Comm. 493 Total Quality Management

Saskatoon Transit Stop Timeliness

Due April 6th, 2020

Matthew Buglass, 11230522

Lukas Cowan, 11238722

Jordan Jungwirth, 11219746

SUMMARY

Saskatoon Transit has been providing transit services to the city of Saskatoon since 1913. Today, the Saskatoon Transit provides roughly 40,000 rides each day. This report will highlight quality issues regarding the timeliness of Saskatoon Transits conventional bus service. Currently, the company operates using a time window of 8 minutes, 3 minutes early (-180 seconds), or 5 minutes late (+300 seconds). These upper and lower bounds are used to test ways to improve the efficiency of the transit services, specifically along route 17. Route 17 provides us a template of a simple, single-loop route that is used by a variety of people each day and can provide us insights into how a typical transit route operates.

With Saskatoon Transit being a public service, there are limits to the amount of data that can be provided for review, such as consumer insight data. However, Saskatoon Transit uses Google Transit's application programming interface. Therefore, we can collect data using this open-source program to collect stop data on route 17. This service provided by Google called GTFS Realtime is normally used to build an application for public transit services, but it also allows for the extraction of data and collection of data points along the route. These points provided us with a trip ID, stop ID, and arrival time, which is then compared to the "Transit: Bus & Train Times" app currently used by Saskatoon Transit, where we can compare scheduled stop times and actual stop times.

After analysing these data sets, it was determined that the Saskatoon transit system is wildly out of control, with several data points falling outside the 3-sigma level, based on the Saskatoon Transits definition of (+5, -3). The results of the research were divided into 3 categories for further analysis; these categories include Aggregate Data, Time Bin Data and Trip Data. After analysing the data in regard to transit timeliness, three areas of improvement for Saskatoon Transit were found. First, improve data collection abilities, second, leverage more aspects of GTFS RealTime data, and finally, move the design of the Saskatoon Transit app in house.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES AND EQUATIONS

PROJECT CHARTER

Business Case

- This project supports the business quality goals by a) providing efficient transit services and reducing variability b) improving customer satisfaction with these services

Opportunity Statement

- An opportunity exists to improve the accuracy of arrival and departure times for the Saskatoon Transit bus services.

Goal Statement

- Reduce the overall variability in bus stop timeliness

Project Scope

- This project will only look at the timeliness of regular Saskatoon Transit bus services and will omit access transit services.

Project Plan

Activity	**Start**	**End**
Define:	3/01	3/03
Measure:	3/05	3/20
Analyze:	3/22	3/26
Improve:	3/28	3/31
Control:	4/01	4/06
Track Benefits:	4/06	-----

Team

- **Team Sponsor:** Michael Moellenbeck
- **Team Members:** Matthew Buglass, Lukas Cowen and Jordan Jungwirth

1 INTRODUCTION – DEFINE

1.1 Objectives

Our objective was to analyze the Saskatoon Transit System in terms of bus arrival times on every stop related to a single route. Through the industry tour, a data value of 8 minutes was given as an acceptable range (+5, -3). These upper and lower bounds will be tested through 3 sigma control limit tests in hopes of finding ways to improve that given 8-minute range and to see where the Saskatoon Transit System can enhance its operations of the bus lines in saskatoon.

1.2 Scope

Looking into the Saskatoon Transit System's ability to deliver service on time, while only looking at route 17, as it is a simple, single loop route, that is frequented by many people of different demographics throughout the day. Due to the nature of this route, each stop is passed by approximately every 10-20 minutes. Because the route doesn't go through downtown, this route can be our best example of what our average transit system could be it does not go through heavy traffic trends.

2 PROCEDURE – MEASURE

Unfortunately, we were not permitted to analyze consumer sentiment data or see any of Saskatoon Transit's historical on-time data. Fortunately, Saskatoon Transit uses Google Transit's open protocol buffer application programming interface (API): GTFS Realtime(Saskatoon Transit, 2020). By writing a custom application (see Appendix A), we were able to collect the stop data for Route 17 on March 6th.

2.1 Data Structure

Saskatoon Transit uses an open-sourced API from Google called GTFS Realtime, which uses a protocol buffer file format to transmit estimated arrival times efficiently, delay, capacity, and GPS data (Google, n.d.; Google, 2019; Saskatoon Transit, 2020). The GTFS Realtime format allows developers to access the transit system's data to build applications and deliver services based on public transportation. Using the public APIs from GTFS Realtime and Apache POI 3.17, we were able to compile Route 17's stop data and export it to excel (protocolbuffers / protobuf; Apache POI, 2017).

2.2 Data Collection

To get an accurate representation of how the bus system performs, we ran the data extraction program all day. We collected 2,425 data points; each data point represents when a bus makes a stop to pick up or drop off passengers. A data point was represented as a node that recorded a trip ID, stop ID, and the timestamp of the stop's arrival time (see Appendix A: LinkedTransitNode). The software collected data from Saskatoon Transit's "Access Trip Updates Here" file form their Open Data from the Saskatoon Transit webpage from 08:30 to 21:00 on March 6th, 2020 (Open data from Saskatoon Transit - Access trip updates here, 2020). Each stop ID represented a different stop along Route 17, and the timestamp was stored in POSIX format; the number of seconds after 00:00:00 January 1st, 1970 UTC (coordinated universal time) (Google, n.d.).

After 21:00, the data extractor exported the data contained in the nodes to an excel sheet for analysis using Apache POI 3.17 (Apache POI, 2017). Because GTFS Realtime stores data in POSIX format, we then converted the POSIX time into Excel's date-time format (days since January 1, 1900) and used the "Transit • Bus & Train Times" app to collect the scheduled times for each stop (Transit App, Inc., 2020). By taking the difference between the scheduled and arrival times, we derived the number of seconds that the bus arrived early (negative) or late (positive) (see Appendix B).

Unfortunately, based on the data stream we collected, it appears that Saskatoon Transit only recorded stop time data when busses physically made a stop to drop off or pick up passengers. Because the times that a bus passed a stop, but did not stop, were not recorded the data for less popular stops is scarce and it becomes difficult to assess the timeliness of a bus as it progresses through its route.

GPS data was investigated to evaluate whether it would be insightful, however, only 50-60% of the buses were GPS enabled, or were broadcasting locations.

2.3 Data Organization and Consolidation

For our analysis, the data was separated into __ categories:

1) **Aggregate:** All 2,425 stops to calculate daily statistics
2) **Time:** Stops grouped into bins of 30 minutes to better analyse performance over time
3) **Stops:** Evaluating stops to identify process disruption through the route

When charting data, the same colour scheme is used for all the charts. Data points are charted in blue, means are plotted in green, the 3 sigma (statistical) limits are plotted in red, and Saskatoon Transit's goals, [-3 minutes, +5 minutes], are plotted in yellow.

3 RESULTS – ANALYSE

Generally, the Saskatoon transit system is wildly out of control. In every scenario that we analyzed, except for the bins of half an hour, several data points fell outside the 3-sigma level and Saskatoon Transit's own definitions of early (3 minutes or -180 seconds) and late (5 minutes of +300 seconds) (Moellenbeck, 2020).

3.1 Aggregate Data

Route 17 averaged 16.83 seconds late, with a standard deviation of 133.65 seconds, and a range of 1,659 seconds.

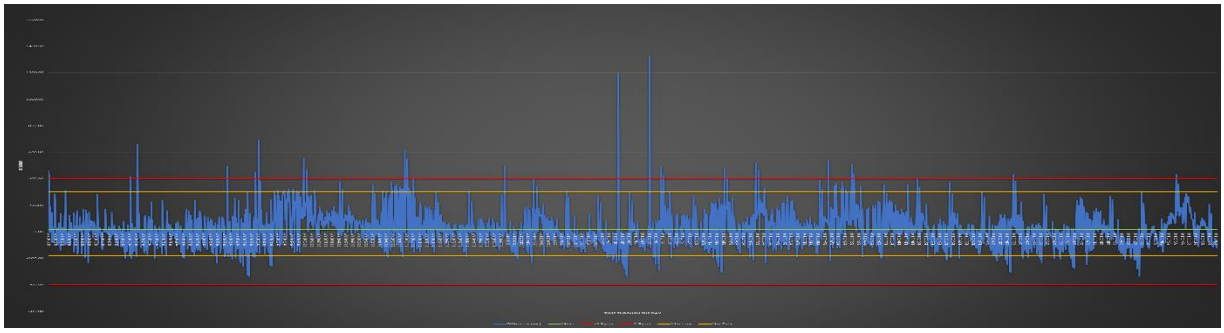


Figure 1: Aggregate stop timeliness control chart of Route 17 throughout March 6th, 2020. A larger format image can be found in Appendix C – Chart.

