word associations. By querying/filtering on “modem” in 2017, we generated 17 records of failures, the same number of failures recorded by the Logistics Center for the same part.

However, on further examination of the descriptions and the action words around the modem summary, we could differentiate between modems that were just “reset” and those which were explicitly “replaced.” Reset modems were not necessarily replaced by a new modem in inventory, so the number of modem failures recorded may have been the same demand estimated from the Logistics Center, but the actual failures of those parts is likely a subset of that total. In addition, parts that failed may not actually be broken, and the Logistics Center tests returned failed parts before restocking serviceable stock at the Depot Logistics Center. By further examining and analyzing the nomenclature and searching across like words – “reset” and “replace” – we might be able to identify additional records or we can eliminate a percentage of these failures by instead recording an aggregate amount as only modem “resets” and not inventory depletion.

7.1.3 Word Descriptions

For the final part analyzed, we searched for the legacy system timecard, also called a “clock” or “CCA” (circuit card assembly). We discovered four records in 2017, just one more than was recorded as 2017 demand using Logistics Center data alone. We generated no records using the words “timecard” and “clock,” but we generated four records, three of which were replacements when querying “CCA.” Since one of the repairs was generated in October, there is a chance that a demand as calculated from the Logistics Center had a timing delay, and one of that year’s failures might have been recorded in the subsequent year.

8 Conclusion

Sometimes in government agencies, cost estimators are presented with incomplete data sets from which they must generate forecasts and conduct predictive analyses. In major government infrastructure capital investments, the FAA cost estimating and business case team generated a methodology for conducting legacy sustainment and obsolescence analyses to estimate a system’s end-of-life and the date by which it might risk loss of service. This analysis was conducted with the premise that the demand data collected to estimate historical parts failures and from which a forecast of future failures could be generated is accurate and not clouded by inconsistent agency practices. However, when estimating parts demand at the FAA, analysts must consider that not all spare parts are returned to the Logistics Center right after failure, and the timing for demand might be skewed depending on the practices and spares inventory at certain locations.

By supplementing one incomplete source of demand data with a second source of data, the team was able to reconcile some parts demand records and improve the fidelity of sustainment analyses. By using topic analysis and other text data analysis techniques, the program office could add context to system failures and parts demand and improve sustainability forecasts.

Adding these techniques to business cases and investments that utilize incomplete data sets, cost estimators and program offices will be able to make better informed decisions about infrastructure investments and the timing of those investments.