Name_____ Date_____

LAB 3C: Random Sampling
Response Sheet

Directions: Record your responses to the **bold** lab questions in the spaces provided.

Learning by sampling

- **Write down two reasons why getting *everyone* in Los Angeles to fill out the survey would be difficult. Also, write a sentence why the DWP might consider using a sample of households instead.**

Loading a population

Load the `cdc` data in `R` and fill in the blanks to take a *convenience* sample of the first 50 people in the data: `s1 <- slice(___, 1:___)`

- **Why do you think we call this method a *convenience* sample?**

Comparing your convenience sample

Using your convenience sample, create a `bargraph` for the number of people in each `grade`.

- **Do you think the distribution of `grade` for your sample would look similar when compared to the whole `cdc` data?**

- **Which groups of people do you think are over or under represented in your convenience sample? Why?**

Create a `bargraph` for `grade` using the `cdc` data.

- **Compare the distributions of the `cdc` data and your convenience sample and write down how they differ.**

Name_____ Date_____

LAB 3C: Random Sampling
Response Sheet

Using randomness

Fill in the blanks below to create a sample by randomly selecting 50 people in the `cdc` data, without replacement.

- **Write a sentence that explains why you think the distribution of `grade` for this *random sample* will look more or less similar to the distribution from the whole `cdc` data.**

Increasing sample size

- **How do the distributions change as the size of the sample increases. Why do you think this occurs?**

`tally()` the proportion of `grades` for your *convenience* sample and all your *random* samples.

- **Which set of proportions looks most similar to the proportions of the population.**

Lessons learned

- **Write down a reason why estimates based on *convenience* samples might not improve even as sample size increases.**