No apparent trends or cycles can be identified in the control chart; however, there are several times that the upper 3 sigma limit (400.95 seconds late) is surpassed. The two outliers in the center of the chart are caused by a bus that was late and fell almost a full route behind.

Using the aggregate data, three statistical tests were conducted to evaluate Saskatoon Transit's performance. The first test evaluated whether the bus system's average difference in scheduled time and stopped time was 0 (i.e. statistically on-time). The hypothesis was set as seen in Equation 1. The calculated t-score of the hypothesis test was 6.20, which surpassed the t critical value of 1.96 (see Appendix C). We reject the null hypothesis and conclude that Saskatoon Transit Route 17 is not statistically on-time.

Equation 1: Hypothesis that Route 17's timeliness is 0 (i.e. on-time)

$$\begin{cases} H_0: \bar{x} = 0 \\ H_1: \bar{x} \neq 0 \end{cases}$$

The other two statistical tests evaluated whether the mean operation of Route 17 fell within Saskatoon Transit's goals of three minutes early

and five minutes late. The hypothesis to evaluate the lower limit is displayed by Equation 2, and the upper limit hypothesis is displayed in Equation 3. Both the calculated t values for the upper and lower bounds tests were less than the critical t values (see Appendix C), hence the null hypotheses cannot be rejected, and the process mean does fall within Saskatoon Transit's limits.

Equation 2: Hypothesis to evaluate whether the process mean is below upper limit

$$\begin{cases} H_0: \bar{x} \leq 300 \\ H_1: \bar{x} > 300 \end{cases}$$

Equation 3: Hypothesis to evaluate whether process mean is above the lower limit

$$\begin{cases} H_0: \bar{x} \geq -180 \\ H_1: \bar{x} < -180 \end{cases}$$

3.2 Time Bin Data

3.2.1 Hourly Bins

In the data that was broken up into hourly bins, we can see that there is some potential for cyclicality throughout the day (see Figure 2). The cycles present do not match what would be expected of a transit route. The peaks (late arrivals) occur after rush hour (10:00, 15:00, 19:00), and the troughs occur during peak traffic hours (08:00, 12:30, 17:30). When grouped into hourly bins, the process seems to be in control, as the statistical limits are within Saskatoon Transit's goals, and the data never falls outside of the 3 sigma limits.

However, when plotted in a range (R) chart, the process is wildly out of control. Peaks are now evident at times of high traffic. From the range chart, it is evident that the process is wildly inconsistent, which is especially harmful in the service industry. The process range mean plots outside of the Saskatoon Transit's range goal (480 seconds), and often plots above the 3-sigma limit (see Figure 3).

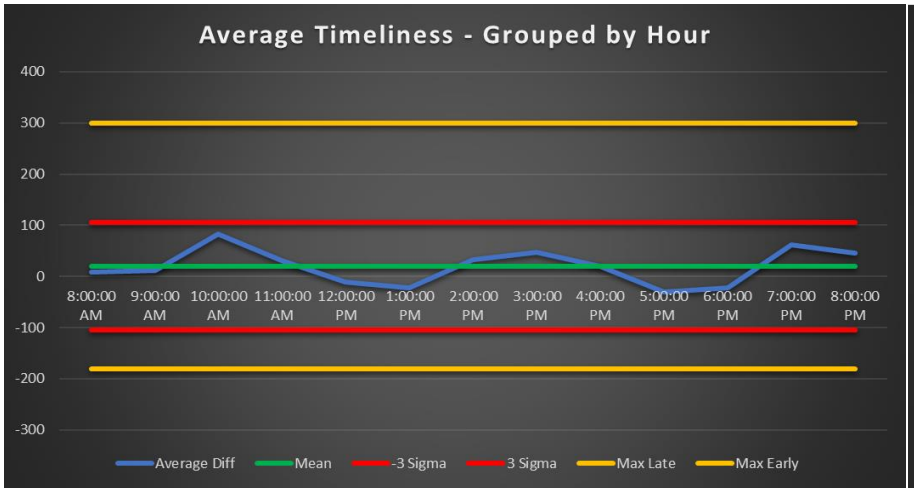


Figure 2: X-Bar chart for Route 17 timeliness, grouped by hourly increments

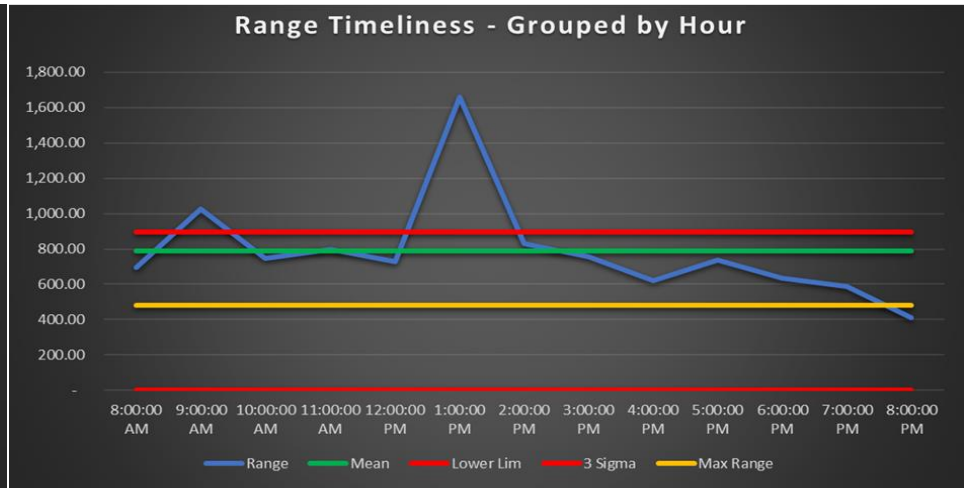


Figure 3: R chart for Route 17 timeliness, stops grouped by h**ourly increments.**

