American Society for Quality
Automotive Division

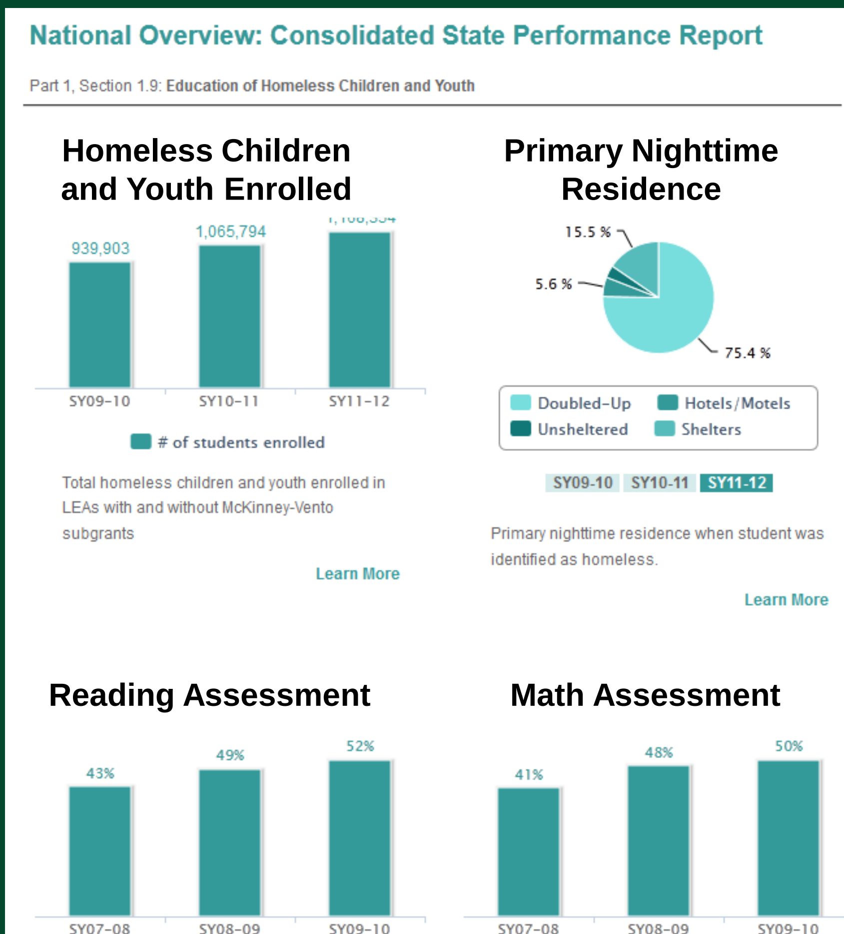
ASQ

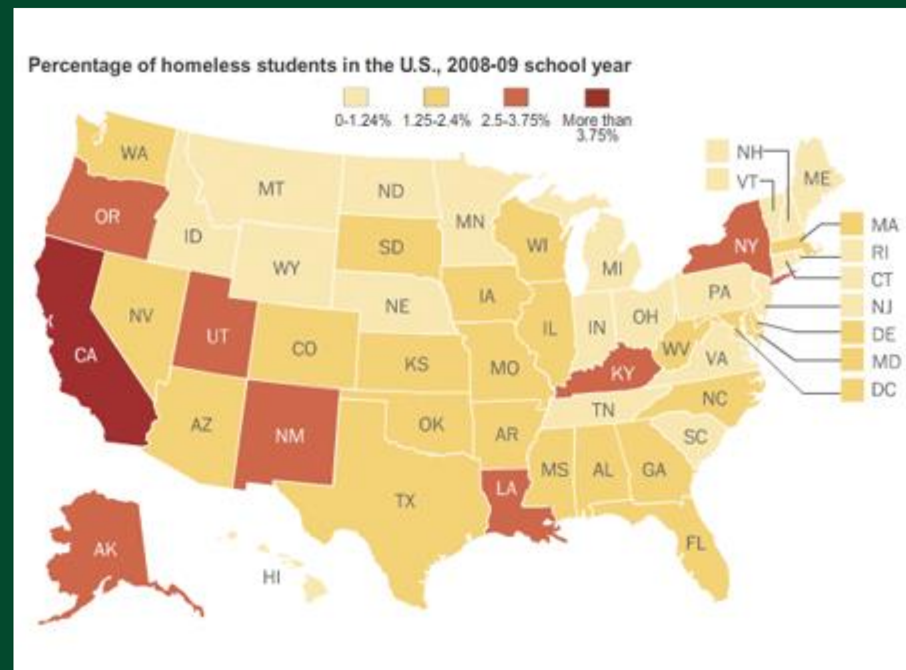
Implementing Predictive Analytics: A Prospective Approach

