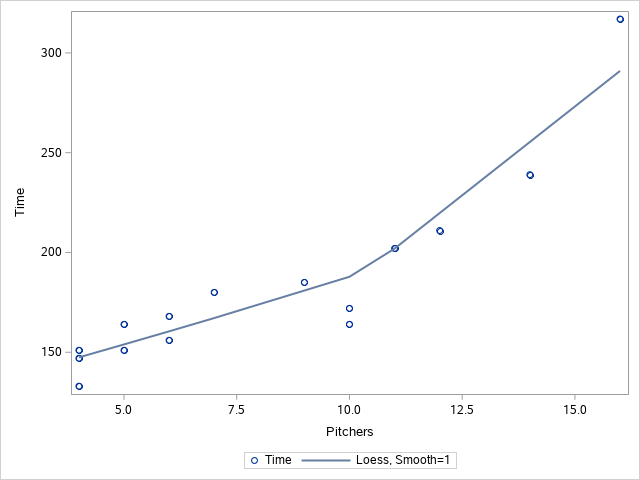
Dylan Smith

1.

a)



It seems to be mostly linear with a possible pull from the final point being a possible outlier.

b)

The Residual vs Quantile plot shows conformity around the center line with the furthest line being the last point. The residual plot shows a downward trend. Again, the last point is the only outlier to this trend. The fit plot for time shows the same trend as the other plots. The final point is on the edge of the prediction line. The model seems adequate if the last point were to be removed.

c)

The transformation that was recommended was that of Y^-1.25.

d)

| **Obs** | **Game** | **League** | **Runs** | **Pitchers** | **Time** | **yt** |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | CLE-DET | AL | 14 | 6 | 168 | .001653345 |
| **2** | CHI-BAL | AL | 11 | 5 | 164 | .001703904 |
| **3** | BOS-NYY | AL | 10 | 11 | 202 | .001313139 |
| **4** | TOR-TAM | AL | 8 | 10 | 172 | .001605423 |
| **5** | TEX-KC | AL | 3 | 4 | 151 | .001889204 |
| **6** | OAK-LAA | AL | 6 | 4 | 133 | .002214040 |
| **7** | MIN-SEA | AL | 5 | 5 | 151 | .001889204 |
| **8** | CHI-PIT | NL | 23 | 14 | 239 | .001064149 |
| **9** | LAD-WAS | NL | 3 | 6 | 156 | .001813820 |
| **10** | FLA-ATL | NL | 19 | 12 | 211 | .001243503 |
| **11** | CIN-HOU | NL | 3 | 4 | 147 | .001953680 |
| **12** | MIL-STL | NL | 12 | 9 | 185 | .001465667 |
| **13** | ARI-SD | NL | 11 | 10 | 164 | .001703904 |
| **14** | COL-SF | NL | 9 | 7 | 180 | .001516734 |
| **15** | NYM-PHI | NL | 15 | 16 | 317 | .000747611 |

e)

The predicted value vs residual plot shows random noise. The residual plot vs quantile plot shows most a pretty strong conformity around the prediction line. The fit plot shows no outliers. There are still questionable points, but they have been slightly improved. Overall there was an improvement but not a big change.

f)

Original R^2=0.7998

Transformed R^2=0.8592

The transformed with the higher R^2 is a better model due to it leaving less variation unexplained.

2.14)

a)

Hypothesis;

H0:B1=0

H1:B1≠0

Test Statistic=58.57

P-Value:<.0001

Conclusion: The P is low therefore, it must go. Reject H0. There is a significant regression relationship.

