

Adjustments Made Easy: A Look at Clustering Time Series from the US Census Bureau Manufacturers' Shipments, Inventories, and Orders



Matthew Barton, Allyson Hineman, Kampbell Howard, Ryan Plumer
Advisor: Dr. Lan Cheng, Department of Mathematical Sciences, SUNY Fredonia
Industrial Liaison: Dr. James Livsey, US Census Bureau



Introduction

The Economic Directorate releases time series data such as new orders, total inventories, and value of shipments for hundreds of different industries on a monthly basis. The US Census Bureau typically analyzes this data univariately when they seasonally adjust it around holiday time each year. The problem with this is that there are many time series to look at and adjust individually. Thus, to make these adjustments easier, by finding time series that follow similar trends, they could perform a multivariate adjustment on the clusters of time series. Through a series of testing, we were able to make suggestions to analysts on which series should be jointly modeled. The results outlined in this poster provide a sample of four from thirteen total suggested clusters using Dynamic Time Warping (DTW) methodology that the Economic Directorate can use to jointly model and relate industries.

Methodology

- Completed data wrangling, with the goal of converting data into time series objects
- Utilized DTW techniques to efficiently cluster related time series
- Produced dendrograms to display the hierarchical clustering
- Learned the syntax of the R programming language
- Created a GitHub Repository to work in a collaborative coding environment

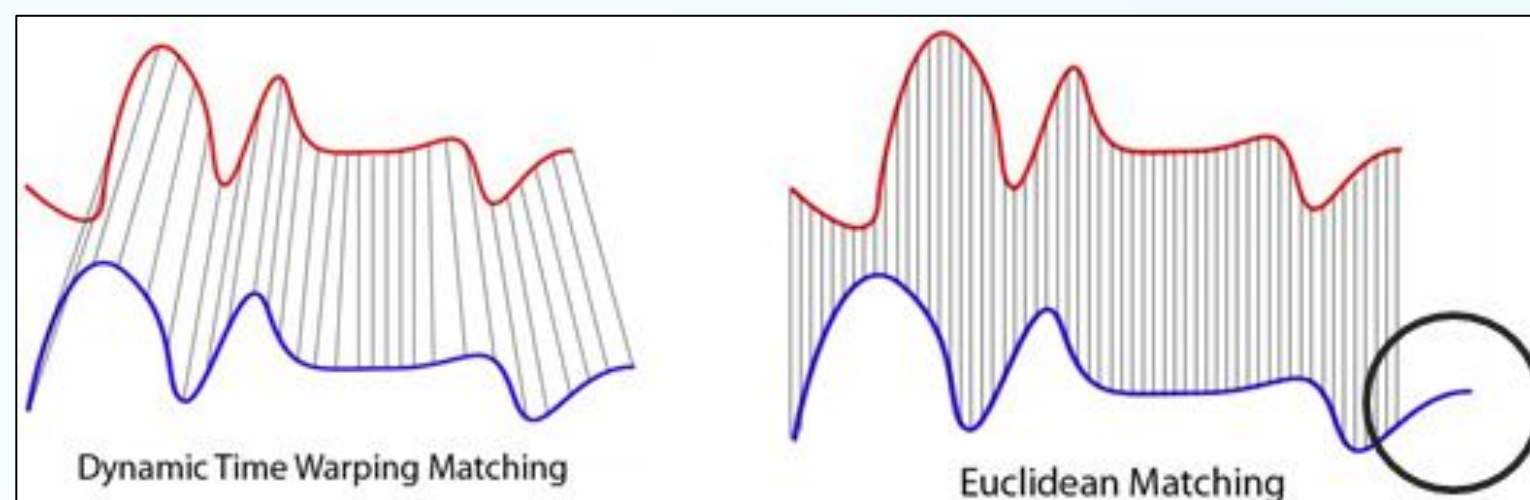


Figure 1. DTW Matching compared to Euclidean Matching.