David J Corliss, PhD

Introduction: Retrospective vs. Prospective

The National Center for Homeless Education (NCHE), a federal agency, compiles data and files an annual report.





Out of date the day it was published

Methodology

- NCHE data by state for the 2009-2010 and 2010-2011
- Socio-economic data – e.g., US Census Bureau
- Two forecasts: total number of homeless students and 12 month % change
- Modeling methodology: linear regression
- Validated by comparison to 2011-2012 actuals
- The factors in these models are analyzed as key drivers of homeless student levels and changes in the population.

Key Findings

- The model forecasts homeless student levels by state for the 2011-2012 school using socio-economic data and NCHE data from previous years

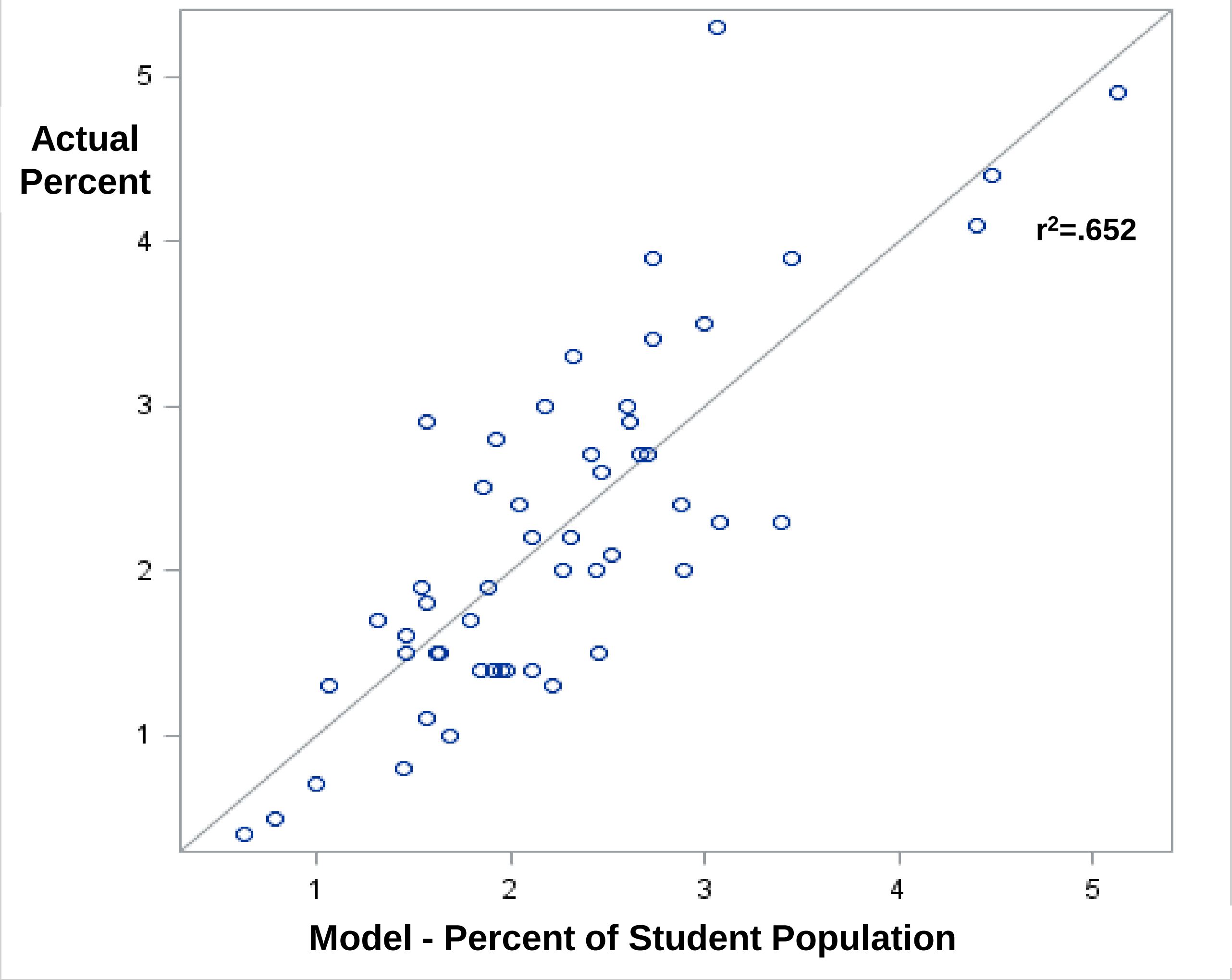
- Key drivers of changes in homeless student levels are income disparity (GINI Index), % of one-parent households, household size, median household income and housing cost

