Capstone Project Impervious Area Update

Final Report

5.14.2021

Andrew Knox - Arthur Wu - Jared Woodard - Matthew Griffith - Charles Pawling

Abstract

The client that our team has been working with this semester is DC Water. DC water is a company that provides different water treatment services for residents and businesses in Washington DC. A tool that DC water uses to manage its service billing is an Impervious Area Change detection app, that uses geographic data to detect if a customer needs to have their bill changed based on the amount of impervious area present in their location. And the app needs to be updated with new data every 3-5 years. Originally, DC Water Wanted our team to find a way to speed up this process of finding the changes in impervious areas by using satellite imaging tools such as ArcGIS. The initial scope of our project consisted of multiple requirements. First of all we were to analyze changes to impervious areas by comparing satellite data in ArcGIS and then determine if customers needed to be billed differently based on our finding. Any new Data we found was to be added to DC Waters' already existing database. We were expected to create a python script that would automate this data analysis process. Finally, our team was expected to document any changes we noticed in a github repository as well as create a manual explaining any code or changes we created. As our project progressed our team had to audible several times and change the scope of our project to better accommodate our capabilities. The final scope of our project had similar but much different requirements. The expectations for our final scope included a Power BI dashboard that could be used to analyze the DC Water impervious area data quality, an image classification tool created through Microsoft Azure, a manual for the Power BI dashboard, and a recommendations report for future groups. Throughout this project our team worked closely with our Project Leader Chad Rogers as well as the Lead Director Hari Karup to accomplish our goals.

Methods

In order to fulfil the project goals and all of the changes throughout the project, we kept in constant contact with Chad Rogers having two meetings a week to discuss our duties, constant communication between each other, and also asking questions whenever we needed to learn a tool or strategy we were not comfortable with.

To conduct the initial analysis of the data and ensure that our data both from DC Water and from outside sources matched up geographically we had to use ArcGIS to manipulate our data from both sources until we got a set that matched and allowed us to compare the variables between both sets of data. In ArcGIS we take data from the city of DC and DC Water and use tools found in ArcGIS to get them to be aligned as they would be in the real world for use in PowerBI for more in depth analysis.

For our recommendation for an image detection tool that could detect changes in impervious areas, we took clips from areas that had changes in impervious areas from ArcMap and converted them to a PNG file. Once we did this, we then put those images into the Microsoft Vision tool and added a tag for images from 2017 and 2019. After this, we ran the tool and checked to see what results we were given.

For Power BI the work was focused mostly on data manipulation prior to creating the visualizations. First we had to import all of the tables we were interested in. Then we had to link them by the appropriate columns to ensure that our data would be represented properly. Next we had to create many measures to display aggregates of the data. Lastly we created visualizations by combining these measures with columns from the tables.

