1. Recommended Course of Action

The result of this program project is a QGIS plugin that can be used to identify trailheads within a specified distance of transit stops. Project sponsors, Turn Off The App, Go Outside (TOTAGO) and The Mountains to Sound Greenway Trust, are able to utilize this data to identify trips to natural areas that utilize transit options. The success of this plugin depends on the quality of the available data. We recommend that the project sponsors acquire data in target markets and administer the plugin to generate a list of potential trailhead locations. These can be further investigated for use in the TOTAGO app. Trailhead data for the contiguous United States were provided as part of the project deliverables. These data were generated through OpenStreetMap (OSM) exports and likely include a margin of error. The quality of the plugin exports will be improved by the use of more accurate and detailed trailhead information. Acquisition of any augmented data should be another priority for project sponsors. A final element for future implementation is to incorporate the use of route distance rather than Euclidean distance in determining how users will get from the transit stop to the trailhead. Collection of route data, including slope, type of path, and distance would prove useful, especially as additional sidewalk information is added to OSM data. We recommend the plugin be used to identify appropriate trailheads for the TOTAGO app, while also acquiring detailed data to improve upon results.
Table of Contents
1. Recommended Course of Action........................................................................................................1
List of Tables ........................................................................................................................................2
List of Figures .....................................................................................................................................2
2. Introduction ......................................................................................................................................3
3. Design & Methods ............................................................................................................................4
   3.1: Project organization ....................................................................................................................4
   3.2: Project Planning ..........................................................................................................................5
   3.3: Project organization ....................................................................................................................5
   3.4: Project development ...................................................................................................................6
      Part 1 – Identification of trailheads that are considered transit accessible.................................6
      Part 2 – Summary of transit service data for stops near trailheads ..............................................6
      Part 3 – Calculation of distance between stops and trailheads......................................................8
4. Results ...........................................................................................................................................8
5. Discussion ......................................................................................................................................12
Literature Cited .................................................................................................................................14

List of Tables
Table 1: Social-Economic Systems Table............................................................................................3
Table 2: Threshold Matrix ....................................................................................................................4
Table 3: Files that comprise a GTFS ....................................................................................................7

List of Figures
Figure 1: Plugin Dialog Box..................................................................................................................9
Figure 2: Example Plugin Output ..........................................................................................................10
2. Introduction

The project sponsors, Turn Off The App, Go Outside (TOTAGO), and The Mountains to Sound Greenway Trust have been working together to increase accessibility to the outdoors and motivate people to explore their surroundings. Their collaboration extends between social and technical components to connect users through a web application. The application is designed to present transit options to access nearby trails. Research has shown that one of the primary barriers for youth accessing the outdoors is lack of transportation (Outdoor Nation, n.d.). It provides the user with information about travel time, elevation gain, and hike type to assist with decision-making. As TOTAGO expands to new metropolitan areas in the United States and beyond, a system of determining appropriate trails would help expand the project more efficiently and effectively. The motivation for this capstone project was to devise a streamlined method of analysis that would allow TOTAGO to determine trailhead locations within a short distance to a transit stop.

Sustainability management is a growing field that tackles issues using problem solving and decision making to carefully consider the ecological, social, and economic impacts of possible solutions. It also aims to improve well-being as part of decision making contexts (Nyerges 2014). A social-ecological systems (SES) table was developed to examine the relationships that exist which impact accessibility to natural areas (Table 1). The SES table is a concise and comprehensive view of the social, ecological, and economic factors affecting three focal scales: the individual; the Seattle public-transit market; and the nation.

