Predictive Analysis of Installation and Operational Qualification Issues vs. Process Severity Events
by
Bidusha Poudyal
B.S., Electrical Engineering
Illinois Institute of Technology, 2014

Submitted to the MIT Sloan School of Management and the Department of Electrical Engineering and Computer Science in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
and
Master of Science in Electrical Engineering and Computer Science

In conjunction with the Leaders for Global Operations Program at the
Massachusetts Institute of Technology
May 2020
© 2020 Bidusha Poudyal. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Author……………………………………………………………………………………….MIT Sloan School of Management
Department of Electrical Engineering and Computer Science
May 8, 2020

Certified by…………………………………………………………………………………………Duane S. Boning
Professor, Department of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by…………………………………………………………………………………………Negin (Nicki) Golrezaei
Assistant Professor of Operation Management
Thesis Supervisor

Accepted by…………………………………………………………………………………………Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Accepted by…………………………………………………………………………………………Maura Herson
Assistant Dean, MBA Program
MIT Sloan School of Management
This page has been intentionally left blank.
Predictive Analysis of Installation and Operational Qualification Issues vs. Process Severity Events
by
Bidusha Poudyal

Submitted to the MIT Sloan School of Management and the Department of Electrical Engineering and Computer Science on May 8, 2020 in Partial Fulfillment of the Requirements for the Degrees of Master of Business Administration and Master of Science in Electrical Engineering and Computer Science

Abstract

As the retail industry grows more popular, ABCD, a world-class electronic commerce (e-commerce) business, is increasingly building new Fulfillment Centers (FCs) to support this rapid demand growth. It is integral for ABCD to validate the installation quality and functionality of Material Handling Equipment (MHE) in these newly built FCs so operations can avoid errors. To achieve this objective, ABCD introduced the Installation and Operational Qualification (IOQ) process in late 2014. While the IOQ process reduces early operational failures, it does not completely eliminate them. Inadequate IOQ and tighter installation timelines are leading to degraded installation quality, resulting in operational issues and costs for ABCD. As the FC network continues to grow, there is a need to improve installation quality to reduce early operational issues and enhance the FC start-up experience.

This project is a part of the ABCD Operation Engineering teams’ effort to improve the existing IOQ process and the FC start-up experience. This initiative consists of three main phases. The first phase – the research phase – is dedicated to understanding the current processes and problem statement. It also includes a study of available data sources to discover failure patterns across different FCs. The second phase involves developing analytical frameworks and machine-learning models to uncover the most problematic equipment in the FC. The third phase focuses on evaluating the effectiveness of the current IOQ process based on Phase 1 and 2 findings, and identifying opportunities to better the process. The thesis summarizes the outcomes from all of these phases.

The project focuses on improving IOQ coverage, efficiently reprioritizing the testing schedule, introducing threshold metric for installation quality, and exploring predictive and preventative maintenance opportunities. This thesis also includes recommendations for refining the data-gathering process to improve future model outcomes. The ultimate goal is to improve FC installation quality and enhance the IOQ process to eliminate start-up issues. The approach taken and the recommendations proposed seek to approximate the ideal state as closely as possible. Incremental adoption of these recommendations will help deliver better-installed FCs, reduce early operational issues, improve start-up experiences, and strengthen ABCD’s infrastructure.

Thesis Supervisor: Duane S. Boning
Title: Professor, Department of Electrical Engineering and Computer Science

Thesis Supervisor: Negin (Nicki) Golrezaei
Title: Assistant Professor of Operation Management, MIT Sloan School of Management
This page has been intentionally left blank.
Acknowledgements

The completion of this thesis would not have been possible without the support of many individuals. I would like to express my gratitude to the following people:

My thesis advisors, Duane S. Boning and Negin (Nicki) Golrezaei, for their support and guidance throughout the project and while writing this document. Since the inception of my project, both took time each week to provide guidance. Their dedicated and consistent involvement throughout my work was instrumental to the project’s successes.

My project champion at ABCD, Kevin Lyons, for sponsoring this project, welcoming me into the organization and considering my input and recommendations.

My industry supervisors, Kieran Higgins and Dan Hartley, and their teams for providing valuable data, input, and thoughtful partnership.

The staff from the MIT Leaders for Global Operations program, who make the LGO program possible and enabled this incredible six-month industry experience.

My classmates who have made the past two years a once-in-a-lifetime experience.

Finally, I owe much of my accomplishments to my family’s love, support, and patience. Thank you to my parents, Bishnu and Muna, for always making my education a priority and for instilling in me the values of lifelong learning and hard work. And to my siblings, Binuda and Bimochan, for supporting me along the way.
Acronyms used

- AR: *ABCD* Robotics
- AFE: *ABCD* Fulfillment Engine
- BOW: Bag of Words
- CC: Control Cabinet
- CLT: Closed Loop Testing
- CTI: Category/Type/Item
- EAM: Enterprise Asset Management
- FC: Fulfillment Center
- HDP: Hierarchical Dirichlet Process
- HMI: Human Machine Interface
- IOQ: Installation and Operational Qualification Process
- IQ: Installation Qualification
- LDA: Latent Dirichlet Allocation
- LPH: Loss in Production Hours
- MHE: Material Handling and Equipment
- OQ: Operational Qualification
- PC1: First principal components
- PC2: Second principal components
- PM: Preventative Maintenance
- RCA: Root Cause Analysis
- RME: Reliability, Maintenance, and Engineering
- ROI: Return on Investment
- SDO: System Description of Operations
- SLAM: (Scan/Label/Apply/Manifest)
- TQ: Throughput Qualification
- VRC: Vertical Reciprocating Conveyor
- WCSS: Within Cluster Sum of Errors
Table of Contents

Abstract ................................................................................................................................. 2
Acknowledgements ............................................................................................................... 4
Acronyms used .................................................................................................................. 6
Table of Contents ................................................................................................................ 8
List of Tables and Figures ................................................................................................. 10
Chapter 1: Introduction ...................................................................................................... 14
  1.1 Background ................................................................................................................ 16
    1.1.1 ABCD Installation and Operational Qualification (IOQ) process ......................... 18
    1.1.2 ABCD Go-Live Process and the Trouble Ticket System ...................................... 20
  1.2 Problem Statement ..................................................................................................... 21
  1.3 Project Objective ........................................................................................................ 22
  1.4 Methodology Overview .............................................................................................. 22
    1.4.1 Existing process validation .................................................................................. 24
    1.4.2 Reactive lens ....................................................................................................... 24
    1.4.3 Preventive lens .................................................................................................... 24
    1.4.4 Predictive lens .................................................................................................... 25
Chapter 2: Data .................................................................................................................. 26
  2.1 Data Sources .............................................................................................................. 26
    2.1.1 Asset Data ......................................................................................................... 26
    2.1.2 Installation and Operational Qualification (IOQ) Red Tags Data ......................... 27
    2.1.3 Trouble Ticket Data .......................................................................................... 28
    2.1.4 Equipment Criticality Data ................................................................................ 30
    2.1.5 Replacement Parts Data .................................................................................... 30
    2.1.6 Root Cause Analysis (RCA) Data ..................................................................... 31
    2.1.7 Wireless Sensor Data ....................................................................................... 32
  2.2 Data Cleaning ............................................................................................................. 33
Chapter 3: Research Methodologies and Findings ............................................................ 35
  3.1 Analysis of existing Installation and Operational Qualification process ...................... 35
  3.2 Improving the Installation and Operational Qualification process .............................. 36
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.1 Installation and Operational Qualification testing coverage</td>
<td>53</td>
</tr>
<tr>
<td>3.2.2 Installation and Operational Qualification testing prioritization</td>
<td>59</td>
</tr>
<tr>
<td>3.3 Exploration of Preventive Opportunities</td>
<td>62</td>
</tr>
<tr>
<td>3.3.1 Leveraging vendor relationship</td>
<td>66</td>
</tr>
<tr>
<td>3.3.2 Study of replacement parts</td>
<td>68</td>
</tr>
<tr>
<td>3.4 Exploration of Predictive Maintenance Opportunities</td>
<td>69</td>
</tr>
<tr>
<td>3.5 Financial Analysis</td>
<td>72</td>
</tr>
<tr>
<td>Chapter 4: Future of Work and Conclusion</td>
<td>75</td>
</tr>
<tr>
<td>4.1 Future of Work</td>
<td>75</td>
</tr>
<tr>
<td>4.2 Conclusion</td>
<td>78</td>
</tr>
<tr>
<td>References</td>
<td>81</td>
</tr>
</tbody>
</table>
List of Tables and Figures

**Figure 1:** FC installation and Go-Live processes prior to ABCD IOQ process introduction...... 17
**Figure 2:** FC installation and go live processes after ABCD IOQ process introduction.......... 18
**Figure 3:** IOQ timelines including the activity of ABCD team and vendors. ......................... 20
**Figure 4:** The focus area of the project – ABCD IOQ and Go-Live processes...................... 23
**Figure 5:** Fixing installation issues during IOQ significantly prevented severity events during operations............................................................................................................... 36
**Figure 6:** Correlation heat map showing a positive correlation between IOQ red tags and Severity Events (circled in green) and strong multicollinearity between variables (circled in blue). .......................................................... 37
**Figure 7:** Graphical model for Latent Dirichlet Allocation [11] ........................................... 40
**Figure 8:** Within Cluster Sum of Errors (WCSS) calculation. ............................................. 44
**Figure 9:** Elbow diagram showing slower reduction in WCSS after about 50 clusters......... 45
**Figure 10:** Silhouette without a clear maximum value; ideal number of topics inconclusive.... 46
**Figure 11:** A few selected topics generated by the LDA model........................................... 47
**Figure 12:** Interactive visualization showing the topics in our topic model fit to severity description data......................................................................................................................... 49
**Figure 13:** Interactive visualization highlighting *Topic 5* and the words in the topic. .......... 51
**Figure 14:** Word frequency analysis from severity description data................................. 52
**Figure 15:** End-to-end E-Prime process to simulate real operational conditions in newly built FCs for testing .................................................................................................................. 56
**Figure 16:** End-to-end CLT process to simulate real operational conditions in newly built FCs for testing. ..................................................................................................................... 58
**Figure 17:** Severity Events, IOQ red tags, and LPH by Control Cabinet for FC3, highlighting the mismatch in testing prioritization. ................................................................. 60
**Figure 18:** Recommended prioritization for IOQ tests......................................................... 61
**Figure 19:** Distribution of installation issues by failure category at FC1, about 85% IOQ complete before FC operation ....................................................................................... 63
**Figure 20:** Distribution of installation issues by failure category at FC2, about 70% IOQ complete before FC operation ....................................................................................... 64
**Figure 21:** Distribution of installation issues by failure category at FC3, about 60% IOQ complete before FC operation ....................................................................................... 65
**Figure 22:** Recommended maximum acceptable threshold for IOQ failures per CC.............. 67
Figure 23: Sensor analysis for predictive maintenance highlighting equipment for new sensor introduction.
This page has been intentionally left blank.
Note on ABCD.com Proprietary Information

In order to protect information that is proprietary to the company, the company is named ABCD throughout this thesis. Further, the data presented throughout this thesis has been modified and does not represent actual values. Data labels have been altered, converted or removed in order to protect competitive information, while still conveying the findings of this project.
Chapter 1: Introduction

In the mid-1990s, when ABCD started as a bookstore, the ABCD warehouse and its office were one and the same. A handful of employees and one lovable dog shared the city space, along with shelves and shelves of books. Over time, as their product portfolio expanded, and demand increased, one warehouse became hundreds, and they are now called ABCD fulfillment centers (FCs). These FCs store millions of items and are responsible for fulfilling customer orders. They serve as distribution centers where associates pick, pack, and ship orders quickly and efficiently [1]. These FCs play a significant role in helping ABCD meet its customer promise of two-day or even same-day delivery in some cities.

