# Outline Solutions for Chapter 2

### 2.1

Time series plot for temperature



Seasonal plot: clear and stable seasonal pattern



### 2.2

Time series plot for Sales ($Billion)



Seasonal plot: December peak



### 2.3

(a) Scatter plot for injuries versus train miles



(b) Time plots for Injuries, Train Miles and Injuries per 100 million train miles







(c) The level of injuries fluctuates but shows an upward trend over time.

### 2.4

The standard Excel output shows that the median is unchanged but the mean is much smaller and the standard deviation much larger. A few outliers can greatly distort the analysis.



The MAD is 7.6 without the outlier and 21.7 with the outlier.

These statistics only make sense for the investor if she has equal amounts invested in each stock.

### 2.5

Means and medians are close suggesting a symmetric pattern. November and the winter months are somewhat more variable. The month-by-month variation is much less than the overall values and the monthly figures are much more useful.

Statistics

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | N | Mean | StDev | Median MAD |  |
| Temperature | 300 | 51.584 | 14.126 | 50.550 12.43 |  |



### 2.6

The standard summary statistics are given in the table. To compute the MAD in each case, calculate the absolute values of the deviations about the mean and compute their average.



|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Train-miles | Injuries | Injuries per T-M |
| MAD |

|  |  |  |
| --- | --- | --- |
| 6.7 |  |  |

 | 142.0 | 110.4 |

Since the variables display a rising trend, the calculation of these averages is not particularly informative.

### 2.7

Correlations

|  |  |  |  |
| --- | --- | --- | --- |
|  | Injuries | Train-miles | Injuries per T-M |
| Train-miles | 0.875 |  |  |
|  | 0.000 |  |  |
|  |  |  |  |
| Injuries per T-M | 0.990 | 0.805 |  |
|  | 0.000 | 0.000 |  |
|  |  |  |  |
| Year | 0.861 | 0.973 | 0.803 |
|  | 0.000 | 0.000 | 0.000 |

*Cell Contents
      Pearson correlation
      P-Value*

 The high correlation between Train miles and Injuries per TM suggests declining safety. The correlations with time reflect the upward trend. The associated P-value is 0.000 to 3 decimal places.

### 2.8

The correlation is 0.955 with a P-value to 3 decimal places of 0.000. As would be expected the consumption pattern is similar from year to year.

### 2.9

Results should agree with Table 2.6 in the text to serve as a check on computing these results.

The GDP forecast is often the baseline from which forecasts are built, as is reflected in the strong correlations across the board. Some variables (e.g. Unemp) show low correlations with variables other than GDP.

Correlations

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | GDP | Privcons | GFCF | Exports | Imports | Govsurp | Consprix |
| Privcons | 0.430 |  |  |  |  |  |  |
|  | 0.032 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| GFCF | 0.601 | 0.295 |  |  |  |  |  |
|  | 0.001 | 0.153 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Exports | 0.459 | 0.262 | 0.192 |  |  |  |  |
|  | 0.021 | 0.206 | 0.357 |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Imports | 0.209 | 0.469 | 0.371 | 0.641 |  |  |  |
|  | 0.315 | 0.018 | 0.068 | 0.001 |  |  |  |
|  |  |  |  |  |  |  |  |
| Govsurp | 0.620 | 0.033 | 0.223 | 0.518 | 0.085 |  |  |
|  | 0.001 | 0.874 | 0.285 | 0.008 | 0.688 |  |  |
|  |  |  |  |  |  |  |  |
| Consprix | 0.317 | 0.192 | 0.336 | 0.210 | 0.382 | 0.179 |  |
|  | 0.123 | 0.358 | 0.101 | 0.313 | 0.060 | 0.391 |  |
|  |  |  |  |  |  |  |  |
| Unemp | -0.363 | 0.073 | -0.295 | -0.273 | -0.087 | -0.317 | 0.063 |
|  | 0.074 | 0.727 | 0.153 | 0.187 | 0.678 | 0.123 | 0.766 |

*Cell Contents
      Pearson correlation
      P-Value*

### 2.10

a. The mean and median are not useful, given the strong trend in the data.

b. There is a clear slowing of percentage growth over time becoming stable around 5 percent. Note the outlier in 2011Q1 as the result of an acquisition.



