The anti_htn data set comes from a parallel group, randomized, double-blind, placebo-controlled phase II clinical trial. The aim of the trial was to determine the effect of a new antihypertensive medication on systolic blood pressure (SBP) in patients with hypertension (i.e., high blood pressure). In this small trial, 46 patients were randomized to receive the experimental medication or a placebo in 1:1 manner. The two treatments were identical in appearance and taste. Three study centers were involved in the trial. Randomization was stratified by study center. Patients took their randomized medication for one year. The primary efficacy outcome of the trial on which treatments were compared was 'change in systolic blood pressure from baseline to 12 months after baseline'. It is a continuous variable. A negative value indicates a decrease in blood pressure from baseline.

Variable	Description
trt	Randomized Treatment (a character variable, entered as "Denerv2", and "Placebo", which stand for Denerv (the name of the treatment) dose 2 and placebo, respectively)
pt	Patient ID (a numeric variable that you can simply ignore)
female	Gender (0 = Male, 1=Female)
site	Study Center ID (there are 3 study centers; the numeric IDs given to the sites to uniquely identify them are 1, 10 and 50)
sbp_base	Baseline SBP (the last SBP measured before randomized treatment started)
sbp_reduction_12m	Change in SBP from baseline to 12 months (a negative value indicates a decrease from baseline)
htn	A dichotomous efficacy outcome indicating whether patient had hypertension at end of study (0=No, 1=Yes). (ALL patients had hypertension at START of the study so this dichotomous outcome helps assess whether the experimental treatment is efficacious at study end)

The study is considered a success if the experimental treatment has a significantly larger decrease than the placebo with respect to the outcome 'change in SBP from baseline to 12 months. State the two sided null and alternative hypothesis of interest.

H_0: $\mu_A = \mu_P$. The change in SBP is the same for both treatment and placebo groups.
H_A: $\mu_A \neq \mu_P$. The change in SBP is not the same in treatment and placebo groups.

Provide appropriate descriptive statistics for site, sex, baseline SBP, change in SBP from baseline to month 12, and hypertension for the overall sample and by treatment group in a table. Please do NOT carry any formal statistical treatment comparison (i.e., no p-values and no confidence intervals) for this question. Present your results in a table using 1 decimal place.

Demographics		**Treatment Group** (n=22)	**Placebo Group** (n=24)	**Total** (n=46)
Sex no. (%)	Male	13 (59.1%)	15 (64.5%)	28 (60.9%)
	Female	9 (41.0%)	9 (37.5%)	18 (39.1%)
Site no. (%)	1	3 (13.6%)	7 (29.2%)	10 (21.7%)
	10	7 (31.8%)	3 (12.5%)	10 (21.7%)
	50	12 (54.6%)	14 (58.3%)	26 (56.5%)
Hypertension no. (%)	No	15 (68.2%)	11 (45.8%)	26 (56.5%)
	Yes	7 (31.8%)	13 (54.2%)	20 (43.5%)
Base SBP mean (stdv)		173.2 (17.6)	175.7 (18.5)	174.5 (17.9)
SBP Change mean (stdv)		-29.8 (15.6)	-16.8 (18.5)	-23.0 (18.2)

Perform an appropriate formal statistical treatment group comparison on the mean of the primary outcome (change in SBP between at 12 months) and test the null hypothesis in question 1. Please adjust for site, the randomization stratification factor. You can assume the variance of the primary outcome is equal between treatment groups. When writing your results, report and interpret your effect estimate reporting 95% confidence intervals and p-value. Explain how the conclusions from the confidence intervals match the conclusions from your formal test of the hypothesis.

An ANOVA test was used to test whether the change in a subject's systolic blood pressure was linearly associated with treatment group and study center. The F-statistic was 2.42 and resulting p-value was 0.0793. With a p-value greater than the α=0.05 significance level, the null hypothesis of there being no linear association between change in systolic blood pressure and the predictor variables was not rejected. There is insufficient evidence to conclude a linear association between a change in systolic blood pressure and either treatment group or study center.

Parameter		**Estimate**	**p-value**	**95% Confidence Interval**	
				Upper Limit	**Lower Limit**
	Intercept	-12.69			
Treatment Group	Treatment	-12.35	0.0252	-23.09	-1.61
	Placebo				
Center	Site 10	-4.26	0.5999	-20.52	12.00
	Site 50	-6.18	0.3487	-19.35	6.98
	Site 1				

A multiple linear regression analysis was used to test whether the change in systolic blood pressure was the same for both treatment and placebo groups, adjusting for study center. The F-statistic was 5.38 and resulting p-value was 0.0252. With a p-value less than the α=0.05 significance level, the null hypothesis of there being no difference in systolic blood pressure change between treatment groups was rejected. There is evidence suggesting that among participants in the same study center, those in the treatment group had a 12.35 mmHg greater decrease in systolic blood pressure than those in the placebo group. The confidence interval for the outcome variable does not include 0, which indicates a statistically significant difference between the two groups, which is the same conclusion from the multiple linear regression analysis.

Repeat the previous analysis but now adjust for baseline SBP. As in question 3, report and interpret your treatment effect estimate. You should gain more significance than in the analysis in question 3. Inspect any relevant data to explain why this is.

