

Homework

You are data scientist at Sports clothing ltd. Your boss wants you to analyze sales data with R and he wants the following:

- Use Rnotebook (html or word) for creating report;
- Load ggplot2, dplyr and ggmap libraries;
- Import bi_product.txt and bi_salesFact.txt. Decimal character is ».« and delimiter is »;«. Be sure to use stringsAsFactors = FALSE argument with read.csv function;
- Show (table) number of sold units and sum of income for every product category;
- What are the mean and median (Revenue) of every category?
- According to mean and median, what can you say about the skewness of distribution?
- Create chart with sales across category and segment (for help refer to Titanic example). Put »category« in facet_grid. Explain the results;
- Optional: Show random 100 sum of unit sales on a US map. For sampling data you can use sample_n(data, 100) and for creating map you can refer to:
<http://amunategui.github.io/ggmap-example/>;

Possible points: 7.5

Final due date: 2017-05-15 23:59

Please submit your homework to email: uros.godnov@gmail.com with subject »data capturing homework«