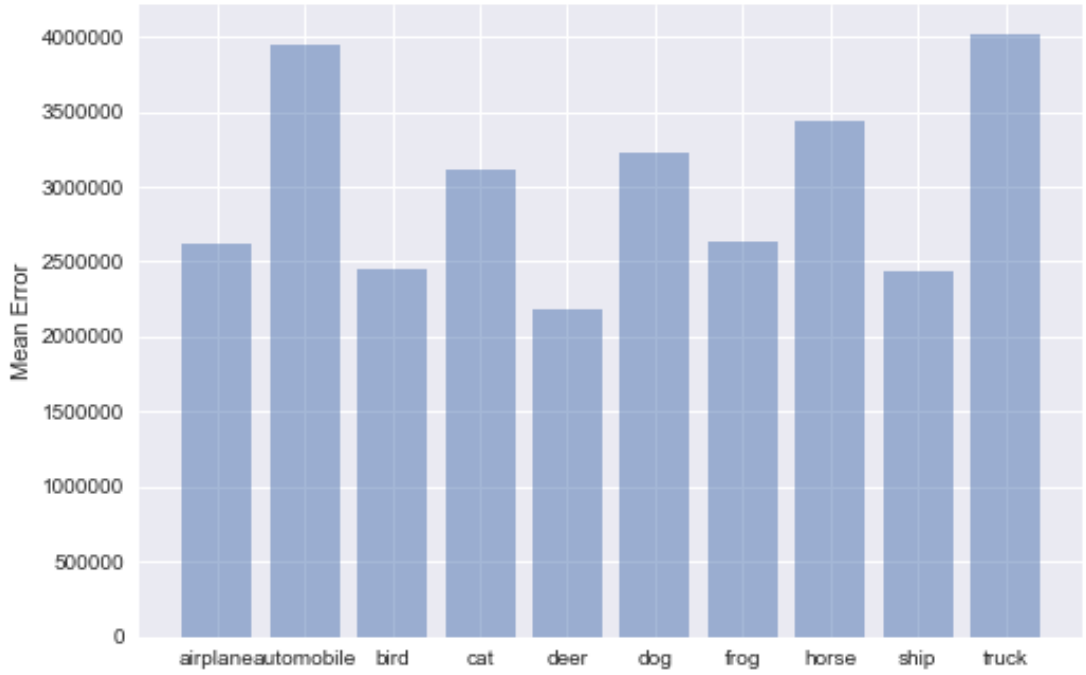
Homework 3 report

1. Error resulting from representing the images of each category using the first 20 principal components against the category



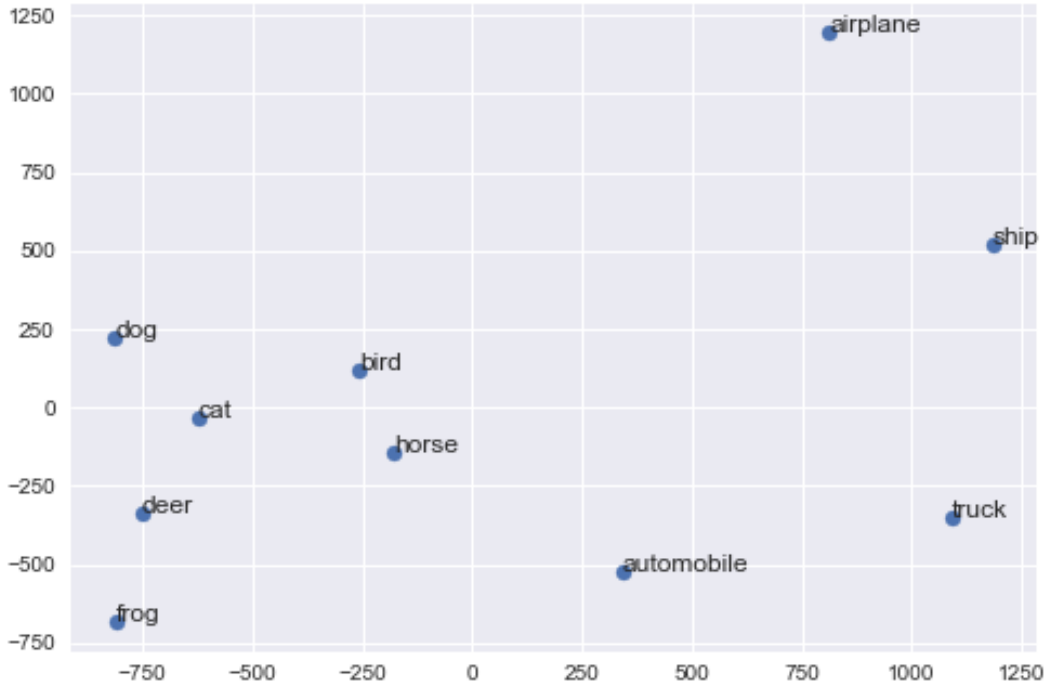
2-1. Distances between mean images for each pair of classes

```
In [160]: DistanceMatrix
```

Out[160]:

	0	1	2	3	4	5	6	7	8	9
0	0.000000	1683.635354	1605.024350	1905.535263	2148.763414	1965.221492	2445.679726	1663.645933	945.541104	1449.094915
1	1683.635354	0.000000	886.236750	1027.649782	1143.081367	1216.079428	1191.191999	950.786078	1303.466548	949.995771
2	1605.024350	886.236750	0.000000	517.311502	601.250335	701.468226	913.747516	418.276306	1557.715017	1416.674733
3	1905.535263	1027.649782	517.311502	0.000000	469.791716	412.181669	677.491969	596.376737	1851.214501	1676.467856
4	2148.763414	1143.081367	601.250335	469.791716	0.000000	617.697141	460.510929	684.346911	2065.621657	1830.740850
5	1965.221492	1216.079428	701.468226	412.181669	617.697141	0.000000	828.581051	843.672091	1897.591821	1880.243772
6	2445.679726	1191.191999	913.747516	677.491969	460.510929	828.581051	0.000000	948.704021	2249.199783	1913.240880
7	1663.645933	950.786078	418.276306	596.376737	684.346911	843.672091	948.704021	0.000000	1660.268080	1347.334102
8	945.541104	1303.466548	1557.715017	1851.214501	2065.621657	1897.591821	2249.199783	1660.268080	0.000000	1066.941632
9	1449.094915	949.995771	1416.674733	1676.467856	1830.740850	1880.243772	1913.240880	1347.334102	1066.941632	0.000000

2-2. 2D map of the means of each categories



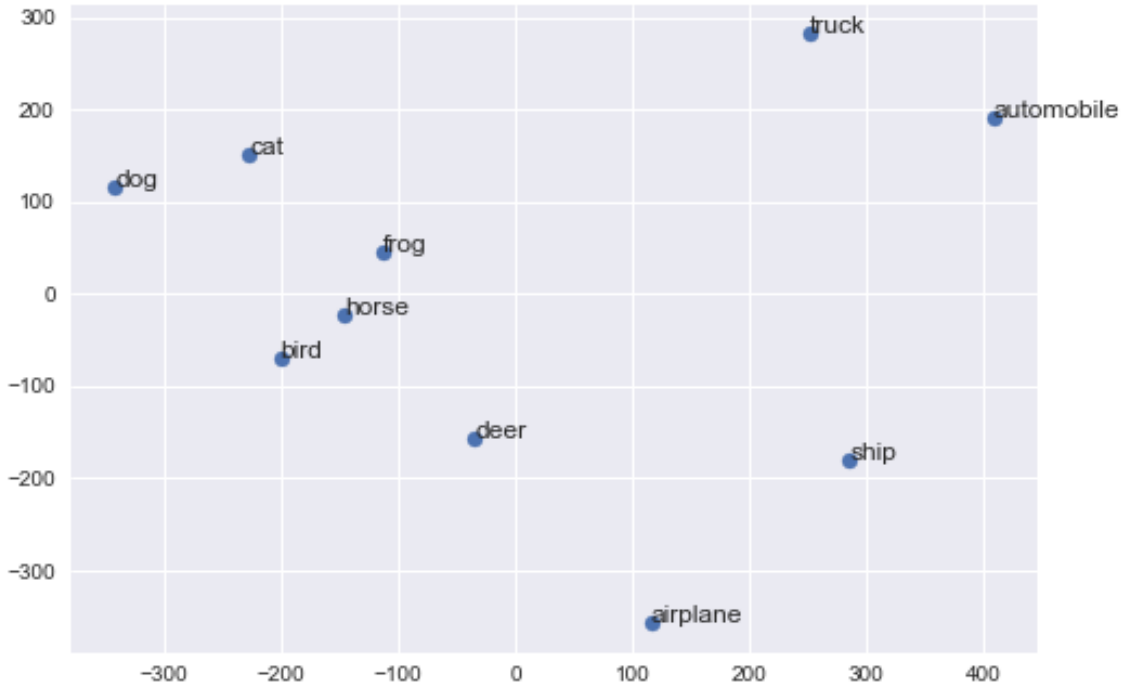
3-1. Distance matrix between classes using definition of (1/2)(E(A | B) + E(B | A)), where E(A | B) to be the average error obtained by representing all the images of class A using the mean of class A and the first 20 principal components of class B.

```
In [194]: pd.DataFrame(DistMatrix)
```

Out[194]:

	0	1	2	3	4	5	6	7	8	9
0	2.620508e+06	3.723684e+06	2.794253e+06	3.306971e+06	2.554937e+06	3.399384e+06	2.949962e+06	3.394522e+06	2.715963e+06	3.794077e+06
1	3.723684e+06	3.950690e+06	3.700088e+06	4.039768e+06	3.450763e+06	4.211536e+06	3.685355e+06	4.230500e+06	3.489308e+06	4.135014e+06
2	2.794253e+06	3.700088e+06	2.447703e+06	2.939465e+06	2.423697e+06	2.967130e+06	2.685703e+06	3.200293e+06	2.813267e+06	3.634823e+06
3	3.306971e+06	4.039768e+06	2.939465e+06	3.116489e+06	2.889815e+06	3.262887e+06	3.028693e+06	3.546528e+06	3.202382e+06	3.897549e+06
4	2.554937e+06	3.450763e+06	2.423697e+06	2.889815e+06	2.180401e+06	2.955137e+06	2.553974e+06	3.041453e+06	2.543438e+06	3.441368e+06
5	3.399384e+06	4.211536e+06	2.967130e+06	3.262887e+06	2.955137e+06	3.231119e+06	3.100522e+06	3.611067e+06	3.386549e+06	4.075478e+06
6	2.949962e+06	3.685355e+06	2.685703e+06	3.028693e+06	2.553974e+06	3.100522e+06	2.630255e+06	3.341750e+06	2.873604e+06	3.648298e+06
7	3.394522e+06	4.230500e+06	3.200293e+06	3.546528e+06	3.041453e+06	3.611067e+06	3.341750e+06	3.441106e+06	3.386185e+06	4.139704e+06
8	2.715963e+06	3.489308e+06	2.813267e+06	3.202382e+06	2.543438e+06	3.386549e+06	2.873604e+06	3.386185e+06	2.440638e+06	3.541629e+06
9	3.794077e+06	4.135014e+06	3.634823e+06	3.897549e+06	3.441368e+06	4.075478e+06	3.648298e+06	4.139704e+06	3.541629e+06	4.021103e+06

3-2. 2D map of the distances of each categories.



This map is different from the previous exercise, because in part two it is Euclidean distances, in two dimension, the distance is :

$$\mathrm{d}(\mathbf{p}, \mathbf{q}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

On the other hand, this exercise define the distance as (1/2)(E(p | q) + E(p | q)), where E(p | q) to be the average error obtained by representing all the images of class p using the mean of class p and the first 20 principal components of class q. Therefore they produces different map.