# 307 Oral Exam script

Say hi and welcome them. Perhaps make a small joke to try and get them somewhat comfortable.

## Describe process, ask if they have any questions before you get started

Share the following bullet points as a document.

* Ask them if they can see your screen
* Note that the format is the same as the first oral exams.
* Ask if they have any questions?

## Share document and ask questions

Switch to sharing the .sas file within SAS Studio (this should show line numbers, also make sure your font is large and your window is mostly taken up by the .sas file) Double check that they can see the document.

Ok, now I'm going to ask you some questions about the program you see.

### Question 1 (2 parts):

#### Option 1

* Consider the DATA step on lines 11 through 21. On line 12 we use a KEEP dataset option. First, what is the purpose of using a KEEP dataset option? (4 pts)
  + *Answer* The KEEP dataset option allows us to KEEP the variables color, spine, width, weight, and y.
  + Not required, but just so we’re aware: our final dataset has different variables, because we’re adding shell, renaming y, and dropping spine.
* We also looked at using a KEEP statement in a data step. What is the big difference between using a KEEP statement and a KEEP dataset option? (6 pts)
  + *Answer* Using KEEP in a dataset option changes the data that’s being copied in and worked with during the data step. Using KEEP as a statement makes changes after the data has been copied.
  + Practically: when you use KEEP as a dataset option, it happens before the other statements.

#### Option 2

* Consider the PROC MEANS step on lines 27 through 30. On line 28 we use a CLASS statement. First, what is the purpose of using a CLASS statement on a PROC MEANS step? (4 pts)
  + *Answer* summarize data by the group defined by CLASS
* We also looked at using a BY statement with PROC MEANS to do a similar creation of summary statistics. What additional step was needed prior to being able to use a BY statement? (6 pts)
  + *Answer* **sort** the data by the categorical variable

### Question 2 (4 pts):

#### Option 1

* On lines 23-25 we run a PROC FREQ step. On line 24, what is the purpose of putting the star or asterisk between **shell** and satellite?
  + *Answer* Create a 2-way contingency table between shell and satellite

#### Option 2

* On lines 43-46 we run a PROC SGPLOT step. On line 44, what does the VBOX statement do?
  + *Answer* Create one vertical boxplot for each category of satellite, displaying summary statistics of width.

### Question 3 (4 pts)

#### Option 1

* Consider the DATA step on lines 11 through 21.On line 19 we have the code IF spine NE 1; What is the purpose of this code? (follow up on filter rows or columns)
  + *Answer* if the observation/row has the spine not equal to 1, we keep the observation; otherwise, we drop the observation.

#### Option 2

* Consider the DATA step on lines 11 through 21. On lines 15 through 18 we use IF ELSE logic. What is the purpose of this code?
  + *Answer* creates a new variable called Shell based on the value of color. If color = 2, Shell = light, if color = 3, shell = medium, if color = 4, shell = dark, if other values of color, shell = darker.

#### Option 3

* On lines 48-52 we run a PROC SGPANEL step. What is the purpose of running a PROC SGPANEL step?
  + *Answer* PROC SGPANEL creates one plot for every level of the categorical variable specified in the PANELBY statement.

#### Option 4

* On **lines 36-41 we run a PROC SGPLOT** step. On line 37, what is the purpose of the WHERE statement?
  + *Answer* select only observations with weight < 3300 for visualization, other observations are not visualized.