Statistical Process Control for Anomalies in Driver Pay

Lexi Gaddy, Chris Oliver, Alex Schussler

University of Arkansas College of Engineering

Abstract

Financial auditing is an important process that can be used to identify anomalies in a payroll. Drivers at J.B. Hunt are paid using an activity-based pay system. This means pay is determined by several factors: miles driven, hours driven, and exception pay. Exception pay is a specialty form of payment that is distributed for various reasons. Some of these reasons include: truck breakdown, yard cleanup, and detention. Due to the lack of standardization in exception pay, auditing is tedious for the managers. We have created a tool that outputs a list of probable outliers.

Introduction

Internal financial auditing is an important operation performed in many businesses. Performing internal auditing allows managers to assess the effectiveness of different processes and systems within the company and potentially find room for improvement. Our team was tasked with creating a new internal pay auditing process for the trucking and transportation company J.B. Hunt. This specific process will assist operations managers in identifying patterns and potential outliers in truck driver pay.

Business Terminology and Background

Board- A group of drivers managed by a single fleet manager. Refer to Figure 1 for a visual representation

Fleet- A group of boards who perform similar functions

Operations Manager- A manager who oversees a group of fleet managers and their boards

Exception Pay- A specific type of pay that occurs for unforeseen reasons, such as breakdowns, cleaning, etc.

Activity-Based Pay- Pay depends on the amount of miles driven, amount of hours worked, stops completed, exceptions, and more. Exception pay includes things such as a truck breaking down, a truck driver having to wait an excessive amount of time at a customer to drop off their load, and many others. Refer to Figure 2 for further detail.

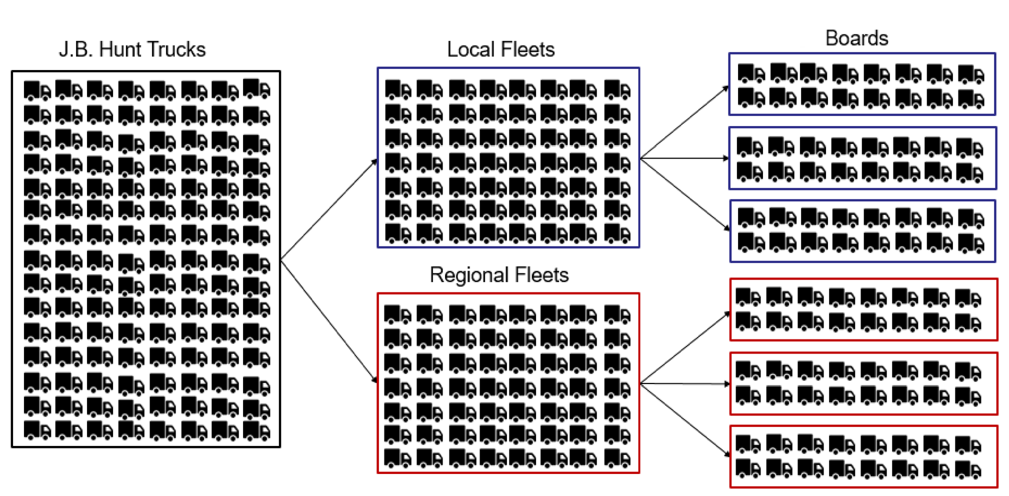


Figure 1: This graphic shows the breakdown of groups of drivers.