- Total Number forecast: good performance (r^2=0.652)
- Rate of Change forecast: moderate performance (r^2=0.518), likely reduced by variation in reporting standards from state to state

- The model increases accuracy over "dead reckoning" (simply carrying over the rate of change from the previous year) by 39.7%

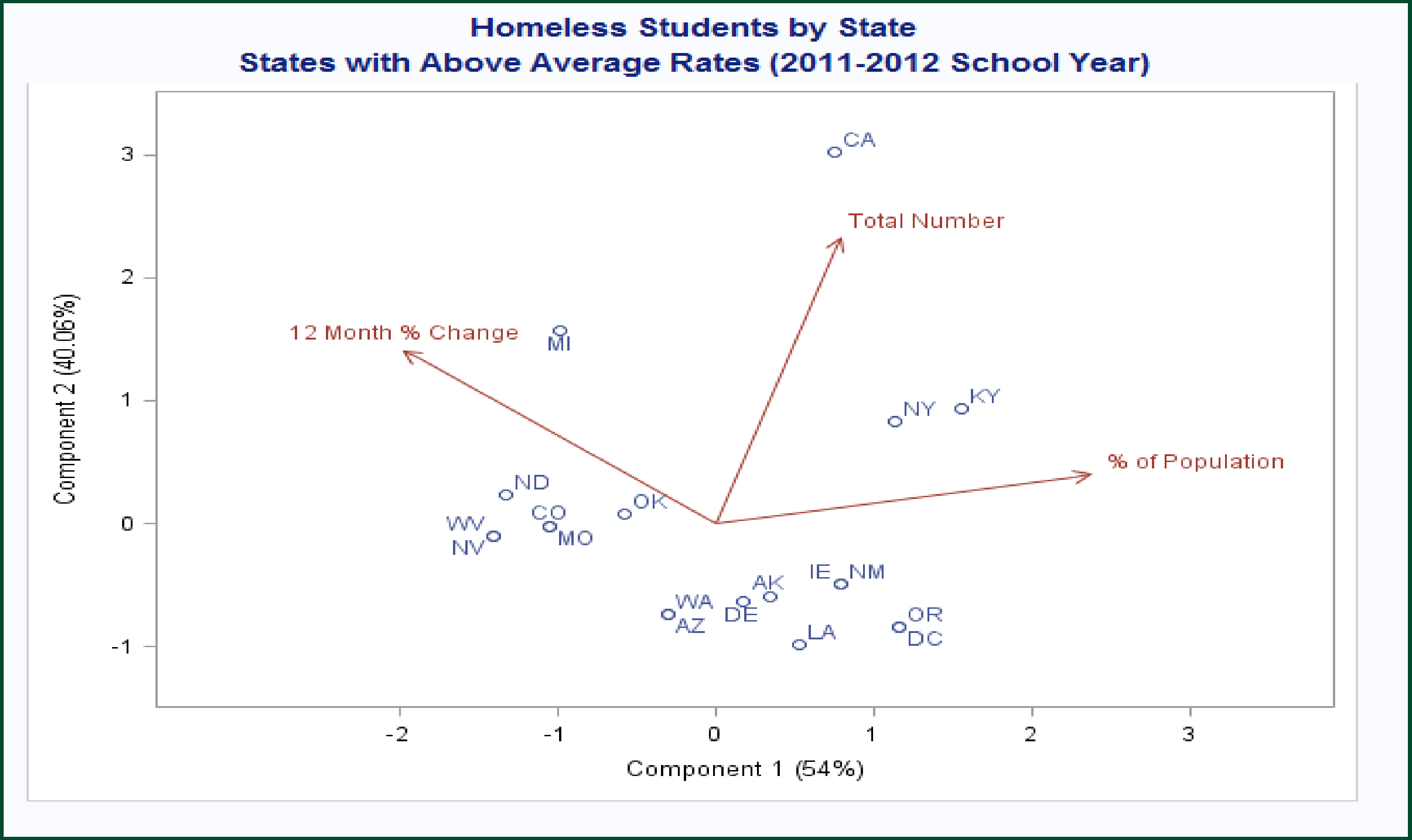
- No forecast is possible for Bureau of Indian Education (BIE) due to lack of complete socio-economic data for Indian Reservations; these areas have the second highest homeless student rate in the nation at 4.9%

Forecast: Number of Homeless Students by State (Percent of Total Student Population)