Given this promise of speedy delivery and the growing customer demand, ABCD is building new FCs every year. Launching a new fulfillment center is a complex, cross-functional effort. Multiple ABCD project leads coordinate with each other and their respective vendors to accomplish their various responsibilities towards the successful launch of the facility as a whole. The Operations Engineering team is responsible for the installation of Material Handling Equipment (MHE), which includes belted and roller conveyance, gravity and powered spirals, high-speed mergers and sorters, trailer loaders, flex conveyors, Vertical Reciprocating Conveyors (VRCs), and ABCD-specific equipment (ARSAW, SmartPac, RWC4, SLAM). Various other teams within ABCD handle storage racks, trash compactors, cardboard augers, lighting, fire sprinklers, dock doors, workstations, office furniture, and ABCD Robotics setup – including pods and drives.
This research focuses on MHE installed in an FC and aims to improve its installation quality to prevent installation-related issues from manifesting in early operations and hindering the performance of newly built FCs. MHE are mechanical equipment used for the movement, storage, control, and protection of materials, goods, and products throughout the process of manufacturing, distribution, consumption, and disposal [2]. Issues related to mechanical equipment are examined for this study. This thesis attempts to leverage analytical techniques on historical MHE data – installation issues and operational issues – and identify opportunities for refining the existing FC installation and qualification process, thereby strengthening ABCD operations.

In the process of performing the analysis, we realized that there is an opportunity for not only fixing problems during the qualification process, but also preventing issues before they arose. The recommendations advocated in this thesis center on improving future FC installation from three different lenses: reactive, preventive, and predictive. Overall, for each area, the existing processes are studied, data sources determined and analyzed, analytical frameworks developed, and solutions recommended.

The purpose of this chapter is to introduce the reader to ABCD, familiarize them with the existing Installation and Operational Qualification (IOQ) process, explain why the current IOQ process needs to be improved, and detail the approach to analyzing the process and providing recommendations. The improvements identified during the study have the potential to significantly reduce operational interruptions in future installations post-launch, saving ABCD an average of 16,000 production hours per FC in the first 90 days of operation alone.
1.1 Background

ABCD was founded in 1994 with the initial vision of becoming “Earth’s Biggest Bookstore” [3]. The company has since expanded to sell electronics, software, video games, apparel, furniture, food, toys, and jewelry. ABCD is continuously striving to be Earth's most customer-centric company.

To achieve the value proposition of becoming Earth's most customer-centric company, ABCD must maintain state-of-the-art operations. In an effort to enhance its operations and provide faster services to its customers, ABCD is expanding its fulfillment network and building many new warehouses, also called FCs. Every year, many new ABCD Robotics (AR) Sortable FCs – one type of FC – are built in North America. With every new build, ABCD is also making its FCs more sophisticated and automated. The company is investing in robotics and other technology to improve productivity, reduce variable costs, improve associate safety, and better serve customers. While ABCD is introducing new technology in newly built FCs, the focus of this study is on the fundamental infrastructure – MHE – in an FC.

Prior to 2014, building an FC involved three distinct processes – Vendor Commissioning and Validation, Operational Readiness Testing, and Operations Go-Live – as outlined in Figure 1. Third-party vendors are contracted to install all the Material Handling Equipment (MHE) in an FC. Since these vendors manufacture those pieces of equipment, they are most familiar with their installation and setup to meet ABCD’s needs and standards. Upon completion of installation, vendors inspect the mechanical, electrical, control, and software components to validate the accuracy of MHE installation. They also run high- and low-volume testing and perform a series of other qualification tests before turning the FC over to ABCD for operations.
Even with the vendor qualification process in place, ABCD experiences operational challenges during the start-up of the FC. The speed at which the new FCs are being built to meet ABCD’s promise of faster delivery introduces installation errors and degrades the installation quality of these new FCs. With an ever-growing number of FCs, ongoing operational issues due to these degraded and faulty installations are costly to ABCD. To overcome these operational challenges caused by installation-related issues and to identify these issues much earlier in the installation process, ABCD introduced its own Installation and Operational Qualification (IOQ) process in 2014, as highlighted in Figure 2 and discussed in more detail below.
Figure 2: FC installation and go live processes after ABCD IOQ process introduction.

1.1.1 ABCD Installation and Operational Qualification (IOQ) process

The Installation and Operational Qualification (IOQ) process is an approach for ABCD to validate that the quality of installation and functionality of MHE meet ABCD’s standards before releasing to Operations. It consists of a checklist of tests that are performed on all the equipment in an FC to identify installation-related issues.

ABCD’s IOQ Process includes three main phases: Installation Qualification, Operations Qualification, and Throughput Qualification. The project focuses on the Installation Qualification and Operations Qualification phases, which are designed to ensure that the installation is error-free and reduce the prospects of operational failures.

Installation Qualification (IQ): Success of the Installation Qualification process is defined as a subsystem that is installed correctly as per the design specification and subsequently qualified and approved by ABCD. Any deficient equipment or issues identified during this process are
added to the project punch list. Punch list items are also called installation issues or IOQ red tags. The IQ process is completed by area or sub-area within an FC.

**Operational Qualification (OQ):** This phase involves a series of tests designed to qualify that the mechanical equipment, controls, and software function properly in an integrated manner as described in the System Description of Operations (SDO). Most checks are performed at the device level to inspect individual components. Software is validated through monitoring of the Human Machine Interface (HMI) or System Logs, as applicable. Success of the Operation Qualification process is defined as the physical system operating as designed in accordance with the SDO. Any deficient items will be added to the project punch list.

**Throughput Qualification (TQ):** TQ testing is conducted following the successful completion of the IQ and OQ phases. The TQ testing process is successful if each component, including sorters and mergers, meets the designated throughput metrics, i.e., the rates per the as-sold contract.

Equipment installation in the FC happens on the Control Cabinet (CC) level, i.e., all the equipment in one CC are installed before the vendor starts on the next CC. IOQ tests are also performed per CC. A CC is a central electrical cabinet containing individual controls for various equipment that works together and completes a single flow within the FC. All the equipment powered by one control cabinet is installed as well as tested together.

The IOQ tests are for ABCD’s validation purposes only and are performed after the vendor IOQ. The objective of the IOQ process is to capture any installation-related issues not caught by the vendor in their inspection process. These issues are classified into two high-level buckets – immediate and punch-listed. Immediate issues can be fixed during the inspection
process and are not documented for future analysis. Punch-listed issues are the ones that cannot be fixed during the inspection and require thorough debugging and problem-solving. Vendors are notified of these punch-listed items so appropriate fixes can be applied. The timeline of the involvement of the ABCD team and vendors during the IOQ process is shown in Figure 3. Once these identified issues are fixed by the vendor and revalidated by ABCD, as shown Figure 3, the FC is handed over for ABCD operations, also called ABCD Go-Live.

![Figure 3: IOQ timelines including the activity of ABCD team and vendors.](image)

1.1.2 ABCD Go-Live Process and the Trouble Ticket System

Upon the completion of the IOQ process, application of associated fixes, and revalidation of those fixes, the FC is opened for operations, i.e., it starts fulfilling customer orders – ABCD Go-Live.
Once the FC is operational, any anomalies in equipment behavior and process failures are reported. Every time an issue is identified, an operations associate at the FC raises a trouble ticket, also called a severity event or an operational issue. A trouble ticket is the way to report problems with a production system at ABCD. The trouble ticket form prompts for all the information necessary, including Category/Type/Item, to begin work on the issue. Each trouble ticket raised in an FC triggers a work order in the Enterprise Asset Management (EAM) tool for the Reliability and Maintenance Engineering (RME) team to resolve the issue. Once the issue is resolved, the technology team in the FC updates EAM with the details of the issue by reading the trouble ticket. This end-to-end process of raising a ticket for an operational issue, tracking it to its resolution, and logging the issue is captured in a system called the Trouble Ticket system.

1.2 Problem Statement

ABCD introduced the IOQ process to validate the quality of installation of an FC and enable a smooth start-up experience. Even with the current IOQ process in place, new ABCD FCs are experiencing degraded installation quality, leading to operational issues. In the ideal scenario, each FC is expected to be operational only after the completion of IOQ tests. However, among the three FCs studied during this internship, FC1, FC2 and FC3 were respectively 85%, 70% and 60% complete with IOQ at their start-up dates. FC2 was under IOQ testing even three months after the initial start-up. Hence, ABCD is looking to review the IOQ process to determine specific areas where it could be augmented or sequenced to improve the final results.
1.3 Project Objective

As newly built FCs are experiencing operational failures related to defective installation, even with the ABCD IOQ in place, the project seeks to analyze these early failures and enhance the IOQ process. Overall, the ultimate objective of the project is to improve FC installation quality to reduce early operational issues. This is achieved through improving IOQ coverage, efficiently reprioritizing the testing schedule, increasing vendor accountability by introducing threshold metric for installation quality, and exploring predictive and preventative maintenance opportunities. The project also provides recommendations for refining the data-gathering process to improve the inputs to the model, yielding more accurate outputs.

1.4 Methodology Overview

To achieve this objective, the first emphasis is on analyzing the ABCD IOQ process and the Go-Live process, as highlighted in Figure 4. Following this, preventive and predictive opportunities are explored to further reduce early operational challenges and enhance the FC startup experience.
Figure 4: The focus area of the project – ABCD IOQ and Go-Live processes.

Analysis of these highlighted processes is performed on data from AR Sortable buildings in North America from the year 2018. Complete electronic data on the installation and IOQ processes are not available for study from years prior to 2018. Further, as AR Sortable is the most commonly built FC type, this category is selected for the study. Three AR Sortable buildings are chosen as they followed the newest installation design, are similar in size, and were installed by two different vendors. Selecting AR Sortable also allows us to access data from 2018 for analysis and model development while also having the opportunity to shadow the actual processes end-to-end. These direct experiences have included FC installation quality inspections, operations Go-Live, early severity events handling, and data collection from the FCs being built during the internship period. Even though the analysis is performed on only these selected buildings, the frameworks developed, and the recommendations proposed can be applied to other regions and FC types. Discussion of these data sources is presented in detail in Chapter 2.
This project begins with an in-depth study of the existing processes followed by an analysis of historical data. Following this, analytical frameworks are developed to identify process-improvement opportunities from three different lenses – reactive, predictive, and preventive. These analyses will be presented in detail in Chapter 3, and are previewed here.

1.4.1 Existing process validation

The IOQ process validates the quality and functionality of MHE in newly built FCs before release to Operations. The initial project emphasis is on understanding this existing process and its ability to reduce early operational issues, as described in Section 3.1.

1.4.2 Reactive lens

Once the value of the process is established, the next focus is on studying historical data to enhance IOQ to capture many of these issues during the qualification phase, instead of after Go-Live. Capturing these issues earlier enables a smoother FC start-up experience. Different techniques are introduced to improve testing coverage, as detailed in Section 3.2.1. A testing prioritization approach is also explored to maximize the impact of the IOQ tests being performed, as detailed in Section 3.2.2.

1.4.3 Preventive lens

Upon identifying ways to improve the IOQ process to capture as many installation issues as possible, different avenues are explored to improve installation quality itself, thereby limiting installation issues. A study to leverage the company’s relationship with vendors to help with installation quality – introducing a maximum threshold of acceptable failures per sub-area of the FC, is presented in Section 3.3.1.
Further, equipment replaced within the first 90 days of operations is studied to identify sensitive and frequently failing equipment, as described in Section 3.3.2. For each of these pieces of equipment, preventative maintenance schedules are studied, gaps identified, and recommendations made for a stricter schedule.