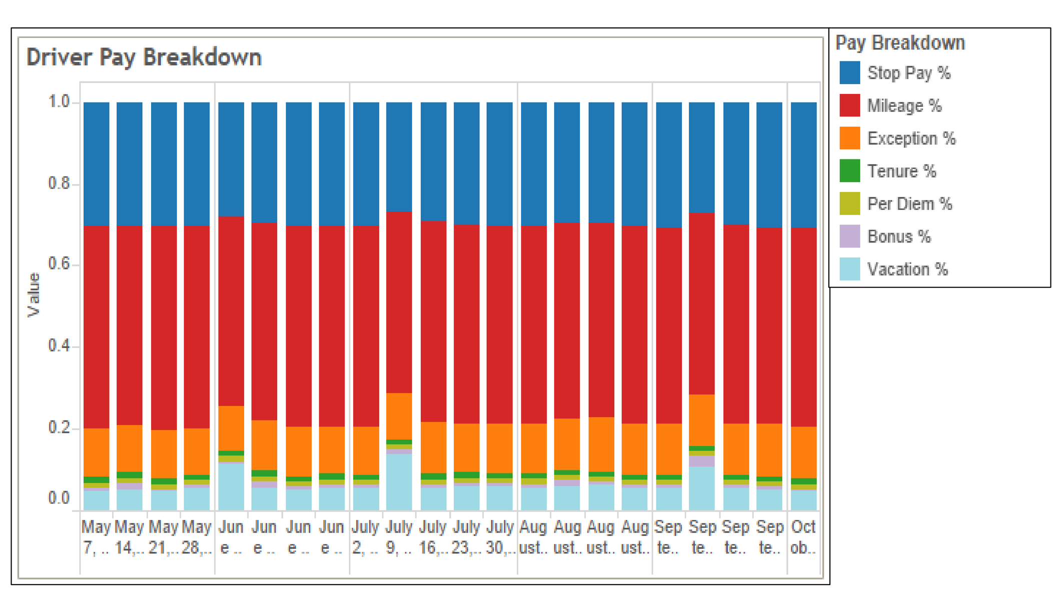


Figure 2: This table above represents the activity based pay program at J.B. Hunt and how they pay their truck drivers.

Deliverable Overview

We created the final concept which is comprised of three key coding components. One portion generates a summary of weekly statistics for each of the operations manager's boards. The second will analyze the input data using several different statistical methods, and then output intervals of reasonable pay for each type of exception within each board. These intervals will then be used to identify potential outliers within the data. Drivers who fall outside of these acceptable intervals will be output into a list containing their name, type of exception pay, and the amount in question that was flagged. The final code component will compare boards within a fleet against each other. Using ANOVA, a list will be generated containing boards that have significantly different average weekly exception pay . The figure below gives an idea of what will be input and output by this tool.

Data Analysis

In order to better understand the data and how to compare boards and drivers within a board, the data had to be thoroughly analyzed. In addition, a method for detecting outliers for each of these problems had to be chosen. It became apparent that the boards vary in distribution despite being in the same fleet. Figure 4 gives an example of how the distributions differ. This tells us that the fleet managers are paying differently. This difference could be based on the specific jobs those boards are performing, differences in location within the city, or differences in policy implementation.