2.15)

a)

mean price=6.4141 – 7.8945

b)

particular =3.0383 – 11.2703

c)

mean price midpoint=7.1561

particular midpoint=7.1543

These two midpoints are almost identical which would make sense if the prediction were accurate with the mean being close a particular 450-page textbook.

d)

mean width=1.4804

particular width=8.232

This makes sense because the mean price is the average and would be closer to the midpoint due to the central limit theorem predicting a build up towards the mid-point whereas the particular value shows the min and the max values.

e)

Probably around 500 due to it being in the middle of

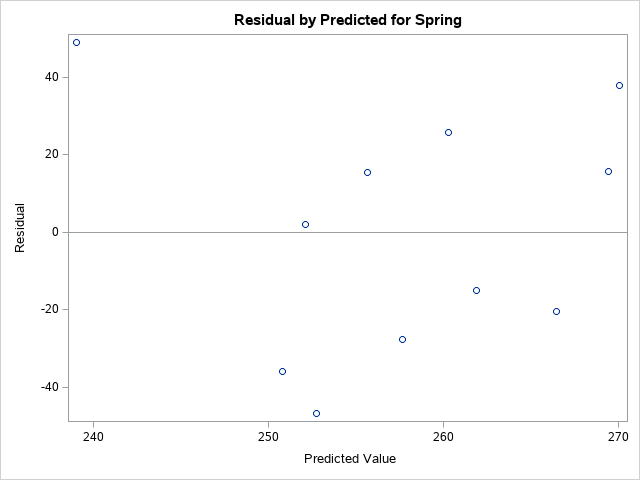
f)

particular prediction interval=12.7943 – 22.6522

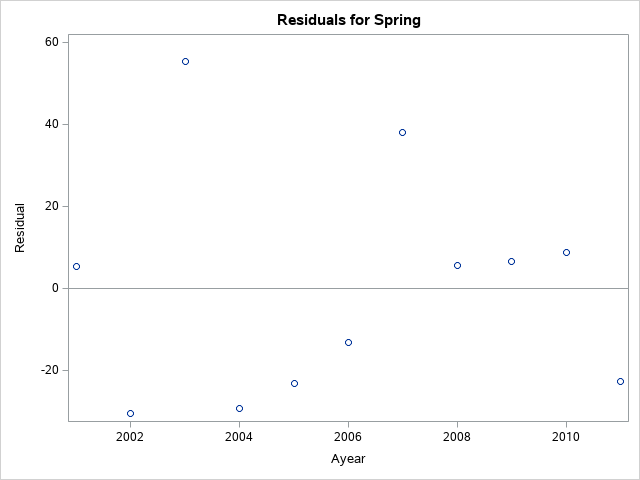
With predicting a number that is so far from the highest available data, the prediction doesn’t really have a 95% confidence interval.

3 (2.23))

a)



The predicted value vs the residual plot shows an almost flipped upside down and backwards image across 0 showing some pattern in the points.



Here it shows again an odd spread with points going from extreme lows to an extreme high then back down to an extreme low.

b)

48.68% of the model is explained.

c)

| **Analysis of Variance** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | | **Sum of Squares** | | **Mean Square** | | **F Value** | | **Pr > F** |
| **Model** | 1 | | 4721.12217 | | 4721.12217 | | 7.59 | | 0.0249 |
| **Error** | 8 | | 4976.47783 | | 622.05973 | |  | |  |
| **Corrected Total** | 9 | | 9697.60000 | |  | |  | |  |
| **Root MSE** | | 24.94113 | | **R-Square** | | 0.4868 | |
| **Dependent Mean** | | 254.80000 | | **Adj R-Sq** | | 0.4227 | |
| **Coeff Var** | | 9.78851 | |  | |  | |

d)

Ho: the slope of the model = 0

H1: the slope of the model ≠ 0

With all given values above considered, the null hypothesis would not be rejected and therefore it shows that there is no significant linear association between spring and fall.

e)

the confidence interval for the slope is 16.7872 – 17.04932

it does not contain zero which is important due to it being another indication that there is no linear association.

3(2.24))

a)

70.2530

b)

70.228 – 70.2779

c)

70.178 – 70.3279

e)

The mean interval (c) due to it having a lower width and it using the available data rather than predicted data.