



**AIR FORCE INSTALLATION CONTRACTING AGENCY
STRATEGIC SOURCING AND CATEGORY
MANAGEMENT THROUGH EXPENDITURE
PROFILING**

THESIS

James T. Okamoto, Capt

AFIT-ENV-MS-18-M-228

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

**DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED..**

The views expressed in this document are those of the author and do not reflect the official policy or position of the United States Air Force, the United States Department of Defense or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States

AFIT-ENV-MS-18-M-228

AIR FORCE INSTALLATION CONTRACTING AGENCY STRATEGIC
SOURCING AND CATEGORY MANAGEMENT THROUGH EXPENDITURE
PROFILING

THESIS

Presented to the Faculty
Department of Systems Engineering and Management
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Cost Analysis

James T. Okamoto, BS
Capt, USAF

22 March 2018

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED..

AIR FORCE INSTALLATION CONTRACTING AGENCY STRATEGIC SOURCING
AND CATEGORY MANAGEMENT THROUGH EXPENDITURE PROFILING

THESIS

James T. Okamoto, BS

Captain, USAF

Committee Membership:

Lt Col Brandon M. Lucas, PhD
Chair

Dr. Bradley C. Boehmke, PhD
Member

Dr. Edward D. White, PhD
Member

Abstract

This thesis examines an approach to characterizing various expenditure profiles for the Air Force Installation Contracting Agency's Operations and Maintenance Appropriated Funds. Using naive, seasonal naive, trailing moving average, exponential smoothing, linear regression, and autoregressive integrated moving average (ARIMA) forecasting methods, the paper evaluates multiple error measures over one fiscal year to find the most precise model for each level of analysis. Levels of analysis included the Air Force enterprise and level 1 category levels, as well as an illustrative approach to Information and Technology spend at the level 2 subcategory, major command, and base levels. Optimal model characteristics were used to compare expenditure profile patterns at the different levels. In general, the more a unit can customize its algorithms, the more accurately it can capture its respective expenditure profile. The more localized the level of spend, the less applicable the aggregate models become, and different sub-groups have more personalized patterns.

Acknowledgements

I would like to thank my faculty advisor, Lt Col Brandon Lucas, for his guidance and support throughout the course of this thesis effort. I would also like to convey my sincere appreciation to Dr. Bradley Boehmke, who heavily invested his time and efforts in mentoring me through the fundamentals of coding and business analytics. Finally, I am in debt to Dr. Edward White for teaching me about statistics and inspiring me to aspire to better thinking.

Contents

Abstract	v
List of Figures	ix
List of Tables	xi
I. Introduction	1
General Issue	1
Problem Statement	2
Research Questions	2
Methodology	2
Assumptions/Limitations	2
Implications	3
Summary	3
II. Literature Review	4
Category Management in the Private Sector	4
The O&M Appropriation	5
AFICA	6
Category Management in the Air Force	7
Transition	7
Spend Analysis	9
Current Analysis	11
Ecological Fallacy	12
Forecasting	13
Forecast Accuracy Measures	14
Summary	15
III. Methodology	16
Data	16
Data Exploration	17
Evaluating Method Performances	22
Benchmarks	23
Naive	23

Trailing Moving Average	24
Tunable Models	24
Exponential Smoothing	25
Regression	25
Autoregressive Integrated Moving Average (ARIMA)	26
Characterizing Expenditure Profiles	27
Summary	28
IV. Results	29
AF Enterprise Expenditure Profile	29
Level 1 Category Expenditure Profiles	30
Level 2 Subcategory Expenditure Profiles	32
MAJCOM Expenditure Profiles	36
Base Expenditure Profiles	39
Outlier Discussion	43
Summary	44
V. Conclusion	45
Research Questions Revisited	45
Research Question 1: How Can AFICA Expect Expenditure Profiles of Each Main Category to Behave at the Aggregate Level?	45
Research Question 2: How Do Expenditure Profile Behaviors of a Category Change Across Different Sub-Categories?	46
Research Question 3: How Do Expenditure Profile Behaviors Change as Categories Disaggregate Into MAJCOM and Base-Level Operations?	47
Implications	49
Limitations	49
Future Research	49
Bibliography	51
Appendix A - Decompositions	54
Appendix B - Model Optimizations	68
Appendix C - Model MASE Tables	96

Appendix D - Error Measure Comparison	99
Appendix E - Base T-Test Results	101
Appendix F - ARIMA Model Exploration	102

List of Figures

1	AFICA Organization Chart	6
2	Government Contract Spend	8
3	AFICA Support to CM	9
4	Level 1 Categories and Subcategories	9
5	Slide from 2004 RAND Spend Analysis	11
6	The Forecasting Process	14
7	Average Dollars by Month Per Transaction	18
8	Average Fiscal Year Dollars by Month Per Transaction	18
9	Total Dollars by Month	19
10	Total Fiscal Year Dollars by Month	19
11	Total Transactions by Month	20
12	Total Fiscal Year Transactions by Month	20
13	Decomposition of Total Dollars Obligated	21
14	Full ARIMA Model	27
15	Full ARIMA Model	27
16	Linear Regression Residuals	30
17	ARIMA Residuals	31
18	Optimal Level 1 Model Counts	31
19	Level 1 Model Counts	32
20	Level 1 Model Performances	33
21	Histogram of MASEs	33
22	ARIMA Residuals	34
23	Optimal IT Level 2 Model Counts	35
24	IT Level 2 Model Counts	35
25	IT Level 2 Model Performances	36
26	IT Level 2 Histogram of MASEs	37
27	Optimal MAJCOM Model Counts	37
28	MAJCOM Model Counts	38
29	MAJCOM Model Performances	38
30	MAJCOM Histogram of MASEs	39
31	Optimal Base Model Counts	40

32	Base Model Counts	41
33	Base Model Performances	41
34	Base Histogram of MASEs	42
35	Outlier Base Plots	43

List of Tables

1	AF Enterprise Model Comparison	29
2	Level 1 Model Summary	46
3	Level 2 Model Summary	47
4	MAJCOM Model Summary	48

Air Force Installation Contracting Agency Strategic Sourcing And Category Management Through Expenditure Profiling

I. Introduction

The Air Force (AF) spends over \$16 billion a year on Operations and Maintenance (O&M) contracts. Recently, the enterprise restructured the way it handles these contracts to do better business. However, there is a gap in understanding how it has historically performed in these new categories. The Air Force is seeking aid in understanding these expenditure profiles, which is the focus of this thesis. This chapter outlines the background to the the way the AF manages its O&M contracts. The defined problem statement, along with focused research questions, then follow. The chapter closes with a brief overview of assumptions, the methodology used in this research, and implications for decision makers.

General Issue

Unlike system acquisition contracts, the AF does not centrally manage installation contracts. These individual contract funds do not represent a large portion of the AF's overall budget, but the aggregated spend total is significant. In 2012, contracted O&M funds were specifically highlighted when the Government Accountability Office (GAO) reported that improved strategic sourcing could save the government over \$50 billion. Strategic sourcing was defined as a “collaborative and structured process of critically analyzing an organization’s spending and using this information to make business decisions about acquiring commodities and services more effectively and efficiently.” In 2014, the Air Force Installation Contracting Agency (AFICA) responded by categorizing installation-support spend and assigning accountable portfolio managers to these defined categories. The logic was that a single manager responsible for an array of similar spend categories would be better positioned to shape consumption and reduce enterprise spend. However, a lack of readily available data on spending trends has made it difficult to identify proper performance levels throughout the fiscal year. Thus, AFICA has not been able to assign dollar figures to each category and properly benchmark performance.

Complicating this challenge, the cyclical nature of how the AF (and other federal entities) funds activities and manages expenditures has always been a discussion within office politics but rarely scrutinized for its impacts. With sustainment funding being one-year money, expenditures typically follow an ebb-and-flow pattern. Resources are typically restrained in the first quarter (largely due to a Continuing Resolution Authority), spike in the second quarter once the President’s Budget is finalized, and drop in the third quarter

before a final large spike in the fourth quarter due to End-of-Year (EOY) spend-out.

Problem Statement

Decision makers at AFICA need to know typical trends in each category to better manage O&M spending. By identifying seasonal spending trends, this research seeks to assist AFICA decision makers project and track category performance throughout the fiscal year.

Research Questions

The purpose of this research is to answer three questions:

1. How can AFICA expect expenditure profiles of each main category to behave at the aggregate level?
2. How do expenditure profile behaviors of a category change across different sub-categories?
3. How do expenditure profile behaviors change as categories disaggregate into major command (MAJCOM) and base-level operations?

Methodology

The thesis used a statistics-based approach to understand AFICA expenditure profiles. The statistics-based approach focuses mainly on variance and the measure of error between projected and actual expenditures. The different methodologies forecasted fiscal year (FY) 2016 expenditures and were compared to the actual expenditure profile of that year as reported by AFICA. Methods that minimized each individual category and location were compared with an aggregate method to characterize expenditure profiles.

Assumptions/Limitations

The nature of forecasting and the DoD budget process guide limitations to this thesis. The foremost assumption in the thesis is that past behavior influences future behavior. Any limitations to the five-year scope of this analysis are compounded by the limitations of the data itself. According to AFICA, any given data reported may include up to 10% errors due to user input. Furthermore, the annual nature of O&M funds limits the window of opportunity to implement any spend analysis application to single-year time frames.

Implications

As AFICA establishes its new category management, this research's approach to understand expenditure profiles may be used in any and all levels of management. Creating performance metrics meant to track category spend throughout each fiscal year may be greatly inaccurate without first understanding the nature of each category's expenditure profile. Knowing how each category differs can help managers take a proactive approach to their respective spending spikes and lulls.

Summary

This chapter explained the background of AFICA's category management. Although O&M bases individually execute small budgets compared to other appropriations, the aggregate-level spend constitutes a significant sum worth investigating. Then, the researchers defined the problem statement, research questions, methodology, assumptions, and implications. With an understanding of the problem, it is next imperative to understand the context of this research.

II. Literature Review

The concept of category management is not new; the private sector uses it extensively across multiple industries and countries. It is first worth looking at how the private sector implements this practice. Then, the chapter looks at the nature of the O&M appropriation that funds the majority of AFICA's activities throughout the fiscal year. A third key concept is AFICA's hierarchy and how category management is being implemented within this hierarchy. To understand the context of how this research aims to help decision makers, it is also important to understand related research to date. When examining category management for AFICA, it is important to understand potential pitfalls and the ecological fallacy. Finally, this literature review will broadly cover the theory of forecasting used.

Category Management in the Private Sector

Although the concept of category management has mainly been practiced in private sector business for decades, its principals of aggregated efficiencies make it attractive to the government, which traditionally writes contracts at the individual organization level. Category management was introduced in the food industry in the late 1980s. Category Management is “the strategic management of spend categories using an array of tools to improve costs and achieve best-in- class category performance” (Muir, Keller, & Knight, 2014). It helped both retailers and suppliers maximize profits by focusing on product organization and the relationship between retailers and suppliers. The concept aided decision makers in applying strategic direction to grouped resources. This improvement in strategic application helped align consumer requirements with supplier capabilities. In the 1980s, there were four essential principles to implementing effective category management: strategic prerequisites, organizational structure, information and technology support, and strong retailer/supplier relationships. According to Zenor (1994), there are tangible profit benefits to aggregating product lines by these four principles, validating the practical utility of the concept.

In 1995, Blattberg further defined more key strategies in implementing category management. He stated that the size of a category is not as important as identifying opportunities. To identify these opportunities, he proposed focusing efforts on several planning steps: define the category, set category roles and goals, develop strategies, and regularly monitor and review the category. These planning steps will require an organization's structure to change to fit these new priorities (Blattberg, 1995). “Through the category management process, opportunities to gain sales and/or profits are systematically identified and acted upon” (Blattberg, 1995).

Blattberg's assertions regarding category management have been supported by follow-on research. In 2000, Gruen and Shah found that, across implementation in these different areas, one of the main drivers

behind the success of category management was a retailer's trust in the new process. They also found that the concept helps decision makers increase total dollar sales by focusing on price, variety, brand, and promotions to increase customer buys (Gruen & Shah, 2000). A year later in 2001, Dhar, Hoch, and Kumar found that the specific role of a category also drives management's ability to identify opportunities to improve business practices. AFICA's new concept, then, is not a one-size-fits-all approach but must be molded to the organization's nature and current expenditures.

Category management can aid decision makers tremendously, but it has its limitations. The performance of each category, for example, largely depends on the makeup of the category. Consolidating products or services into a single portfolio does not enhance productivity by itself. The theory is enticing, but it does not guarantee better business for the Air Force without wise application.

The O&M Appropriation

As proposed in the 2014 Concept of Operations for Category Management in the Air Force, category cost forecasts could be integrated into an installation-support process. Because these category expenditure profiles are the focus of this research, it is key to first understand the O&M funds that are used to cover these installation costs. The O&M appropriation funds the cost of operating and maintaining equipment at a state of readiness. Although individual bases use a host of appropriations to pay for expenditures, the O&M appropriation dominates the majority of these costs and represents a realistic scope of leaders' decision authority.

The Budget Control Act of 2011 aimed to encourage Congress to decrease federal spending by setting ambitious budget cut goals for the next decade. When Congress failed to meet these goals, sequestration was enacted, and the federal budget was cut across major categories of spend. Although the Bipartisan Budget Act of 2013 and 2015 modified the Budget Control Act of 2011 to raise budget caps, the government, including the Air Force's O&M appropriation, suffered severely in FY 2013. Because an annual appropriation, O&M is arguably one of the most sensitive to the fiscal season of the Air Force's spend; next-year's funds are never guaranteed.

Because of its annual rhythms, the O&M appropriation is one of the most dynamic of the federal government. Its spending is riddled with policies that restrict management's spending options. Apportionment, for example, trickles funds down to bases by handing out authority each quarter rather than all at once. This policy aims at preventing individual installations from spending all of their authority too quickly. The 80-20 rule requires 80% of total budgetary authority to be executed by 31 July, further restricting management's

Category Management in the Air Force

O&M funds may be spent by contracts, government purchase cards, or a number of other transaction vehicles. In FY 2014, the federal government spent over \$428 billion in contract spend, which accounted for about 12% of the entire federal budget (“Government-wide CM Guidance_May 2015,” 2015). This amount far exceeds that of many Fortune 100 companies, making the federal government the world’s largest buyer (“Government-wide CM Guidance_May 2015,” 2015). With this significant spend, the federal government has an opportunity to leverage its buying power with industry to maximize the benefits and use of taxpayer dollars (“Government-wide CM Guidance_May 2015,” 2015).

Transition

In May 2005, the Office of Management and Budget (OMB), Office of Federal Procurement Policy (OFPP) released a memorandum announcing the introduction of strategic sourcing as a government-wide requirement for all federal agencies (“Government-wide CM Guidance_May 2015,” 2015). Although guidance from OFPP has required strategic sourcing of goods and services within all federal agencies since May 2005, a more organized, systematic, and collaborative approach to strategic sourcing across the entire federal government was critical for maximizing value for spend (“Government-wide CM Guidance_May 2015,” 2015). Currently, there are more than 3,300 contracting units across the federal government with minimal sharing of information and best practices (“Government-wide CM Guidance_May 2015,” 2015). This degree of fragmentation and lack of coordination drives costly redundancies and inefficiencies in procurement actions, contracting vehicles, and overall acquisition efforts (“Government-wide CM Guidance_May 2015,” 2015). In addition to duplication and unnecessary complexity, there is also failure to leverage spend and knowledge (“Government-wide CM Guidance_May 2015,” 2015). There are huge price differences for the exact same item — sometimes as much as a 300% price difference (“Government-wide CM Guidance_May 2015,” 2015). As agencies continue to see decreases of acquisition personnel, the loss of subject matter expertise and gaps in data and information transfer may exacerbate these inefficiencies (“Government-wide CM Guidance_May 2015,” 2015).

In December 2014, OFPP released the memo “Transforming the Marketplace: Simplifying Federal Procurement to Improve Performance, Drive Innovation, and Increase Savings,” which cites a critical need for the federal government to shift from managing purchases and price individually across thousands of procurement units to managing entire categories of common spend and total cost through category management (“Government-wide CM Guidance_May 2015,” 2015). Category management aims to help

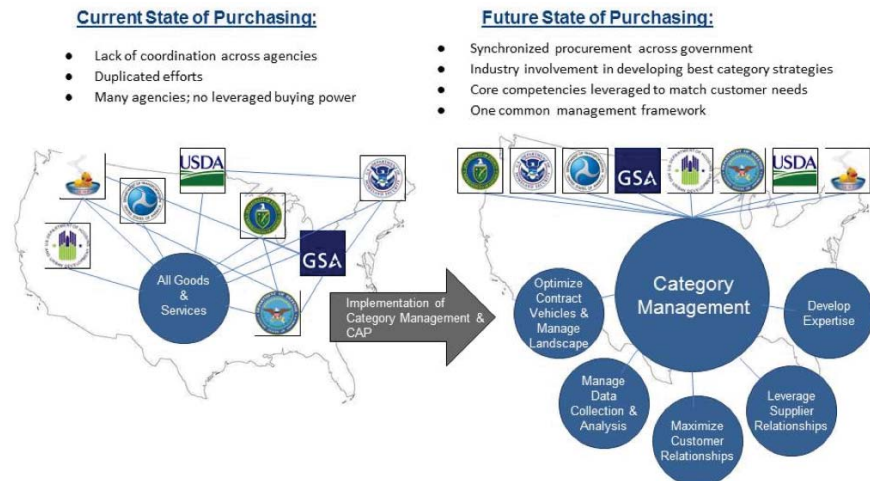


Figure 2: Government Contract Spend

the federal government act as one, buy as one, and reap the benefits as one by aggregating volumes of commonly purchased goods and services to achieve best-in-class pricing, developing common specifications and service levels, expanding supplier relationships, and leveraging shared solutions (“Government-wide CM Guidance_May 2015,” 2015). This vision of moving the government towards better purchasing through category management can be seen in Figure 2 (“Government-wide CM Guidance_May 2015,” 2015).

The Office of Management and Budget (OMB) published a memo in 2014 emphasizing the growing need to focus efforts on strategic sourcing specifically by implementing the concept of category management toward its federal contracts (Office, 2014). Around the same time, the AF began implementing Air Force Installation and Mission Support Center (IMSC), an attempt to better manage base-level support functions through centralization at one location. The AF married the two efforts, establishing category leads at the IMSC to oversee each consolidated grouping of expense.

Although category management receives the most attention in the packaged foods industry today, it has the potential to help the AF better manage its contract funds. While retailers use the concept to focus on increasing total dollar sales, portfolio managers will focus on their respective category’s dollars obligated. Instead of aiming to increase customer buys like in the private sector, portfolio managers will use category management to focus on strategic sourcing and contract savings. Integrating category managers, called portfolio managers for the level 1 categories, is visually represented in Figure 3.

Again, the intended audience for this research is AFICA/KA, which is responsible for analyzing and interpreting contract execution data. This unit is also required to assist portfolio managers (the level 1 category managers) in analyzing portfolio and category cost information. Ultimately, then, this research aids

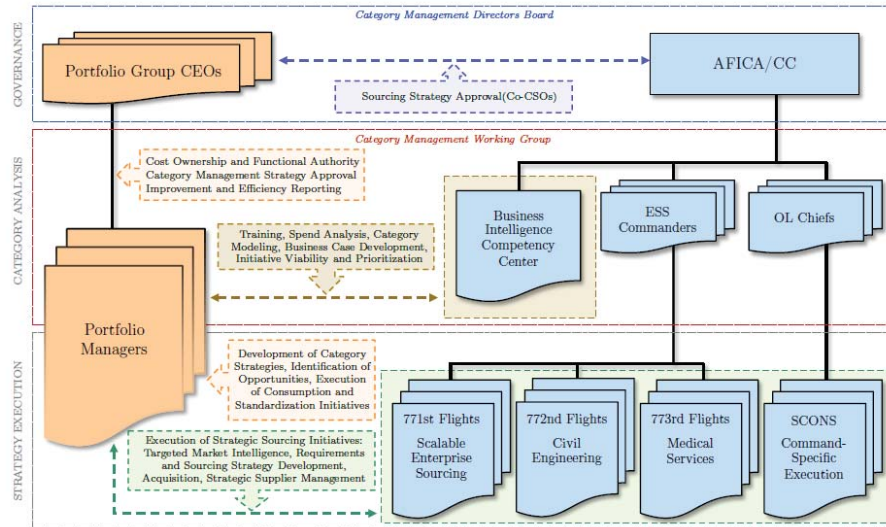


Figure 3: AFICA Support to CM

Common Government Spend Categories 1-10 (total FY 2014 spend \$275B)									
1. IT – \$49.9B	2. Professional Services – \$61.9B	3. Security and Protection – \$5.5B	4. Facilities & Construction – \$75.7B	5. Industrial Products and Services – \$10.5B					
1.1 IT Software	2.1 Business Administration Services	3.1 Security Animals & Related Services	4.1 Construction Related Materials	5.1 Machinery & Components					
1.2 IT Hardware	2.2 Legal Services	3.2 Security Systems	4.2 Construction Related Services	5.2 Equipment					
1.3 IT Consulting	2.3 Management Advisory Services (Excludes R&D 17.0)	3.3 Security Services	4.3 Facility Related Materials	5.3 Hardware & Tools					
1.4 IT Security	2.4 Marketing and Distribution		4.4 Facility Related Services	5.4 Test & Measurement Supplies					
1.5 IT Outsourcing	2.5 Public Relations and Professional Communications Services		4.5 Facilities Purchase & Lease	5.5 Industrial Products					
1.6 Telecommunications	2.6 Real Estate Services			5.6 Install/Maintenance/Repair/Rebuild					
	2.7 Trade Policy and Services			5.7 Basic Materials					
	2.8 Technical and Engineering Services (non-IT) (Excludes 1.0)			5.7 Oils, Lubricants, and Waxes					
	2.9 Financial Services								
	2.10 Social Services								
6. Office Management – \$1.9B	7. Transportation and Logistics Services – \$26.8B	8. Travel and Lodging – \$2.7B	9. Human Capital – \$4.1B	10. Medical – \$36.0B					
6.1 Office Management Products	7.1 Package Delivery & Packaging	8.1 Passenger Travel	9.1 Alternative Educational Systems	10.1 Drugs and Pharmaceutical Products					
6.2 Office Management Services	7.2 Logistics Support Services	8.2 Lodging	9.2 Educational Facilities	10.2 Medical Equipment & Accessories & Supplies					
6.3 Furniture	7.3 Logistics Civil Augmentation Program	8.3 Travel Agent & Misc. Services	9.3 Educational Institutions	10.3 Healthcare Services					
	7.4 Transportation of Things		9.4 Specialized Educational Services						
	7.5 Motor Vehicles (non-combat)		9.5 Vocational Training						
	7.6 Transportation Equipment		9.6 Human Resources Services						
	7.7 Fuels								

Figure 4: Level 1 Categories and Subcategories

both AFICA/KA and portfolio managers in understanding expenditure profile trends.

As of 2016, category managers for all of the 10 common government categories were appointed. A breakout of the 10 categories (level 1) along with their respective subcategories (level 2) are shown in Figure 4. The “Government-Wide CM Guidance” of 2015 cited this structuring shift as a multi-year roll-out process. Templates for category strategies and other pertinent documents were still in draft at the time. Correspondence with AFICA/KA revealed no spend analysis like this research has been accomplished to date.

Spend Analysis

Part of AFICA/KA’s assistance includes spend analysis, a high priority for most procurement organizations today (Muir, William A.; Keller, Richard S.; Knight, 2014). Spend analysis is a process for analyzing and interpreting past expenditure patterns (Pandit & Marmanis, 2008). It provides expenditure visibility and aids with its control (Pandit & Marmanis, 2008). Spend analysis is a high priority for most procurement

organizations (Limberakis, 2012). One of the simplest aims of spend analysis is to help leaders identify top spend categories and potential areas for savings (Limberakis, 2012).

In a 2003 report, the GAO identified that spend analysis could save the DoD and other federal agencies up to \$50 billion. It studied five private companies, including IBM, Chevron, and Delta Airlines, all of which reported billions of dollars in savings since implementing company-wide spend analysis strategies. Although the report admitted the difficulties the DoD faced in adopting such practices, such as its large and complex range of fragmented expenditure data across multiple information systems, the private companies that successfully implemented spend analyses overcame arguably equally difficult challenges (U. S. G. A. Office, 2003a). Based upon the DoD's state in 2003, the GAO report detailed recommendations to move toward spend analysis capabilities, mostly highlighting efforts in improving information systems and accounting practices.

Half a year after that report was published, the GAO released another report focusing on DoD contract spend (U. S. G. A. Office, 2003b). While the report found that each military department had management structures in place to review contract expenditures at the aggregate level, there were major flaws; for example, while the structures allowed periodic review of spend, none of them were conducive to identifying opportunities for savings. The report encouraged the DoD to focus on strategies for procurement across multiple departments. The DoD concurred with GAO findings and continued to work on its contracting management structures to incorporate spend analysis while developing multi-organization procurement strategies (U. S. G. A. Office, 2003b). These efforts were the beginnings of the federal-level category management transition that shaped the context of this research.

In 2004, the GAO broadened its focus outside the DoD, seeing spend analysis as key to taking a strategic approach to procurement at the federal level; it focused on reviewing management practices of five different federal agencies. The GAO recommended the three federal agencies that used spend analysis at the time — Veterans Affairs, Health and Human Services, and Agriculture — continue efforts to improve those practices; namely, the GAO encouraged executives to emulate the best practices of leading private sector companies and ensure supporting structures, processes, and efforts to use spend analysis to improve decision makers' knowledge and identify opportunities for leveraged buying (U. S. G. A. Office, 2004). At this point, the ground was set for federal agencies to adopt a new way of doing business.

The same year, the Research and Development Corporation (RAND) completed a spend analysis specifically for the AF to identify opportunities in purchasing and supply management. They conducted the spend analysis on DD350 data, which was the most complete, centralized source of direct AF expenditure

Having Many Contracts in the Same Federal Supply Class Indicates Prospective Contract Leverage Targets

FSC	# Ctrx	% POs	# Ctr ID Codes	% Ctrx Sole Srce	\$M	% \$ Sole Srce	# PO Cds
Passenger air charter service	770*	97	38	<1	353	<1	2
Maintenance/other miscellaneous buildings	483	23	349	35	243	23	72
RDTE/other defense-applied research	449	<1	269	3	310	1	21
Office furniture	403	24	256	2	95	1	91
Maintenance/office buildings	387	23	302	33	196	51	74
Automated data processing equipment (ADPE) system configuration	344	33	277	19	342	4	76
Other professional services	331	39	273	37	306	23	66
Other automated data processing (ADP) and telecommunications	330	16	277	20	485	8	39
Gas turbines and jet engines aircraft and components	329	17	105	27	2,084	73	5
Airframe structural components	311	47	133	58	1,137	69	26

Source: FY02 DoD-wide DD350 data.

*745 are coded as purchase orders against master solicitations.

RAND

DB-434-AF-16 07/03

Figure 5: Slide from 2004 RAND Spend Analysis

data at the time (RAND, 2004; Moore, Cook, Grammich, & Lindenblatt, 2004). The report cited a number of potential areas of savings to the AF, including consolidating the nearly 240 purchase office codes to reduce transaction costs and consolidating localized base operating support services to improve efficiencies across bases. Because the available DD350 data at the time was limited in scope and quality for viewing aggregate-level AF expenditures, the report stated its limitations, as well as what would improve the data for future analyses. Apart from increasing the quality of what was already available at the time, RAND recommended incorporating multiple data sources for supplier, market, and internal AF requirements information. Although the 2003 GAO report found DoD to have sufficient management structures in place for incorporating spend analyses for better purchasing, the AF found otherwise as it tried to implement this report's recommendations. Figure 5 shows one of the slides from that RAND report as an example of the type of analysis done at the time.

Current Analysis

The 2014 Concept of Operations for AFICA's Category Management mentioned the benefit of spend analysis. It included brief overviews of each portfolio spend for FY 2009-FY 2013, including lower-level subcategories. Rather than providing the much-needed spend analysis, these descriptions were intended to serve as a springboard for future research.

In 2017, AFICA released a series of Prioritized Spend Profile Reports (PSPR) as the primary source of historical contract spending information for category managers. For each portfolio, these PSPRs provide initial spend analysis for both level 1 categories and their respective subcategories. These reports are vital in giving category leads an idea of what their portfolio spend profiles look like over time.

None of these reports, however, details any spend past annual totals. As previously mentioned, the single-year O&M appropriation provides the majority of AFICA's funds. Therefore, this research seeks to supplement these PSPRs by describing each portfolio's spend profile on a monthly basis.

Ecological Fallacy

Analysis to date involving DoD and AF costs enterprise-wide primarily focuses on the aggregate level. Research from the 1950s to current day has consistently supported the notion that statistical analysis involving aggregate data loses legitimacy when applied to the group's individuals, also known as ecological fallacy (Clark & Avery, 1976; Freedman, 1999; Garrett, 2003; Lubinski & Humphreys, 1996; Orcutt et al., 1968; Robinson, 1950). Analysis in the AF supports this concept as well; in 2015, Boehmke et al. found that, specific to the AF enterprise, aggregation masked significant differences in installation cost behaviors.

In 1968, Orcutt et al. studied national survey information and demonstrated that aggregation of information up to the highest macrolevel loses a significant amount of estimation power. Freedman focused on similar 1995 national survey information to draw ecological inferences. At the national level, there appeared to be a strong correlation between foreign-born persons and literacy rates. However, running the correlation at the individual level revealed a negative correlation. He concluded aggregate data is often easier to obtain than that of individuals and that it may offer valuable insight into individual behavior. But ecological fallacies will, in consequence, continue to be made with aggregation bias. In 2010, Clark and Avery continued studying ecological fallacies in social sciences by studying varying characteristics of the Los Angeles metropolitan area, finding similar conclusions as Freedman.

Not only do inferences from aggregate patterns disrupt sound decision making for individuals, but Lubinski, Humphreys (1999), and Garrett (2002) found that statistical models differ across degrees of data aggregation as well. Robinson (2009) supported these findings and showed that it is a mistake to use ecological correlations as substitutes for individual correlations. Even quantitative data can be dangerous when inferring from the wrong level.

The concept of macrolevel patterns changing at microlevels, also known as Simpson's Paradox, further illustrates the danger of ecological fallacy (Blyth 1972). Simpson's Paradox states that individual patterns

and trends tend to either disappear or even reverse when grouped together. In 2013, Couto used Simpson's Paradox for the DoD to research training and genetic opportunities with respect to military working dogs. Boehmke et al.'s research in both 2015 and 2016 supported the existence of different cost patterns depending on the hierarchy level analyzed, further supporting the application of Simpson's Paradox in the DoD. As AFICA establishes standards and performance metrics for individual categories and lower levels of spend management, it is imperative to realize the pitfalls of looking solely to aggregate-level data to understand individual category and installation expenditure patterns. Simpson's Paradox and the ecological fallacy both suggest that AFICA's spending trends at the enterprise level will not hold for localized usage.

Forecasting

Spend analysis includes establishing baselines for categories and sourcing strategy as well as category performance measurement ("Government-wide CM Guidance_May 2015," 2015). Established baselines, or predictions, are important - nearly all Fortune 500 companies use them (Muir, William A.; Keller, Richard S.; Knight, 2014). Predicting the future, based on the type of data collected and purpose of the prediction, can either be qualitative or quantitative in nature. When possible, quantitative forecasting is preferred because it usually does a better job than qualitative forecasting at removing bias (Hyndman, 2012). Because historical data on AFICA's O&M contract spend is available, quantitative forecasting is used in this analysis.

Quantitative forecasting focuses on predicting the future as accurately as possible using time series data. Therefore, quantitative forecasting methods require a basic understanding of the different components of time series data. The trend is the long-term increase or decrease in the data. The seasonal pattern occurs when a time series is affected by seasonal factors, such as the time of the year or the day of the week. The level is the average value around which observations vary.

As in all data analysis, the process of forecasting begins with goal definition. Data is then collected, cleaned, and explored using visualization tools. A set of potential forecasting methods is selected, based on the nature of the data. The different methods are applied and compared in terms of forecast accuracy and other measures related to the goal to choose a "best" method (Shmueli & Lichtendahl, 2016).

Of course, the process does not end once forecasts are generated; forecasting is typically an ongoing goal. Hence, forecast accuracy is monitored, and sometimes the forecasting method is adapted to accommodate changes in the goal or the data over time. A diagram of the forecasting process is shown in Figure 6 (Shmueli & Lichtendahl, 2016).

The two sets of arrows indicate that parts of the process are iterative. For instance, once the series is



Figure 6: The Forecasting Process

explored, one might determine that the series cannot achieve the required goal, leading to the collection of new or supplementary data. Another iterative process takes place when applying a forecasting method and evaluating its performance. The evaluation often leads to tweaking or adapting the method, or even trying out other methods (Shmueli & Lichtendahl, 2016).

Forecast Accuracy Measures

Many measures of forecast accuracy exist, and there is no shortage of recommendations about which measure should be used when comparing different forecasting methods. Accuracy studies go as far back as 1969 and are highlighted most notably by the Makridakis competitions, or M-competitions (Makridakis et al., 1982). Three separate occasions in 1982, 1993, and 2000 each included over 15 forecasting methods, a multitude of different time series data, multiple time horizons, and a number of subject matter experts. These competitions aimed to be exhaustive empirical studies regarding the accuracy of different methods applied to different types of time series data (Makridakis & Hibon, 2000).

All studies used multiple accuracy measures and found four overall findings (Makridakis & Hibon, 2000):

1. Complicated methods do not necessarily produce more accurate forecasts than simple ones
2. The accuracy measure used to compare methods makes a difference in the relative rankings of methods
3. Combining different forecast methods generally improves forecast accuracy when compared to using individual methods
4. Forecast accuracy largely depends on the forecast's time horizon.

These findings from the M-competitions impact this research with the following effects (Makridakis & Hibon, 2000):

1. If methods are relatively close in accuracy, the simpler method would be preferred
2. Multiple accuracy methods will be used to compare methods used in this research to validate findings
3. Methods will be combined where possible in the hopes of improving forecast accuracy
4. Forecasts will be limited to one fiscal year to maintain maximum forecast accuracy.

In 2006, Hyndman and Koehler took an objective look at the M-competitions and each accuracy measure used to compare forecasting methods. They concluded that, although many textbooks and the M-competition mainly used the Mean Absolute Percentage Error (MAPE) measurement, this MAPE only holds true when the data is all positive and far from zero (Hyndman & Koehler, 2006). When using a single data set, they found the Mean Absolute Error (MAE) measure sufficient and even preferred due to its simplicity.

