City Semester 13-14 Ms. Nelson

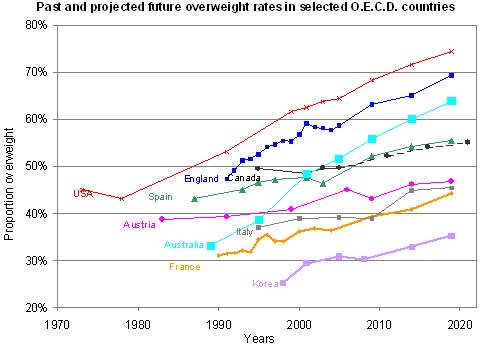
# PROBLEM SET 2-2: Review Univariate/Linear Modeling with Data Sets

1. The solution file for PeoplePerCar has been posted in the Unit 1 folder.

* Compare that to your work on that problem and identify where you could have improved. Now look at your Knowledge Check (SubwayStops) **and add detail** where necessary / **fix mistakes**.
* If you would like a fresh copy of the Knowledge Check to work with, it’s posted in the Unit 1 folder.

1. We all know Mayor Bloomberg made it a personal quest to control what New Yorkers eat in an effort to improve health – largely motivated by statistics like this: *the percentage of New York State adults who are overweight or obese increased from 42% in 1997 to 60% in 2008.*   
   (source: NYS Dept of Health)

New Yorkers compare slightly more favorably to US citizens as a whole, but not much. The graph below shows the percent of the population in various countries that are overweight.



*Note: this includes both those that are overweight (BMI>25) and obese (BMI>30).*

Source: *Global Nutrition Market, Obesity and World Health*. Global Sherpa, n.d. Web. 25 Feb. 2013. <http://www.globalsherpa.org/nutrition-market-obesity-malnutrition>.

* 1. According to the graph, what was the approximate percentage of the US population that was overweight in 1995? Is your answer an interpolation or an extrapolation?
  2. The graph is cut off at about 2013.  Based on the portion of the graph that is shown, about what percent of the US population do you think will be overweight in 2020? Is your answer an interpolation or an extrapolation?
  3. What portion of the US population do you think will be overweight in 2100? How did you arrive at that answer? (hint: linear means constant rate of change, so use the slope) Are you more or less confident in this answer than in part b?

1. 🖳 Average Hourly Wage for workers in NYC 2009 to 2013
   1. Open the Fathom data file “AvgHourlyEarnings-NYC” and **fit a linear model to the data** with “Months since Jan09” as the input and “AverageHourlyEarning” as the output.
   2. Right-click on the graph after you’ve plotted your function and choose “Make Residual Plot” (near the bottom). What is the residual plot showing you?
   3. Looking at the residual plot, do you think a linear model is appropriate for this data? Do you think your model is a useful predictor of data values? Explain.