<table>
<thead>
<tr>
<th>BIOPHYSICAL</th>
<th>SOCIAL</th>
<th>ECONOMIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above Focal Scale +1 (National)</td>
<td>Emissions resulting from automobiles tends to be higher than emissions from mass transit by household. In addition to the transition towards fuel efficient automobiles, transit accessibility has the greatest impact towards reducing greenhouse gas emissions at the household level (Zahabi, Miranda-Moreno, Patterson &amp; Barla 2013).</td>
<td>Engaging in Civic ecology can lead to social learning and community resilience in youth (Smith, DuBoid, &amp; Krasny 2013)</td>
</tr>
<tr>
<td>Discourage hikers from using unofficial trails in protected areas to prevent erosion and degradation. The Mountains to Sound Greenway Trust is working on an environmental restoration program that involves volunteers and conservation crews improve natural areas and remove invasive species (Mountains to Sound Greenway Trust n.d.).</td>
<td>Cities are often home to the largest concentrations of poverty, and due to sprawl can be the areas farthest removed from natural and scenic areas.</td>
<td>Millennials value mobile devices over a car (Zipcar 2013)</td>
</tr>
<tr>
<td>Experiencing the outdoors can foster environmental stewardship. This can lead to better preservation for natural areas. Green spaces have been shown to mitigate the health hazards of urban living, lower stress levels, and reduce mental health risks (Nisbet &amp; Lem 2015).</td>
<td>Youth participation in outdoor activities has declined since 2006, but those with greater accessibility were more likely to participate: 21% greater participation for those with nearby walking routes and 25% greater for those with nearby biking routes (Children and Nature Network 2012)</td>
<td>While many outdoor activities require equipment that can be costly, hiking is an outdoor activity that has no associated cost for the individual beyond transportation</td>
</tr>
</tbody>
</table>

Table 1: Social-Economic Systems Table

Some conclusions from the SES table include that there are trends towards decreased car ownership and that accessibility is key to increased engagement with outdoor natural areas; engagement with the outdoors can lead to environmental stewardship. The threshold matrix seen in Table 2 exhibits the competing variables at play within the context of connecting people to the outdoors. The characteristic state displays the generalized current state, while the alternate state provides insight into the state’s resiliency.
The goal of the capstone project prompted an investigation of various software and mapping programs. Major objectives of the project were to 1) identify trailheads that are transit accessible and add service attributes to trailhead data; 2) identify service information for transit stops within specified distance of trailheads; 3) identify the distance between trailheads and transit stops and gather information on the path from the transit stop to the trailhead; 4) explore the project’s mission in the context of sustainability management. The project was successful in many aspects of these objectives but ultimately leaves room for growth and further study.

### 3. Design & Methods

An investigation of available software programs was performed to better understand what method would be the most appropriate course of action. The approach to this project was largely guided by the sponsor's interest in using open source software. The sponsoring organization expressed the desire to use PostGIS to manage spatial data, as they already are familiar with and currently use PostgreSQL, the platform on which PostGIS is constructed.

Using open source data was also a priority of the project, with transit data coming from General Transit Feed Specification (GTFS) files. The source of trailhead data, central to the sponsor's application, was also considered. While TOTAGO does have some existing trailhead data, OpenStreetMap (OSM) data was identified as a way to increase the number of trailheads in their database.

#### 3.1: Project organization

After discussing the project with the sponsors, and analyzing the requirements, the group researched possible ways to approach the project. The use of QGIS was evaluated, and identified as part of the solution. This open-source GIS has a variety of powerful, open-source
tools that would be able to perform the geoprocessing required. Additionally, the software works easily with PosGIS data.

Trailhead and transit stop data can be accessed in PostGIS through exporting the data in QGIS. In QGIS create a new PostGIS connection to your PostGRES database under PostGIS >> New Connection. Once the connection is established and shown under the PostGIS tree root, DB Manager can be used to export the data. DB Manager is located under the Database tab in QGIS. Users can navigate to the PostGIS database and click the icon for Import Layer/File. QGIS offered an important feature which became a central part of the solution, the ‘plugin’. A plugin is a way to organize scripting to perform a series of tasks. It is programmed using Python scripting, which is open source, and familiar to both members of the group. Ultimately, this proved to be an excellent way to organize the project.

3.2: Project Planning

Development of the project began by planning the different activities that would need to be completed to identify trailheads near transit stops, and summarize transit information for these stops. These were organized logically, and this was developed into a scope of work. The project was initially divided into three separate parts:

- Identification of trailheads that are considered transit accessible
- Summary of transit service data for stops near trailheads
- Calculation of distance between stops and trailheads

Lastly, while not separately identified in the scope, the data needs to be exported into files that can be easily consumed by the TOTAGO app. Dividing the project in this manner made it possible to develop the different pieces asynchronously.

