1. **Let's assume that you have World Bank data on financial, economic and social indicators for several countries. You want to measure the factors that affect the level of development in these countries. To this end, you decide to use per capita income as a proxy for the development level, which is defined as the national income divided by the population. You want to use some features in your dataset to predict per capita income. Is this task a classification or a regression task? Why?**

This will be a regression task.  
The income per capita by itself is continuous variable. There is not set numbers which can be classified. Additionally, Some features which may have relation to ‘per capita income’ may also be continuous. A continuous variable requires a regression task in order to predict this datasets ‘per capita income’

1. **Which of the following arguments are false and why?**
   * *OLS is a special type of linear regression models*

False, The OLS is not a type of model but is a algorithm used in estimating the coefficients.

* + Regression models become useless if they don’t meet the assumptions of linear regression.

True, Without these assumptions, the results of the regression model will be inaccurate

* + *Estimation and prediction are the same thing in the context of linear regression models.*

False, estimation uses an unknown parameter and prediction uses a guess random value.

* + Linear regression is only one kind of regression model. Regression problems can also be solved with other kind of models like Support Vector Machines or Random Forests.

True, there are several kinds of regression models which we can choose and experiment from.

1. **Assume that your project manager wants you to discover which free services your company offers make your customers buy more of your paid services. Formulate this task as a regression problem and write down the potential outcome and features that you’d like to work on.**

Initially, we gather the data on the free services provided by the company. This should also include data that provides features such as conversion to paid service or time/amount spent using the app, as well as age or demographics of users.

Users who spend more time on the service most likely will want more premium services and are more willing to pay for such

Demographics of users will provide us information on the users as working adults are more likely to pay for an service rather than students or children.

Features, if the user is an artist, features may target this individual to use the service more.

Conversion rate, by looking at how the user spends time with the service and what prompted them to pay for the service instead of just the free version.