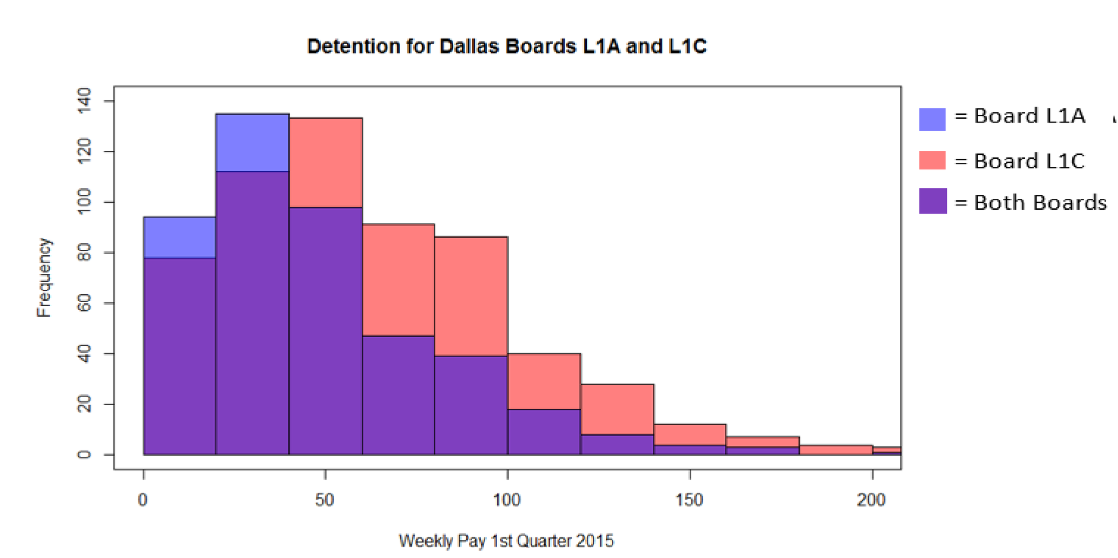


Figure 3: This figure is comparing the distribution of a specific type of exception pay for two different boards. These graphs are showing weekly detention pay for each driver over the course of a quarter. Even though these are both local boards located in Dallas, their distributions are significantly different.

Comparing Differences in Pay by Board: As we talked about earlier, some fleet managers pay exception pay differently. This is why we must compare boards to see who might be paying exception pay a certain way. For example, one fleet manager may pay detention only during the time window that the driver is supposed to be at the customer. In that case, if a driver shows up early before their designated time window, the driver may not be compensated for the time that they are there before the time window. On the other hand, some fleet managers will pay their board detention for when they show up early to customers. Comparing boards may reveal these differences in how fleet managers pay boards differently. Refer to Figure 5 for a visual representation of how boards are paying differently.

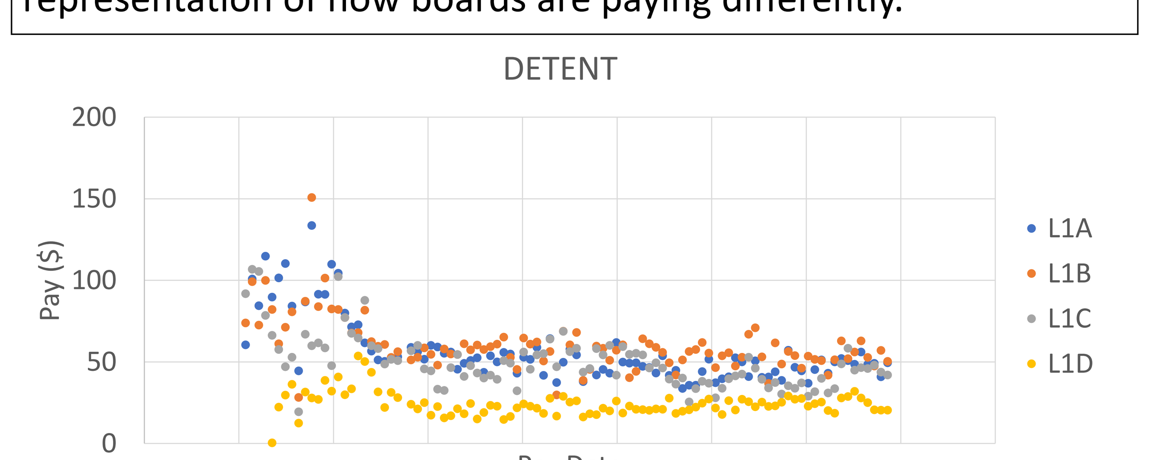


Figure 4. Average pay per driver across different boards shows how fleet managers might be paying their drivers differently.

Outlier Detection Methods

To determine how we would detect outliers in our data we researched different techniques for detecting outliers. These techniques were then implemented into our code, and we performed analysis to determine which statistical test fits this data best.