Summary

The research aims to improve control of category expenditure profiles under AFICA by both describing historical spend patterns throughout the fiscal year, as well as prescribing forecasting models that yield confidence intervals of predicted spend patterns in the upcoming fiscal year. Several factors play into these results. First, the federal government must adopt the concept of category management from the private sector. The nature of the O&M appropriation that funds the majority of AFICA activities also affects the extent to which category managers can control spending throughout the fiscal year. Specific to the AF, this research's recommendation will be tailored to the way category management is implemented into AFICA's organizational hierarchy. Its aim is to supplement current analysis already published in the PSPRs for each level 1 category. While doing so, it is imperative to avoid the ecological fallacy of applying statistical inferences from aggregate analysis to lower levels of spend. Finally, this analysis borrows heavily from forecasting methodology.

III. Methodology

Chapter 3 seeks to characterize expenditure profiles at the main level 1 category level and disaggregated level 2 subcategory, MAJCOM, and base levels. The chapter begins by defining and exploring the data. It then explains the six different forecasting methods before describing how those methodologies were compared. Next, the chapter maps out the process for applying the modeling methodology to each level to characterize and compare expenditure profiles.

Data

AFICA provided this data from the Air Force Business Intelligence Tool (AFBIT), a tool that allows users to retrieve spend information and details about any commodity or service the AF purchases. It was developed by AFICA/KA to aid in category management. The original data set pulled on 17 April 2017 has 148 variables; 729,660 observations; and 22,953,255 missing values; it is in a comma-separated text file format and includes several federal agencies and appropriation lines of accounting other than O&M. Because this analysis solely addresses the AF's O&M appropriation, all data was filtered to show expenditure observations with this line of accounting, leaving 334,997 observations from the original data set and 11,941 missing values (approximately 6% of the filtered data).

In order to answer the previously mentioned research questions, only eight of the original 148 variables are required. A supplemental spreadsheet provided by AFICA titled "AF DODAAC MAPPING" matches the "Contracting Office ID" to a base and MAJCOM name. Per AFICA guidance, it is assumed that the contracting office would be the same location as the funding office for the majority of O&M contracts.

The "Signed Date" field was changed to a date figure and then to a single numeric value representing the month in which each observation's expenditure was made. These calendar months were reordered to follow the fiscal year pattern (i.e., October = 1, September = 12, etc.). This allows for comparisons of expenditure patterns each fiscal year.

Values in the "Dollars Obligated" column range from -\$59,890,987 to \$391,048,312 with a mean of \$268,184.50. The "Signed Date" column ranges from 2011-10-01 to 2017-03-31, and the "Fiscal Year" column ranges from 2012 to 2017. Per AFICA, although the data contains part of FY 17 data, it will be omitted in analysis because of policy changes that render this latest data inaccurate. There are no missing values in the data set.

Data Exploration

Before analysis, the researchers looked to validate the data. Almost 25% of observations have a “Dollars Obligated” value of \$0, which would provide little insight into actual expenditure profiles and will be omitted in further analysis. Also, approximately 10% of the data represents actions taken with negative “Dollars Obligated” values. These observations, for the most part, represent legitimate deobligations and will be kept in further analysis.

Although the “Unique Transaction ID” variable should represent a single observation, there are many such values with multiple observations. By the recommendation of AFICA, “Unique Transaction ID” was combined with the “Instrument Number” field of the original data set to see if this improved the uniqueness of each observation. This did not significantly reduce the number of repeated unique id observations; approximately 10% of the data remained without unique identification. This 10% was explained by AFICA as either administrative error or AF policy changes that occurred for FY 17 that affect the unique identifiers.

Because bases seldom obligate dollars every day in every category of spend, the data was aggregated to a monthly basis. For time series analysis, this yielded 60 time periods of observation (12 months per year for five years). The general idea that spending spikes at the end of the fiscal year is visually tested by plotting average dollars spent by month across all fiscal years of the data. Figure 7 and Figure 8, grouped together on the following page, represent these average expenditures from 2012 to 2017 by month — the second graph breaks these expenditures out by each fiscal year. Spending appears to spike, on average, around the end of the first quarter (December). There is a small spike in September at the end of the year but not nearly as much of a spike as expected. Visually, it appears there could be some relationship between spending and the fiscal seasons — spikes of varying sizes depending on the fiscal year occur around the end of each quarter. Although the average spend by month does not show end-of-year spikes as expected, the bar graphs grouped together in Figures 9 and 10, showing total spend per month, paint a picture more in line with expectations. This contrast shows that, while the AF spends more money per transaction on average in December and March, it consistently spends more total dollars in September.

The contrasts between average and total spend per month shown previously suggests that more transactions occur in the month of September than other months. The bar graphs grouped together in Figures 11 and 12 support this hypothesis, showing that the AF consistently completes more transactions in September and October than other months during the fiscal year. Knowing workload increases in these two months may benefit AFICA leaders with respect to manpower decisions.

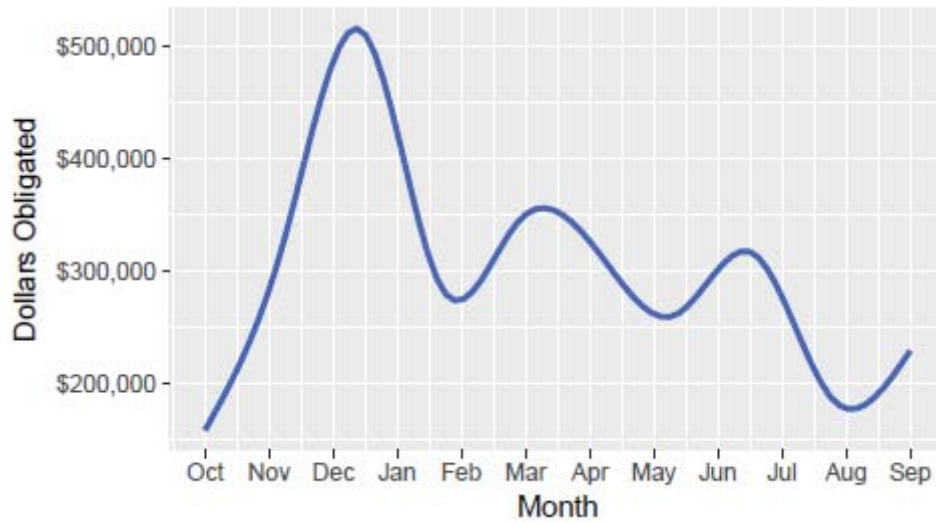


Figure 7: Average Dollars by Month Per Transaction

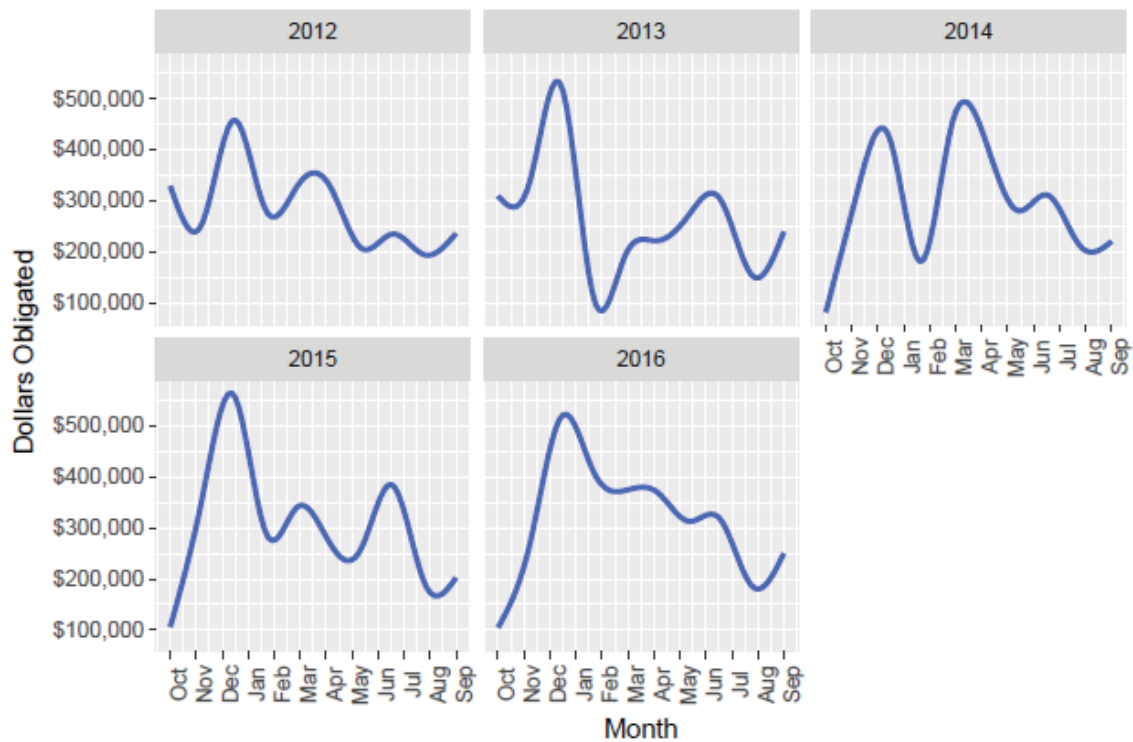


Figure 8: Average Fiscal Year Dollars by Month Per Transaction

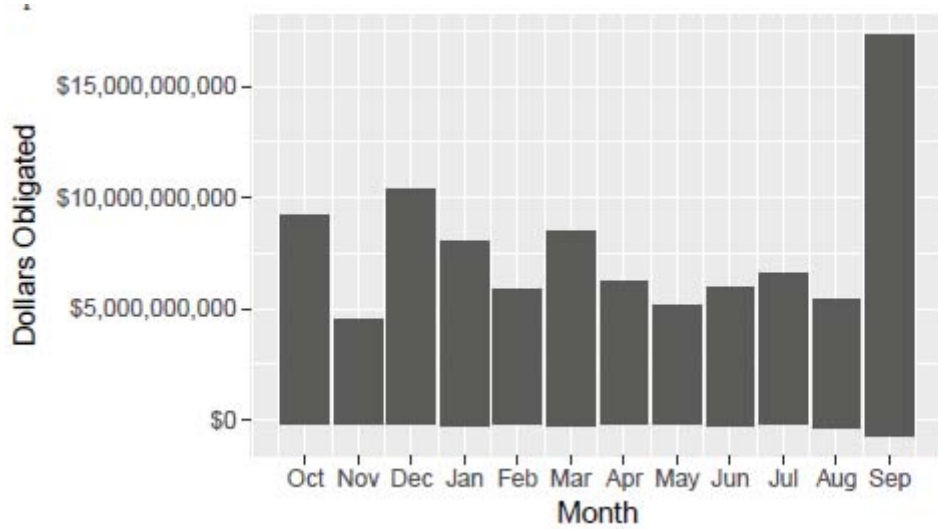


Figure 9: Total Dollars by Month

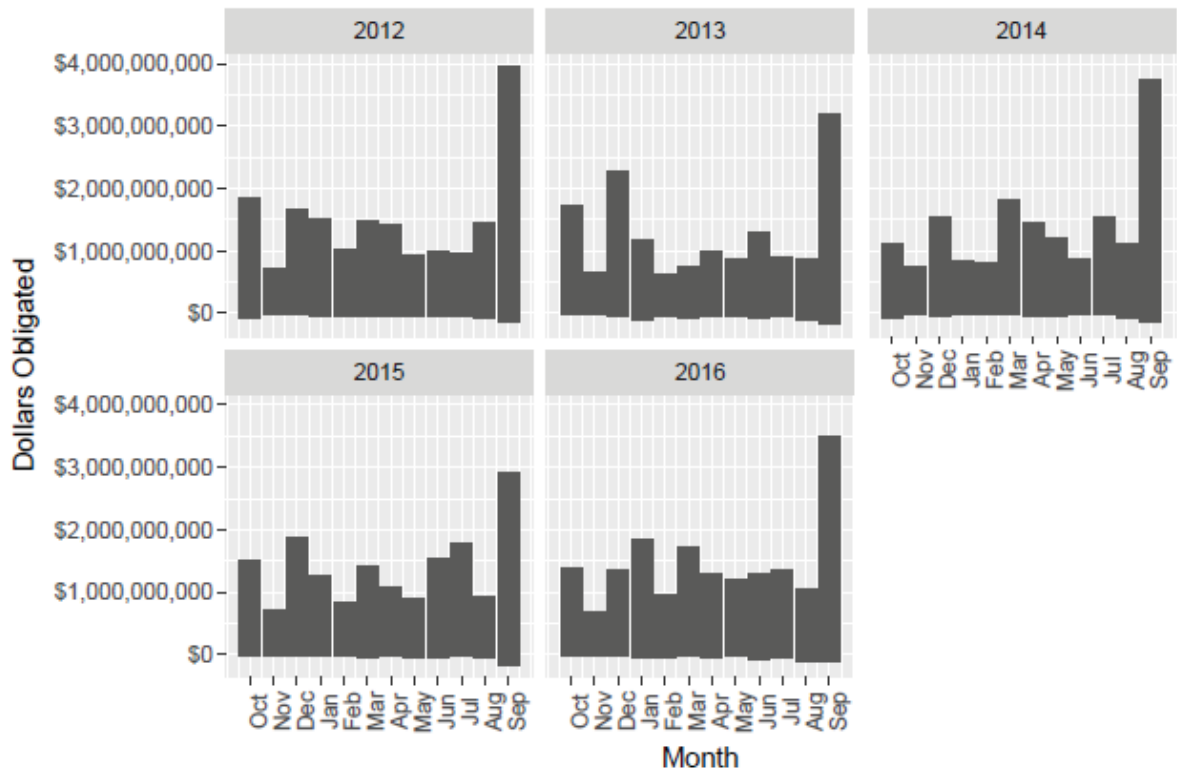


Figure 10: Total Fiscal Year Dollars by Month

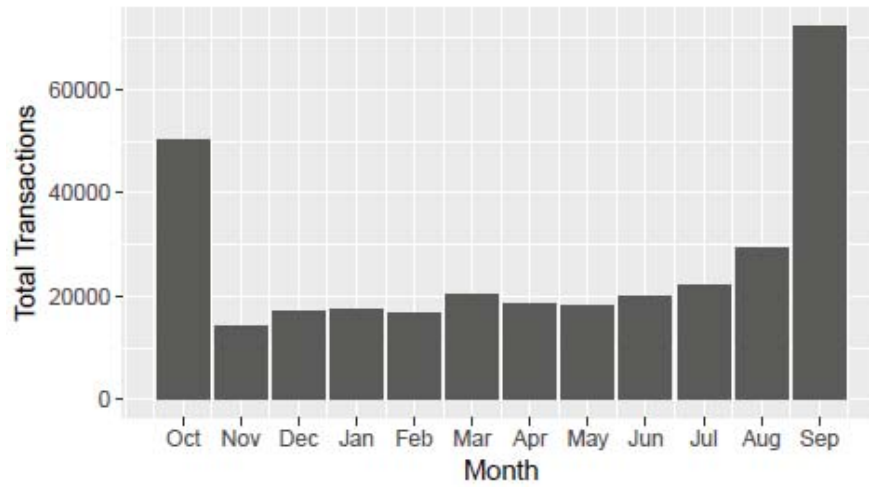


Figure 11: Total Transactions by Month

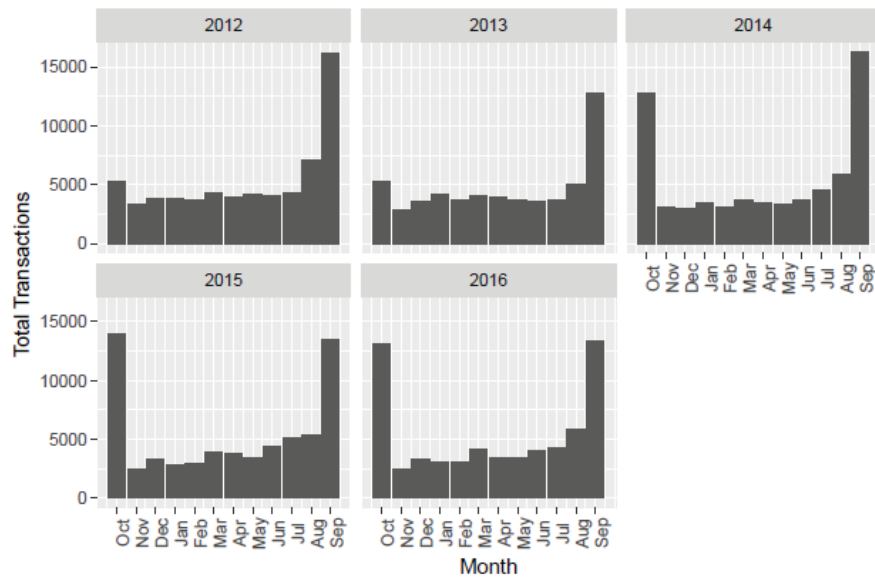


Figure 12: Total Fiscal Year Transactions by Month

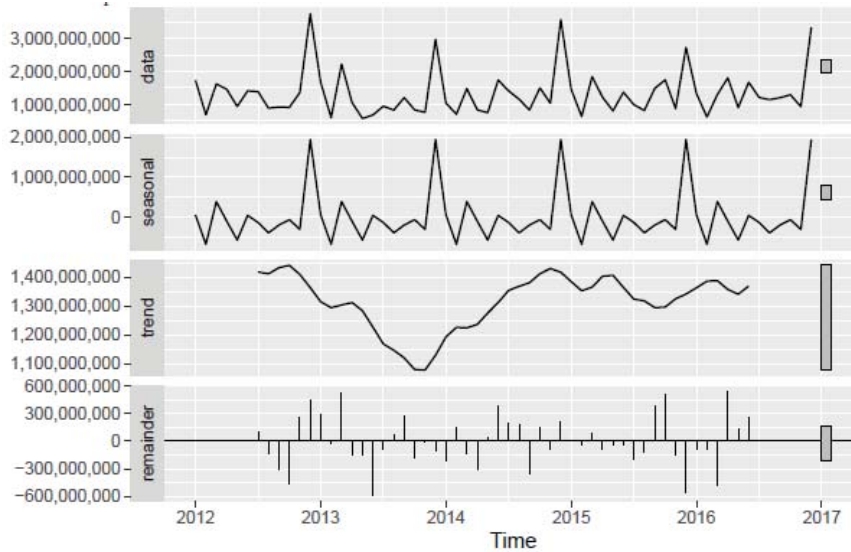


Figure 13: Decomposition of Total Dollars Obligated

After looking at fiscal year patterns at the aggregate level of spend, it can be useful to dive into the level 1 categories that constitute AFICA’s O&M spending. Per AFICA guidance, only 10 of the 19 level 1 categories will be analyzed as the “Common Government Spend Categories.” Approximately 88% of spend over the past five years can be attributed to three categories: Professional Services, Facilities & Construction, and IT. These same three categories account for nearly 95% of the total number of transactions. In general, two or three level 2 categories in each aggregated level 1 category account for the majority of spend. This can aid AFICA leaders focus efforts on the accounts of expense that make a difference.

Total dollars obligated was analyzed for seasonality and trend. A seasonal pattern occurs when a time series is affected by factors such as the time of the year or the day of the week. The trend is the long-term increase or decrease in the data. The decomposition in Figure 13 confirms small seasonal spikes in each quarter followed by larger spikes at the end of each fiscal year. There may be a slight increasing trend overall, but these hints are overshadowed by the large dip in FY 13 spending; there is no obvious long-term trend in overall spending from FY 12-16. There is no apparent pattern with noise, suggesting that all systematic patterns have been captured.

While conducting analysis, missing data was discovered at the MAJCOM level, with the AF Operational Test and Evaluation Center (AFOTEC) and the AF Reserve Command (AFRES) missing monthly entries. No clear time period with consistently higher missing entries than other time periods were found. Although AFRES shows missing data further in the past, there are a couple missing points in recent years as well. Like AFOTEC, this MAJCOM does not display consistent patterns of missing data. Both AFOTEC and

AFRES were removed from the data set, leaving 314,120 of the 334,910 data points remaining. Some of the rarer MAJCOM categories were also ignored (“NULL”, “SAF”, and “NA” MAJCOM values), leaving 11 MAJCOMs to analyze. This excluded 87 data points.

There are 14 bases from ACC, AETC, USAFE, AFMC, and AFSPC MAJCOMs with missing entries. Like the MAJCOMs with missing data, these 14 bases were removed from the data in further analysis, decreasing the total observation count from 314,033 to 305,269. Base entries were then consolidated; “NULL” and “NA” entries were both changed to “None” and excluded from base analysis. Some entries, such as “Wright Patterson AFB, OH” represented the same location as another entry, “Wright-Patterson AFB, OH” with slight differences in spelling. This occurred with Davis Monthan AFB, Joint-Base San Antonio, and Ramstein AB, as well.

FY 17 data was excluded from this analysis for two reasons. First, FY 17 is only partially represented; a full seasonal cycle is not available with the given data set. Second, per correspondence with AFICA/KA, a policy change in FY 17 rendered the way the FPDS database accounted for certain expenditures inaccurate; data previous to this policy change will be kept. Also, any impacts from inflation would be captured by the trend attributes in the models, so the data is left in current year dollars.

Evaluating Method Performances

In order to evaluate the performance of different forecasting methods on the data, values were partitioned into “train” and “test” sets. Observations from FY 12-15 were included in the train set, and those in FY 16 were included in the test set, yielding an 80/20 split in the data (48 and 12 monthly observations in each set, respectively). By applying methods to the first four years of data, the model performances were evaluated based on their predictive accuracies of 2016 spend compared to the actual data. In all cases, lower values are desired for each measure of predictive accuracy.

Three popular measures of predictive accuracy were used in evaluating method performance on the test set. RMSE (root mean squared error) maintains the same units as the data and represents the magnitude of the average absolute error. This measure is popular when evaluating model performance on a single set of data but cannot be used to compare model performance across data sets of different magnitudes. MAPE (mean absolute percentage error) gives a percentage score of how predictions deviate from actual values and provides a means by which model performance comparisons may be made across series of differing scales. Although it is another popular measure, MAPE tends to fluctuate heavily, even yielding infinite values, when series values approach zero. MASE (mean absolute scaled error) is an alternative to using percentage errors

when comparing forecast accuracy across series on different scales. Instead of comparing performance of a model to the actual data, this measure compares model performance to that of an average naive forecast computed on the training data. MASE will not approach infinity as series values get close to zero.

Although RMSE and MAPE, two of the most popular error measures used in practice, were recorded, MASE was the primary measure minimized in selecting optimal models for each time series. MAPE, though useful in most cases, yields infinite values for some base-level calculations. These series have true zero monthly expenditures in some cases; it is understandable that some bases would have lulls in spending certain months on certain categories. Thus, MAPE was rejected as the primary measure.

RMSE provides utility for each individual series modeled, but it has limited use when trying to compare models across different series. The purpose of this research was to understand expenditure profiles across different levels, categories, and locations. Although this research could have used RMSE to compare models for each series and then MASE to compare model performance across different series, selecting models by RMSE often yielded different results than selecting by MASE. To make it easier for the user, MASE was chosen as the primary measure of both selecting a model for each series and comparing model performance across different series.

Benchmarks

Unlike more advanced methods, benchmark methods are not flexible — they do not use parameters that are adjustable based on the data. Data feeds into set equations to create these models. Though they cannot be customized, they do provide a good standard to see if other, customizable methods actually improve forecasting performance.

Naive

The most recent information may be the most relevant for predicting the future. With this idea in mind, naive, or random walk, forecasts are simply the most recent values of a series ($F_{t+k} = y_t$). For seasonal series, the seasonal naive forecast is the value from the most recent identical season (e.g., forecast September using last September's value). Naive forecasts can be used as actual forecasts or as a baseline for evaluating predictive performance of alternative methods.

Trailing Moving Average

Trailing moving averages reduce noise and uncover patterns in data over a specified time period. A moving average of order m can be written as

$$\hat{T}_t = \frac{1}{m} \sum_{j=-k}^k y_{t+j},$$

where $m = 2k + 1$. That is, the estimate of the trend-cycle at time t is obtained by averaging values of the time series within k periods of t . Observations that are nearby in time are also likely to be close in value, and the average eliminates some of the randomness in the data, leaving a smooth trend-cycle component. This is called an m -MA, meaning a moving average of order m . The selection of a time window, then, is a balance between reducing noise and maintaining systematic trends. For forecasting purposes, a trailing moving average utilizes this concept for a time window in the past to predict a future average. For the five most recent time periods, the function of a trailing moving average is:

$$\hat{y}_{t+1} = \frac{y_{t-4} + y_{t-3} + y_{t-2} + y_{t-1} + y_t}{5},$$

where $\hat{y}_{t+1}$ represents the future response predicted. For the purpose of this research, a time window of 12 months was used for the benchmark model. Although smaller windows such as quarterly may be useful, 12 months matches the underlying nature of AFICA's O&M active lifespan.

Tunable Models

With three benchmark models used (naive, seasonal naive, and trailing moving average), three “tunable” models were also tested in optimizing models to capture expenditure profiles. These models, exponential smoothing, linear regression, and auto-regressive integrated moving average (ARIMA), all use parameters that may be adjusted to better capture individual series patterns. In theory, this ability to customize models to each series should perform better than all benchmark models for every series. However, if a particular series includes a significant amount of noise, a fixed model may outperform a custom model in practice.

All tunable models for each optimization were automated for the purpose of this research. For each of the models, there is a corresponding function in R, a language and environment for statistical computing, that estimates the best values for applicable parameters based on set algorithms. Though far from perfect, these functions are useful for the purpose of this analysis and allow optimization efficiently and effectively.

Further individual “tuning,” or adjusting, these model parameters might yield better performances, but, for the purposes of this research, models were customized solely by the automated algorithms available. This helps maintain replicability as well.

Exponential Smoothing

The first “tunable” model used was the exponential smoothing model. The theory of moving averages drives the foundation of exponential smoothing. This method uses weighted averages of past observations to predict future values, with weights decreasing exponentially as the observations get older. The more recent the observations, the higher the associated weight they are given.

The rate at which the weights decrease is controlled by the parameter α between 0 and 1. If α is small (i.e., close to 0), more weight is given to observations from the more distant past. If α is large (i.e., close to 1), more weight is given to the more recent observations.

Adding a trend to the forecast adds a β^* variable to represent the trend level of the forecast. β^* is the smoothing parameter for the trend for values between 0 and 1. Holt (1957) and Winters (1960) extended Holt’s method to capture seasonality, adding a final γ parameter to the equation to account for the season. The Holt-Winters seasonal method comprises the forecast equation and three smoothing equations — one for the level ℓ_t , one for the trend b_t , and one for the seasonal component s_t , with smoothing parameters α , β^* , and γ . h represents the number of periods for forecasting. The robust equation for exponential smoothing, then, can be expressed as:

$$\begin{aligned}\hat{y}_{t+h|t} &= \ell_t + hb_t + s_{t-m+h_m^+} \\ \ell_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \\ s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}\end{aligned}$$

Regression

Another popular “tunable” method that captures any trend, seasonality, and other patterns is based on linear regression. This model captures global trends, or ones that apply to the entire series of data (versus local trend), of any shape with functions; linear, quadratic, higher-order polynomial functions, and even trigonometric functions help explain the systematic patterns of time series. Any type of trend shape can be fit as long as it has a mathematical representation and is expected to continue in the future. Any other

component of the series, such as seasonality, must also be a global pattern and simply adds to the function. A time series with a simple linear trend may be modeled with regression in the following equation:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t,$$

where each β variable represents a distinct systematic pattern in the series and ε_t represents an error term. This method is capable of incorporating external information into the time series by including additional β variables into the equation — a capability absent in other models. Also, because regression models assume global patterns, they are capable of handling missing data and inconsistencies that might disrupt other methods. With one variable representing one patternal component each, these models are also relatively easy to interpret and translate for users.

Regression models are “tunable” not only with respect to their individual parameter estimates but also with a parameter known as lambda, λ . While not part of the model’s equation itself, λ transforms the response logarithmically when model residual values appear heteroscedastic. The λ parameter may take on any value between 0 and 1 to fully logarithmically transform response values, partially transform responses, or not transform the responses at all.

Autoregressive Integrated Moving Average (ARIMA)

Other than exponential smoothing, auto-regressive integrated moving average (ARIMA) models may be the most widely used approach to time series forecasting (Hyndman & Athanasopoulos, 2012). Rather than focusing on capturing trend and seasonality in data like exponential smoothing or linear regression, ARIMA models look to explain the relationship between a variable’s value and its surrounding values. Ordinary regression models do not account for correlation between values in different periods, which in cross-sectional data is assumed to be absent (what is called “independent observations”). Yet, in the time series context, values in neighboring periods tend to be correlated (autocorrelation).

ARIMA models involve two models — an auto-regressive model for capturing autocorrelation of past responses and a moving average model for capturing autocorrelation of the past forecast errors. The auto-regressive model or order AR(p) is shown in the following equation:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t,$$

where c is a constant and ϕ_i represent the “tunable” auto-regressive parameters. These parameters can

$$\text{ARIMA} \quad \underbrace{(p, d, q)}_{\substack{\uparrow \\ \left(\begin{array}{c} \text{Non-seasonal part} \\ \text{of the model} \end{array} \right)}} \quad \underbrace{(P, D, Q)_m}_{\substack{\uparrow \\ \left(\begin{array}{c} \text{Seasonal part} \\ \text{of the model} \end{array} \right)}}$$

Figure 14: Full ARIMA Model

$$\begin{array}{ccccccc} (1 - \phi_1 B) & (1 - \Phi_1 B^4) & (1 - B) & (1 - B^4) y_t & = & (1 + \theta_1 B) & (1 + \Theta_1 B^4) e_t. \\ \uparrow & \uparrow & \uparrow & \uparrow & & \uparrow & \uparrow \\ \left(\begin{array}{c} \text{Non-seasonal} \\ \text{AR}(1) \end{array} \right) & \left(\begin{array}{c} \text{Seasonal} \\ \text{AR}(1) \end{array} \right) & \left(\begin{array}{c} \text{Non-seasonal} \\ \text{difference} \end{array} \right) & \left(\begin{array}{c} \text{Seasonal} \\ \text{difference} \end{array} \right) & & \left(\begin{array}{c} \text{Non-seasonal} \\ \text{MA}(1) \end{array} \right) & \left(\begin{array}{c} \text{Seasonal} \\ \text{MA}(1) \end{array} \right) \end{array}$$

Figure 15: Full ARIMA Model

capture a wide array of different patterns in autocorrelation. The moving average model of order MA(q), shown here:

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q},$$

similarly uses a c constant and θ_i “tunable parameters” to capture a wide range of different patterns in forecast error autocorrelations.

Before consolidating equations into a single ARIMA model, there is another piece to the puzzle — accounting for seasonality. While p , d , and q parameters measure the order of auto-regression, degree of differencing (computing the difference between consecutive observation values), and order of moving average, respectively, the uppercase P , D , and Q parameters similarly represent those corresponding seasonal aspects to the ARIMA model. These ARIMA model parameters are displayed in the format shown in Figure 14, and Figure 15 shows the consolidated ARIMA equation. In tuning this model, p , d , q , P , D , and Q parameters can all be adjusted in integers from 0 up; typical values range from 0 to 3. The best ARIMA model while varying these parameters minimizes the Akaike Information Criterion (AIC), a measure of the information loss in a model.

Characterizing Expenditure Profiles

Understanding the different models used in optimizing each data series allowed this research to continue on to characterizing expenditure profiles. First, all six models were applied to the AF enterprise-level of

spend to find an “aggregate” level model to compare against individually optimized models at lower levels. This process was repeated for each of the ten level 1 main categories. With these models, both the aggregate and individually optimized models were applied to each level 1 category, and the MASE measure distributions were compared. It can be useful to decision makers to know which of the level 1 categories yield individually optimized models mirroring the aggregate model more so than the other six models.

Once the level 1 spend was analyzed, this same process was applied within the level 1 IT category to level 2 subcategories, MAJCOMs, and bases. In other words, the level 1 IT category was used in three different sets of comparisons: level 1 with level 2 subcategories, level 1 with MAJCOMs, and level 1 with individual bases. In these three separate deep dive cases, the “aggregate” level model referred to the model that minimized MASE for the level 1 IT category. Although this research focused on characterizing expenditure profiles within the level 1 IT category, the approach modeled may be applied to any other level 1 category/deep dive desired.

There are 10 level 1 categories, six IT level 2 subcategories, 11 MAJCOMs, and 77 bases. Because the first three levels of analysis involve few different data series, only the base level analysis included statistical tests to compare model performances. Because MASE values from both aggregate model performance and individual model performance are dependent upon the base expenditure profile to which they are applied, the test to compare differences was a paired difference t-test. The null hypothesis is that the true difference between aggregate model MASE values and individual model MASE values is 0. A p-value below 0.05 in this test would infer there is sufficient evidence to reject the null hypothesis in favor of the alternative hypothesis that the true difference is not equal to 0. This test will involve validating the two main assumptions of a paired difference t-test: The two sets of errors have equal variances and are normally distributed. If the null hypothesis were not rejected, customizing models at the base level would not benefit AFICA decision makers any more than applying an aggregate model to all would.

Summary

This thesis explores the FPDS data set from AFICA. It compared the six forecasting methodologies in order to select an optimal, MASE-minimizing model for each category of spend. These six methodologies were naive, seasonal naive, trailing moving average, exponential smoothing, linear regression, and ARIMA. The chapter then described the process of using this model-selection process to characterize and compare expenditure profiles of different levels of AFICA’s O&M.

IV. Results

This chapter shows the results of applying each methodology to each category of spend. The chapter describes both the aggregate and individual optimal models selected and MASE distribution differences at each level. A summary of the chapter briefly reviews the material covered.

AF Enterprise Expenditure Profile

Three benchmark models (naive, seasonal naive, and trailing moving average) were tested against three automated modeling methods (exponential smoothing, linear regression, and ARIMA). By MASE, an ARIMA model (ARIMA(1,0,0)(1,1,0)[12]) appears to best capture the aggregate-level expenditure profile patterns. The linear regression model, however, yields a lower RMSE.