An ANCOVA test was used to test whether the change in a subject's systolic blood pressure was linearly associated with treatment group, study center, and baseline systolic blood pressure. The F-statistic was 3.49 and resulting p-value was 0.0153. With a p-value less than the α=0.05 significance level, the null hypothesis of there being no linear association between change in systolic blood pressure and the predictor variables was rejected.

Parameter		Estimate	p-value	95% Confidence Interval	
				Upper Limit	Lower Limit
Intercept		45.06			
Treatment Group	Treatment	-13.33	0.0117	-23.54	-3.12
	Placebo				
Center	Site 10	-3.27	0.6711	-18.70	12.16
	Site 50	-5.27	0.3997	-17.76	7.23
	Site 1				
baseline SBP		-0.33	0.0200	-0.61	-0.06

A multiple linear regression analysis was used to test whether the change in systolic blood pressure was the same for both treatment and placebo groups, adjusting for study center and baseline systolic blood pressure. The F-statistic was 6.96 and resulting p-value was 0.0117. There is evidence suggesting that among participants in the same study center and with the same baseline systolic blood pressure, those who received the experimental treatment had a 13.33 greater decrease in systolic blood pressure than those in the placebo group. The confidence interval for the outcome variable does not include 0, which indicates a statistically significant difference between the two groups, which is the same conclusion from the multiple linear regression analysis.

The increase in significance from the ANOVA test is due to baseline SBP and reduction in SBP having a Pearson correlation coefficient of -0.3079, indicating a moderate correlation between the two variables. Adjusting for baseline covariates that are found to be correlated with the outcome variable of interest gives more precise treatment effect estimates.

Using the SAS output from question 4, write down the corresponding linear model with the parameter estimates. Interpret the intercept, the parameters for center and the parameter for baseline SBP.

$$\hat{Y} = 45.06 - 13.33trt - 3.27site10 - 5.27site50 - 0.33base\_sbp$$

The intercept, 45.06, is the average decrease in SBP for those in the placebo group at site 1 when the baseline SBP is 0.

The treatment effect, -13.33, is the difference in SBP decrease between treatment and placebo groups for those at the same study center and same baseline SBP.

The parameter for Site 10, -3.27, is the difference in SBP decrease between those at Site 10 and those at Site 1 for subjects in the same treatment group with the same baseline SBP.

The parameter for Site 50, -5.27, is the difference in SBP decrease between those at Site 50 and those at Site 1 for subjects in the same treatment group with the same baseline SBP.

The parameter for baseline SBP, -0.33, is the difference in SBP decrease for every 1 mmHg increase in baseline SBP for those in the same treatment group and study center.

Given the results your compiled above, is the study "successful"?

The study can be considered successful because the experimental treatment decreased SBP by an additional 13.33 mmHg than the placebo did, adjusting for study center and baseline SBP.

A secondary endpoint of this study is the binary outcome 'hypertension by the end of the study at month 12. The sponsor wants to show that the experimental mediation has lower percentage of individuals with hypertension than the placebo. State the two-sided null and alternative hypothesis of interest for this secondary endpoint.

H_0: $\pi_A = \pi_P$. $OR = 1$. The odds of having hypertension at the end of the study is the same in both treatment groups.

H_A: $\pi_A \neq \pi_P$. $OR \neq 1$. The odds of having hypertension at the end of the study is not the same in both treatment groups.

Use logistic regression to test the null hypothesis and present odds ratios for hypertension for the experimental treatment versus placebo with confidence intervals, while adjusting for site, the randomization stratification factor. Given these results, is the study "successful" on this secondary outcome?

A multiple logistic regression analysis was used to test whether having hypertension at the end of the study was linearly associated on the natural logarithmic scale with treatment group and study center. From the likelihood ratio test, the chi-squared statistic was 4.5698 with 3 degrees of freedom and the resulting p-value was 0.2062. With a p-value greater than the $\alpha=0.05$ significance level, the null hypothesis of there being no linear association between having hypertension and the two regressors was not rejected. There is insufficient evidence suggesting that there is a linear association on the natural logarithmic scale between having hypertension at the end of the study and the two predictor variables, treatment group and study center.

A Wald test was used to test whether having hypertension at the end of the study was linearly associated on the logarithmic scale with treatment group, adjusting for study center. The Wald chi-squared statistic was 2.3444 with 1 degree of freedom and the resulting p-value was 0.1257. With a p-value greater than the $\alpha=0.05$ significance level, the null hypothesis of there being no linear association on the natural logarithmic scale between having hypertension at the end of the study and treatment group, adjusting for study center was not rejected. There is insufficient evidence the suggest the experimental drug is more efficient at reducing the percentage of individuals with hypertension. The study cannot be considered a success on this outcome.

A colleague suggests to further adjust for baseline SBP, because baseline SBP is likely to be correlated with hypertension at the end of the study. Without running any further model, please explain if and why you agree or disagree with the suggestion.

If it is known that baseline SBP is correlated with hypertension, then it should be included in the final model. Adjusting for covariates that are correlated with the outcome of interest will give a less precise treatment effect estimate.