

HW 13 – Math 58B

LAST HOMEWORK! due Thursday, April 30, 2026

your name here

```
library(tidyverse)
library(mosaic)
library(infer)
library(broom) # includes the tidy() function
library(praise)
```

Assignment Summary (Goals)

- Multiple Regression
 - choosing variables
 - confidence intervals for mean response
 - prediction intervals for individual response
 - R^2 and adjusted R^2

RailTrails

Consider the RailTrail data available in the mosaic package.

```
# to see the variables: ?RailTrail
data("RailTrail")
```

Q1. Build

a linear model with `cloudcover` used to predict the `volume` of riders. For the process: make a scatter plot (include a regression line), run the linear model (find the coefficients), check to see if `cloudcover` is discernible, and calculate the adjusted R^2 .

Hint: R^2 and adjusted R^2 are printed when the linear model is piped into `glance()`.

Q2. Check

Are the technical conditions met for the model?

Q3. Model

Find a model which predicts `volume` of ridership using the following set of variables: `hightemp` + `lowtemp` + `avgtemp` + `spring` + `summer` + `fall` + `cloudcover`. What happened?

Q4. Compare

With R^2 and adjusted R^2 (use: `glance()`) compare the following two models:

Model 1: `volume ~ avgtemp + cloudcover`

Model 2: `volume ~ spring + summer + cloudcover`

Q5. More variables

Find a model which predicts `volume` of ridership using some set of the explanatory variables which seem best to you. Use coefficient p-values as well as adjusted R^2 to make your determination.

Q6. Intervals

Find the CI and PI for the `volume` of riders using the model with `avgtemp` and `cloudcover`. Unfortunately, the intervals need to be generated separately, pipe linear model to: (1) `augment(interval = "confidence")` and (2) `augment(interval = "prediction")`. Consider the following day (which is in the dataset!): `avgtemp = 47F` and `cloudcover = 1.7` oktas.

- Interpret the CI for average volume at `avgtemp = 47F` and `cloudcover = 1.7` oktas.
- Interpret the PI for and individual day's volume at `avgtemp = 47F` and `cloudcover = 1.7` oktas
- Did that day's volume fall into the prediction interval?

Hint: you might want to `filter()` the `augment()`ed data by to find the days that were 47F.

```
praise()
```

```
[1] "You are flawless!"
```