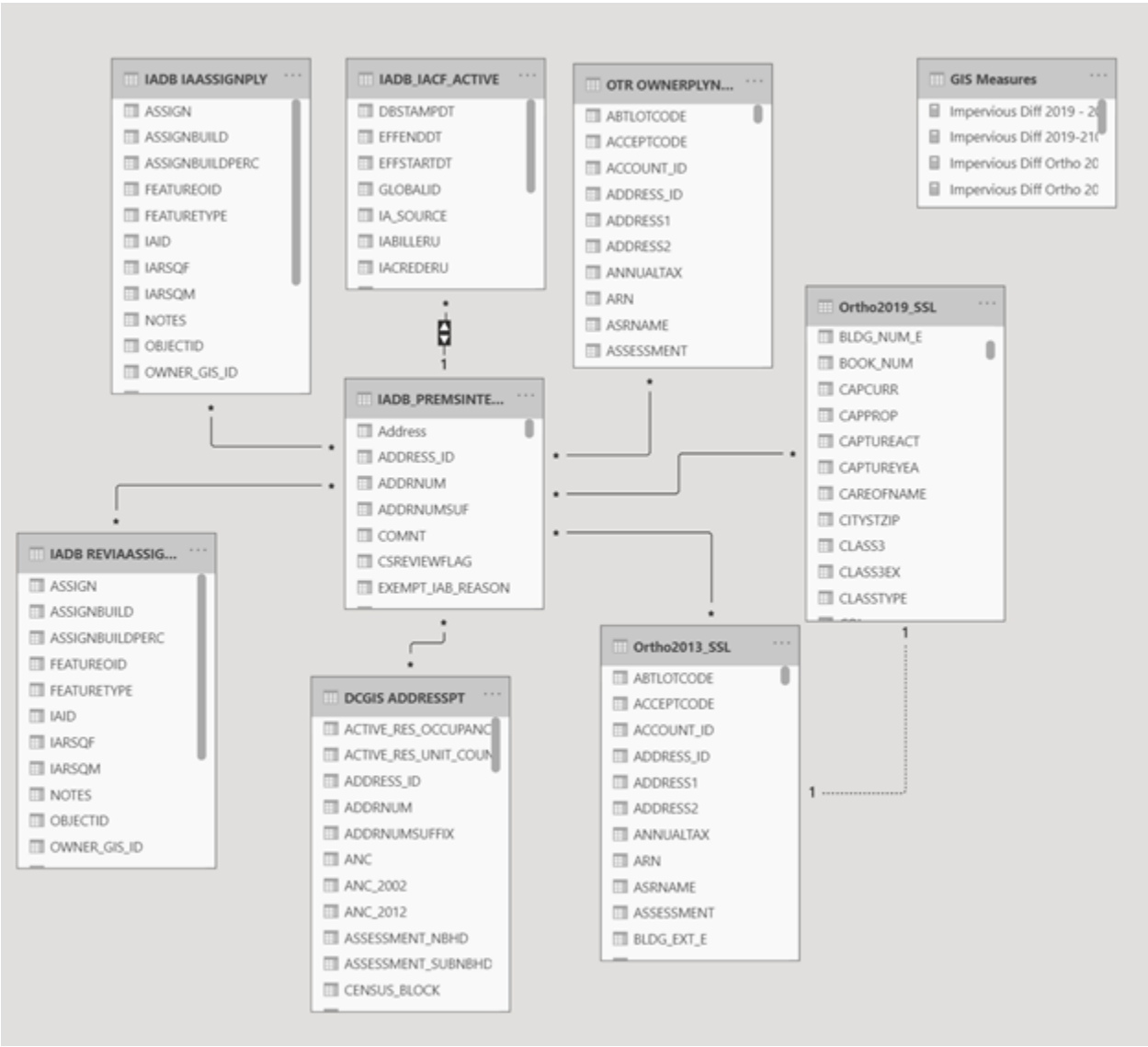
Some challenges that we had was that our project goals and deliverables were constantly changing and we were using tools that we were not totally comfortable with. Due to the limited amount of time from changing tasks, we did not have enough time to learn all of the new tools to the capability they wanted. We then spoke with Chad Rogers about how we can change our scope and deliverables to not only allow us to have a realistic goal, but also one that could still benefit DC Water. In the end, we created a Power BI dashboard with visualizations and tables that they could use to identify possible changes in impervious areas.

Deliverables

Our project had five deliverables for DC Water. Our first deliverable was updating their database and creating two new tables. We titled these tables

Ortho2013_SSL and Ortho2019_SSL. We created these tables by combining ortho tables from DC Open Data with DC Water tables containing an SSL column. DC Water uses SSL to identify buildings or groups of buildings. By combining these ortho tables with SSL we would also DC Water to compare their data to the DC government data.

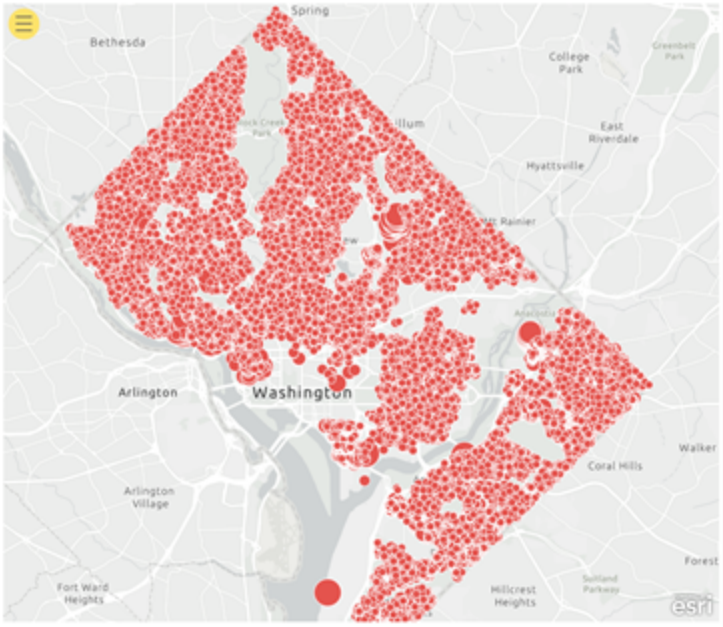
After creating these tables we needed to create a way for DC Water to compare their data to DC government data. We achieved this by creating a Power BI dashboard. The first step was to import our two new tables as well as DC Water's tables to create a data model.



