ISOM 670 Business Statistics Fernbank Analysis

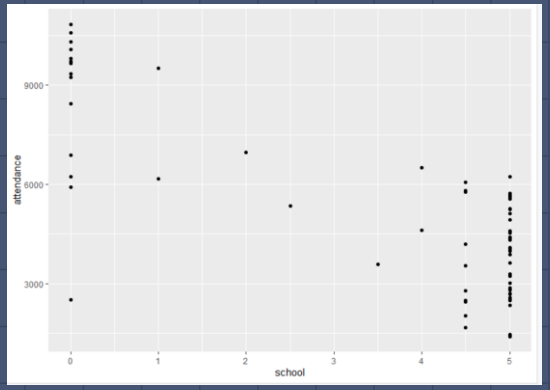
Group 5
Carl Xi, Stella Guo, Ivy Zhou, Jake Arendsen

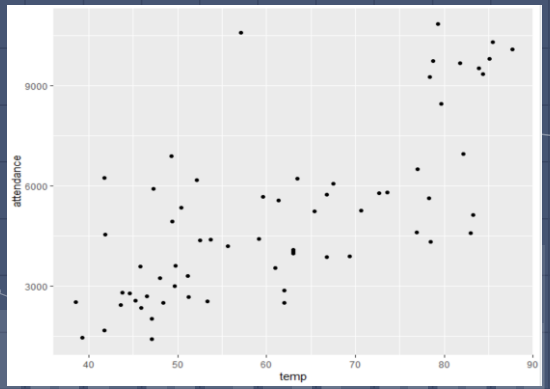
Attendance vs Key Variables

1. **What can you tell Colette about the relationship of attendance to these known variables?**

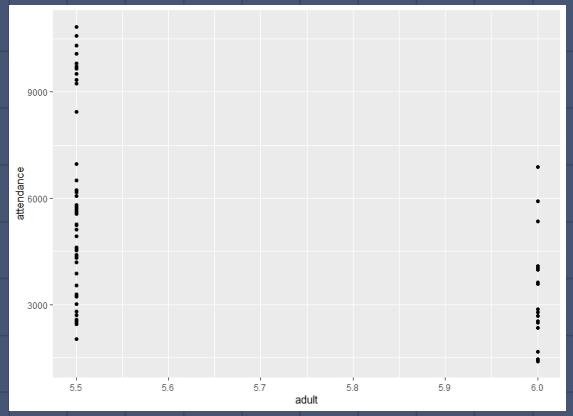
A: We see that attendance is strongly correlated with school and temp, moderately correlated with adult and weakly correlated with the remaining variables.

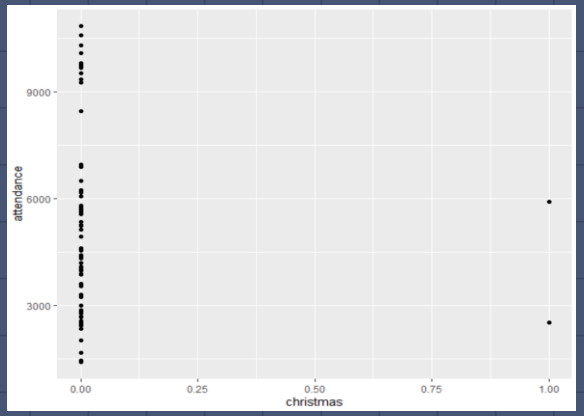
	attendance
attendance	1.0000
adult	-0.4045
school	-0.7735
snow/ice	-0.2006
temp	0.7106
sunshine	0.2385
christmas	-0.0628
advertising	-0.1195