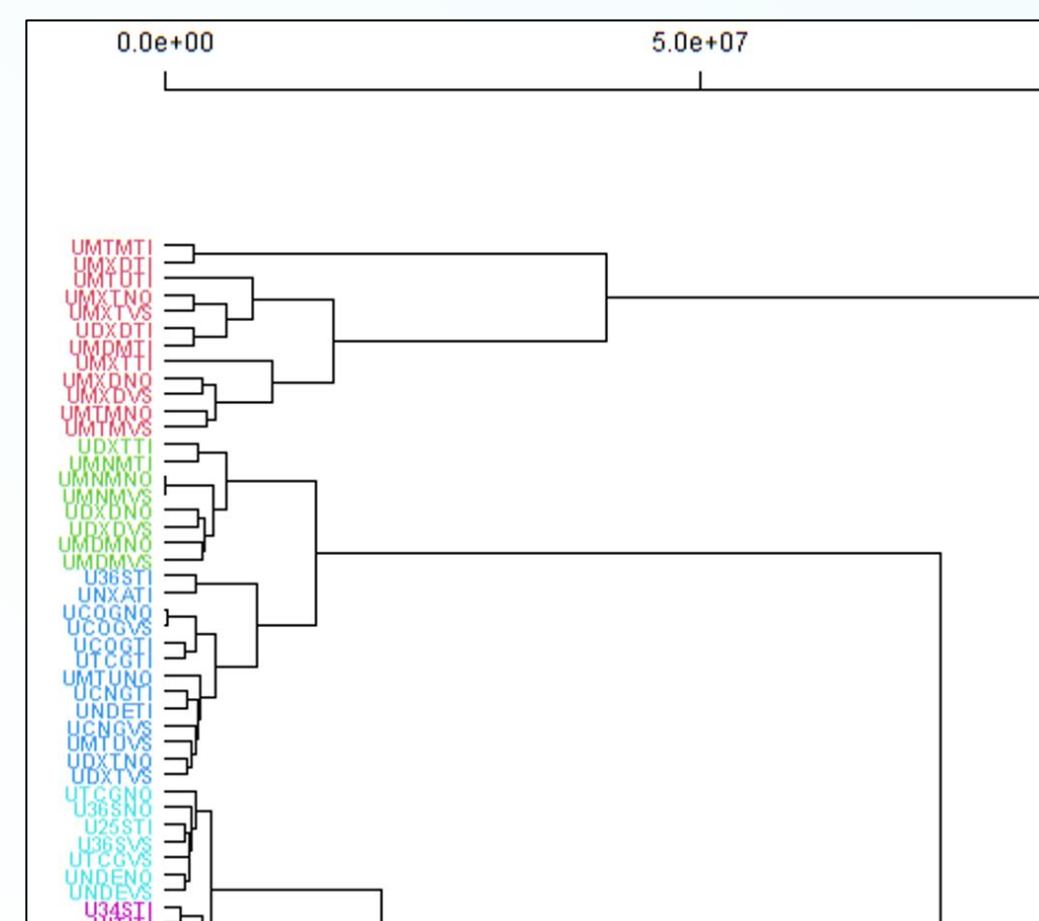


Figure 2. Section of completed dendrogram in millions of dollars from DTW clustering method that suggests which time series should be jointly modeled by inspection.

Results

Unadjusted Domestic Manufacturing Sector Clusters from the US Census Bureau Economic Directorate