Table 1: AF Enterprise Model Comparison

Model	RMSE	MAPE	MASE
Naive	1484433472	131.992	3.569
Seasonal Naive	362161440	20.479	0.754
Trailing Moving Avg	663188501	29.515	2.108
Exponential Smoothing	292051909	14.596	0.503
Linear Regression	279041338	14.379	0.497
ARIMA	290250847	13.917	0.485

The residuals of both give a good picture of which model better captures systematic patterns. Both residual plots, shown in Figures 16 and 17, respectively, show fairly constant variances. The auto-correlated function (ACF) plots show no obvious auto-correlations of errors, which would otherwise suggest that systematic patterns remain uncaptured by a model. Distribution of residuals does not look normal in either case. The linear regression model yields a Ljung-Box test (null hypothesis that residuals are white noise) p-value of 0.0003279 (less than 0.05), suggesting that the residuals exhibit some systematic patterns uncaptured. The ARIMA model yields a Ljung-Box test p-value of 0.2064, failing to reject the notion that residuals exhibit no autocorrelation. The ARIMA model with parameters (1,0,0)(1,1,0)[12] was selected to represent the aggregate-level expenditure profile. Manually tuning this ARIMA model with other p, d, q, P, D, and Q parameters did not significantly improve MASE performance, and the automatically-selected model was utilized moving forward. While little impact to AFICA, this aggregate model provides a baseline model

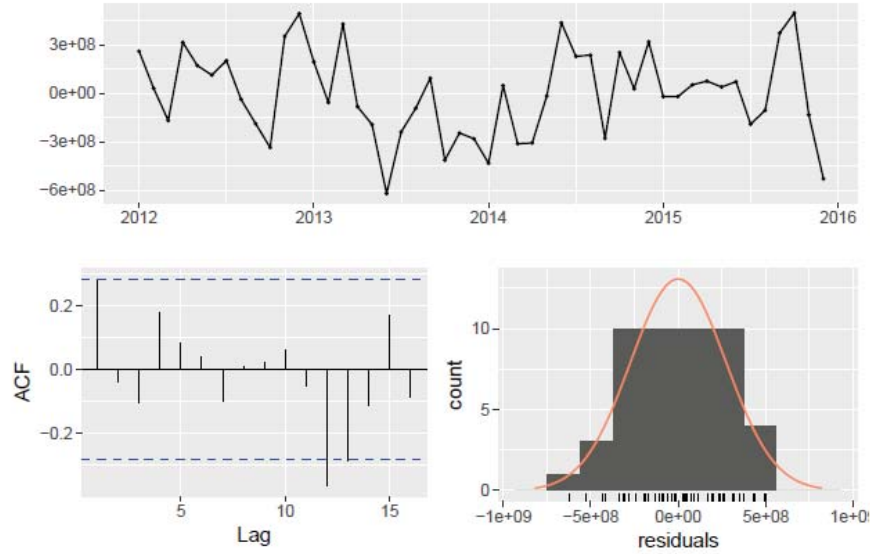


Figure 16: Linear Regression Residuals

to compare to at lower levels.

Level 1 Category Expenditure Profiles

See Appendix A for level 1 spend patterns and decompositions. The aggregate enterprise’s optimal ARIMA model was applied to each level 1 category’s spend data in addition to the six forecasting models. The model that minimized MASE out of the seven available was selected as the optimal model for each specific level 1 category. Appendix D covers some of the differences when a model is selected based on minimizing RMSE versus minimizing MASE (though the same model is selected by both measures the majority of the time).

Figure 18 illustrates the frequency of which each of the seven forecasting models (aggregate model and six original models) minimized a level 1 category’s MASE. Out of the 10 level 1 categories, the aggregate enterprise ARIMA model minimized MASE five times, or 50% of the time. The naive and exponential smoothing models minimized MASE twice each; the linear regression model minimized MASE once; and the seasonal naive, trailing moving average, and automated ARIMA models minimized MASE for none of the level 1 categories. Appendix B shows model error measures for each level 1 category.

Figure 19 groups the three benchmark models together and the four “tunable” models together in a frequency plot to show how many times a fitted model outperformed the fixed models and vice versa. Here, “tunable” models better capture expenditure profiles of level 1 categories 80% of the time. This suggests that

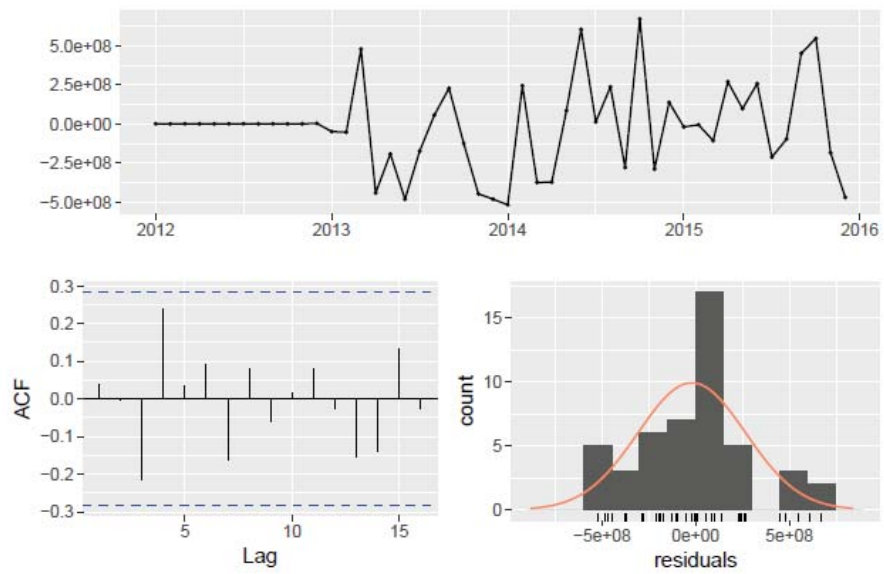


Figure 17: ARIMA Residuals

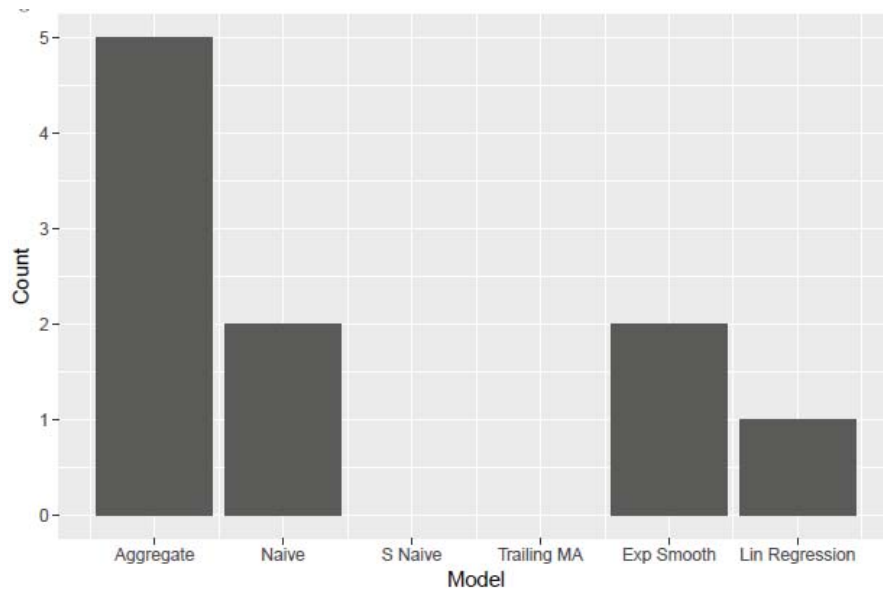


Figure 18: Optimal Level 1 Model Counts

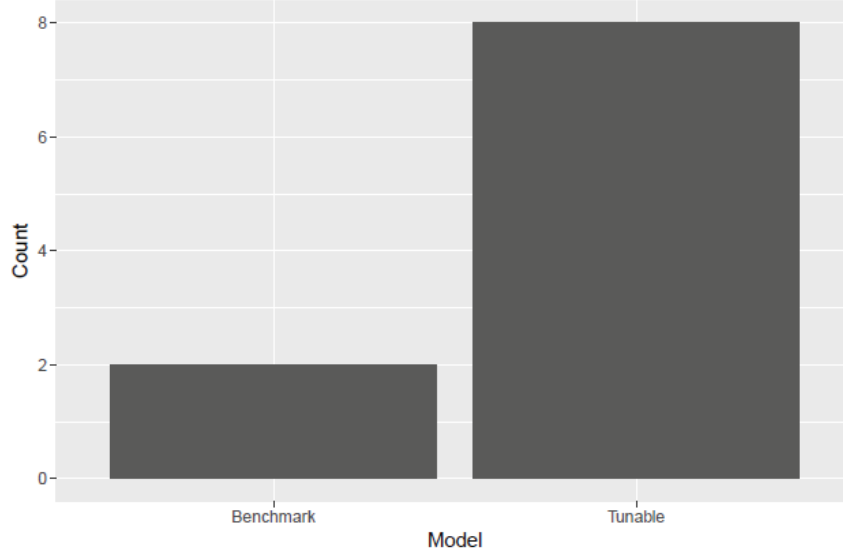


Figure 19: Level 1 Model Counts

the more level 1 category managers can customize their algorithms, the more accurately they can capture their expenditure profiles.

After visually comparing the optimal models selected for each of the 10 level 1 categories, the aggregate enterprise ARIMA model and individually optimized models were each applied to the expenditure profiles to compare the difference and thus the worth of customizing models. These MASE values may be found in Appendix C. First, the box-plots in Figure 20 show the distribution of MASE values for the aggregate model applied to each of the 10 categories next to the distribution for each customized model of the 10. Since half of the categories minimized MASE with the aggregate enterprise ARIMA model, it is understandable to see the median MASE values, represented by the thick black line, to be relatively equal. Although the individual category models show a larger range in MASE values, the distribution of MASE values appears to be slightly lower than the aggregate model's distribution of MASE values. The same distribution of MASE values is shown in Figure 21. Thirty percent of the aggregate model MASE values appear around 0.9, while the MASE values for individual category models appear to be dispersed in a wider range. Interpretation of patterns across the level 1 categories is difficult with so few data points.

Level 2 Subcategory Expenditure Profiles

To illustrate an approach that AFICA can follow for other categories, this thesis specifically focuses on the level 1 category of IT. For this particular level 1 category, the aggregate enterprise ARIMA model performed the best by MASE. Its parameters were $(1,0,0)(1,1,0)[12]$. The model's residuals in Figure 22

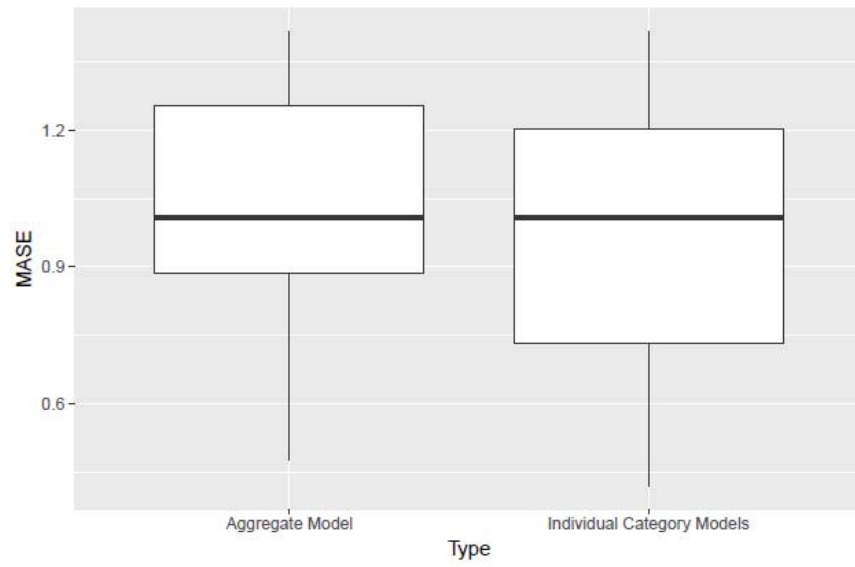


Figure 20: Level 1 Model Performances

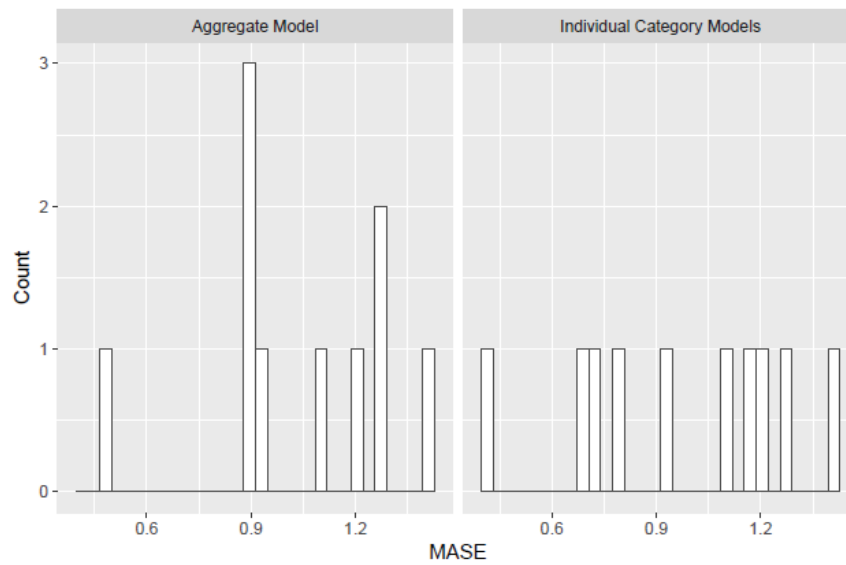


Figure 21: Histogram of MASEs

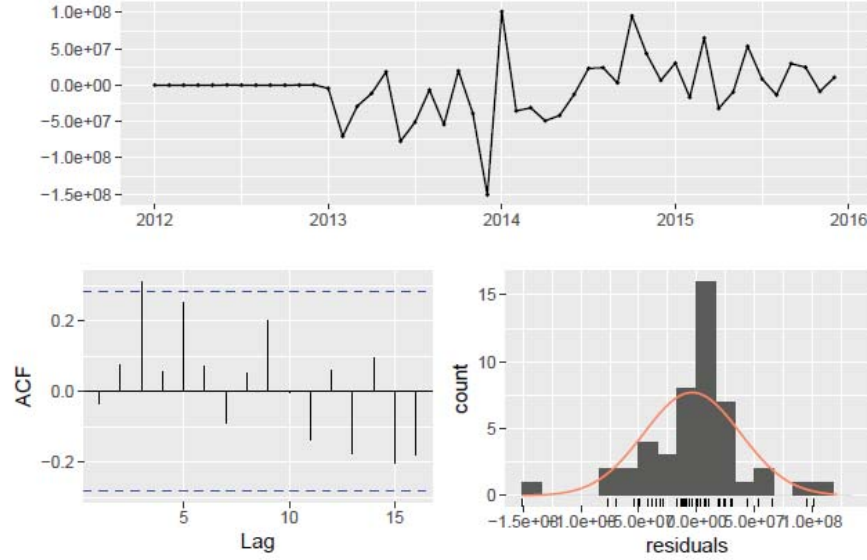


Figure 22: ARIMA Residuals

show heteroscedastic variance and suggest that some uncaptured systematic patterns remain. This means that further efforts to customize may yield a more accurate model. The residuals appear to be normally distributed, and there do not appear to be any significant autocorrelations remaining.

See Appendix A for decomposition of each level 2 subcategory within the IT level 1 category. Each of the six subcategories display annual seasonality, but trend patterns differ. Similar to the process for comparing expenditure profiles at the level 1 category level, level 2 subcategory optimal models are displayed in Figure 23. Out of the six subcategories, the aggregate ARIMA and seasonal naive models minimized MASE twice each; linear regression and automated ARIMA models minimized MASE once each; and the naive, trailing moving average, and exponential smoothing functions minimized MASE for none of the IT level 2 subcategories. Appendix B shows model error measures for each level 2 subcategory.

Figure 24 groups the three benchmark models together and the four “tunable” models together in a frequency plot to show how many times a fitted model outperformed the fixed models and vice versa. Here, “tunable” models better capture expenditure profiles of IT level 2 subcategories 67% of the time. Like comparison at level 1, this suggests that the more level 2 subcategory managers can customize their algorithms, the more accurately they can capture their respective subcategories’ expenditure profiles.

After visually comparing the optimal models selected for each of the six level 2 subcategories, the IT level 1 ARIMA model and individually optimized models were each applied to the expenditure profiles to compare the difference, and thus the worth, of customizing models. These MASE values may be found in Appendix C. First, Figure 25 shows the distribution of MASE values for the level 1 model applied to each

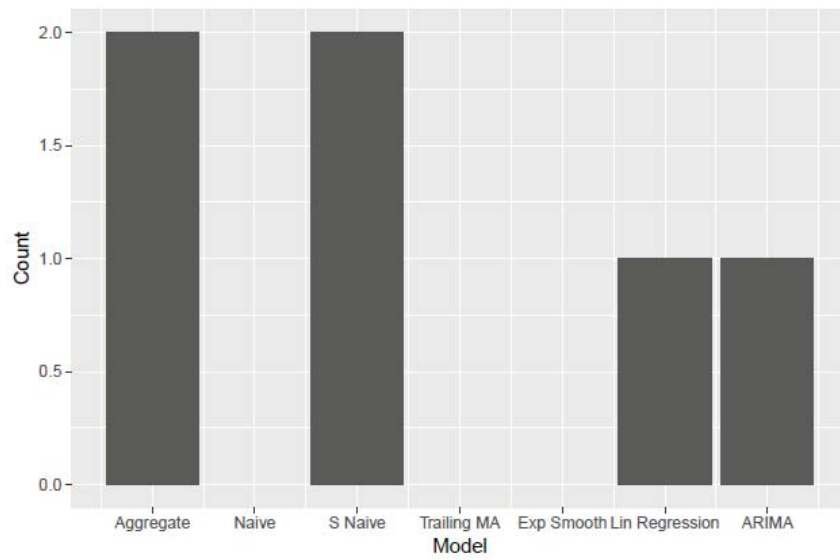


Figure 23: Optimal IT Level 2 Model Counts

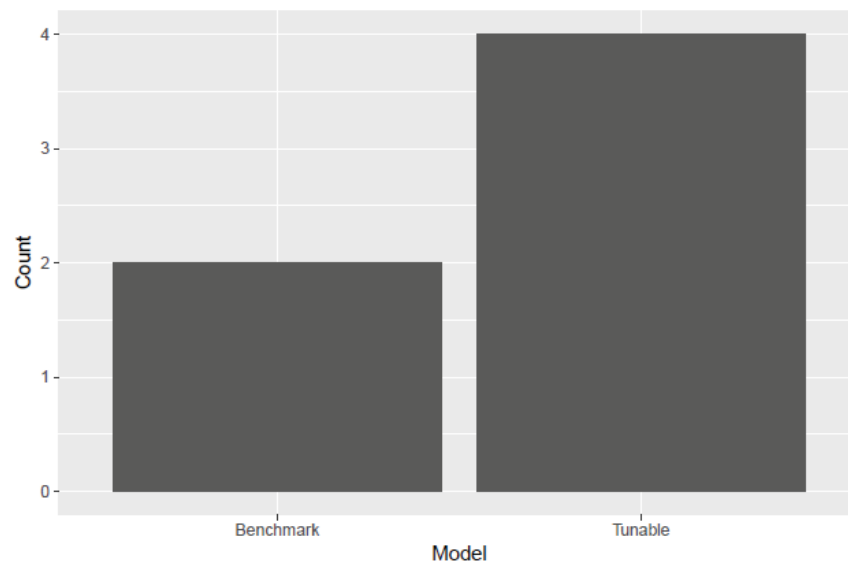


Figure 24: IT Level 2 Model Counts

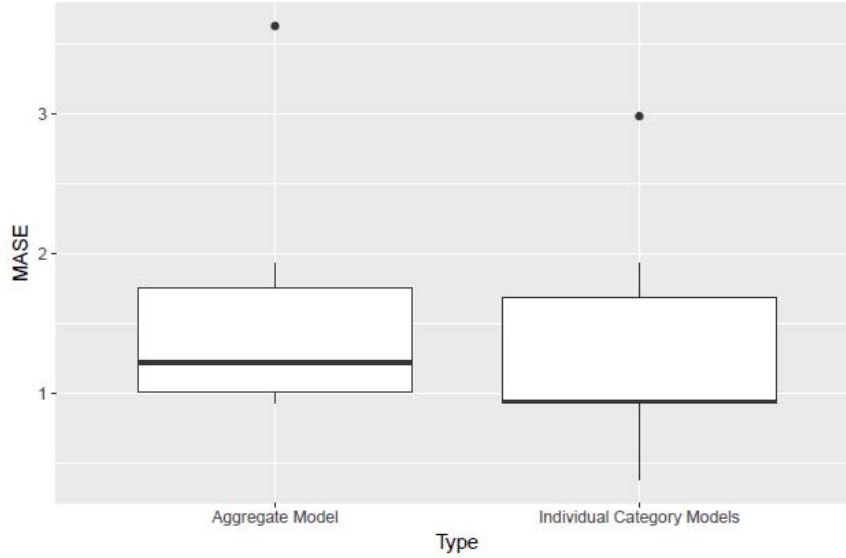


Figure 25: IT Level 2 Model Performances

of the six level 2 subcategories next to the distribution for each customized model of the six. The median MASE value for individual models is visibly lower than the IT level 1 model's median MASE value. The same distribution of MASE values is shown in Figure 26. Fifty percent of the individual model MASE values are the same around 0.94, while 67% of the MASE values for level 1 category models appear to hover around 1 as well.

MAJCOM Expenditure Profiles

See Appendix A for MAJCOM expenditure decompositions within the IT level 1 category. Similar to the process for comparing expenditure profiles at the level 2 subcategory level, MAJCOM optimal models are displayed in Figure 27. Out of the 11 MAJCOMs, the exponential smoothing and linear regression models minimized MASE thrice each; the aggregate ARIMA model and seasonal naive models minimized MASE twice each; the automated ARIMA model minimized MASE once, and the naive and trailing moving average models minimized MASE for none of the 11 MAJCOMs. Appendix B shows model error measures for each MAJCOM.

Figure 28 groups the three benchmark models together and the four “tunable” models together in a frequency plot to show how many times a fitted model outperformed the fixed models and vice versa. Here, “tunable” models better capture expenditure profiles of MAJCOMs 82% of the time. Like comparison at level 1 and level 2, this suggests that the more MAJCOM leaders can customize their algorithms, the more accurately they can capture their respective MAJCOMs' expenditure profiles.

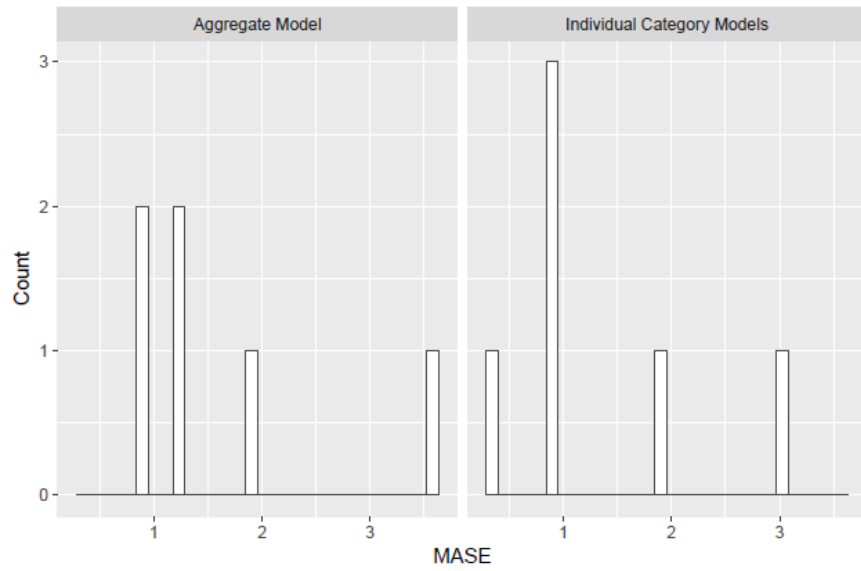


Figure 26: IT Level 2 Histogram of MASEs

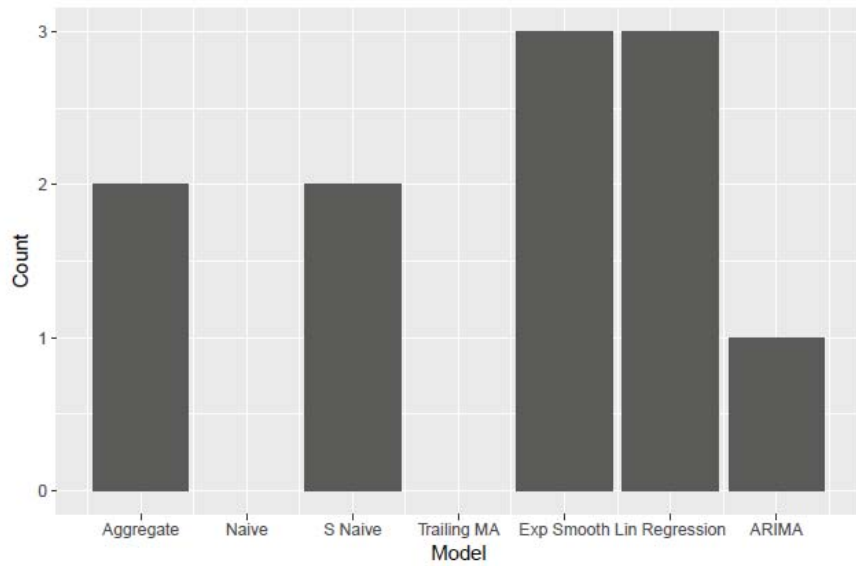


Figure 27: Optimal MAJCOM Model Counts

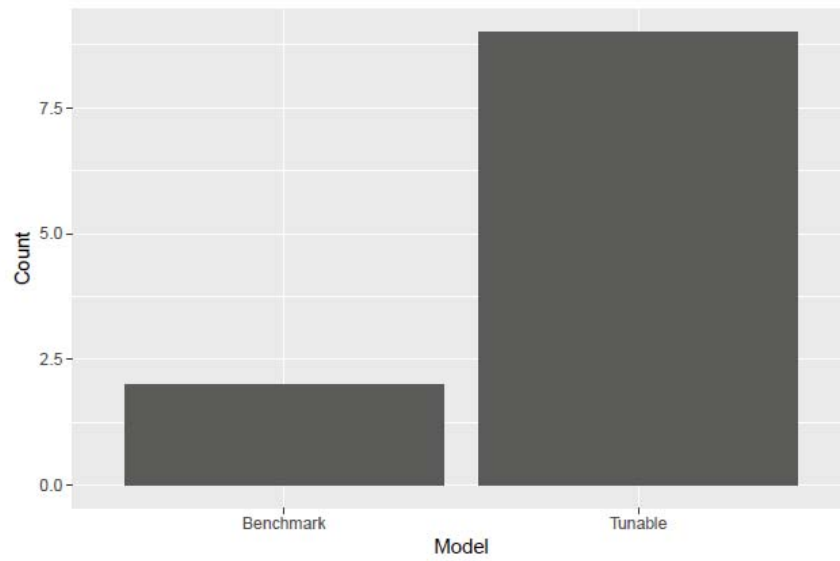


Figure 28: MAJCOM Model Counts

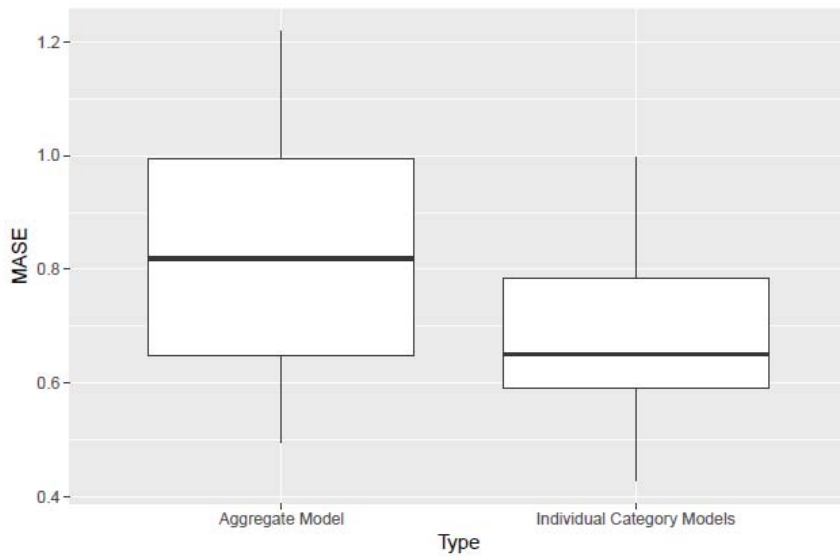


Figure 29: MAJCOM Model Performances

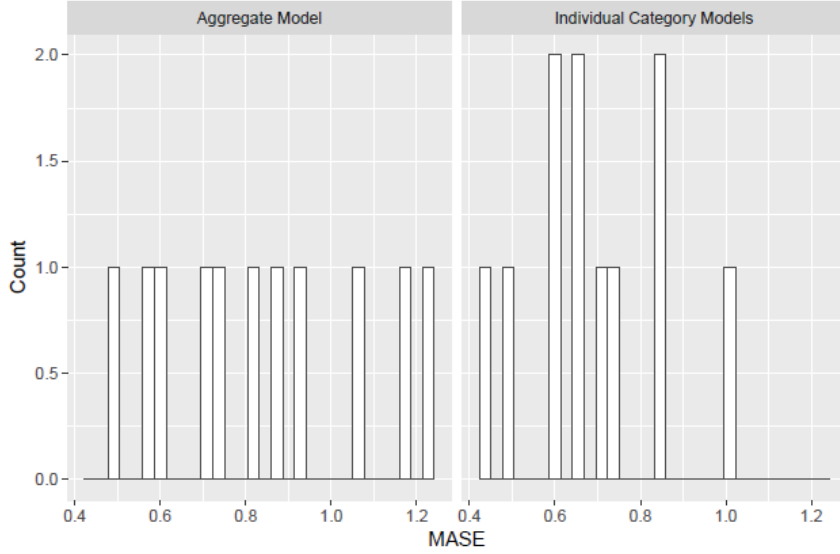


Figure 30: MAJCOM Histogram of MASEs

After visually comparing the optimal models selected for each of the 11 MAJCOMs, the IT level 1 ARIMA model and individually optimized models were each applied to the expenditure profiles to compare the difference, and thus the worth, of customizing models. These MASE values may be found in Appendix C. First, box-plots in Figure 29 show the distribution of MASE values for the level 1 model applied to each of the 11 MAJCOMs next to the distribution for each customized model of the 11. The distribution of MASE values as a whole is visibly lower for individual models than the distribution of MASE values from the aggregated IT level 1 ARIMA model. The same distribution of MASE values is shown in Figure 30. MASE values from the aggregate model’s performance appear evenly distributed between roughly 0.5 and 1.3. While a similar range of values, the distribution of MASE values for individual MAJCOMs appears slightly more concentrated with 73% of values between 0.6 and 0.82.

Base Expenditure Profiles

Similar to the processes for comparing expenditure profiles at the level 2 subcategory and MAJCOM levels, base optimal models are displayed in Figure 31. Out of the 77 bases, MASE was minimized by a linear regression model 28 times (36% of bases), by a seasonal naive model 16 times (21% of bases), by the aggregate IT level 1 ARIMA model 13 times (17% of bases), by an exponential smoothing model 11 times (14% of bases), by an automated ARIMA model eight times (11% of bases), by a naive model once (1% of bases), and by a trailing moving average model 0 times. Appendix B shows model error measures for each base.

Figure 32 groups the three benchmark models together and the four “tunable” models together in a

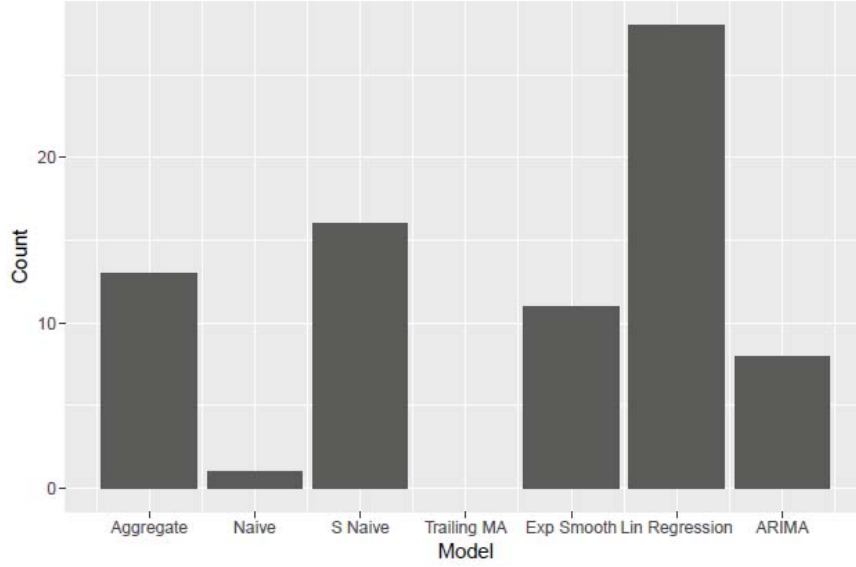


Figure 31: Optimal Base Model Counts

frequency plot to show how many times a fitted model outperformed the fixed models and vice versa. Here, “tunable” models better capture expenditure profiles of bases 78% of the time. Like comparison at the level 1, level 2, and MAJCOM levels, this suggests that the more leaders can customize their algorithms, the more accurately they can capture their respective bases’ expenditure profiles.

After visually comparing the optimal models selected for each of the 77 bases, the IT level 1 ARIMA model and individually optimized models were each applied to the expenditure profiles to compare the difference, and thus the worth, of customizing models. These MASE values may be found in Appendix C. First, the box-plots in Figure 33 show the distribution of MASE values for the level 1 model applied to each of the 77 bases next to the distribution for each customized model of the 77. The distribution of MASE values as a whole is visibly lower for individual models than the distribution of MASE values from the aggregated IT level 1 ARIMA model. The same distribution of MASE values is shown in Figure 34. MASE values from the aggregate model’s performance concentrate between 0.6 to 1, while individual model MASE values concentrate around 0.5 to 0.8.

