1. 2 sample vs. paired
    a. *In .R file*
    b. In .R file
    c. In .R file
    d. Welch's sample

Welch Two Sample t-test

data: A and B

t = -1.3186, df = 17.105, p-value = 0.2047

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-1.8958707 0.4370999

sample estimates:

mean of x mean of y

0.1772973 0.9066827

Paired

One Sample t-test

data: diffs

t = -1.0797, df = 9, p-value = 0.3083

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

-2.2575069 0.7987361

sample estimates:

mean of x

-0.7293854

Both tests reach the same conclusion, that the null hypothesis can't be rejected. The test statistics are similar; the paired t-test indicates a less extreme observation, which is also reflected in the p-value. The point estimates are quite similar: a mean difference of ~0.7.

e. Welch's sample

Welch Two Sample t-test

data: A and C

t = -1.3424, df = 16.041, p-value = 0.1981

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-1.8272308 0.4101305

sample estimates:

mean of x mean of y

0.1772973 0.8858474

Paired

Paired t-test

data: A and C

t = -2.7164, df = 9, p-value = 0.02375

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-1.2986192 -0.1184811

sample estimates:

mean of the differences

-0.7085501

In this case, the tests disagree. The 2-sample t-test indicates that the null hypothesis can't be rejected, while the paired t-test indicates that we should reject the null hypothesis. The point estimates for the difference in means is roughly equivalent between the two, but the confidence interval of the paired t-test does not include zero.

f. Samples A and B are generated from the same distribution, with a population mean of 0. Sample B is then shifted with respect to sample A, but not sufficiently to result in a rejection of the null hypothesis.
Given that sample C is dependent on sample A, there exists a one-to-one mapping from observations in A to observations in C, and hence a paired t-test is appropriate. The

paired t-test results in narrower confidence intervals with respect to Welch's two-sample t-test, given the greater standard error. This narrower confidence interval excludes zero, allowing us to reject the null hypothesis of equivalent means.

2. Language transcriptionist
    a. The data is paired, as there are an equal number of observations in each sample and there is a one-to-one mapping from each observation in the first sample to another observation in the second sample. The source of variation would be the language proficiency, or fluency, in each English and French.
    b. Hypothesis test: The null hypothesis can't be rejected, and the transcriptionists mean transcription time can't be judged to be different between the two languages (paired t-test, df = 6, t = -1.63, p-value = 0.1537).

        Paired t-test

        data: q2_data[, "e"] and q2_data[, "f"]

        t = -1.6324, df = 6, p-value = 0.1537

        alternative hypothesis: true difference in means is not equal to 0

        95 percent confidence interval:

        -12.851666 2.565952

        sample estimates:

        mean of the differences

        -5.142857

    c. The transcriptionist can't be considered more or less proficient in either language, English of French, given the results (and conclusion drawn) above.
    d. There does appear to be a practical significance: the 95% confidence interval is heavily shifted towards English and the point estimate of the mean difference is over 5 minutes faster for English transcription.
3. Chemist
    a. No, this data does not appear to be paired. In this case, other than sequencing, there doesn't appear to be any clear one-to-one mapping from an observation in group A to an observation in group B. The independence of each observation ensures that the sequencing is irrelevant. Sources of variation would include the chemical reactions themselves, and potentially any (unintentional) variation in the environment, despite the chemist's attempts to the contrary.
    b. Welch's 2-sample t-test

        Welch Two Sample t-test

        data: q3_data[, "e"] and q2_data[, "f"]

t = 7.3953, df = 7.5082, p-value = 0.0001058

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

242.5519 466.0552

sample estimates:

mean of x mean of y

391.87500 37.57143

c. There is strong evidence that the mean amount of residue of chemical reactions A and B are not the same (Welch's 2-sample t-test, df = 7.5, p-value = 1.06e-4). It is estimated that the mean residue remaining from reaction A is 391.9 ug and from reaction B 37.6 ug. With 95% confidence it is estimated that the mean residue of A is 242.6 to 466.1 ug higher than that of reaction B.