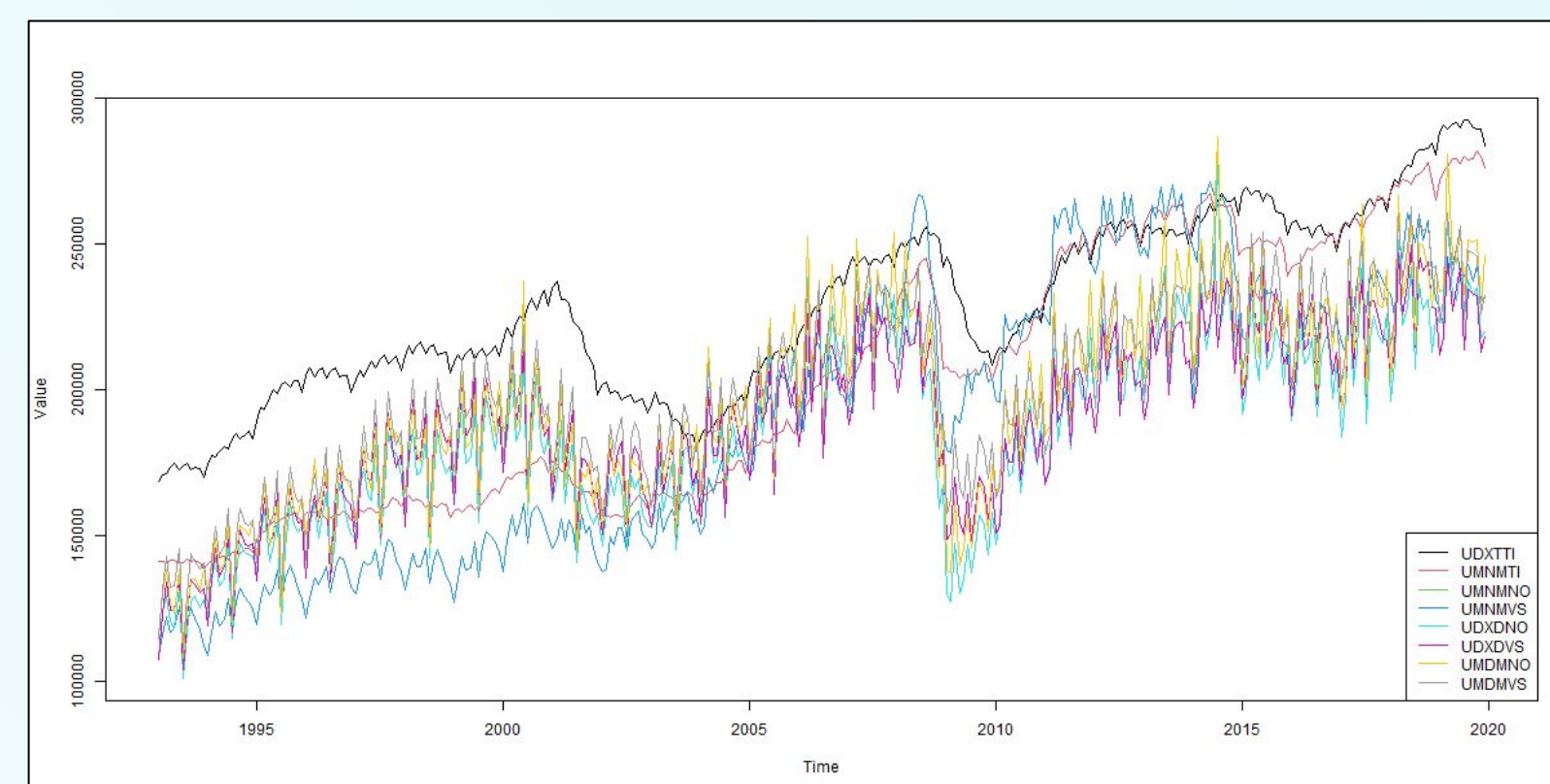


Figure 3. Includes: durable goods excluding transportation total inventory (UDXTTI), non-durable goods total inventory/new orders/value of shipments (UMNMTI/NO/VIS), durable goods excluding defense new inventories/value of shipments (UDXDNO/VIS), and durable goods new orders/value of shipments (UMDMNO/VIS) in millions of dollars from Jan. 1993 to Dec. 2019.