1.4.4 Predictive lens

Sensors can help detect anomalies in equipment behavior and provide opportunities for predictive maintenance. Predictive maintenance using wireless sensors on FC equipment can reduce costs, minimize the risk of catastrophic failures, and maximize system availability. With this objective in mind, all of the equipment from the FCs studied during the project is analyzed to identify equipment that could benefit from the introduction of sensors for continuous monitoring as discussed in Section 3.4.

For each of these aforementioned approaches formulated to aid ABCD in minimizing the operational challenges during FC startup, we did a study to evaluate their value proposition to ABCD Business. This financial analysis is presented in Section 3.5. Following this, Chapter 4 presents the immediate next steps and also highlights the opportunities that could be explored in the future.
Chapter 2: Data

The development of the three lenses defined in Chapter 1 requires detailed analysis of a variety of data sources, including Operations Engineering’s internal databases and the EAM platform. Once gathered, the data is subsequently cleaned and processed in order to improve its usability.

2.1 Data Sources

To identify data sources that would assist in this research, it is important to understand the entire flow – from FC installation, to IOQ, to operational Go-Live – and learn how data are captured at each point and how these data could be useful. After extensive research and real-time shadowing, the following data sources are determined to be necessary and sufficient for the purposes of this study. They have been collected from all three FCs studied and are used in all the approaches implemented throughout the internship.

The Asset dataset and IOQ red tags dataset are stored in the North America Operations Engineering teams’ databases. The rest of the datasets are stored in Enterprise Asset Management (EAM) databases.

2.1.1 Asset Data

Third-party vendors are contracted to install the MHE in the FCs based on the design developed by internal ABCD teams. CC equipment data are systematically logged by the vendor during this installation phase. This process captures information on the equipment, also called assets, which is of interest for this research. “Assets” and “equipment” are used interchangeably within ABCD and in this thesis.
This data is sourced by the vendors during the installation process and validated by
ABCD during the asset verification process to ensure that all required assets have been installed. This dataset contains detailed information including asset description, asset id (a unique alpha-numeric identifier), equipment id (a unique numeric identifier), asset model, control cabinet where the asset is installed, some additional metadata, and the asset’s EAM profile. This profile, introduced by the EAM team, is a grouping of many similar assets.

2.1.2 Installation and Operational Qualification (IOQ) Red Tags Data

The IOQ process follows after the completion of vendor installation. Once the list of equipment for the study was identified, IOQ was shadowed for two weeks to get a better sense of how the qualification tests are run and how data are captured in case of test failures. As the equipment are grouped per CC, IOQ tests are also performed per CC. A series of predefined tests are run and, for each failed test in a piece of equipment, a separate red tag is created, and the failure recorded. Detailed information on the nature of failure, problem identified, and cause of failure and any pictures, if applicable, is collected and stored as Red Tag data in the NA Operations Engineering team’s database. Shadowing the IOQ process has also highlighted improvement opportunities for data gathering. During IOQ, each piece of equipment is tested in isolation and not as a part of an integrated system. Hence, only unit failures are captured as opposed to both unit and system failures. Further, since logging each test failure takes about five minutes, the associates performing the IOQ often choose to fix many identified issues rather than logging the failure information.

The IOQ red tags dataset contains all relevant information for each installation issue identified during the ABCD IOQ process. This data is sourced by an ABCD associate validating the equipment installation quality in an ABCD FC. During IOQ, the ABCD associate runs a series
of tests for each piece of equipment. If any of these tests fail or the equipment does not run as expected, this is a red flag that the equipment installation might not be correct. In case of these installation issues, the associate hangs a QR code–embedded red tag in the equipment and records the test findings. Since a physical tag is placed on the equipment to highlight potential installation related issues, this dataset is referred to as IOQ red tags or punch-listed items. For each punch-listed item, the associate is expected to capture a unique punch list identifier (QR code of the red tag), a problem-failure-cause category, an asset id, and some additional metadata that could be used by the vendors to apply relevant fixes.

2.1.3 Trouble Ticket Data

The next piece of information that is critical for the study is the data on operational failures experienced during the startup of the FC. Once IOQ is complete and the vendor applies the relevant fixes to identified issues, the FC is handed over to Operations. While in an ideal situation the FC is expected to have a smooth start-up and flawless operation, all newly built FCs experience many operational challenges. This dataset, sourced by the operation associate working on the FC, contains information on all the operational hurdles encountered. The Go-Live process was shadowed for a week to understand what qualifies as an operational challenge, how these challenges arise and are flagged, and how the process of maintenance is triggered and prioritized.

All ABCD fulfillment centers are equipped with Andon lights. Andon is a manufacturing term referring to a system to notify management, maintenance, and other workers of a quality or process problem. The alert can be activated manually by a worker using a pull cord or button, or it may be activated automatically by the production equipment itself. At ABCD, operational challenges are flagged and visually communicated with the help of these Andon lights. If an
associate working on a station notices any problems, they use these lights to call for attention. Alternatively, some systems also automatically trigger these lights to indicate a problem.

For each identified problem, the associate working on the station creates an online ticket – called a trouble ticket – that triggers a work order for the Reliability, Maintenance, and Engineering (RME) team to come and fix the issue. FC start-up, or Go-Live, is a major, well-planned event and all the teams are on site to provide the immediate support needed. These teams include third-party vendors (installing MHE in the FCs), Operations Engineering (overseeing the building and installation of the FC), Operations (operating the FC once live), and RME (fixing any operational issues). When a ticket is raised, it calls attention from all the teams on site. Any update on the problem or resolution of the issue is communicated on the online ticket. Once the issue is resolved, the technology team in the FC updates the EAM database with the details of the issue from the trouble ticket.

The trouble ticket dataset contains a short description of the issue as observed by the associate working on the FC floor, the asset id of the equipment causing the failure, a unique trouble ticket number associated with the issue, the severity level of the issue (how impactful the issue is to business), and some additional metadata. While the associate is expected to log the problem-failure-cause category information for each identified issue, this information is not always captured due to new, inexperienced associates in a recently built FC and their lack of knowledge about operations. Issues with severity level one or two also include associated loss in production hours. This dataset is also referred to as Severity Events or Operational Issues. As the thesis focuses on installation-related operational issues, only failures experienced during the first 90 days of operation are used, under the assumption that these issues would capture most of the flaws in installation.
2.1.4 Equipment Criticality Data

When multiple issues are raised, the RME team prioritizes the fix for one issue over the other. In this thesis research, we have sought to understand the basis for this prioritization, and have found the following factors. In case of failures, some equipment has a higher impact on business in terms of lost production hours and stress introduced in the FC network due to reallocation of shipments. Certain equipment is, therefore, more critical to business compared to other equipment. This information on equipment criticality could help enhance this process around sensitive equipment and prioritize their testing, and so is included in the research for this study.

An operations team in each of the FCs is responsible for assigning a criticality value to all equipment from 1 (most critical) to 3 (comparatively less critical) in their particular FC. The criticality numbers are reversed for all calculations to align them with the human thought process of a larger number meaning more critical. As the local operations team sources this data, this information is not standardized across the FC network, e.g., a piece of equipment considered Criticality 1 at a particular FC could be considered Criticality 3 at another FC. This dataset contains criticality information for most of the equipment available in the FC. Even though operation teams’ experiences factor heavily into this data, it is critical to incorporate this data in the study while developing the models and developing informed recommendations.

2.1.5 Replacement Parts Data

When an issue is identified during the operation of an ABCD Fulfillment Center and a trouble ticket is raised, the Reliability, Maintenance, and Engineering (RME) team is notified to analyze and fix the issue. Any fix requiring replacement of parts is logged for inventory
management. This information, stored as Replacement data in EAM, is thus included for further study during this research.

This dataset contains detailed information of all the parts that are replaced in the FC. The thesis focuses on the parts replaced in the first ninety days of operation. Parts that are replaced frequently in early operations can provide insight about how failure-prone the equipment is and allow installation effort to be allocated appropriately.

2.1.6 Root Cause Analysis (RCA) Data

The RME team revisits a select few severe operational challenges that cause a significant loss in production hours to perform a deeper analysis and identify the root causes. The detailed analysis of these events and associated findings, also called Root Cause Analysis (RCA) data, is included in this research to get a better understanding of the failures happening in an FC and propose more informed recommendations for process improvement.

A member from the RME team performs an analysis similar to Five-Whys on these events to get to the depth of the issue. For each of the chosen events, a study is performed to answer the questions below and the findings are logged.

- What happened?
- How did it happen?
- What was the fix?
- What were the lessons learned?
- What are the action items to prevent it from happening again?
As RCA is done very infrequently, the dataset is extremely sparse. However, this dataset has been analyzed to understand ABCD’s initiative on improving their processes and to develop recommendations for further process improvement.

2.1.7 Wireless Sensor Data

To identify early signals of equipment failures and thereby potentially avert these highly severe operational issues from happening, the RME team has introduced sensors in various equipment within the FC. Preventative maintenance schedules are introduced for those equipment based on the data reported by the sensors. The equipment on which the sensors are introduced is based on the RME team’s experience from fixing severity issues reported in an FC. Sensor data is included in this thesis research for further analysis to make informed decisions on identifying additional equipment that would benefit from the introduction of sensors. This effort would significantly help in the reduction operational issues.

These sensors monitor the physical and environmental conditions, including the temperature and vibration of the equipment. Based on the information captured by these sensors, if an anomaly is detected in the equipment’s behavior, maintenance is performed on the equipment. There are two different datasets that contain information on wireless sensors. One dataset contains all the equipment in each FC on which sensors are installed. The second dataset contains information on all the instances in which failures were averted by doing predictive maintenance using the data captured by the sensors. Both of these datasets have been studied to understand the impact of wireless sensors in successfully identifying early signals of potential failures and averting them from ever happening. Additional equipment has been identified that would benefit from the introduction of these sensors.
2.2 Data Cleaning

The information required for this research is derived from a combination of these datasets. As many of the available datasets are incomplete, effort is invested in cleaning the data and tying the different datasets together to form a better picture. Once all the data for the study is collected, the next phase is ensuring the quality of the data gathered.

Incorrect or inconsistent data can lead to false conclusions and misdirected investments. Decisions informed by the use of data analytics methods are only as good as the data on which they are based. Data quality has been shown to have significant impacts on both the tangible and intangible aspects of business if not properly managed [4]. Poor data quality may lead to poor decisions with irreversible consequences for a company. Hence, it is important to clean the available datasets, remove irrelevant data, and source critical missing information before using them for further analysis and model development.

All the available datasets are exported into a Python script where they are cleaned and merged together. The list below highlights some of the data cleaning and manipulation performed on the available datasets to provide as complete and accurate data as possible for any future analysis.

- For any asset in the Asset dataset missing an EAM Profile, the data is sourced from the Enterprise Asset Management (EAM) database.
- For equipment across different datasets with a miscategorized EAM Profile, the misclassified profile values are corrected based on information from the Enterprise Asset Management (EAM) database.
• Equipment id is a numerical identifier that uniquely identifies each piece of equipment in the FC. This information is missing in the IOQ red tagged data and so is sourced from the Asset data. Equipment criticality and EAM Profile information are sourced from the equipment criticality dataset and the Asset dataset, respectively.

• Red tags data missing an equipment id are ignored from analysis as there is no way to associate the issue with any equipment.