Two assumptions drive the statistical test to determine whether or not there is a difference between the aggregate level 1 model’s performance and the customized models’ performance across the bases. The first assumption is that the two sets of errors have equal variances. The test of variance in Appendix E shows a p-value of 0.4325 against the null hypothesis that the true ratio of variances is equal to 1. Therefore, we fail to reject the null, and the variances are assumed to be equal. The second assumption is that each sample of errors is approximately normally distributed. With sample sizes of 77, both error sets are subject to the

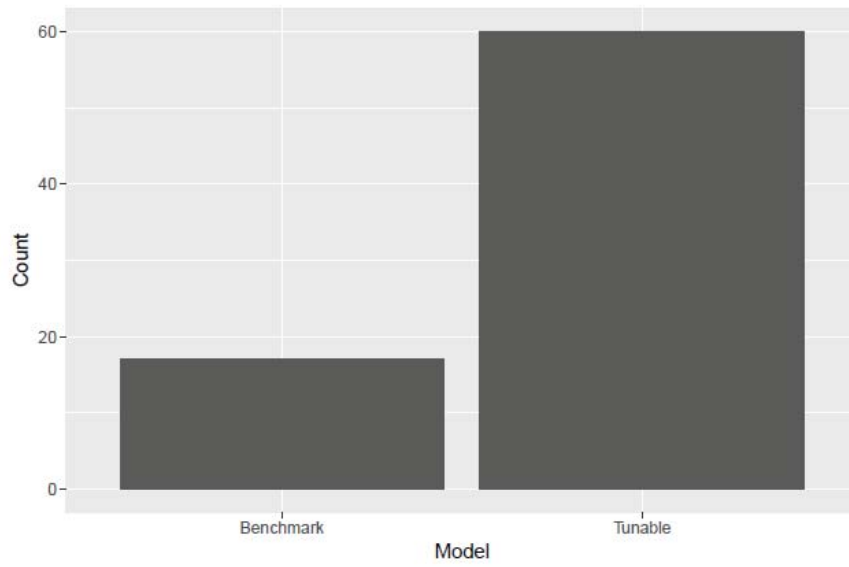


Figure 32: Base Model Counts

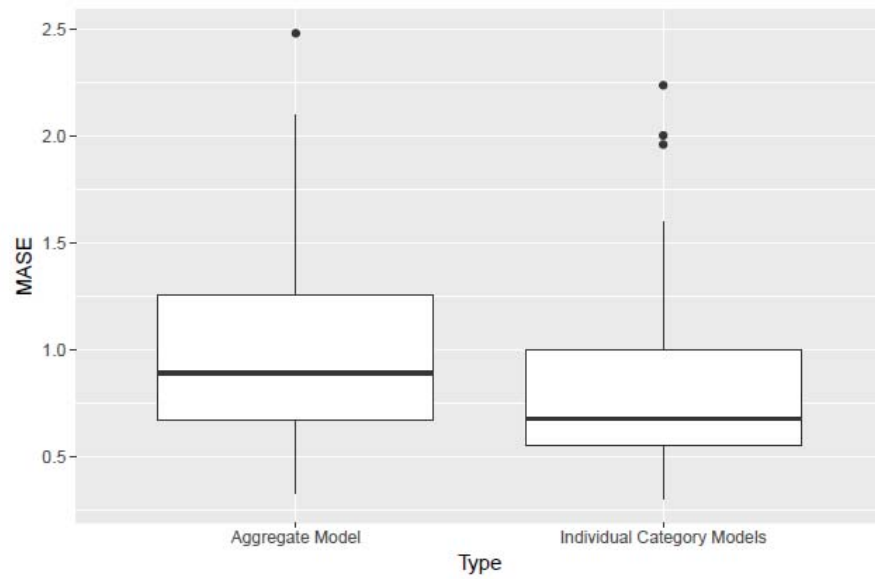


Figure 33: Base Model Performances

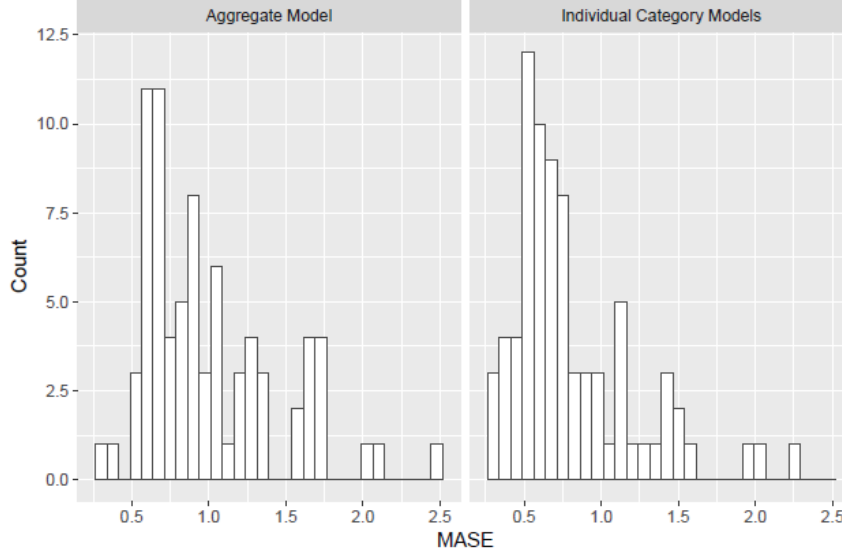


Figure 34: Base Histogram of MASEs

Central Limit Theorem (valid for sample sizes greater than 30) and are assumed to be normal. The two sets of error measures are dependent upon each other; error measures tend to spike or dip depending on the base the two different models are applied to. Therefore, the paired difference of the two error measures at each base was tested.

The paired t-test in Appendix E of the difference between the two error measures at each base (one-sample test of 77 observations) yields a p-value of $7.569\text{e-}12$ against a null hypothesis that the true difference in error set means is equal to 0. Therefore, we reject the null in favor of the alternative hypothesis that the true difference in means is not equal to 0 (mean difference of 0.1936211). Finally, to ensure robust findings, a non-parametric Wilcoxon signed-rank test, shown in Appendix E, yields a p-value of $3.61\text{e-}12$, and we once again reject the null hypothesis that the true difference in means is not equal to 0. Testing the paired difference between aggregate and individual model performances two separate times with the same result of rejecting the null hypothesis presents strong evidence that there is a difference in model performance at the base level between using the same aggregate model and using individually optimized models for each.

Finally, a look at optimized ARIMA models for each base, regardless of whether or not the ARIMA model is the optimal model, can further characterize each base's expenditure profile. ARIMA models can help distinguish bases whose spending is more reactive from those whose spending is more deliberate or consistent. Any base with an automated, optimal ARIMA model with $(0,0,0)(0,0,0)$ attributes exhibit no systematic pattern in spending. Those bases with a 1 attribute, especially when it is a seasonal $\text{AR}(1)$ or first order seasonal differencing, implies a base spends largely based on the previous season's spending. This infers

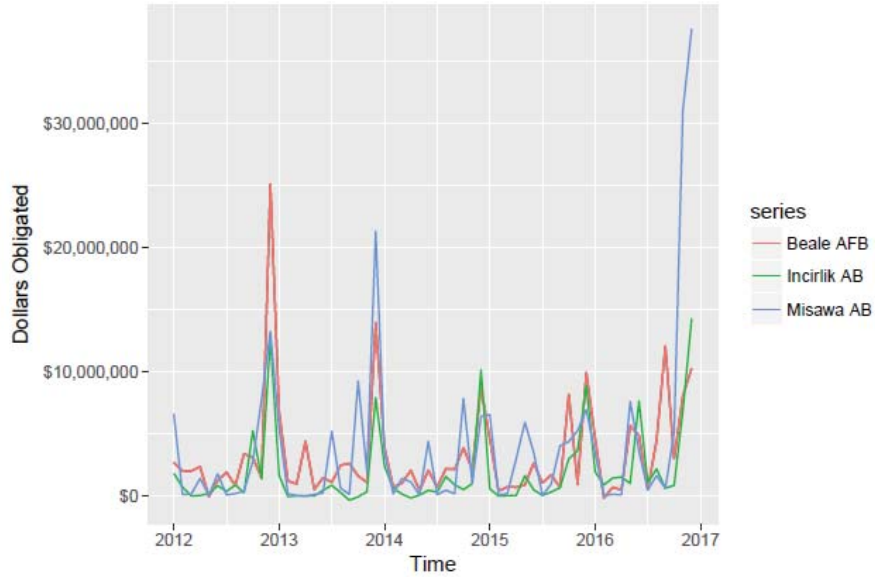


Figure 35: Outlier Base Plots

a base's spending is largely dependent on what it spent that same month last year.

Appendix F shows the results of this exploratory ARIMA optimization. The most common attribute across all bases is a first order seasonal AR term followed by first order seasonal differencing. This suggests that most base-level expenditures can be predicted by using information from the preceding season. Heath AFB, Robins AFB, Tinker AFB, and Wright-Patterson AFB exhibited optimized ARIMA models of $(0,0,0)(0,0,0)$, suggesting these bases spend their O&M contract dollars reactively.

Outlier Discussion

Analysis at the level 1 category, level 2 subcategory, and MAJCOM levels all yielded MASE error measures falling within 2 standard deviations of the mean. However, out of the 77 models optimized for each base, 3 models yielded MASE measures greater than 2 standard deviations from the mean MASE of all base models. These 3 bases may be outliers and represent expenditure profiles that are more difficult to model than may be worth the effort, or they may be bases whose spend require models other than those tested in this research.

Figure 35 plots all 3 base expenditure profiles. High error measures could be attributed to the large, uncharacteristic spike mid-year through the test year of FY 2016. Additional effort may be necessary to create accurate metrics for these bases' spends.

Summary

“Tunable” models outperformed benchmark models 88% of the time, overall. At lower levels of spend, the aggregate models become less applicable; different sub-groups have personalized patterns. Visually and statistically with the base level t-tests, evidence supports the notion that customizing models to fit individual expenditure profiles outperforms aggregate models applied to these same profiles. The trailing moving average benchmark model was never selected as the optimal model for any category or location.

V. Conclusion

The analysts review the research questions outlined in Chapter 1 and provide answers interpreted from Chapter 4. Limitations in the data and research methods are then discussed. Next, potential topics for future research are covered. The chapter then summarizes the research.

Research Questions Revisited

By capturing expenditure profiles, this research seeks to assist AFICA decision makers in better tracking category performance throughout the fiscal year. A total of six different forecasting models were applied to each expenditure data series of interest, and the one that minimized MASE was selected to best represent each expenditure's profile. Expenditure profile models were compared, and an ARIMA model was found to best represent the AF's enterprise-level expenditure profile. This aggregate model was added to the six models tested for each lower level.

Research Question 1: How Can AFICA Expect Expenditure Profiles of Each Main Category to Behave at the Aggregate Level?

The enterprise ARIMA model best captured patterns in half of the level 1 expenditure profiles — five of the ten level 1 category expenditure profiles behave similarly to the enterprise expenditure profile. This profile includes both predictable annual seasonality from year to year as well as autocorrelations, or related spending month-to-month. Two of the categories exhibit expenditure profiles similar to random walks, or spending with no apparent systematic patterns. Attempts to accurately predict cost performance in these categories may require other methods outside the scope of this research. The three remaining categories exhibit systematic seasonal patterns best modeled by other forecasting methods, including exponential smoothing and linear regression. These categories' spends change from year to year in predictable patterns, but do not exhibit relationships month to month like the enterprise spend profile does. The results of analysis for this research question are summarized in Table 2.

Table 2: Level 1 Model Summary

Level 1 Category	Agg Seasonal	Alt Seasonal	No Seasonal
Professional Services	X		
Facilities & Construction		X	
IT	X		
Transportation & Logistics Services	X		
Industrial Products & Services		X	
Human Capital		X	
Office Management			X
Security & Protection	X		
Travel & Lodging			X
Medical	X		

Research Question 2: How Do Expenditure Profile Behaviors of a Category Change Across Different Sub-Categories?

The level 1 ARIMA model for IT best captured patterns in 33% of the level 2 expenditure profiles — two of the six level 2 subcategory expenditure profiles behave similarly to the aggregate level 1 IT expenditure profile (which also happens to be one of the five other level 1 categories that behave much like the AF enterprise expenditure profile). These expenditure profiles exhibit seasonal patterns year to year and relationships between spending month to month. Another two of IT’s level 2 subcategories exhibit benchmark model patterns, which suggests extra effort in customizing algorithms for these may not be worthwhile or that other methods may be required. Although these expenditure profiles exhibit seasonality year to year, these seasonal patterns are not consistent. The remaining two level 2 subcategories exhibit patterns best modeled by either linear regression or other automated ARIMA models. These categories’s spends follow predictable seasonal changes year to year but not relationships month to month like the enterprise-level spend. This means that up-front planning of resources at the beginning of each fiscal year based on prior year spending is more effective than mid-year adjustments to budgets based on recent expenditures. The results of analysis for this research question are summarized in Table 3.

Table 3: Level 2 Model Summary

Level 2 Subcategory	Agg Seasonal	Alt Seasonal	No Seasonal
IT Consulting			X
IT Hardware		X	
IT Outsourcing	X		
IT Security		X	
IT Software	X		
Telecommunications			X

Research Question 3: How Do Expenditure Profile Behaviors Change as Categories Disaggregate Into MAJCOM and Base-Level Operations?

With respect to MAJCOMs, 2 the 11 expenditure profiles follow similar patterns to the aggregate level 1 category of IT expenditure profile. Like previously mentioned, these MAJCOMs exhibit predictable seasonal changes year to year as well as relationships month to month in spending. Another two of the profiles behave according to benchmark model patterns, suggesting extra effort in modeling these profiles may not be worthwhile or that other methods may be required. None of the systematic patterns seen at the aggregate category level are present in these profiles. The remaining seven MAJCOMs exhibit patterns best modeled by other methods between exponential smoothing, linear regression, and automated ARIMA. These commands show predictable seasonal changes year to year but not the month-to-month relationships seen in the aggregate level 1 expenditure profile. The results of analysis for this research question are summarized in Table 4.

Table 4: MAJCOM Model Summary

MAJCOM	Agg Seasonal	Alt Seasonal	No Seasonal
AFMC	X		
AFSPC		X	
ACC		X	
AETC		X	
AMC		X	
PACAF		X	
USAFE		X	
AFGSC		X	
AFDW			X
USAFA	X		
AFSOC			X

With respect to bases, 17% follow similar patterns to the aggregate level 1 category of IT expenditure profile. One percent of bases exhibit no predictable patterns seen in the aggregate category expenditure profile. These bases' spend may be more difficult to accurately predict. Twenty one percent of bases do not exhibit any spend patterns captured by the seven models, suggesting, as previously stated, that extra effort in modeling these profiles may not be worthwhile. The remaining 61% of bases exhibit predictable seasonal changes in spending patterns best modeled by other methods, including exponential smoothing, linear regression, and automated ARIMA. These base spends do not show the same relationships in spending month-to-month as seen at the aggregate category level. Analysis rejects the notion that an aggregate IT model performs similarly to customized models for each base. See Appendix B for all optimized base models. The research questions support the conclusions that, in general, the more a unit can customize its algorithms, the more accurately it can capture its expenditure profile. The more localized the level of spend, the less applicable the aggregate models become, and different sub-groups have more personalized patterns. Further exploration with optimized ARIMA models showed that most base-level expenditures can be predicted by using information from the preceding season. Four of the 77 bases, however, exhibited reactive expenditure profiles rather than deliberate, consistent spending behaviors.

Implications

By using MAPE, where available, it is possible to put a dollar value on the implications of customizing expenditure profile models rather than applying an aggregate model to all disaggregates. Although this may be applied at any level, for demonstrative purposes, this research highlights two of the level 2 subcategories of IT: IT Hardware and IT Consulting. With a five-year historical spend total of 2.63 billion dollars, IT Hardware may be modeled with the same aggregate ARIMA model that best captures the behaviors of the aggregate level 1 IT category. Using this model as a means of tracking spend performance through the years, multiplying its MAPE error measure by the total spend yields a rough estimate for how off the model can be expected to track performance from actual obligations. The difference between this and the same total dollar error using an individualized model equates to a 1% improvement, or a difference of \$21.6 million. This difference is how much better, in dollars, AFICA can do in tracking spend performance of IT Hardware. Deviations will occur from any metric tracking spend, but this can give a more realistic baseline off of which AFICA may require justification from organizations who deviate significantly.

Similarly, with respect to the IT Consulting subcategory, the difference over these same five years between using the aggregate ARIMA model and an individualized model equates to a 135% error improvement, or \$315 million. Again, deviations will occur from any metric tracking spend, but this can give a more realistic baseline off of which AFICA may require justification from organizations who deviate significantly. More accurate metrics would also allow AFICA to better allocate resources each fiscal year for O&M execution.

Limitations

Quality of analysis is limited to the quality of the data input. AFICA admits a threshold of error in its FPDS database, and any imperfections may adversely affect the results and conclusions from this research. Also, with data limited to a five-year timespan, it is possible that other external events from 2012—2016 affected the research; sequestration enacted in 2013, for example, may have influenced certain expenditure profiles. Although only a few MAJCOMs and bases were removed before analysis, their removal may also affect the results of this research.

Future Research

Additional forecasting methods may add to this research and make findings more robust. Modeling methods outside of forecasting may also prove beneficial; neural networks, hierarchical linear models, and

other methods using cross-sectional data may yield important findings. Repeated analysis using updated data, including FY 17 obligations, may update findings to be more relevant to today's spending environment. For example, inflation impacts would be captured by the trend attribute in the models. It is interesting that the models did not exhibit any trend, suggesting no inflation. This may be due to the dip in obligations in FY 13 with sequestration, and repeating this analysis on updated data may result in new trends. Future research may also include conducting similar analysis for the decision makers in Financial Management with the capability of affecting the management of O&M funds. Analysis of other appropriations may also be worthwhile; the applicability of the approach in this research is only limited by the number of analysts willing to use it.

Bibliography

- AF Category Manager Support Office Charter. (2017).
- Blattberg, R. C. (1995). *Category Management, Guides*, 1–5. Washington DC: Food Marketing Institute.
- Blyth, C. R. (1972). On Simpson's Paradox and the Sure-Thing Principle. *Journal of the American Statistical Association*, 67(338), 364–366. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/01621459.1972.10482387>
- Boehmke, B. C., Johnson, A. W., White III, E. D., Weir, J. D., & Gallagher, M. A. (2015). A multilevel understanding of tooth-to-tail. *IIE Annual Conference and Exposition 2015*, (September), 38–47. <https://doi.org/10.13140/RG.2.1.1708.2964>
- Boehmke, B. C., Johnson, A. W., White, E. D., Weir, J. D., & Gallagher, M. A. (2016). The influence of operational resources and activities on indirect personnel costs: A multilevel modeling approach. *The Engineering Economist*, 61(4), 289–312. <https://doi.org/10.1080/0013791X.2016.1155247>
- Boehmke, B. C., Johnson, A. W., White, E. D., Weir, J. D., & Gallagher, M. A. (2015). Bending the Cost Curve: Moving the Focus from Macro-level to Micro-level Cost Trends with Cluster Analysis. *Journal of Cost Analysis and Parametrics*, 8(2), 126–148. <https://doi.org/10.1080/1941658X.2015.1064046>
- Clark, W. A. V., & Avery, K. L. (2010). The Effects of Data Aggregation in Statistical Analysis. *Geographical Analysis*, 8(4), 428–438. <https://doi.org/10.1111/j.1538-4632.1976.tb00549.x>
- Couto, C. G. (2013). Integrative Lifecourse and Genetic Analysis of Military Working Dogs. Retrieved from <http://www.dtic.mil/docs/citations/ADA604395>
- Dhar, S. K., Hoch, S. J., & Kumar, N. (2001). Effective category management depends on the role of the category. *Journal of Retailing*, 77(2), 165–184. [https://doi.org/10.1016/S0022-4359\(01\)00045-8](https://doi.org/10.1016/S0022-4359(01)00045-8)
- Freedman, D. a. (1999). Ecological Inference and the Ecological Fallacy. *International Encyclopedia of the Social & Behavioral Sciences*, 6(4027–4030), 1–7. <https://doi.org/10.1073/pnas.96.1>
- Gardner, N. R. (2015). *Forecasting Foreign Currency Exchange Rates for Air Force Budgeting*.
- Garrett, T. A. (2003). Aggregated versus disaggregated data in regression analysis: Implications for inference. *Economics Letters*, 81(1), 61–65. [https://doi.org/10.1016/S0165-1765\(03\)00149-6](https://doi.org/10.1016/S0165-1765(03)00149-6)
- Government-wide CM Guidance_May 2015. (2015).
- Gruen, T. W., & Shah, R. H. (2000). Determinants and outcomes of plan objectivity and implementation in category management relationships. *Journal of Retailing*, 76(4), 483–510. [https://doi.org/10.1016/S0022-4359\(00\)00041-5](https://doi.org/10.1016/S0022-4359(00)00041-5)

- Hyndman, R. J., & Athanasopoulos, G. (2012). *Forecasting: principles and practice* | OTexts. Retrieved from <https://www.otexts.org/fpp>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Limberakis, C. G. (2012). Spend Analysis: Best-in-Class. *Supply Chain Management Review*, 16(April). Retrieved from [https://mycourses.aalto.fi/pluginfile.php/388468/mod_resource/content/1/Limberakis, C..pdf](https://mycourses.aalto.fi/pluginfile.php/388468/mod_resource/content/1/Limberakis_C..pdf)
- Lubinski, D., & Humphreys, L. G. (1996). Seeing the forest from the trees: When Predicting the Behavior or Status of Groups, Correlate Means. *Psychology, Public Policy, and Law*, 2(2), 363–376. <https://doi.org/10.1037//1076-8971.2.2.363>
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., . . . Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1(2), 111–153. <https://doi.org/10.1002/for.3980010202>
- Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, 16(4), 451–476. [https://doi.org/10.1016/S0169-2070\(00\)00057-1](https://doi.org/10.1016/S0169-2070(00)00057-1)
- Moore, N. Y., Cook, C., Grammich, C., & Lindenblatt, C. (2004). Using a Spend Analysis to Help Identify Prospective Air Force Purchasing and Supply Management Initiatives. Summary of Selected Findings. Santa Monica, CA. <https://doi.org/RAND>
- Muir, William A.; Keller, Richard S.; Knight, L. E. (2014). Category Management a Concept of Operations for Improving Costs At the Air Force Installation, 1–69.
- Office, E., & The, O. F. (2014). Transforming the Marketplace: Simplifying Federal Procurement to Improve Performance, Drive Innovation, and Increase Savings.
- Office, U. S. G. A. (2003a, June 9). Best Practices: Improved Knowledge of DOD Service Contracts Could Reveal Significant Savings. Retrieved from <http://www.gao.gov/products/GAO-03-661>
- Office, U. S. G. A. (2003b, September 10). Contract Management: High-Level Attention Needed to Transform DOD Services Acquisition. Retrieved from <http://www.gao.gov/products/GAO-03-935>
- Office, U. S. G. A. (2004, September 16). Best Practices: Using Spend Analysis to Help Agencies Take a More Strategic Approach to Procurement. Retrieved from <http://www.gao.gov/products/GAO-04-870>
- Orcutt, G. H., Watts, H. W., & Edwards, J. B. (n.d.). Data Aggregation and Information Loss. *The American Economic Review*. American Economic Association. <https://doi.org/10.2307/1815532>
- Pandit, K., & Marmanis, H. (2008). Spend Analysis: The Window into Strategic Sourcing. Fort Lauderdale,

Florida: J. Ross Publishing.

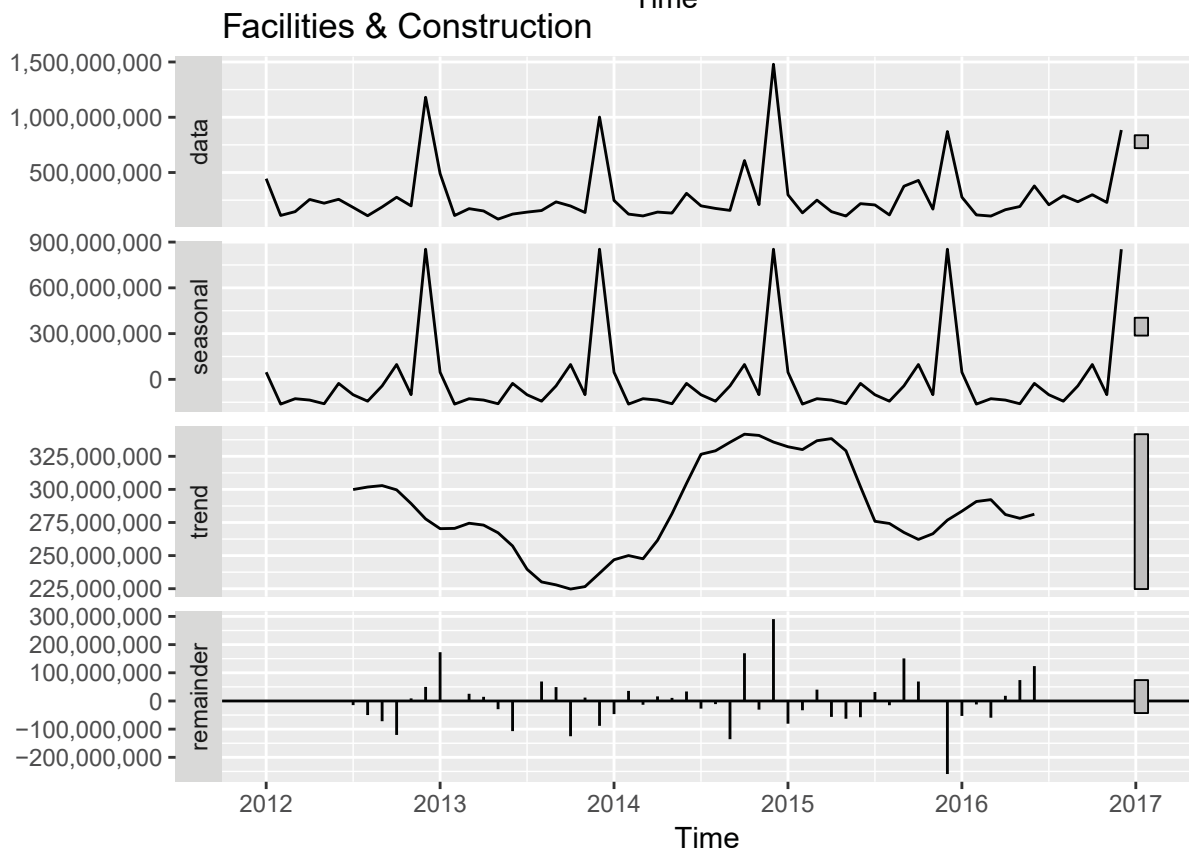
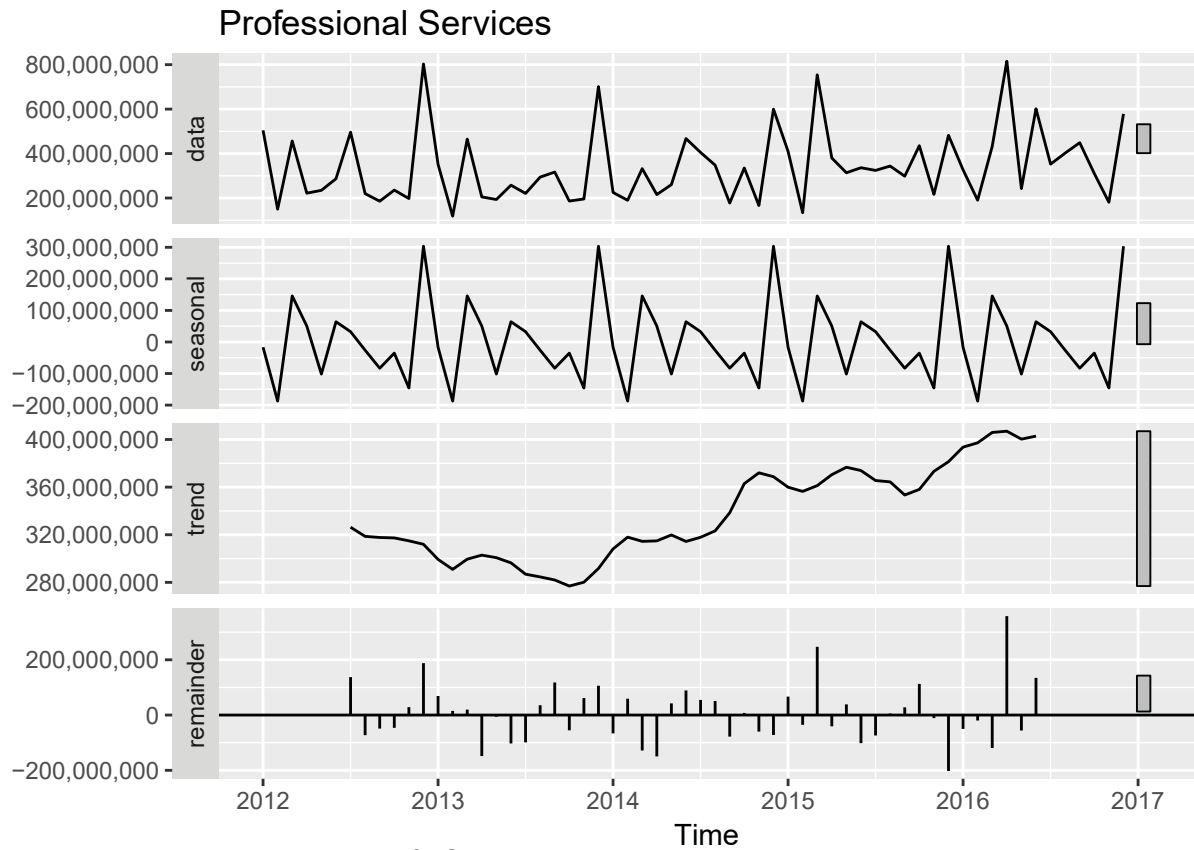
Robinson, W. (2009). Ecological Correlations and the Behavior of Individuals*. *International Journal of Epidemiology*, 38(2), 337–341. <https://doi.org/10.1093/ije/dyn357>

Shmueli, G., & Lichtendahl Jr., K. C. (2016). *Practical Time Series Forecasting With R: A Hands-On Guide*. Axelrod Schnall Publishers.

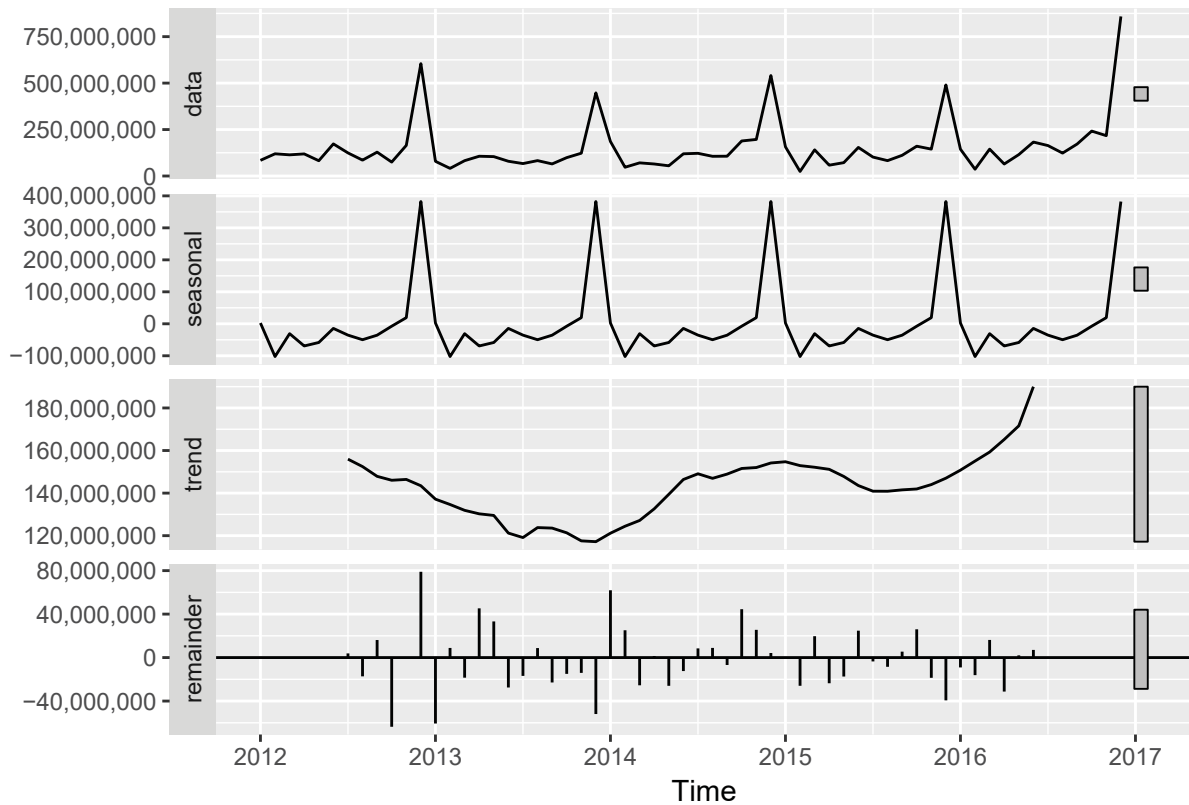
United States Government Accountability Office. (2012). *Strategic Sourcing: Improved and Expanded Use Could Save Billions in Annual Procurement Costs*. United States Government Accountability Office - Report, GAO-12-919(September), 1–65.

