36-315: Statistical Graphics and Visualization

Homework 3
Date: January 29, 2002 **Due: start of class February 3, 2002**

---

1. Download `lab3.csv` and source `lab3.r` just like in lab. Load the data into R, extract the column named `MEDRENT` (the median gross rent per census tract), and split according to population density.

   (a) Make a boxplot comparison between high and low population density (city versus country).

   (b) Now overlay the full distributions, and emphasize differences using a diff or ditch plot.

   (c) Looking at the boxplots, one might conclude: "The rents in cities are shifted upward relative to the country." Critically evaluate this statement.

   (d) Make a Q-Q plot comparison between high and low population density. According to this plot, the rent distribution for high population density can be produced by transforming the low-population-density rents. In this transformation, some rents are increased and others decreased. Which ones?

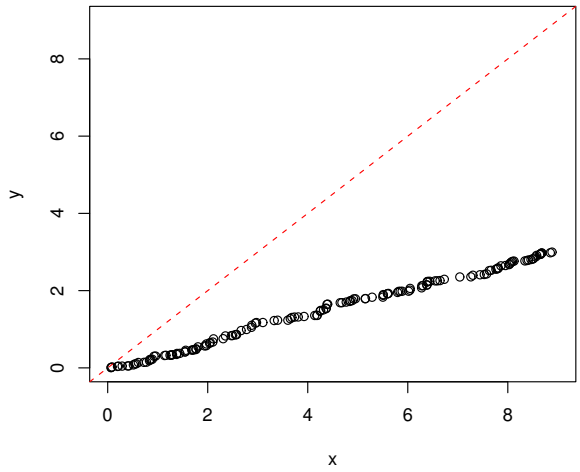
2. From CMU to Shadyside there are two different bus routes, stopping at nearby places on campus (call them A and B) at similar times. Wilbur wants to try using the "bus distribution" idea to decide which bus stop to wait at. So he collects arrival times for both routes, circa 7pm.

   What plot should Wilbur use to determine when to wait at stop A versus stop B? Sketch an example and indicate when he would wait at stop A and when he would wait at stop B. (Hint: which distribution comparison task is this problem most similar to?)

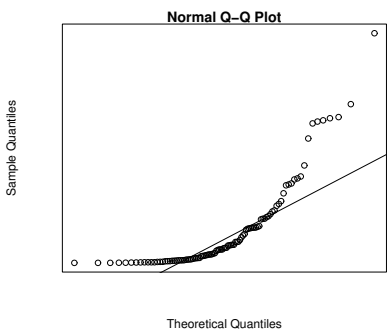
3. Below find a quantile-quantile plot relating a variable x to y.

   (a) What mathematical transformation, when applied to the x data, would produce y's distribution?

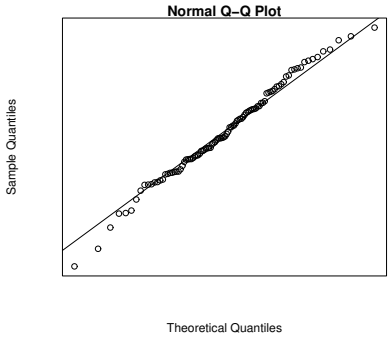
   (b) Sketch what an overlay of the two distributions might look like.



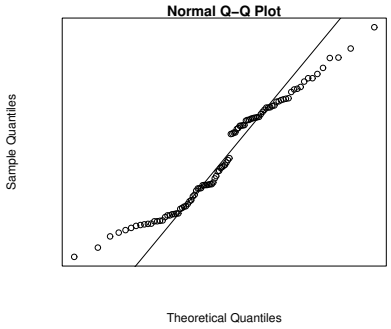
4. A normal Q-Q plot and histodot plot were made for five datasets. However, the plots got mixed up. Match each normal Q-Q plot with its corresponding histodot plot.



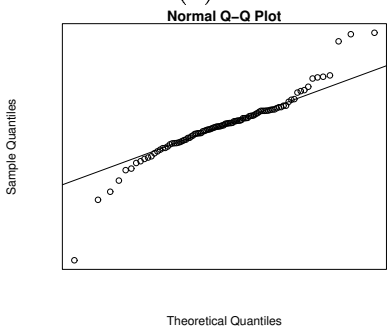
(a)



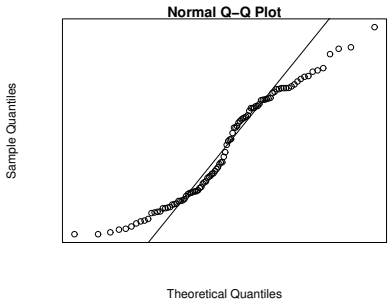
(b)



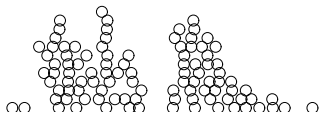
(c)



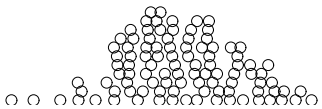
(d)



(e)

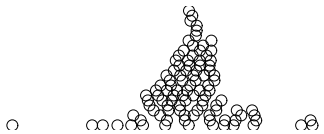


(a)

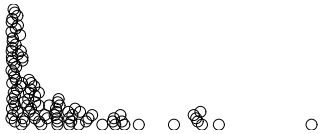


(b)

(c)



(d)



(e)