

# ETC3250 Lab 7

*Di Cook*

*8 September 2015*

## Making data plots and cleaning data

### Assignment 7

On the PISA data from last week. Make a plot to answer this question:

- Does truancy affect math score, on average?
- And pick one other interesting question based on the data dictionary description information to answer using a plot.

Write a couple of sentence explaining why you chose the particular plot design, and how it helps to answer the question.

Turn in two items: a .Rmd document, and the output .pdf from running it. No need to include the R output and plots in your pdf, but the code should be in the Rmd file.

For truancy:

Below you have the means and standard deviations of math scores against frequency of skipping school. Skipping school decreases the score on average by almost 100 points once students skip school 5 or more times per month. There is a lot of variability, though, so some students who skip school 5 or more times per month still do better than some students who never skip school.

```
## [1] 14481      80
```

```
## Source: local data frame [4 x 3]
```

```
##
```

```
##   truancy      m      s
```

```
##   (fctr)    (dbl)  (dbl)
```

```
## 1      0 510.0515 96.83555
```

```
## 2      1 471.2836 90.10048
```

```
## 3      2 445.0754 92.83297
```

```
## 4      3 428.6922 98.79920
```