• As the research focuses on Material Handling Equipment (MHE) in an FC, operational issues raised on equipment outside of MHE are removed from analysis.

• In the trouble ticket dataset, EAM Profile, Control Cabinet, and Equipment Criticality are sourced from the Enterprise Asset Management (EAM) database, the Asset dataset, and the Equipment Criticality dataset, respectively.

• A Python program was written to infer missing problem codes in the trouble ticket dataset from the free-text descriptions of the severity events. If any word in the description of the severity message maps to a problem code or failure code from the IOQ Red Tag data, the severity event is categorized with the same problem code as the IOQ red tag.
Chapter 3: Research Methodologies and Findings

Data gathered from various sources, as explained in Chapter 2, are studied thoroughly to gain better insight into the installation problems and operational issues and to develop concrete recommendations for future process improvement. This chapter describes in detail the analyses performed and various approaches explored during the research phase of this project.

3.1 Analysis of existing Installation and Operational Qualification process

Once the data is cleaned, the initial emphasis is on understanding the existing IOQ process and the value it adds in reducing the early operational issues. Preliminary analysis is done on IOQ red tags (installation issues) and Severity Events (FC startup issues, early operational failures) collected from the trouble ticket database to establish if there is a business case for IOQ checks. The data from the three FCs studied support that there is indeed a reduction in start-up severity incidents if IOQ is performed and installation issues identified prior to Go-Live, as shown in Figure 5.
Figure 5: Fixing installation issues during IOQ significantly prevented severity events during operations.

For FC1, the IOQ process was 85% complete prior to Go-Live: 4,800 hours were invested to identify any issues in installation and apply the fixes. For FC2 and FC3, IOQ process was 70% and 60% complete and 4,620 and 4,480 hours were invested, respectively. At FC2 and FC3 IOQ timelines were constantly shifted due to delays in installation, ultimately reducing the hours spent on the tests. As a result, many installation issues manifested as operational issues.

Figure 5 shows the increase in severity events as effort spent on IOQ decreases and installation issues fail to be captured and fixed prior to Go-Live.

3.2 Improving the Installation and Operational Qualification process

While the IOQ process is seen to help reduce early operational failures, it does not completely eliminate them. The next step is to understand why.
The correlation between installation and operational issues in each control cabinet is studied. Both installation and operational issues are broken down into four high-level problem categories – mechanical, electrical, control, and software– before the correlation calculation. The results indicate a positive correlation between IOQ red tags and severity events in each of the three FCs, i.e., control cabinets in which we discover more issues during IOQ also have more severity events during the early operation phase. Data also shows a strong multicollinearity between variables. Multicollinearity is a state of very high inter-correlations or inter-associations among the independent variables. Figure 6 presents the correlation results from one of the FCs analyzed.

**Figure 6**: Correlation heat map showing a positive correlation between IOQ red tags and Severity Events (circled in green) and strong multicollinearity between variables (circled in blue).
Based on this finding, three testable hypotheses are formulated to explain the observed correlation.

**Hypothesis 1:** Complex systems have many possible points of failures – some caught during IOQ and others not until operations. More complex systems are therefore more likely to have failures both during IOQ and early operations.

**Hypothesis 2:** Even if issues are caught during installation and fixes are applied, these fixes are insufficient, and issues resurface as operational issues.

**Hypothesis 3:** The testing conditions do not adequately simulate real operational conditions.

Free-text descriptions from severity events data and root cause analysis data could help explore these hypotheses further and develop recommendations for IOQ process improvement. To discover the most problematic equipment and identify high failure-prone areas within FCs, two techniques were explored: topic modeling and word frequency analysis.

**Topic Modeling**

Topic modeling is one of the most powerful techniques for text mining, latent data discovery and finding relationships among data and text documents [5]. This unsupervised learning model helps discover hidden topical patterns across the collection of textual documents. Unlike manual tagging, which is effort intensive and requires expertise in the document subject matter, topic modeling is an automated process. Relying on the assumption that each document in a collection refers to a small number of topics, it extracts bags of words attributable to these topics [6]. Researchers have published many articles in the field of topic modeling and have applied it in various applications – for instance to infer the research interest of a faculty member.
given the title of only a few research papers written by him [7], to deduce user persona by deriving the topics in each user’s self-produced tweets and retweets [8], to extract topics in source code and perform visualization of software similarity [9], etc. Overall, topic modeling is a powerful, smart technique that is widely applied in natural language processing for inferring patterns and hidden structures in gigantic unstructured data. There are various methods for topic modelling, among which Latent Dirichlet Allocation (LDA) is one of the most popular.

**Latent Dirichlet Allocation (LDA)**

LDA is a generative statistical model that allows sets of observations to be explained by unobserved groups that describe why some parts of the data are similar. Approaches that explicitly or implicitly model the distribution of inputs as well as outputs are known as generative models, because by sampling from them it is possible to generate synthetic data points in the input space [10]. This means we can create documents with a mixture of topics and a mixture of words based on those topics. LDA suggests that each document is a mixture of a small number of topics and that each word's presence is attributable to one of the document's topics. Documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over all the words. **Figure 7** shows the graphical representation of the LDA generative process for a corpus with $D$ documents, each with length $N$. 
Figure 7: Graphical model for Latent Dirichlet Allocation [11]

*D*: Number of documents in the corpus, *K*: Number of topics, *N*: Number of words in each document

Given the amount of textual information captured when an operational challenge is experienced in an FC, topic modeling is leveraged for our study to infer the most problematic equipment and to identify failure-prone areas within FCs. Each severity description is analogous to a text document. The collection of these descriptions is thus similar to a collection of documents that are usually analyzed for topic modeling. Among several existing algorithms one can use to perform topic modeling, we used Latent Dirichlet Allocation (LDA) for the project. We did the execution in Python using the Gensim implementation. LDA creates topics as groups of words with respective frequencies of appearance and classifies each severity description as a member of multiple topics in specific proportions. Each topic is a collection of dominant keywords with their own specific frequencies. The words in each topic can potentially be used to identify the relevant equipment and/or failure mode for that topic. The relevance of output topics...
can be attributed to factors including the input data, the quality of text processing, variety of topics the severity description talks about, number of topics used for modeling, and the algorithm tuning parameters.

In the section below, we illustrate the LDA method in the context of analyzing textual data extracted from the issue-tracking system at ABCD – the trouble ticket system. Once the data is extracted, the next step is the text-cleaning process.

**Preprocessing the data**

The purpose of text cleaning is to simplify the text data, eliminating language-dependent factors as much as possible. Articles are written in natural language for humans to understand. But in text mining, those data are not always easy for computers to process [12]. Below are the five steps in text-cleaning applied for this study.

*Tokenization:* The words or tokens should be split out of the input text in order to eventually count them. Severity description data is cleaned by removing all punctuation and unnecessary characters and each description is then partitioned into a list of words called tokens.

*Removing stop words:* Some words appearing in texts are not useful in topic analysis; such words are called stop words. It is common in natural language processing and information retrieval systems to filter out stop words before building a model. Stop words such as "the", "if", "and" etc. frequently occurring in human communication and not relevant to the desired analysis are removed.

*Bigram and Trigram model:* Bigrams are two consecutive words frequently occurring together in the document, and trigrams are three words frequently occurring. A bigram predicts the next
word based on the one before, and a trigram predicts based on two words. To enhance the output of the model, bigrams and trigrams are created. Gensim’s Phrases model is used to build and implement the bigrams and the trigrams. The two important arguments to Phrases are minimum count and threshold. The higher the values of these parameters, the harder it is for words to be combined.

**Lemmatizing Words:** Lemmatizing is the process of reducing words to their original root. English language has multiple forms and tenses. Hence, it is a common practice to lemmatize the words before using them for building the model.

For the study, each word is lemmatized to its root form to remove inflectional endings and to return the base or dictionary form of the word. For example: organize, organized, organizes, and organizing all belong to the same root – organize. Thus, if a sentence contains any of these words, on lemmatizing, it will be resolved it to the same root form – organize. Lemmatizing words to their base form helps reduce redundancy, decreases the vocabulary sizes, and enhances model output.

**Common word removal:** Most severity descriptions contain some common words such as the name of the FC where the issue happened, the word “issues”, etc. These common words are often skipped because they muddle topic summaries and make interpretation more difficult.

**Building the model**

Once the severity descriptions were preprocessed, data was prepared via bigrams and trigrams, and lemmatization was done, the next step was to index descriptions for use as documents within an LDA implementation.
A simple vector representation, Bag of Words (BOW), was generated using all the cleaned and tokenized words from each description. Next, the list of descriptions was converted into lists of vectors, all with length equal to the vocabulary in BOW. These vectors and the BOW served as the inputs for training the LDA model.

Algorithm tuning parameters α (alpha) and η (eta) for the LDA model affect sparsity of the document-topic (theta) and topic-word (lambda) distributions. Default values of 1.0/number of topics are used for α and η in generating our model.

“Chunksize” and “passes” are two other input parameters for the model. Chunksize controls how many severity descriptions are processed at a time in the training algorithm. Increasing the value of chunksize will speed up training, at least as long as the chunk of severity descriptions easily fit into memory. The passes parameter controls how often we train the model on the entire corpus. Chunksize is set to 100 and passes 10 for this study.

The LDA model also requires the number of topics as an input parameter. There are more than 30 methods to find the appropriate numbers of clusters/topics in a dataset, including two conventional approaches – the elbow method and the average silhouette method. The elbow method looks at the within cluster sum of errors (WCSS) for different cluster sizes. For each cluster, the distance between each point in the cluster and the centroid is calculated. WCSS is thus the summation of these distances for all the clusters, as shown in the example in Figure 8.
Figure 8: Within Cluster Sum of Errors (WCSS) calculation.

WCSS calculated for different number of clusters is then plotted, which generates an elbow shaped graph. As we increase the number of topics (also called clusters), WCSS always decreases. However, the rate of drop in WCSS starts to decrease as we increase the number of clusters. The number of clusters from where the rate of drop in WCSS does not drop substantially, or where the rate of drop is much less, is the best number of topics that should be created. The elbow diagram is drawn for our severity description data to infer the best number of topics as shown in Figure 9. The elbow cannot always be unambiguously identified [13], which is also true in our case.
The silhouette method looks at the average distance between points in a cluster, which is known as cohesion. This number is compared to the average distance between groups, which is a measure of separation. Overall, in this method we look for the number of clusters where the difference between separation and cohesion is maximized [14]. As with the elbow method, we plot these points for the severity description data, as shown in Figure 10. With this approach we hope to get a clear maximum value that can be easily identified programmatically, which was not true in our case. The number of clusters corresponding to the maximum silhouette score would be the ideal number of topics.

**Figure 9:** Elbow diagram showing slower reduction in WCSS after about 50 clusters.
Figure 10: Silhouette without a clear maximum value; ideal number of topics inconclusive.

Next, in order to deduce the best number of topics for the LDA model, a Hierarchical Dirichlet Process (HDP) model is explored. The HDP model is a natural nonparametric generalization of Latent Dirichlet Allocation, where the number of topics can be unbounded and learned from data[15]. The number of output topics generated by the HDP model, which is in the range of 25 to 150 for the three FCs studied, is then used as an input parameter to the LDA model. Both HDP and LDA models generate similar topic per word distribution output when using the number of topics identified by the HDP for LDA model. However, there are built-in Python tools developed for visualizing LDA models which do not support the HDP models.
Hence, to visualize the output of the model and better understand the topic distribution, an LDA model was developed by using the number of topic output from HDP model.

