An analysis of SEPTA Regional Rail On-Time Performance

Presented by Will Entriken
25 January 2009

Executive Overview

Data was collected from the SEPTA website and analyzed over the past three months. The collected data provides visibility into which trains are late and why. This report shows how to change the train schedules to improve On-Time Performance and it provides a recommended schedule for each rail line. Also, some statistics are given for various rail lines, which highlight trains that require attention. The recommended changes could improve On-Time Performance from 86% to over 90%. All data, statistics, and updated schedules in this report were collected automatically using a computer program which can run on a standard web server. The methods in this report could also be applied to other SEPTA services.

Introduction

A computer program was written which can run on a standard LAMP web server. This program automatically collects data from the SEPTA website and automatically generates statistics based on Regional Rail performance.

Some statistical definitions and methods are shown. Following that is a description of output statistics, and results and conclusions.

To make meaningful statistics, we need at least 30 data points. Each day is a data point. All data used in this report is from 13 November 2008 until 10 January 2009, which is sufficient for this report.

This report only discusses weekday train performance. Weekend performance could be analyzed but that would require more data to be collected. Specifically we would need 15 weeks for a Saturday/Sunday schedule and 30 weeks for a schedule that only applies on one day.

Definitions

On-Time Performance (OTP) is defined as the percentage of trains on a line which are less than or equal to 5:59 late in reaching their destination. Trains are said to be traveling inbound until they reach Suburban Station, then they are traveling outbound to a different destination. Inbound trains' destination are Suburban Station, for OTP purposes.

A new metric, Modified On-Time Performance (MOTP), is defined. This is the percentage of times a train is less than or equal to 5:59 late in reaching a stop, for all stops on a line.

For clarity, the lines R7 and R8 are broken up into R7 Chestnut Hill East, R7 Trenton, R8 Chestnut Hill West, and R8 Fox Chase.

Lateness is defined as the delay experienced by passengers for a given stop. This can be expressed as an average lateness, or can be express as a standard deviation.

Percentiles provide information on a set of numbers. To make a statement of "the 10th percentile lateness of a stop is 2.3 minutes" is to specifically mean "lateness is less than 2.3 minutes for less than 10% of the time." Here, this also specifically means "lateness is at least 2.3 minutes for greater than 90% of the time."

Method

Train lateness data is collected from the SEPTA train view website. Also, schedules are collected from the Regional Rail website. These data and schedules are automatically interpreted by the software. Several sets of statistics are output by the software, and some of these tables are attached to this report.

Rail OTP Breakdown

For each rail, the OTP may be expressed as a percentage. It may also be expressed as an absolute number. For example: during the report period, the R1 Airport Line traveling inbound was late 139 times. This table breaks down that lateness by train. For example: 20 of those latenesses were caused by the 9152 train. In this case, over half of the latenesses are caused by only 8 trains. These tables can be used to focus on trains that are damaging the OTP.

Average Lateness

Average lateness is shown for each stop of each train on a given line. This table is color coded for clarity.

Recommended Schedule

A recommended schedule is included. This is generated by taking the existing schedule and delaying some of the stops. Stops are recommended to be delayed when a train does not make that stop on time for a vast majority of the time. Specifically, each recommended time is set to the 10th percentile (rounded down) of the time the train actually arrives at that stop. In other words, any recommended schedule change may cause a train to delay one or two times a month. But for all the other days, the train would be closer to the scheduled time.

This 10th percentile was chosen to have a minimal risk of negative effectives on OTP and MOTP. Each schedule shows the current OTP and MOTP, as well as the estimated OTP and MOTP if the proposed schedule were accepted.

Each recommended change could be accepted or rejected independently and still have the desired effect.

Standard Deviation Lateness

A table showing the standard deviation lateness for each stop. This will indicate a train that arrives at a stop with wildly varying times.

Percentile Lateness

A table showing the 10 percentile (or any other percentile) lateness for all stops.

Results

Outputs from the program are included at the end of this report:

- Average lateness for all stops on all lines
- OTP breakdown for R1 and R2, both directions
- Recommended schedule changes for all lines (shows original and proposed OTP/MOTP)

Tables showing standard deviation and percentile lateness are not included.

Conclusions

The OTP during the report period was 86%. Making the very conservative changes recommended in this report are estimated to increase this to 90%. A percentile higher than 10% could be used, or attention could be given to the top one train listed in each OTP breakdown. These could increase overall OTP to 93%.