Name:	ID:

Write **1–3 sentences** for each short answer. Answer **in English, not R code.** Explain any pictures you draw. Try to answer every question to maximize partial credit.

1. **[12 pts]** For each of the items below, give a short definition and a specific example.

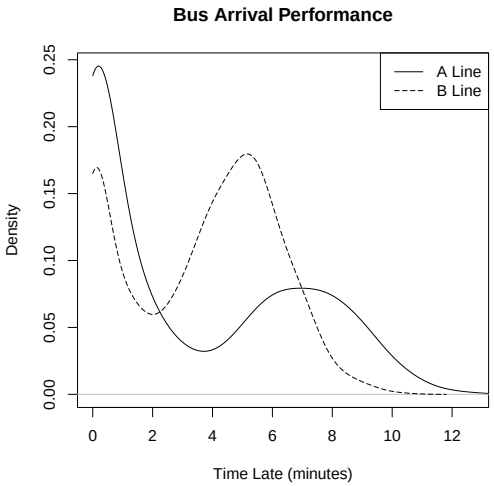
   a) numerical data

   b) ordinal data

   c) mosaic plot

   d) function

2. **[12 pts]** The density plot below shows the arrival performance of two bus lines. Answer each question and *justify your answers* by pointing out details from the plot.



   a) Which distribution is more spread out?

   b) Which bus is more likely to be more than 2 minutes late?

   c) Suppose both buses are more than 2 minutes late. Which bus is more likely to arrive first?

3. **[12 pts]** The table below shows body weight statistics (in pounds) for three groups of cats, each fed a different kind of food.

	Percentile			
Food	25th	50th	75th	Mean
Cruncy Chicken	8.0	12.0	15.0	6.4
Smokey Salmon	7.5	9.4	12.0	9.6
Tasty Tuna	8.4	8.9	9.3	14.0

   a) Which group has the widest distribution of body weights?

   b) Do any of the groups appear to contain outliers? How can you tell?

   c) If someone wants their cat to stay slender, which food is best? Why?

4. **[8 pts]** Explain what it means for a statistic to be "robust."

5. **[12 pts]** Sketch a box plot and label each of the items below.

   a) median

   b) upper whisker

   c) 25th quantile

   d) outlier

6. **[12 pts]** The table below shows a data frame named `cars`. For each command listed, write down the result if the command is valid or "error" if the command is not valid.

```
> cars
            vehicle  mpg  hp weight
1       Camaro Z28 13.3 245   3.84
2          Valiant 18.1 105   3.46
3       Duster 360 14.3 245   3.57
4    Maserati Bora 15.0 335   3.57
5 Pontiac Firebird 19.2 175   3.85
```

   a) `subset(cars, hp <= 200)`

   b) `cars[2, 1]`

   c) `cars[3]`

   d) `cars[3, ]`

7. **[12 pts]** For each of the pairs of variables below, suggest an appropriate plot to determine whether there's a relationship between them.

   a) age (years) and class level (freshman, sophomore, …)

   b) time (days) and price (dollars)

   c) vehicle type (car, truck, …) and vehicle color (red, blue, …)

8. **[20 pts]** The plots below show Richter scale magnitude versus depth in kilometers for earthquakes (left) and tsunami (right). Identify **five** formatting problems and explain how each could be corrected.