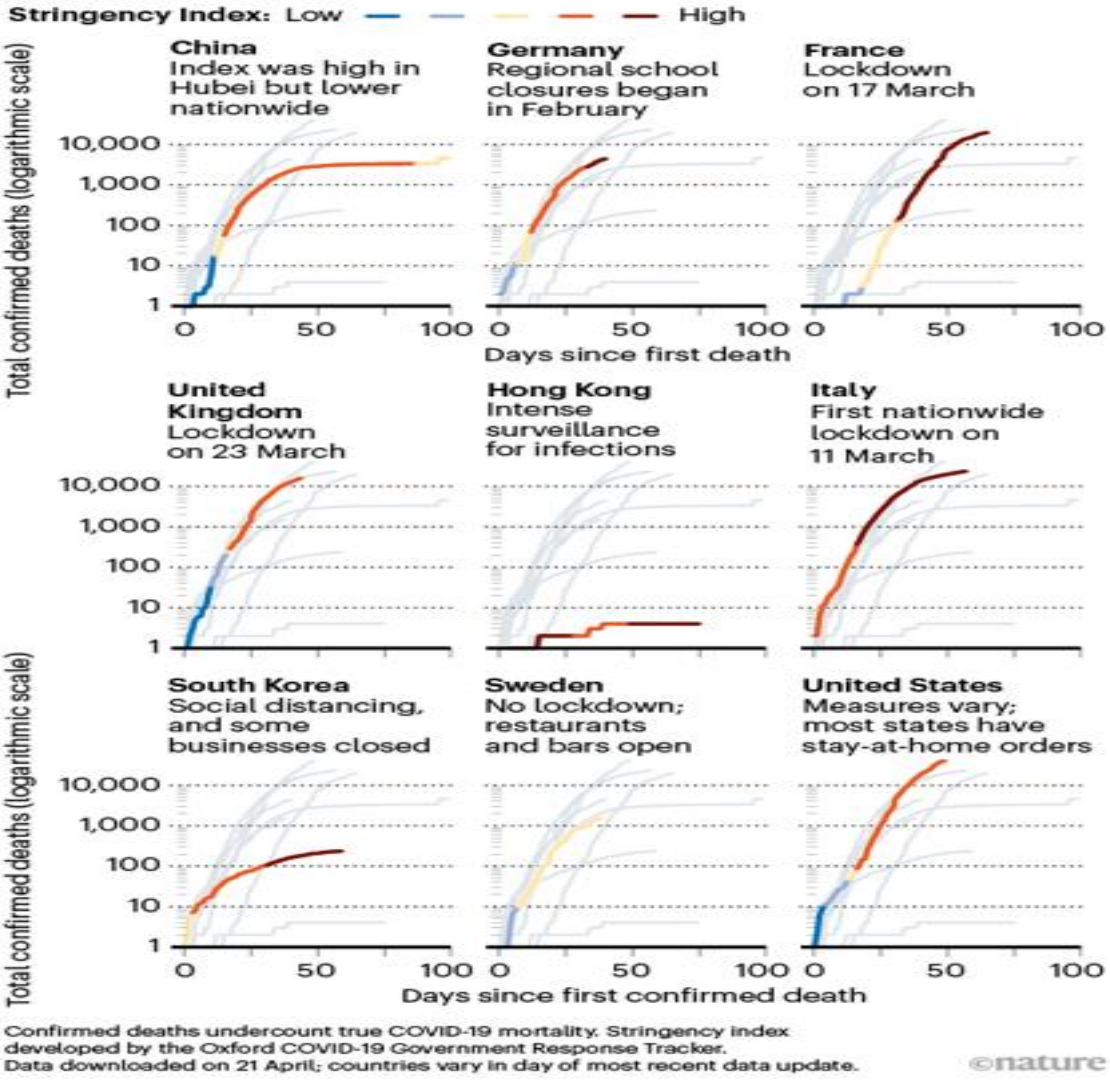


Figure 1: Confirmed death undercount true COVID-19 mortally.