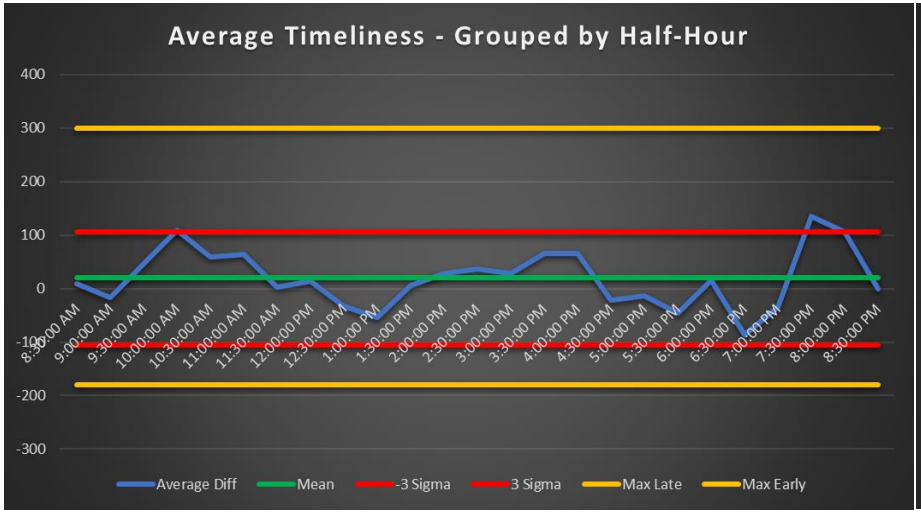


Figure 4: X-Bar chart for Route 17 timeliness, grouped by half-hour increments

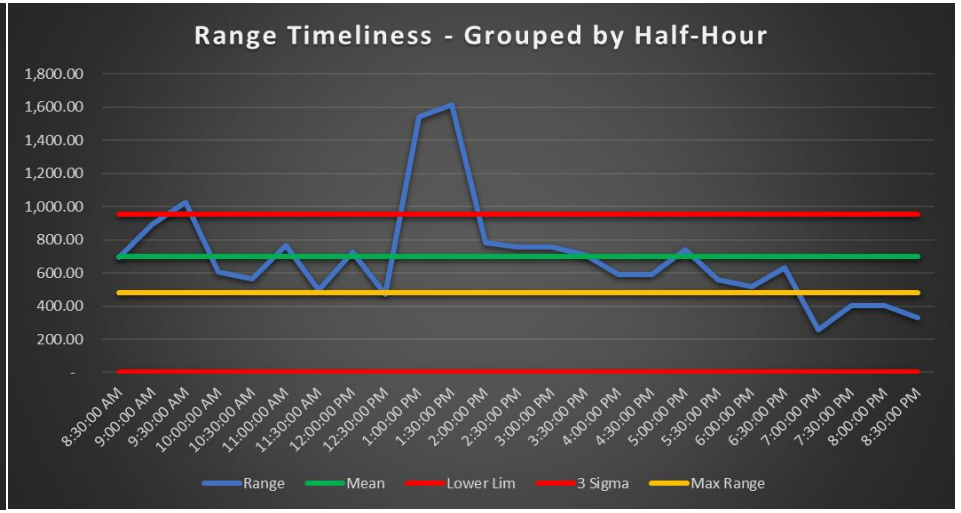


Figure 5: R chart for Route 17 timeliness, stops grouped by hourly increments.

To investigate further, the data was broken down into half-hour increments. The half-hour increments exhibit similar behavior, but with a higher degree of variation. The X-bar chart (Figure 4) demonstrates that the process does fall beyond the statistical limit but remains within Saskatoon Transit's goal time difference.

The half-hourly range chart (Figure 5) confirms the results from the hourly data that the process has a range mean above the desired range of Saskatoon Transit, and the process frequently falls outside of statistical control limits. For data tables, refer to Appendix D.

3.3 Stop Data

Because of the number of data points collected, and the dispersed nature of the data, the analysis by stop ID was performed in two different ways: Most popular stops, and sampled stops.

3.3.1 Most Popular Stops

The five most popular stops for Route 17 are Hunter Road and Manor Remple Cres. (id: 3715), Hunter Road, and Preston Ave. (id: 3716) Cornish Road and Lewin Way (id: 3719), Cornish Road and MacInnes (id: 5475), Cornish Road, and Willis Cres. (id: 5476), and Stonebridge Blvd. and Cornish Road (id: 5482). There was no valuable data gained from these stops, as they are all consecutive along the North of Hunter Road, and hence, all exhibited similar timeliness. It was strange to find that these stops had the highest number of pickups, as none of them are terminals. Because the terminals did not have the highest number of stops, which they should because busses must always stop at terminals, it lowers the confidence level of Saskatoon Transit's data collection and publishing.

3.3.2 Sampled Stops

Six stops were sampled at various points along Route 17 to generate a picture of how the bus's timeliness is affected through its route. The six sampled stops are: Hunter Road and Manor Remple Cres. – University (id: 3715), Taylor St. and Clarence Ave. – University (id: 4153), Aird St. and Clarence Ave. – University (id: 4139), 12th St. and Clarence Ave. – Stonebridge (id: 4145), Isabella St. and Clarence Ave – Stonebridge (id: 4145), Stonebridge Blvd. and Wellman Cres. – Stonebridge (id: 5626).

To piece the route together, as if one bus was driving it all day, a starting time of 08:54 at stop id 3715 was used and then the times of subsequent stops were plotted in Figure 6 (see

Appendix E for a sample of the data). When the data was arranged in this way, the statistical upper limit matches Saskatoon Transit's upper limit of +300 seconds. Several times, this route drops below Saskatoon Transit's lower limit, but remains in statistical control. No consistent process interferences were detecting by analysing the data in this way.

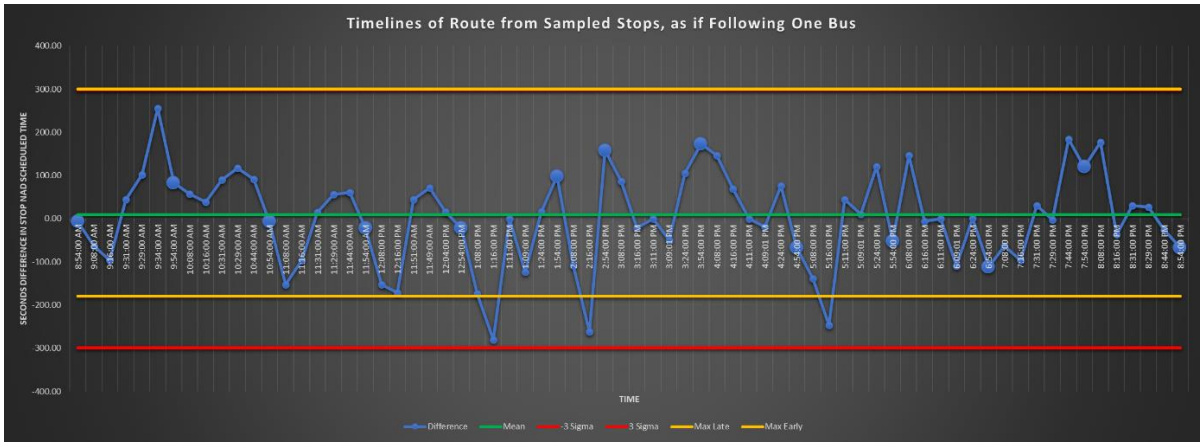


Figure 6: Timeliness of Route 17 from sampled stops

4 DISCUSSION – INFERENCE

From the data analysis conducted, it has been determined that the route 17 bus operations are an out of control process. The data analysis led to the realization that the bus usually runs late. By operating on a consistent late arrival basis, it can lead to upset and dissatisfied customers as they expect the bus to arrive at the time scheduled through the transit app. The lateness experienced by route 17 buses is not primarily caused by construction or route change; otherwise, it could be inferred that the bus would then arrive on time more frequently. Along with this, the analysis of hourly bins data showed similar issues with regards to the timeliness of bus operations. It showed the inconsistency of bus operations, which again is related to customer expectations and dissatisfaction with the service provided by Saskatoon Transit. Without consistency and consistent lateness of arrival, the customer will be more apt to use other transport methods, and this could lead to a financial impact on the operations of Saskatoon Transit.

Along with the consistent late arrivals of buses, some valuable insights were obtained from this data. For example, the data was also able to provide information related to the most popular stops on route 17. From the information gathered, there are five popular stops along the route, including Hunter Road and Manor Remple Cres., Hunter Road, and Preston Ave., Cornish Road and Lewin Way, Cornish Road and MacInnes, Cornish Road, and Willis Cres., and Stonebridge Blvd. and Cornish Road. The reason this is important is that although the

bus is consistently late, not all stops need to be improved to ensure accurate arrival time. Instead, the focus of the improvements should be made on finding ways to more accurately predict arrival time at these stops, as most users along route 17 will board a bus at these stops. Therefore, this can improve customer satisfaction for the majority of the route and also allow Saskatoon Transit to work towards achieving their targets stop time window and provide exceptional customer service to their patrons.

