How would you analyze the data in the following examples?

The answers in red suggest ways to analyze the data, but they are not the only right answers.

1. You wonder whether a new cycling incentive will make people want to ride their bike to work more. You survey a group of Stanford employees about their willingness (on a scale from 1 to 5) to bike to work before the incentives are put into place, and a week after the new incentives are advertised you survey another group of employees.

Independent-samples t test (Compare the mean willingness to bike to work before the incentives are put into place with the mean willingness after the incentives are put into place.)

2. You wonder whether the incentives for biking are even working. You count the number of people present on the Oval in a 10-minute period at 10am one week before the incentives are in place and one week after they are in place. Each time you note how many people are riding their bike.

Z test for the difference in proportions or a bivariate chi-square test looking at the relationship between time period (before/after) and biking (biking/not biking)

3. You want to find out the impact of temperature on violence. You obtain, for three consecutive years, daily outside temperature and daily reported violence incidents in New York City.

Correlation between temperature and violence or regression predicting violence from temperature (but beware of correlated errors)

4. You are told by a law enforcement consultant that violent events are also linked to the street price of crack cocaine. You want to test the link between temperature and violence independently of the price of drug (which you assume is independent from temperature).

Semi-partial correlation between violence controlling for drug price and temperature

5. You want to figure out the relationship between the price and sales of cookies. Specifically, you want to know whether cookies that cost more are sold less. However, when looking at the data, the relationship between these two variables does not appear to be linear.

Apply a transformation to improve linearity and then run a simple regression. Another option is to run a multiple regression with price and price^2 (centered) as predictors of sales.

6. You tested the life satisfaction of four groups: married women, married men, single men and single women. Half of your respondents in each group were asked how satisfied they were; half of your respondents were asked how dissatisfied they were.

2 (gender) x 2 (marital status) x 2 (question type) ANOVA (You would also call this a three-way ANOVA.)

7. You wonder how perceptions of parental strictness vary depending on who you ask. You ask 40 teenagers how strict their parents are. You then ask their mothers how strict they are.

Paired-samples t test to compare the mean strictness per group

8. You think children are more likely to remember words pronounced by their mother than by a stranger <u>because</u> they pay more attention to their mother. You have 2 groups of children in the lab, one hearing words from their mother, another hearing words from a stranger. For each child you record how many words they remember at test, as well as the proportion of time they spend looking at their mother at learning.

Use mediational analysis to test the full model. A partial correlation could tell you whether there is a relationship between condition (mother vs. stranger) and memory controlling for time spent looking.

9. You are interested in whether there is a relationship between whether a person has had their bikelight stolen and whether they would steal someone else's bikelight. You take ratings of willingness to steal a bikelight on a 7-point scale. The variance of the responses of people who have had their bikelights stolen in the past is much larger than the variance of the responses of people who have not had their bikelights stolen before.

You could try a transformation (natural log, square root) before running an independent-samples t test.

10. A relationship between how cold it is outside and how students score on a test has been shown. You hypothesize that when it is cold, people will stay inside more, so they will study for longer. You believe there is no direct link between temperature outside and the test score. You collect information about the temperature outside when the test was taken, the test score, and how much the person studied.

To determine that there is no direct link between temperature and test score, run a partial correlation between temperature and test score controlling for study time. To test the whole model, run a mediational analysis.

11. A study was run in which participants judged how much leniency they would show a person who had possibly committed a crime. Participants read about one defendant and saw a picture of this person. In the picture, the defendant was either smiling in one of three ways (false, felt, miserable) or he had a neutral expression. The question of interest is whether there is a relationship between smiling and leniency ratings. (See an old section handout for pictures of the facial expressions.)

    (a) Assume that the researcher had no hypothesis about which means would differ.

Run a one-way ANOVA and then run a posthoc test (like Tukey HSD).

(b) Suppose the researcher's main hypothesis was that the neutral condition would differ from the other three. The researcher was also interested in whether a miserable smile differed from the other two types of smiles and whether the false and felt smiles differed.

Look at contrasts (1 1 1 -3), (1 1 -2 0), and (1 -1 0 0).

(c) What if the researcher had the hypotheses of part b, but the method of the study was changed so that each participant saw a picture of all four facial expressions?

Run a repeated-measures analysis and then look at those contrasts.