Provide your answers and the code used for each question in the text boxes below. The boxes are approximately sized to match the amount of space you will need for each answer, but you are welcome to expand or shrink the text boxes as needed. You may submit answers as text or screenshots. Note: The text has been set to blue inside the text boxes. This is intentional and will make it easier for the TA’s to see your answers.

For this Homework, we will be using the Age-Religion dataset that we used in the lab.

1. Create a correlation matrix for the following variables in the data set:
   1. Depression\_04
   2. Satisfied\_04
   3. SelfWorth\_04
   4. FearDeath\_04

Attach a screenshot of the correlation matrix and include the code below. **[1 point]**

A screenshot of a graph

Description automatically generated

cor\_q1 <- data |>

dplyr::select(Depression\_04, Satisfied\_04, SelfWorth\_04, FearDeath\_04) |>

psych::corr.test()

cor.plot(cor\_q1$r)

1. On the correlation matrix associated with Question 1, mark/bold/color any correlation coefficient that was significant (*p* < 0.05). You can copy the matrix here and highlight the values. **[1 point]**

Depression\_04 Satisfied\_04 SelfWorth\_04 FearDeath\_04

Depression\_04 **1.00** **0.36 0.28** 0.04

Satisfied\_04 **0.36** **1.00 0.39 0.14**

SelfWorth\_04 **0.28** **0.39 1.00 0.52**

FearDeath\_04 0.04 **0.14 0.52 1.00**

cor\_q1 <- data |>

dplyr::select(Depression\_04, Satisfied\_04, SelfWorth\_04, FearDeath\_04) |>

psych::corr.test()

1. Which two variables were most strongly (and significantly) correlated? **[1 point]**

SelfWorth\_04 and FearDeath\_04

1. Create three scatter plots where FearDeath\_04 is on the y-axis (in all three plots) and the x-axis is i. Depression\_04, ii. Satisfied\_04, iii. SelfWorth\_04. Make sure the graphs are clean, properly labeled, and that the y-axis range and scaling is consistent across graphs. Include a line of best fit on each graph. Put the three graphs in a panel. **[3 points]**

**Extra Credit [1 point]:** Make the line green if the association was significant or red if the association was not significant.

A graph of different points

Description automatically generated with medium confidence

long\_data <- data |>

dplyr::select(Depression\_04, Satisfied\_04, SelfWorth\_04, FearDeath\_04) |>

tidyr::pivot\_longer(cols = -FearDeath\_04,

names\_to = "variable",

values\_to = "value")

long\_data\_plot <- long\_data |>

tidyr::nest(data = -variable) |>

dplyr::mutate(

fit = purrr::map(data, ~lm(FearDeath\_04 ~ value, data=.) |> broom::tidy())

) |>

tidyr::unnest(fit) |>

dplyr::filter(term != "(Intercept)") |>

tidyr::unnest(data)

long\_data\_plot |>

ggplot(aes(x = value, y = FearDeath\_04)) +

geom\_point(alpha = 0.5) +

facet\_wrap(~variable) +

geom\_smooth(method = "lm", fullrange = T) +

labs(x = NULL) +

xlim(0, 40) +

theme\_bw()

1. Fit a multiple regression model where you predict the frequency of FearDeath\_04 with Depression\_04, SelfWorth\_04 and Satisfied\_04 variables. Write the equation of the estimated regression model. **[1 point]**
2. From the equation above, interpret the effect on FearDeath\_04 for a one-unit increase in each of the three predictors. Also mention if they are significant or not. **[2 point]**

A one-unit increase in Depression\_04 is associated with a decrease of 0.069 in FearDeath\_04 when controlling for Satisfied\_04 and SelfWorth\_04. This effect is statistically significant. Similarly, a one-unit increase in Satisfied\_04 is associated with a decrease of 0.056 in FearDeath\_04 when controlling for Depression\_04 and SelfWorth\_04, and this effect is also statistically significant. Finally, a one-unit increase in SelfWorth\_04 is associated with an increase of 1.38 in FearDeath\_04, when controlling for Depression\_04 and Satisfied\_04 and this effect is also statistically significant.

1. What percent of the variation in FearDeath\_04 is accounted for by the three varibles Depression\_04, Satisfied\_04 and SelfWorth\_04? **[1 point]**

28.9%