5 RECOMMENDATIONS – IMPROVE

5.1 Ineffective Data Collection

Their failings in data collection can be attested to the fact that not all their busses are currently GPS enabled. We ran a program that has the ability location and analyse the GPS feeds, on average 50-60% of the busses had a GPS location, which leads to our inference that the entire fleet is not GPS enabled. The GTFS Realtime API can put GPS positioning on actual stops, not just busses. Doing this can now allow for the bus and stop GPS's to interact with each other, when i bus physically passes the stop it will mark down the data. Currently the data can only be collected when the bus physically stops at the stop. This leads to ineffective data collection as it can leave glaring holes in the data set, where some busses only stop as low as 14 times during their trip, and others as much as 38 times (see Appendix F). Due to these factors we cannot get a consistent process to measure.

5.2 Lack of Technical Knowledge

The data has been collected through a server system, GTFS Realtime. This is the universal server that most transit systems world-wide use. This server has many features to aid in the operations of transit. The Saskatoon transit system does not utilize this service to its full capabilities. The main service that the transit system is neglecting to use is their 'Delay' service that the GTFS Realtime offers. This service will allow the transit app to relay valuable information to the user, it will be state 'this is the scheduled time; our bus is currently this far behind'. The improvements to data collection will lead to the ability to draw valuable conclusion, with the data presented and available to us, along with the data gathered and reported on, it is evident that we don't have there is a lack of technical knowledge in data collection.

5.3 Resistance

The Saskatoon Transit System does not manage the transit app, it is operated by a separate company. Since the app is universal in all transit system and linked up too many other transit

system servers, it's likely that the app will already have the architecture in place to implement the recommended additions. If the app can't accommodate the implementation of these processes, look in to building a Saskatoon Transit app. Due to the nature of the latter there will be a longer time to implement, but can warrant more customizability. You can utilize local tech incubators for local talent while developing and implementing the application. Since the Saskatoon Transit system has close contacts with the Regina and Prince Albert systems, you can coordinate with them for Saskatchewan wide transit app.

6 CONCLUSION

- Managing customer expectations
    - Can't improve process right now but can improve your customer expectation right now
- Ineffective data collection
    - This due to the stuff talked about in 5.1
- Inconclusive data leading to the conclusion
- get better data collection methods

REFERENCES

Apache POI. (2017, September 8). *Apache POI >> 3.17.* Retrieved March 5, 2020, from MVN Repository: https://mvnrepository.com/artifact/org.apache.poi/poi/3.17

Google. (2019, July 3). *GTFS realtime overview*. Retrieved March 2, 2020, from Google Transit APIs > Realtime Transit: https://developers.google.com/transit/gtfs-realtime/

Google. (n.d.). *GTFS Realtime reference*. Retrieved March 3, 2020, from Google Transit APIs > Realtime Transit: https://developers.google.com/transit/gtfs-realtime/reference

Google. (n.d.). *Protocol buffers*. Retrieved March 2, 2020, from Googel Developers: https://developers.google.com/protocol-buffers/

Google. (n.d.). *protocolbuffers / protobuf.* Retrieved March 3, 2020, from GitHub: https://github.com/protocolbuffers/protobuf

Moellenbeck, M. (2020, March 2). Saskatoon Transit COC - Tour. (M. Buglass, L. Cowan, & J. Jungwirth, Interviewers)

Open data from Saskatoon Transit - Access trip updates here. (2020, March 6). Retrieved March 6, 2020, from Saskatoon Transit: http://apps2.saskatoon.ca/app/data/TripUpdate/TripUpdates.pb

Open data from Saskatoon Transit - Access vehicle positions information here. (2020, March 3). Retrieved March 3, 2020, from Saskatoon Transit: http://apps2.saskatoon.ca/app/data/Vehicle/VehiclePositions.pb

Saskatoon Transit. (2020). *Open data from Saskatoon Transit*. Retrieved March 2, 2020, from Saskatoon Transit: https://transit.saskatoon.ca/about-us/open-data-saskatoon-transit

Saskatoon Transit. (n.d.). *Route 17 Stonebridge University.* Retrieved March 6, 2020, from Transit Sakstoon: https://transit.saskatoon.ca/sites/default/files/route_17_2020_covid.pdf

Transit App, Inc. (2020). *Transit • Bus & Train Times*. Retrieved April 1, 2020, from App Store Preview: https://apps.apple.com/ca/app/transit-bus-train-times/id498151501

APPENDIX A – ORIGINAL SOFTWARE FOR DATA EXTRACTION

All code is the intellectual property of Matthew Buglass. No replication or use of the following source code is permitted without written consent of Mr. Matthew Buglass.

LinkedTransitNode

```
package AuxStructures;

public class LinkedTransitNode {
    String tripID, stopID;
    String timeStamp;

    LinkedTransitNode next;

    public LinkedTransitNode(String stopID, String tripID, String timeStamp) {
        this.stopID = stopID;
        this.tripID = tripID;
        this.timeStamp = timeStamp;
    }

    /**
     * Compares the contents of the object
     *
     * @return 0 if the contents are identical of the time difference is less than
60seconds, 1 otherwise
     */
    public int compareTo(LinkedTransitNode o) {
        // check if all fields are identical
        if (this.getTimeStamp().compareTo(o.getTimeStamp()) == 0
                && this.getStopID().compareTo(o.getStopID()) == 0
                && this.getTripID().compareTo(o.getTripID()) == 0) {
            return 0;
        }
        else {
            // if the stop and route are the same, and the time difference is less
than 60sec, they are essentially the same
            if (Math.abs(Float.parseFloat(this.timeStamp) -
Float.parseFloat(o.timeStamp)) < 60000
                    && this.getStopID().compareTo(o.getStopID()) == 0
                    && this.getTripID().compareTo(o.getTripID()) == 0) {
                return 0;
            }

            return 1;
        }
    }


    public String getTimeStamp() {
        return timeStamp;
    }

    public String getTripID() {
        return tripID;
```

```
    }

    public void setTimeStamp(String timeStamp) {
        this.timeStamp = timeStamp;
    }

    public void setTripID(String tripID) {
        this.tripID = tripID;
    }

    public String getStopID() {
        return stopID;
    }

    public void setStopID(String stopID) {
        this.stopID = stopID;
    }

    public void setNext(LinkedTransitNode next) {
        this.next = next;
    }

    public LinkedTransitNode getNext() {
        return next;
    }

    public String toString() {
        return "Trip ID " + tripID + " arrived at stop " + stopID + " at " +
timeStamp + ".\n";
    }
}
```

LinkedTransitList

```
package AuxStructures;

import java.util.LinkedList;

public class LinkedTransitList extends LinkedList<LinkedTransitNode> {

    public boolean contains(LinkedTransitNode n) {
        for (LinkedTransitNode o : this) {
            if (o.compareTo(n) == 0) {
                return true;
            }
        }

        return false;
    }
}
```