Zenor, M. J. (1994). The Profit Benefits of Category Management. *Journal of Marketing Research*, 31(2), 202. <https://doi.org/10.2307/3152194>

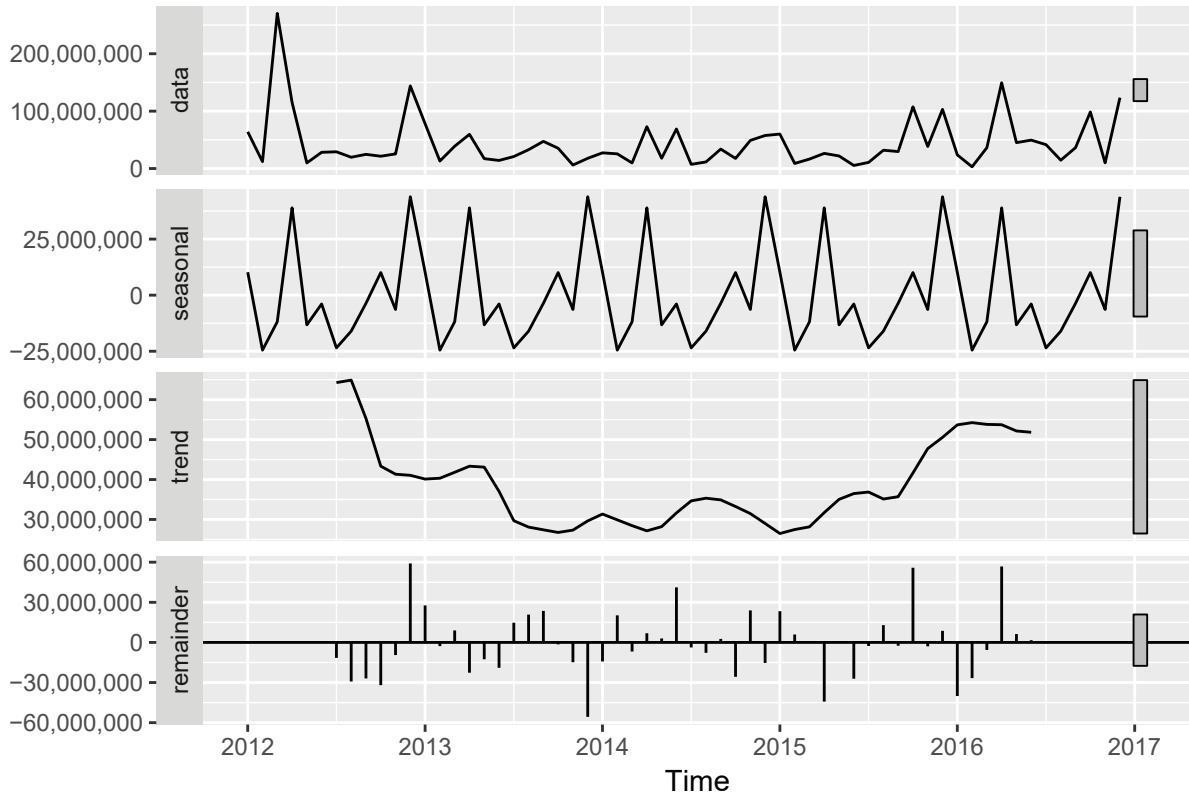
Appendix A - Decompositions



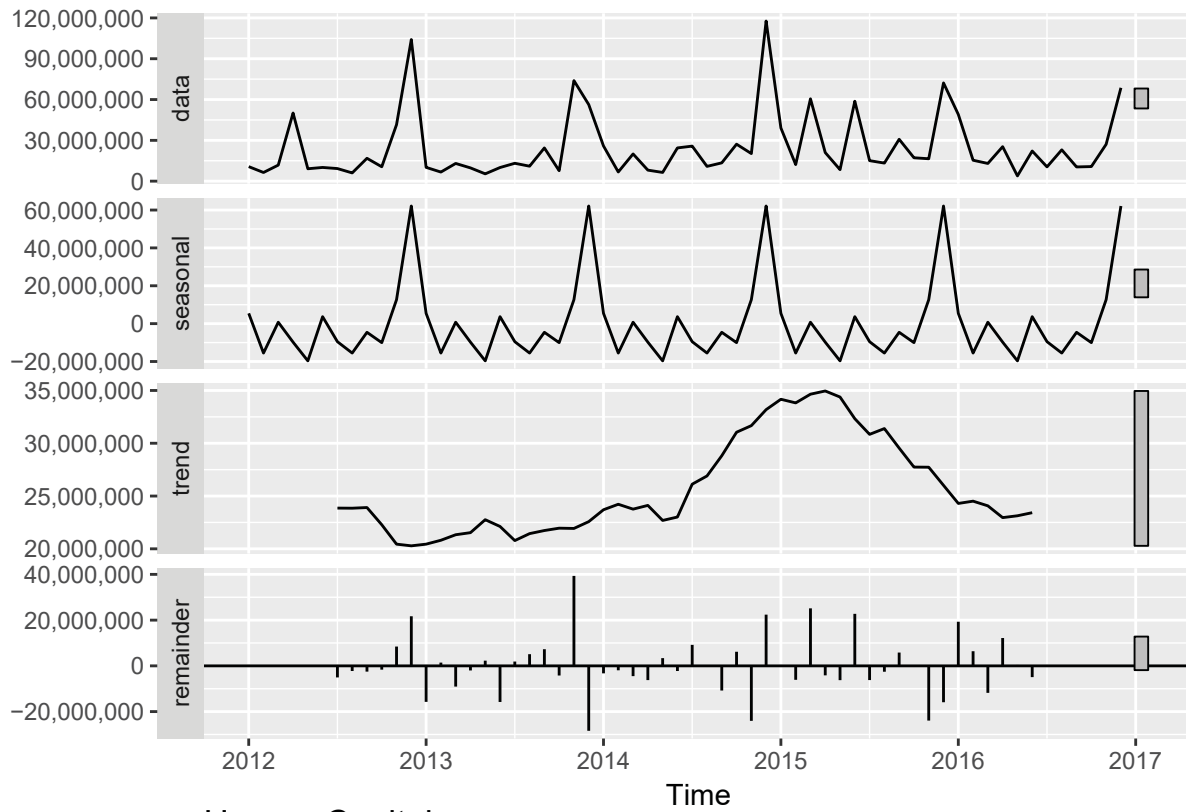
IT



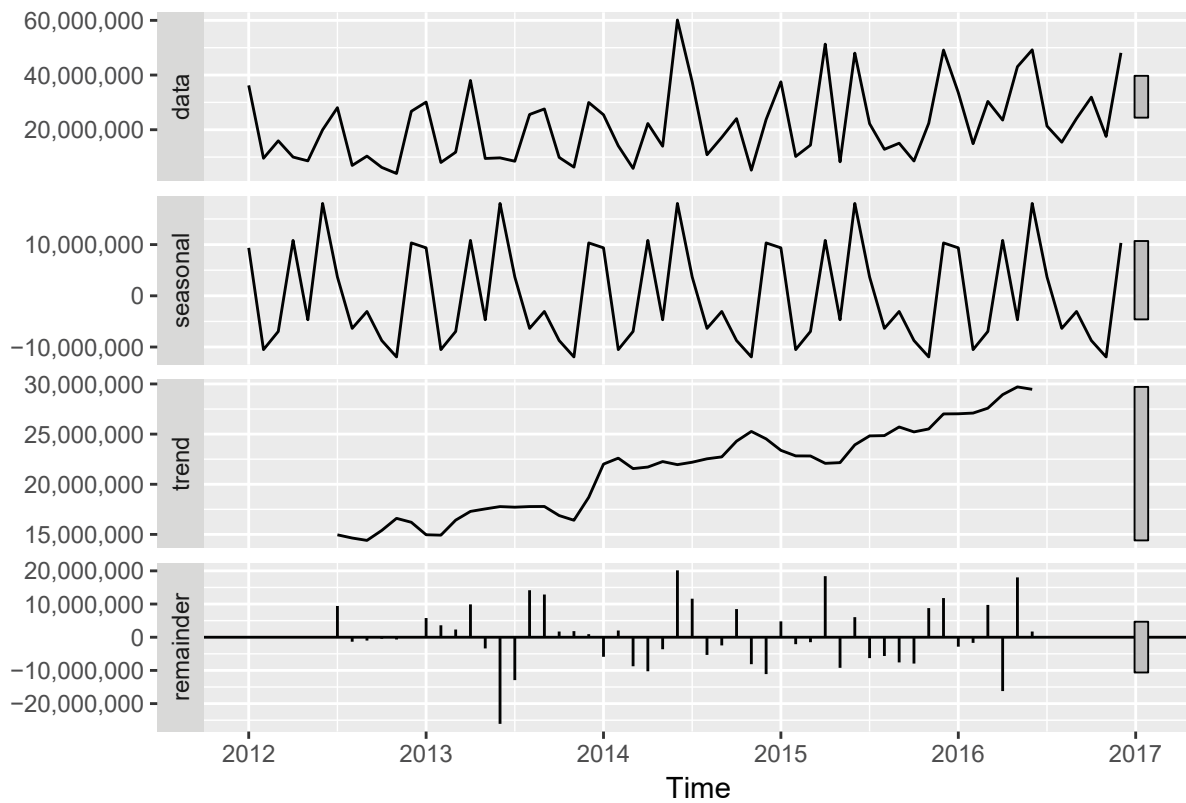
Transportation and Logistics Services



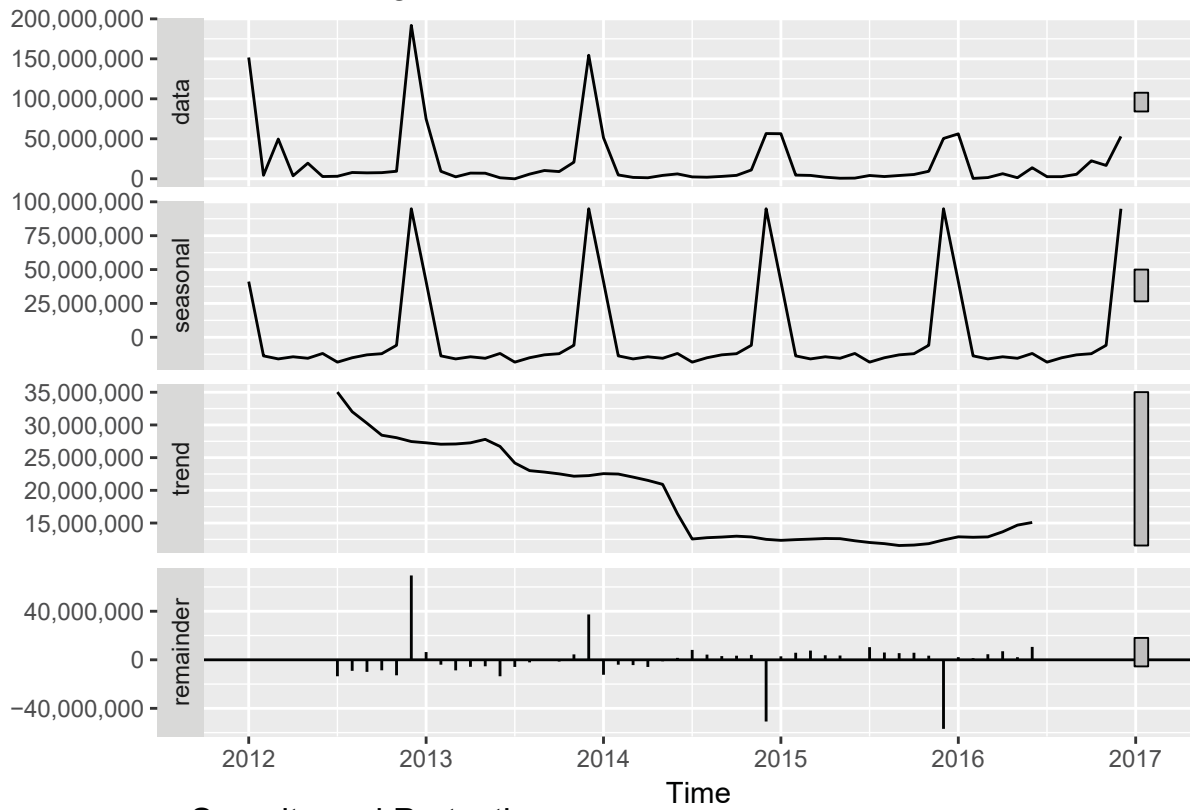
Industrial Products & Services



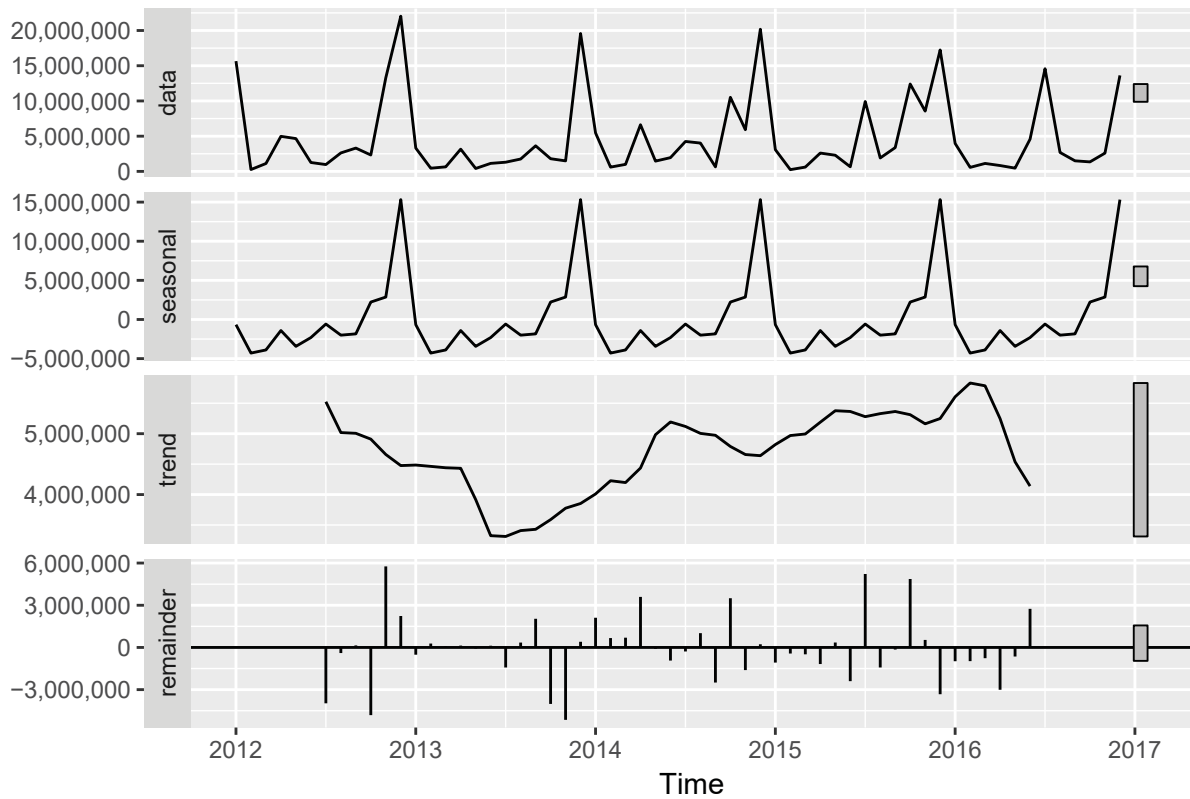
Human Capital



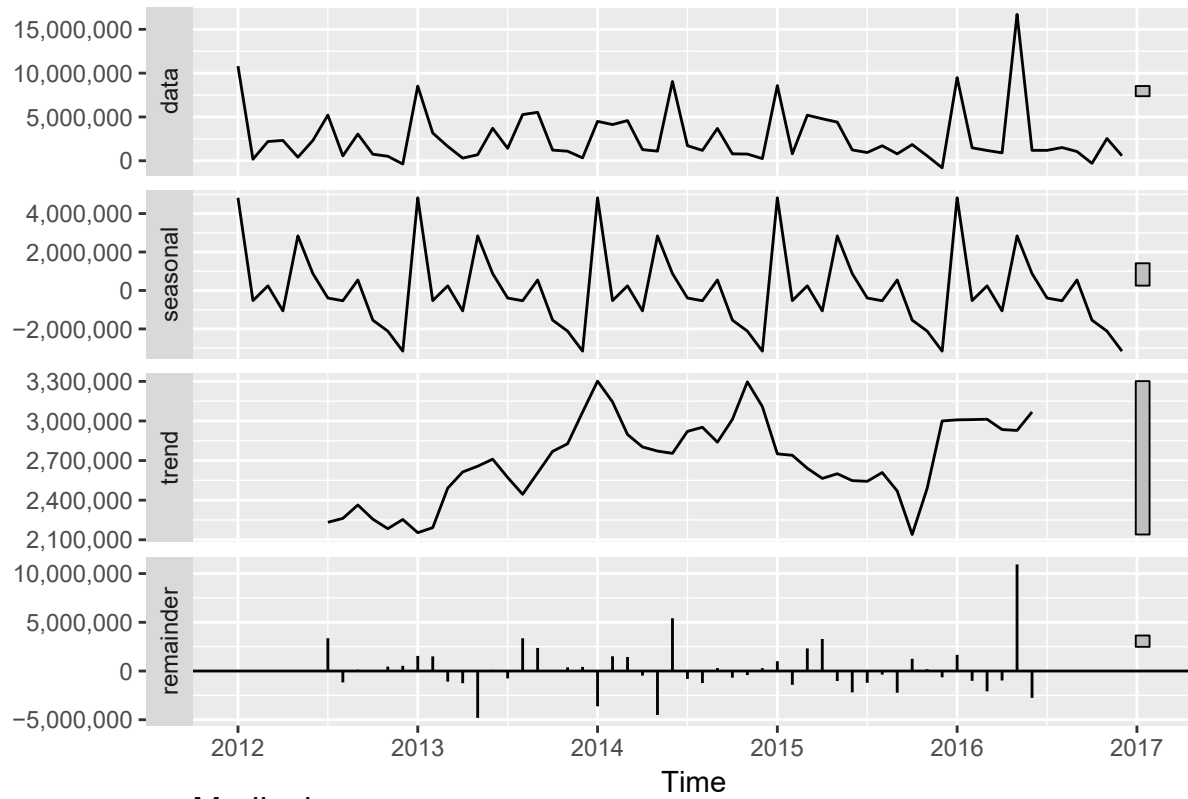
Office Management



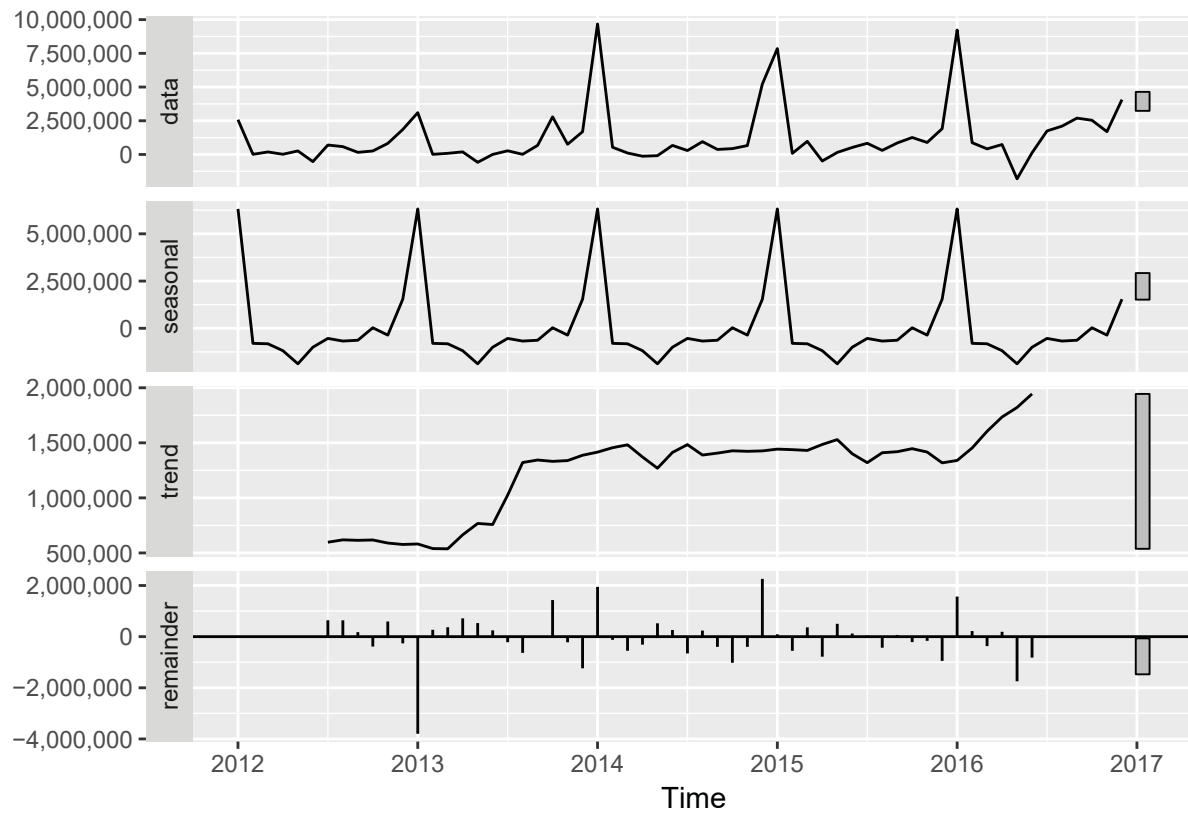
Security and Protection



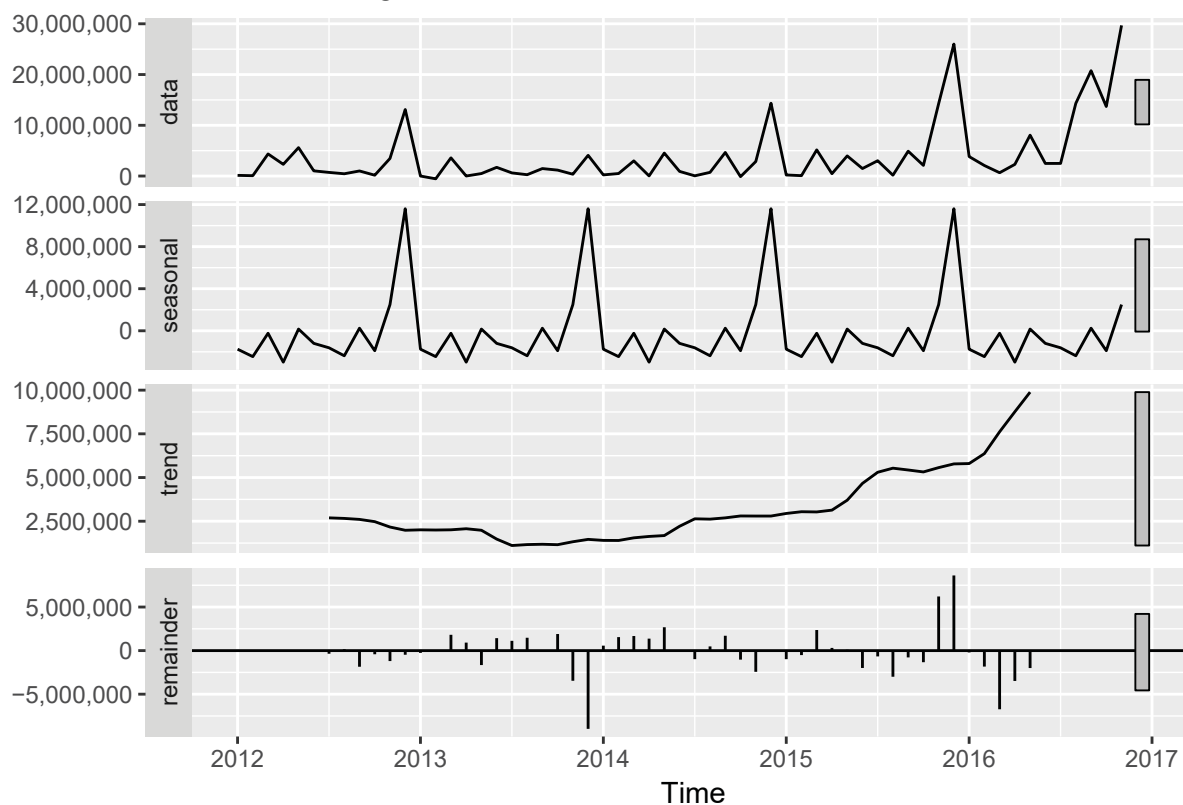
Travel & Lodging



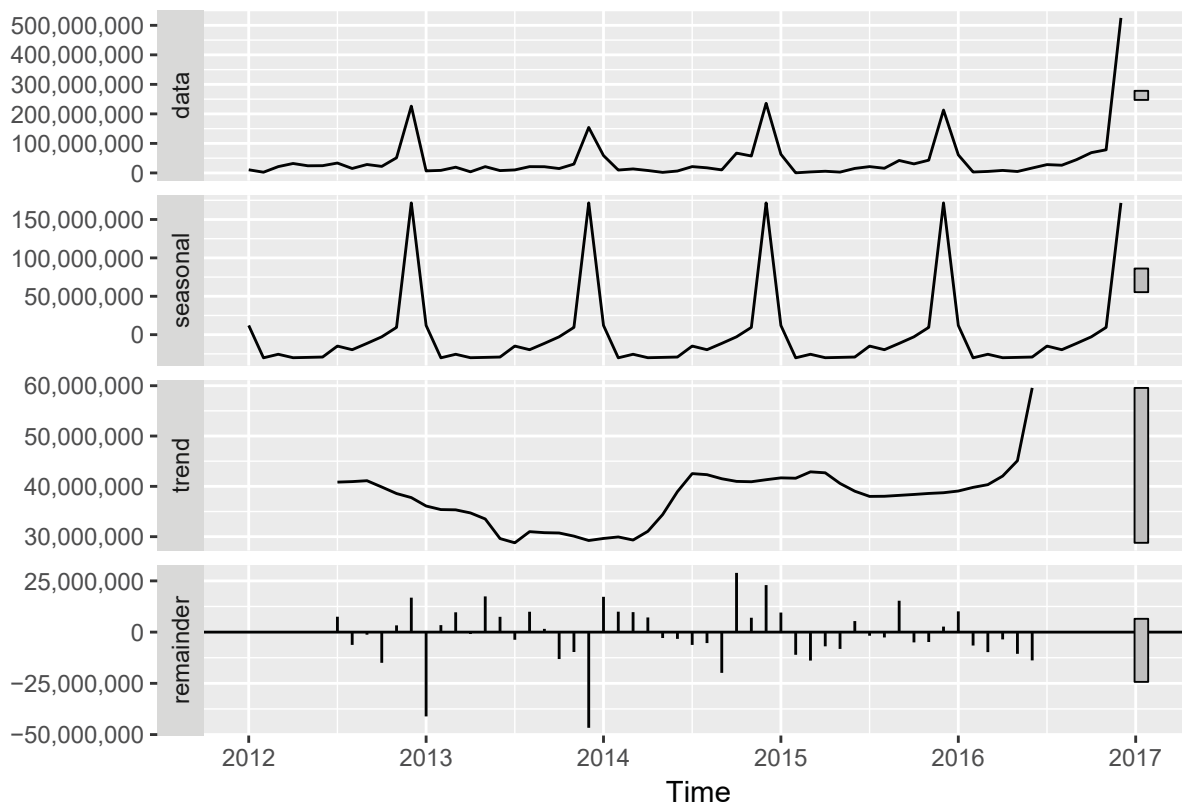
Medical