Method	Equation	Normality Assumed	Benefits
Standard Deviation	$\bar{x}$ ± 2 SD or $\bar{x}$ ± 3 SD	Yes	Most common method used in industry
Tukey's Boxplot	Inner Fences:[Q1-1.5 IQR, Q3+1.5IQR] Outer Fences: [Q1-3 IQR, Q3+3 IQR	No	Applicable to skewed data
MAD_e	Median ± 2 MAD_e Median ± 3 MAD_e	No	Unaffected by extreme values in data set
Median Rule	Median ± 2.3 IQR	No	Applicable to skewed data

Table 1: This table outlines the key benefits for each outlier detection method. Three of the four methods outlined include robust tests that account for skewness in data. Each of these tests was implemented into Python code. Adapted from (Seo 2002).

Modeling

We received several years worth of payroll data from J.B. Hunt, but no true outliers were defined in that data. In order to properly test which statistical method would be the most effective we modelled new payroll data using the statistical software R. Figure 5 gives an example of the data we created. There were around 100 outliers and 500 normal data points in this set.

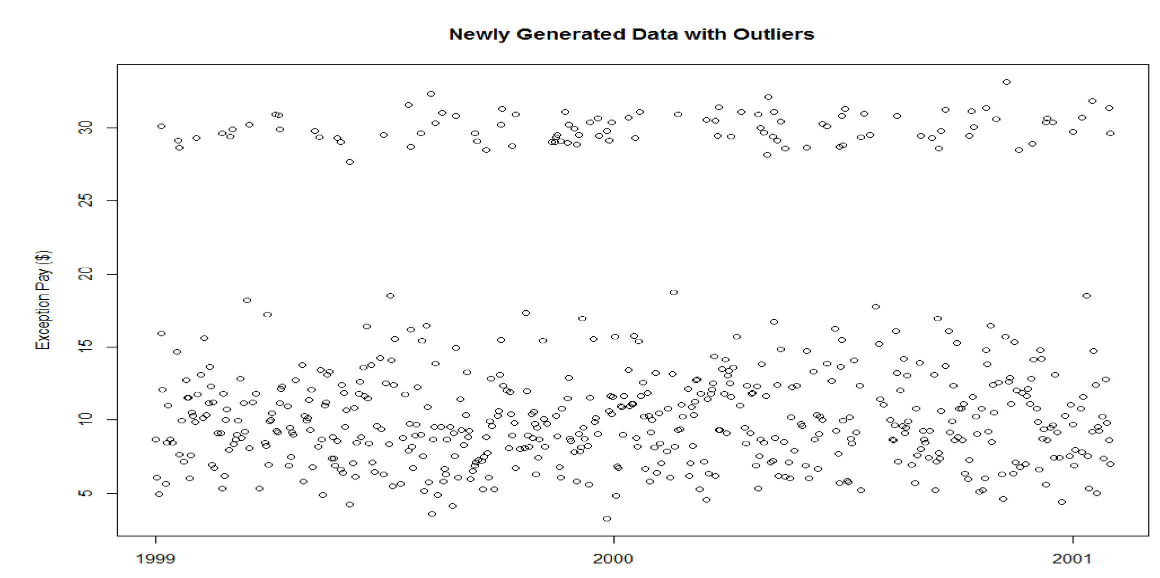


Figure 5: Generated in R, this graph shows the exception pay for each of the 600 theoretical drivers. The newly generated data in this case contained a gamma distribution, while the outliers had a normal distribution.

Modeling Continued

After testing the tool and determining that all was functioning properly, a Receiver Operating Characteristic curve, or ROC curve, was created using the output from the modelled data. An ROC curve is a graph that plots the true positive rate of a data set versus the false positive rate. In this situation, true positives were identified as the artificial outliers generated. False positive outliers in this case are false alarms that could potentially cost the operations manager time and resources to investigate.

