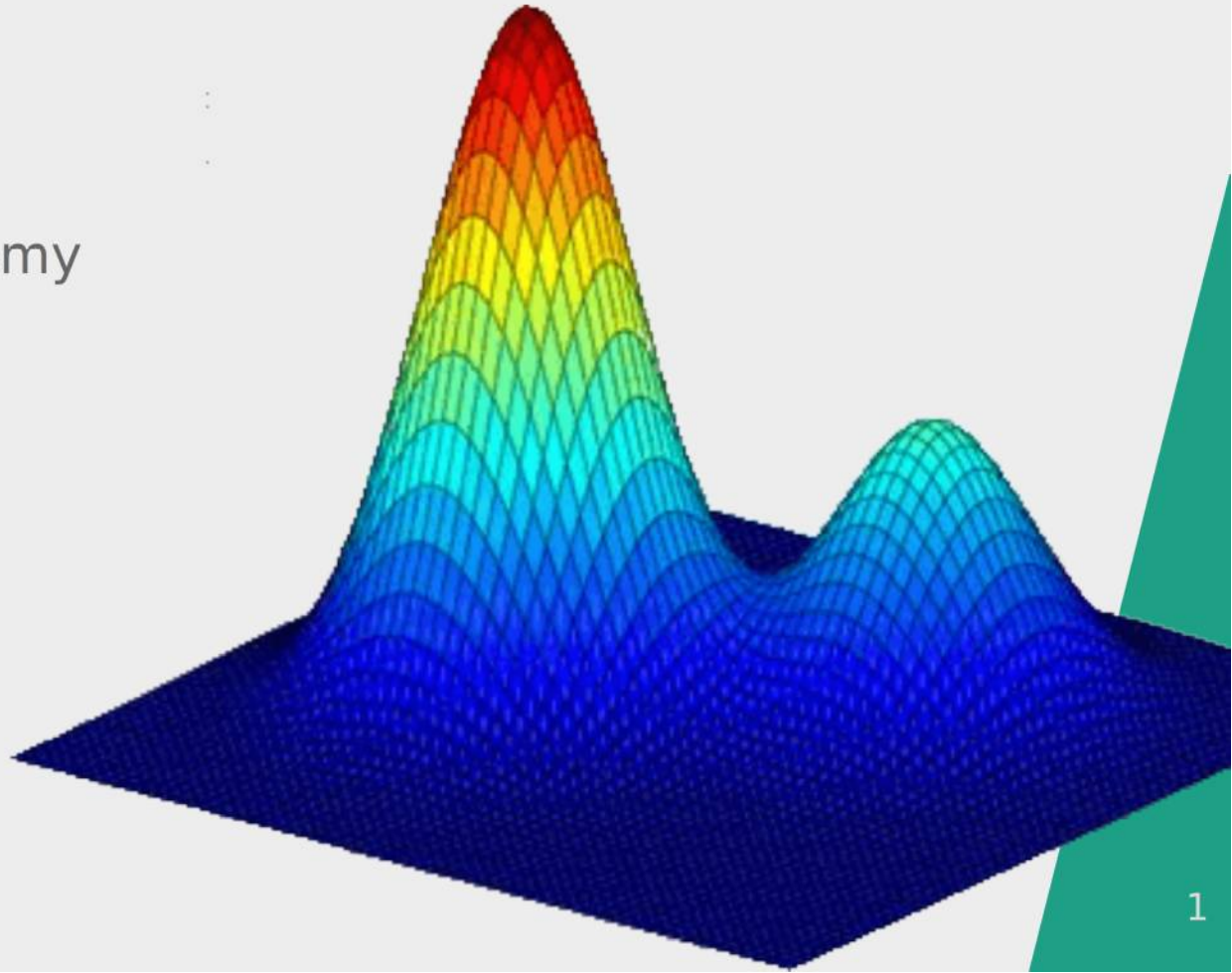
Outlier Detection

with

Gaussian Mixture Model

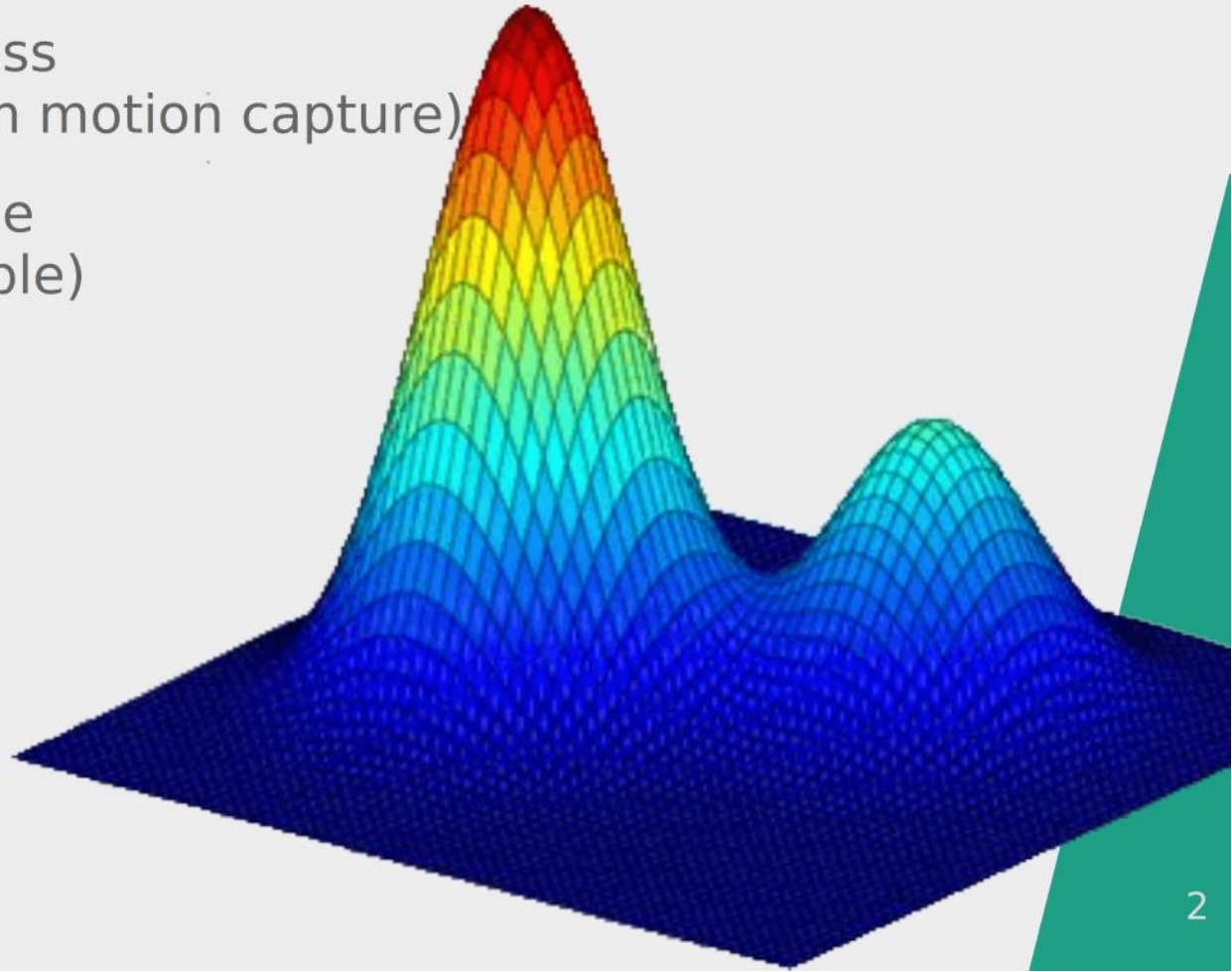
Thomas @ Spiced Academy

Berlin, September 13th, 2021



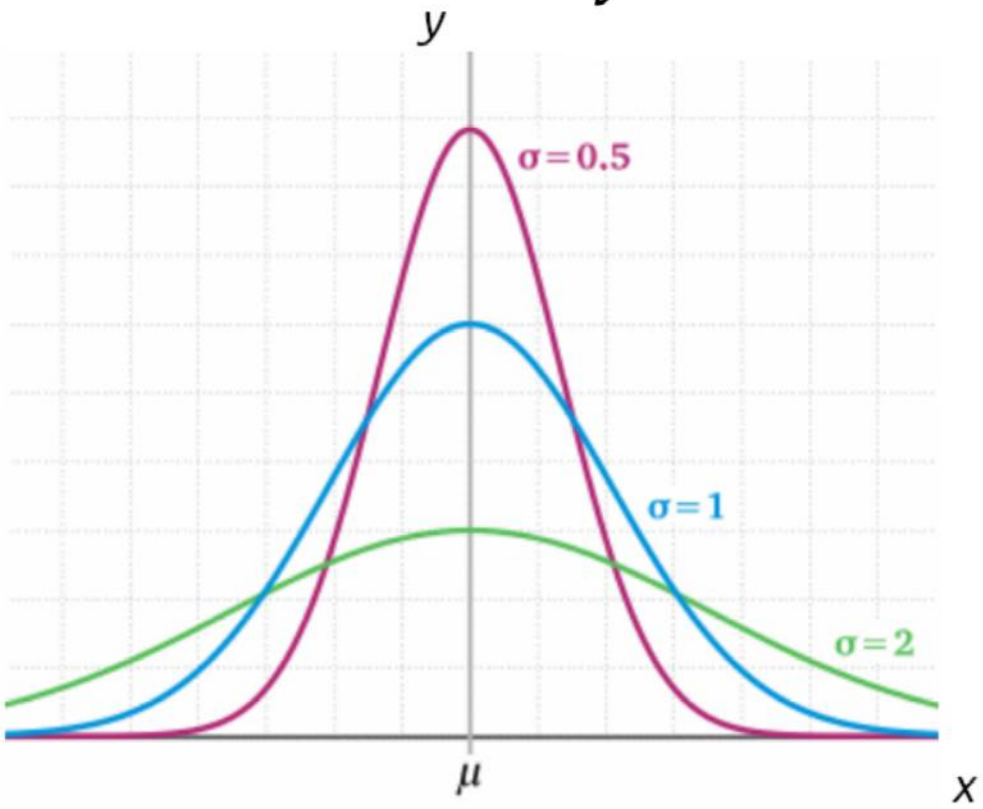
Content of the lecture

- explain the functionality of GMM
  (repeat histogram, normal distribution)
- motivate the usefulness
  (real life example from motion capture)
- present a running code
  (applied on the example)

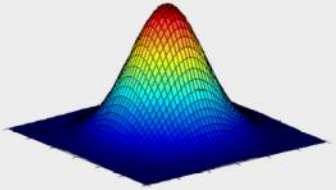


Normal Distribution

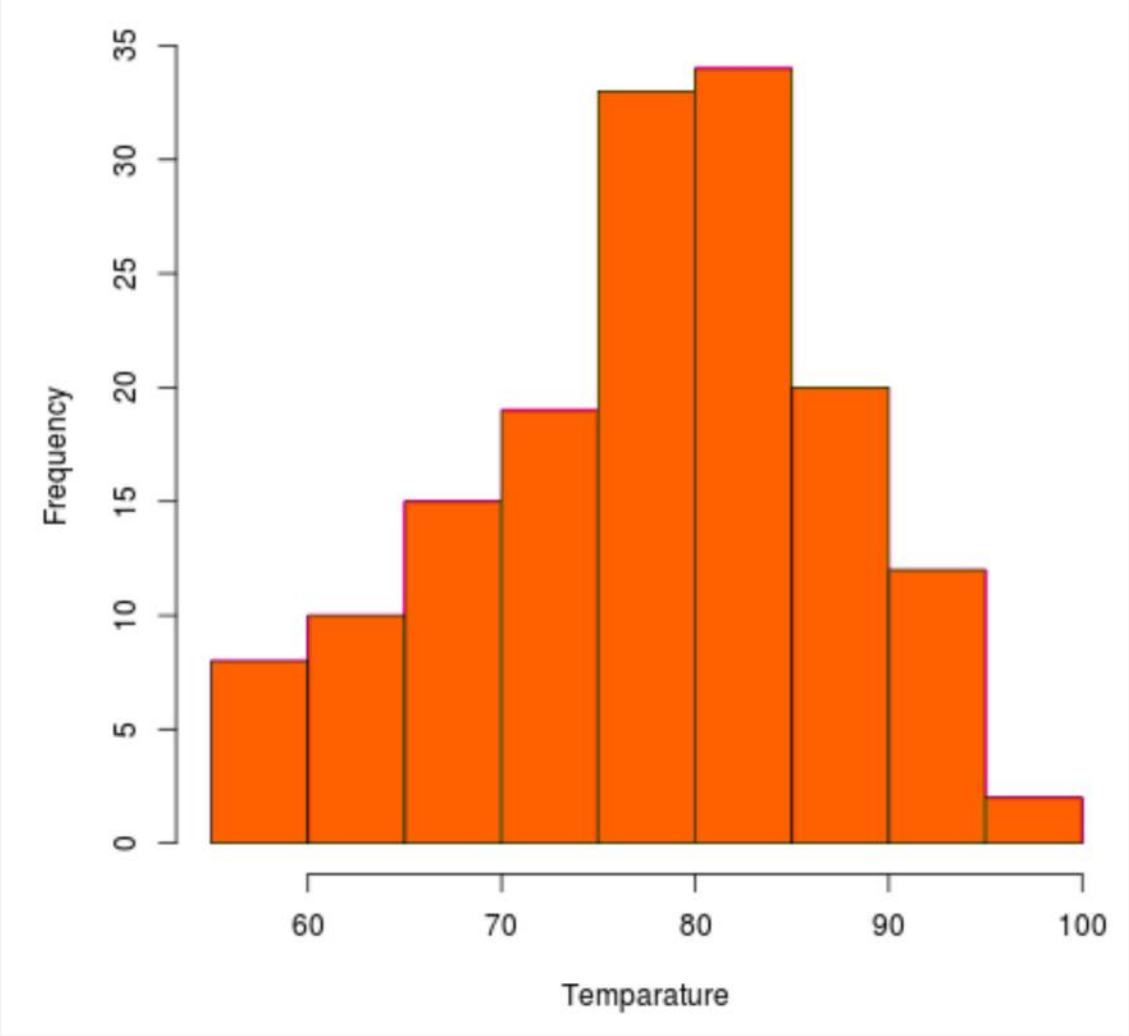
$$y = \mathrm{a} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



also available in 3D, 4D, ... nD



Histogram

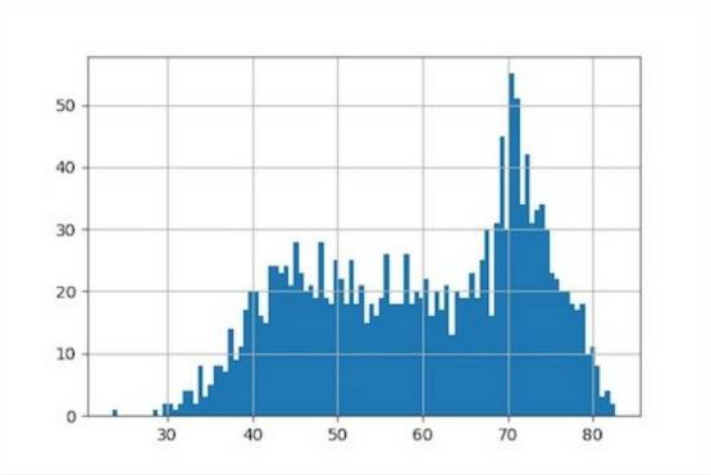


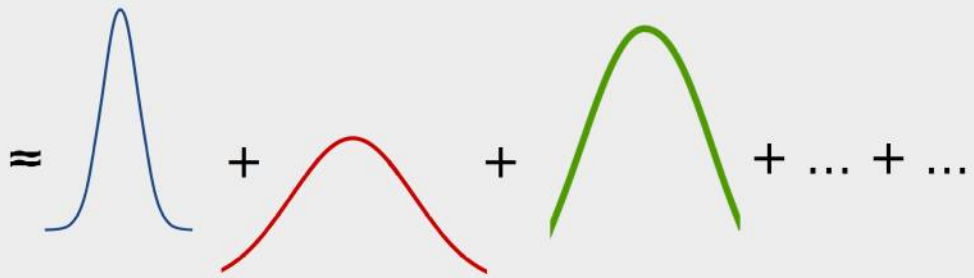
Gaussian Mixture Model

... a probabilistic model for representing the presence of subpopulations.

or: ... a sum of different normal distributions with different means, SD and frequencies

aim: to determine the probability of a given data point belonging to one of the groups



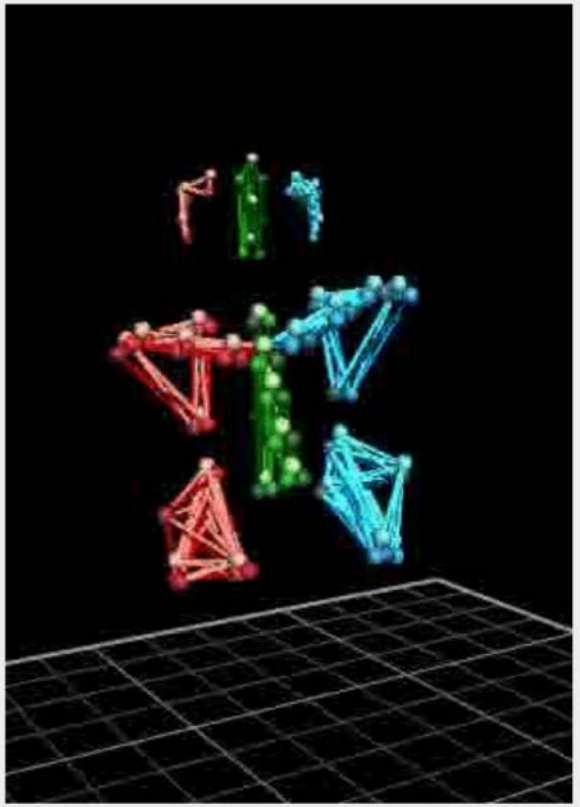


instanciation via:

```
GaussianMixture(n_components=1,covariance_type='full', tol=0.001)
```

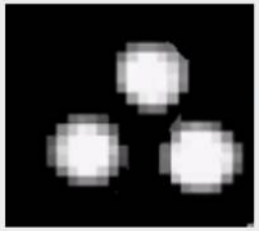
Upper Body Motion Capture

Subject with reflecting markers

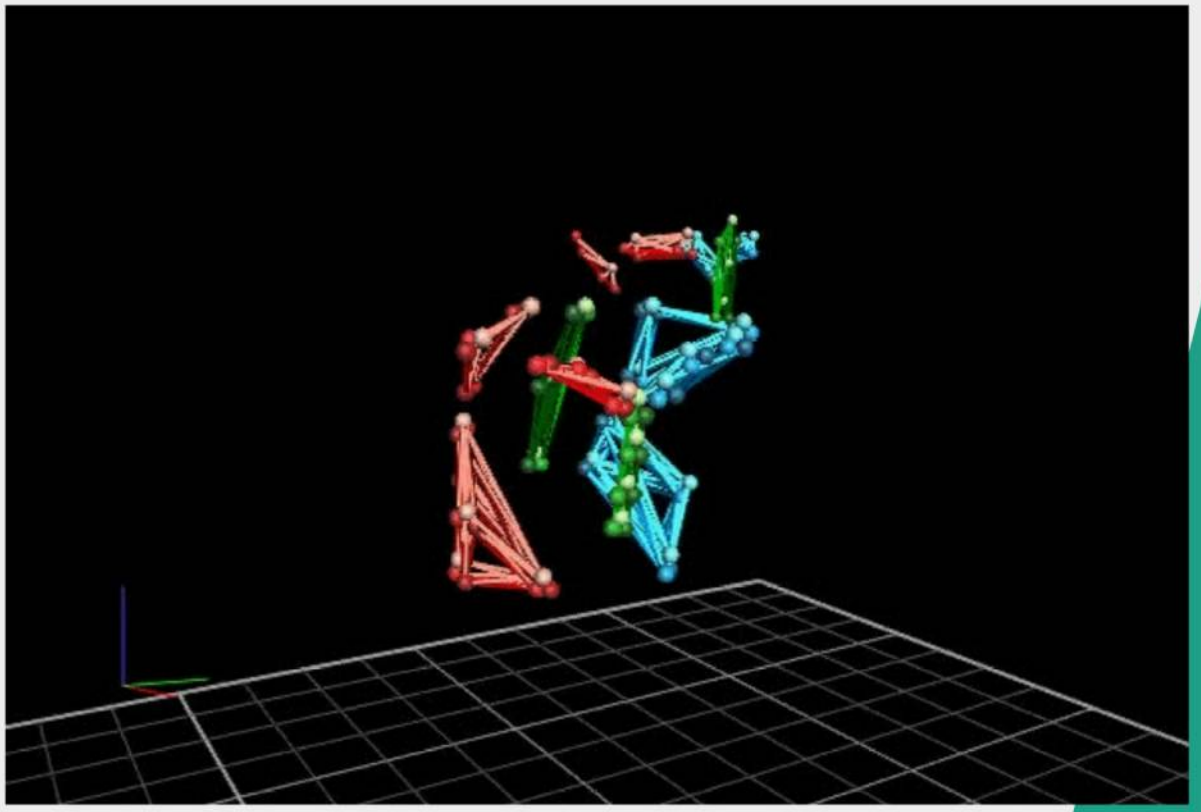


Digitised upper body

Infrared Camera

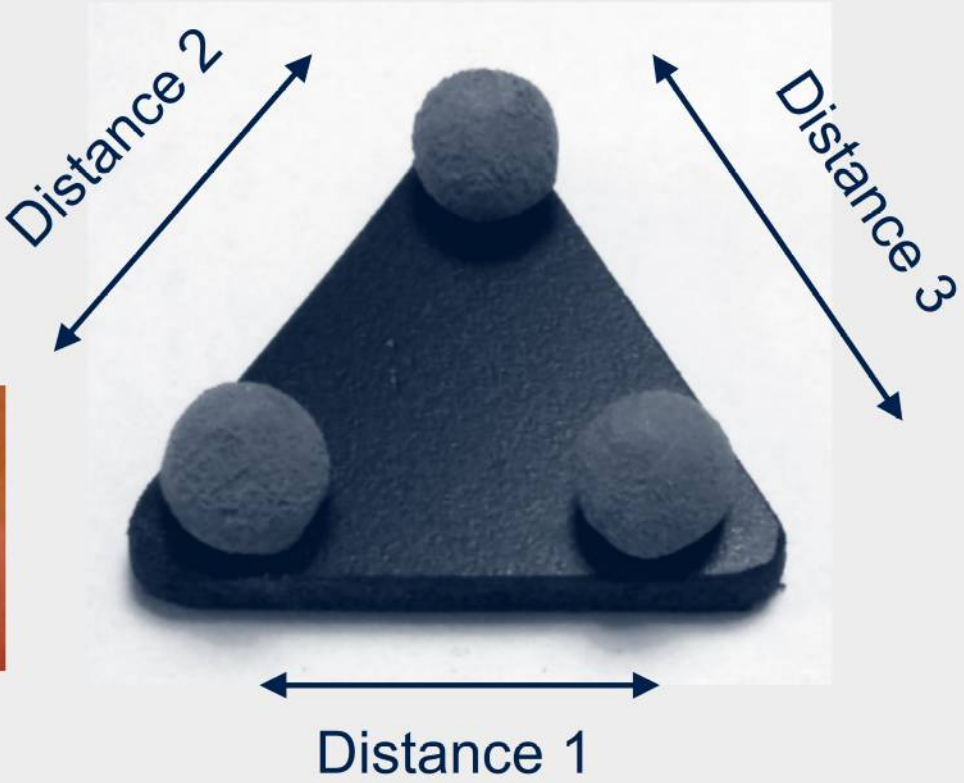


Digitisation



Digitised model **but with flickering markers**

Idea for detecting wrongly digitised coordinates?



let‘s have a look at the code …

References

Spiced course material:

https://krspiced.pythonanywhere.com/chapters/project_movie_recommender/unsupervised_learning/README.html#anomaly-detection

explanation of the method:

https://jakevdp.github.io/PythonDataScienceHandbook/05.12-gaussian-mixtures.html

sklearn documentation:

https://scikit-learn.org/stable/modules/mixture.html#gmm

class parameters, methods and attributes:

https://scikit-learn.org/stable/modules/generated/sklearn.mixture.GaussianMixture.html

explanation of the different available covariance types:

https://stats.stackexchange.com/questions/326671/different-covariance-types-for-gaussian-mixture-models

code and document:

https://github.com/th-under/GaussianMixtureModel

Thanks for your attention