**Interpreting the results**

Below is a sample topic generated by the model. As we can see, the topic is a combination of keywords – rebin, belt, collector, stop, empty, stacker, frozen, high, middle, replace – and each keyword contributes a certain weight to the topic as indicated by the number prior to each word.

**Topic 5** Words: $0.519 \times \text{"rebin"} + 0.177 \times \text{"belt"} + 0.109 \times \text{"collector"} + 0.036 \times \text{"stop"} + 0.024 \times \text{"empty"} + 0.020 \times \text{"stacker"} + 0.017 \times \text{"frozen"} + 0.013 \times \text{"high"} + 0.007 \times \text{"middle"} + 0.005 \times \text{"replace"}

**Figure 11** shows six topics generated by the model. The words shown are the top 10 ranking words for each of these topics, the most heavily allocated words in the topic.

<table>
<thead>
<tr>
<th>Topic 1</th>
<th>Topic 2</th>
<th>Topic 4</th>
<th>Topic 5</th>
<th>Topic 18</th>
<th>Topic 24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station</td>
<td>SmartPac</td>
<td>Flat</td>
<td>Rebin</td>
<td>Divert</td>
<td>Need</td>
</tr>
<tr>
<td>ARSAW</td>
<td>Require</td>
<td>Sorter</td>
<td>Belt</td>
<td>Main</td>
<td>Printer</td>
</tr>
<tr>
<td>Arm</td>
<td>Sealing</td>
<td>Fault</td>
<td>Collector</td>
<td>Sorter</td>
<td>Eye</td>
</tr>
<tr>
<td>Destacker</td>
<td>Replacement</td>
<td>Motor</td>
<td>Stop</td>
<td>Unable</td>
<td>Photo</td>
</tr>
<tr>
<td>Function</td>
<td>Flex</td>
<td>Tran</td>
<td>Empty</td>
<td>Routing</td>
<td>Label</td>
</tr>
<tr>
<td>RME</td>
<td>Part</td>
<td>Repair</td>
<td>Stacker</td>
<td>Turn</td>
<td>Prod</td>
</tr>
<tr>
<td>Description</td>
<td>Switch</td>
<td>Show</td>
<td>Frozen</td>
<td>Inoperable</td>
<td>Help</td>
</tr>
<tr>
<td>Calibration</td>
<td>Cutting</td>
<td>Tracking</td>
<td>High</td>
<td>Come</td>
<td>Detect</td>
</tr>
<tr>
<td>Verify</td>
<td>Wait</td>
<td>Functioning</td>
<td>Middle</td>
<td>Cognex</td>
<td>Faa</td>
</tr>
<tr>
<td>Gift</td>
<td>Jack</td>
<td>Inactive</td>
<td>Replace</td>
<td>Sor</td>
<td>bad</td>
</tr>
</tbody>
</table>

**Figure 11**: A few selected topics generated by the LDA model.
The words in the same topic tend to be similar; i.e., the terms are associated with each other. For example, topic 1 indicates an issue with the *ABCD* Fulfillment Engine (AFE) system, topic 2 is about sealing issues in the SmartPac, topic 4 indicates a motor issue in the flat sorter that needs repair, and topic 18 indicates issue routing/diverting the packages at the Main Routing Sorter. The topics generated provide a way to identify systems experiencing frequent operational challenges and areas susceptible to failures in the FC.

Each topic is assigned a number by the LDA model. However, the order in which the topic number is assigned is arbitrary and has no meaning. If the LDA is re-run, you would get similar topics but in different orders.

**Visualizing the results**

One of the most commonly used ways to visualize the information contained in a topic model is by using a visualization tool called LDAVis. The visualization tool requires five inputs:

- A $K \times W$ matrix that contains the estimated probability mass function over the $W$ terms in the vocabulary for each of the $K$ topics in the model.
- A $D \times K$ matrix that contains the estimated probability mass function over the $K$ topics in the model for each of the $D$ descriptions in the corpus. The corpus is the entire set of severity descriptions.
- The number of tokens, $n_d$, observed in severity description $d$, where $d = 1 \ldots D$.
• A vector of length $W$ containing the terms in the vocabulary.

• The frequency of term $w$ across the entire corpus.

Python has a library available for interactive topic model visualization called pyLDAVis. This library was designed to help users interpret the topics in a topic model that has been fit to a corpus of text data. This package extracts information from a fitted LDA topic model to inform an interactive web-based visualization[16]. Figure 12 shows the interactive visualization for our topic model, using the available Python library.

Figure 12: Interactive visualization showing the topics in our topic model fit to severity description data.
A good topic model will have fairly big, non-overlapping bubbles scattered throughout the chart instead of being clustered in one quadrant. A model with too many topics will typically have many overlapping small bubbles clustered in one region of the chart.

**The left panel**, labeled Intertopic Distance Map, represents topics as circles. Similar topics appear closer together and the dissimilar topics farther apart. The size of a topics circle in the plot corresponds to the relative frequency of the topic in the corpus: the larger the bubble, the more prevalent is that topic. When no circle is selected, the default topic – topic 0 – is selected as shown in the top left corner. An individual topic may be selected for closer scrutiny by clicking on its circle or entering its number in the “selected topic” box in the upper-left.

The two axis labels PC1 and PC2 are first and second principal components. PC1 represents the maximum variance direction in the data as a linear combination of the underlying variables. The second principal component (PC2) is oriented such that it reflects the second largest source of variation in the data, while being orthogonal to PC1.

**The right panel** includes the bar chart of the top 30 terms. When no topic is selected in the plot on the left, the bar chart shows the top-30 most “salient” terms in the corpus. A term’s saliency is a measure of both how frequent the term is in the corpus and how “distinctive” it is in distinguishing between different topics. If one of the bubbles on the left-hand plot is selected by hovering the mouse over it, the words and bars on the right-hand side will update. These words are the salient keywords that form the selected topic. An example is shown in the Figure 13.
As mentioned earlier, a good topic model will have fairly big, non-overlapping bubbles scattered throughout the chart instead of being clustered in one quadrant, as we can see in our graph. The model generated using the severity descriptions at each of the FCs studied produces similar graphs with extremely overlapping bubbles. While the topics outputted by the model give us an indication about the problematic issues and areas, given the limited number of data points, each with extremely short descriptions, the model output does not define clear, well-differentiated topics. While the current model output is overlapping, future outcomes could be better with more thorough descriptions, and the resulting model could be beneficial for future iterations.

**Figure 13:** Interactive visualization highlighting *Topic 5* and the words in the topic.
To supplement the findings from topic modeling above in inferring patterns of common equipment failures from severity descriptions, the next approach explored is word frequency analysis using the same BOW approach explained above. Free-text input is was cleaned, non-alphanumeric characters and stop words are removed, and words are lemmatized before creating a BOW. The frequency of each word is then computed using count vectorizer and graphed. 

**Figure 14** shows the output from one of the FCs studied.

![Figure 14: Word frequency analysis from severity description data.](image)

The following keywords (presented in order of failure frequency) appeared most frequently in the free-text descriptions from all the FCs studied: AFE, sorter, jam, pakivaa, slam, induct, smart pack, rebin, divert, belt, conveyor, VRC, power, tray/tote, merge, MRS, and routing. This list includes two distinct categories of failures: equipment failures and process/functionality failures. The top 10 to 15 most frequently repeating words, representing
either a problematic equipment or a frequently failing mode, are prioritized for further analysis in an effort to improve the IOQ process currently in place, as discussed below.

3.2.1 Installation and Operational Qualification testing coverage

Once the frequently failing equipment and/or processes have been identified, the focus is in identifying ways to incorporate this information in enhancing the existing IOQ process. The IOQ checklist is thoroughly scanned to ensure that it provides coverage for the top 10-15 frequently failing equipment identified above. Concentrating the efforts on these few equipment helps uncover the fact that the current IOQ process does not have tests for four of them – Slam, SmartPac, Rebin, and VRC.

New tests, based on the failure descriptions, are proposed for each of these pieces of equipment missing IOQ coverage. These tests have already been included in the IOQ checklist for future FCs. Below are the tests that are proposed.

**SmartPac**

<table>
<thead>
<tr>
<th>Safety</th>
<th>Functionality</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Ensure there are no loose parts. All necessary components are completely and correctly installed.</td>
<td>• Check E-Stop operation. Make sure the tension on the e-stop core is within the approved weight range.</td>
<td>• Ensure Smart Pac ASIN (ABCD Standard Identification Numbers) response time meets requirement.</td>
</tr>
<tr>
<td>• Check loose wiring terminations - check for loose or missing wires.</td>
<td>• Verify the functionality of photo eye on conveyor.</td>
<td>• Perform high volume and low volume tests.</td>
</tr>
<tr>
<td>• Ensure necessary conveyor guarding is installed (4&quot;, 6&quot;).</td>
<td>• Ensure Smart Pac is sealing properly.</td>
<td></td>
</tr>
<tr>
<td>• Ensure that the corners are smoothed out – for safety reasons and to ensure packages do not get stuck.</td>
<td>• Ensure Smart Pac prints label correctly.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ensure Smart Pac sealer jaw/blade cuts packages properly.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ensure labels do not get stuck or stick to each other.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ensure Smart Pac film roll sensors work properly.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ensure that the Andon lights to the machine are working as expected.</td>
<td></td>
</tr>
</tbody>
</table>
**Settings**

- Ensure correct settings for Smart Pac cameras and printers.

**SLAM (Scan/Label/Apply/Manifest)**

**Safety**

- Ensure there are no loose parts. All necessary components are wired correctly.
- Check E-Stop operation. Make sure the tension on the e-stop core is within the approved weight range.

**Functionality**

- Ensure SLAM line slams high volume packages without falling out of S.
- Ensure SLAM reflectors, scanners, and printers are functioning properly.
- Ensure SLAM scales are functioning properly.
- Ensure there are no jams on Slam. Verify no box sizes get stuck on Slam, i.e., test with all available box sizes. Ensure we test all alignments including boxes arriving sideways.
- Verify fake shipments have been created for testing each Auto Slam.

**Performance**

- Verify upstream speeds of conveyance going into the Alignment Bed.
- Ping every device in each Auto SLAM line to check connectivity in to the network.
- Run throughput test which should achieve a minimum of 2K packages in 8 hours shift.

**Setting**

- Check Proper Alignment for SLAM Scales: Check that nothing is contacting scale which would affect measurements.
- Ensure that the conveyor alignment is correct.

**Rebin**

**Functionality**

- Ensure Rebin rollers are functioning properly.
- Ensure flashing white-light indicating order complete and red-light indicating order not complete are working properly on both sides of the wall.
- Ensure Rebin scanners read the totes.
- Ensure Rebin wall motors (motor driven roller) are working properly.
- Ensure trays and totes are diverting off the Rebin lane.

**VRC**

As the severity descriptions for VRC failures do not contain information on the failure, new tests capturing all the failures could not be formulated properly. It is recommended that correct descriptions be included in the severity tickets. However, the two tests below have been included in the checklist to test some functionality of the equipment.
<table>
<thead>
<tr>
<th>Functionality</th>
<th>• Ensure the VRC doors close properly.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>• Perform high volume and low volume pod transfer tests to ensure VRCs don’t go down.</td>
</tr>
</tbody>
</table>

For the remaining equipment (AFE, sorter, belt, conveyor, power, tray/tote, MRS) and functional failure modes (jam, induct, merge, divert, routing), scanning the available tests reveals that the unit-test based IOQ process do not adequately capture the real FC operational scenarios, i.e., it lacks integration, functional, stress, and load testing. As IOQ tests are not conducted in a live operational environment or with any real order items, the FC is not subjected to a realistic load. Current testing does not replicate the actual operational scenarios where separate components work together as an integrated system. This inadequacy in IOQ was also noted while shadowing IOQ and Go-Live process at one of the AR Sortable buildings from 2019. For example, during IOQ, a few bags of water bottles or a handful of totes and trays were used at a particular orientation to test different equipment, while during the actual Go-Live, packages of different shapes and sizes passed through the same equipment in different orientation, speed, and frequencies causing a jam.