After tweaking this model we used it and new measures to create visualizations that would allow DC Water to see how their data compared to the government data. It also allowed them to compare their data to their older data to see if it was more accurate.

SSL	SUM Impervious Ortho 2013 (SQFT)	SUM Impervious Ortho 2019 (SQFT)
	54,547,793.86	
0005 0855	3,317.68	2,333.20
0006 0042	22,744.55	21,985.84
0008 0809	13,908.36	46,699.84
0008 0812	24,539.63	48,418.56
0013 0044	22,169.30	3,719.37
0014 0019	8,012.52	2,972.56
0014 0047	1,356.66	1,443.51
0014 0048	889.73	748.79
0014 0051	2,019.70	2,068.07
0014 0054	2,900.66	2,280.82
0014 0055	1,363.23	1,078.71
0014 0058	954.22	1,066.21
0014 0059	1,098.09	1,305.21
0014 0070	5,799.95	1,637.86
0014 0071	4,985.23	5,957.37
0014 0072	1,488.49	1,450.39
0014 0073	16,471.74	10,807.21
0015 0004	23,380.66	10,313.26
0015 0005	216.94	207.20
0015 0030	16,049.25	1,447.20
0015 0049	13,154.29	8,733.54
0015 0050	15,035.07	15,206.08
0015 0051	377.00	3,122.22
0015 0052	12,010.80	9,138.82
0015 0827	6,848.77	12,891.14
0016 0030	505.09	510.20
Total	**292,828,078.91**	**236,516,601.59**

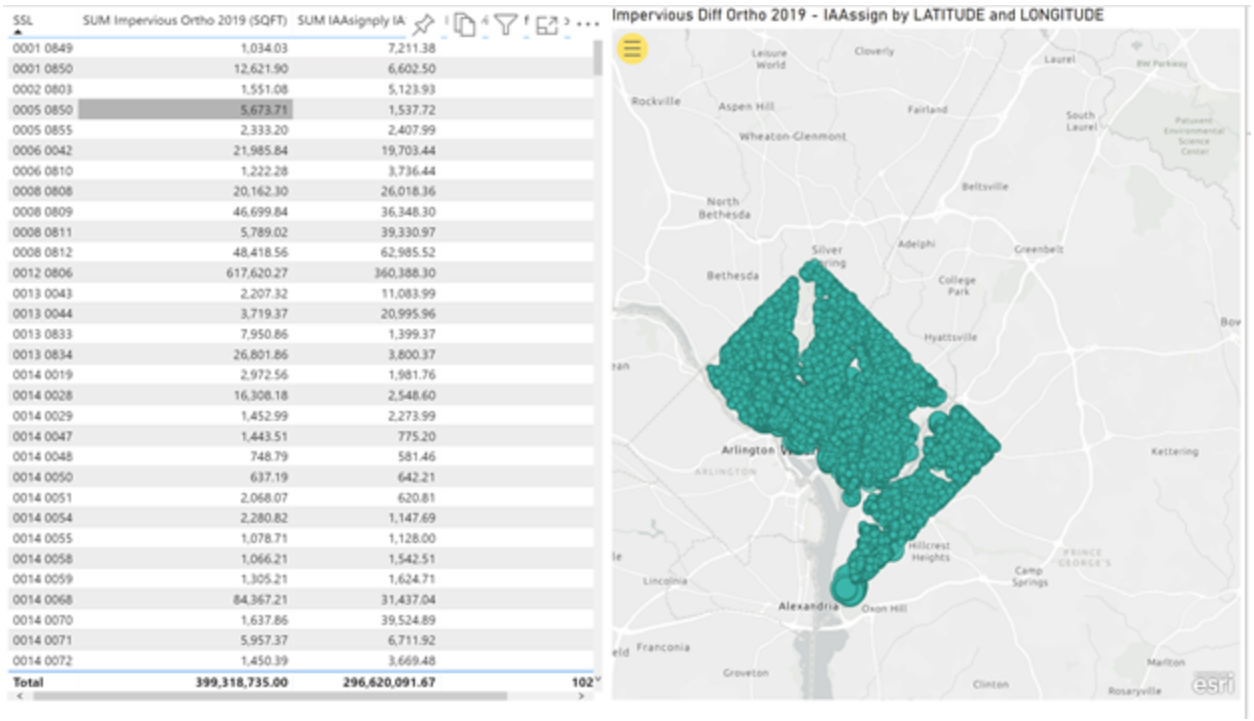
Impervious Difference (2013-2019) (SQFT) by LATITUDE, LONGITUDE and RES_TYPE



