Lisa Silverstein Graph Analysis

#1

```
Analyzing TEMP with month;
proc anova data=AirQualityT;
      class month;
      model TEMP = month;
      ods select OverallANOVA FitStatistics ModelANOVA;
run;
proc reg data=AirQualityT;
  model TEMP = month;
  output out=diagnostics cookd= cd;
run;
```

The F-value for the model is highly significant (P < .0001), indicating that the month variable is a significant predictor of TEMP and that the model explains more variance in TEMP than what would be expected by chance.

The R-squared value is 0.0231, suggesting that the model explains about 2.31% of the variance in TEMP. This indicates a relatively low explanatory power, suggesting that other variables not included in the model might also influence TEMP.

The residuals are not randomly scattered around the zero line, and there is a discernible pattern. This indicates potential non-linearity or presence of other variables influencing the residuals. There are some negative residuals and some positive, indicating that at times the observed value is less than the predicted value and at others the opposite.

#2

```
*Analyzing PRES with month;
proc anova data=AirQualityT;
      class month;
      model PRES = month;
      ods select OverallANOVA FitStatistics ModelANOVA;
run;
proc reg data=AirQualityT;
  model PRES = month;
  output out=diagnostics cookd= cd;
run;
```

The F-value for the model is highly significant (P < .0001), indicating that the month variable is a significant predictor of PRES.

The R-squared value is 0.0012, suggesting that the model explains only about 0.12% of the variance in PRES. This is very low and suggests that the model's explanatory power for PRES is minimal.

The residuals exhibit a cyclical pattern, similar to the pattern observed in the previous dataset for TEMP. This suggests that there might be a seasonal effect or other cyclical factor influencing PRES that the current model isn't capturing. Unlike TEMP, PRES stays generally constant with some seasonal changes.

#3

```
*Analyzing PM25 with month;
proc anova data=AirQualityT;
      class month;
      model PM25 = month;
      ods select OverallANOVA FitStatistics ModelANOVA;
run;
proc reg data=AirQualityT;
  model PM25 = month;
  output out=diagnostics cookd= cd;
run;
```

The F-value for the model is highly significant (P < .0001), indicating that the month variable is a significant predictor of PM25.

The R-squared value is 0.0011, suggesting that the model explains only about 0.11% of the variance in PM25. This is very low and suggests that the model's explanatory power for PM25 is minimal.

The pattern for PM25 is much less cyclical. Instead, there is a concentration of residuals around zero, suggesting that the model's predictions are close to the observed values. There are a few significant outliers, specifically toward the right side of the graph.

#4

```
*Analyzing PM10 with month;
proc anova data=AirQualityT;
      class month;
      model PM10 = month;
      ods select OverallANOVA FitStatistics ModelANOVA;
run;
proc reg data=AirQualityT;
  model PM10 = month;
  output out=diagnostics cookd= cd;
run;
```

The F-value for the model is highly significant (P < .0001), indicating that the month variable is a significant predictor of PM10.

The R-squared value is 0.0000134, suggesting that the model explains only about 0.00134% of the variance in PM10. This is *very* low and suggests that the model's explanatory power for PM10 is essentially nothing.

The pattern for PM10 is much less cyclical. Instead, there is a concentration of residuals around zero, suggesting that the model's predictions are close to the observed values. There are a few outliers, specifically toward the right side of the graph.

#5

```
*Analyzing SO2 with month;
proc anova data=AirQualityT;
      class month;
      model SO2 = month;
      ods select OverallANOVA FitStatistics ModelANOVA;
run;
proc reg data=AirQualityT;
  model SO2 = month;
  output out=diagnostics cookd= cd;
run;
```

The F-value for the model is highly significant (P < .0001), indicating that the month variable is a significant predictor of SO2.

The R-squared value is 0.0642, suggesting that the model explains only about 6.42% of the variance in SO2. This suggests that the model's explanatory power for SO2 is higher than other variables so far, but still small.

The pattern for SO2 is much less cyclical. Instead, there is a concentration of residuals around zero, suggesting that the model's predictions are close to the observed values. There are a few outliers, specifically toward the right side of the graph. The model isn't capturing certain patterns or changes towards the end of the observation period, or the final months of the year.

#6

```
*Analyzing NO2 with month;
proc anova data=AirQualityT;
      class month;
      model NO2 = month;
      ods select OverallANOVA FitStatistics ModelANOVA;
run;
proc reg data=AirQualityT;
  model NO2 = month;
  output out=diagnostics cookd= cd;
run;
```

The F-value for the model is highly significant (P < .0001), indicating that the month variable is a significant predictor of NO2.

The R-squared value is 0.0035, suggesting that the model explains only about 0.35% of the variance in NO2. This is low and suggests that the model's explanatory power for NO2 is minimal.

The pattern for NO2 is much less cyclical. Instead, there is a concentration of residuals around zero, suggesting that the model's predictions are close to the observed values. There are a few outliers throughout the graph. The model may not be capturing certain patterns or changes towards the end of the observation period, or the final months of the year.

#7

```
*Analyzing CO with month;
proc anova data=AirQualityT;
      class month;
      model CO = month;
      ods select OverallANOVA FitStatistics ModelANOVA;
run;
proc reg data=AirQualityT;
  model CO = month;
  output out=diagnostics cookd= cd;
run;
```

The F-value for the model is highly significant (P < .0001), indicating that the month variable is a significant predictor of CO.

The R-squared value is 0.0042, suggesting that the model explains only about 0.42% of the variance in CO. This is low and suggests that the model's explanatory power for CO is minimal.

The pattern for CO is much less cyclical. Instead, there is a concentration of residuals around zero, suggesting that the model's predictions are close to the observed values. There are a few outliers throughout the graph. The model may not be capturing certain patterns or changes towards the end of the observation period, or the final months of the year.

#8

```
*Analyzing O3 with month;
proc anova data=AirQualityT;
      class month;
      model O3 = month;
      ods select OverallANOVA FitStatistics ModelANOVA;
run;
proc reg data=AirQualityT;
  model O3 = month;
  output out=diagnostics cookd= cd;
run;
```

The F-value for the model is highly significant (P < .0001), indicating that the month variable is a significant predictor of O3.

The R-squared value is 0.0143, suggesting that the model explains only about 1.43% of the variance in O3. This is low and suggests that the model's explanatory power for O3is minimal.

The pattern for O3 is much less cyclical. Instead, there is a concentration of residuals around zero, suggesting that the model's predictions are close to the observed values. There are a few outliers on the right side of the graph. The model may not be capturing certain patterns or changes towards the end of the observation period, or the final months of the year.