### 2.11

The results for Forecast\_1 (previous year) and Forecast\_2 (long-run average) are

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Statistics

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | N | Mean | Median |
| Error\_1 | 48 | 0.217 | 0.150 |
| MAE\_1 | 48 | 3.188 | 2.500 |
| RMSE\_1 | 48 | 4.023 | 2.500 |
| Error\_2 | 48 | 0.710 | 0.354 |
| MAE\_2 | 48 | 2.408 | 2.178 |
| RMSE\_2 | 48 | 3.088 | 2.178 |

 |  |  |

The RelMAE values (Forecast\_2 relative to Forecast\_1) are 0.755 (mean) and 0.871 (median). The median measures are generally smaller but the patterns are consistent.

The use of the means produces apparently better results, but it should be recognized that the means include the observations actually being forecast. Only prior information should be used. The previous year values have a smaller bias.

### 2.12

Statistics

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | N | Mean | StDev | Minimum | Q1 | Median | Q3 | Maximum | Skewness | Kurtosis |
| Temperature | 18 | 36.11 | 10.97 | 7.00 | 32.00 | 35.00 | 42.50 | 55.00 | -0.78 | 2.03 |

The mean is around 36 degrees and the standard deviation is 11 degrees. With 18 observations, the 95% t-value from tables is 2.11 so the prediction interval is . In other words, be prepared for almost any kind of weather. Clearly much better forecasts are possible closer to the event!!

### 2.13

Normal uses mean +/- 1.645\*SD for 90 percent and mean +/- 1.96\* SD for 95 percent.

Empirical uses the 5.0 percentile for 90 percent and the 1.67 percentile (conservative) for 95 percent.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Empirical-90%** | 1 | 2 | 3 | 4 | 5 |
| Lower | -6 | -9 | -8 | -8 | -10 |
| Upper | 3 | 5 | 7 | 11 | 10 |
| **Normal-90%** | 1 | 2 | 3 | 4 | 5 |
| Lower |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -5.13 | -7.27 | -9.17 | -10.37 | -11.18 |
| 4.06 | 5.94 | 7.37 | 7.83 | 8.31 |

 |
| Upper |
| **Empirical-95%** | 1 | 2 | 3 | 4 | 5 |
| Lower | -9 | -10 | -11 | -9 | -13 |
| Upper | 3 | 8 | 11 | 12 | 12 |
| **Normal-95%** | 1 | 2 | 3 | 4 | 5 |
| Lower |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| -6.01 | -8.54 | -10.76 | -12.11 | -13.05 |
| 4.94 | 7.21 | 8.96 | 9.58 | 10.18 |

 |
| Upper |

The empirical 95 percent limits are mostly a bit wider; they are not monotone increasing. The 90 percent limits are close to each other.

### 2.14

The normal intervals for the one-year-ahead forecasts are +/- 8.55 and for the overall average +/- 6.01. In both cases the intervals cover 95 percent of cases.

The empirical intervals cover 95 percent by construction (no calculations needed!).

Because the intervals are based upon the error distributions rather than out-of-sample, these results are not quite legitimate. Why?

### 2.15

The RMSE is 211 and the value from tables is z = 1.645. The intervals are very wide and more data on monthly variations would be useful.

Period Forecast Lower PI Upper PI

Jan 790 442.90 1137.10

Feb 810 462.90 1157.10

Mar 680 332.90 1027.10

Apr 500 152.90 847.10

May 520 172.90 867.10

Jun 810 462.90 1157.10

Jul 1120 772.90 1467.10

Aug 1840 1492.91 2187.10

Sep 1600 1252.91 1947.10

Oct 1250 902.90 1597.10

Nov 740 392.90 1087.10

Dec 610 262.90 957.10

2.16

Let S = sales and I = inventory. Then the cost will be:

 

Given the probabilities, we can compute the expected cost for given values of C and I. The solution may be determined numerically by various means. As expected, when C=1, the optimal level of inventory to hold is I = 100 and the total expected cost is 5.24. When C=3, the optimal level of inventory to hold is I = 105 and the total expected cost is 7.86.

To guarantee a service level of 90 percent we would need to stock108 units.

**Minicase 2.1**

Look for non-linearities in the plots. Is 12 years a reasonable cut-off?

**Minicase 2.2**

Summarize from a graphical perspective, looking for trends. Relate to Minicase 3.2 later.

**Minicase 2.3**

Note that the data are NOT a regular time series. However, the series of durations and gaps might be treated as such to see if there is any dependence within or between the series. 