MIXED USE | NON RESIDENTIAL | RESIDENTIAL

PROPTYPE	Shape_Ar_1	SHAPEAREA	Impervious Diff 2019...
Residential-Cooperative (Verti	200,209.26	252,416.12	
Residential-Conversion (Less T	1,192,058.25	1,210,402.21	
Cooperative-Mixed Use (Vertica	35,196.62	52,938.92	
Restaurants	45,378.66	59,174.43	
Residential-Condominium (Garag	2,986.47	12,778.62	
Commercial-Office (Small)	203,569.89	212,662.35	
Store-Department	35,218.76	41,445.91	
Residential-Conversion (5 Unit	32,313.86	37,275.87	
Vacant-Unimproved Parking	23,330.55	25,430.75	
Fraternity/Sorority House	2,216.70	4,254.14	
Commercial-Retail (Condominium	136.88		
Club-Private	25,679.70	25,525.51	
Industrial-Truck Terminal	31,585.78	31,429.24	
Industrial-Warehouse (Condomin	209.69		
Residential-Cooperative (Horiz	262,462.41	262,246.17	
Cooperative-Mixed Use (Horizon	2,998.66	2,650.99	
Residential-Single Family (Mis	37,768.08	37,183.07	
Store-Restaurant	129,934.17	129,264.93	
Residential-Mixed Use	20,463.52	19,780.03	
Industrial-Raw Material Handli	1,843.54	1,082.19	
Tourist Homes	3,568.33	2,706.48	
Special Purpose (Memorial)	1,164.02		
Commercial-Planned Development	4,488.69	3,084.38	
Store-Barber-Beauty Shop	30,518.52	29,063.04	
Residential-Conversion (More T	44,336.36	41,873.92	
Total	**70,608,438.15**	**53,088,213.39**	

Impervious Diff 2019 - 2013 (SQFT) by PROPTYPE



SSL	SUM Impervious Ortho 2019 (SQFT)	SUM IAAsignply IA
0001 0849	1,034.03	7,211.38
0001 0850	12,621.90	6,602.50
0002 0803	1,551.08	5,123.93
0005 0850	5,673.71	1,537.72
0005 0855	2,333.20	2,407.99
0006 0042	21,985.84	19,703.44
0006 0810	1,222.28	3,736.44
0008 0808	20,162.30	26,018.36
0008 0809	46,699.84	36,348.30
0008 0811	5,789.02	39,330.97
0008 0812	48,418.56	62,985.52
0012 0806	617,620.27	360,388.30
0013 0043	2,207.32	11,083.99
0013 0044	3,719.37	20,995.96
0013 0833	7,950.86	1,399.37
0013 0834	26,801.86	3,800.37
0014 0019	2,972.56	1,981.76
0014 0028	16,308.18	2,548.60
0014 0029	1,452.99	2,273.99
0014 0047	1,443.51	775.20
0014 0048	748.79	581.46
0014 0050	637.19	642.21
0014 0051	2,068.07	620.81
0014 0054	2,280.82	1,147.69
0014 0055	1,078.71	1,128.00
0014 0058	1,066.21	1,542.51
0014 0059	1,305.21	1,624.71
0014 0068	84,367.21	31,437.04
0014 0070	1,637.86	39,524.89
0014 0071	5,957.37	6,711.92
0014 0072	1,450.39	3,669.48
Total	**399,318,735.00**	**296,620,091.67**

These are just some examples of visualizations we created.