DataExtraction – Main Program

```
import java.io.IOException;
import java.net.URL;

import AuxStructures.LinkedTransitList;
import AuxStructures.LinkedTransitNode;

import com.google.transit.realtime.GtfsRealtime;
import com.google.transit.realtime.GtfsRealtime.FeedEntity;
import com.google.transit.realtime.GtfsRealtime.FeedMessage;

import java.io.File;
import java.io.FileOutputStream;

import java.util.Map;
import java.util.Set;
import java.util.TreeMap;

import org.apache.poi.ss.usermodel.Cell;
import org.apache.poi.xssf.usermodel.XSSFRow;
import org.apache.poi.xssf.usermodel.XSSFSheet;
import org.apache.poi.xssf.usermodel.XSSFWorkbook;

public class DataExtraction {


    public static void main(String[] args) throws Exception {
        LinkedTransitList ll = new LinkedTransitList();
        // March 6th, 2020 20:00
        long stopTime = 1583550000;

        // Have it run for a number of seconds
        long practiceTime = System.currentTimeMillis()/1000 + 30;

        while (System.currentTimeMillis()/1000 < stopTime) {
            URL url = new
URL("http://apps2.saskatoon.ca/app/data/TripUpdate/TripUpdates.pb");
            try {
                FeedMessage feed = FeedMessage.parseFrom(url.openStream());
                long currentTime = System.currentTimeMillis() / 1000;

                for (FeedEntity entity : feed.getEntityList()) {
                    if
(entity.getTripUpdate().getTrip().getRouteId().compareTo("10816") == 0) {
                        GtfsRealtime.TripUpdate update = entity.getTripUpdate();
                        for (GtfsRealtime.TripUpdate.StopTimeUpdate su :
update.getStopTimeUpdateList()) {

                            if (su.getArrival().getTime() < currentTime) {
                                LinkedTransitNode node = new
LinkedTransitNode(su.getStopId(), update.getTrip().getTripId(),
                                        "" + su.getArrival().getTime());
                                if (!ll.contains(node)) {
                                    ll.addLast(node);
                                }
                            }
                        }
                    }
```

```
                }
            }
            catch (IOException e) {
                System.out.println(e.getLocalizedMessage());
            }
        }

        System.out.println(ll.toString());

        //Create blank workbook
        XSSFWorkbook workbook = new XSSFWorkbook();

        //Create a blank sheet
        XSSFSheet spreadsheet = workbook.createSheet( "Route 17 info");

        //Create row object
        XSSFRow row;

        //This data needs to be written (Object[])
        Map < String, Object[] > empinfo = new TreeMap < String, Object[] >();
        empinfo.put( "1", new Object[] {
                "Trip ID", "Stop ID", "Time Stamp" });

        int rowNum = 2;

        // Add all the node's data
        for (LinkedTransitNode node : ll) {
            empinfo.put(String.valueOf(rowNum), new Object[]{
                    node.getTripID(), node.getStopID(), node.getTimeStamp()});
            rowNum ++;
        }

        //Iterate over data and write to sheet
        Set < String > keyid = empinfo.keySet();
        int rowid = 0;

        for (String key : keyid) {
            row = spreadsheet.createRow(rowid++);
            Object [] objectArr = empinfo.get(key);
            int cellid = 0;

            for (Object obj : objectArr){
                Cell cell = row.createCell(cellid++);
                cell.setCellValue((String)obj);
            }
        }
        //Write the workbook in file system
        FileOutputStream out = new FileOutputStream(
                new File("./../../../../Route17Data.xlsx"));

        workbook.write(out);
        out.close();
        System.out.println("Route17Data.xlsx written successfully");
    }

}
```

APPENDIX B – SAMPLE OF TIMELINESS DATA

Sample of the 2,425 data point collected on March 6th, 2020 from 08:30 to 21:00.

Trip ID	Stop ID	Time Stamp	Stop Time	Scheduled Time	Difference (sec)
1064471	5531	1583505522	8:38:42 AM	8:31:00 AM	462.00
1064471	5674	1583505604	8:40:04 AM	8:33:00 AM	424.00
1064563	3715	1583505233	8:33:53 AM	8:34:00 AM	-7.00
1064471	5626	1583505428	8:37:08 AM	8:34:00 AM	188.00
1064563	3716	1583505294	8:34:54 AM	8:35:00 AM	-6.00
1064563	3719	1583505336	8:35:36 AM	8:35:00 AM	36.00
1064471	5483	1583505445	8:37:25 AM	8:35:00 AM	145.00
1064522	4138	1583505275	8:34:35 AM	8:36:00 AM	-85.00
1064522	4144	1583505306	8:35:06 AM	8:36:00 AM	-54.00
1064563	5475	1583505394	8:36:34 AM	8:36:00 AM	34.00
1064522	4146	1583505346	8:35:46 AM	8:37:00 AM	-74.00
1064563	5476	1583505449	8:37:29 AM	8:37:00 AM	29.00
1064471	5532	1583505474	8:37:54 AM	8:37:00 AM	54.00
1064471	5554	1583505705	8:41:45 AM	8:37:00 AM	285.00
1064471	5675	1583505574	8:39:34 AM	8:37:00 AM	154.00
1064509	5678	1583505400	8:36:40 AM	8:37:00 AM	-20.00
1064522	4130	1583505485	8:38:05 AM	8:38:00 AM	5.00
1064522	4153	1583505387	8:36:27 AM	8:38:00 AM	-93.00
1064563	5482	1583505508	8:38:28 AM	8:38:00 AM	28.00
1064509	5676	1583505413	8:36:53 AM	8:38:00 AM	-67.00
1064522	4128	1583505441	8:37:21 AM	8:39:00 AM	-99.00
1064470	4142	1583505607	8:40:07 AM	8:39:00 AM	67.00
1064470	4154	1583505522	8:38:42 AM	8:39:00 AM	-18.00
1064470	4155	1583505559	8:39:19 AM	8:39:00 AM	19.00
1064470	4163	1583505397	8:36:37 AM	8:39:00 AM	-143.00
1064471	5520	1583505677	8:41:17 AM	8:39:00 AM	137.00
1064509	5680	1583505456	8:37:36 AM	8:39:00 AM	-84.00
1064563	3465	1583505615	8:40:15 AM	8:40:00 AM	15.00
1064470	4123	1583505630	8:40:30 AM	8:40:00 AM	30.00
1064470	4125	1583505634	8:40:34 AM	8:40:00 AM	34.00
1064522	4133	1583505530	8:38:50 AM	8:40:00 AM	-70.00
1064470	4120	1583505660	8:41:00 AM	8:41:00 AM	0.00

APPENDIX C – AGGREGATE DATA

Statistics

Summary				
Mean	16.83		StdDev	133.65
Range	1,659.00		Max Late	300
Count	2,425.00		Max Early	-180

Significantly Different from 0

X-bar	16.83		t-crit	1.96
u	0		t-calc	6.20
s	133.65		Alpha	0.05
n	2,425.00			
df	2,424.00			

Reject the null hypothesis. The bus system is not on time.

$$\begin{cases} H_0: \bar{x} = \mu \\ H_1: \bar{x} \neq \mu \end{cases}$$

Significantly under 5 mins late

X-bar	16.83		t-crit	1.645
u	300		t-calc	- 104.34
s	133.65		Alpha	0.05
n	2,425.00			
df	2,424.00			

Do not reject the null hypothesis. The bus is statistically less that 5 mins late.