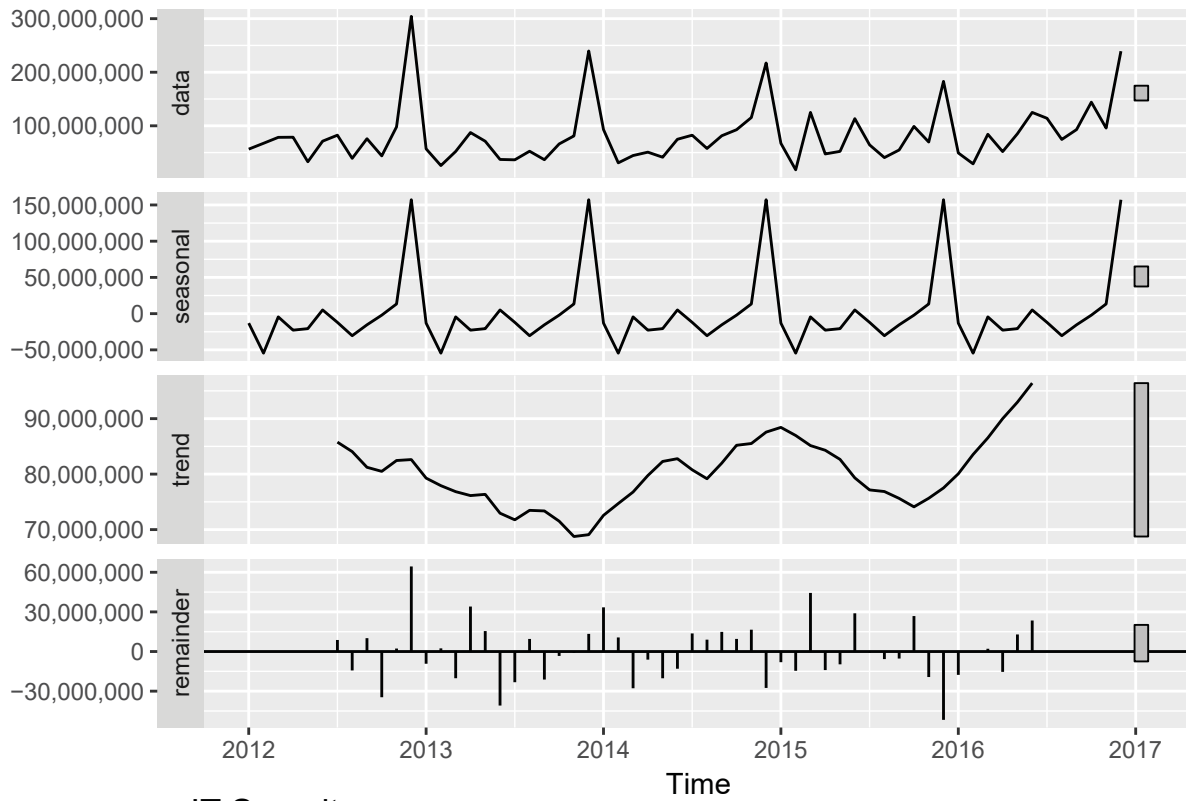
IT Consulting



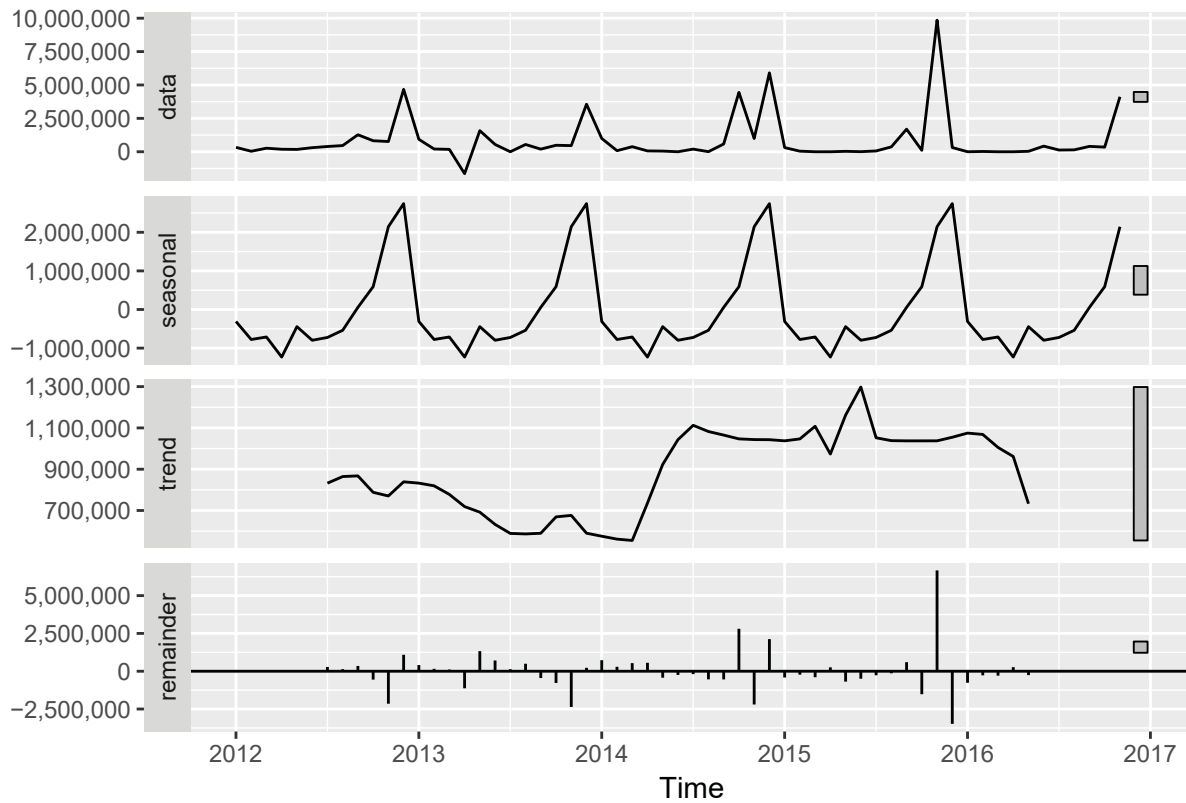
IT Hardware



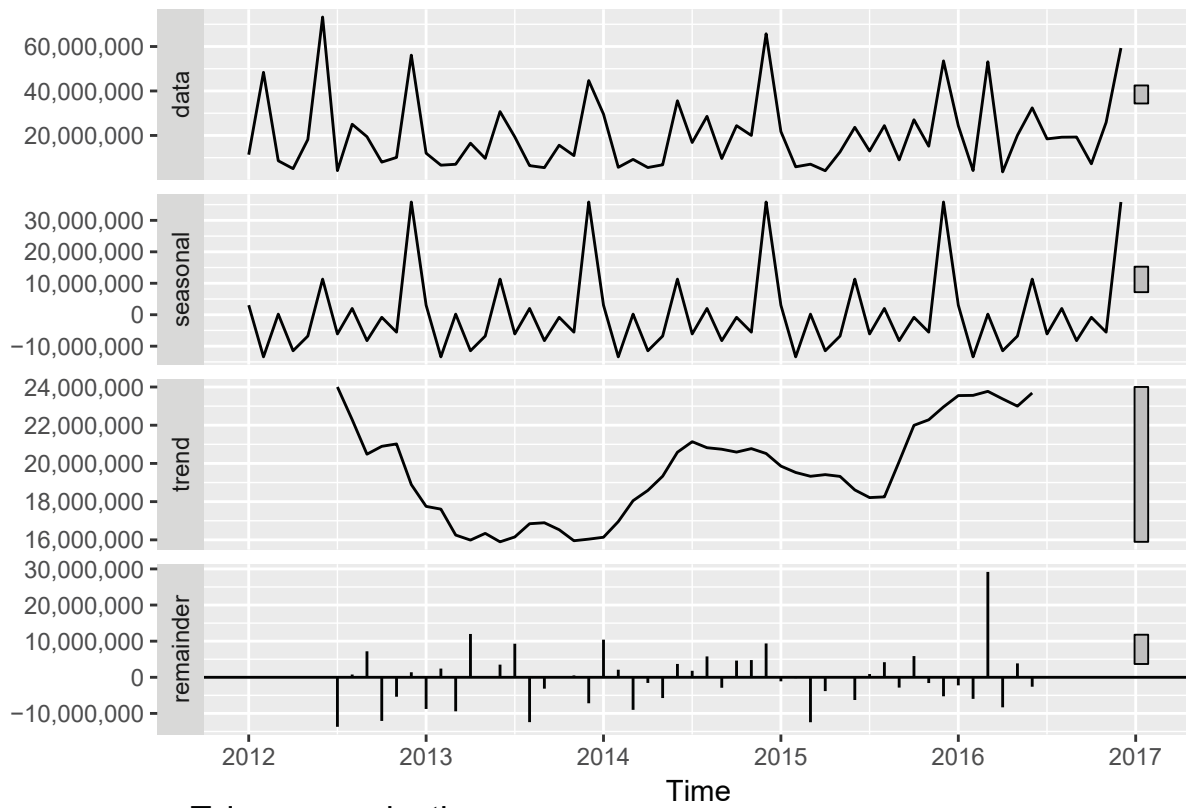
IT Outsourcing



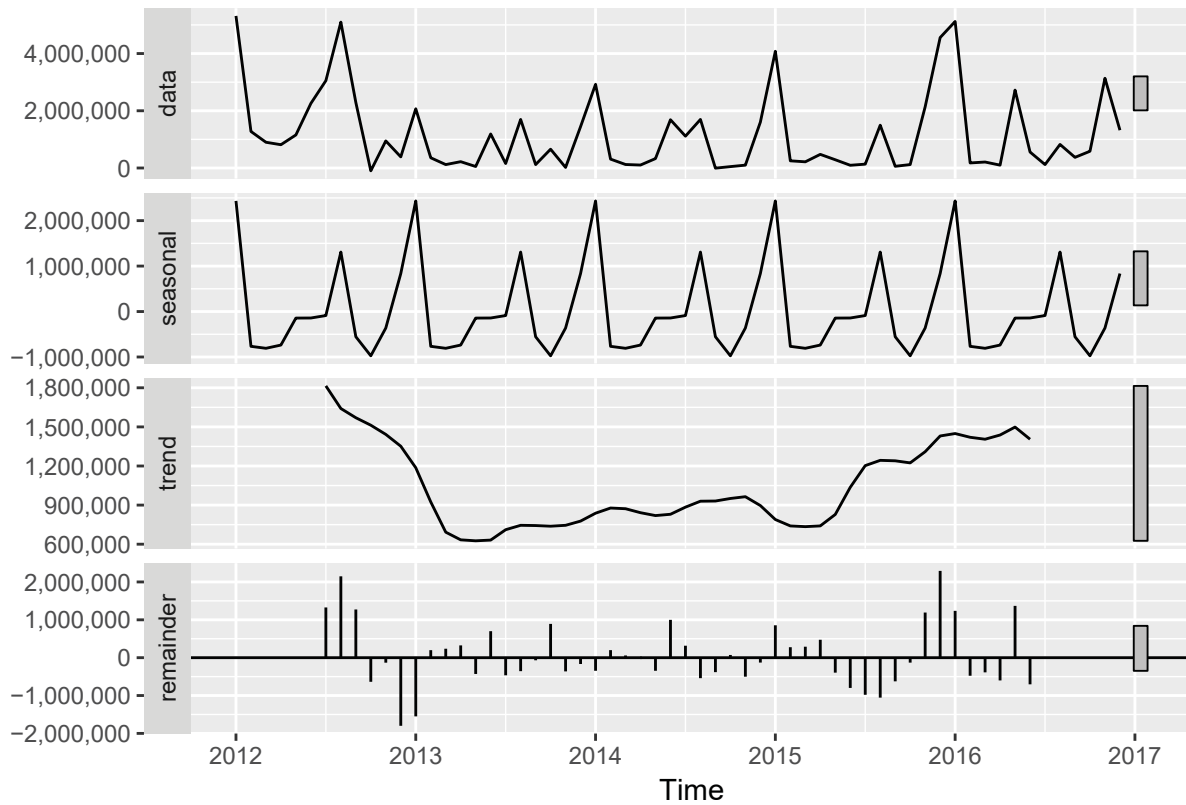
IT Security



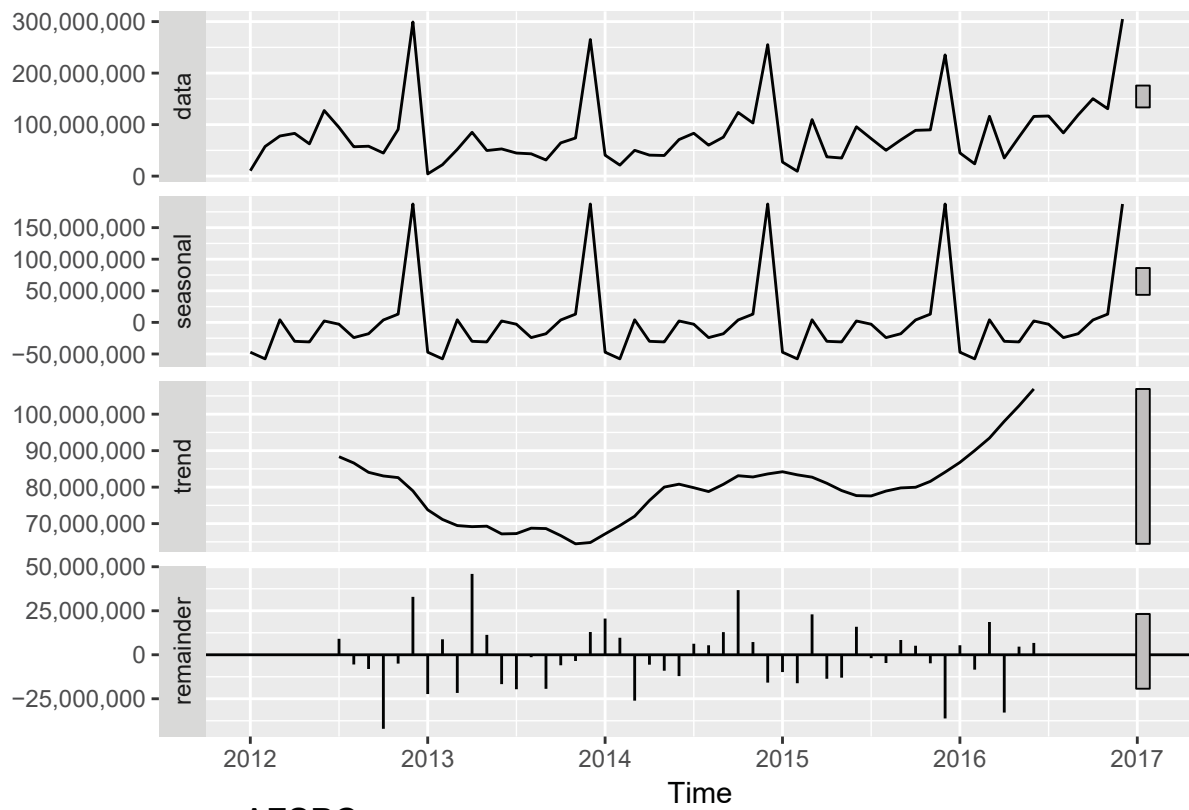
IT Software



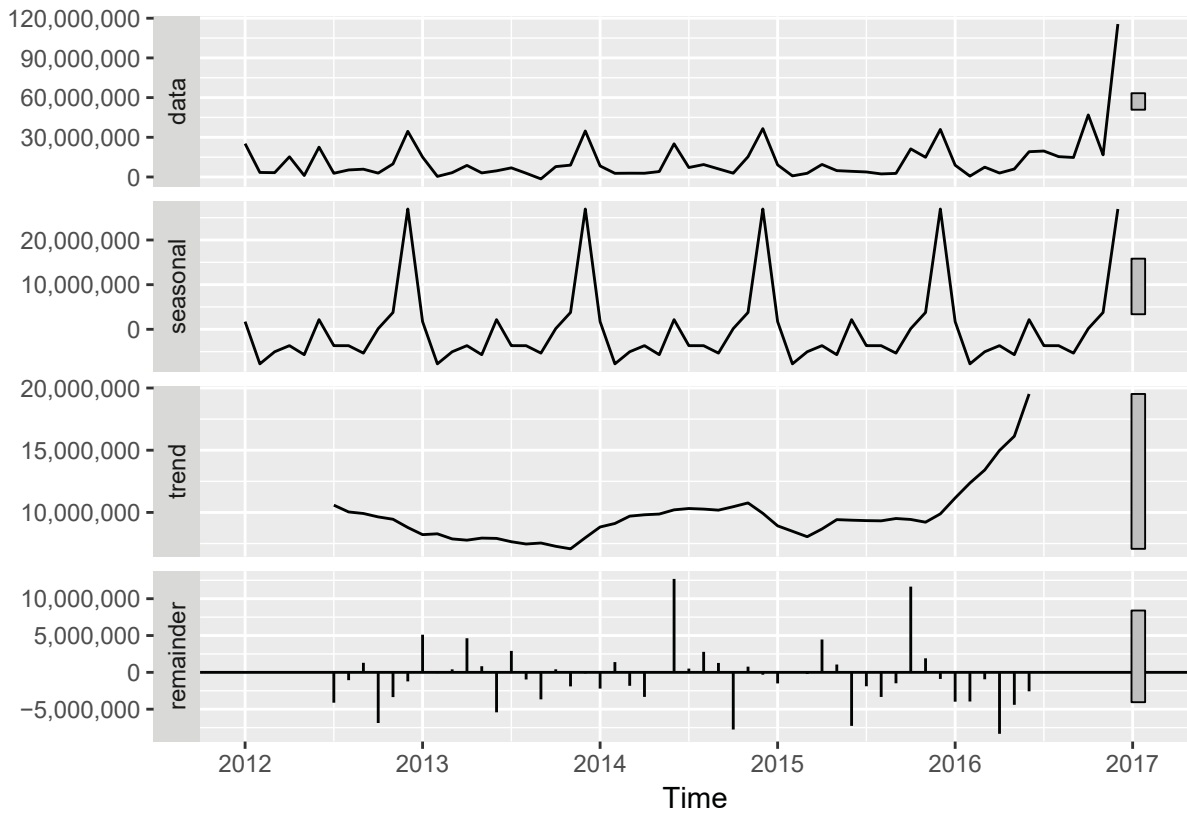
Telecommunications



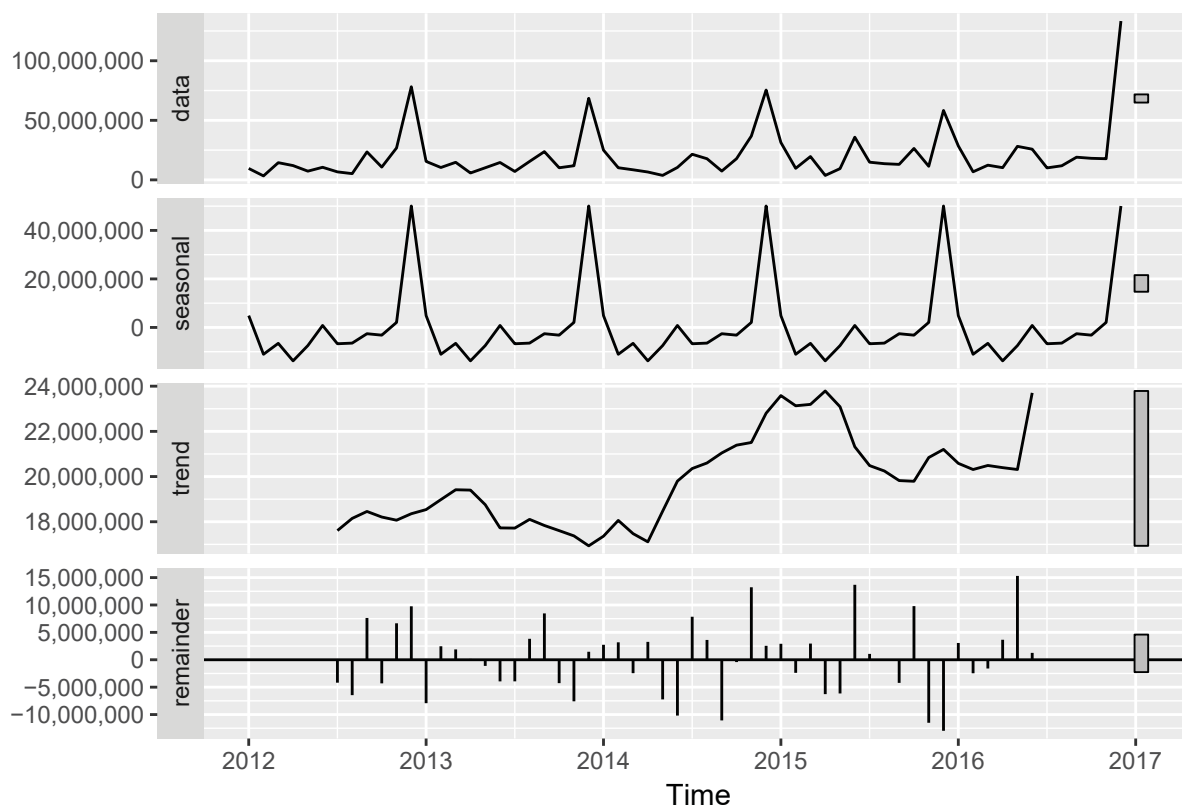
AFMC



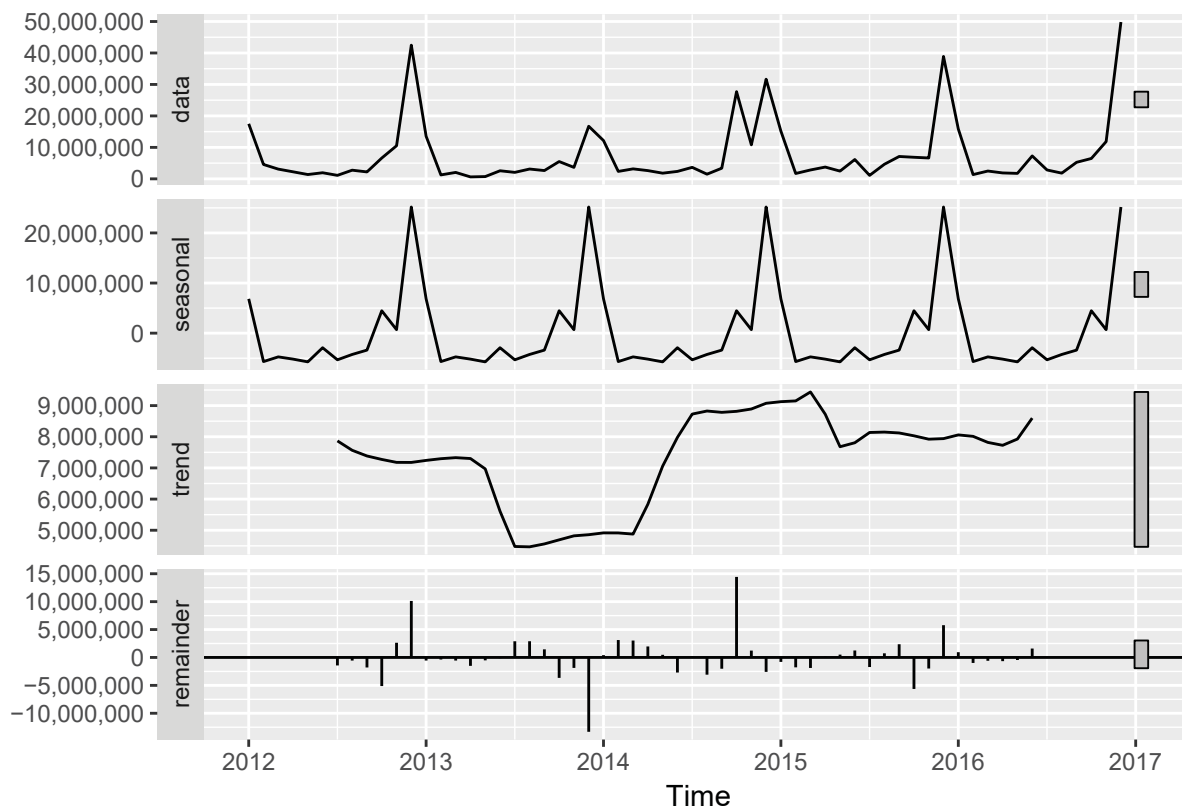
AFSPC



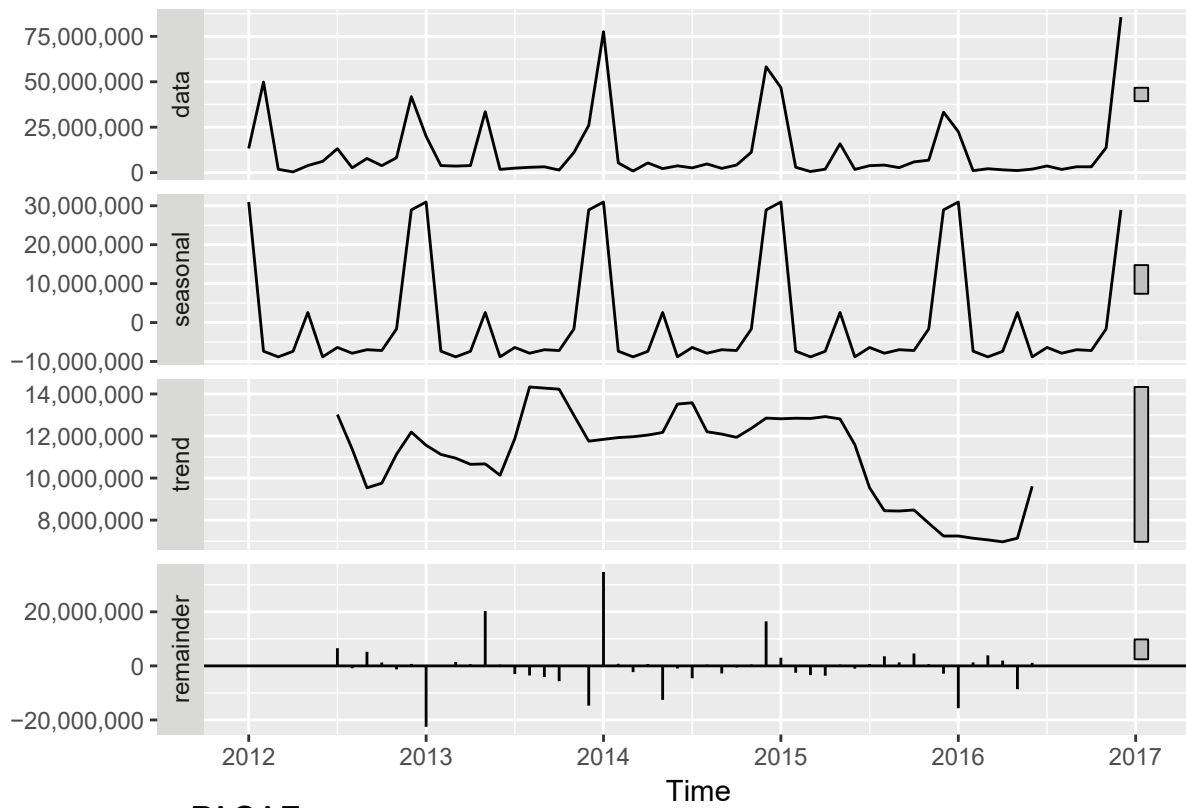
ACC



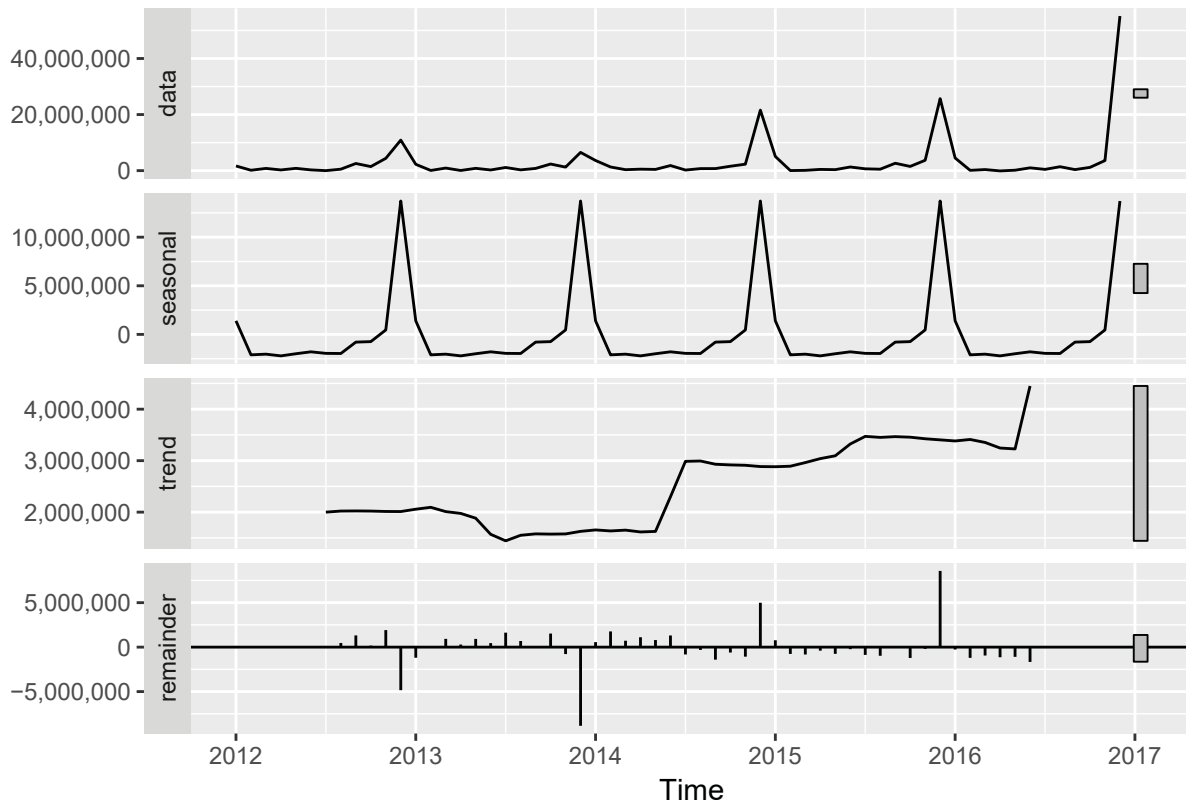
AETC



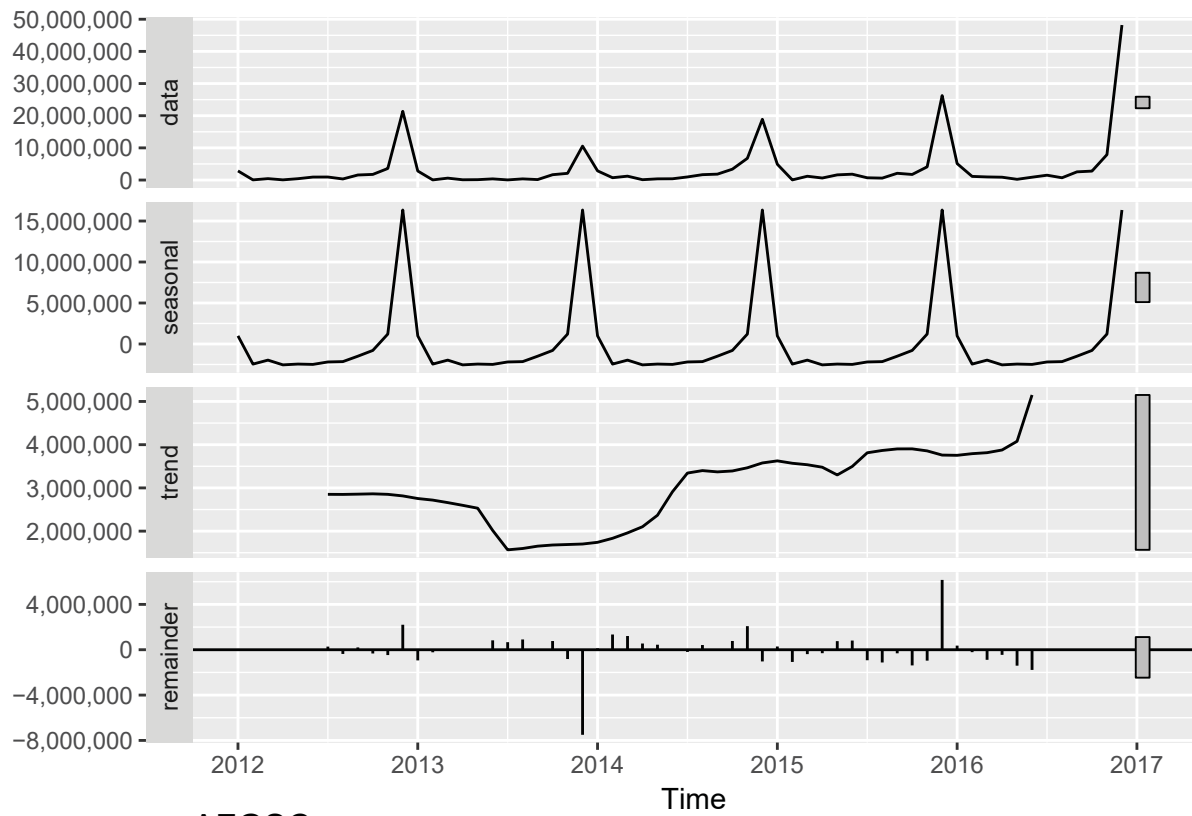
AMC



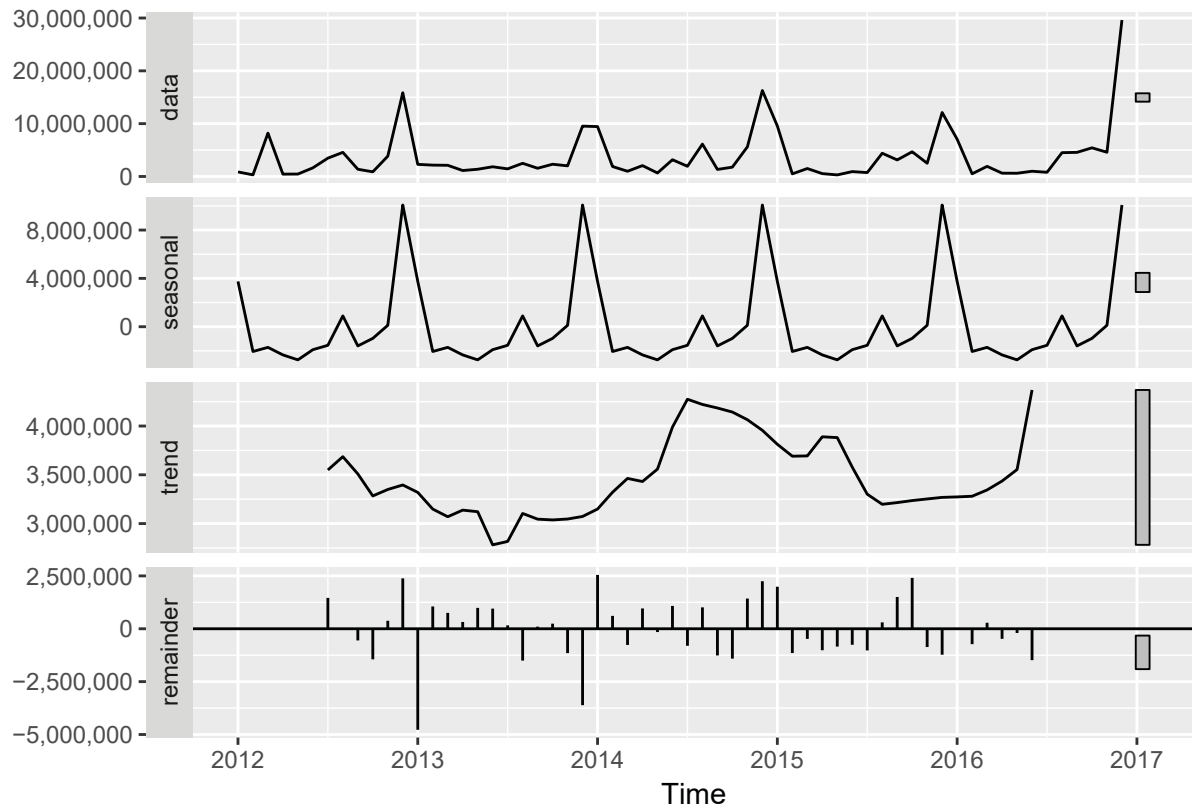
PACAF

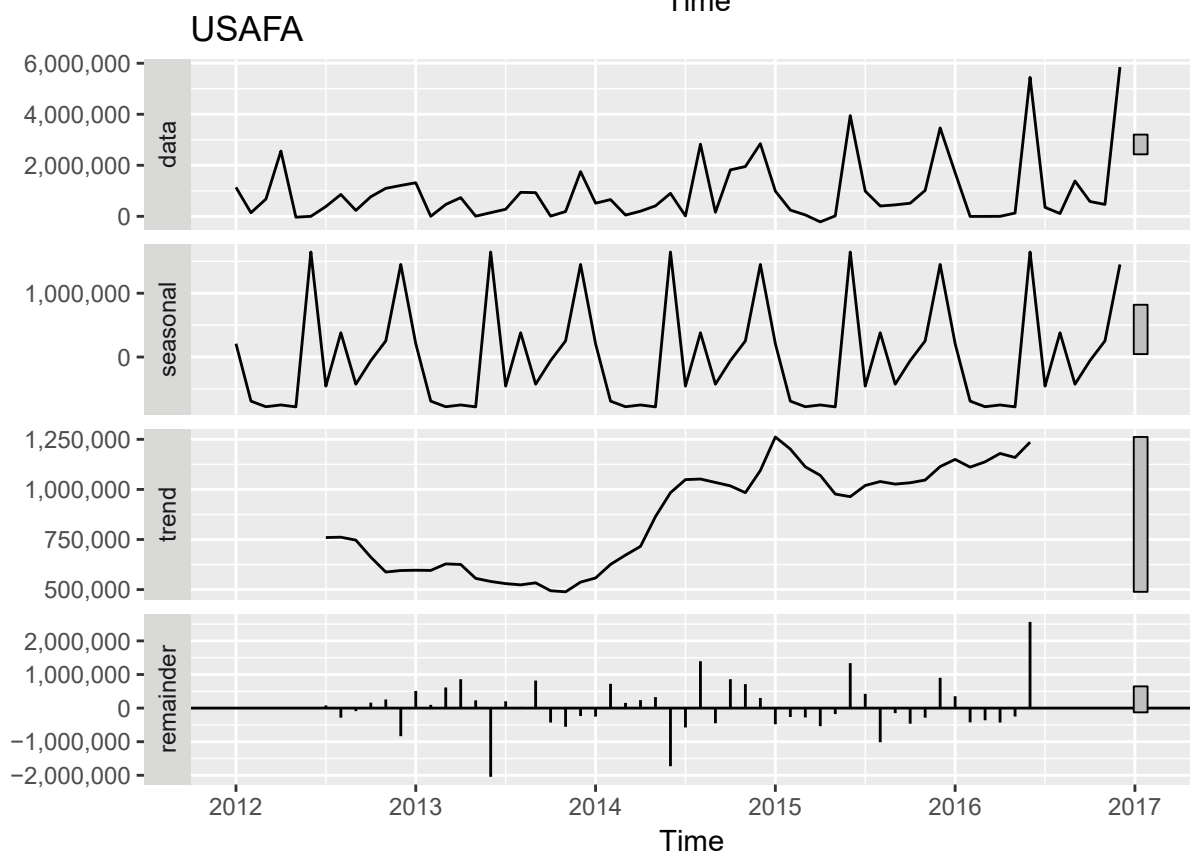
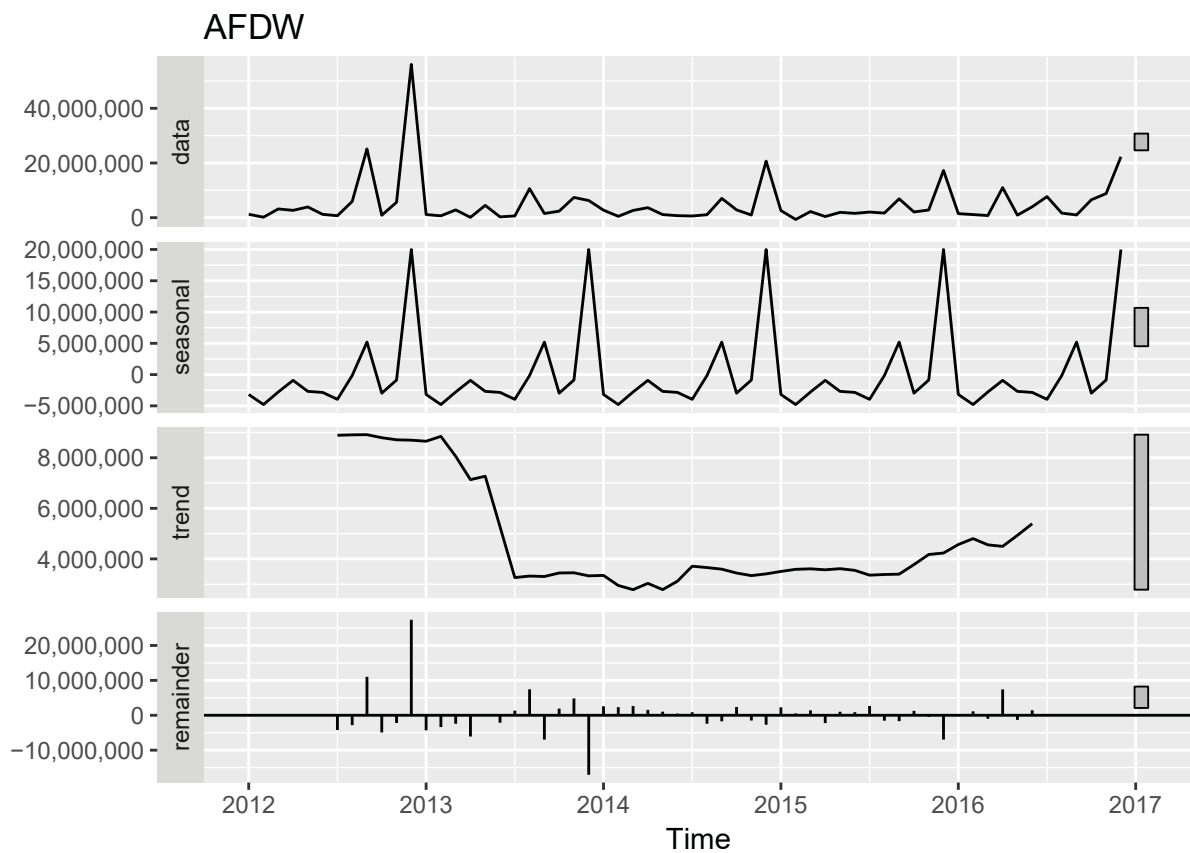


