Agenda

Test corrections rtuffahl

1) correct answer: d, c
2) correct answer: b, d
5) correct answer: d, c

6) [choose - explore payments made vs. APR @ time of payment. I would make APR the x variable as it changes over time, and payments the y because we are examining its dependence on x.
[fit - find estimate for the slope and intercept. data can be plugged into R using an lm function which evaluates the SLR equation $Y = \beta_0 + \beta_1 X + \varepsilon$. in this case, the intercept would be β_0 and the slope $\beta_1 x$.
[assess - after calculating slope & intercept, I would create a histogram, followed by a residuals vs. fitted plot and a normal quantile plot. these are to assess the linearity, constant variance, normality, and the zero mean. I would also check the p-values & R^2 to see if they fit with my original hypotheses.
[answer - I would create a confidence interval for the slope based on the observed data, plots, and calculations. from there I would make decisions about credit card spending.

Agenda

7) based on the data in this image, we can understand that $SS_{mod} \approx SSE$. since the total SSE = the SSE of the model and the SSE added, when the R^2 value is as low as 0.06, this shows that the SS_{Total} must be comprised of low values added together. therefore, SS_{mod} and SSE must not be very far apart numerically in order for neither of them to equal zero and for both of them to add up to 0.06.

8) ⓐ we need to check for the normality, constant variance, and linearity of the model. based on the normal probability plot, the distribution is mostly normal - no strong skewing. however, though there is a zero mean, the variance is not constant - as it gets further from the mean, the variance increases. the plots are pretty linear, based on the n.p.p. though we cannot infer independence or randomness from these plots, we can investigate how the data was collected to learn more about these conditions.

Agenda

9) ⓒ two sided t-test (comparing mean/avg burning of runner vs others in age group)

I would use a two sided t test. ~~to~~ to find a confidence interval at 95%, meaning the CI would have a 95% chance of including the difference in means. after checking conditions to make sure the data was normal, i would note the H_0 = there is no difference between the runner's avg & the others' avg ($\mu_1 = \mu_2$), and $H_a = \mu_1 \neq \mu_2$. from there i would complete the t-test and find the CI. ~~in the chart it appears to~~