

# Homework 1

STAT 425

Due Sep 6, 2024

## Measurements of ship size, capacity, crew, and age for 158 cruise ships.

Use the following code to read a fixed width formatted data file on cruise ship characteristics:

```
winner.url = "https://users.stat.ufl.edu/~winner/data/"
tmp.url = paste0(winner.url, "cruise_ship.dat")
st = c(1,21,41,49,57,65,73,81,89)
en = c(20,40,48,56,64,72,80,88,97)
wds = diff(c(st, en[length(en)]))
cship = read.fwf(tmp.url, widths = wds, strip.white = T)
colnames(cship)=c("ShipName", "CruiseLine", "Age", "Tonnage", "passengers",
                  "Length", "Cabins", "pdensity", "crew")
```

The data includes the following variables:

Variables/Columns

Ship Name 1-20

Cruise Line 21-40

Age (as of 2013) 46-48

Tonnage (1000s of tons) 50-56

passengers (100s) 58-64

Length (100s of feet) 66-72

Cabins (100s) 74-80

Passenger Density 82-88

Crew (100s) 90-96

- 1.(2 pts) Use the `summary` function to compute summary statistics of the the number of passengers and length of the ship. Be sure to interpret the output.
- 2.(2 pts) Plot the relationship between the length of the ship (x-axis) and the number of passengers (y-axis) (include appropriate axis labels). Describe the relationship between the variables.
3. (2 pts) Use the `cor` function to compute the correlation between the number of passengers and length of the ship.
4. (2 pts) Fit the least squares regression line with number of passengers as the response and length of the ship as the explanatory variable. Interpret the slope parameter.
5. (2 pts) Estimate the mean number of passengers for ships that are 800 feet long.
6. (2 pts) What is the estimate of the variance of the error  $e_i$ ?
7. (2 pts) Explain whether or not the assumption of constant error variance is reasonable.
8. (2 pts) What is the coefficient of determination  $R^2$ ?

**L<sup>A</sup>T<sub>E</sub>X**

9. (2 pts) Use L<sup>A</sup>T<sub>E</sub>X to write the equation,  $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + e_i$ .
10. (2 pts) Find the derivative of  $E[y]$  for the model in (9) with respect to  $x$ . (Use L<sup>A</sup>T<sub>E</sub>X to write your answer).