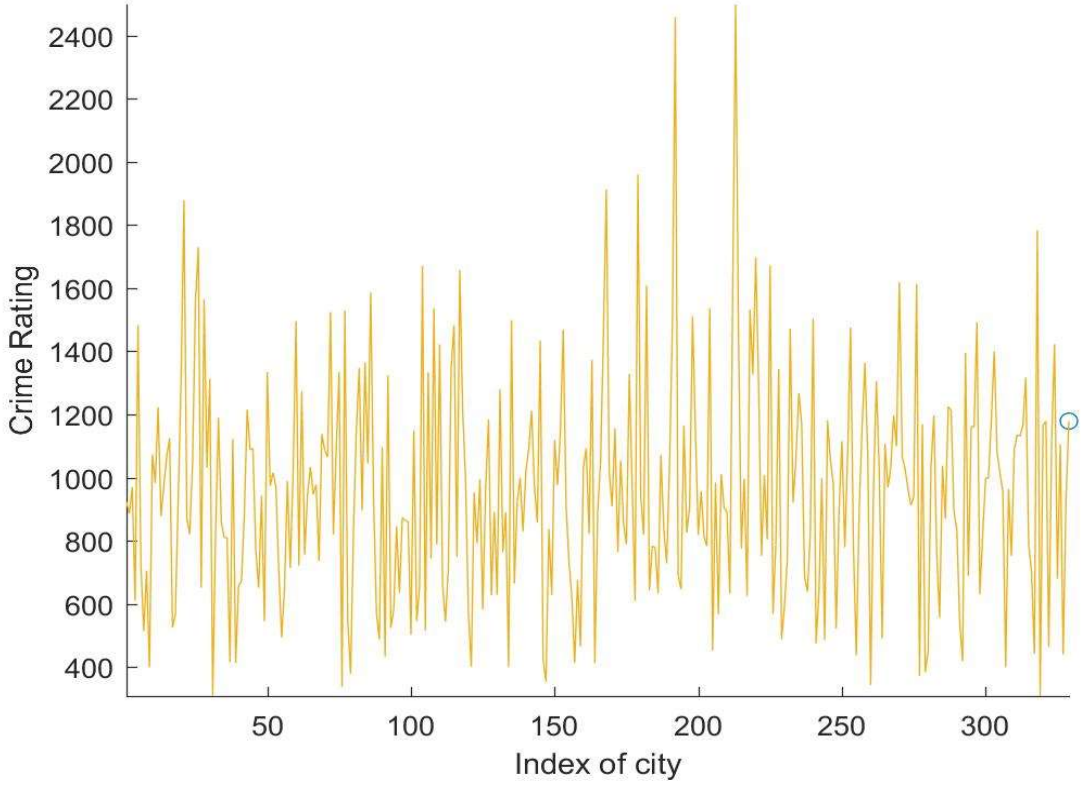
MP1

Following codes are in mp1.m

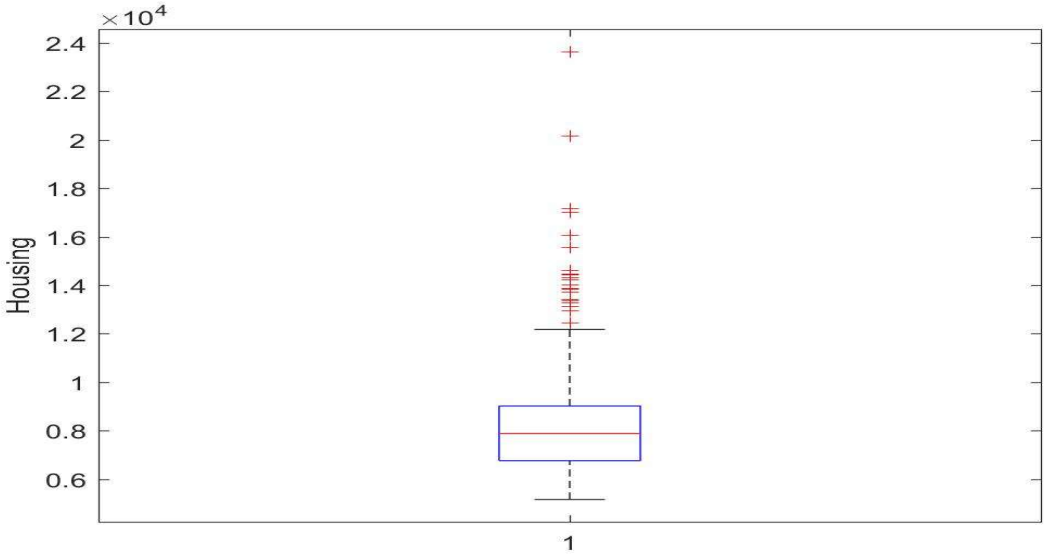
1.1



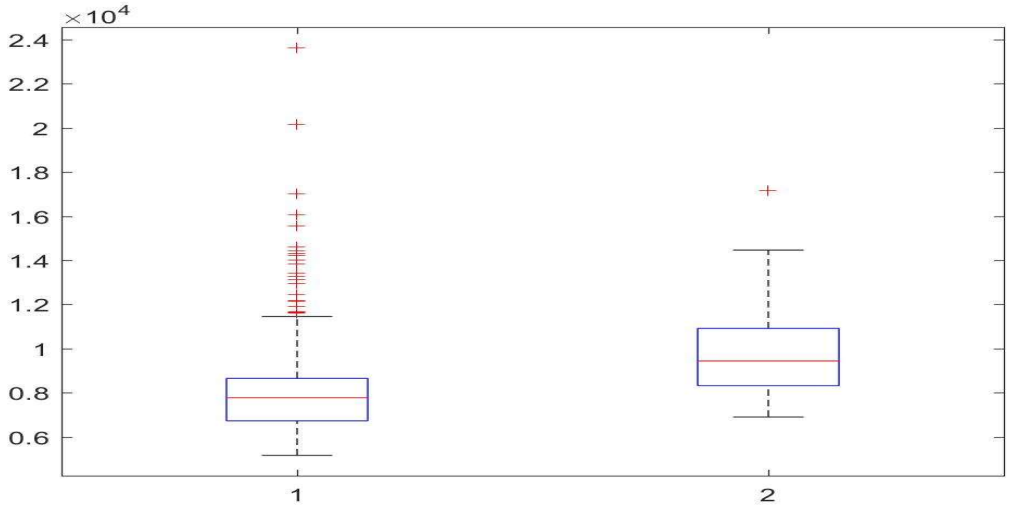
1.2

```
>> for i = 1:329
       if ratings(i,4) == max(ratings(:,4))
          names(i,:)
       end;
end;

ans =

New York, NY
```

2.1

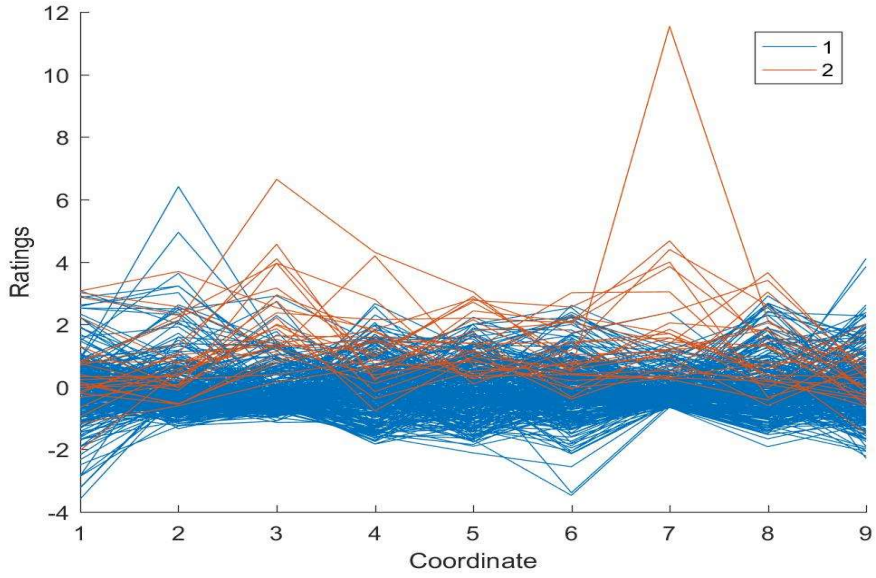


2.2



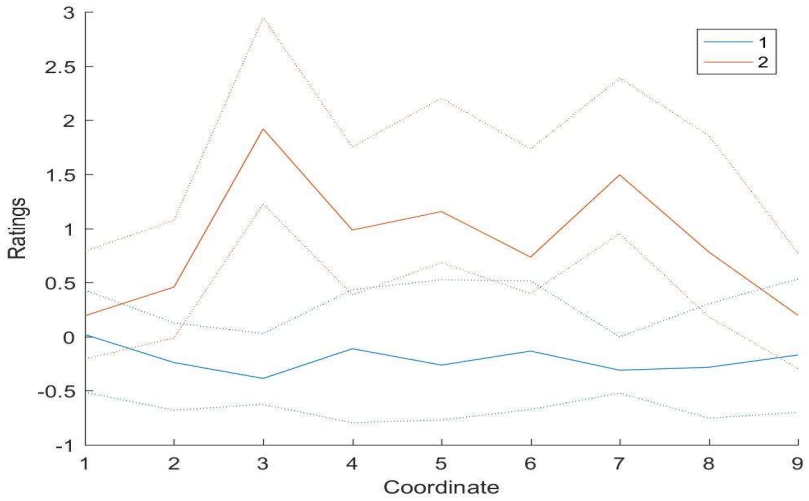

```
>> a = [ratings,group];
boxplot(a(:,2),a(:,10))
>>
```

3.1



3.2

```
>> parallelcoords(ratings, 'group', group,'standardize', ...
'on','quantile',.25);
ylabel('Ratings');
```



3.3

Health is the coordinate 3 and education is 6,

The red line (group 2) is better both in health and education.

MP3

3.a

```
function [V,v] = Edit_Dist(str1,str2)
m=length(str1);
n=length(str2);
v=zeros(m+1,n+1);
for i=1:1:m
    v(i+1,1)=i;
end
for j=1:1:n
    v(1,j+1)=j;
end
for i=1:m
    for j=1:n
        if (str1(i) == str2(j))
            v(i+1,j+1)=v(i,j);
        else
            v(i+1,j+1)=1+min(min(v(i+1,j),v(i,j+1)),v(i,j));
        end
    end
end
V=v(m+1,n+1);
End
```

```
%question 1%
fileID = fopen('customers.csv');

C = textscan(fileID,'%s %s %s %s %s %s %s',...
'Delimiter',',','EmptyValue',-Inf);
fclose(fileID);

Lastname = C{4};
Dist = zeros(401,401);
for i = 2:401
    for ii = 2:401
        Dist(i,ii) = Edit_Dist(Lastname{i,1},Lastname{ii,1});
    end;
end;
```

3.b

```
customerid = C{1};
fid=fopen('MP3.xyang105.txt','wt');
for i = 2:401

    fprintf(fid,'%s\t',customerid{i,1});
    for ii = 2:401
        if Dist(i,ii) < 3 && Dist(i,ii)~=0
            fprintf(fid,'%s\t',customerid{ii,1});
        end;
    end;
    fprintf(fid,'\n');
end;
fclose(fid);
```