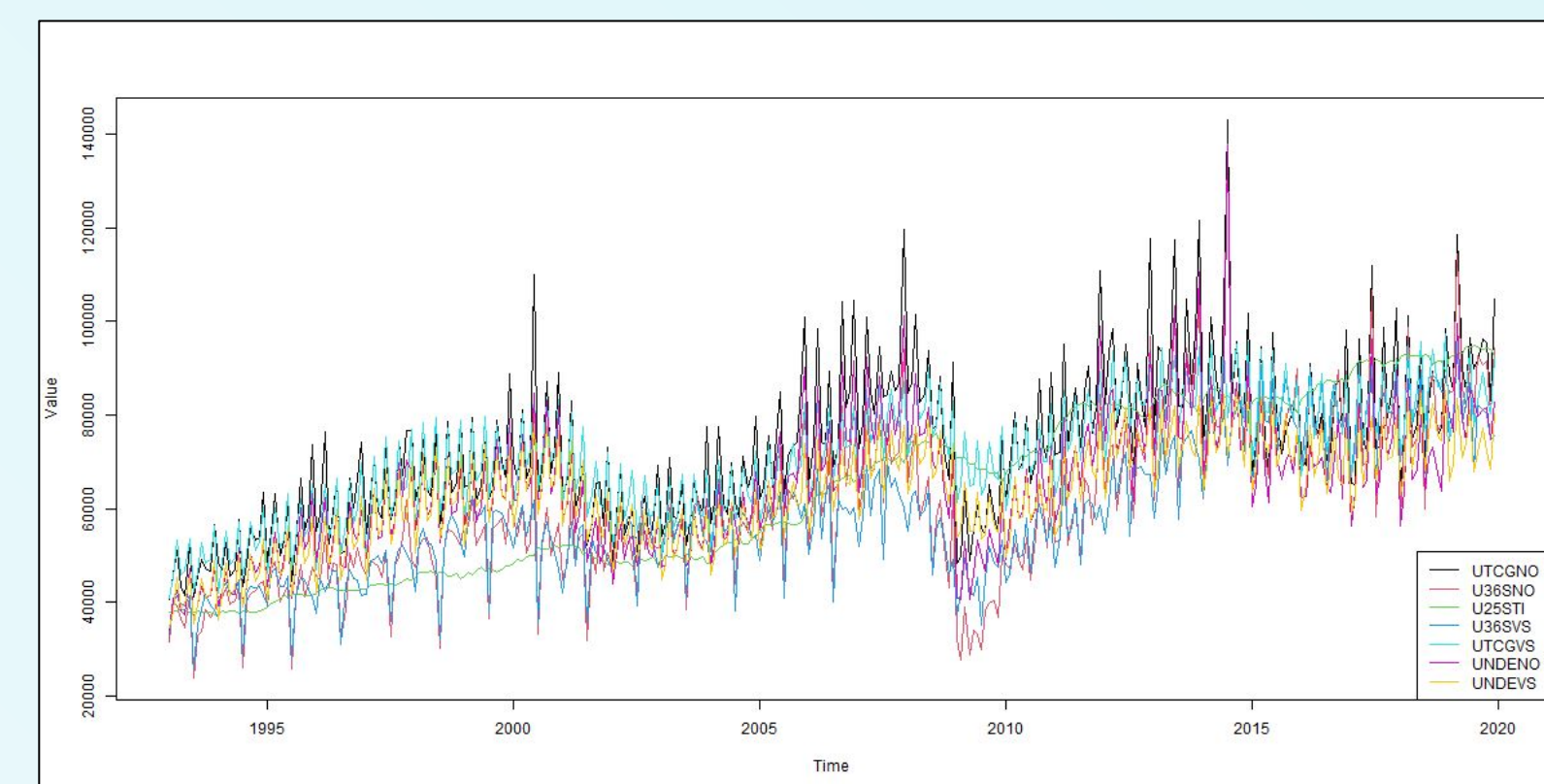


Figure 4. Includes: capital goods new orders (UTCENO), transportation equipment new orders (U36SNO), chemical products total inventory (U25STI), transportation equipment value of shipments (U36SVS), capital goods value of shipments (UTCGVS), and non-defense capital goods new orders/value of shipments (UNDENO/VIS) in millions of dollars from Jan. 1993 to Dec. 2019.

