In the file `data.csv`, you will find a dataset from Compustat, consisting of all US firms that have assets above $10M, with the following variables:

`gvkey`	Compustat identifier
`fyear`	Fiscal year (2015 only in this dataset)
`tic`	Stock ticker
`conm`	Company name
`oiadp`	Operating income after depreciation (a.k.a. EBIT, earnings before interest and taxes)
`at`	Total assets
`ind`	Industry code: 10 Energy 15 Materials 20 Industrials 25 Consumer Discretionary 30 Consumer Staples 35 Health Care 40 Financials 45 Information Technology 50 Telecommunication Services 55 Utilities

Using your favorite statistical analysis package (e.g., Stata or R), please perform the following exercises on the above data set:

1. Define the following variables:
   - Return on assets: `roa = oiadp / at`
   - The (natural) logarithm of assets: `lat = log(at)`

   Then regress (using ordinary least squares) `roa` onto `lat`
   a. What is the regression coefficient? Is it statistically significant?
   b. Describe the meaning of the regression coefficient in one sentence.

2. Now perform the same regression, but include industry fixed effects:
   a. Which industry sector is on average most profitable?
   b. What is the regression coefficient of `lat`? Is it statistically significant?
   c. Describe the meaning of the regression coefficient in one sentence.
   d. In a few sentences, explain the meaning of the difference of this coefficient compared to the previous one (in exercise 1).

3. Now add the squared term of the log-assets as a regressor to the regression of exercise 1 (no fixed effects), i.e. regress `roa` onto `lat` and `lat^2`
   a. What are the regression coefficients of `lat` and its square? Are they statistically significant?
   b. Is there a level of assets for which, on average, profitability is at a minimum or maximum? If so, at which level?
   c. In a few sentences, provide a potential explanation for why the relation between RoA and company size is as you observe.