Ultimately, the project scripting was managed using GitHub, which facilitates collaboration in coding projects. The project was not migrated to GitHub until late in the process, and this delay made development more challenging. Until the migration, code was inefficiently shared back and forth using a folder on Google Drive.

3.3: Project organization

All scripting work completed for this project has been organized into a QGIS plugin. After the plugin is added to QGIS, it is available as a tool on the toolbar. When this is executed, a dialog box appears that lets the user enter various parameters, including the type and location of input trailhead data, and the location of the folder with the GTFS data. See Figure 1 in Results for a screenshot of the plugin’s dialog box.

The basic plugin structure is created from another plugin, Plugin Builder, which is generally included with QGIS installations. All scripting is from Python. An additional software product, Qt Creator, was used to easily customize the dialog box.

When a plugin in first set up, a folder is created in the plugin folder in the user’s QGIS working folder, for example: C:\Users\(user name)\qgis2\python\plugins. All files used by the plugin are placed in this directory. The plugin can be shared by simply zipping into a compressed folder, and distributing to other users. Alternately, if the plugin is developed using GitHub, it can be installed from this location as well.
3.4: Project development

The scope was divided into different parts to allow asynchronous development. The different parts were developed separately, and then integrated in a later stage of development.

Part 1 – Identification of trailheads that are considered transit accessible

To begin the process, a geographic data layer showing transit stops included in the GTFS is created by QGIS. Then, trailhead data from either a PostGIS database or a shapefile is analyzed by QGIS. Points with the user specified distance are identified as transit accessible, and placed in a new data layer.

The default distance used to determine whether trailheads are transit accessible is set to 400 meters. This parameter can be changed by the user of the plugin, however, 400 meters is a generally accepted distance that the average American will walk rather than drive (Yang & Diez-Roux 2012).

One issue that had to be overcome is that stops and trailhead data is stored as latitude and longitude coordinates, using decimal degrees. However, the group wished to use a consistent distance in meters to identify nearby transit stops. To facilitate this, the data was converted on the fly to Web Mercator, so that distance calculation could be made in meters. This was also necessary to get the distances, discussed below in Part 3.

Part 2 – Summary of transit service data for stops near trailheads

A large part of the project is the script that is used to summarize transit service, as presented in a GTFS. A GTFS is designed to include a transit agency’s schedule, and including information of the routes that are operated, the stops that are served, the times the various routes stop at each transit stop, and in some cases, route traces and fare information.

While made up of text files, a GTFS has some of the characteristics of a relational database. Files can be related based on keys contained in the data. Table 3, below, shows the different files that can make up a GTFS. While many files are required, there are several files where inclusion is optional. By design, optional files were excluded from use in this project.

<table>
<thead>
<tr>
<th>Filename</th>
<th>Required</th>
<th>Defines</th>
</tr>
</thead>
<tbody>
<tr>
<td>agency.txt</td>
<td>Required</td>
<td>The transit agencies that provide the data in the feed.</td>
</tr>
<tr>
<td>stops.txt</td>
<td>Required</td>
<td>Stop locations where vehicles pick up and/or drop off passengers.</td>
</tr>
<tr>
<td>routes.txt</td>
<td>Required</td>
<td>The transit routes that are operated by a transit agency.</td>
</tr>
<tr>
<td>trips.txt</td>
<td>Required</td>
<td>The trips that are made each route. This is the specific service operated by the agency.</td>
</tr>
<tr>
<td>File Name</td>
<td>Required/Optional</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>stop_times.txt</td>
<td>Required</td>
<td>Times that a transit vehicle arrives and/or departs from an individual stop(s).</td>
</tr>
<tr>
<td>calendar.txt</td>
<td>Required</td>
<td>Dates the GTFS is valid.</td>
</tr>
<tr>
<td>calendar_dates.txt</td>
<td>Optional</td>
<td>Exceptions to the calendar file. Note that this file may be specified instead of calendar.txt.</td>
</tr>
<tr>
<td>fare_attributes.txt</td>
<td>Optional</td>
<td>Fare information for the transit agency’s service.</td>
</tr>
<tr>
<td>fare_rules.txt</td>
<td>Optional</td>
<td>The rules that apply to the agency’s transit fares.</td>
</tr>
<tr>
<td>shapes.txt</td>
<td>Optional</td>
<td>The geometry of the routes operated by a transit agency</td>
</tr>
<tr>
<td>frequencies.txt</td>
<td>Optional</td>
<td>Headway for routes with variable frequency of service, often used for rail or Bus Rapid Transit (BRT) service.</td>
</tr>
<tr>
<td>transfers.txt</td>
<td>Optional</td>
<td>Rules for connections at transfer points between different routes.</td>
</tr>
<tr>
<td>feed_info.txt</td>
<td>Optional</td>
<td>Additional information about the feed itself.</td>
</tr>
</tbody>
</table>

