This paper sets out to extract any business cycle information that might exist in four terabytes of weekly scanner data. This dataset has two unique features: the Great Recession of 2008 is included in the sample, and observations for both price and quantity.

The main challenge is to handle the volume, variety, and characteristics of the data within the constraints of our computing environment. While the available subsampling algorithms are computationally efficient, the sampling scheme may not be desirable from the viewpoint of economic analysis. This data also require the researcher to perform seasonal filtering, a task that is usually undertaken by government agents. The problem requires a particular type of low rank decomposition. In short, economic data have specific needs and require specific tools. Cross-disciplinary work is needed to develop improved alternatives.