In this visualization researcher provided a 'stringency index' for total severity of coronavirus outbreak response from the countries. 'stringency index' created by oxford COVID-19 Government Response tracker. oxford COVID-19 Government

Response tracker provided excellent interactive visualizations for COVID-19 are given below:

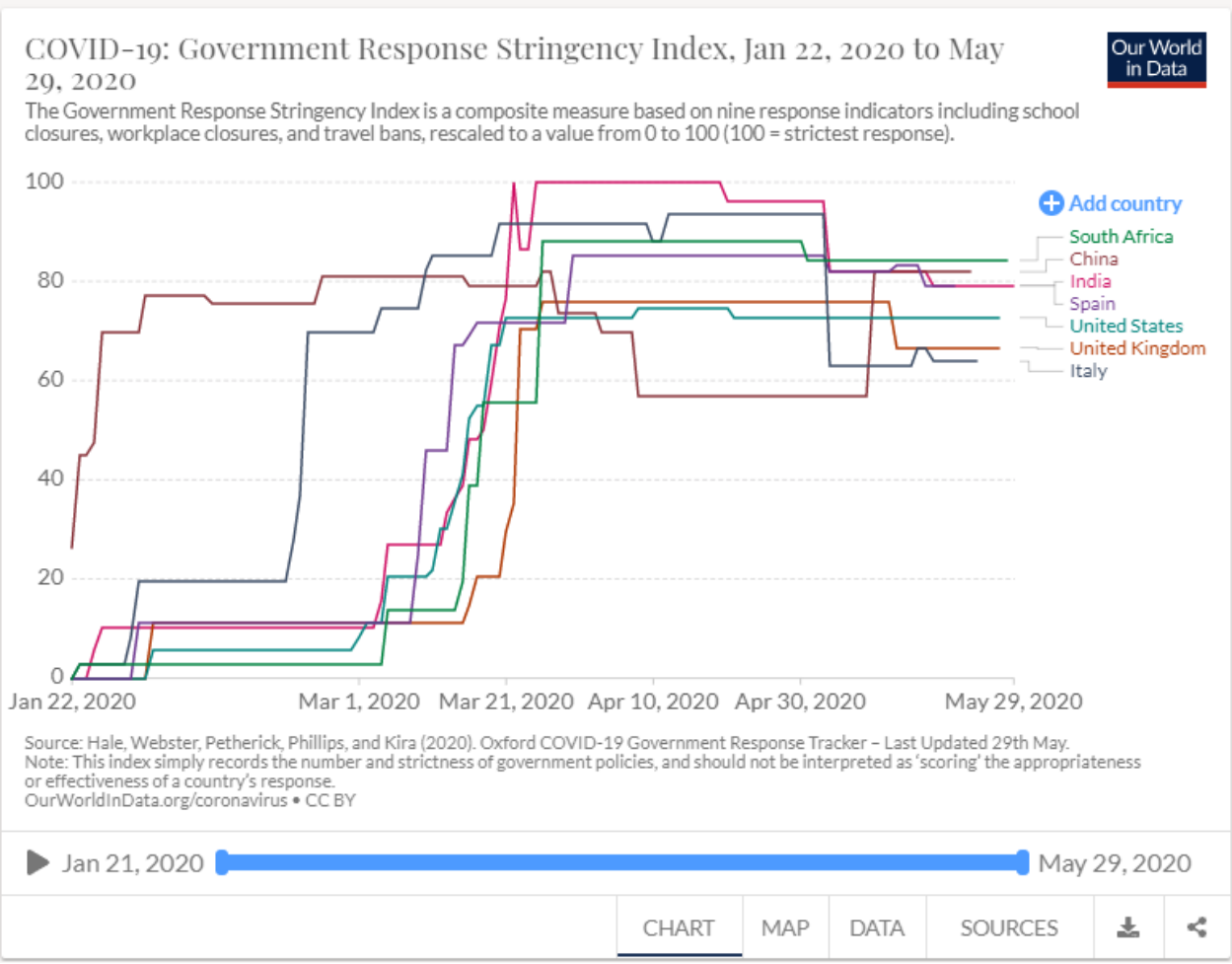


Figure 2: Interactive visualization of Chart and Map of COVID-19 Outbreak.

Problem C2: Lie Factor

By comparing the ratio of the biggest to the smallest, assume say the area of smallest object is Q, and the ratio of width: biggest/smallest=50/6, therefore biggest object's area is $((50/6)^2)Q$

Increase in Area $=\frac{((49/5)^2)Q}{Q-1} = (49/5)^2 - 1$

The numbers grow: 50/6- 1

Therefore, it gives:

LIE FACTOR = $((49/5)^2 - 1)/(50/6- 1) = 2.536$ which is greater than 1.05, hence the graphics distort.

LOG(LF) = log (2.536) = 0.404 which is greater than zero, hence it is overstating.