USAFE

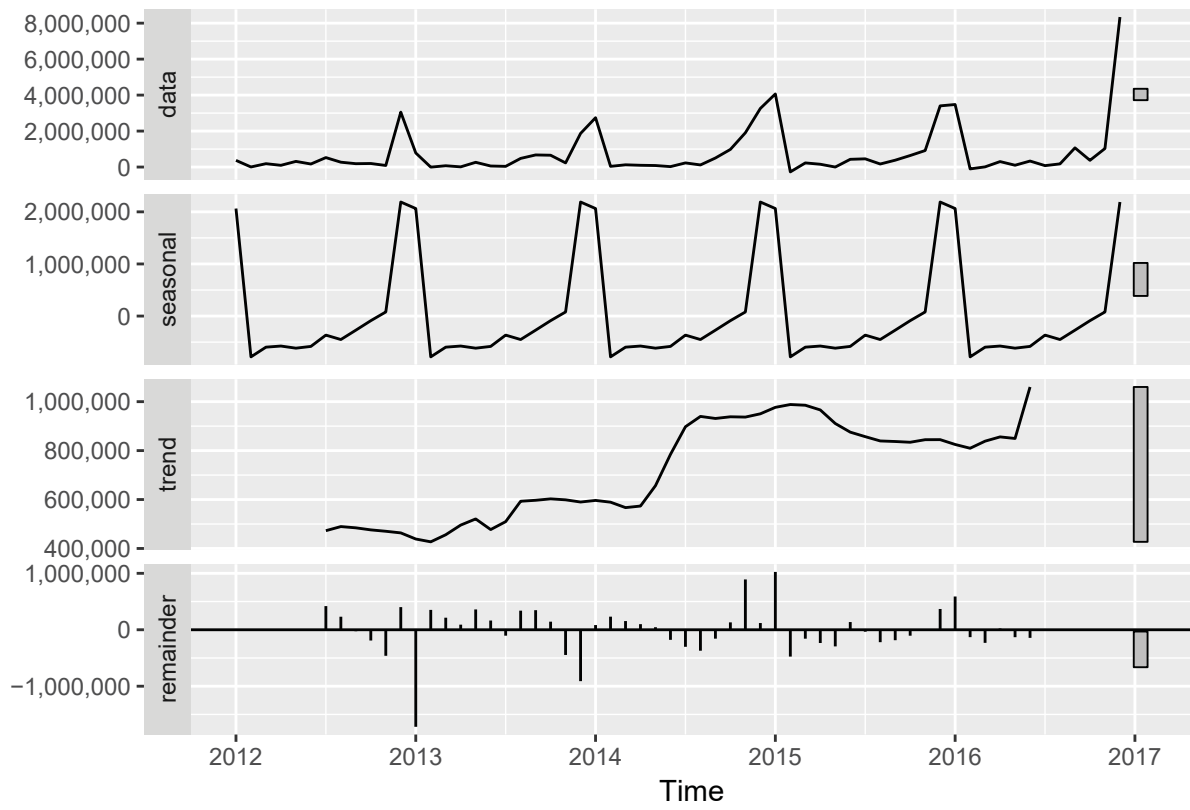


AFGSC





AFSOC



Appendix B - Model Optimizations

Professional Services

	RMSE	MAPE	MASE
Aggregate Model	176847718	22.78301	1.09650
Naive	193837348	56.64253	1.63008
Seasonal Naive	188898808	32.84739	1.40674
Trailing Moving Avg	182076986	38.32231	3.34882
Exponential Smoothing	189918589	24.29707	1.21527
Linear Regression	193531448	22.53668	1.14687
ARIMA	176830916	22.76123	1.09685

Facilities & Construction

	RMSE	MAPE	MASE
Aggregate Model	149243611	30.12579	0.88416
Naive	608709576	336.43439	5.40613
Seasonal Naive	99561303	39.41571	0.74309
Trailing Moving Avg	192531024	54.14464	1.81794
Exponential Smoothing	138391915	31.03056	0.88545
Linear Regression	103581324	27.49570	0.71715
ARIMA	149631590	29.83864	0.87817

IT

	RMSE	MAPE	MASE
Aggregate Model	108310449	25.87481	1.41665
Naive	344870103	319.22542	7.46024
Seasonal Naive	116337127	27.91062	1.45843
Trailing Moving Avg	213388364	55.84477	4.35671
Exponential Smoothing	115337127	34.35045	1.58489
Linear Regression	111882555	34.13006	1.57696
ARIMA	108346413	25.89562	1.41718

Transportation and Logistics Services

	RMSE	MAPE	MASE
Aggregate Model	41051202	106.25224	0.91826
Naive	67613460	499.03846	1.89555
Seasonal Naive	42818354	95.39196	0.93200
Trailing Moving Avg	46762921	160.54637	2.80117
Exponential Smoothing	46274813	110.52832	1.04781
Linear Regression	37829987	102.43171	0.96000
ARIMA	45903386	173.67741	0.98131

Industrial Products & Services

	RMSE	MAPE	MASE
Aggregate Model	15105839	79.32802	0.88557
Naive	51345731	429.40906	3.21561
Seasonal Naive	19206513	91.54917	0.89219
Trailing Moving Avg	19492765	147.47177	2.95537
Exponential Smoothing	15925068	73.24762	0.91300
Linear Regression	12349725	54.76884	0.68048
ARIMA	15409601	95.49252	0.84568

Human Capital

	RMSE	MAPE	MASE
Aggregate Model	13265203	34.89373	0.88062
Naive	22994069	96.94321	1.67456
Seasonal Naive	15546791	38.18210	0.90550
Trailing Moving Avg	12297573	33.52511	2.62223
Exponential Smoothing	13900610	36.35674	0.93350
Linear Regression	14088415	34.71879	0.91791
ARIMA	11814526	31.01576	0.77896

Office Management

	RMSE	MAPE	MASE
Aggregate Model	7304940	145.5204	0.47525
Naive	39711845	2033.2395	3.29894
Seasonal Naive	6904771	129.3119	0.41869
Trailing Moving Avg	18943661	447.1787	1.72840
Exponential Smoothing	24199923	403.6952	1.64375
Linear Regression	20198598	255.5081	1.12209
ARIMA	7563874	251.6069	0.55457

Security and Protection

	RMSE	MAPE	MASE
Aggregate Model	4264231	168.1243	1.21305
Naive	14080022	1169.0599	5.29851
Seasonal Naive	4242583	174.2235	1.22722
Trailing Moving Avg	4860271	305.0096	4.25465
Exponential Smoothing	4539179	178.8420	1.44987
Linear Regression	4405850	150.6669	1.33263
ARIMA	4242583	174.2235	1.22722

Travel & Lodging

	RMSE	MAPE	MASE
Aggregate Model	4320418	148.6054	1.28159
Naive	6211605	165.1724	1.99917
Seasonal Naive	4029151	154.6150	1.16980
Trailing Moving Avg	4772404	166.5473	7.11027
Exponential Smoothing	4687024	139.7411	1.40182
Linear Regression	4627343	135.7871	1.29629
ARIMA	4254541	132.7498	1.21258

Medical

	RMSE	MAPE	MASE
Aggregate Model	1247650	94.84930	1.26808
Naive	2591458	206.12301	1.91357
Seasonal Naive	1411301	103.39715	1.44309
Trailing Moving Avg	2699388	146.30714	4.14126
Exponential Smoothing	1420121	75.59245	1.33983
Linear Regression	1641245	63.28321	1.48182
ARIMA	1826788	106.91233	1.64745

IT Consulting

	RMSE	MAPE	MASE
Aggregate Model	10056512	256.5140	3.629093
Naive	19242914	820.8104	7.730452
Seasonal Naive	9041669	121.9302	2.984101
Trailing Moving Avg	9867399	137.3036	5.960836
Exponential Smoothing	11021428	489.5273	4.269752
Linear Regression	11054504	117.2539	3.467666
ARIMA	11632260	528.2114	4.491553

IT Hardware

	RMSE	MAPE	MASE
Aggregate Model	86834542	40.16997	1.933882
Naive	196436213	1875.73429	11.224205
Seasonal Naive	91083151	36.74996	2.029456
Trailing Moving Avg	142396757	290.78316	6.945169
Exponential Smoothing	103996413	69.21081	2.474036
Linear Regression	92921312	69.34774	2.306881
ARIMA	86883451	39.34735	1.930331

IT Outsourcing

	RMSE	MAPE	MASE
Aggregate Model	30325658	27.96105	0.9350960
Naive	95211403	145.18730	3.1909137
Seasonal Naive	35164652	33.88092	1.1108785
Trailing Moving Avg	56287201	43.53787	3.8699078
Exponential Smoothing	33369192	29.75566	0.9717886
Linear Regression	33195721	28.85999	0.9400781
ARIMA	53060707	37.47335	1.3987537

IT Security

	RMSE	MAPE	MASE
Aggregate Model	1934009.4	Inf	1.2023925
Naive	1167548.2	Inf	0.4696658
Seasonal Naive	1774802.1	Inf	0.6554274
Trailing Moving Avg	1311689.8	Inf	3.6034875
Exponential Smoothing	1241091.5	Inf	0.9037940
Linear Regression	567356.7	Inf	0.3838939
ARIMA	1006992.8	Inf	0.6076852

IT Software

	RMSE	MAPE	MASE
Aggregate Model	15128492	49.31138	0.9468986
Naive	33374047	345.81690	3.0259599
Seasonal Naive	15933118	53.55105	1.0797239
Trailing Moving Avg	17848292	93.62074	3.7362530
Exponential Smoothing	15756991	81.68800	1.1517207
Linear Regression	15061014	67.68268	1.0546407
ARIMA	15933118	53.55105	1.0797239

Telecommunications

	RMSE	MAPE	MASE
Aggregate Model	1295314	333.01355	1.241797
Naive	3626119	1358.10262	3.847419
Seasonal Naive	1253882	97.22056	0.939257
Trailing Moving Avg	1668687	482.49416	4.112775
Exponential Smoothing	3286002	1178.40468	3.382243
Linear Regression	1211736	179.85837	1.094982
ARIMA	3112662	974.52409	3.230704

AFMC

	RMSE	MAPE	MASE
Aggregate Model	202366954	14.75343	0.4962440
Naive	630808283	93.70069	2.1929698
Seasonal Naive	247034916	24.39503	0.7710601
Trailing Moving Avg	312853003	33.46084	1.9891459
Exponential Smoothing	248138080	17.95027	0.5863542
Linear Regression	207398127	15.43534	0.5122148
ARIMA	313569521	32.35524	0.9150031

AFSPC

	RMSE	MAPE	MASE
Aggregate Model	70651115	35.32708	0.9398463
Naive	153175363	180.84544	3.1648111
Seasonal Naive	63788231	42.01358	1.0558543
Trailing Moving Avg	99216172	63.90427	3.8900660
Exponential Smoothing	71292529	27.69006	0.8634516
Linear Regression	69068628	27.53942	0.8388226
ARIMA	71660629	35.67877	0.9492524

ACC

	RMSE	MAPE	MASE
Aggregate Model	66125507	30.84827	1.2185448
Naive	136362271	131.15145	3.1290745
Seasonal Naive	70982777	29.57561	1.1822880
Trailing Moving Avg	105572867	41.32432	3.4362506
Exponential Smoothing	51283738	25.95841	0.8581576
Linear Regression	47903351	25.26828	0.8681556
ARIMA	104435677	44.13091	1.5120521

AETC

	RMSE	MAPE	MASE
Aggregate Model	30269258	25.82360	0.5706436
Naive	76252817	123.18942	1.7532266
Seasonal Naive	29690524	30.12483	0.6055570
Trailing Moving Avg	43197454	47.03237	2.4960488
Exponential Smoothing	34807924	21.21960	0.5748158
Linear Regression	25352559	16.40036	0.4274971
ARIMA	42531033	39.88527	0.8017432

AMC

	RMSE	MAPE	MASE
Aggregate Model	20042525	29.34142	0.7067173
Naive	224275172	914.41228	11.5297918
Seasonal Naive	18792764	31.79262	0.7466589
Trailing Moving Avg	77122676	119.94807	6.9677572
Exponential Smoothing	21613041	33.61662	0.7569086
Linear Regression	17675666	26.44697	0.6504199
ARIMA	19845841	29.33014	0.7033873

PACAF

	RMSE	MAPE	MASE
Aggregate Model	23179815	41.57031	1.0506663
Naive	94409242	420.02643	6.3468591
Seasonal Naive	26219186	44.04403	1.1513706
Trailing Moving Avg	49234202	66.15784	3.0210335
Exponential Smoothing	18285415	41.94620	0.9958208
Linear Regression	23893876	42.16182	1.1301837
ARIMA	26219186	44.04403	1.1513706

USAFE

	RMSE	MAPE	MASE
Aggregate Model	23741963	65.63483	0.8196597
Naive	63922547	671.46746	3.5648068
Seasonal Naive	26777985	85.31279	0.9945806
Trailing Moving Avg	36038727	175.36555	5.3546863
Exponential Smoothing	25141634	50.40344	0.7141535
Linear Regression	27416005	59.01424	0.8831066
ARIMA	26058255	57.52595	0.8389898

AFGSC

	RMSE	MAPE	MASE
Aggregate Model	30378610	48.06147	1.1668064
Naive	66486668	332.22390	4.7135726
Seasonal Naive	11501946	41.23337	0.7372459
Trailing Moving Avg	19436792	58.43460	1.7060898
Exponential Smoothing	20206897	39.51477	0.9515558
Linear Regression	17411694	36.28676	0.8502770
ARIMA	9396464	33.41467	0.5917267

AFDW

	RMSE	MAPE	MASE
Aggregate Model	14887396	153.4703	0.7297111
Naive	55837196	683.9119	3.1642203
Seasonal Naive	12525113	152.3165	0.6485959
Trailing Moving Avg	26383472	189.6455	4.0121571
Exponential Smoothing	20837002	114.7757	0.7032711
Linear Regression	18918478	121.3499	0.6904058
ARIMA	14887515	153.4692	0.7297115

USAFA

	RMSE	MAPE	MASE
Aggregate Model	8213764	307.2159	0.5878490
Naive	36078521	5320.3648	4.1551348
Seasonal Naive	9055689	135.9737	0.6195996
Trailing Moving Avg	11400774	1229.4822	4.7580567
Exponential Smoothing	8856417	311.2140	0.7861183
Linear Regression	8382655	321.6223	0.7130262
ARIMA	7730589	354.3067	0.6005144

AFSOC

	RMSE	MAPE	MASE
Aggregate Model	3417559	142.35798	0.8659930
Naive	21440770	2485.98260	6.1081258
Seasonal Naive	3447909	49.86898	0.7317169
Trailing Moving Avg	8852365	599.67584	4.3791005
Exponential Smoothing	3685713	146.88225	0.8493454
Linear Regression	3397386	132.88556	0.7581783
ARIMA	3944408	130.02150	0.7738289

Al Dhafra AB, AE

	RMSE	MAPE	MASE
Aggregate Model	2263191	160.9250	1.344024
Naive	6687188	1539.7216	5.746946
Seasonal Naive	2465366	186.8485	1.407078
Trailing Moving Avg	1572482	271.4186	2.877982
Exponential Smoothing	2015717	151.3712	1.115558
Linear Regression	2049571	142.1741	1.109680
ARIMA	2259082	156.6488	1.333171

Al Udeid AB, QA

	RMSE	MAPE	MASE
Aggregate Model	1886120	2618.625	1.205151
Naive	2807187	9174.430	2.575642
Seasonal Naive	1909150	2803.357	1.235483
Trailing Moving Avg	1897001	4489.877	4.190826
Exponential Smoothing	2115235	1839.540	1.272147
Linear Regression	1904647	1883.892	1.141155
ARIMA	1909150	2803.357	1.235483

Ali Al Salem AB, KW

	RMSE	MAPE	MASE
Aggregate Model	587615.8	124.8861	0.5552548
Naive	1245860.9	724.5405	1.5209177
Seasonal Naive	719756.1	129.9453	0.6485232
Trailing Moving Avg	947465.5	497.2913	3.5699375
Exponential Smoothing	850701.8	473.2483	1.0529890
Linear Regression	575910.5	172.4752	0.5360750
ARIMA	863299.8	483.9894	1.0698448

Altus AFB, OK

	RMSE	MAPE	MASE
Aggregate Model	554477.3	42.93140	0.3803332
Naive	5080991.7	771.04824	4.1772382
Seasonal Naive	1517353.7	71.86179	0.8320562
Trailing Moving Avg	2782092.4	133.23490	2.2737530
Exponential Smoothing	643132.6	31.98350	0.3823444
Linear Regression	470696.8	30.85859	0.3159010
ARIMA	1931010.0	68.36195	0.8084126

Anderson AB, GU

	RMSE	MAPE	MASE
Aggregate Model	713368.4	242.1238	0.6928373
Naive	3098349.4	1600.3113	3.8051966
Seasonal Naive	1710228.8	139.9028	1.1527406
Trailing Moving Avg	2410046.0	380.4185	3.4364614
Exponential Smoothing	1344264.3	177.2994	1.0157696
Linear Regression	1284577.8	158.6883	0.9395671
ARIMA	1710228.8	139.9028	1.1527406

Andrews AFB, MD

	RMSE	MAPE	MASE
Aggregate Model	9277637	263.5662	0.6196287
Naive	27819556	1525.8335	2.7283128
Seasonal Naive	3886290	160.8009	0.3038240
Trailing Moving Avg	19225702	502.7929	9.9715390
Exponential Smoothing	19255245	430.1027	1.2013521
Linear Regression	15311529	152.7611	0.8779601
ARIMA	5398274	193.2538	0.3726441

Ramstein AB, GE

	RMSE	MAPE	MASE
Aggregate Model	21596594	93.50169	0.7934088
Naive	36687796	560.43379	2.2475262
Seasonal Naive	23970417	119.94896	0.9760251
Trailing Moving Avg	28928050	202.09204	5.7709318
Exponential Smoothing	28906471	204.09860	1.3086099
Linear Regression	25120351	86.85823	0.8910820
ARIMA	21633306	93.70735	0.7986069

Aviano AB , IT

	RMSE	MAPE	MASE
Aggregate Model	905999.9	294.9035	0.7670745
Naive	7304952.8	5984.3433	8.3786330
Seasonal Naive	644160.6	145.6279	0.5380042
Trailing Moving Avg	1964176.2	1032.8758	2.9458154
Exponential Smoothing	665047.5	325.7404	0.5878855
Linear Regression	624094.1	251.0362	0.5411196
ARIMA	907900.8	295.6559	0.7694761

Barksdale AFB, LA

	RMSE	MAPE	MASE
Aggregate Model	3801763	84.54125	1.0897295
Naive	9385418	548.69545	3.4567655
Seasonal Naive	5401787	90.50430	1.5834065
Trailing Moving Avg	6680807	205.29293	5.0511875
Exponential Smoothing	3373717	74.37084	0.9928711
Linear Regression	3423675	69.59437	0.9996648
ARIMA	5521707	109.22001	1.6149393

Beale AFB, CA

	RMSE	MAPE	MASE
Aggregate Model	4801067	90.96162	2.099925
Naive	6606934	1014.20763	3.897422
Seasonal Naive	4494823	87.71239	1.963298
Trailing Moving Avg	4237425	281.83080	4.218967
Exponential Smoothing	4284938	103.43543	2.085629
Linear Regression	4170865	153.04640	2.078007
ARIMA	4617200	127.93033	2.143595

Bolling AFB, DC

	RMSE	MAPE	MASE
Aggregate Model	7551833	127.6938	0.6185407
Naive	30979066	624.4404	3.0308088
Seasonal Naive	10310027	149.9534	0.8626005
Trailing Moving Avg	10656054	168.4272	2.0180635
Exponential Smoothing	7313357	107.1538	0.5968943
Linear Regression	8061686	131.7732	0.6737471
ARIMA	8562775	138.7285	0.7130226

Buckley AFB, CO

	RMSE	MAPE	MASE
Aggregate Model	1000483.2	166.2800	0.7090073
Naive	1431572.4	1542.3001	1.0220632
Seasonal Naive	1033305.9	188.7931	0.6572532
Trailing Moving Avg	1271333.5	1050.9966	1.4827793
Exponential Smoothing	858732.5	185.1366	0.6459545
Linear Regression	751350.0	162.7344	0.5513653
ARIMA	1134488.3	746.8991	0.7053805

Cannon AFB, NM

	RMSE	MAPE	MASE
Aggregate Model	1714400	99.60742	0.9311769
Naive	12336680	1382.39580	8.3799497
Seasonal Naive	1793098	94.09765	0.9296849
Trailing Moving Avg	3453277	251.44052	5.3127487
Exponential Smoothing	1678177	86.84317	0.8482601
Linear Regression	1545470	74.94845	0.7548237
ARIMA	1690999	97.69466	0.9062932

Columbus AFB, MS

	RMSE	MAPE	MASE
Aggregate Model	4070024	65.16379	0.6549592
Naive	4411092	1084.64245	1.0626853
Seasonal Naive	4598038	61.52515	0.7914429
Trailing Moving Avg	4370590	783.06503	1.6309488
Exponential Smoothing	4444310	716.86887	0.8524295
Linear Regression	4260705	1099.33337	0.6880578
ARIMA	4598038	61.52515	0.7914429

Davis Monthan AFB, AZ

	RMSE	MAPE	MASE
Aggregate Model	4943603	49.89010	0.9947261
Naive	14898412	1042.04465	6.1527530
Seasonal Naive	2596653	75.38338	0.7703201
Trailing Moving Avg	3610433	132.71252	1.8961641
Exponential Smoothing	3312122	44.50827	0.8283214
Linear Regression	2966091	49.47353	0.7598612
ARIMA	4942537	49.71175	0.9880343

Dover AFB, DE

	RMSE	MAPE	MASE
Aggregate Model	6468679	109.1489	1.1536215
Naive	29351310	30101.4270	15.3145606
Seasonal Naive	6587839	114.5956	1.2433581
Trailing Moving Avg	3220313	3424.3007	2.4623894
Exponential Smoothing	1446780	314.1384	0.3932439
Linear Regression	2465819	219.1767	0.5637249
ARIMA	6587839	114.5956	1.2433581

Dyess AFB, TX

	RMSE	MAPE	MASE
Aggregate Model	11126330	512.4918	1.2616819
Naive	9509304	3228.1659	1.8675799
Seasonal Naive	3528923	106.3870	0.3733534
Trailing Moving Avg	3469969	610.4344	0.7701635
Exponential Smoothing	8339874	274.5378	0.8936749
Linear Regression	7203971	321.1200	0.7801640
ARIMA	3528923	106.3870	0.3733534

Edwards AFB, CA

	RMSE	MAPE	MASE
Aggregate Model	2684753	51.10690	1.294304
Naive	4944963	247.86737	2.853921
Seasonal Naive	2655159	55.96935	1.298429
Trailing Moving Avg	3327156	82.74532	4.107613
Exponential Smoothing	2516374	42.45148	1.139092
Linear Regression	2473671	40.00132	1.107723
ARIMA	2599656	60.56099	1.349306

Eglin AFB, FL

	RMSE	MAPE	MASE
Aggregate Model	8010090	61.40939	0.6883962
Naive	14423496	187.10426	1.4004653
Seasonal Naive	8762942	61.06800	0.7227340
Trailing Moving Avg	9282133	64.78701	1.2839099
Exponential Smoothing	8977557	67.38373	0.7176647
Linear Regression	8582578	91.03905	0.7055832
ARIMA	7892972	58.26659	0.6214084

Eielson AFB, AK

	RMSE	MAPE	MASE
Aggregate Model	2459289	229.0597	1.2791083
Naive	3060850	2271.0305	1.5833974
Seasonal Naive	2336385	123.4217	1.0761454
Trailing Moving Avg	1863883	1243.8406	1.6696055
Exponential Smoothing	2341235	492.2339	1.0519993
Linear Regression	2322841	116.6629	1.0456171
ARIMA	2027437	697.6873	0.9991734

El Segundo, CA

	RMSE	MAPE	MASE
Aggregate Model	42168273	310.7026	1.3157864
Naive	33877625	107.7530	0.8387606
Seasonal Naive	45557258	352.2475	1.4469758
Trailing Moving Avg	36045768	294.2895	4.9752342
Exponential Smoothing	35648807	251.6784	1.2136519
Linear Regression	32724275	228.6280	1.1125871
ARIMA	45557258	352.2475	1.4469758

Ellsworth AFB, SD

	RMSE	MAPE	MASE
Aggregate Model	1492299	65.23493	1.0408783
Naive	3818983	1433.49004	3.2479854
Seasonal Naive	1357846	81.65966	0.8208456
Trailing Moving Avg	1551261	474.07482	1.6434352
Exponential Smoothing	1743641	98.69744	1.1661718
Linear Regression	1466612	104.04207	1.1042569
ARIMA	1521811	102.38483	1.0661180

F E Warren AFB, WY

	RMSE	MAPE	MASE
Aggregate Model	4537763	86.35318	0.8876704
Naive	2706111	259.90672	0.7992309
Seasonal Naive	3142134	189.55762	0.8784702
Trailing Moving Avg	2779226	141.85433	1.0795023
Exponential Smoothing	2725508	161.33577	0.6219969
Linear Regression	1725919	94.09523	0.5523808
ARIMA	2668358	184.18066	0.6461069

Fairchild AFB, WA

	RMSE	MAPE	MASE
Aggregate Model	2739166	419.0202	0.8479146
Naive	22774987	4713.7619	9.0754078
Seasonal Naive	3451083	640.4248	0.9163711
Trailing Moving Avg	7476906	778.5501	3.3538605
Exponential Smoothing	3338963	510.5075	1.1519991
Linear Regression	3010958	201.1975	0.7778214
ARIMA	2753001	435.9711	0.8199938

Goodfellow AFB, TX

	RMSE	MAPE	MASE
Aggregate Model	1578771	126.3549	0.9477824
Naive	6140496	1019.4734	4.6351704
Seasonal Naive	1377958	115.7025	0.8683319
Trailing Moving Avg	2659993	233.4617	3.9831375
Exponential Smoothing	1704426	175.5186	1.0009527
Linear Regression	1243303	130.1594	0.7192528
ARIMA	1531407	138.5038	0.9725519

Grand Forks AFB, ND

	RMSE	MAPE	MASE
Aggregate Model	2427680	212.2303	0.5843155
Naive	13893064	2764.4286	5.4559836
Seasonal Naive	1924576	264.3111	0.6295659
Trailing Moving Avg	7919249	395.1532	3.1985279
Exponential Smoothing	1779498	189.0997	0.5997153
Linear Regression	1539754	158.2420	0.4960391
ARIMA	2427692	212.2306	0.5843175

Hanscom AFB, MA

	RMSE	MAPE	MASE
Aggregate Model	20338298	24.55267	0.6004600
Naive	40371989	170.21123	1.7492314
Seasonal Naive	23670240	28.71577	0.7063044
Trailing Moving Avg	29051558	93.99901	5.1142948
Exponential Smoothing	20125132	33.90269	0.7393914
Linear Regression	16699854	26.27387	0.5708108
ARIMA	20278349	43.30552	0.8087270

Heath, OH

	RMSE	MAPE	MASE
Aggregate Model	997782	126.45978	0.3309560
Naive	3083983	459.07844	1.1256571
Seasonal Naive	1433361	61.69498	0.4065903
Trailing Moving Avg	1400175	181.06085	3.2327688
Exponential Smoothing	1382455	202.49750	0.4619223
Linear Regression	1174429	158.29242	0.3694915
ARIMA	1387966	187.29554	0.4768311

Hill AFB, UT

	RMSE	MAPE	MASE
Aggregate Model	23824893	50.58706	0.8910779
Naive	63815028	183.29089	2.5077932
Seasonal Naive	28530861	47.85179	0.9197237
Trailing Moving Avg	39874873	50.10561	2.3041689
Exponential Smoothing	19631925	39.59438	0.7429829
Linear Regression	17004069	41.17305	0.6514140
ARIMA	38988171	57.74824	1.1718209

Hurlburt Field, FL

	RMSE	MAPE	MASE
Aggregate Model	2447836	35.13074	0.6855231
Naive	9531886	570.20860	3.4795597
Seasonal Naive	3555598	48.28635	0.9721744
Trailing Moving Avg	5871460	153.11255	4.7792911
Exponential Smoothing	2966808	48.21706	0.8682303
Linear Regression	2917615	52.95066	0.9199347
ARIMA	4011913	52.36599	0.9510466

Incirlik AB, TU

	RMSE	MAPE	MASE
Aggregate Model	2935143	62.66758	2.015175
Naive	6812019	532.03286	6.549580
Seasonal Naive	2990169	86.86718	2.265633
Trailing Moving Avg	4358985	61.45183	4.562906
Exponential Smoothing	2923457	64.14897	2.005755
Linear Regression	2981871	71.24786	2.107935
ARIMA	2932134	63.93181	2.027912

JB - San Antonio

	RMSE	MAPE	MASE
Aggregate Model	24656842	29.04700	0.5554157
Naive	37555820	97.78584	1.0289439
Seasonal Naive	24823759	30.32598	0.5611050
Trailing Moving Avg	32692545	54.35723	2.3886697
Exponential Smoothing	25334118	23.00946	0.4613190
Linear Regression	21436405	30.02435	0.4978118
ARIMA	25844932	43.05797	0.6182113

JB Charleston, SC

	RMSE	MAPE	MASE
Aggregate Model	7119027	258.9842	1.758810
Naive	21008214	6395.8016	9.041299
Seasonal Naive	7141236	260.5291	1.600509
Trailing Moving Avg	12228590	1075.1120	8.687026
Exponential Smoothing	7108334	263.1720	1.695347
Linear Regression	7200921	300.1013	1.749821
ARIMA	7691634	373.0842	1.847962

JB Elmendorf - Richardson, AK

	RMSE	MAPE	MASE
Aggregate Model	5559363	384.3531	0.5786570
Naive	10433186	861.7413	1.0899590
Seasonal Naive	5665787	359.6970	0.5408305
Trailing Moving Avg	6706739	406.1108	2.3442683
Exponential Smoothing	9814476	274.2999	1.0027609
Linear Regression	7816162	419.9850	0.7243214
ARIMA	4850613	400.9763	0.4649233

JB Langley-Eustis, VA

	RMSE	MAPE	MASE
Aggregate Model	3672719	151.4784	0.5652773
Naive	22570630	2937.5256	4.5115072
Seasonal Naive	4384259	148.2508	0.6963530
Trailing Moving Avg	9217959	767.8863	1.7017284
Exponential Smoothing	4190548	168.0715	0.6483149
Linear Regression	3986170	178.9711	0.6195142
ARIMA	3587010	153.8935	0.5506362

JB Pearl Harbor - Hickam, HI

	RMSE	MAPE	MASE
Aggregate Model	14734703	170.5921	0.9328580
Naive	34374533	3593.1139	4.3192090
Seasonal Naive	15853447	190.4813	1.0153594
Trailing Moving Avg	21425547	802.2995	2.0874774
Exponential Smoothing	21388306	865.6229	1.3859541
Linear Regression	16614648	171.7881	0.9504335
ARIMA	15853447	190.4813	1.0153594

JB- McGuire-Dix-Lakehurst, NJ

	RMSE	MAPE	MASE
Aggregate Model	7667505	243.6258	1.734655
Naive	12736171	656.4376	3.502490
Seasonal Naive	7769692	239.4279	1.829361
Trailing Moving Avg	7463083	240.4010	8.633675
Exponential Smoothing	6607955	225.6523	1.456649
Linear Regression	6892470	236.4747	1.515662
ARIMA	7769692	239.4279	1.829361

Kadena AB, JA

	RMSE	MAPE	MASE
Aggregate Model	2847472	440.3017	1.0387565
Naive	11957159	9108.2183	5.8219216
Seasonal Naive	2285341	269.2728	0.6608521
Trailing Moving Avg	4223751	1700.8395	4.2852187
Exponential Smoothing	3162487	247.6145	0.9308922
Linear Regression	2820501	506.7302	0.8520546
ARIMA	4146775	2184.3707	1.5724390

Keesler AFB, MS

	RMSE	MAPE	MASE
Aggregate Model	7479335	255.4256	0.8353319
Naive	5578378	176.3540	0.7406936
Seasonal Naive	6950638	151.4017	0.7343365
Trailing Moving Avg	5584020	185.7890	3.2494927
Exponential Smoothing	5582816	168.5225	0.7280961
Linear Regression	6339394	161.7159	0.7214349
ARIMA	5578376	176.2323	0.7404894

Kirtland AFB, NM

	RMSE	MAPE	MASE
Aggregate Model	9276802	39.67561	1.632564
Naive	31974570	496.49944	8.140600
Seasonal Naive	9652582	46.34116	1.760613
Trailing Moving Avg	8434338	62.82306	4.623558
Exponential Smoothing	8296913	40.42689	1.461116
Linear Regression	8084240	40.86324	1.436763
ARIMA	9652582	46.34116	1.760613

Lackland AFB TX

	RMSE	MAPE	MASE
Aggregate Model	32908241	313.8366	1.0363483
Naive	64253099	7057.0163	2.8184010
Seasonal Naive	25679345	307.8472	0.9137065
Trailing Moving Avg	19918257	2494.2508	1.5652107
Exponential Smoothing	25864872	1923.3112	0.9365408
Linear Regression	23844702	1298.3316	0.8742386
ARIMA	32929344	194.8296	1.0302523

Lajes Field Azores, PT

	RMSE	MAPE	MASE
Aggregate Model	551887.8	Inf	0.7249830
Naive	888088.0	Inf	1.5616014
Seasonal Naive	698472.9	Inf	0.7733802
Trailing Moving Avg	674586.8	Inf	1.8241427
Exponential Smoothing	1159978.7	Inf	2.1218085
Linear Regression	655550.2	Inf	0.7554684
ARIMA	620368.2	Inf	0.8649436