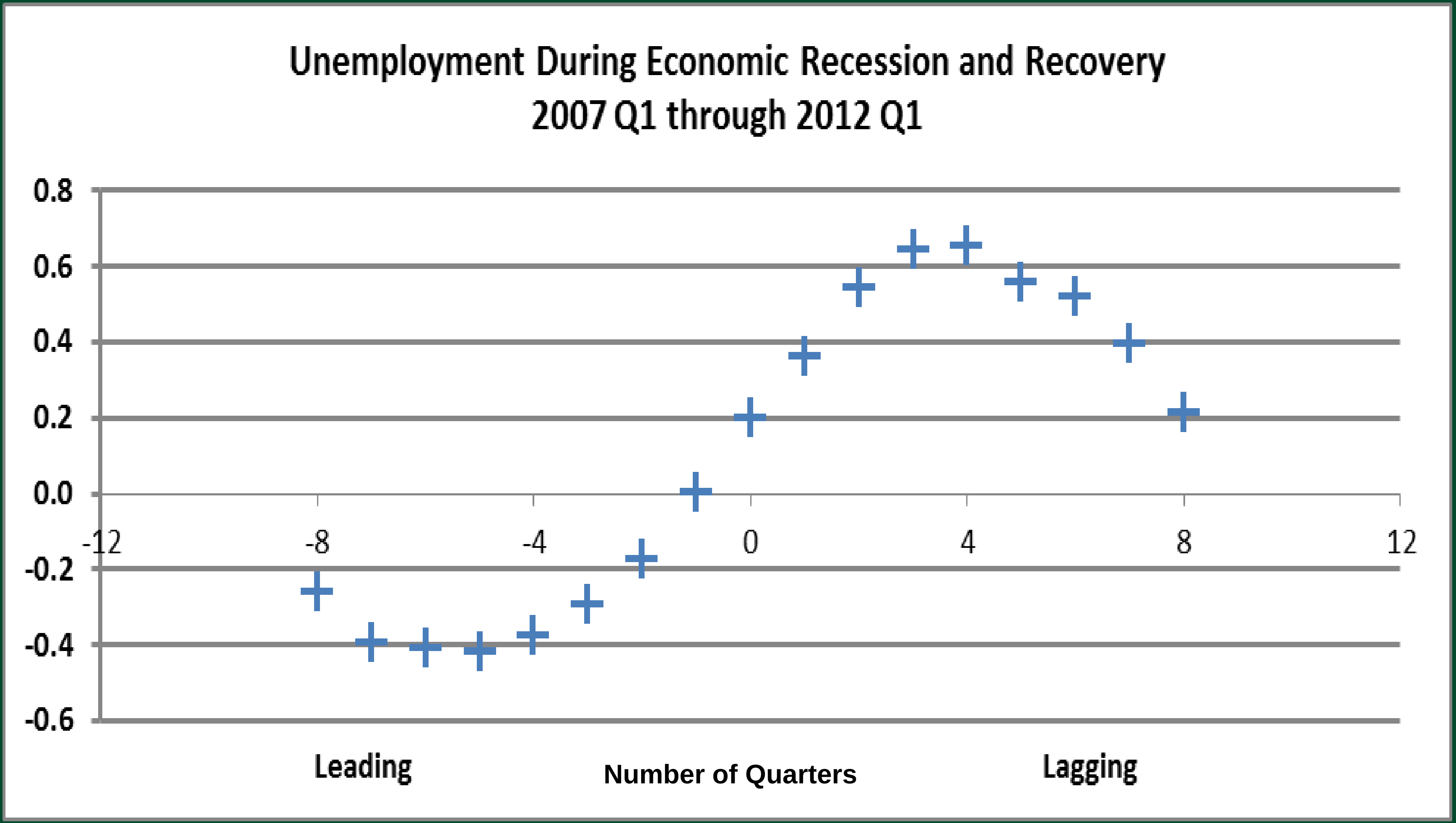
Solid performance of the model across the range from low to high homelessness states indicates consistency of factors correlated with the number of homeless students

Principal Component Analysis