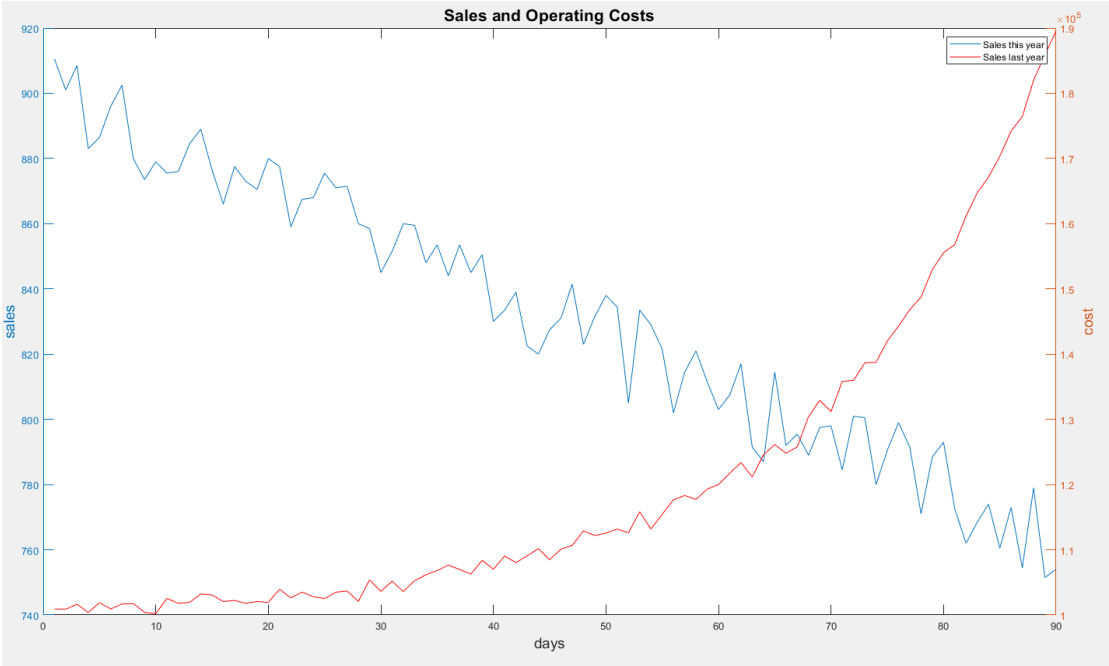
Problem C3: Chartjunk and Data-Ink Ratio

***codes are in the index section of few last pages of this report.**

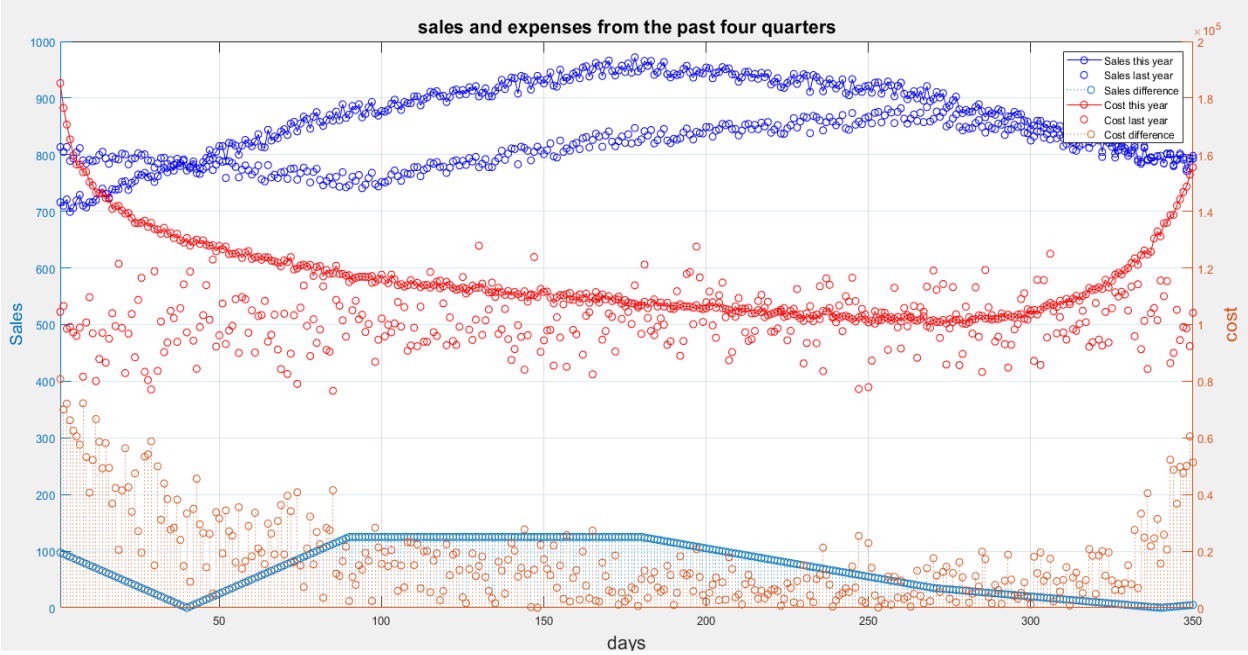
Problem C3:

a)

Firstly, we have created toy data and 90 days linear sales data added with it and performed uniformed distribution. Here we can see the case data increased exponentially with noise with the normal distribution.



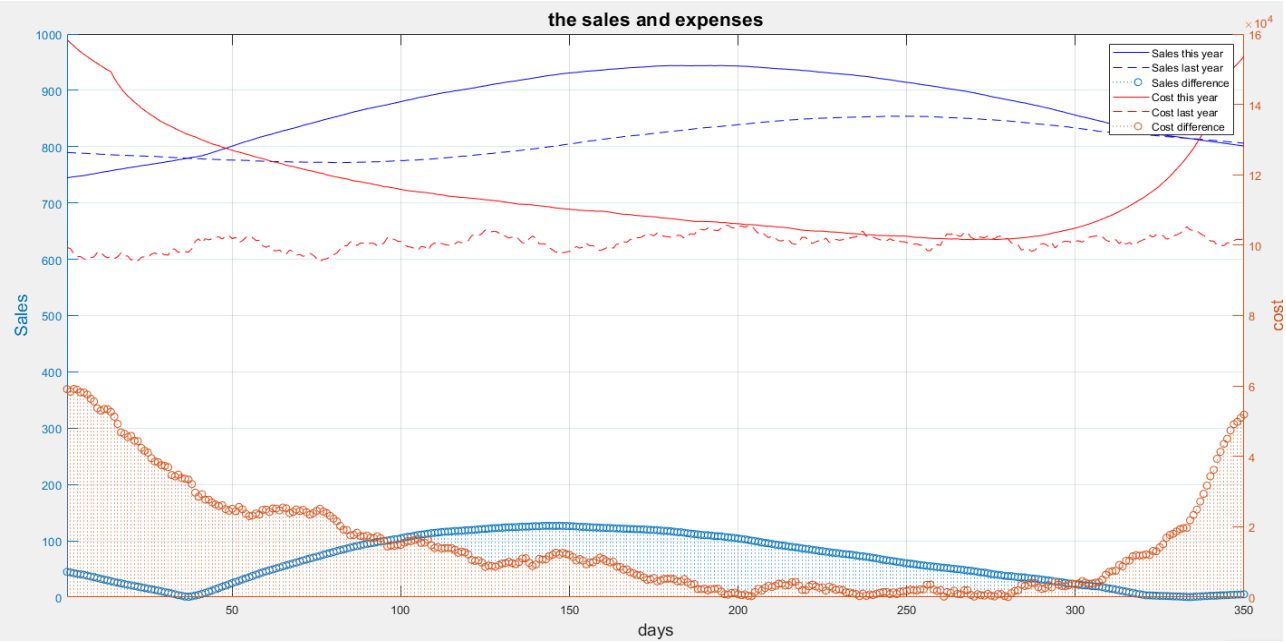
b)

