

Name _____

Fall 2021

STAT 587-2

Final exam
(50 pts)

Instructions:

1. Full credit will be given only if you show your work.
2. The questions are not necessarily ordered from easiest to hardest.
3. You are allowed to use any resource except aid from another individual.
4. Aid from another individual will automatically earn you a 0.

1. Use the scatterplots on the page titled “Scatterplots” to answer the following questions. We will consider 3 possible regression models:

- no transformation needed (x - y)
- log explanatory variable only ($\log(x)$ - y)
- log response variable only (x - $\log(y)$)

For each plot, indicate which of the following models will fit the data best using the code in parentheses, e.g. x - $\log(y)$. If none of these models will fit the data well, then write “none”. (1 pt each)

• a

• b

• c

• d

• e

• f

• g

• h

• i

• j

2. Use the code on the page titled “Food’s Carbon Footprint” to answer the following questions.

(a) How many observations are in the data set? (1 pt)

(b) How many levels of food category are in the data set? (1 pt)

(c) What is the reference level for food category? (1 pt)

(d) If the design is balanced (with respect to food categories), how many countries are represented in the data set? (1 pt)

(e) What is the estimate for the residual standard deviation? (1 pt)

(f) Provide a point estimate for the CO₂ emissions for Fish when consumption is 10 kg/person/year. Show your work for full credit. (5 pts)

3. Use the data file `SRM_1540e_Fibrous_Glass_Board.csv` to answer the following questions. In this dataset, we evaluate the thermal conductivity (W/mK) of SRM 1540e Fibrous Glass Board at various temperatures (K) and air pressure (kPa).

(a) What values of temperature exist in the experiment? (2 pts)

(b) Is the design of the experiment complete? Why or why not? (2 pts)

(c) Is the design of the experiment balanced? Why or why not? (2 pts)

(d) Is this a replicated experiment? Why or why not? (2 pts)

(e) What is the correlation between temperature and air pressure in these data? Explain why this is the correlation. (2 pts)

Important: Throughout this question, treat temperature and air pressure as categorical (rather than continuous) variables.

(f) Conduct an F-test to determine whether an interaction between temperature and air pressure is needed.

- F-statistic (1 pt)

- p-value (1 pt)

- conclusion (1 pt)

(g) With the **additive model**, provide a prediction with uncertainty for the thermal conductivity at temperature of 320 K and air pressure of 100 kPa.

- point estimate (1 pt)

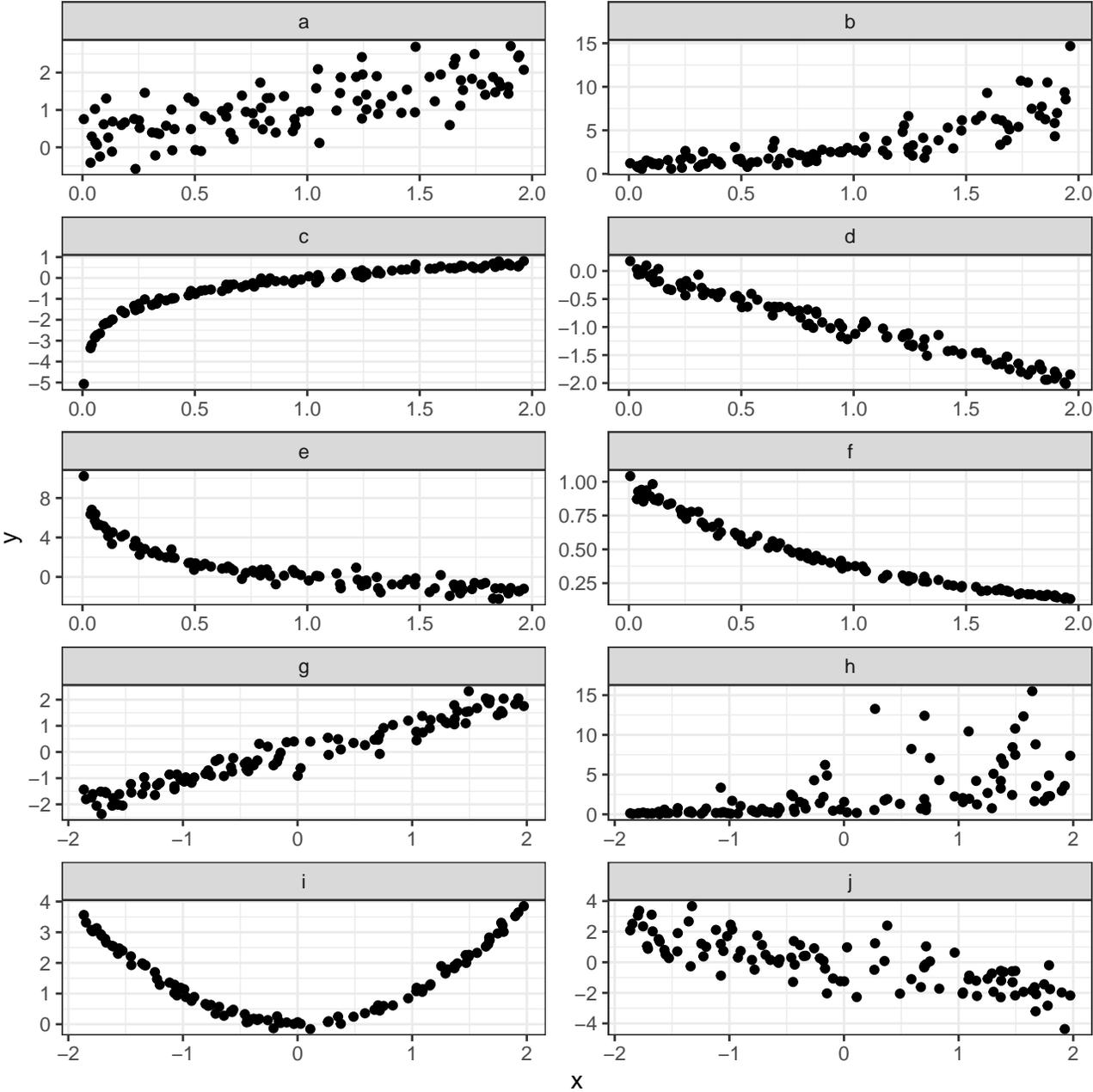
- 90% Prediction interval (2 pts)