While some states have many homeless student mostly due a large overall population, a separate group is distinctive for a high percentage of homelessness

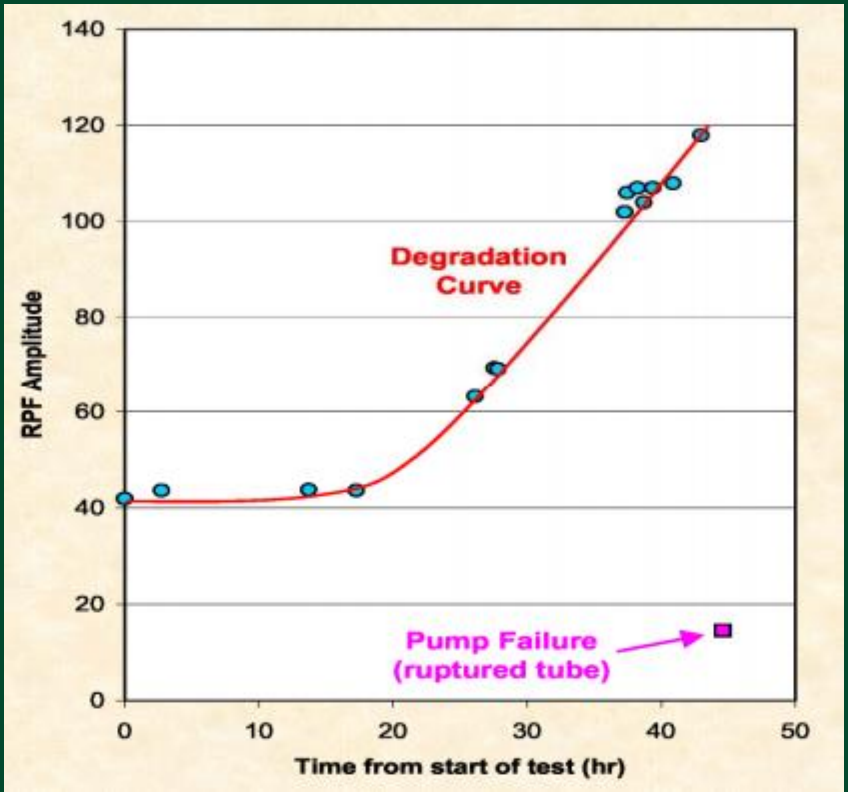
Prospective Analysis Using Leading Indicators