To simulate real operational conditions in newly built FCs for testing their installation quality, two end-to-end testing approaches, i.e., soft Go-Live programs – E-Prime and Closed Loop Testing – are proposed. Two different papers have been written and submitted to ABCD outlining the scope of each program, desired state after its implementation, implementation details, cost and value proposition, dependencies, risks and blockers, success metrics, stakeholders involved, and recommendations for program sustainability. The sections below present a high-level overview of each of these testing approaches.
E-Prime Program: An Exclusive Prime Experience for *ABCD* Employees

E-Prime is a one-week end-to-end testing program proposed to simulate real operational conditions in newly built FCs. The testing timeline and how it fits into the current installation and Go-Live schedule is shown in **Figure 15**.

![Diagram](image)

**Figure 15**: End-to-end E-Prime process to simulate real operational conditions in newly built FCs for testing.

The estimated time for the inventory build-up phase is one week, with two 10-hour shifts per day. Continuous inventory refill will occur during the E-Prime week to ensure demand is fulfilled. As actual orders will be received and fulfilled, Active ASIN’s – *ABCD* Standard Identification Numbers – corresponding to actual products will be utilized to build the inventory.
Three weeks prior to Go-Live, the FC will open for a week-long virtual Go-Live allowing pre-selected employees to purchase available inventory items at a discount. These discounts are provided to employees to create desired volume for testing. During the week of E-Prime, the FC under test will receive real orders all day but fulfill the demand in only 10 hours per day, mimicking an ABCD work shift. The remaining 14 hours will be dedicated to fixing the issues identified in the aggressive 10-hour testing period. Taking orders throughout the day but fulfilling it in the 10-hour window will also create the volume and stress that a live FC experiences for the 10-hour testing period. Further, E-Prime week will be treated as a revenue-generating event, and the Operations team on-site will be expected to work through the failures and errors and make every attempt at two-day delivery. However, the program will be limited to ABCD employees to protect ABCD branding in case the delivery estimate accuracy is not met. This one-week program will allow ABCD to test the entire system, identify installation issues that could manifest in operations, resolve these issues, and retest the system to ensure that identified issues are fully fixed prior to the actual Go-Live.

**Closed Loop Testing**

Closed Loop Testing (CLT) is a condensed version of the E-Prime program with slight modifications. The motivation behind CLT is to create a real operational scenario for testing without actual shipment of products. As no products are being shipped, the inventory in CLT will purely consist of expired and obsolete ASINs (ABCD Standard Identification Numbers) gathered from live FCs across North America. Pre-programmed customer requests using hypothetical package destinations for the available inventory items will be generated to test outbound routing features. Each round of testing ends once all the programmed packages are at the end of outbound dock ready to be shipped. All of these packages will be consolidated and run
through the inbound and stored as inventory for the next round of testing. Further, recirculating the packages from outbound back to inbound after each day of testing will allow for re-testing of inbound systems. As the packages never leave the system and are circulated between inbound and outbound, this approach is called closed loop testing. Figure 16 shows the timelines for CLT which is similar to E-Prime, but condensed.

**Figure 16:** End-to-end CLT process to simulate real operational conditions in newly built FCs for testing.

This three-day program will allow ABCD to test the entire system, resolve issues, and retest the system to ensure that identified issues are fixed without the potential risk of delayed package shipments. As this program has minimum risk, we recommend implemented in the years prior to E-Prime to test its feasibility. CLT could be used to set up infrastructure required to
support this new testing approach, and when ready the program can be enhanced to introduce live products, take real employee order requests, and ship the actual products, i.e., evolve into E-Prime.

The next significant contributor to early operational issues is the tight timeline leading to incomplete IOQ. As mentioned earlier, in the ideal scenario each FC is expected to go live only after the completion of IOQ tests. However, the timelines and speed of installation to support the growing need for capacity do not always allow full IOQ completion prior to Go-Live. For example, FC1 was about 85% complete, FC2 was about 70% complete, and FC3 was about 60% complete with IOQ at their respective first receive dates. FC2 was under IOQ testing even after three months of Go-Live. As it would not be possible to extend the timelines of IOQ prior to Go-Live, we instead look for ways to maximize the impact of IOQ tests performed within the allocated time.

### 3.2.2 Installation and Operational Qualification testing prioritization

An in-depth study is performed to understand the distribution of severity issues per Control Cabinet (CC) and the criticality of equipment in the CC. IOQ red tags data is also explored in relation to severity issues to understand if the current IOQ process timeline is prioritized effectively. **Figure 17** presents an aggregate view of installation issues, severity issues, and Loss in Production Hours (LPH) per CC at one of the analyzed sites, FC3. To incorporate all three attributes in the same graph, LPH has been scaled down at a 1:50 ratio. This graph reveals that for the three CCs with the most severity events, there is either minimal to no IOQ performed, or no issues are identified during the IOQ process. After speaking with the IOQ team, it has been verified that the former is true. While the LPH for these CCs is on the lower
end, the graph presents an opportunity for \textit{ABCD} to reduce unwanted noise in severity events raised. A similar pattern is observed for CC411 which experienced the highest LPH and no trace of IOQ testing. In addition, CC411 is one of the first few control cabinets to be installed and should have been fully IOQ tested. However, a mismatch in IOQ testing prioritization caused over 6000 hours in LPH during FC operation which translates to over $138,000 from one control cabinet alone.

![Graph showing severity events, IOQ red tags, and LPH by control cabinet for FC3.](image)

**Figure 17**: Severity Events, IOQ red tags, and LPH by Control Cabinet for FC3, highlighting the mismatch in testing prioritization.

Given the tight timelines and observed pattern of incomplete IOQ testing across the FC network, it is critical to introduce a testing priority for different CCs in order of business impact to maximize the value of IOQ performed during the available time frame. A Python program has been written to recommend the testing priority for control cabinets. The program considers the number of assets per CC, IOQ red tags, severity events, LPH concentration within a specific CC, and CC criticality. The proposed recommendations, based on the FCs analyzed, are shown in Figure 18. The detailed description of the control cabinets and additional program outputs have
been hidden to protect proprietary information. A Python program to aggregate different data sources and create the visual representation of installation issues, operational issues, and associated impact to business as shown in Figure 17 has also been developed and submitted.

<table>
<thead>
<tr>
<th>Control Cabinets</th>
<th>IOQ Timelines</th>
<th>Required Human Resources (40hrs/week)</th>
<th>Total required hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC920 &amp; CC921</td>
<td>3 days</td>
<td>2</td>
<td>240</td>
</tr>
<tr>
<td>CC412</td>
<td>3 days</td>
<td>2</td>
<td>240</td>
</tr>
<tr>
<td>CC950 &amp; CC951</td>
<td>6 days</td>
<td>2</td>
<td>480</td>
</tr>
<tr>
<td>CC910</td>
<td>3 days</td>
<td>4</td>
<td>480</td>
</tr>
<tr>
<td>CC512</td>
<td>4 days</td>
<td>2</td>
<td>320</td>
</tr>
<tr>
<td>CC521</td>
<td>2 days</td>
<td>4</td>
<td>320</td>
</tr>
<tr>
<td>CC410</td>
<td>3 days</td>
<td>4</td>
<td>480</td>
</tr>
<tr>
<td>CC530</td>
<td>5 days</td>
<td>4</td>
<td>800</td>
</tr>
<tr>
<td>CC220</td>
<td>3 days</td>
<td>2</td>
<td>240</td>
</tr>
<tr>
<td>CC914</td>
<td>2 days</td>
<td>2</td>
<td>160</td>
</tr>
<tr>
<td>CC510</td>
<td>1 day</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>CC940</td>
<td>3 days</td>
<td>4</td>
<td>480</td>
</tr>
<tr>
<td>CC915</td>
<td>2 days</td>
<td>2</td>
<td>160</td>
</tr>
<tr>
<td>CC411</td>
<td>1 day</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>CC921</td>
<td>3 days</td>
<td>2</td>
<td>240</td>
</tr>
<tr>
<td>CC946</td>
<td>2 days</td>
<td>2</td>
<td>160</td>
</tr>
<tr>
<td>CC916</td>
<td>2 days</td>
<td>2</td>
<td>160</td>
</tr>
<tr>
<td>CC211</td>
<td>1 day</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>CC213</td>
<td>1 day</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>CC430, 431</td>
<td>2 days</td>
<td>2</td>
<td>160</td>
</tr>
<tr>
<td>CC201</td>
<td>2 days</td>
<td>2</td>
<td>160</td>
</tr>
<tr>
<td>CC507</td>
<td>1 day</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>CC223</td>
<td>1 day</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>CC917</td>
<td>2 days</td>
<td>2</td>
<td>160</td>
</tr>
</tbody>
</table>

**Figure 18**: Recommended prioritization for IOQ tests.

The program assumes that the installation of these control cabinets is complete when IOQ is expected to begin, which in theory should be true but in reality is not always the case. Installation timelines and priorities shift very often. Hence, to provide the operations engineering team flexibility to deviate from the recommended prioritization above as needed, **Figure 18** also includes the estimated hours – provided by the IOQ planning team – it takes to fully complete
the IOQ in any control cabinet. IOQ testing prioritization along with the time required for complete IOQ allows the Operations Engineering team to be smart about their time allocation for testing different control cabinets and enhance the value added to ABCD’s business. Further, the program empowers the team to dynamically adjust the testing prioritization every year, as changes in FC design, vendor, and other installation parameters may introduce new patterns of installation errors.

The two analyses above – expanding the IOQ testing coverage and introducing IOQ testing prioritization – focus on improving the existing IOQ process to achieve our project objective of reducing early operational issues thereby enhancing early FC operation. While the recommendations for IOQ process improvements have the potential to significantly reduce early operational issues, this project has also explored other processes and possibilities that would further reduce operational challenges in a newly built FC and improve its startup experience. In particular, preventive and predictive opportunities are considered which is discussed in the next section.

3.3 Exploration of Preventive Opportunities

The fundamental cause of early operational issues is sub-optimal installation. Year-over-year, ABCD is experiencing degrading/lower quality of vendor FC turnover. In each of the FCs analyzed, a significant number of installation issues were recorded, as shown in Figure 19, Figure 20, and Figure 21, during the ABCD IOQ process – a qualification process post vendor qualification – indicating potential issues with vendor installation quality. Early attempts to analyze the vendor IOQ results reveals that ABCD does not have visibility into vendors’ IOQ checklists and the test results. ABCD performs the IOQ process independently of vendor IOQ.
The next phase of the project focuses on ways to integrate efforts between ABCD and the vendors for IOQ testing, and to increase vendor accountability in delivering higher quality FC. Overall, the motivation is to prevent, or at least reduce, installation flaws.

Figure 19: Distribution of installation issues by failure category at FC1, about 85% IOQ complete before FC operation
Figure 20: Distribution of installation issues by failure category at FC2, about 70% IOQ complete before FC operation
Figure 21: Distribution of installation issues by failure category at FC3, about 60% IOQ complete before FC operation
3.3.1 Leveraging vendor relationship

The thesis research seeks to find ways to improve the installation quality itself. If installation is good, early operations will not have any problem even if IOQ is not thorough. These improvements would thus need to start with the first stage – vendor installation.

