Homework 3

# Created by Wei Ye

# Date: June 21, 2022

# Due on June 22, 23:59

## Question 1

Please see the attached code in 'census\_data\_analysis\_hw.m' in the email.

I checked my new file on June 21, it worked. The main change i did is adding additional data into my original data set x in which i don't need to combine data anymore.

## Question 2

y\_obs=[7 3.5 3 1.75 1.25 0.2 -1.2];

y\_pred=my\_linear\_eq(6,-1,y\_obs)

x=linspace(-3,8,size(y\_obs,2));

subplot(2,1,1);

plot(x,y\_pred,"r")

title('Prediction figure')

xlabel('x')

ylabel("y\\_pred")

subplot(2,1,2);

plot(x,y\_obs,"b-");

hold on

plot(x,y\_pred,"r")

title('Figure combined')

square\_errors=(y\_pred-y\_obs).^2;

sse\_total=sum(square\_errors);

MSE=mean(sse\_total)

No, it's not a good fit.

Obvious if we increase the slope, it can fitt the data better (As the the slop is negative), for example, change a=5,b=-0.7

y\_pred1=my\_linear\_eq(5,-0.7,y\_obs);

MSE1=mean(sum((y\_pred1-y\_obs).^2))

It decreases! Since our data is limited, it's very easy to overfit. But in this case, we know we can change the MSE if we change parameters.