Changes in GDP lead the unemployment rate by 11 months

Leading Indicators and Future Part Failure

Behavioral Changes



Arthur Zingher & Mike Silverman, OPS A La Carte
Oak Ridge National Laboratory

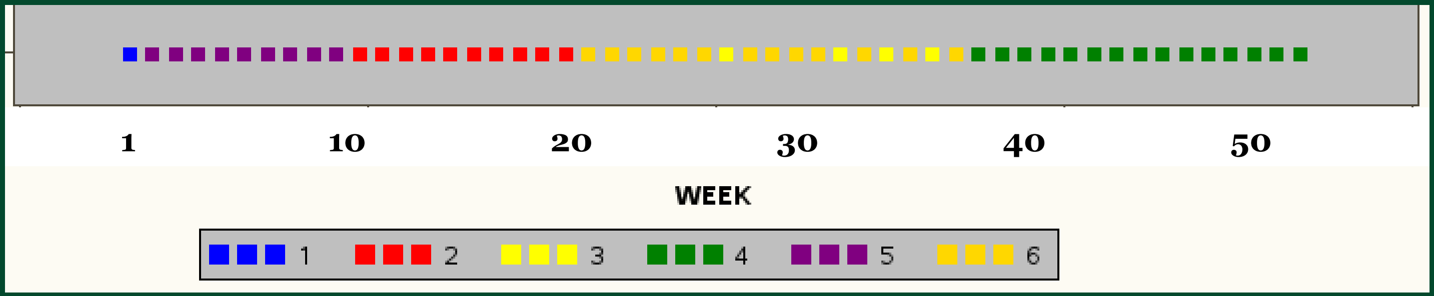
Changes in vibration indicate future part failure

Leading Correlations

Part Number	Correlation
ABN-0.0118	0.978
ABN-0.3619	0.925
ABL-0.8491	0.685
ABN-0.3916	0.679
ABN-0.2541	0.567
.	.
.	.
.	.
ABL-0.0082	0.216
ABL-0.5483	0.140
ABL-0.1425	0.109
ABL-0.6054	0.075

The rate of early failure in a minor part may correlate to early failure in a major component