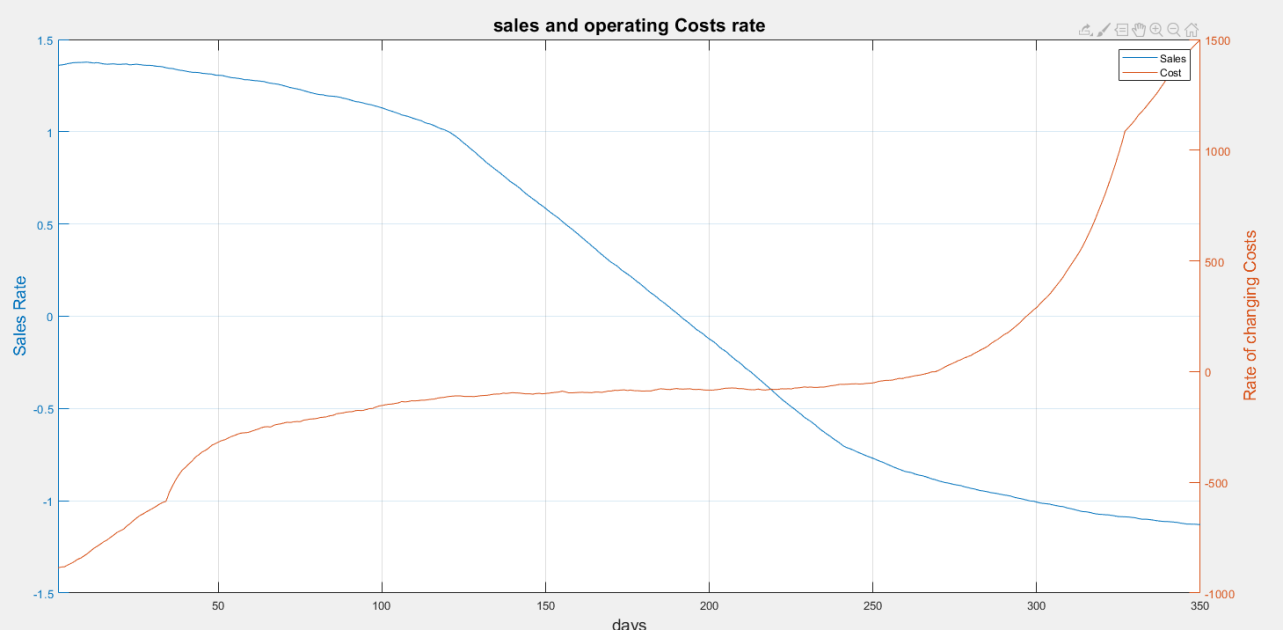


A visualization of the sales and expenses from the past four quarters have been created that makes the situation look less dramatic. We have used chart junk to display the data.

c)

Kenneth and Jeffrey got busted by the SEC for providing incorrect information about the company's financial situation to the shareholders. A truthful visualization of the sales and expenses has been created below by improving using MATLAB to see the standard graph and improved one.





Problem C4: Improving Real-world Visualizations

a. The visualization of the figure is not good, I think. The graphics causes misleading inferences about the data because percentage is not decreasing in but look like it is decreasing. Difference between day and night have not showed properly.

b. Proposed improvement of the first figure:

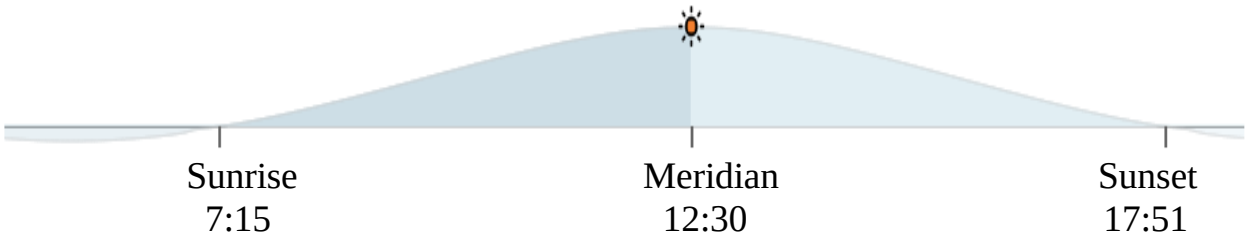


Figure 3: Daylight hours

Second Figure:

