

**STAT 1124**  
**Project**  
**Relating Alcohol Consumption to Blood Alcohol Content**

**Instructions:**

- It is recommended that you do this project in pairs. If you do so please hand in only one paper, with both your names, student no., and section number on it.
- Please hand in only 1 file in PDF format.

Two universities conducted experiments to investigate the relationship between Blood Alcohol Content (BAC) and amount of alcohol consumed, as well as other factors such as gender, weight, and age.

**At Ohio State University, USA**, 16 students volunteered for the experiment. Each student blew into a breathalyzer to indicate that his/her initial BAC was zero. They then drew tickets at random from a bowl. Each ticket showed a number, between 1 and 9, indicating the number of glasses of beer that they had to drink. Thirty minutes after their last drink, their BAC was measured again. A police officer also administered a sobriety test to each volunteer before and after consuming alcohol. The test consisted of coordination demonstration, such as balancing on one foot, touching one's nose with the tip of the forefinger, etc. Test score is from 1 to 10, 10 being the best. The police officer was not aware of how much each subject had drunk.

**At the University of Western Sidney, Australia**, 22 staff and students volunteered for the experiment. Like the OSU experiment, all subjects were also tested before the start to make sure that their BAC was zero. White wine was offered in 120 ml glasses; each subject drank as many as he or she wished. Snack food and water was available at all times. After 45 minutes the subjects were told to stop drinking. After an additional 15 minutes each had his/her BAC measured again, and the number of glasses of wine drunk recorded.

The data from the studies have been prepared for analysis in StatGraphics. You may download them from the course page. The file name is BAC\_OSU.sgd (If the file fails to load in SG Stratus, try loading the Excel version: BAC\_OSU.xlsx)

The variables are as follows.

Gender	= (male, female)
Weight	= weight (in Kg)
Beers	= number of 12-ounce beers consumed (OSU data only)
BAC_OSU	= blood alcohol content (in gr. of ethanol per 100 ml of blood.)
AWR	= alcohol consumption per kg of body weight $\left( = \frac{Beer}{Weight} \right)$
Sobriety1	= score on road sobriety test before consuming alcohol (OSU data only)
Sobriety2	= score on road sobriety test after consuming alcohol (OSU data only)
Sobriety_Diff	= Sobriety1 – Sobriety2

Note, data from the Australian study have been excluded due to file size limitation in SG Stratus.

1. Both studies constituted an experiment, since treatments (drink alcohol) was imposed on subjects in both cases. There was a difference, however, in that in OSU, the treatment was randomized by getting subjects to draw tickets from bowl to determine how many beers they were to drink. In AUS each subject chose for him/herself how many glasses of wine to drink. Discuss the practical and scientific advantage and disadvantage of randomization in these cases. [4 marks]
2. The other difference between the two was that in the AUS case, the experiment was conducted in a 'party atmosphere,' with snacks and water available at all time. Was this an advantage or disadvantage? Briefly explain your answer. [3 marks]
3. In the OSU experiment, when the police officer administered the second sobriety test, he/she was prevented from knowing how many beers the subject had consumed. Why was this done? [2 marks]
4. Regress BAC\_OSU against number of beers drunk for the OSU study.

SG Centurion:

Relate > One Factor > Simple Regression. Place BAC\_OSU as the response (Y) variable, and Beers as the explanatory (X) variable. Under Regression Options, make sure that Type of Model is set to Linear, with "Include constant" and "no alternative fit" checked. Under Tables and Graphs, select Analysis Summary, Plot of Fitted Model, and Residuals vs. X.

SG Stratus:

Relate > Simple Regression. Place Place BAC\_OSU as the response (Y) variable, and Beers as the explanatory (X) variable. Click Output.

Write down the correlation coefficient between these two variables. But do not include the rest of the output from this analysis in your report. [1 mark]

5. The same number of drinks has a greater impact on a smaller person than on a larger person. To account for this difference, we should enter weight as part of the equation. The variable AWR (alcohol to weight ratio) is defined as the number of beers drunk per kg. of body weight. Now let's regress BAC on AWR instead of Beer.

SG Centurion:

Return to the regression analysis pane. Click the *Input Dialog* icon (the left-most on top of the analysis pane.) Change the X variable from Beers to AWR. Right-click the scatter plot, under *Pane Options* deselect Prediction and Confidence Limits. This will get rid of the curves around the regression line that we do not use. Click on the Residual Plot, under Pane Options, select Residuals instead of Studentized Residuals.