Seasonal Variations



Cluster analysis of gas prices reveals and quantifies annual seasonal changes

Blue: Post-holiday lull

Purple: Winter Season

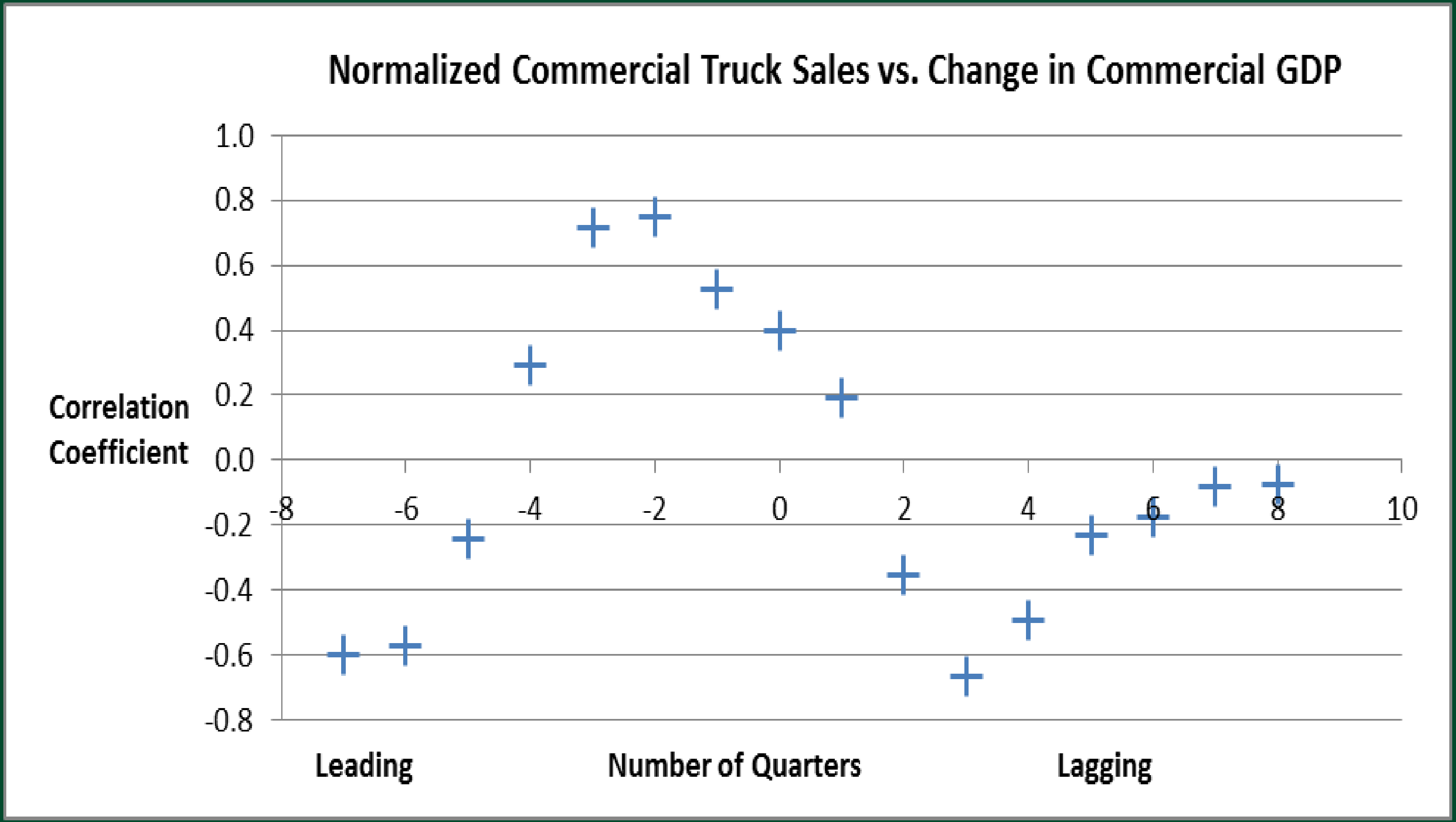
Red: Spring Run-up

Amber: Summer Driving Season

Yellow: Holiday Spikes

Green: Fall Season

Analysis Used To Forecast Future Economic Conditions



Changes in commercial truck sales lead the economy by 7 months

Conclusions

Reports filed over time can provide the data needed to develop a forecast that will facilitate decisions based on present, rather than historical, circumstances

In the example from homeless students, the accuracy of predictive modeling was 39.7% higher than an assessment based on historical reporting

A prospective approach using advanced analytic methods can predict future part failures, costs, sales, and economic conditions impacting the automotive industry

Questions

David J. Corliss
david.corliss@wayne.edu