# Homework №4

## Problem 1

910 randomly sampled registered voters from Tampa, FL were asked if they thought workers who have illegally entered the US should be (i) allowed to keep their jobs and apply for US citizenship, (ii) allowed to keep their jobs as temporary guest workers but not allowed to apply for US citizenship, or (iii) lose their jobs and have to leave the country. The results of the survey by political ideology are shown below.

|  |  | Political ideology | | | |
| --- | --- | --- | --- | --- | --- |
|  |  | Conservative | Moderate | Liberal | Total |
|  | (i) Apply for citizenship | 57 | 120 | 101 | 278 |
| Response | (ii) Guest worker | 121 | 113 | 28 | 262 |
|  | (iii) Leave the country | 179 | 126 | 45 | 350 |
|  | (iv) Not sure | 15 | 4 | 1 | 20 |
|  | Total | 372 | 363 | 175 | 910 |

(a) What percent of these Tampa, FL voters identify themselves as conservatives?

(b) What percent of these Tampa, FL voters are in favor of the citizenship option?

(c) What percent of these Tampa, FL voters identify themselves as conservatives and are in favor of the citizenship option?

(d) What percent of these Tampa, FL voters who identify themselves as conservatives are also in favor of the citizenship option? What percent of moderates share this view? What percent of liberals share this view?

(e) Do political ideology and views on immigration appear to be associated? Explain your reasoning.

## Problem 2

(This problem is inspired by 2.172 in the textbook.) The website TED.com offers free short presentations, called TED Talks, on a variety of interesting subjects. One of the talks is called "The Happy Planet Index," by Nic Marks.[1] Marks comments that we regularly measure and report economic data on countries, such as Gross National Product, when we really ought to be measuring the well-being of the people in the countries. He calls this measure Happiness, with larger numbers indicating greater happiness, health, and well-being. In addition, he believes we ought to be measuring the ecological footprint, per capita, of the country, with larger numbers indicating greater use of resources (such as gas and electricity) and more damage to the planet.

You can find a tidy version of the 2012 Happy Planet Index (`hpi-tidy.csv`) on the course webpage (An untidy version can be downloaded from http://www.happyplanetindex.org/data/). In this homework problem you will continue to explore the Happy Planet Index data using R. Please copy all plots from RStudio into a typesetting program to reduce wasted paper.

---

[1]Marks, N. "The Happy Planet Index," www.TED.com/talks, August 29, 2010.

A basic description of all of the variables included in the data set is given below:

| Variable | Description |
| --- | --- |
| HPIRank | HPI rank for the country |
| Country | Name of country |
| LifeExpectancy | Average life expectancy (in years) |
| Wellbeing | "Ladder of Life" index from the Gallup World Poll (0 = worst possible life, 10 = best possible life) |
| HappyLifeYears | Index variable combining life expectancy and well-being |
| Footprint | Ecological footprint–a measure of the per capita ecological impact |
| HappyPlanetIndex | Happy Planet Index (0–100 scale) |
| Population | Population (in millions) |
| GDPcapita | Gross Domestic Product (per capita) |
| GovernanceRank | Governance ranking (1 = highest) |
| Region | Region of the world |

(a) Compute the following summary statistics for the Happy Planet Index by Region: minimum, Q1, median, Q3, maximum, mean, and standard deviation. Report your results as a table.

(b) Create side-by-side boxplot of life expectancy by region and briefly describe what you learn from the plot.

(c) Create overlaid density plots of life expectancy by region. What aspects of the distributions are easier to see using this these density plots than the boxplots you created in the previous part? What aspects are harder/impossible to see?

(d) Create a scatterplot of happy life years (on the $y$-axis) against ecological footprint (on the $x$-axis). Describe the relationship between happy life years and ecological footprint, commenting on the direction, form, and strength, along with any outlying observations.

(e) Create another scatterplot of happy life years against ecological footprint, but this time use different colors and shapes to represent the region for each country. This can be done by using the following argument for `geom_point`:

```
mapping = aes(x = Footprint, y = HappyLifeYears, shape = Region, colour = Region)
```

What do you learn by adding this extra information to the plot?