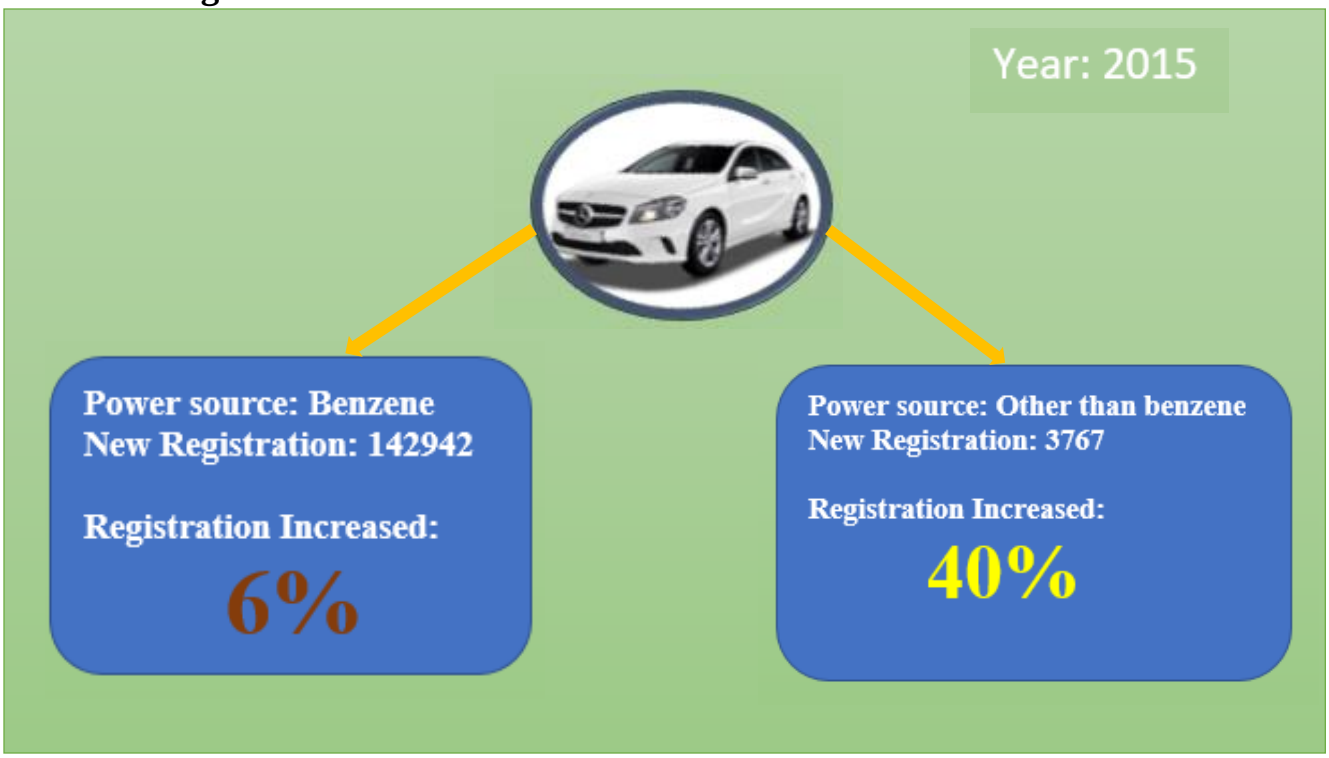


Figure 4: First-time registrations for different power source cars

Third Figure:

Figure 5: A lot of activity in department stores. According to statistics Finland retail sales grew 2.4 percent in August compared to August 2015.

c. This figure represents amount of unemployed job seekers and amount of open jobs from 2007 to 2016. The graph is proving good visualization result.

Index

Problem C3: Chartjunk and Data-Ink Ratio

a

In [ ]:

```
rng(24);
days=1:90;
noise=randi([-15,15],1,90);
sales=-1.5*[1:90]+noise+898;
cost=(exp(days/20)+randn([1,90])+100)*1000;
figure1 = figure;
yyaxis left
plot(days,sales)
ylabel('sales','FontSize',14)
hold on
yyaxis right
plot(days,cost,'r')
ylabel('cost','FontSize',14)
xlabel('days','FontSize',14)
title('Sales and Operating Costs','FontSize',16)
saveas(figure1,'figure1.png')
```


Sales and Operating Costs
Sales this year
Sales last year
sales
cost
days
×10^5
920
900
880
860
840
820
800
780
760
740
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1
0
10
20
30
40
50
60
70
80
90

b

In [ ]:

