BITL1

Lab 2

11 January 2022

Introduction to Visualization

---

1. **Warm up:** Go to `twitter.com/jburnmurdoch/status/1479218302954221570` and look at the picture.

   (a) What do you notice about it? What is your first reaction?

   (b) Find something that you think has been done well, and something that you think is missing.

2. **Daily Python:** Download the file `lab2-names.csv`.

   (a) Load it with `f = open('lab2-names.csv','r')` and `data = list(csv.reader(f))`

   (b) How many letters `e` are there in the column `firstname`? For every string `s`, use:

   - `s.count('e')` to count the number of occurences of `'e'` in `s`
   - `s.lower()` to make every character in `s` lowercase

3. **Main task:** Go to each of the following websites, explore it, and download one data set.

   (a) *Atvērto datu portāls:* `data.gov.lv`

   (b) *European data portal:* `data.europa.eu`

   (c) *Kaggle:* `kaggle.com/datasets`

   (d) *OpenfMRI:* `openfmri.org/dataset`

   How are the data sets different? Which websites are easy / difficult to use?

4. Take one data set from part 3. above that has at least 100 numbers.

   (a) Find the maximum, minimum, average values. If your data has several columns, find the min / max / avg for one column.

   (b) Pair up with a classmate and share your data with them.

   i. Do they find the same min / max / avg values?

   ii. Do you see any "outliers" for your data? Discuss with your partner what that means to both of you.

   iii. Decide which of your two data sets you prefer, and use that for part 3. below.

5. Take your chosen data set from part 4. above and import it into Python.

   (a) Plot your data in two different ways. Use `python-graph-gallery.com` for inspiration.

   (b) Submit your image in ORTUS.

6. **For next week:** Find a data visualization on the internet, post it in the `#labs` channel on Slack, and say what you like about it. Be sure to give the source!

   Some places to start looking: Twitter `#dataviz`, Economist "Graphic Detail" blog