*Table 3: Files that comprise a GTFS*

Source: Google Developers 2016

While there are several required files and data elements, the format of a GTFS is flexible, and there is some variation in files created by different agencies. This presented some challenges during development of the coding. As feeds might include optional fields, a user cannot assume that each field referenced will be in the same position in every feed. To account for this, the scripting was set up so that each field referenced was located programmatically.

To compile a complete summary of transit service for stops in proximity to trailhead, information was pulled from the following GTFS files:

- agency.txt
- calendar.txt
- routes.txt
- stops.txt
- stop_times.txt
- trips.txt

A series of dictionaries and dictionary lists were used to summarize various data points, including the transit routes that serve a particular stop and the departures available at a particular stop. Many GTFS include service operated by different agencies. By using an agency id in the output data, the agency operating a particular trip can be identified.

This transit service summary is exported to a file, service_summary.txt, which is created in the same folder as the GTFS files. This information is designed to be used in conjunction with the file that contains information about transit accessible trailheads.
Part 3 – Calculation of distance between stops and trailheads

To show how close each stop is to a particular trailhead, the script calculates the Euclidian distance between these features. Some trailheads have several stops that are within walking distance, and a distance is calculated for each. This information is included in the GeoJSON file that is exported.

When all of the analysis is complete, a GeoJSON is exported that contains a record for each trailhead. The record includes the trailhead id (from the GIS data), the name, the stops that are within walking distance, and the distance between each stop and the trailhead.

4. Results

The group successfully developed a QGIS plugin that accomplished the major tasks identified by the sponsors:

- Identify trailheads that are proximate to transit stops
- Compile basic transit schedule information in a format that is compatible with the trailhead data
- Estimate distance between trailheads and the transit stops

Overall, the group feels that it successfully completed the project. The plugin was successfully developed and deployed. The sponsor’s major goals, noted above, were successfully implemented in a single, open source tool. This tool should give TOTAGO an easy and cost-effective way to increase the amount of detailed information available through their web application.

To use the plugin, a user would download it from GitHub (https://github.com/totago/qgis-gtfs-trails-plugin) and unzip it into QGIS’s plugin folder. The plugin can then be installed in QGIS through the plugin manager. An icon will appear on the tool bar. When selected, the user will be presented with a dialog box with various parameters to complete, as shown in Figure 1. Among the required parameters are the location of the GTFS files, the format and location of trailhead data, and an output directory. The user has the option to add the processed data to the map canvas. A default buffer size of 400 meters is used, though this can be changed by the user.

Trailhead data may be derived from either shapefiles or a PostGIS table. A field with a unique id for trailheads must also be specified. If using PostGIS, details for the connection to the database must be entered. If QGIS cannot connect to the database, the program will terminate without results. The program assumes that PostgreSQL and PostGIS are installed and available.
After successfully running the plugin in QGIS, five new output layers will be generated:

1. Trailhead\_points\_WM – The trailhead points converted to Web Mercator.
2. Trailhead\_buffer\_layer – A polygon file, based on the buffer distance specified in the tool dialog, surrounding each trailhead point.
3. TransitStops\_NonDissolved\_layer – A point file that identifies any transit stop within a Trailhead\_buffer\_layer polygon. If a transit stop is within the buffer range of two different trailheads, it will be counted twice, with two separate point entries.
4. Trailhead\_buffer\_d – A polygon file that is the result of dissolving the Trailhead\_buffer\_layer polygons. A single polygon will remain, which removes all overlap.
5. TransitStops_NonDissolved_layer – A point file that identifies any transit stop within the Trailhead_buffer_d polygon. Even if a transit stop is within the buffer range of two different trailheads, only one entry will exist in this file.

