Attachment A - Recommendations from the NCSU Department of Statistics

Dr. Johnson's Recommendations for Assigning Groups

Assigning a Station to a Seasonal Group

It will usually be best to assign a Station to the Seasonal Group with which its counts are most highly correlated over days and months. To help with the visual comparisons, compare the traffic counts at a Station with the inverse of the Index for each Seasonal Group. This makes both counts and Indexes begin with relatively low values in January and increase to relatively high values during the summer. Standardize the values of both the Indexes and the Counts and put the results on a common scale in plots of the values versus Month. Standardize the values by subtracting the mean and dividing by the standard deviation. This transformation is often called the Z transformation.

Although it is usually best to assign a Station to the Seasonal Group with which the readings are most highly correlated, several groups have similar profiles and there may be other reasonable choices. Practice with a few stations has shown that two or three correlations may be very close. This means that with good judgement, an engineer may use the assignments of neighboring stations to choose among the reasonable choices for Seasonal Group to which to assign a given station. We do not have a formal automated procedure to make such comparisons. However, the information contained in the recommended reports for each station will support the informed engineering judgement for such assignments.

Sampling Traffic Counts at Stations as the Basis for Assigning a Group

When the Group for assignment is clear from past history or the surrounding stations, data from three or four 48-hour sampling periods might be sufficient to estimate the scaling multiplier. The scaling multiplier is the number by which the Seasonal Group Index is multiplied to get the estimated profile for the station. However, when the Seasonal Group assignment is in doubt, additional 48 hour sampling periods may be required to have reasonable confidence in the assignment.

Refer to the figure Inverse Seasonal Index Patterns by Month to see differences in patterns that you are trying to find with the sampling procedure. See the table Inverse Seasonal Index Pattern by Month for the Pairwise Correlation among the seven Seasonal Groups. Profiles of Groups 6 and 7 are highly correlated (0.9635), which means it would take a lot of observations to determine whether a Station should be assigned to Group 6 or to Group 7. However, because Groups 6and 7 are so highly correlated, this distinction should not be critical. Groups 3 and 7 are also highly correlated (0.9342). The main difference between Group 3 and Group 7 seems to be that there is more variation between days in Group 3. Therefore, readings for entire weeks would be required to distinguish between Groups 3 and 7.

Groups 1 and 2 are highly correlated (0.9387) with the main distinction being more of a drop in September for Group 2 than for Group 1.

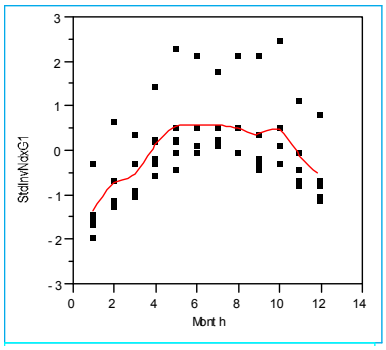
Groups 4 and 5 have visually unique profiles and are not highly correlated with any other Group. The distinctions for Group 4 are the drop between April and May, the bimodal pattern (with peaks in April and September) and the great variation between days of the week in the last half of the year. The distinctions for Group 5 are the drop from August to September, the increase from September to December and the drop from December to January and February.

To find these distinctive differences, it would be well to begin with a 48-hour sampling period in each of the even numbered months 2, 4, 6, 8, 10 and 12. To measure the variation between days of the week, a full week of readings should be included in month 6 or month 8. To identify groups 3,4 and 5, additional sampling periods should be added in months 5, 7 and 9. If these results do not establish an assignment with sufficient confidence, additional sampling periods should be assigned to the months that would most reduce the uncertainty in assignment.