SG Stratus:

Click Data Input. Change the Y and X variables accordingly. Under Tables & Graphs make the following selections:

Tables/Graphs All Clear		Options										
<input checked="" type="checkbox"/> Analysis Summary												
<input checked="" type="checkbox"/> Plot of Fitted Model Titles and Scaling		Plot: <input checked="" type="radio"/> All models <input type="radio"/> Least squares fit only <input type="radio"/> Alternative fit only X-axis resolution: 101 Include: <input type="checkbox"/> Prediction limits <input type="checkbox"/> Confidence limits Confidence level: 95 % <input checked="" type="radio"/> Two-sided interval <input type="radio"/> Lower bound <input type="radio"/> Upper bound										
<input type="checkbox"/> Forecasts		Forecast at X: <table border="1"> <tr><td></td><td></td></tr> <tr><td></td><td></td></tr> <tr><td></td><td></td></tr> <tr><td></td><td></td></tr> <tr><td></td><td></td></tr> </table> Type of limits: <input checked="" type="radio"/> Two-sided interval <input type="radio"/> Lower bound <input type="radio"/> Upper bound Confidence level: 95 %										
<input type="checkbox"/> Observed versus Predicted Titles and Scaling												
<input type="checkbox"/> Lack-of-Fit Test												
<input type="checkbox"/> Comparison of Alternative Models												
<input type="checkbox"/> Unusual Residuals		<input checked="" type="radio"/> Residuals <input type="radio"/> Studentized residuals <input type="radio"/> Residuals from alternative model Smoother: <input checked="" type="radio"/> None <input type="radio"/> Running means <input type="radio"/> Running lines <input type="radio"/> LOWESS <input type="radio"/> Robust LOWESS Smoothing fraction: 50 %										
<input type="checkbox"/> Residuals versus Predicted Titles and Scaling		<input type="radio"/> Residuals <input checked="" type="radio"/> Studentized residuals <input type="radio"/> Residuals from alternative model Smoother: <input checked="" type="radio"/> None <input type="radio"/> Running means <input type="radio"/> Running lines <input type="radio"/> LOWESS <input type="radio"/> Robust LOWESS Smoothing fraction: 50 %										
<input type="checkbox"/> Residuals versus Row Number Titles and Scaling		<input type="radio"/> Residuals <input checked="" type="radio"/> Studentized residuals <input type="radio"/> Residuals from alternative model Smoother: <input checked="" type="radio"/> None <input type="radio"/> Running means <input type="radio"/> Running lines <input type="radio"/> LOWESS <input type="radio"/> Robust LOWESS Smoothing fraction: 50 %										
<input type="checkbox"/> Influential Points												

Include the text output, the Scatter Plot, and the Residual Plot from this analysis. [3 marks]

6. Predict the BAC of a 150 lbs person after he/she has had 5 beers. Use an interval such that your prediction will be correct 95% of the time. Note that SG reports slope of the regression line ( $b_1$ ) as estimate of the slope,  $b_0$  as estimate of the intercept. These two numbers are located in Estimate column. The RMSE is reported as “Standard Error of Est.” To minimize round-off error carry as many significant digits as possible in your calculation. [4 marks]

7. In BC, it is illegal to drive with a BAC of more than 0.08. Assuming all regression assumptions are satisfied, calculate the chance that a 150-lb person who drinks 4 beers exceeds the legal limit. Assume all regression assumptions are met. [4 marks]

8. Can we generalize the results of this experiment to all adults? What about to all university-age adults? Justify your answer. [3 marks]

9. Before hand-held breathalyzer became available, the police gauged a person’s level of impairment based on the sobriety test. How well does the test agree with BAC? Regress the Blood Alcohol Content on the difference in sobriety score. Use the same commands as before, but now use BAC\_OSU as the dependent variable, and Sobriety\_Diff as the independent variable. Include the Scatter Plot and the correlation coefficient (and nothing else in your report). [2 marks]

10. Based on the data determine whether there is a linear correlation between BAC and the decrease in sobriety score. Include a decision point in your report. [3 marks]
11. Identify the outlier in the lower right hand corner. Is this a male or female, how much does he/she weigh, how many beers he/she drank, etc? Why is he/she an outlier? Hint: go to the Databook in SG Centurion, or Display Data in SG Stratus. [4 marks]
12. The police's main concern is to keep the streets safe. As far as determining whether a person is too impaired to drive safely, which variable do you think they should base their judgment on, BAC or sobriety score? Discuss the logical as well the practical limitation of each. [4 marks.]