Name:

Collaborators:

Instructions:

You must submit your worksheet individually by end-of-class or end-of-day. Your name must exist in your worksheet and the names of your collaborators.

Worksheets are marked mostly on completion, and partially on correctness. It will be marked either pass or fail, there will no detailed feedback on worksheets, and no opportunities for revisions and make-up.

Identify Types of Variables & Compute Descriptive Statistics

1. Air pollution and birth outcomes, study components.

Researchers collected data to examine the relationship between air pollutants and preterm births in Southern California. During the study air pollution levels were measured by air quality monitoring stations. Specifically, levels of carbon monoxide were recorded in parts per million, nitrogen dioxide and ozone in parts per hundred million, and coarse particulate matter (PM_{10}) in $\mu g/m^3$. Length of gestation data were collected on 143,196 births between the years 1989 and 1993, and air pollution exposure during gestation was calculated for each birth. The analysis suggested that increased ambient PM_{10} and, to a lesser degree, CO concentrations may be associated with the occurrence of preterm births.

a. Identify the main research question of the study.

b. Who are the subjects in this study, and how many are included?

c. What are the variables in the study? Identify each variable as numerical or categorical. If numerical, state whether the variable is discrete or continuous. If categorical, state whether the variable is ordinal.

2. Understanding Variable Types and Computing Descriptive Statistics

You are tasked with analyzing and comparing four data sets. Follow the steps below to identify the variable types, calculate descriptive statistics, and explain your reasoning.

Set	Data
X	$3, 5, 5, 5, 8, 11, 11, 11, 13$
Y	$3, 5, 5, 5, 8, 11, 11, 11, 20$
A	$cat, cat, cat, dog, dog, mouse, mouse$
B	$dog, mouse, cat, cat, cat, dog, cat$

a. Identify the type of variable and its sub type for each set.

b. Calculate descriptive statistics. For numerical data, compute the mean and median. For categorical data, calculate the frequency and proportion for each category.

c. Provide a short written explanation for your findings. Be sure to justify why the chosen descriptive statistics are appropriate for the given variable type.