$$\begin{cases} H_0: \bar{x} \leq \mu \\ H_1: \bar{x} > \mu \end{cases}$$

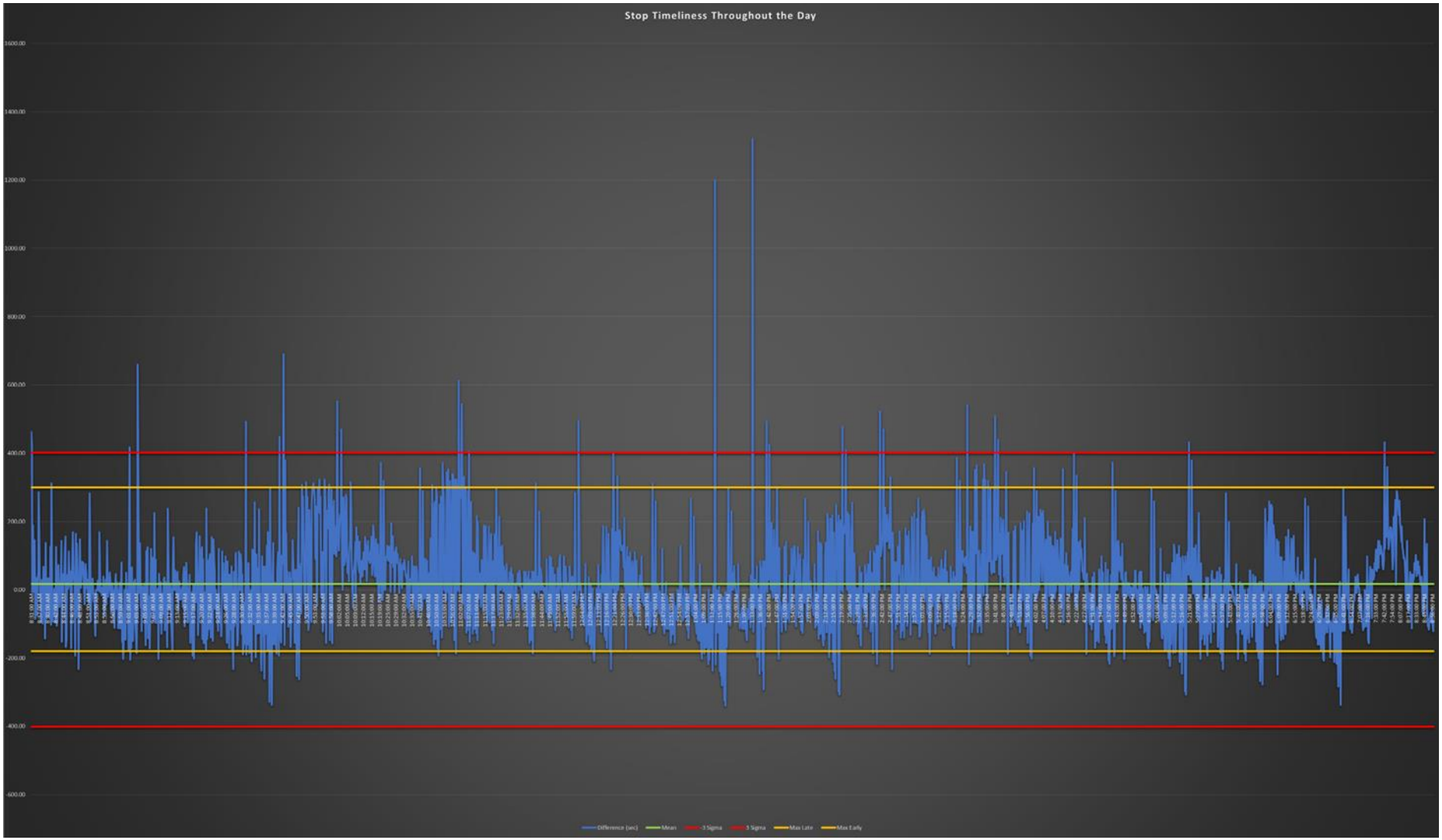
Significantly over 3 min early

X-bar	16.83		t-crit	-1.645
u	-180		t-calc	72.52
s	133.65		Alpha	0.05
n	2,425.00			
df	2,424.00			

Do not reject the null hypothesis. The bus is statistically less that 3 mins early.

$$\begin{cases} H_0: \bar{x} \geq \mu \\ H_1: \bar{x} < \mu \end{cases}$$

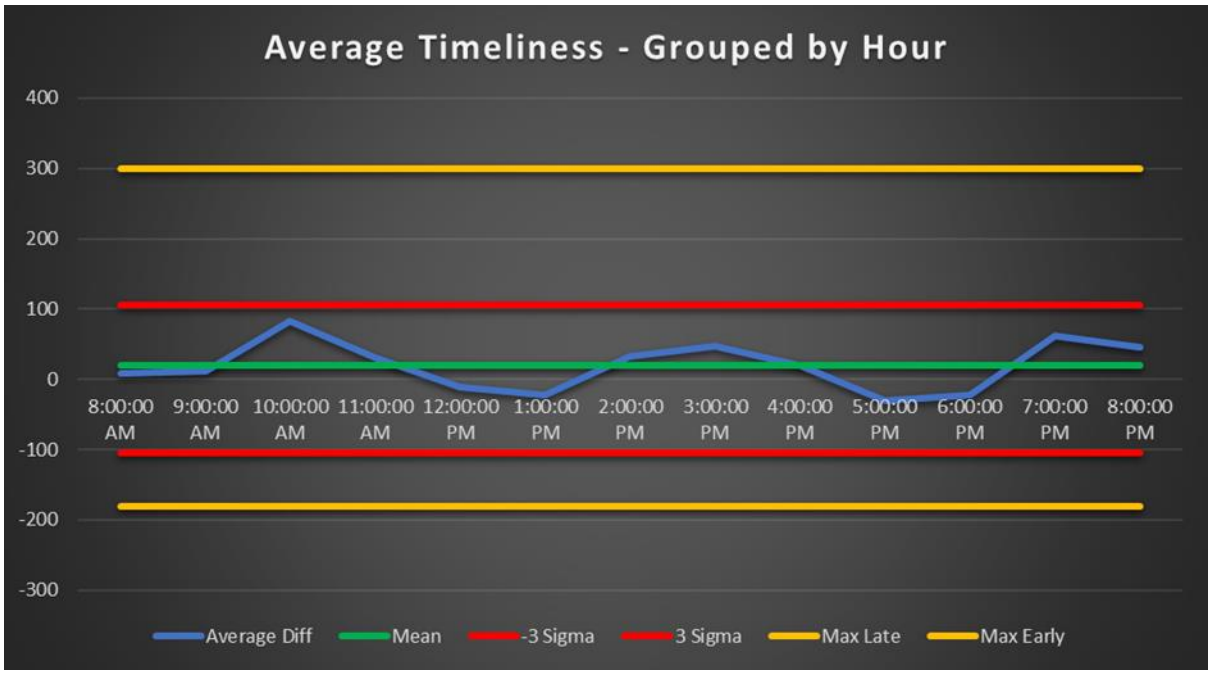
Chart


Stop Timeliness Throughout the Day
1600.00
1400.00
1200.00
1000.00
800.00
600.00
400.00
200.00
0.00
-200.00
-400.00
-600.00
Difference (sec)
Mean
-3 Sigma
3 Sigma
Max Late
Max Early

APPENDIX D – STOP TIMELINESS, TIME BIN DATA

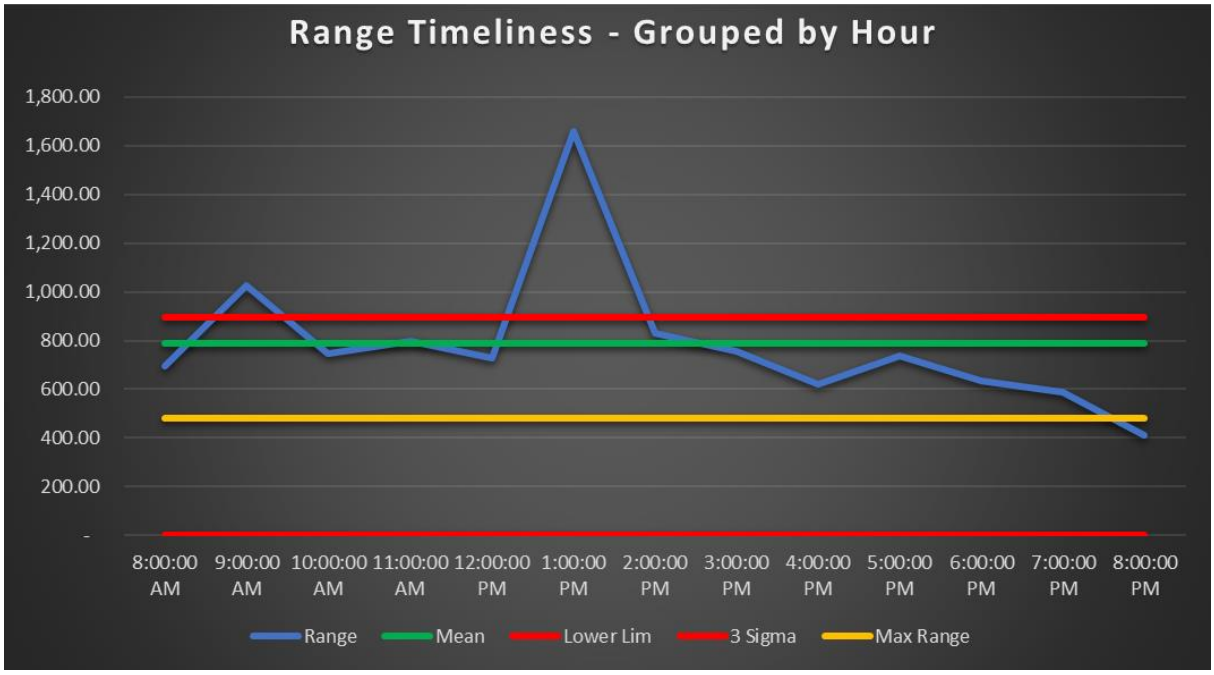
X-Bar data grouped by hour

Time	Average Diff	Mean	-3 Sigma	-2 Sigma	-1 Sigma	1 Sigma	2 Sigma	3 Sigma	Max Late	Max Early
8:00:00 AM	9.149068357	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
9:00:00 AM	11.40761843	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
10:00:00 AM	83.27089599	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
11:00:00 AM	30.96357368	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
12:00:00 PM	-10.46480174	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
1:00:00 PM	-22.66847981	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
2:00:00 PM	32.95680731	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
3:00:00 PM	47.80902099	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
4:00:00 PM	20.33214878	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
5:00:00 PM	-30.63254415	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
6:00:00 PM	-22.95430002	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
7:00:00 PM	62.04378292	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
8:00:00 PM	45.15022343	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180