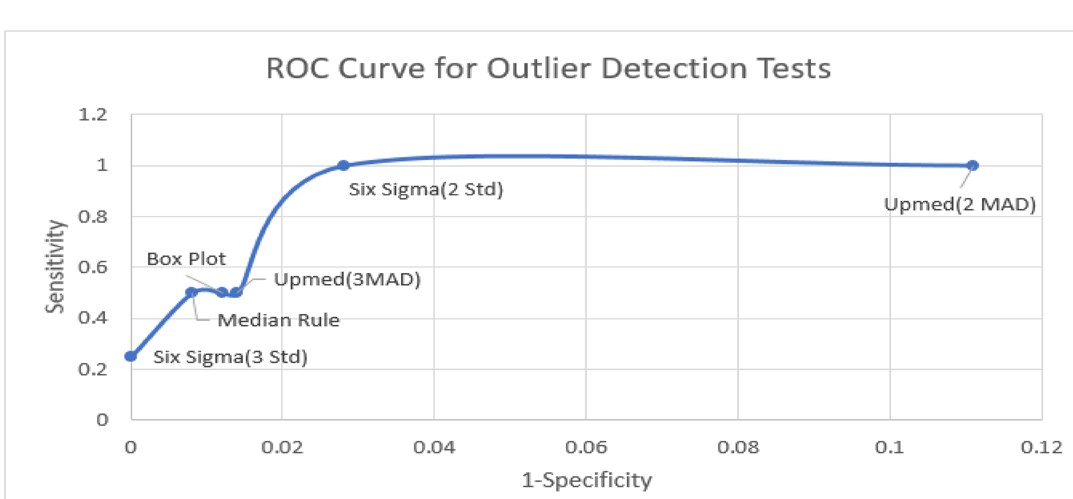


Figure 6: This ROC curve depicts the sensitivity of each method of outlier detection versus each method's false positive rate, or 1-specificity. The tests that are best suited for this situation require the lowest possible false positive rate; indicating that the three-standard deviation test will be the best fit.

Cost Analysis

We generated output based on real-world data for an operations manager who was overseeing 12 boards. This output contained 12 outliers. After receiving feedback from an operations manager, 8 of these outliers were identified as being actionable items. All 12 outliers flagged were in fact outliers, but 4 of these outliers had justifiable reasoning. From this data, we made the assumption that for every 12 boards, there will be 8 actionable items, in other words 2 outliers for every 3 boards.

	Before	After
Determining which drivers to research	20 minutes	0 minutes
True outliers identified	1/10	6/10
Total time spent per true outlier	100 minutes	13.33 minutes

Figure 7: The figure above describes the time savings associated with each identified outlier. This is based off of the feedback given by the operations manager interviewed.

$$Savings\ of\ Auditing = \overbrace{\left(\frac{2}{3}\right)}^{\text{Proportion of outliers per board } 8/12 \rightarrow 2/3} * \underbrace{(\$72.20)}_{\text{Savings estimate per outlier}} * 230\ boards + \overbrace{\left(\frac{\$35}{hr}\right)}^{\text{Hourly pay estimate for operations managers}} * \underbrace{\left(\frac{86}{60}\right)}_{\text{Time savings per outlier}} * \overbrace{230}^{\text{Average number of outliers identified}} = \$22{,}609\ per\ quarter$$

Figure 8: The graphic above shows our calculations for quarterly savings based on the assumptions stated.

Final Cost Savings = $22,609 per quarter

Conclusion: J.B. Hunt was in need of an effective payroll auditing tool that would help operations managers efficiently identify outliers. To remedy this problem, our team created a tool using the programming language Python. This tool outputs a spreadsheet with a straightforward list of most the probable outliers. Test output was sent to an operations manager to verify the outliers flagged. All in all, we received positive feedback from both the operations managers and the engineering team. For future work, this code may be implemented into a more comprehensive auditing tool with an easily accessible user interface.