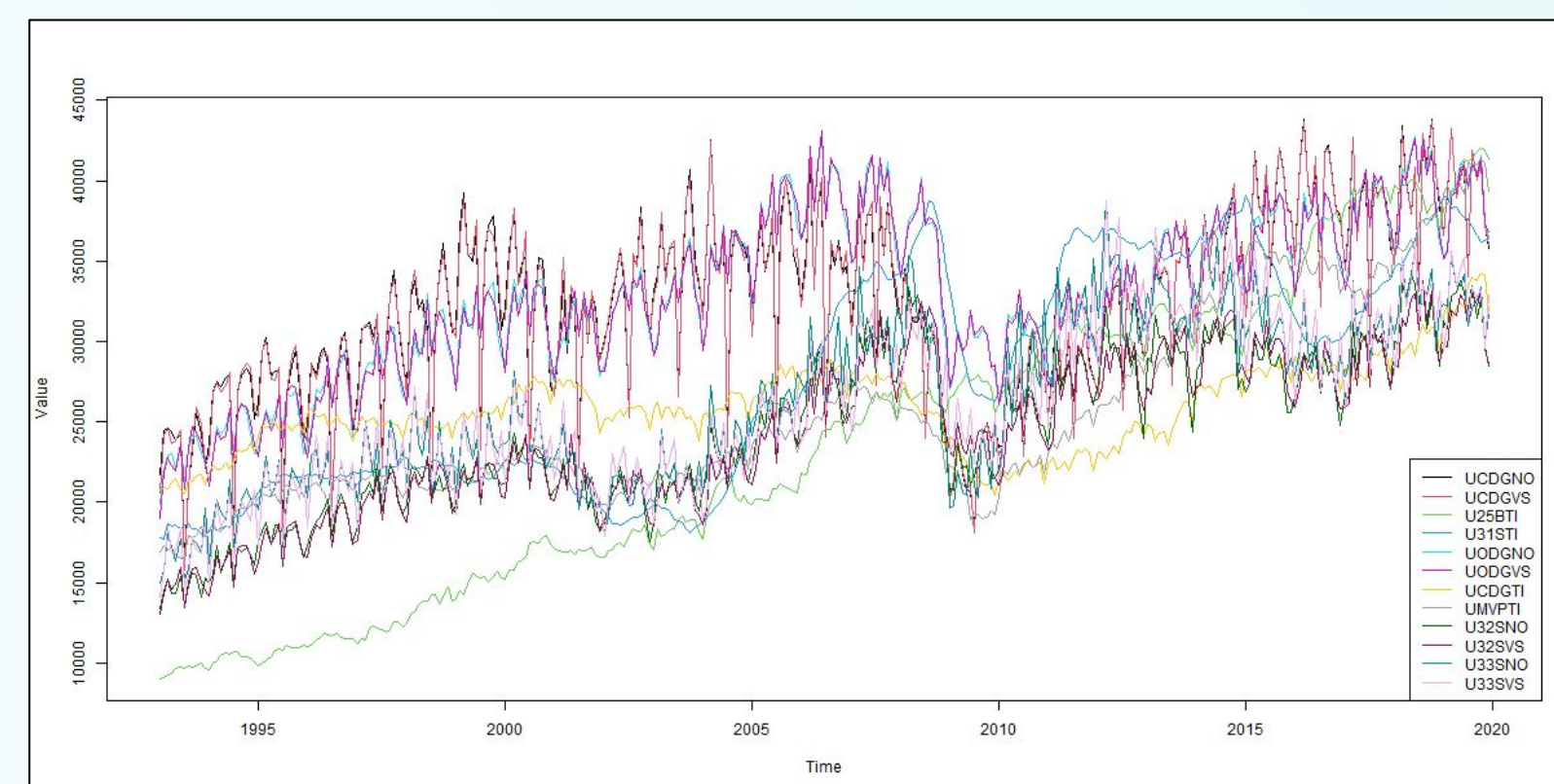


Figure 5. Includes: consumer durable goods new orders/value of shipments/total inventory (UCDENO/VIS/TI), pharmaceutical and medicine manufacturing total inventory (U24BTI), primary metals total inventory (U31STI), other durable goods new orders/value of shipments (UODGNO/VIS), motor vehicles parts total inventory (UMDVPTI), fabricated metal products new orders/value of shipments (U32SNO/VIS), and machinery new orders/value of shipments (U33SNO/VIS) in millions of dollars from Jan. 1993 to Dec. 2019.