Laughlin AFB, TX

	RMSE	MAPE	MASE
Aggregate Model	2119983	66.62002	0.6297295
Naive	6231410	409.40653	2.2671684
Seasonal Naive	2363199	61.55439	0.6149680
Trailing Moving Avg	2148280	111.87726	2.5590565
Exponential Smoothing	2266884	86.65283	0.6756506
Linear Regression	2110998	56.69856	0.6893593
ARIMA	2363199	61.55439	0.6149680

Liittle Rock AFB, AR

	RMSE	MAPE	MASE
Aggregate Model	8098966	767.6142	1.7395934
Naive	19833152	10711.6113	6.9233386
Seasonal Naive	3169171	151.0740	0.6096876
Trailing Moving Avg	18715713	2700.9541	8.9866841
Exponential Smoothing	4002489	302.7523	0.8728976
Linear Regression	5659371	314.4490	1.0640363
ARIMA	4588794	822.4357	1.0774610

Luke AFB, AZ

	RMSE	MAPE	MASE
Aggregate Model	4208434	102.91030	1.0873537
Naive	7316295	376.87392	2.4624162
Seasonal Naive	5364185	115.25859	1.1411493
Trailing Moving Avg	2623686	137.87927	2.4527627
Exponential Smoothing	4213739	68.11572	0.8382301
Linear Regression	3485971	63.91218	0.7593227
ARIMA	4199202	101.11538	1.0988902

MacDill AFB, FL

	RMSE	MAPE	MASE
Aggregate Model	7185935	109.34577	1.256120
Naive	35768487	1856.38080	8.957772
Seasonal Naive	6700728	127.16638	1.332956
Trailing Moving Avg	11363910	328.97096	5.116258
Exponential Smoothing	5650583	82.53028	1.107967
Linear Regression	7049056	83.89026	1.069187
ARIMA	9179154	443.30060	2.111549

MAFB - Gunter Annex, AL

	RMSE	MAPE	MASE
Aggregate Model	9674479	43.74370	0.5310710
Naive	17678617	208.52628	1.0626904
Seasonal Naive	9924128	47.24701	0.5689979
Trailing Moving Avg	13735696	89.46436	1.5718304
Exponential Smoothing	12574676	108.76662	0.6095667
Linear Regression	11313368	55.89755	0.5899524
ARIMA	13767288	61.28246	0.7426028

Malmstrom AFB, MT

	RMSE	MAPE	MASE
Aggregate Model	3689467	79.28403	1.657040
Naive	2888042	127.13388	1.653029
Seasonal Naive	2756029	41.30504	1.097375
Trailing Moving Avg	3110897	49.52728	3.025216
Exponential Smoothing	11449590	445.37996	6.722535
Linear Regression	3545505	52.22819	1.496502
ARIMA	2866203	39.18385	1.159158

Maxwell AFB, AL

	RMSE	MAPE	MASE
Aggregate Model	12114328	100.83696	0.8456631
Naive	28082114	776.30016	2.7577389
Seasonal Naive	14970073	155.27199	0.9053462
Trailing Moving Avg	7211885	213.77581	1.8967416
Exponential Smoothing	6138046	165.21886	0.5440507
Linear Regression	7923554	79.29575	0.5805995
ARIMA	7401768	157.29786	0.5742688

McConnell AFB, KS

	RMSE	MAPE	MASE
Aggregate Model	3614276	68714.04	1.752315
Naive	19313545	213208.15	13.001906
Seasonal Naive	3766287	25818.29	1.404289
Trailing Moving Avg	2828105	25325.08	7.072317
Exponential Smoothing	3258921	52856.42	1.509246
Linear Regression	3229681	46377.91	1.436670
ARIMA	3614223	68715.54	1.752238

Minot AFB, ND

	RMSE	MAPE	MASE
Aggregate Model	4906616	111.82350	0.9270299
Naive	4421286	666.13003	1.9464215
Seasonal Naive	2489458	156.51029	0.7599455
Trailing Moving Avg	1833572	294.29035	1.1753405
Exponential Smoothing	4756064	123.31753	0.9692345
Linear Regression	3106030	94.96801	0.6721391
ARIMA	1654767	211.26645	0.5969632

Misawa AB, JA

	RMSE	MAPE	MASE
Aggregate Model	11998473	352.0972	2.484108
Naive	12235909	2500.0487	3.525057
Seasonal Naive	11676880	329.4768	2.387744
Trailing Moving Avg	12932043	1196.5613	8.452146
Exponential Smoothing	10633290	203.7430	2.240518
Linear Regression	10926925	199.2874	2.300870
ARIMA	12248215	664.7265	2.581735

Moody AFB,GA

	RMSE	MAPE	MASE
Aggregate Model	5538169	84.89712	1.0282259
Naive	12128484	955.98742	3.8856238
Seasonal Naive	5525970	83.73340	1.0209008
Trailing Moving Avg	7201239	254.67769	2.8257049
Exponential Smoothing	2879376	62.96138	0.5593462
Linear Regression	1444301	49.94522	0.3604037
ARIMA	5544933	99.70819	1.0111846

Mountain Home AFB, ID

	RMSE	MAPE	MASE
Aggregate Model	1386222	101.13402	0.6873646
Naive	5011471	3633.45810	3.1813961
Seasonal Naive	1520582	82.60182	0.7305947
Trailing Moving Avg	2726143	910.25952	2.3407125
Exponential Smoothing	1387819	159.51120	0.7350300
Linear Regression	1106266	82.69217	0.4808502
ARIMA	1776826	440.21321	0.8016761

Nellis AFB, NV

	RMSE	MAPE	MASE
Aggregate Model	5835422	102.01344	0.8938284
Naive	28630275	868.66827	6.0628522
Seasonal Naive	3802311	80.41512	0.6791559
Trailing Moving Avg	11181253	187.34415	5.7652372
Exponential Smoothing	4879661	79.17739	0.7793337
Linear Regression	4280925	72.47984	0.6929814
ARIMA	5798266	102.81972	0.9016419

Newport News, VA

	RMSE	MAPE	MASE
Aggregate Model	50551146	96.40909	1.687369
Naive	55419052	69.79357	1.632887
Seasonal Naive	47464769	83.55785	1.492502
Trailing Moving Avg	51133441	101.17746	5.932139
Exponential Smoothing	50940014	103.19097	1.720329
Linear Regression	46863016	83.65094	1.502628
ARIMA	51735159	95.49802	1.700187

Offutt AFB, NE

	RMSE	MAPE	MASE
Aggregate Model	9514028	26.61674	0.6379244
Naive	51011567	239.83799	4.2013384
Seasonal Naive	9665245	29.88632	0.6791600
Trailing Moving Avg	21612934	42.44647	3.2011385
Exponential Smoothing	9568412	33.84440	0.7136582
Linear Regression	8428321	29.69088	0.6298521
ARIMA	10854404	31.16982	0.7072762

Patrick AFB, FL

	RMSE	MAPE	MASE
Aggregate Model	21539009	58.22030	0.9276321
Naive	64513508	354.11004	3.9943443
Seasonal Naive	23848578	60.82532	1.0623281
Trailing Moving Avg	18877269	82.09063	3.8689262
Exponential Smoothing	14733649	45.88105	0.7416266
Linear Regression	15130230	43.09838	0.6928348
ARIMA	21342549	58.04196	0.9220957

Peterson AFB, CO

	RMSE	MAPE	MASE
Aggregate Model	63953048	177.5779	1.578385
Naive	60679764	342.8765	1.613884
Seasonal Naive	63694738	105.7983	1.469364
Trailing Moving Avg	68346455	230.4675	3.877626
Exponential Smoothing	60832155	399.1422	1.493371
Linear Regression	63036392	138.4446	1.564506
ARIMA	66172966	251.3664	1.675936

RAF Alconbury, UK

	RMSE	MAPE	MASE
Aggregate Model	327604.7	352.6600	0.6414594
Naive	1385110.1	3044.1482	3.1019747
Seasonal Naive	292781.7	200.3410	0.5170040
Trailing Moving Avg	1102577.1	1258.9316	4.8193926
Exponential Smoothing	270308.5	467.1960	0.5513051
Linear Regression	206892.2	324.2145	0.3981363
ARIMA	320549.0	347.9339	0.6299798

RAF Lakenheath , UK

	RMSE	MAPE	MASE
Aggregate Model	1322461	1209.0400	0.6738013
Naive	5059339	12957.3205	3.1342293
Seasonal Naive	1096331	2018.5993	0.6015089
Trailing Moving Avg	3402648	4553.8741	3.4235787
Exponential Smoothing	1291771	604.4428	0.6758499
Linear Regression	1321921	1017.3343	0.6819415
ARIMA	1130773	1372.5799	0.6221395

Robins AFB, GA

	RMSE	MAPE	MASE
Aggregate Model	93079967	45.64220	0.6674143
Naive	115274177	108.54485	0.9130421
Seasonal Naive	135195656	65.70652	0.9119845
Trailing Moving Avg	110637234	98.14434	2.6516204
Exponential Smoothing	110640829	48.57699	0.7762067
Linear Regression	97902638	47.51035	0.7282252
ARIMA	108161325	74.68168	0.7704148

Rome, NY

	RMSE	MAPE	MASE
Aggregate Model	8297245	Inf	0.7644495
Naive	73003438	Inf	7.9991962
Seasonal Naive	6003100	Inf	0.5165305
Trailing Moving Avg	22169143	Inf	4.7406831
Exponential Smoothing	9747440	Inf	0.8484888
Linear Regression	5209493	Inf	0.4909689
ARIMA	6577061	Inf	0.6204405

Schriever AFB, CO

	RMSE	MAPE	MASE
Aggregate Model	10171665	126.9608	1.347959
Naive	45940008	2543.3930	9.253794
Seasonal Naive	9456116	102.1757	1.109529
Trailing Moving Avg	7766153	430.8160	4.567331
Exponential Smoothing	12527638	123.3428	1.414767
Linear Regression	11928581	101.8094	1.440794
ARIMA	8881157	137.3302	1.189174

Scott AFB, IL

	RMSE	MAPE	MASE
Aggregate Model	18606979	113.7592	1.172426
Naive	43963623	1531.8893	4.873091
Seasonal Naive	17492544	168.6547	1.181852
Trailing Moving Avg	26867812	343.2241	6.344110
Exponential Smoothing	21313608	147.2246	1.255512
Linear Regression	19741442	174.8979	1.238588
ARIMA	18710938	162.4198	1.188254

Seymour Johnson AFB, NC

	RMSE	MAPE	MASE
Aggregate Model	7872572	584.2941	1.2039236
Naive	5124319	7730.0076	1.3250775
Seasonal Naive	5259742	698.2187	0.9070032
Trailing Moving Avg	5670511	1718.7428	1.4398929
Exponential Smoothing	5436832	242.7147	0.9009220
Linear Regression	5573595	1030.9595	0.9864223
ARIMA	5057182	3805.5514	1.0074846

Shaw AFB, SC

	RMSE	MAPE	MASE
Aggregate Model	5274147	111.9427	0.7897627
Naive	8731216	777.5031	1.3468869
Seasonal Naive	7474001	112.5172	0.8375741
Trailing Moving Avg	12926538	219.2545	8.1112772
Exponential Smoothing	6532083	496.6261	1.1532548
Linear Regression	2994081	110.1934	0.4654009
ARIMA	7501820	219.3018	0.8416740

Sheppard AFB, TX

	RMSE	MAPE	MASE
Aggregate Model	6663245	97.55281	0.5859775
Naive	8670191	244.39729	0.7580551
Seasonal Naive	9191747	110.58395	0.6808106
Trailing Moving Avg	6945020	330.79170	2.1845911
Exponential Smoothing	7345921	301.69817	0.6564843
Linear Regression	6969507	75.78271	0.5634457
ARIMA	6880800	337.29404	0.6251999

Spangdahlem AB, GE

	RMSE	MAPE	MASE
Aggregate Model	3002489	243.6276	1.614836
Naive	12668151	6054.0878	10.372258
Seasonal Naive	3288625	294.8238	1.694006
Trailing Moving Avg	2218924	724.1421	2.471182
Exponential Smoothing	2368255	164.2507	1.290957
Linear Regression	2553472	211.1790	1.415874
ARIMA	3003209	241.9921	1.617429

Tinker AFB, OK

	RMSE	MAPE	MASE
Aggregate Model	56976069	24.51955	0.5931110
Naive	68082791	44.71669	0.7709992
Seasonal Naive	74185335	40.45918	0.8828134
Trailing Moving Avg	67848604	30.79919	2.1783161
Exponential Smoothing	55644745	25.11785	0.5667767
Linear Regression	59990272	27.02369	0.6303578
ARIMA	64715943	33.17177	0.6933981

Travis AFB, CA

	RMSE	MAPE	MASE
Aggregate Model	2161069	2308.1770	0.9203884
Naive	13176688	23078.3470	5.7556250
Seasonal Naive	1828172	730.6226	0.7185467
Trailing Moving Avg	4725819	3832.0166	2.5062063
Exponential Smoothing	1394650	1970.4745	0.5542288
Linear Regression	1310398	1342.6552	0.4926681
ARIMA	2167977	2316.6481	0.9224524

Tyndall AFB FL

	RMSE	MAPE	MASE
Aggregate Model	16704049	87.48100	1.007443
Naive	20930654	270.09999	1.560857
Seasonal Naive	17164280	92.27449	1.040489
Trailing Moving Avg	15293035	135.44951	2.822768
Exponential Smoothing	19944524	94.60716	1.204846
Linear Regression	16715349	91.99172	1.037951
ARIMA	15587258	103.97259	0.999010

USAFA, CO

	RMSE	MAPE	MASE
Aggregate Model	8213764	307.2159	0.5878490
Naive	36078521	5320.3648	4.1551348
Seasonal Naive	9055689	135.9737	0.6195996
Trailing Moving Avg	11400774	1229.4822	4.7580567
Exponential Smoothing	8856417	311.2140	0.7861183
Linear Regression	8382655	321.6223	0.7130262
ARIMA	7730589	354.3067	0.6005144

Vandenberg AFB CA

	RMSE	MAPE	MASE
Aggregate Model	7792895	103.35758	0.6746775
Naive	13824177	263.03798	1.5549938
Seasonal Naive	8863532	147.92093	0.8280668
Trailing Moving Avg	5196543	97.31566	1.8828568
Exponential Smoothing	8120530	138.31804	0.7104193
Linear Regression	8365670	136.46642	0.7486280
ARIMA	6230410	112.80263	0.5605969

Whiteman AFB, MO

	RMSE	MAPE	MASE
Aggregate Model	3189886	210.2731	1.671199
Naive	7322919	918.2471	4.849682
Seasonal Naive	2527080	236.4272	1.508051
Trailing Moving Avg	2212183	271.4375	2.226641
Exponential Smoothing	3135018	182.5830	1.585176
Linear Regression	3031328	175.7666	1.526356
ARIMA	2268123	239.1679	1.358105

Wright Patterson AFB, OH

	RMSE	MAPE	MASE
Aggregate Model	114688297	41.31496	0.7402126
Naive	240921263	129.27334	1.7206867
Seasonal Naive	129219997	47.21796	0.7481494
Trailing Moving Avg	119141898	50.29521	2.4366742
Exponential Smoothing	125935840	53.29319	0.8837982
Linear Regression	119059516	44.67818	0.8031508
ARIMA	118499891	48.79217	0.7875289

Yokota AB, JA

	RMSE	MAPE	MASE
Aggregate Model	2913063	57.05014	0.6710483
Naive	31709349	2337.83841	10.6802685
Seasonal Naive	2785679	61.94516	0.6887981
Trailing Moving Avg	7180668	257.48572	2.0626570
Exponential Smoothing	3092666	64.45055	0.7044847
Linear Regression	3378517	66.80270	0.6841942
ARIMA	2785679	61.94516	0.6887981

Appendix C - Model MASE Tables

lvl1	Individual Category Models	Aggregate Model
Professional Services	1.0964981	1.0964981
Facilities & Construction	0.7171546	0.8841635
IT	1.4166475	1.4166475
Transportation and Logistics Services	0.9182604	0.9182604
Industrial Products & Services	0.6804809	0.8855707
Human Capital	0.7789602	0.8806200
Office Management	0.4186944	0.4752471
Security and Protection	1.2130530	1.2130530
Travel & Lodging	1.1697956	1.2815852
Medical	1.2680826	1.2680826

ITlvl2	Individual Category Models	Aggregate Model
IT Consulting	2.9841010	3.6290933
IT Hardware	1.9303314	1.9338820
IT Outsourcing	0.9350960	0.9350960
IT Security	0.3838939	1.2023925
IT Software	0.9468986	0.9468986
Telecommunications	0.9392570	1.2417972

MAJCOM	Individual Category Models	Aggregate Model
AFMC	0.4962440	0.4962440
AFSPC	0.8388226	0.9398463
ACC	0.8581576	1.2185448
AETC	0.4274971	0.5706436
AMC	0.6504199	0.7067173
PACAF	0.9958208	1.0506663
USAFE	0.7141535	0.8196597
AFGSC	0.5917267	1.1668064
AFDW	0.6485959	0.7297111
USAF	0.5878490	0.5878490
AFSOC	0.7317169	0.8659930

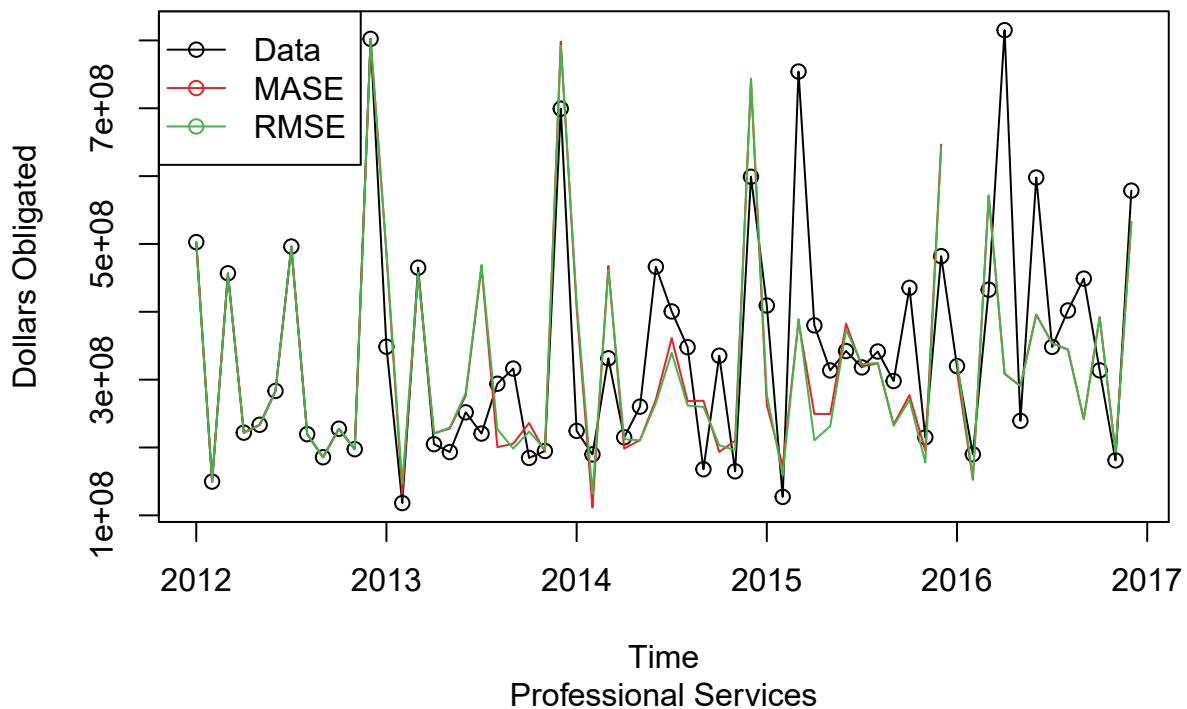
Base	Individual Category Models	Aggregate Model
Al Dhafra AB, AE	1.1096798	1.3440238
Al Udeid AB, QA	1.1411550	1.2051506
Ali Al Salem AB, KW	0.5360750	0.5552548
Altus AFB, OK	0.3159010	0.3803332
Anderson AB, GU	0.6928373	0.6928373
Andrews AFB, MD	0.3038240	0.6196287
Ramstein AB, GE	0.7934088	0.7934088
Aviano AB , IT	0.5380042	0.7670745
Barksdale AFB, LA	0.9928711	1.0897295
Beale AFB, CA	1.9632983	2.0999249
Bolling AFB, DC	0.5968943	0.6185407

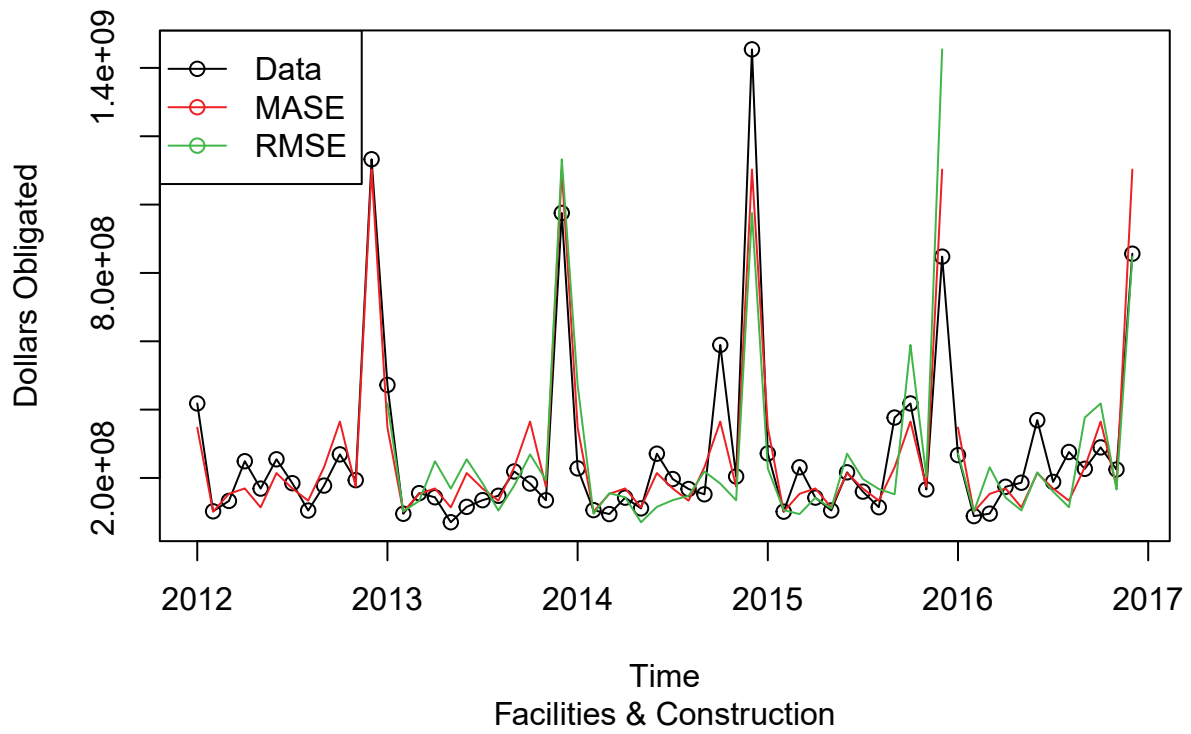
Base	Individual Category Models	Aggregate Model
Buckley AFB, CO	0.5513653	0.7090073
Cannon AFB, NM	0.7548237	0.9311769
Columbus AFB, MS	0.6549592	0.6549592
Davis Monthan AFB, AZ	0.7598612	0.9947261
Dover AFB, DE	0.3932439	1.1536215
Dyess AFB, TX	0.3733534	1.2616819
Edwards AFB, CA	1.1077233	1.2943042
Eglin AFB, FL	0.6214084	0.6883962
Eielson AFB, AK	0.9991734	1.2791083
El Segundo, CA	0.8387606	1.3157864
Ellsworth AFB, SD	0.8208456	1.0408783
F E Warren AFB, WY	0.5523808	0.8876704
Fairchild AFB, WA	0.7778214	0.8479146
Goodfellow AFB, TX	0.7192528	0.9477824
Grand Forks AFB, ND	0.4960391	0.5843155
Hanscom AFB, MA	0.5708108	0.6004600
Heath, OH	0.3309560	0.3309560
Hill AFB, UT	0.6514140	0.8910779
Hurlburt Field, FL	0.6855231	0.6855231
Incirlik AB, TU	2.0057548	2.0151751
JB - San Antonio	0.4613190	0.5554157
JB Charleston, SC	1.6005088	1.7588103
JB Elmendorf - Richardson, AK	0.4649233	0.5786570
JB Langley-Eustis, VA	0.5506362	0.5652773
JB Pearl Harbor - Hickam, HI	0.9328580	0.9328580
JB- McGuire-Dix-Lakehurst, NJ	1.4566490	1.7346552
Kadena AB, JA	0.6608521	1.0387565
Keesler AFB, MS	0.7214349	0.8353319
Kirtland AFB, NM	1.4367629	1.6325644
Lackland AFB TX	0.8742386	1.0363483
Lajes Field Azores, PT	0.7249830	0.7249830
Laughlin AFB, TX	0.6149680	0.6297295
Liittle Rock AFB, AR	0.6096876	1.7395934
Luke AFB, AZ	0.7593227	1.0873537
MacDill AFB, FL	1.0691866	1.2561198
MAFB - Gunter Annex, AL	0.5310710	0.5310710
Malmstrom AFB, MT	1.0973747	1.6570402
Maxwell AFB, AL	0.5440507	0.8456631
McConnell AFB, KS	1.4042888	1.7523153
Minot AFB, ND	0.5969632	0.9270299
Misawa AB, JA	2.2405178	2.4841076
Moody AFB,GA	0.3604037	1.0282259
Mountain Home AFB, ID	0.4808502	0.6873646
Nellis AFB, NV	0.6791559	0.8938284
Newport News, VA	1.4925019	1.6873690
Offutt AFB, NE	0.6298521	0.6379244
Patrick AFB, FL	0.6928348	0.9276321
Peterson AFB, CO	1.4693644	1.5783852
RAF Alconbury, UK	0.3981363	0.6414594
RAF Lakenheath , UK	0.6015089	0.6738013
Robins AFB, GA	0.6674143	0.6674143
Rome, NY	0.4909689	0.7644495

Base	Individual Category Models	Aggregate Model
Schriever AFB, CO	1.1095293	1.3479594
Scott AFB, IL	1.1724264	1.1724264
Seymour Johnson AFB, NC	0.9009220	1.2039236
Shaw AFB, SC	0.4654009	0.7897627
Sheppard AFB, TX	0.5634457	0.5859775
Spangdahlem AB, GE	1.2909573	1.6148359
Tinker AFB, OK	0.5667767	0.5931110
Travis AFB, CA	0.4926681	0.9203884
Tyndall AFB FL	0.9990100	1.0074434
USAF, CO	0.5878490	0.5878490
Vandenberg AFB CA	0.5605969	0.6746775
Whiteman AFB, MO	1.3581047	1.6711993
Wright Patterson AFB, OH	0.7402126	0.7402126
Yokota AB, JA	0.6710483	0.6710483

Appendix D - Error Measure Comparison

The following two graphs illustrate how well two different models selected by RMSE and MASE fit the data in the Professional Services category and the Facilities and Construction category. Numerically, when two different models are chosen between RMSE and MASE, the difference in error measures is minimal. Visually, too, there is no apparent pattern that optimal RMSE models exhibit compared to optimal MASE models. To reiterate, though RMSE and MAPE values are captured, this thesis focuses on optimizing models and comparing model performance across different series solely with MASE.





Appendix E - Base T-Test Results

```
##
## F test to compare two variances
##
## data: i.errorB$MASE and a.errorB$MASE
## F = 0.83462, num df = 76, denom df = 76, p-value = 0.4325
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 0.5305627 1.3129238
## sample estimates:
## ratio of variances
## 0.8346187

##
## Paired t-test
##
## data: MASE by Type
## t = 8.085, df = 76, p-value = 7.569e-12
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 0.1459241 0.2413181
## sample estimates:
## mean of the differences
## 0.1936211

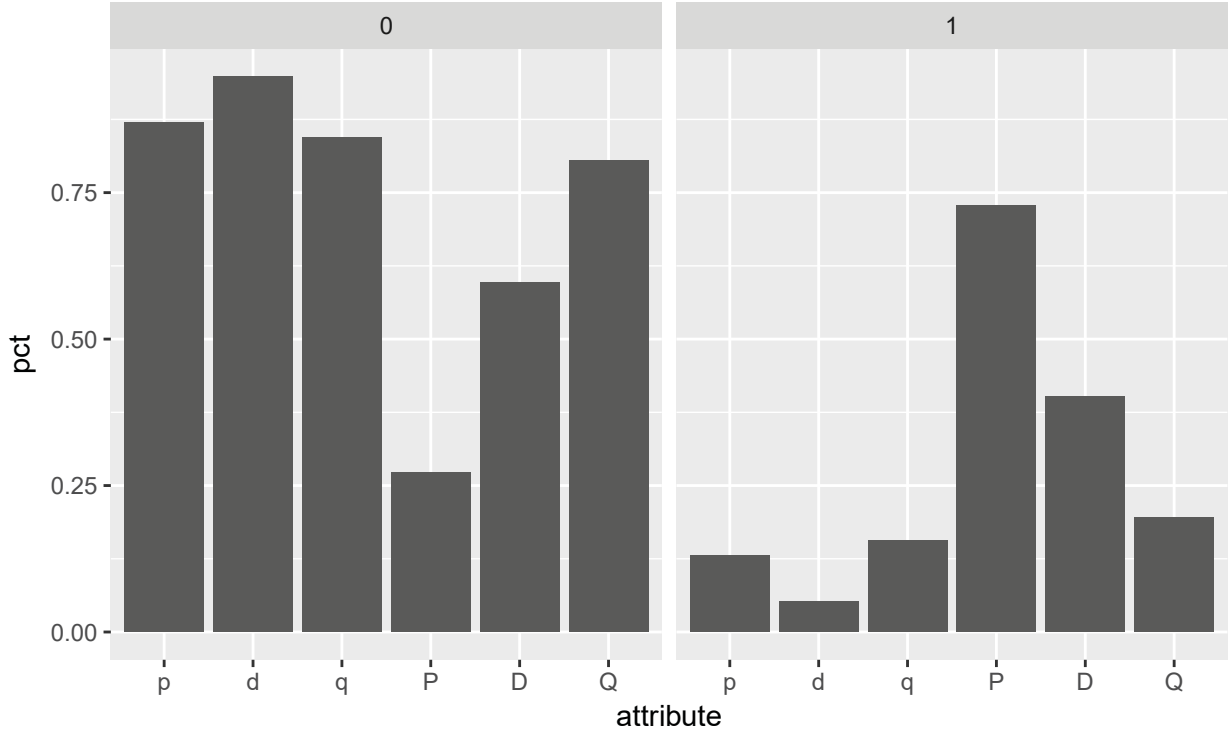
##
## Paired t-test
##
## data: MASE by Type
## t = 8.085, df = 76, p-value = 3.784e-12
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
## 0.1537437 Inf
## sample estimates:
## mean of the differences
## 0.1936211

##
## Wilcoxon signed rank test with continuity correction
##
## data: MASE by Type
## V = 2080, p-value = 3.61e-12
## alternative hypothesis: true location shift is not equal to 0
```

Appendix F - ARIMA Model Exploration

Common model attributes

The most common attribute across all bases is a first order seasonal AR term followed by first or This suggests that most base-level expenditures can be predicted by using information from the



model	n	pct	cumpct
c(0, 0, 0, 1, 0, 0)	27	0.3506494	0.3506494
c(0, 0, 0, 1, 1, 0)	14	0.1818182	0.5324675
c(0, 0, 0, 0, 1, 1)	7	0.0909091	0.6233766
c(0, 0, 0, 0, 0, 0)	4	0.0519481	0.6753247
c(1, 0, 0, 1, 0, 0)	4	0.0519481	0.7272727
c(0, 0, 1, 1, 0, 0)	3	0.0389610	0.7662338
c(1, 0, 0, 1, 1, 0)	3	0.0389610	0.8051948
c(0, 0, 0, 0, 1, 0)	2	0.0259740	0.8311688
c(0, 0, 0, 1, 0, 1)	2	0.0259740	0.8571429
c(0, 0, 1, 0, 0, 1)	2	0.0259740	0.8831169
c(0, 1, 1, 0, 1, 1)	2	0.0259740	0.9090909
c(0, 0, 0, 0, 0, 1)	1	0.0129870	0.9220779
c(0, 0, 1, 0, 0, 0)	1	0.0129870	0.9350649
c(0, 0, 1, 0, 1, 0)	1	0.0129870	0.9480519
c(0, 1, 1, 1, 0, 0)	1	0.0129870	0.9610390
c(1, 0, 0, 1, 1, 1)	1	0.0129870	0.9740260
c(1, 0, 1, 1, 0, 0)	1	0.0129870	0.9870130
c(1, 1, 1, 0, 1, 0)	1	0.0129870	1.0000000

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 22-03-2018		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From - To) 14 Aug 16 - 22 Mar 18	
4. TITLE AND SUBTITLE Air Force Installation Contracting Agency Category Management Through Expenditure Profiling				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Okamoto, James T, Capt				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENV-MS-18-M-228	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Installation Contracting Agency 1940 Allbrook Drive Area A, Building 1, Door 7B Wright-Patterson AFB, OH 45433				10. SPONSOR/MONITOR'S ACRONYM(S) AFICA/KA	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.					
14. ABSTRACT This thesis examines an approach to characterizing various expenditure profiles for the Air Force Installation Contracting Agency's Operations and Maintenance Appropriated Funds. Using naive, seasonal naive, trailing moving average, exponential smoothing, linear regression, and autoregressive integrated moving average (ARIMA) forecasting methods, the paper evaluates multiple error measures over one fiscal year to find the most precise model for each level of analysis. Levels of analysis included the Air Force enterprise and level 1 category levels, as well as an illustrative approach to Information and Technology (IT) spend at the level 2 subcategory, major command, and base levels. Optimal model characteristics were used to compare expenditure profile patterns at the different levels. In general, the more a unit can customize its algorithms, the more accurately it can capture its respective expenditure profile. The more localized the level of spend, the less applicable the aggregate models become, and different sub-groups have more personalized patterns.					
15. SUBJECT TERMS Category Management, Expenditure Profiling, Forecasting, Decision Support, Air Force Installation Contracting Agency					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 115	19a. NAME OF RESPONSIBLE PERSON Brandon Lucas, AFIT/ENV
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code) (937)255-3636 x4441