References:

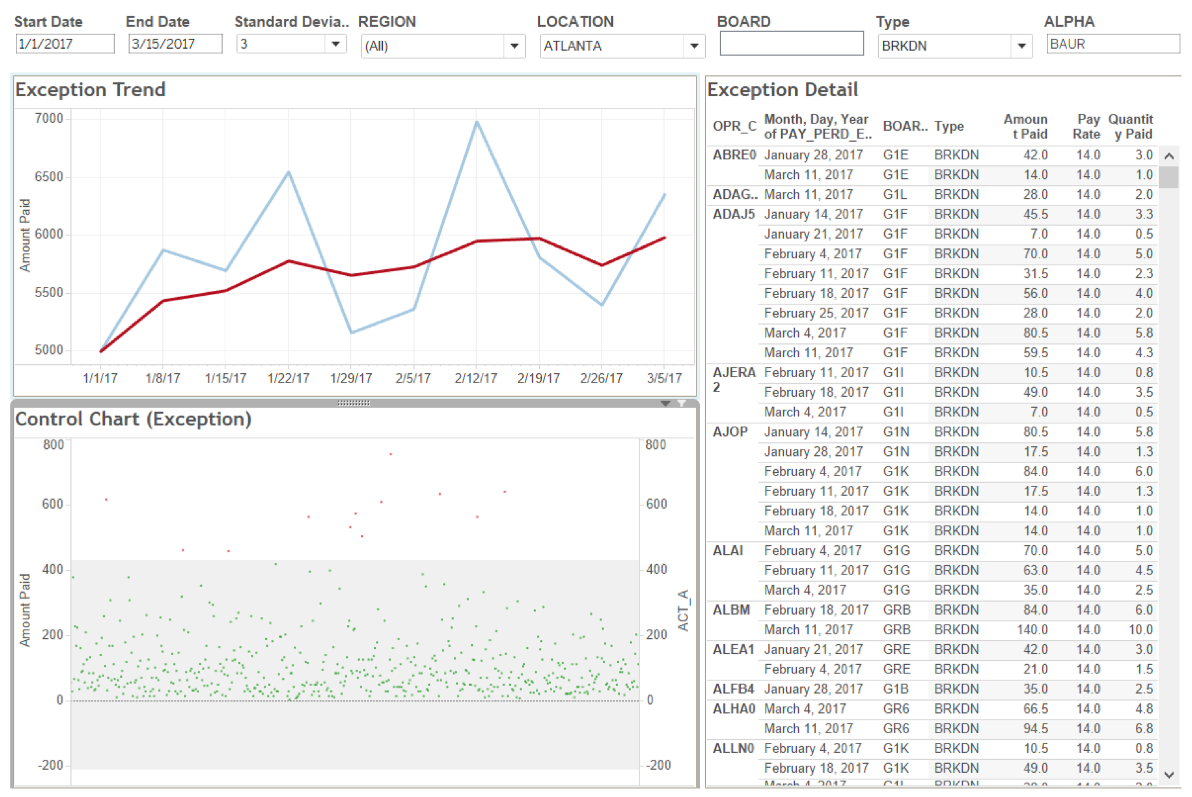
Seo, Songwon. (2002). A Review And Comparison Of Methods For Detecting Outliers In Univariate Data Sets. Graduate. University of Pittsburgh Graduate School of Public Health. Retrieved March 20, 2017, from http://d-scholarship.pitt.edu/7948/1/Seo.pdf