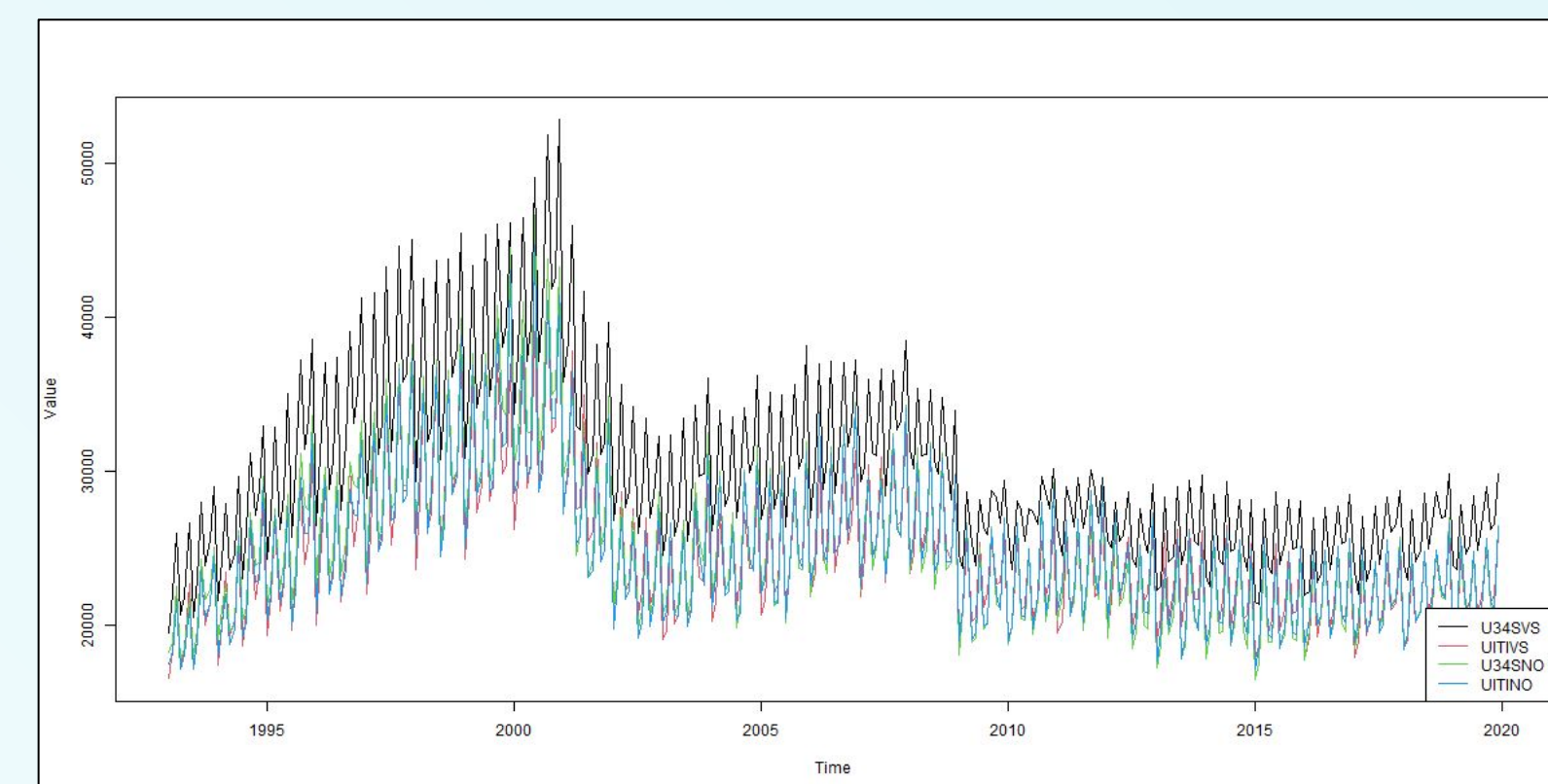


Figure 6. Includes: computer and electronic products value of shipments/new orders (U34SVS/NO), and information technology industries value of shipments/new orders (UITIVS/NO) in millions of dollars from Jan. 1993 to Dec. 2019.

Conclusion

No matter the industry being analyzed it can be seen that the suggested clusters from the DTW dendrogram provides clear and reliable information. By plotting and verifying the time series data it can be concluded that there are endless ways to inspect the data and then create relevant clusters. DTW is one sufficient method.

Future Work

Some possible areas for future work include:

- Manipulating these suggested clusters to create “stronger” associations
- Cluster using other methods besides DTW
- Compiling more time series data from different directorates
- Applying factor models on the clustered data sets to forecast.

Acknowledgments

PIC Math is a program of the Mathematical Association of America (MAA) and the Society for Industrial and Applied Mathematics (SIAM). Support is provided by the National Science Foundation (NSF grant DMS-1722275).

Special thanks to our industrial liaison Dr. James Livsey and advisor Dr. Lan Cheng for their continued support with the project. Also, thank you to Dr. H. Joseph Straight for his mentorship.

