

# Assignment 3

Statistics 4420/8426

Exploratory Data Visualization and Quantification

Due on: March 16, 2017 by midnight

1. **Criticizing visual display:** Visit [dadaviz.com](http://dadaviz.com) and explore the many data graphics they display. Not all their graphics are good in terms of our principle. Now answer the following questions.
  - a) Pick and provide the links for two of the best graphics from this web site. Describe the reasons for your pick.
  - b) Identify and provide the links for two of the bad graphics from this web site and explain why you think they are bad. Clearly mention which principle is not satisfied. Also mention an alternative display that would be preferable.
  - c) Provide the links for two of the graphics that used some unconventional genre and specifically describe why you think they are not conventional.
  - d) In each of the six graphics you have identified above, determine the visual discourse community.
2. **Organizing data for display:** In class we presented how to reorganize data to display and generated a scatter plot of French fries data to show the relationship between replication 1 and 2. Generate and provide a similar plot that shows relationships between time 1 and 10 for each of the sensory attributes. Also provide your codes to demonstrate how you generated the plot.
3. **Colors palettes in graphics:** The following R codes will generate same plot with two different color schemes. Explain what are the differences between those two plots. Which plot do you prefer for display? Explain why.

```
library(ggplot2)
dsamp <- diamonds[sample(nrow(diamonds), 1000), ]
d <- ggplot(data=dsamp, aes(carat, price, colour=clarity)) + geom_point()
d + scale_colour_brewer(palette="Blues")
d + scale_colour_brewer(palette="Set1")
```

**Note:** Please don't try to copy the code to save time. Instead, type them.

4. Carefully read the article available in the link below and answer the following questions.  
<http://cran.r-project.org/web/packages/colospace/vignettes/hcl-colors.pdf>
  - a) Describe the differences between qualitative, sequential and diverging palette.
  - b) Give example situations for each of these palettes for which they are suitable.
5. Diastolic blood pressure (DBP) was measured 5 times for each of the two treatments (TRT) group of subject. The data is provided in `diastolicBP.csv`. Generate a good display of the data. Answer the following questions based on the plot you have generated
  - a) What is the main message?
  - b) What is the sub message?
  - c) What numerical summery of the data we glean from the plot?