Our third deliverable was a report describing everything we did in Power BI. We hope this will help DC water understand our methods and allow DC Water to replicate them in the future. It may also help future iConsultancy students if they continue to work on it.

Our fourth deliverable is a recommendations report. It lists what we think would be the next steps for this project. We included this report because we hoped it would help, but also because we weren't able to complete the project goal ourselves.

Our last deliverable is this final report. It will clarify the work we've done as well as provide important feedback to DC Water.

These deliverables all work towards the original goal of the project, to improve the amount of time it takes to update impervious areas and customer's water bills. Our first two deliverables worked more directly towards that goal. The new tables did so by providing alternative ways that DC water could update their data. The dashboard also goes hand in hand with that by providing DC Water with ways to see if their data is accurate. We'd hope that it could be used to even achieve the final goal of the project if tweaked a bit. The last three deliverables contribute to the final goal of the project more indirectly. They are mainly used to

provide DC Water with information on how to complete the project with another group. They also provide our feedback on ways to improve their future projects via better communication and planning.

Recommendations

Our Future recommendations for DC Water are listed in our recommendations report which is delivered separately from the final report. We will restate our recommendations again for comprehensivity.

Microsoft Vision

DC Water wants to find new tools they could use to automate impervious area change detection. A tool they asked us to investigate is Microsoft Vision. The end goal of using Microsoft Vision was to see if it could tell when there was a difference between two ortho images of the same area. After developing a model of their data and examining the results, we noticed that this tool is meant for detecting and recognizing items within images. In our model, our two tags were 2017 and 2019, so it would try to detect images that were from 2017 and 2019 rather than observing changes between the two years. If this is a tool still of interest, we would recommend using tags that would describe the picture more such as construction and non-construction, but this would just be labelling those images rather than detecting change. Vision also recommends using at least 50 images per tag, so more clips would need to be created. In the future, we would recommend DC Water to look into tools in ArcMap since it would work with the tools and datasets already in ArcMap and would also be more useful in detecting change in ortho images.

PowerBI

Data Updates

Currently, there are still some discrepancies within the data and tables loaded into the PowerBI dashboard. A specific example of this is the existence of empty SSL rows in the Ortho2013 table. With the existence of these empty SSL rows, there are 55 million square feet of impervious area not assigned to an SSL location. This throws off the balance between the Ortho2013 and Ortho2019 layers which will inevitably cause errors in the dashboard. To continue refining the dashboard, we would recommend to continue looking for and fixing

discrepancies in the data. Since our visualizations are effective, increasing the quality of the data is a major way to raise the quality of the dashboard.

The dashboard should also be updated with new data once every few years to make sure that the dashboard is up to date. This can be done manually by following our dashboard manual, or a data gateway can be established in the dashboard to automate updates in the dashboard. Advice to do this can also be found in our dashboard manual.

Analysis

Future groups should develop data analysis reports on the data presented in the dashboard to draw conclusions and make decisions. This will allow for greater decision making ability using the PowerBI dashboard. The reports should focus on studying how DC Open Data has changed between 2013 and 2019 in terms of impervious area, and how DC Water's data compares in data quality to the DC Open Data 2019 dataset.

Conclusion

Our team was tasked to assist the DC Water Impervious Area Team by assessing their current impervious area data and suggesting ways to improve their data analysis process. To do this, our team developed a Power BI dashboard that allows the team to visualize the differences between 2013 and 2019 impervious area data fetched from DC Open Data, and compare DC Water's current impervious area data with data from the 2019 DC Open Data dataset. We also developed a dashboard manual allowing future groups to update and replicate the dataset in the future. We believe that through our efforts, DC Water obtained a better understanding of their impervious area data and possible future options in their attempts to automate data quality checks and analysis. We have also left DC Water with the ability to continue our work through future groups by suggesting further data analysis and data updates following our dashboard manual. If there are any further questions or concerns, you can reach our team at dcwiaup21@gmail.com and we will do our best to respond as soon as possible.