(h) With the **additive model**, calculate a contrast to compare the difference in thermal conductivity between temperatures 360 and 280 averaged over air pressure.

- point estimate (2 pts)

- 99% Confidence/credible interval (2 pts)

Scatterplots



Food's Carbon Footprint

The following regression model analyzes the relationship between carbon dioxide emissions (kg CO₂/person/year) and food consumption (kg/person/year) across a large number of countries and food categories.

```
table(food_consumption$category)
##
## Beef Eggs Fish Pork Poultry
## 130 130 130 130 130
m <- lm(co2_emission ~ I(consumption-10) * category, data = food_consumption)

summary(m)
##
## Call:
## lm(formula = co2_emission ~ I(consumption - 10) * category, data = food_consumption)
##
## Residuals:
## Min 1Q Median 3Q Max
## -0.0053348 -0.0024859 -0.0003105 0.0024896 0.0053201
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.086e+02 2.616e-04 1179562 <2e-16 ***
## I(consumption - 10) 3.086e+01 2.656e-05 1161701 <2e-16 ***
## categoryEggs -2.994e+02 3.774e-04 -793256 <2e-16 ***
## categoryFish -2.926e+02 3.786e-04 -772841 <2e-16 ***
## categoryPork -2.732e+02 3.790e-04 -720716 <2e-16 ***
## categoryPoultry -2.978e+02 4.155e-04 -716745 <2e-16 ***
## I(consumption - 10):categoryEggs -2.994e+01 5.746e-05 -521012 <2e-16 ***
## I(consumption - 10):categoryFish -2.926e+01 2.979e-05 -982228 <2e-16 ***
## I(consumption - 10):categoryPork -2.732e+01 3.117e-05 -876342 <2e-16 ***
## I(consumption - 10):categoryPoultry -2.978e+01 3.186e-05 -934733 <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002913 on 640 degrees of freedom
## Multiple R-squared: 1, Adjusted R-squared: 1
## F-statistic: 3.219e+11 on 9 and 640 DF, p-value: < 2.2e-16

anova(m)
## Analysis of Variance Table
##
## Response: co2_emission
## Df Sum Sq Mean Sq F value Pr(>F)
## I(consumption - 10) 1 1717832 1717832 202497070305 < 2.2e-16 ***
## category 4 13733627 3433407 404728154460 < 2.2e-16 ***
## I(consumption - 10):category 4 9121848 2280462 268819640074 < 2.2e-16 ***
## Residuals 640 0 0
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```