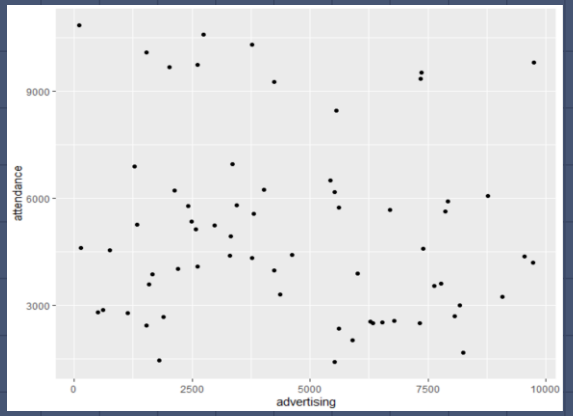


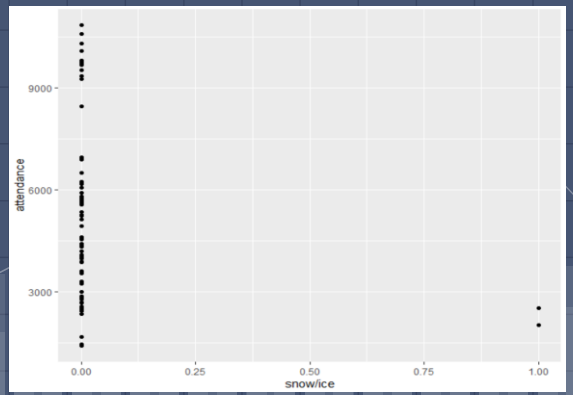


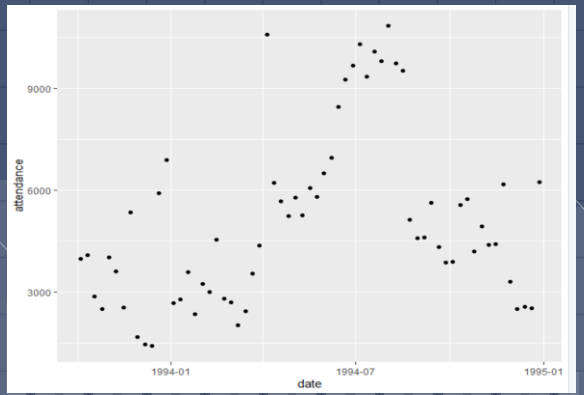
Attendance vs Remaining Variables

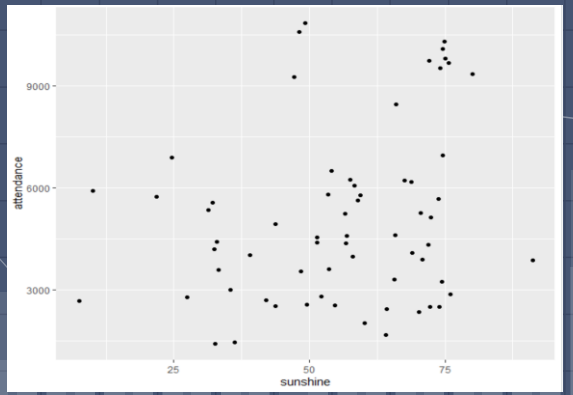












Forecasting Model

2. Colette would like a forecasting model that uses no more than 5 predictors. What would you recommend?

A: We recommend a forecasting model with 5 variables: school, temp, snow/ice, adult, sunshine.

Formula: Attendance = 17470 -781*school + 71*temp - 3001*snow/ice - 2334*adult - 13*sunshine

Our thought process is as follows:

- We started with a model with school because school and attendance had the highest correlation of -0.7735
- After assessing the coefficients and p-values, we continued to add variables that are highly correlated with the residuals
- We ended up with a model that has a residual standard error of 889.5 and an adjusted R-square of 0.8773. We felt that any additional variables will not constructively add value to our model

```
Call:
lm(formula = attendance ~ school + temp + `snow/ice` + adult +
    sunshine)

Residuals:
     Min       1Q   Median       3Q      Max
-1550.71  -646.03     2.75   519.41  2476.51

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 17470.282   3544.815   4.928 7.07e-06 ***
school       -780.907     58.782 -13.285  < 2e-16 ***
temp           71.355     10.047   7.102 1.82e-09 ***
`snow/ice`  -3001.336    688.991  -4.356 5.36e-05 ***
adult       -2333.670    571.353  -4.084 0.000135 ***
sunshine      -12.733      7.245  -1.758 0.084005 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 889.5 on 59 degrees of freedom
Multiple R-squared:  0.8868,    Adjusted R-squared:  0.8773
F-statistic: 92.49 on 5 and 59 DF,  p-value: < 2.2e-16
```

Price vs Attendance

3. What can you tell Colette about the relationship between price and attendance?

A: We see a moderate negative relationship between price and attendance, which is supported by a statistically significant negative coefficient in our model. That being said, price does not seem to be the strongest single predictor of attendance. Additionally, the fact that only two two price points have been tested in our data and one of these price points was tested in only winter months leaves us concerned with some potential multicollinearity this data could contain.