Statistical Process Control for Anomalies in Driver Pay

Lexi Gaddy, Chris Oliver, Alex Schussler

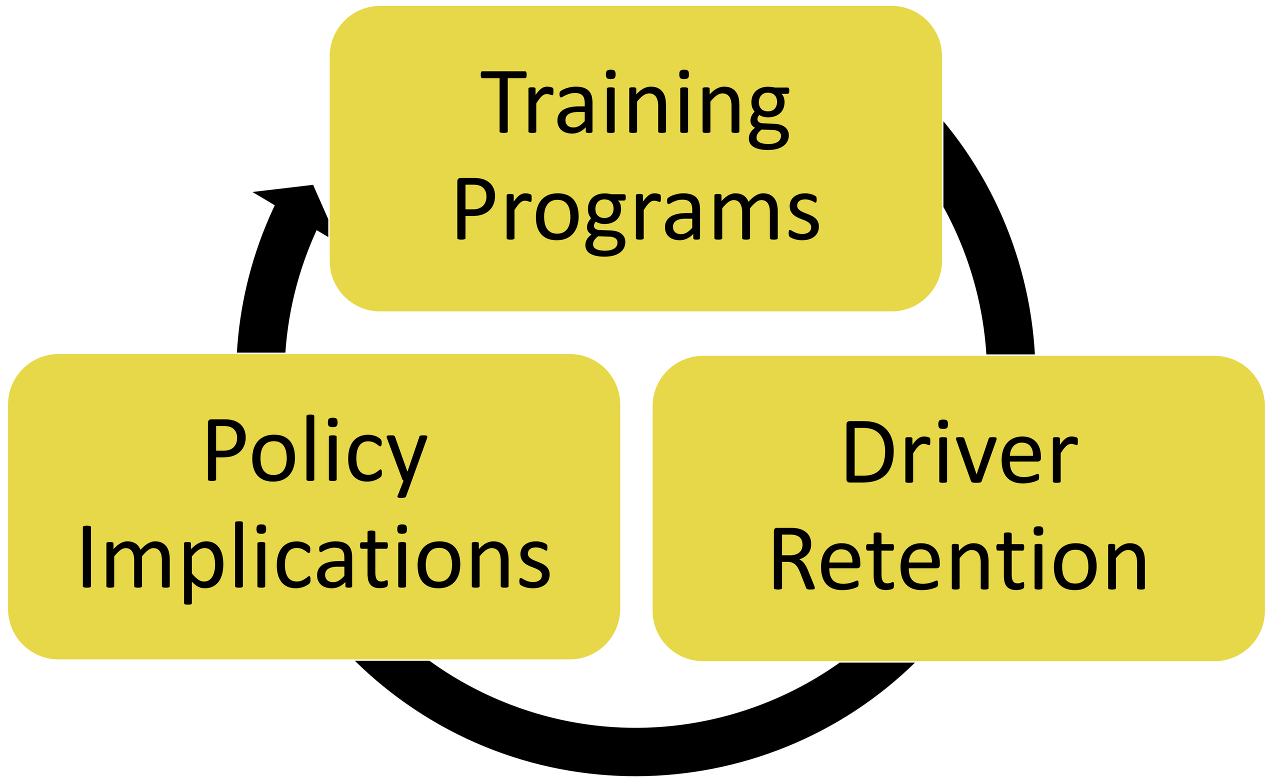
Before Tool

- Described as “cumbersome and overwhelming”
- Operations managers used their “gut feeling” when choosing drivers to audit
- Operations managers had to decide which types of exception pay to look into
- Current tool requires training in Tableau



Importance of Tool

- Hard Savings: $22,609 per quarter
- Soft Savings: If there is a 0.5% increase in driver retention, there is a potential savings of $100,000 annually



After Tool

Summary Statistics

Gives the big picture of what is going on in each board on a weekly basis including: average miles driven, average hours worked, and average pay