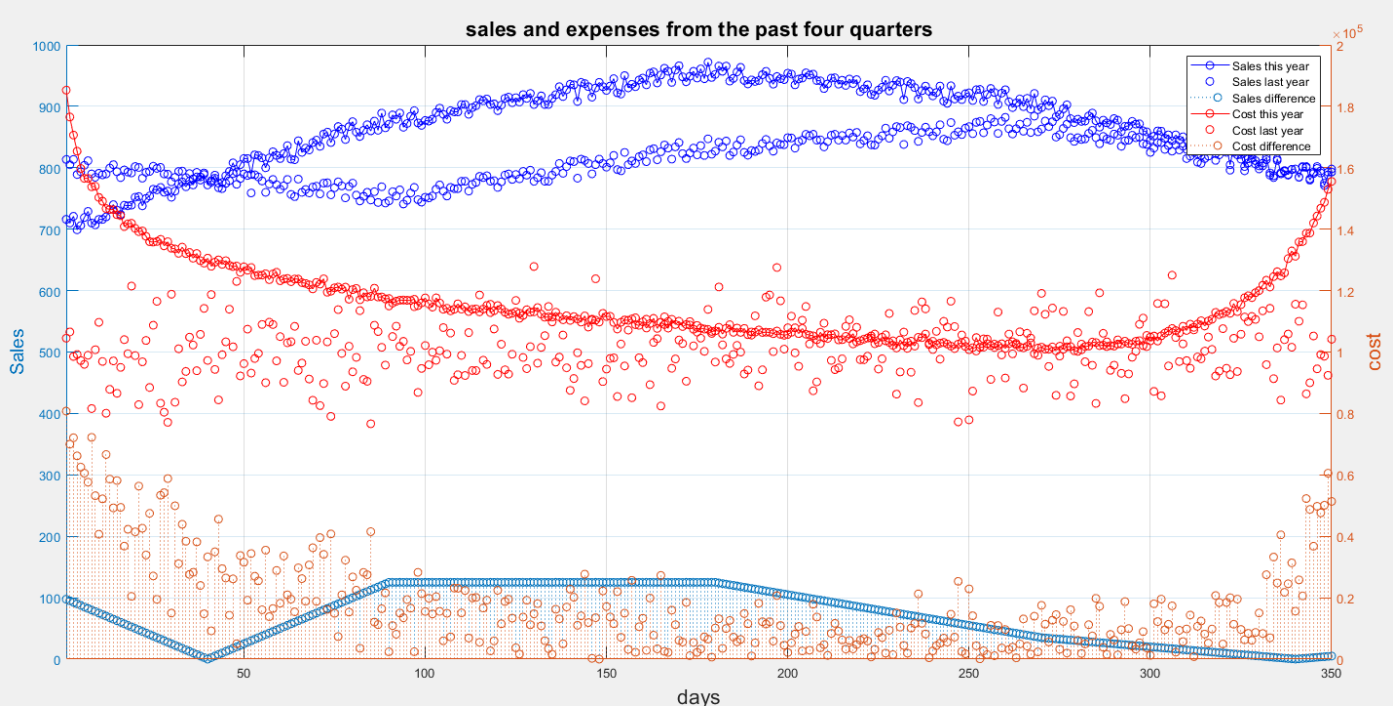
```
rng(24);
days=1:360;
noise=randi([-15,15],1,90)
sales_1Q=2*[1:90]+noise+700;
sales_2Q=[1:90]+noise+sales_1Q(end);
sales_3Q=-0.5*[1:90]+noise+sales_2Q(end);
sales_4Q=-1.5*[1:90]+noise+sales_3Q(end);
sales=[sales_1Q sales_2Q sales_3Q sales_4Q];

sales_1Q_L=-0.5*[1:90]+noise+800;
sales_2Q_L=[1:90]+noise+sales_1Q_L(end);
sales_3Q_L=0.5*[1:90]+noise+sales_2Q_L(end);
sales_4Q_L=-1*[1:90]+noise+sales_3Q_L(end);
sales_L=[sales_1Q_L sales_2Q_L sales_3Q_L sales_4Q_L];

noise2=randn([1,90]);
cost_Q4=(exp([1:90]/20)+noise2+100)*1000;
cost_Q123=(-15*log([1:270]/2)+randn([1,270])+100)*1000+75000;
cost=[cost_Q123 cost_Q4];

cost_L=normrnd(100000,10000,[1,360]);

figure2 = figure;
yyaxis left
plot(days,sales,'o-b')
hold on
plot(days,sales_L,'ob')
hold on
stem(abs(sales-sales_L))
ylabel('Sales','FontSize',14)
yyaxis right
plot(days,cost,'o-r')
hold on
plot(days,cost_L,'or')
hold on
stem(abs(cost-cost_L))
grid on
xlim([1 350])
ylabel('cost','FontSize',16)
xlabel('days','FontSize',16)
title('sales and expenses from the past four quarters','FontSize',16)
legend(['Sales this year'],['Sales last year'],['Sales difference'],['Cost this yea
r'], ['Cost last year'],['Cost difference'])
saveas(figure2,'Figure1.png')
```


sales and expenses from the past four quarters
Sales this year
Sales last year
Sales difference
Cost this year
Cost last year
Cost difference
Sales
cost
days
×10^5
0
100
200
300
400
500
600
700
800
900
1000
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
50
100
150
200
250
300
350