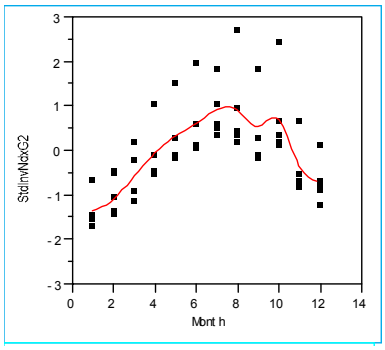
Inverse Seasonal Index Pattern by month

(fit with spline: lambda = 0.1)

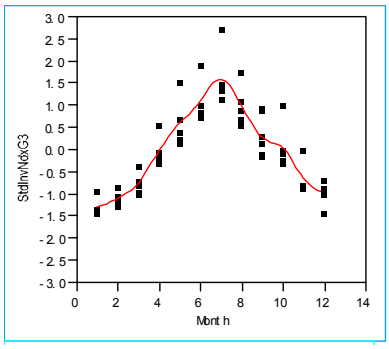
Group 1



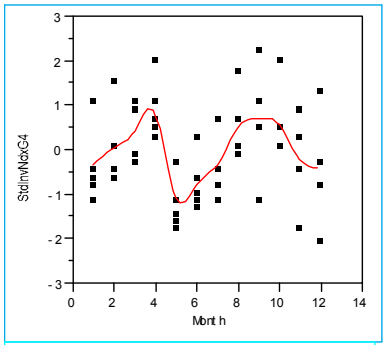
Group 2



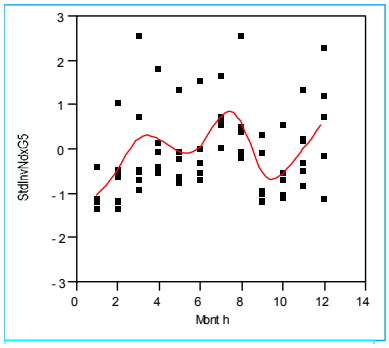
Group 3



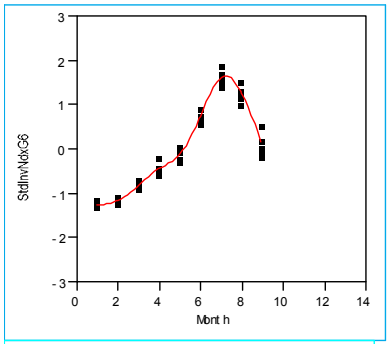
Group 4



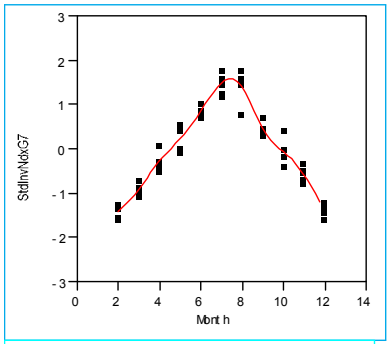
Group 5



Group 6



Group 7



Inverse Seasonal Index Pattern by month
(fit with spline: lambda = 0.1)

Pairwise Correlations

Variable	by Variable	Correlation	Count	Signif Prob
InvNdxG2	InvNdxG1	0.9387	60	0.0000
InvNdxG3	InvNdxG1	0.7764	60	0.0000
InvNdxG3	InvNdxG2	0.8834	60	0.0000
InvNdxG4	InvNdxG1	0.5602	60	0.0000
InvNdxG4	InvNdxG2	0.4595	60	0.0002
InvNdxG4	InvNdxG3	0.1544	60	0.2389
InvNdxG5	InvNdxG1	0.6715	60	0.0000
InvNdxG5	InvNdxG2	0.6152	60	0.0000
InvNdxG5	InvNdxG3	0.4832	60	0.0001
InvNdxG5	InvNdxG4	0.4862	60	0.0001
InvNdxG6	InvNdxG1	0.6058	45	0.0000
InvNdxG6	InvNdxG2	0.7826	45	0.0000
InvNdxG6	InvNdxG3	0.9180	45	0.0000
InvNdxG6	InvNdxG4	-0.0179	45	0.9073
InvNdxG6	InvNdxG5	0.4659	45	0.0013
InvNdxG7	InvNdxG1	0.5781	55	0.0000
InvNdxG7	InvNdxG2	0.7648	55	0.0000
InvNdxG7	InvNdxG3	0.9342	55	0.0000
InvNdxG7	InvNdxG4	0.0573	55	0.6778
InvNdxG7	InvNdxG5	0.2533	55	0.0621
InvNdxG7	InvNdxG6	0.9635	40	0.0000