Shadowing one of the FCs from the year 2019 during the IOQ process indicated that the vendor testing might have been skipped altogether. Many issues were evident to ABCD associates purely through visual inspection even before running any IOQ tests. These issues should have been caught during the vendor inspection process prior to FC turnover. An effort to study the tests that were performed and understand why these issues were not caught during vendor testing reveals that ABCD does not have the visibility into vendor test checklist and their test results. This meant that there is no way for ABCD to ensure that installation is tested before turnover to ABCD. It is therefore recommended that ABCD demands the results of vendor installation testing in the future. Visibility into the results of the vendor checks will ensure that the vendor IOQ is indeed complete before FC turnover to ABCD. Further, knowing the outcomes of the tests performed by the vendor will help ABCD adjust their own IOQ checklist by avoiding duplicate tests and focusing ABCD’s energy on equipment not well-tested by the vendor. This opportunity to integrate the effort has been highlighted to ABCD.

The shadowing experience also helped uncover the lack of a threshold metric for installation quality. Multiple repeated issues are identified and logged at the same CC. A maximum acceptable failure count per CC could be introduced in the vendor contract to enforce a quality standard at delivery. If the number of issues identified by the ABCD IOQ team in a CC hits a predefined threshold, a broad supplier/vendor recheck should immediately be triggered.
The number of disqualified CCs could be noted and incorporated into future vendor allocation decisions.

A Python program has been developed to recommend the maximum acceptable installation failures per CC. The program considers the number of assets per CC, criticality of those assets, and an acceptable failure percentage based on the criticality information. The acceptable percentage of failures is a configurable property that can be modified as needed. For the study, acceptable failure percentages are 5%, 10%, 12%, and 15% for most critical, critical, medium critical, and less critical control cabinets, respectively, as recommended by the Operations Engineering team. Control cabinets with an overall calculated criticality – average of the criticality of all assets in the control cabinet – greater than 2.6 are considered most critical, between 2.2 and 2.6 critical, between 2.0 and 2.2 medium critical, and below 2 less critical. The proposed recommendation, based on the FCs analyzed, is shown in Figure 22.

<table>
<thead>
<tr>
<th>Control Cabinets</th>
<th>Maximum Threshold</th>
<th>Control Cabinet Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC915</td>
<td>38</td>
<td>Lower AFE 3.0 Rebin (Lanes 19-24)</td>
</tr>
<tr>
<td>CC912</td>
<td>37</td>
<td>Lower AFE 3.0 Rebin (Lanes 1-6)</td>
</tr>
<tr>
<td>CC913</td>
<td>36</td>
<td>Lower AFE 3.0 Rebin (Lanes 7-12, JP)</td>
</tr>
<tr>
<td>CC914</td>
<td>36</td>
<td>Lower AFE 3.0 Rebin (Lanes 13-18)</td>
</tr>
<tr>
<td>CC942</td>
<td>18</td>
<td>Upper AFE 2.0 Upper Rebin (Lanes 1-8)</td>
</tr>
<tr>
<td>CC943</td>
<td>16</td>
<td>Upper AFE 2.0 Rebin (Lanes 9-16, JP)</td>
</tr>
<tr>
<td>CC201</td>
<td>14</td>
<td>Receiving</td>
</tr>
<tr>
<td>CC510</td>
<td>14</td>
<td>Shipping Singulator</td>
</tr>
<tr>
<td>CC944</td>
<td>13</td>
<td>Upper AFE 2.0 Rebin (Lanes 17-22)</td>
</tr>
<tr>
<td>CC651</td>
<td>12</td>
<td>Singles Pack Lines (1, 2, 3)</td>
</tr>
<tr>
<td>CC916</td>
<td>11</td>
<td>LWR AFE 3.0 OB (OB 1, 2) (from walls 1-12)</td>
</tr>
<tr>
<td>CC414</td>
<td>10</td>
<td>Routing TA 21 – 39</td>
</tr>
<tr>
<td>CC917</td>
<td>10</td>
<td>LWR AFE 3.0 OB (OB 3, 4) (from walls 13 - 24)</td>
</tr>
<tr>
<td>CC223</td>
<td>9</td>
<td>RSP 3 (S) - Level 3 (dock side) (PH1)</td>
</tr>
<tr>
<td>CC946</td>
<td>8</td>
<td>LWR AFE 3.0 OB (OB 1, 2) (from walls 1-12)</td>
</tr>
</tbody>
</table>

**Figure 22:** Recommended maximum acceptable threshold for IOQ failures per CC (omitting control cabinets where maximum acceptable threshold < 8).
Having a preset and agreed upon metric for installation quality would compel the vendors to improve the pre-handover processes to ensure that this threshold is met. This quality threshold can be iteratively upgraded to improve the quality of installation year-over-year, thereby limiting the number of installation related issues that impact operations.

Another analysis from a preventive lens includes study of frequently failing parts. Operational challenges caused due to equipment failures that require part replacement can provide insight into their sensitivities. Further, preventative maintenance schedules could help identify gaps in maintenance. Based on this information, robust qualification could be introduced on these parts during the IOQ process, and a stricter preventative maintenance schedule could be enforced to avoid part failures altogether and thereby improve FC startup. This study is described next.

**3.3.2 Study of replacement parts**

Data on parts replaced in the first 90 days of operations is examined to identify the most frequently failing equipment. For each of the replaced equipment, Loss in Production Hours (LPH), and equipment Preventative Maintenance (PM) schedule data are gathered. A Python program has been written to highlight the most frequently replaced parts in an FC using the word frequency analysis described in Section 3.2, using the replaced parts description from this dataset. This program also generates an output file that provides the ability to view equipment that has been replaced and that did not have a regular maintenance program.

Belts, rollers, and motors surface as the most problematic parts. While all these parts are tested during the IOQ process, without enough volume and real operational load, this equipment
is not exposed to the stress level that causes damage to the parts. Adoption of the E-Prime or CLT programs mentioned in Section 3.2.1 is strongly recommended.

Data also shows that preventative maintenance does not cover all of the frequently failing parts identified above. Stricter preventative maintenance schedules could be introduced for these parts. While the objective of this study is to use early failures to avert later failures in these pieces of equipment at the FCs analyzed, this information could also be used when creating a PM schedule for future sites.

3.4 Exploration of Predictive Maintenance Opportunities

Based on reactive and preventive analysis above, the last phase of this project looks at predictive maintenance prospects to enhance FC early operations. Sensors can capture failure signals and equipment anomalies in an FC before they become severity issues. Real-time data collected via wireless sensors can help detect when a critical piece of equipment is close to failure, and influence decisions about when and how to carry out maintenance operations to lessen the likelihood of a severity event.

All the FCs under study already have wireless sensors installed on some pieces of equipment. An analysis is performed to identify additional equipment without existing sensors that might benefit from having them so that predictive maintenance would be possible and severity issues could be avoided. An optimization model has been written to identify equipment for which sensors should be introduced. The model also computes the Return on Investment (ROI) timelines for the sensors if installed.

The purpose is to perform a cost-benefit analysis of introducing additional sensors. The parameters that drive the output include cost of the sensors, probability of equipment failure,
severity issues on each equipment, criticality of the equipment to business, and the effectiveness of sensors in predicting potential failures. Cost of the sensors – one-time setup and installation and ongoing yearly maintenance – has been provided by the RME team. The approach and calculation details are explained below.

**Impact to business calculation:**

Impact to business in the first 90 days (from new sensor installation) = Probability of the equipment failing × Impact to business if the equipment fails × Probability of a sensor capturing the failure before it happens

**Step 1: Calculating the probability of equipment failing**

Probability of equipment failing = \( \frac{\text{Total number of times the equipment failed in the FC}}{\text{Total number of possible failures in the FC}} \)

- Total number of times the equipment failed in the FC = Severity issues reported
- Total number of possible failures in the FC = Average failure per equipment × Total number of equipment in the FC

Average failure per equipment: As this value is unavailable, average failure per sensor-installed equipment is calculated and used as a best estimate.

- Average failures per sensor-installed equipment = \( \frac{\sum \text{Total possible failures for each equipment with sensors}}{\text{Count (Equipment in the FC with sensors installed)}} \)

- Total possible failures for each equipment with sensors = Number of times the sensors detected an event/failure before happening + Actual failure that happened despite the sensors

**Step 2: Calculating the impact to business if an equipment fails**

Impact to business if the equipment fails = Total Loss in Production Hour (LPH) for the given equipment × $23/hour

**Step 3: Calculating the probability of a sensor capturing the failure before it happens**

- Network wide average probability of a sensor capturing the failure before it happens = \( \frac{\sum \text{probability of installed sensor catching the event for each equipment across network}}{\text{Count (Equipment with sensors installed across all FC in the network)}} \)
For each equipment with sensors installed, probability of installed sensor catching the event = 
\[
\frac{\text{Number of times the sensors detected an event/failure before happening}}{\text{Total number possible failures for the equipment (Events avoided + event that happened)}}
\]

Output from the calculations in step 1, 2 and 3 are used in the overall equation to compute the business impact of installing new sensors in certain piece of equipment. The results for the three FCs are presented below. Column **ROI timeline** is the time in years after which the sensors pays for itself. Column **Impact to Business in the first 90 days** presents the value to business in dollars in the first 90 days of operation.

### FC1

<table>
<thead>
<tr>
<th>Equipment</th>
<th>ROI timeline</th>
<th>Impact to Business in the first 90 days (in $)</th>
<th>Asset Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000132639</td>
<td>2</td>
<td>87</td>
<td>Intelligrated Soft Touch Sorter</td>
</tr>
</tbody>
</table>

### FC2

<table>
<thead>
<tr>
<th>Equipment</th>
<th>ROI timeline</th>
<th>Impact to Business in the first 90 days (in $)</th>
<th>Asset Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000141410</td>
<td>1</td>
<td>240</td>
<td>SL2 Upper AFE 2.0 Sorter</td>
</tr>
<tr>
<td>1000141973</td>
<td>1</td>
<td>120</td>
<td>SL2 Lower AFE 3.0 Sorter</td>
</tr>
<tr>
<td>1000146601</td>
<td>1</td>
<td>82</td>
<td>SL2 Shipping Sorter</td>
</tr>
<tr>
<td>1000146739</td>
<td>3</td>
<td>35</td>
<td>Cross belt Sorter</td>
</tr>
<tr>
<td>1000141653</td>
<td>3</td>
<td>34</td>
<td>SL2 Routing Sorter</td>
</tr>
<tr>
<td>1000140136</td>
<td>4</td>
<td>23</td>
<td>Lower AFE 3.0 Merge/Recirc/Inbound</td>
</tr>
</tbody>
</table>

### FC3

<table>
<thead>
<tr>
<th>Equipment</th>
<th>ROI timeline</th>
<th>Impact to Business in the first 90 days (in $)</th>
<th>Asset Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000112550</td>
<td>1</td>
<td>601</td>
<td>Tote Routing FlexSorter</td>
</tr>
<tr>
<td>1000112533</td>
<td>1</td>
<td>262</td>
<td>Upper AFE 2.0 Sorter</td>
</tr>
<tr>
<td>1000113549</td>
<td>1</td>
<td>240</td>
<td>2340 Transition Bed</td>
</tr>
<tr>
<td>1000115829</td>
<td>1</td>
<td>134</td>
<td>Intralox Sorter</td>
</tr>
<tr>
<td>1000115864</td>
<td>1</td>
<td>96</td>
<td>SC3 Cross belt Sorter</td>
</tr>
<tr>
<td>1000115956</td>
<td>1</td>
<td>78</td>
<td>SL2 Upper AFE 2.0 Sorter</td>
</tr>
<tr>
<td>1000113557</td>
<td>2</td>
<td>60</td>
<td>SL2 Lower AFE 3.0 Sorter</td>
</tr>
<tr>
<td>1000115865</td>
<td>3</td>
<td>35</td>
<td>SL2 Shipping Sorter</td>
</tr>
<tr>
<td>1000112538</td>
<td>5</td>
<td>20</td>
<td>2340 Transition Bed</td>
</tr>
<tr>
<td>1000113550</td>
<td>5</td>
<td>19</td>
<td>Lower AFE 3.0 Rebin (Lanes 13-18)</td>
</tr>
</tbody>
</table>

**Figure 23:** Sensor analysis for predictive maintenance highlighting equipment for new sensor introduction.

Using wireless sensors on various equipment in an FC for predictive maintenance techniques can reduce maintenance costs, minimize the risk of catastrophic failures, and
maximize system availability. The computer program that generated the output above has been shared with the RME team. The majority of the equipment in the list above was already being considered for new sensor installation. This program provides a repeatable data-driven approach to confirm their intuition on which equipment needs sensors.

