

Task1:

In this case the two axes I choose is release year and max resolution.

At first I performed statistical analysis on individual variables:

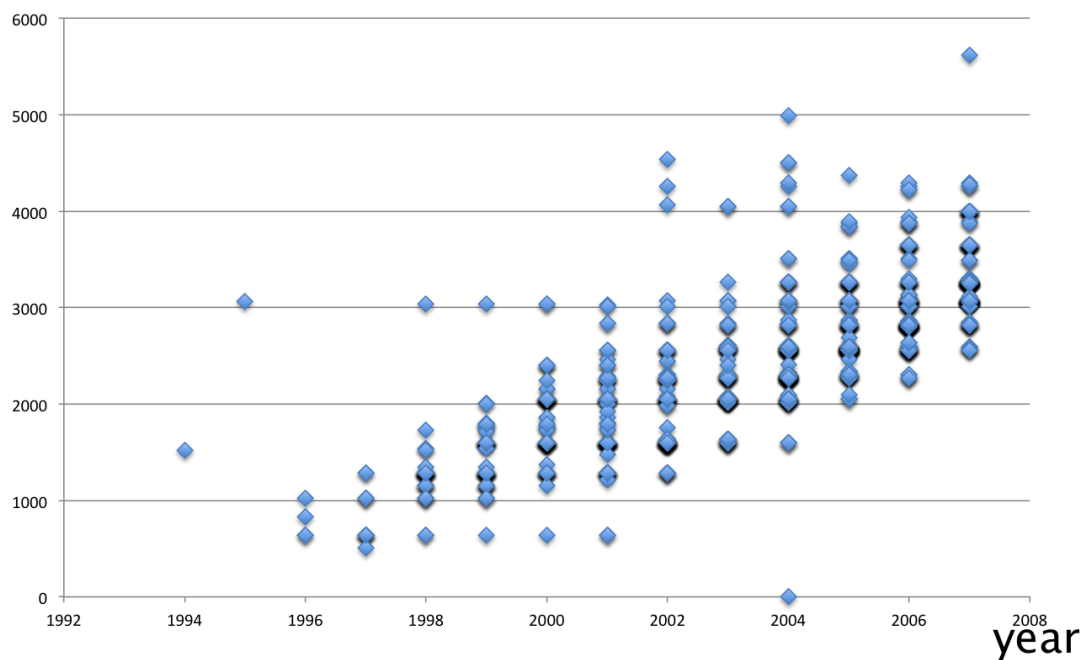
For zoom wide, zoom tele, normal focus range, macro focus range, these four types of data mainly changes based on different series of different brand. For example from the data set I can see the Canon EOS series got 0 on these four columns of data.

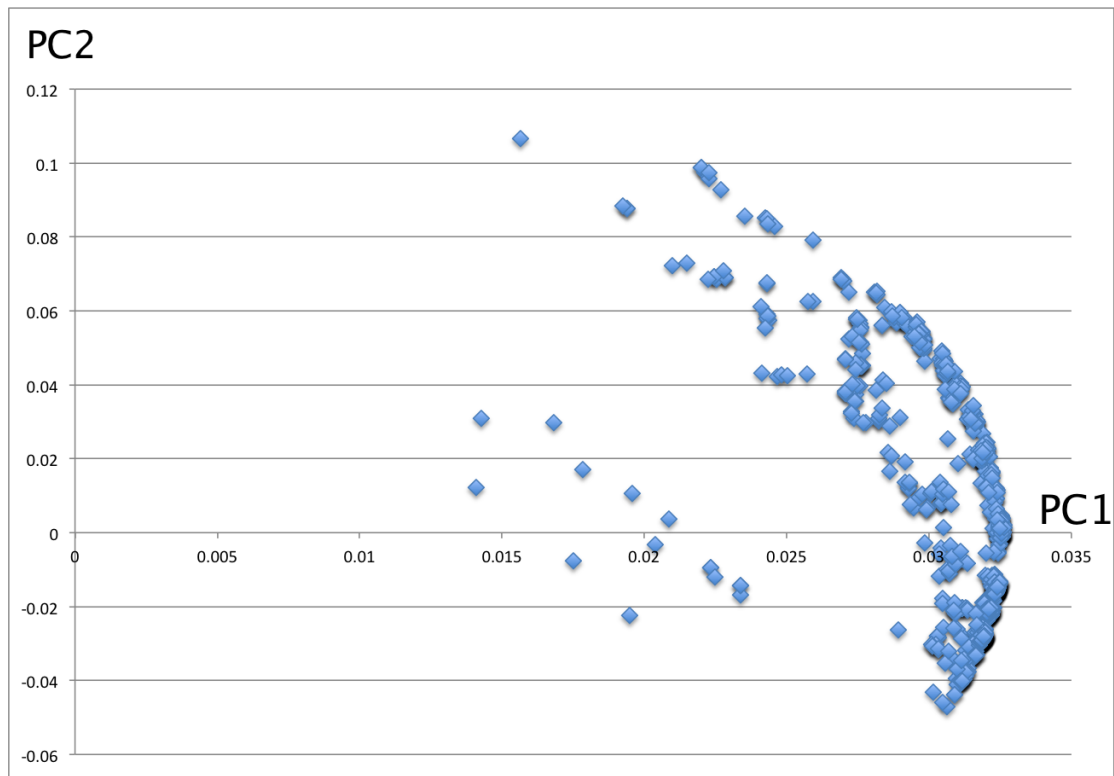
For storage included, weight, effective pixels and dimensions, I think people are not really looking for these variables.

For max and low resolution I use excel to combine them to release data separately and generate two scatter graphs. The graphs show that both of them increase through the increase of release year. Compared by low resolution, I think max resolution is more important in this dataset. So I choose release year and max resolution.

Here is the graph I generate:

max resolution



Task2:

I name the dataset x and the matlab command are as follows:

```
x=x';
stdr=std(x);
[n,m]=size(x)
sddata=x./stdr(ones(n,1),:);
[p,princ,egenvalue]=princomp(sddata)
p2=p(:,1:2)
sc=princ(:,1:2)
egenvalue
per=100*egenvalue/sum(egenvalue)
```

The final per I get is a set of numbers which show the percentage of contribution. The biggest two are 90.4 and 5.2, which belong to release year and max resolution. So these two are the principle components. Then I use the coefficient of principle components to generate the graph, which is shown above.