These five outputs can be used for further mapping analysis in QGIS and are supported by the GeoJSON export and service_summary located in the specified directory.

Figure 2 displays an example result in the Seward Park neighborhood of Seattle. The five outputs described above are shown in the QGIS legend. The dark green outlines represent the 400 meter buffer. Light green points are trailheads and orange points are transit stops that are located within the buffer zone. This clearly highlights how transit stops are identified as part of the plugin. These stops can then be explored further as potential trip destinations in the TOTAGO app.

Figure 2: Example Plugin Output

Toward the end of the project, TOTAGO indicated that they would like to be able to use the tool to bulk process data. This was accommodated by the group, who were able to make adjustments that permitted the processing of single or multiple datasets.

While the major goals of the project were successfully realized, some elements that were proposed were ultimately omitted from the final project. The two project components that fall into this category include:

- Using OpenStreetMap data to identify the network distance between stops and trailheads
- Adding transit service data directly to the GeoJSON file output

Different factors motivated each change.
Toward the end of the quarter, it became apparent to the group that the time remaining in the quarter would not be sufficient to complete the project as originally intended. The largest remaining outstanding part of the project was using OpenStreetMap data to compute network distances. This part of the project was designed to provide a more realistic distance measurement between a trailhead and a transit stop, and potentially, to evaluate the quality of path between the two points by analyzing the availability of sidewalks. As a substitute, the Euclidian distance was computed and reported in place of a network distance.

It was unclear how much additional effort would be needed to script the computation of network distances, a serious issue considering the short amount of time that remained. Additionally, a closer look at OpenStreetMap data indicated that the benefits of this work might be marginal at this time. The group concluded that very few streets are flagged as having sidewalks in the OpenStreetMap data. While the number of streets identified as having sidewalks will likely increase in the future, adding this functionality to the tool was unlikely to have much of an impact at this time. Additionally, only transit stops within a fairly small distance of a trailhead are identified as transit accessible. This small distance (400 meters) means that in most cases, the Euclidian distance should provide a suitable indication of how close a transit stop is to a trailhead.

Lastly, and perhaps the most important factor in the decision to omit this part of the project was that this was proposed by the group, and not the project sponsor. While the sponsors do see the benefit of using a network distance, it was never a core part of the project. It would make an excellent item for future development.

Another project modification was a decision not to add the summarized transit information directly to the trailhead data in the final exported GeoJSON file. Originally, the group proposed that transit service data, such as the route numbers serving a particular stop, and even all daily departures from the stop could be added to the final output file. Practically however, this has the potential to be problematic. Some stops have very high levels of service, which leads to some very long record lengths when this information is summarized in a file. Many trailheads are near more than one transit stop, and when combined, lead to even longer record lengths. It was unclear how well these long strings could be accommodated in a GeoJSON file. Conversely, some transit stops are near more than one trailhead, and would have lead to duplicate information stored in the file.

To consume this transit service data in the TOTAGO app, information will have to be extracted from both the service summary file with the trailheads in the GeoJSON file, but the sponsor confirmed that this could be done. Considering all of these factors, it seemed that storing transit data with the trailhead data inside the GeoJSON file had no advantages, and many disadvantages.

Overall, the group considers these modification small, and believes that they do not diminish the overall usefulness of the project. These could be added in a future update of the project, if useful.
5. Discussion

Overall, the group enjoyed working on this project and found the project sponsor's easy to work with and very supportive. Both group members had the desire to learn more about open-source GIS products, and this project proved an excellent way to do this.

A number of factors made this project challenging to complete, however. One of the largest hurdles was the group’s relative unfamiliarity with the open-source products that the sponsor wished to use. While both group members had some basic familiarity with QGIS, neither had used Python scripting in conjunction with the software. Additionally, group members had only a superficial knowledge of PostGIS at best. It took more time than expected to learn basic workflows, and begin scripting.

Researching the plugin format, and setting up a plugin took a fair amount of time at the beginning of the project. After a false start, the group realized that a solution was easily accessible, in the form of Plugin Builder. This QGIS plugin can be used to easily develop a new plugin, and is standard with many, if not all QGIS installations. Once this solution was identified, development of the plugin began.