These reactive, preventive, and predictive analyses provide a robust understanding of the current processes and available data, and have helped in formulating systematic approaches to improving the FC startup experience. The programs developed throughout this study will empower ABCD to continuously improve their processes and achieve their target operational efficiency at FC startup. The next section considers the potential cost savings from each of the recommendations developed during the study.

3.5 Financial Analysis

During the internship, different approaches have been formulated to aid ABCD in minimizing the operational challenges during FC startup, as detailed above. The study in this section presents the value proposition of each of these recommendations in the first 90 days of FC operations alone. Data from the three FCs analyzed during the study are used for calculating the value of each recommendation to ABCD’s business. While the results are presented here, the detailed calculations – as shared with ABCD – are not included for confidentiality reasons. Note that the cost savings of each recommendation are not mutually exclusive as there is some overlap in the benefits gained.

Three recommendations have been made based on the analysis from the reactive lens: new tests for equipment not covered by current IOQ process, E-Prime and CLT programs to simulate real operational conditions in newly built FCs for testing, and IOQ testing prioritization.
Including tests for Slam, SmartPac, VRC, and Rebin will save ABCD roughly $24K per FC. These tests will also significantly reduce noise in the severity events, allowing the RME team to allocate their effort elsewhere, especially in predictive and preventative maintenance.

Introducing the E-Prime and CLT programs has the biggest prospect to eliminate startup issues and strengthen ABCD operations. However, there is a cost associated with implementing and running programs at this scale. A thorough study of the required investment and the cost savings from eliminating severity issues shows that the savings of about $375K per FC are offset by the logistics costs for running these programs. While the financial incentive in the first 90 days is neutral, these processes are highly recommended as they pay for themselves in 90 days of FC operation, while strengthening the ABCD FC infrastructure for the rest of its lifetime.

Establishing a testing prioritization will enable the North Americas Ops Engineering team to be most effective in the available time and secure the cost savings of about $350K per FC.

Recommendations from the preventive lens study include introducing and enforcing a quality threshold per CC to improve the installation quality itself, and introducing stricter preventive maintenance schedules for sensitive equipment in the FC. Increasing vendor accountability by setting the threshold will not only help reduce severity issues but also help save time during IOQ inspection and red tags data logging. Further, as the team performing ABCD IOQ process travels to the site for IOQ, significant logistics cost reduction is possible if the time to perform IOQ tests and to log the data can be reduced. Introducing the quality threshold can therefore save ABCD $100K’s per FC. While introducing stricter preventive maintenance helps ABCD avoid severity issues, thereby strengthening ABCD operations, the calculated financial benefit offsets the cost of doing these regular maintenances. However, it is recommended to enforce stricter schedules, as the stress associated with fixing the issue after it happens is
significantly higher than performing similar fixes as a part of regular maintenance. This could easily lead to additional mistakes in other areas.

Recommendations from the predictive lens include introduction of wireless sensors on certain equipment in the FC. Based on the probability of these new sensors correctly reporting anomaly signals so predictive efforts can be taken before an event takes place, there is an opportunity for ABCD to save over $100K per FC.

The data show that there is a business value in implementing each of the recommendations proposed in this thesis.
Chapter 4: Future of Work and Conclusion

In summary, our thesis research resulted in the development and implementation of a variety of new processes to improve the FC installation quality, thereby limiting startup operational challenges. This section will provide recommendations on strategic future direction for ABCD based on observations and information gathered throughout the project, ideas for future studies, and concluding remarks.

4.1 Future of Work

As ABCD continues to grow, its distribution network will expand and become more complex. The ABCD team must, therefore, continue to nurture its obsession with testing and invention in new FCs, while having the patience to see the results through to the end. The work compiled during this internship project provides a strong foundation for enhancing installation quality and improving the FC start-up process. The immediate next phase should involve a pilot program including all of the deliverables presented in this thesis. A parallel effort should be invested in incorporating the recommendations below for refining the data-gathering process to improve future model outcomes.

Currently, the testing is done on the asset level – unit testing. Hence, severity issues are also reported on the asset level. System and process-level checks could be introduced to support functional testing within the FC. Moving to system-level tests will help future analyses target different functional areas as needed. As an initiative to facilitate this, critical functional areas – AFE.SHOE, MOD, PCK, RTE.SHOE, SHP.CB, SHP.MAN, SHP.SHOE, TSHP.ARB, TSHP.POPUP, RCV, SHP.ARB, TSHP.SHOE, RCV.POPUP – at AR sortable sites were
identified during the internship. Data on installation and operational issues grouped by these functional areas will help provide opportunities for more focused study.

The NA Ops Engineering team should also investigate ways to seamlessly access data available to different teams within ABCD. This enables programmatic extraction of data for analysis, and, in doing so, automates the process by eliminating the manual steps of querying and extracting the data. A lack of a system user account to access data from the EAM databases forced the introduction of manual interventions for each of the programs written during the internship. SQL queries were manually run, data was downloaded locally, and programs were run on this data.

Data discrepancy was one of the biggest issues faced during the internship. Two different data sources – NA Ops Engineering Databases and EAM Databases – had conflicting information on the Material Handling Equipment (MHE) installed at each FC. Information on these physical assets that are installed in the FCs should be synchronized between these two critical data sources. In addition, where possible, these two systems should use the same vocabulary. For example, when an issue is raised at an FC – installation or severity – a closing code (problem-failure-cause code) is captured. Problem, failure, and the cause codes are a set of acronyms that help explain the issue on a high level. However, the vocabulary for the closing code used while capturing an installation issue is significantly different than that used during severity events. Even though both systems started with the same set of closing codes, over time, they have diverged significantly. The EAM databases that store data on severity events on assets have about 345 distinct closing codes, almost 10 times as many as NA Ops Engineering databases, which capture information on installation issues. A one-time data sync should be considered to bring both systems to a common ground. Further, proper checks and restrictions
should be introduced to prevent similar data divergence in the future. In addition to the discrepancy in data between these two different systems, equipment criticality data is not uniform across all the FCs in ABCD’s Network. This data is sourced from the home operations team at each FC, so divergent experiences might play a factor in collecting this data. Non-uniform network-wide data impacts the analyses. The NA Ops engineering team should direct effort to understand what experiences go into collecting the asset criticality and if there are data that would inform better standard values. Uniform asset criticality data should be established across the network as an immediate next step to enhance the outcome of ongoing and future analyses.

ABCD invests heavily in collecting quality data so continuous learning and improvement is a possibility. They should continue investing in enhancing the quality of the data being collected. The better your data’s quality, the more you can get out of it. With the advent of various machine learning algorithms, the better the data a machine learning model has, the faster these algorithms can produce results, and the better those results will be. During this research, different machine learning algorithms have been used on the severity descriptions to identify the most frequently failing equipment and the problematic failure modes in the FC. The quality of the model outcome thus relies on these severity descriptions captured. Currently, when severity tickets are raised by an associate, the description usually indicates the visually observed issues. More qualitative descriptions could help improve the results. An initiative should be taken to enrich this initially captured description with additional information as the issue gets addressed. It is recommended that the associate provide a post-fix “updated description” explaining what the actual problem was, why it happened, and what the fix was. This is similar to root cause analysis before closing the issue. This will serve three purposes. First, it aids in teaching
associates and helps them recognize different modes of failures and applicable fixes. This equips them to be more descriptive and thoughtful when raising similar future issues. Second, it provides a more descriptive explanation of all issues that get raised in an FC. Various natural language processing algorithms can be performed on the data to identify patterns and trends. Third, it ensures that severity events are raised only for issues that actually need attention, and in doing so, eliminates white noise in the severity events raised and improves the developed model’s outcomes.

Beyond these aforementioned recommendations to enhance the future model outcomes using descriptive data, the logical next step would be to identify ways to incorporate other forms of data in the analyses. Many digital images and pictures are collected during the installation inspection process. Images should also be captured during severity events. Different machine learning techniques for image recognition can then be explored to infer valuable insights from these pictorial data. Image recognition gives a machine the ability to interpret the input received through computer vision and categorize what it sees. This powerful feature can be leveraged to catch visual cues in the pictorial data and automatically detect defects in installation.

4.2 Conclusion

This project set out to rectify a problem in ABCD’s supply network. Slipping FC installation quality has led to costly severity issues during the start-up phase after FC Go-Live, reducing the early operational efficiency of new FCs.

These new FCs are being built under tight timelines to support the growing fulfillment demand. The time pressure increases the number of installation mistakes and prevents the completion of Installation and Operational Qualification (IOQ) process. Inadequate IOQ
coverage further exacerbates the situation, as even if the tests are completed, they do not hit certain key areas and possible failures. Improving checks in the areas of AFE, sorter, belt, conveyor, power, tray/tote, MRS and adding checks for SmartPac, Slam, Rebin, and VRC will have an immediate positive impact to improving performance and customer experience.

Various machine learning techniques have been applied to the descriptions of the identified early operational issues to identify patterns of failure and ways to correct them. Topic grouping alone has proved to be insufficient at the moment due to a lack of verbose descriptions, but combined with word frequency analysis, has successfully revealed missing tests and gaps in the existing testing approach.

To mitigate these identified gaps, new tests have been proposed and incorporated in the IOQ checklist for the four frequently failing pieces of equipment – SmartPac, Slam, Rebin, and VRC. E-Prime and Closed Loop Testing (CLT) approaches have been developed to simulate real operational conditions in an FC prior to go live so integration, load, and stress test can be performed on critical equipment in the FC. Detailed implementation steps have also been prepared to pilot these two testing approaches. To empower the NA Ops Engineering team to deliver higher installation quality FCs within the available timeframe, a program has been written to efficiently reprioritize the testing schedule. In addition, a programmatic data-driven method has been developed to allow the NA Ops Engineering team to seamlessly leverage their vendors in enhancing the quality of the FC turned over to ABCD. Finally, programmatic ways to establish a stricter preventative maintenance schedule and to identify opportunities for introducing sensors on new equipment have been delivered. These computer programs will allow ABCD to perform better preventive and predictive maintenance in order to avoid operational challenges.
This holistic approach will address the problem on all fronts and enable smoother FC Go-Live. It will eliminate an average of 16K lost production hours per FC in the first 90 days of operations. This will save $ABCD$ over $3.6$ Million across the new AR Sortables in the pipeline for the upcoming year.
References