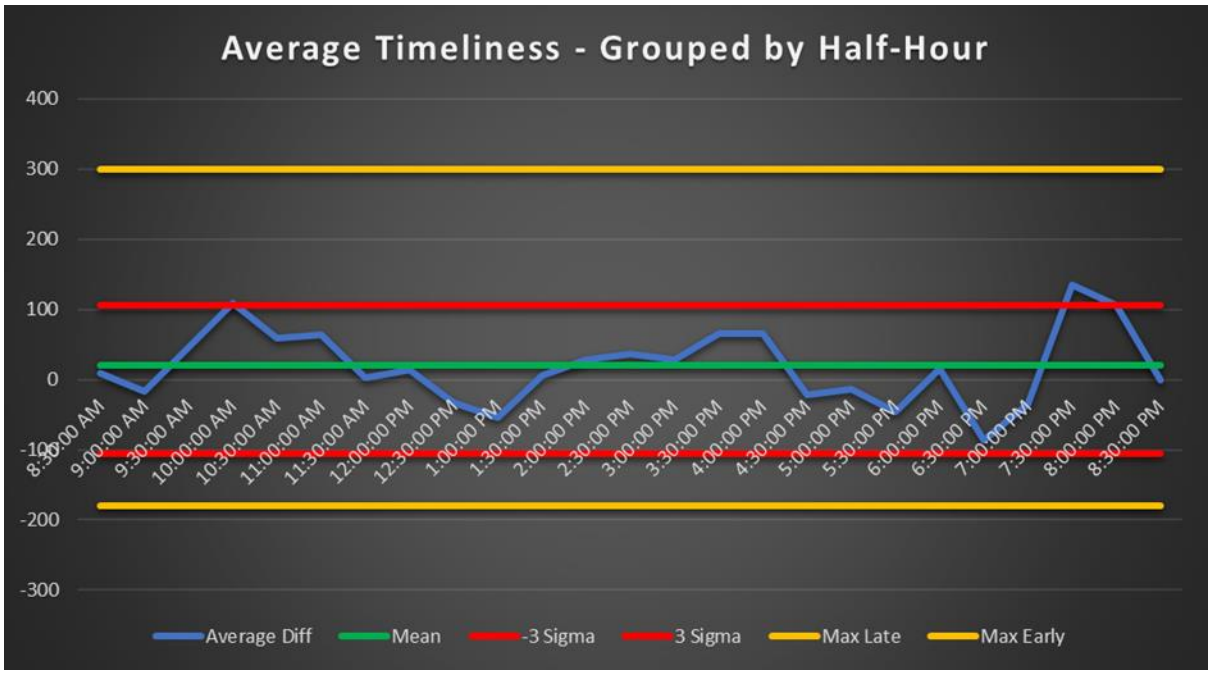
R data grouped by hour

Time	Range	Mean	Lower Lim	3 Sigma	Max Range	Max Early
8:00:00 AM	694.00	786.85	-	895.82	480	-180
9:00:00 AM	1,027.00	786.85	-	895.82	480	-180
10:00:00 AM	745.00	786.85	-	895.82	480	-180
11:00:00 AM	799.00	786.85	-	895.82	480	-180
12:00:00 PM	727.00	786.85	-	895.82	480	-180
1:00:00 PM	1,659.00	786.85	-	895.82	480	-180
2:00:00 PM	829.00	786.85	-	895.82	480	-180
3:00:00 PM	757.91	786.85	-	895.82	480	-180
4:00:00 PM	619.00	786.85	-	895.82	480	-180
5:00:00 PM	739.00	786.85	-	895.82	480	-180
6:00:00 PM	634.00	786.85	-	895.82	480	-180
7:00:00 PM	589.06	786.85	-	895.82	480	-180
8:00:00 PM	410.00	786.85	-	895.82	480	-180



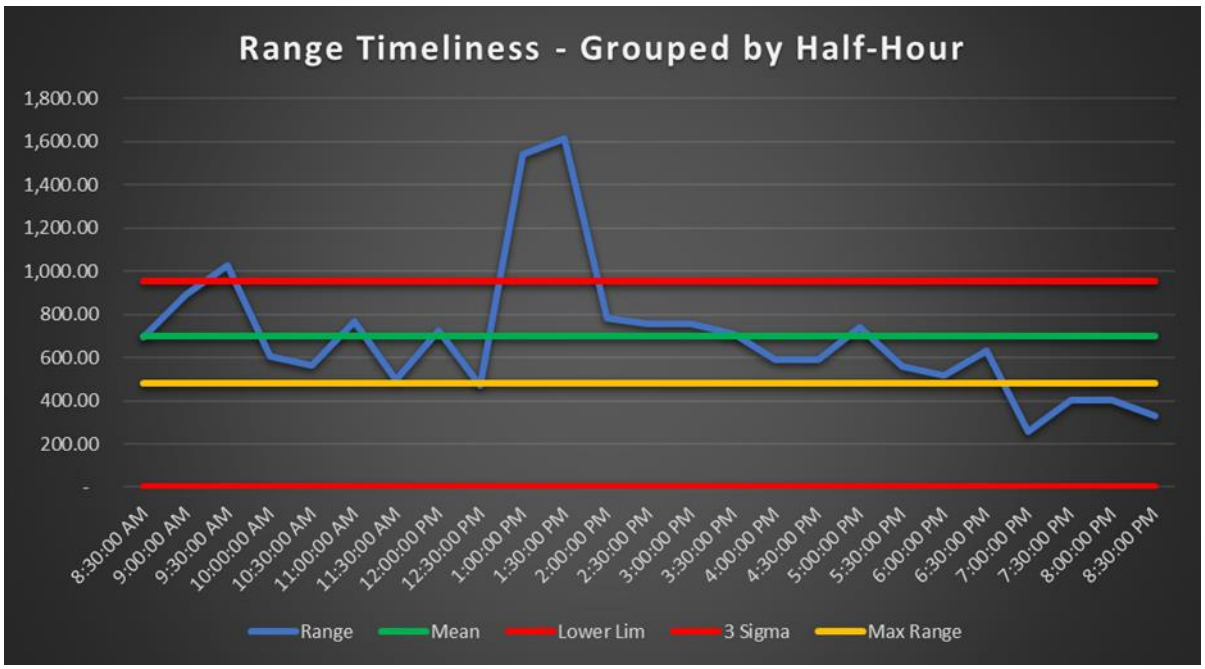
X-Bar data grouped by half-hour

Time	Average Diff	Mean	-3 Sigma	-2 Sigma	-1 Sigma	1 Sigma	2 Sigma	3 Sigma	Max Late	Max Early
8:30:00 AM	9.149068357	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
9:00:00 AM	-17.27918368	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
9:30:00 AM	46.7173271	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
10:00:00 AM	108.5437499	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
10:30:00 AM	58.4829043	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
11:00:00 AM	63.71907431	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
11:30:00 AM	3.060739819	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
12:00:00 PM	14.01747157	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
12:30:00 PM	-32.61542997	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
1:00:00 PM	-54.68331867	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
1:30:00 PM	5.382617093	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
2:00:00 PM	28.85407073	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
2:30:00 PM	36.5861512	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
3:00:00 PM	28.05299431	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
3:30:00 PM	65.68352131	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
4:00:00 PM	64.68897325	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
4:30:00 PM	-21.46562813	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
5:00:00 PM	-12.80651398	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
5:30:00 PM	-46.57312882	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
6:00:00 PM	15.8091659	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
6:30:00 PM	-86.68406601	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
7:00:00 PM	-34.46666852	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
7:30:00 PM	134.4266215	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
8:00:00 PM	105.5517233	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180
8:30:00 PM	-0.945658056	19.72	- 105.40	- 70.27	- 35.13	35.13	70.27	105.40	300	-180



R data grouped by half-hour

Time	Range	Mean	-3 Sigma	3 Sigma	Max Late	Max Early
8:30:00 AM	694.00	696.63	-	952.20	480	-180
9:00:00 AM	892.00	696.63	-	952.20	480	-180
9:30:00 AM	1,027.00	696.63	-	952.20	480	-180
10:00:00 AM	604.00	696.63	-	952.20	480	-180
10:30:00 AM	565.00	696.63	-	952.20	480	-180
11:00:00 AM	769.00	696.63	-	952.20	480	-180
11:30:00 AM	499.00	696.63	-	952.20	480	-180
12:00:00 PM	727.00	696.63	-	952.20	480	-180
12:30:00 PM	469.00	696.63	-	952.20	480	-180
1:00:00 PM	1,539.00	696.63	-	952.20	480	-180
1:30:00 PM	1,612.00	696.63	-	952.20	480	-180
2:00:00 PM	784.00	696.63	-	952.20	480	-180
2:30:00 PM	754.94	696.63	-	952.20	480	-180
3:00:00 PM	757.91	696.63	-	952.20	480	-180
3:30:00 PM	709.00	696.63	-	952.20	480	-180
4:00:00 PM	589.91	696.63	-	952.20	480	-180
4:30:00 PM	589.00	696.63	-	952.20	480	-180
5:00:00 PM	739.00	696.63	-	952.20	480	-180
5:30:00 PM	559.00	696.63	-	952.20	480	-180
6:00:00 PM	514.88	696.63	-	952.20	480	-180
6:30:00 PM	634.00	696.63	-	952.20	480	-180
7:00:00 PM	255.00	696.63	-	952.20	480	-180
7:30:00 PM	402.06	696.63	-	952.20	480	-180
8:00:00 PM	402.00	696.63	-	952.20	480	-180
8:30:00 PM	328.07	696.63	-	952.20	480	-180



APPENDIX E – ROUTE PROGRESSION THROUGH THE DAY (SAMPLE)

Stop ID	Scheduled time	Difference	Mean	-3 Sigma	3 Sigma	Max Late	Max Early
3715	8:54:00 AM	-7.00	9.89	-299.19	299.19	300.00	-180.00
4153	9:08:00 AM	-63.00	9.89	-299.19	299.19	300.00	-180.00
4139	9:16:00 AM	-96.00	9.89	-299.19	299.19	300.00	-180.00
4120	9:31:00 AM	45.00	9.89	-299.19	299.19	300.00	-180.00
4154	9:29:00 AM	102.05	9.89	-299.19	299.19	300.00	-180.00
5626	9:34:00 AM	256.00	9.89	-299.19	299.19	300.00	-180.00
3715	9:54:00 AM	83.00	9.89	-299.19	299.19	300.00	-180.00
4153	10:08:00 AM	57.00	9.89	-299.19	299.19	300.00	-180.00
4139	10:16:00 AM	39.00	9.89	-299.19	299.19	300.00	-180.00
4120	10:31:00 AM	90.00	9.89	-299.19	299.19	300.00	-180.00
4154	10:29:00 AM	116.90	9.89	-299.19	299.19	300.00	-180.00
5626	10:44:00 AM	91.01	9.89	-299.19	299.19	300.00	-180.00
3715	10:54:00 AM	-7.00	9.89	-299.19	299.19	300.00	-180.00
4153	11:08:00 AM	-153.00	9.89	-299.19	299.19	300.00	-180.00
4139	11:16:00 AM	-96.00	9.89	-299.19	299.19	300.00	-180.00
4120	11:31:00 AM	15.00	9.89	-299.19	299.19	300.00	-180.00
4154	11:29:00 AM	56.75	9.89	-299.19	299.19	300.00	-180.00
5626	11:44:00 AM	61.01	9.89	-299.19	299.19	300.00	-180.00
3715	11:54:00 AM	-22.00	9.89	-299.19	299.19	300.00	-180.00
4153	12:08:00 PM	-153.00	9.89	-299.19	299.19	300.00	-180.00
4139	12:16:00 PM	-171.00	9.89	-299.19	299.19	300.00	-180.00
4120	11:51:00 AM	45.00	9.89	-299.19	299.19	300.00	-180.00
4154	11:49:00 AM	71.70	9.89	-299.19	299.19	300.00	-180.00
5626	12:04:00 PM	16.01	9.89	-299.19	299.19	300.00	-180.00
3715	12:54:00 PM	-22.00	9.89	-299.19	299.19	300.00	-180.00
4153	1:08:00 PM	-174.00	9.89	-299.19	299.19	300.00	-180.00
4139	1:16:00 PM	-280.00	9.89	-299.19	299.19	300.00	-180.00
4120	1:11:00 PM	0.00	9.89	-299.19	299.19	300.00	-180.00
4154	1:09:00 PM	-123.50	9.89	-299.19	299.19	300.00	-180.00
5626	1:24:00 PM	16.00	9.89	-299.19	299.19	300.00	-180.00
3715	1:54:00 PM	98.00	9.89	-299.19	299.19	300.00	-180.00
4153	2:08:00 PM	-108.00	9.89	-299.19	299.19	300.00	-180.00
4139	2:16:00 PM	-261.00	9.89	-299.19	299.19	300.00	-180.00
3715	2:54:00 PM	158.00	9.89	-299.19	299.19	300.00	-180.00
4153	3:08:00 PM	87.00	9.89	-299.19	299.19	300.00	-180.00
4139	3:16:00 PM	-21.00	9.89	-299.19	299.19	300.00	-180.00
4120	3:11:00 PM	0.00	9.89	-299.19	299.19	300.00	-180.00
4154	3:09:01 PM	-48.80	9.89	-299.19	299.19	300.00	-180.00
5626	3:24:00 PM	106.00	9.89	-299.19	299.19	300.00	-180.00
3715	3:54:00 PM	173.00	9.89	-299.19	299.19	300.00	-180.00
4153	4:08:00 PM	147.00	9.89	-299.19	299.19	300.00	-180.00
4139	4:16:00 PM	69.00	9.89	-299.19	299.19	300.00	-180.00
4120	4:11:00 PM	0.00	9.89	-299.19	299.19	300.00	-180.00
4154	4:09:01 PM	-18.95	9.89	-299.19	299.19	300.00	-180.00
5626	4:24:00 PM	76.00	9.89	-299.19	299.19	300.00	-180.00
3715	4:54:00 PM	-67.00	9.89	-299.19	299.19	300.00	-180.00
4153	5:08:00 PM	-138.00	9.89	-299.19	299.19	300.00	-180.00
4139	5:16:00 PM	-246.00	9.89	-299.19	299.19	300.00	-180.00
4120	5:11:00 PM	45.00	9.89	-299.19	299.19	300.00	-180.00
4154	5:09:01 PM	10.90	9.89	-299.19	299.19	300.00	-180.00
5626	5:24:00 PM	121.00	9.89	-299.19	299.19	300.00	-180.00
3715	5:54:00 PM	-52.00	9.89	-299.19	299.19	300.00	-180.00
4153	6:08:00 PM	147.00	9.89	-299.19	299.19	300.00	-180.00
4139	6:16:00 PM	-6.00	9.89	-299.19	299.19	300.00	-180.00
4120	6:11:00 PM	0.00	9.89	-299.19	299.19	300.00	-180.00

APPENDIX F – NUMBER OF DATA POINTS BY STOP ID