

Data acquisition and visualisation

Autumn 2015

Your name:

Provide answers for the eleven questions below and submit them no later than the 16th of December.

Question 1

According to Roger D. Peng in *The Art of Data Science*, what are the six types of question?

..... (6 points)

Question 2

What is Stephen Few's fundamental argument in *Our Irresistible Fascination with All Things Circular*?

..... (3 points)

Question 3

Read the documentation on the World Bank's indicator API:

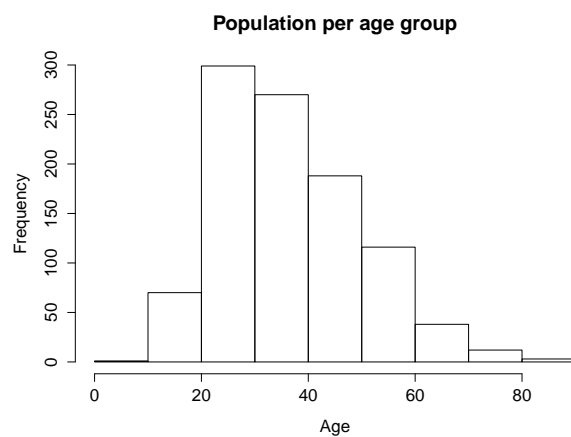
<http://data.worldbank.org/developers/api-overview/indicator-queries>

What is the sum of GDP in current US dollars (indicator NY.GDP.MKTP.CD) for the United States from 1960 to 1969?

..... (5 points)

Question 4

What type of chart is this? Pick one.



- Histogram
- Boxplot
- Area chart

(2 points)

Question 5

What projection is recommended by Landmælingar Íslands when mapping Iceland?

..... (3 points)

Question 6

Highlight and describe four things in this data table that should be addressed in order to tidy the data.

Country	Indicator	2004	2005	2006
United States	FG1	NA	3.4	5.6
United Kingdom	FG1	1.1	NA	2.0
Iceland ^a	FG1	NULL	0.1	0.2
U.S.	VY5	1.1	1.1	1.1
United Kingdom	VY5	4.0 ^b	NA	4.0 ^b
Iceland	VY5	3.2	3.6	3.8

(8 points)

Question 7

Which best describes a large scale map? Pick one.

- The larger the object appears on the map, the larger the map's scale
- The smaller the object appears on the map, the larger the map's scale
- More than 100km is shown on each axis

(3 points)

Question 8

You're creating a map showing the diversity in age of a city's buildings. What is the best colour scheme? Pick one.

- Sequential
- Diverging
- Qualitative

(2 points)

Question 9

Provide three best practices, as suggested by Jonathan Gray et al in *The Data Journalism Handbook*, when making a freedom of information request to a government.

..... (6 points)

Question 10

In *Tidy Data*, what does Hadley Wickham list as the five most common problems with messy datasets?

..... (5 points)

Question 11

There are seven problems in the three charts shown in figures 1–3. Highlight and describe each problem.

(7 points)

Figure 1

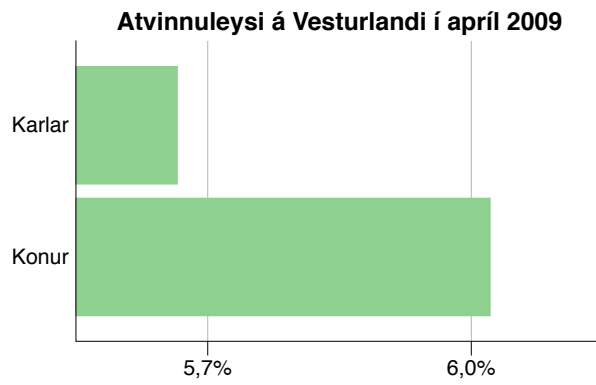


Figure 2

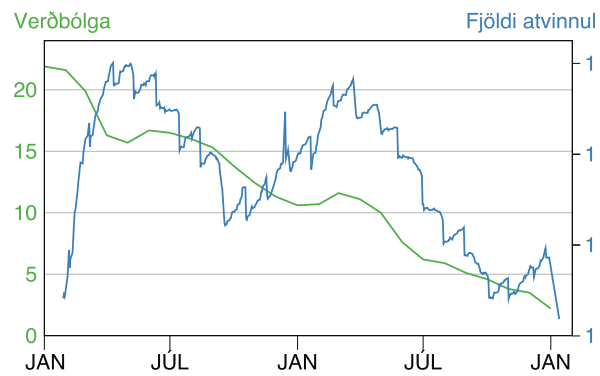


Figure 3