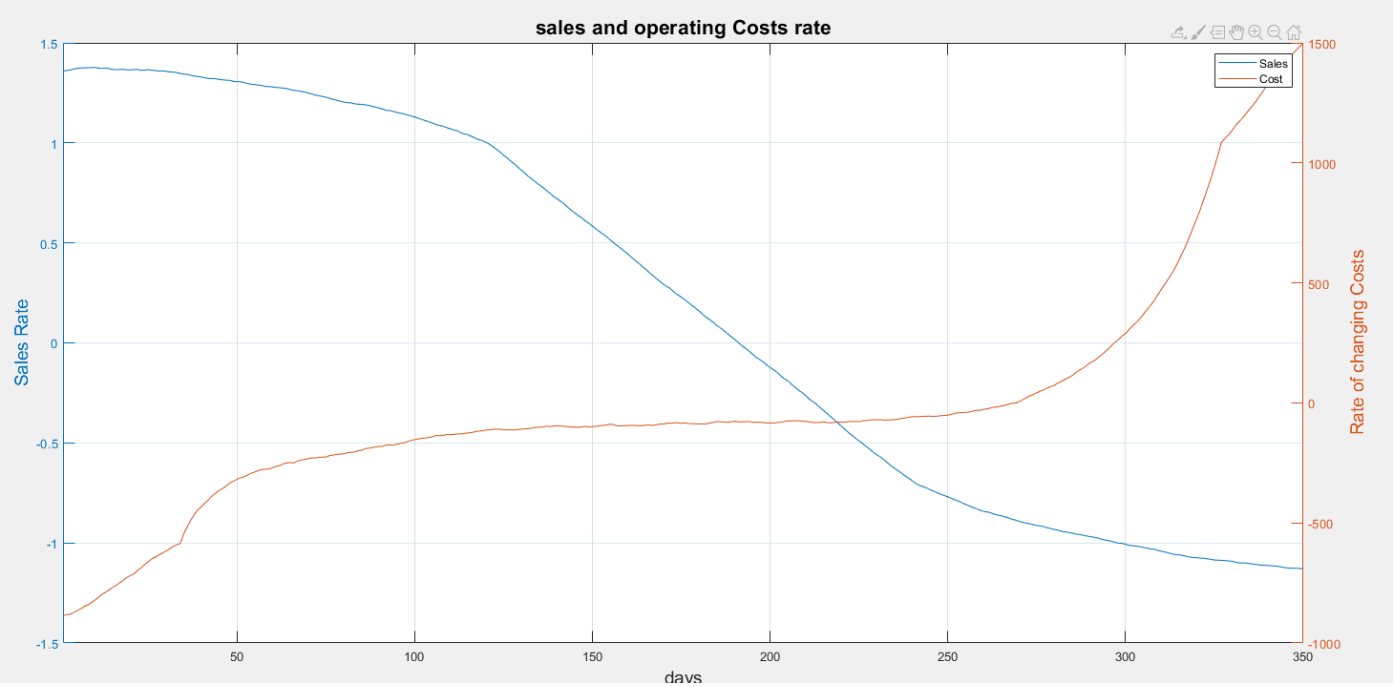
c

In [ ]:

```
figure3 = figure;
yyaxis left
plot(days,smoothdata(sales),'b')
hold on
plot(days,smoothdata(sales_L),'--b')
hold on
stem(abs(smoothdata(sales)-smoothdata(sales_L)))
ylabel('Sales','FontSize',14)
yyaxis right
plot(days,smoothdata(cost),'r')
hold on
plot(days,smoothdata(cost_L,'movmean',20),'--r')
hold on
stem(abs(smoothdata(cost)-smoothdata(cost_L,'movmean',20)))
grid on
xlim([1 350])
ylabel('cost','FontSize',14)
xlabel('days','FontSize',14)
title('the sales and expenses','FontSize',16)
legend(['Sales this year'],['Sales last year'],['Sales difference'],['Cost this yea
r'], ['Cost last year'],['Cost difference'])
saveas(figure3,'Figure3.png')
```

In [ ]:

```
figure4 = figure;
yyaxis left
plot(smoothdata(diff(smoothdata(sales))))
hold on
ylabel('Sales Rate','FontSize',14)
yyaxis right
plot(smoothdata(diff(smoothdata(cost))))
xlim([1 350])
grid on
ylabel('Rate of changing Costs','FontSize',14)
xlabel('days','FontSize',14)
title('sales and operating Costs rate','FontSize',16)
legend(['Sales'],['Cost'])
saveas(figure4,'Figure4.png')
```



In [ ]: