Lab 2

# Task 1

1. **When can you use linear regression?**  
   You can use linear regression to make a forecast of upcoming data points or get a mean value of the already existing data. To get a good result, the data must be correlated in a linear pattern. Linear regression assumes the following:

* The relationship between X and Y is linear
* Y is distributed normally at each value of X
* The variance of Y at every value of X is the same (homogeneity of variances)
* The observations are independent

1. **How can you generalize linear regression models to account for more complex relationships among the data?**You can draw from an array of different distributions to find the “best” fit model. (GLM – Generalized Linear Models: Poisson regression, normal regression and binomial regression)
2. **What are the basis functions?**The basis functions are the functions that are used to project our one-dimensional values into higher dimensions to find more complex relationships between x1 and x2.
3. **How many basis functions can you use in the same regression model?**We did not find a good answer for this question, but it should be
4. **Can overfitting be a problem? And if so, what can you do about it?**

Yes, overfitting can create misleading data (even outside the min/max boundaries). An example in the beginning of the graph in the figure below. If you test your model with different datasets, use cross-validation or/and stop early in the fitting, you can prevent it.