As noted, the plugin was developed using Python scripting. In Python, indentation levels are highly important, and they must be consistent. QGIS proved unforgiving when indentation was not completely consistent within a block of code, for example, when spaces and tabs were used interchangeably. When this occurs, the plugin simply does not load in QGIS until the error is resolved. This proved frustrating at times.

Over the later phases of the project, several seemingly small issues arose, and these often took more time to resolve than would have been expected. For example, the project scripting relies on accessing and using the geometry of each trailhead in the dataset. This is easily done when the data comes from a shapefile, but not when the data came from a PostGIS database. In this case, the data comes into QGIS as multipoint data, instead of single points when using a shapefile. Researching a solution to this problem required valuable time.

One additional issue was that the project sponsors hoped that they would be able to gain additional trailhead data from the project. While the group was happy to help, this was not understood to be a core part of the project. Ultimately, the group developed a script that used OpenStreetMap extracts to create trailhead data. As the group was unfamiliar with QGIS workflows, and there was not much time available to complete this, a script using ESRI geoprocessing was developed. As the sponsor’s access to ArcGIS was unknown, trailheads were created for the entire ‘lower 48’ states. While some refinement will be necessary, the group hopes that this will be helpful to TOTAGO.

An alternate objective of the project was for the group members to explore options in open source software. This capstone course is part of a larger Masters of Geographic Information Systems (MGIS) program that balanced proprietary and open source software as part of the course instruction. This project allowed for a more in-depth exploration of open source products: how they function alone and work together. The following list includes programs that were investigated to some extent, even if they were not eventually used to create project deliverables: QGIS, PostgreSQL, PostGIS, Heroku, Ruby, Qt, and GitHub. Exposure and examination led to a better understanding of how a combination of Open Source software can be used to achieve mapping and development goals.
6. Business Case and Implementation Plan

The project’s resulting plugin should prove useful to TOTAGO and their work with the Mountains to Sound Greenway Trust. If the necessary data is in place, the plugin could theoretically locate any trailhead within the provided parameters to further TOTAGO’s mission of connecting people to the outdoors. The trailhead data prepared by the project team provides the opportunity to begin exploration in the contiguous United States. Acquiring transit data for other focus areas will be an important step to implementing the plugin to its full potential.

With any data currently available, the first step will be to generate results using the plugin and evaluate the results for use in the TOTAGO app. Following this first implementation, it will be important to identify target areas, and acquire the best possible data for analysis. This would include acquiring local GTFS data and locating trailhead data more accurate than what was provided. Following data acquisition, exports should be regenerated using the plugin. Further investigation of the trails would be used to update the Hike Summary and Trail Details available on the TOTAGO app.

The current plugin calculates the Euclidean distance between identified trailhead points and their corresponding transit stops. Time constraints limited the original goal of analyzing route distance over Euclidean distance. Review of the OSM data revealed that path and sidewalk data may not be readily available in many areas of the United States. This is likely to improve over time as more contributions are made to OSM. In the meantime, Euclidean distance was meaningful, but future implementation should involve the consideration of route distance. This can be achieved in a variety of ways. One method would be to update the current plugin code to run this calculation as part of the export. A potentially faster course of action would be to utilize some of the plugins currently available, such as pgrouting, to analyze the route distance. This would create a more accurate hike summary available to users on the app, and could possibly be used to incorporate data such as slope, type of path, and distance information.

Future improvements to the plugin could include reviewing the code to make the plugin run more quickly and cleanly. Python allows for multiple methods to achieve the same goal. There are likely to be alternative approaches that could make the code more efficient. Other future improvements would likely be driven by the development of TOTAGO. As their mission expands and identifies new project needs, the plugin could be updated accordingly.

This plugin will assist the project sponsors’ current work activities by providing an efficient method of identifying trailheads for further investigation. This process contributes to TOTAGO’s overall mission of connecting people to the outdoors. The cost-benefit scenario favors the use of the plugin for contributing to project objectives. There will be time required to run the plugin, and acquire any necessary data, but the payoff allows for an efficient collection of trailhead data that can be used to enhance the trips promoted and available on TOTAGO.
Literature Cited