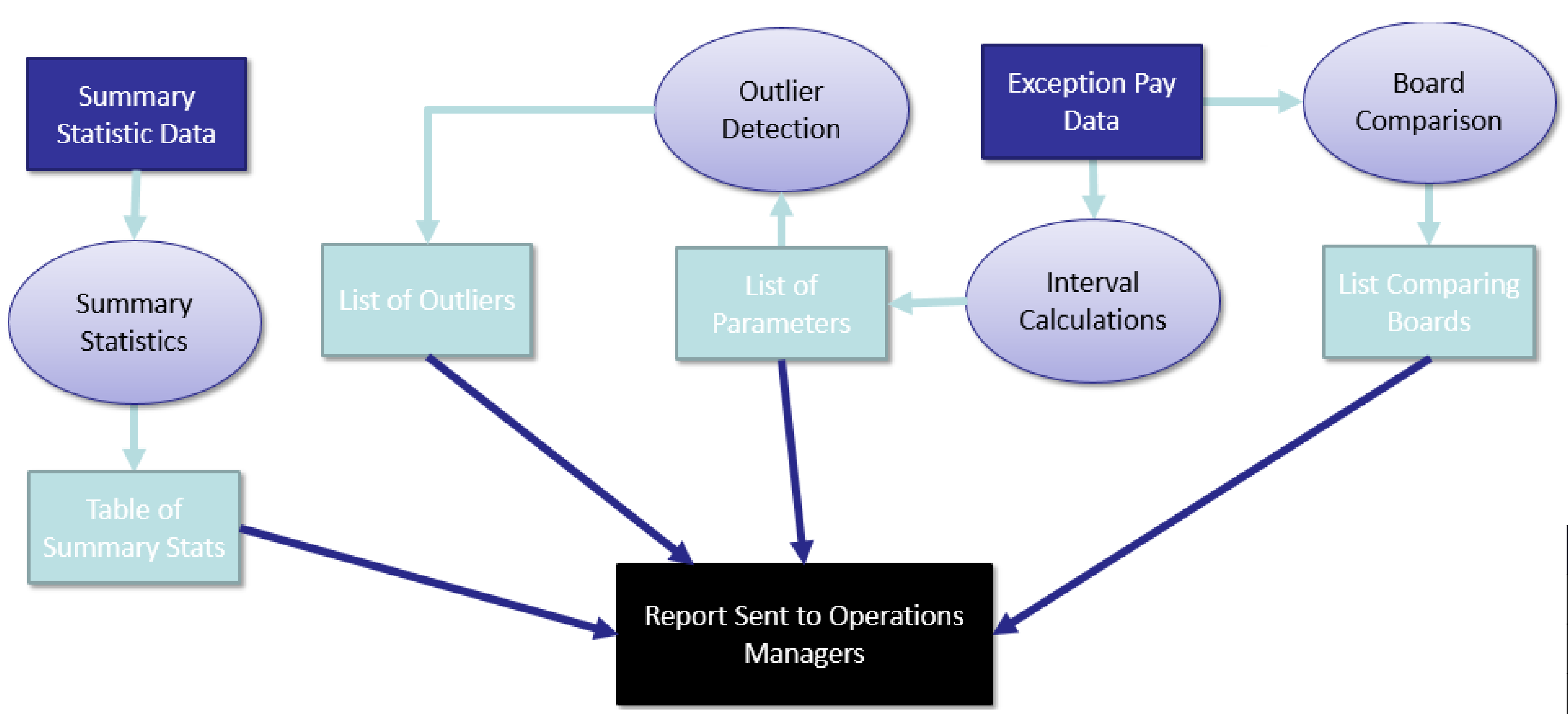
DSP_BRD_C	ULTILIZATION	AvgWkMiles	AvgWkHours	AvgWkPay
G1A	LOCAL	1113	39	902
G1B	LOCAL	1056	41	867
G1C	LOCAL	1020	37	852
G1D	LOCAL	925	37	789

Interval Calculations

The limit, based on the three standard deviation method outlined in the previous poster, is listed for every board and each type of exception pay

DSP_BRD_C	ACT_SUB_TYP	3 Standard Deviation Limit
W1#	BRKDN	336
W1#	CONGE	1177
W1#	CROSS1	98
W1#	DETENT	2454

Overview of Code



Outlier Detection

Provides a list of outliers including the driver’s name, board, type of exception and pay that was flagged

Driver	Board	Exception	Pay
ADKJ9	W2B	STRAIN	202.5
AGUJ0	W2A	EMOVEH	71.25
AGUJ0	W2A	STRAIN	90
ALLS39	W2A	RSC	209

Board Comparison

Shows which boards are paying exception pay differently

Board 1	Board 2	Different
W2A	W2E	TRUE
W2A	W2F	TRUE
W2A	W2G	TRUE
W2A	WR2	TRUE
W2B	W2D	TRUE