Practice exam

This is a practice exam covering part of the topics:
modeling, visualization in R and
basic skills in python.

The final exam includes all topics: make, docker, ...

Some tips)

1. Make your Rstudio clean by doing `rm(list = ls())` or by opening a new project.

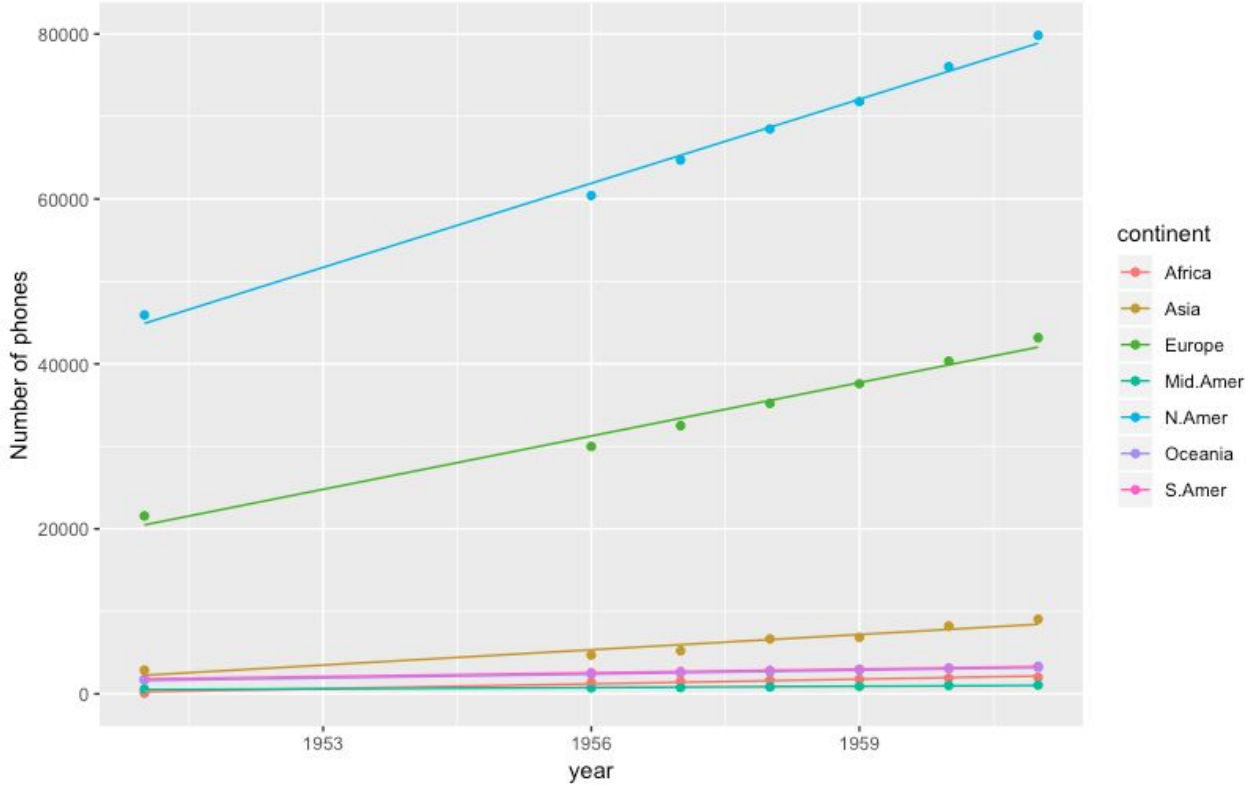
2. Try to knit your Rmd in the middle, not once at the last minute.

Part I R

15 minutes

modeling, visualization

1. reproduce the following graph using `WorldPhones` data in R.
   (Each continent has a linear fitted line.)



2. Clustering and visualization

(a) Use `attitude` data to do K-means clustering.
(K is the smallest integer that gives tot.withinss < 12,000,
where tot.withinss is the within-cluster sum of squares.)

(b) Draw a scatter plot using x = learning, y = complains, and col = cluster.

(c) Using the previous part (b), draw a scatter plot, also having the centers of clusters (centers should have black points with size=3).

Part II

15 minutes

python

1. Write a python function called “prefix_str” that takes in a string and a number x and return the first x letters of that string.
    Eg. prefix_str(“abcde”, 2) should return “ab”

2. Write a function called “col_constraint” that takes in a pandas dataframe, a column name, and a number n. This function should output the rows with values of that column less than n.
    Eg. col_constraint(df, “BMI”, 20) should return the rows of df with BMI less than 20

3. Write a python function called “list_delete” that takes in a list and a number i and output the list without the ith element of that list
    Eg. list_delete([1,2,3,4,5], 2) should return [1,3,4,5]